Institute of Measurement Science
Slovak Academy of Sciences
Dúbravská cesta 9
841 04 Bratislava, Slovakia

MEASUREMENT 2009

Proceedings

MEASUREMENT 2009

7th International Conference on Measurement

Smolenice, Slovakia, May 20 - 23, 2009

The 7th International Conference MEASUREMENT 2009 is organized by the Institute of Measurement Science of the Slovak Academy of Sciences, Bratislava, Slovakia.

7th International Conference on Measurement

MEASUREMENT 2009

Proceedings

Congress Centre of SAS, Smolenice Castle, Slovakia
May 20 -23, 2009

MEASUREMENT 2009

Editors: Milan Tyšler, Ján Maňka and Viktor Witkovský
Publisher: Institute of Measurement Science
Slovak Academy of Sciences
Dúbravská cesta 9, 841 04 Bratislava

ISBN 978 - 80 - 969672 - 1 - 6

Printed in Slovakia
by Komprint s.r.o., Bratislava

ORGANIZERS OF THE CONFERENCE

Organizer

Institute of Measurement Science, Slovak Academy of Sciences, Bratislava

Co-organizers

Institut für Fertigungstechnik, Technische Universität Wien, Austria
Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic
Faculty of Biomedical Engineering, Czech Technical University in Prague, Kladno, Czech Republic
Faculty of Mechatronics and Machine Building, Kielce University of Technology, Poland
Faculty of Mechanical Engineering, Technical University Košice, Slovakia
Faculty of Electrical Engineering and Information Technology, STU Bratislava, Slovakia
Slovak Institute of Metrology, Bratislava, Slovakia

Technical sponsorship:

IEEE, Czechoslovakia section
Slovak Metrological Society
Slovak National Committee U.R.S.I.
Slovak National Committee IMEKO
Association of Slovak Scientific and Technological Societies

Conference Chairman

Milan Tyšler, Slovakia

Steering Committee

S. Adamczak, Poland
S. Ďuriš, Slovakia
I. Frollo, Slovakia
V. Havlíček, Czech Republic
H. Osanna, Austria
J. Vajda, Slovakia
J. Živčák, Slovakia

Local Organizing Committee

M. Tyšler
J. Maňka
V. Witkovský
P. Andris
M. Hain
E. Bukovenová
K. Cimermanová
V. Juráš
S. Karas
M. Škrátek
S. Štolc
M. Teplan

International Program Committee

J. Maňka, Slovakia, Chairman
J. Bartl, Slovakia
V. Boháč, Slovakia
Z. Domotor, USA
D. Eidukas, Lithuania
Y. Gao, Hong Kong, China
D. Hofmann, Germany
J. Hozman, Czech Republic
R. Hudák, Slovakia
Y. V. Chugui, Russia
E. Kawate, Japan
G. Kozmann, Hungary
L. Lhotská, Czech Republic
R. Maniewski, Poland
K. Meigas, Estonia
L. Michaeli, Slovakia
P.P.L. Regtien, The Netherlands
K. Roubík, Czech Republic
V. Smieško, Slovakia
D. Šimšík, Slovakia
S. Takayama, Japan
M. Tarata, Romania
L.I. Titomir, Russia
E. Tkacz, Poland
G. Wimmer, Slovakia
V. Witkovský, Slovakia
E.G. Woschni, Germany
J. Zakrzewski, Poland

Preface

The international conferences MEASUREMENT have always fostered intensive knowledge interchange among various disciplines within the broad field of measurement science and technology. Traditionally, participants are scientists dealing with theoretical problems of measurement as well as engineers applying new measuring methods and systems in research, industry, or medicine.

This book contains Proceedings of the 7th International Conference MEASUREMENT 2009 that was organized as a continuation of successful events in the years 1997 to 2007. It covers topics from theoretical problems of measurement, measurement in biomedicine, measurement of physical quantities and specific measuring systems. The aim of the Conference was not only to continue in the exchange of information among experienced specialists in the disciplines of measurement science, metrology, and selected application areas but also to attract young students and specialists and to challenge them by the competition for the Young Investigator Award sponsored by the Czechoslovak section of the IEEE.

The conference was developed to its current form thanks to Professor Ivan Frollo, founder and chairman of the previous MESUREMENT conferences who started the meetings of the international community more than 12 years ago and converted the former local workshops to international conferences. Since then, every two years, the attractive conference center in the Smolenice Castle offers the participants the possibility to meet and exchange their ideas form effective co-operations and build new professional and friendly contacts.

The conference was traditionally organized by the Institute of Measurement Science of the Slovak Academy of Sciences with substantial and inspiring support of all co-organizers and technical sponsors. Representatives of them were involved in preparation of the conference program and proceedings, what gives a guarantee that the conference would maintain its high scientific level.

We hope that this year's conference will become again both, a memorable scientific meeting and an enjoyable social event. The organizers also believe that the proceedings reflect the professional enthusiasm during the conference and the book will be a quoted reference in the field of Measurement Science.

Milan Tyšler, Ján Maňka and Viktor Witkovský
editors

CONTENTS

Theoretical Problems of Measurement

Measurement in Biomedicine

Specific Measuring Systems

Measurement of Physical Quantities

Theoretical Problems of Measurement

Errors with Direct Process -Coupled Digital Sensors

E.-G. Woschni

ACATECH München and University of Technology Chemnitz D-09197 Chemnitz, Germany

Email: E.-G.Woschni@infotech.tu-chemnitz.de

Abstract. *Today increasingly sensors with direct digital output are used. Because of the direct coupling with the process often an anti-aliasing filtering is not possible. In the paper it is shown in detail that these errors are depending on the signal processing after sampling. As an example with great importance in practice it is shown that in mean-value measurement – for instance in surface measuring – no aliasing errors occur even if the sampling theorem is violated. On the other hand processing with a differential algorithm leads to increasing errors.*

Keywords: Sampling Theorem; Signal Processing; Aliasing Errors

1. Introduction

Today and the more in future because of the advantages of direct computer-coupling sensors with direct digital output will be used. In these cases an anti-aliasing filtering often is not possible: The sensors are direct-connected to the signal source of the process to be measured and if the signal is e.g. a nonelectric quantity as for instance a force an anti-aliasing filtering is not possible.

The paper therefore deals with the problem of both cut-off and aliasing errors with and especially without anti-aliasing filtering. Of special interest with respect to the field of intelligent measuring instruments is the fact, that these errors both are depending on the algorithms of signal processing after sampling.

2. Errors without Signal Processing after Sampling

Using the mean-square error definition often closed mathematical solutions are gained because of the validity of Euclidean geometry [1]. As well-known the cut-off-error is with $S_{xx}(\omega)$ = spectral power density of the signal [2]

$$\overline{\varepsilon_1{}^2} = 2\int_{\omega_{s/2}}^{\infty} S_{XX}(\omega)d\omega. \tag{1}$$

The same value has the aliasing-error if only the side-band of first order has to be taken into consideration. The general case of a non-ideal anti-aliasing filtering $G_a(j\omega)$ as shown in Fig. 1 yields [3]

$$\overline{e^2} \approx \overline{e_1^2} + \overline{e_2^2} + \overline{e_3^2} =$$
$$2\left[\int_0^{\omega_S/2} S_{XX}(\omega)\left|1-G_a(j\omega)\right|^2 d\omega + \int_{\omega_S/2}^{\infty} S_{XX}(\omega)d\omega + \int_{-\omega_S/2}^{\omega_S/2} S_{XX}(\omega+\omega_s)\left|G_a\left[j(\omega+\omega_s)\right]\right|^2 d\omega\right] \tag{2}$$

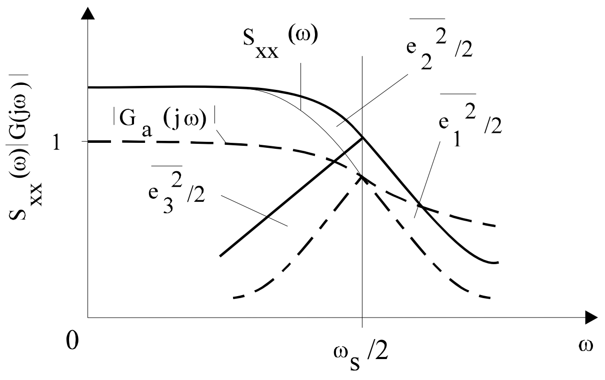


Fig. 1. $\overline{e_1^2}$ = damping error; $\overline{e_2^2}$ = cut-off error; $\overline{e_3^2}$ = aliasing error

If mirrored side-bands of higher orders fall into the range $-\omega_S/2<\omega<\omega_S/2$ the last term in (2) is

$$\overline{e_3^2}=\sum_{r=1}^{\infty}\int_{-\omega_S/2}^{+\omega_S/2} S_{XX}(\omega+r\omega_S)\left|G_a(j[\omega+r\omega_S])\right|^2 d\omega \tag{3}$$

3. Errors with Signal Processing after Sampling

$G_{id}(j\omega) = F\{O_{p_{id}}\}$ may be the ideal and $G_{real}(j\omega) = F\{O_{p_{real}}\}$ the real linear processing procedure. If the sampling Theorem is fulfilled linear system theory yields the mean-square error

$$\overline{e^2(t)}=2\left[\int_{\omega_S/2}^{\infty} S_{XX}(\omega)\left|G_{id}(j\omega)\right|^2 d\omega+\int_{0}^{\omega_S/2} S_{XX}(\omega)\left|G_{id}(j\omega)-G_{real}(j\omega)\right|^2 d\omega\right]. \tag{4}$$

If the sampling Theorem is not fulfilled, e. g. including cut-off and aliasing errors one obtains taking into consideration also side-bands of higher order as with equ. (3)

$$\overline{e^2(t)}=2\left[\begin{array}{l}\int_{\omega_s/2}^{\infty} S_{XX}(\omega)\left|G_{id}(j\omega)\right|^2 d\omega+\int_{0}^{\omega_s/2} S_{XX}(\omega)\left|G_{id}(j\omega)-G_{real}(j\omega)\right|^2 d\omega \\ +\sum_{r=1}^{\infty}\int_{-\omega_s/2}^{+\omega_s/2} S_{XX}(\omega+r\omega)\left|G_{real}(j\omega)\right|^2 d\omega\end{array}\right]. \tag{5}$$

The application of this theory to typical model signals leads to solutions with great importance to problems of intelligent measurement. In most cases the power density of the signal is not known because of missing a-priori-information. Then $S_{XX}(\omega)$ is to be estimated - e. g. caused from a band-limited white noise by a low-pass of first order –

$$S_{XX}(\omega)=\frac{S_O}{1+(\omega/\omega_O)^2}; \qquad S_O=0 \;\; for \;\; \omega\geq\omega_{XO}. \tag{6}$$

For this example it follows

a) without signal processing
   - with anti-aliasing filtering

$$\overline{\varepsilon^2} = 2\int_{\omega_s/2}^{\omega_{x0}} \frac{S_0}{1+(\omega/\omega_0)^2} d\omega = 2S_0\omega_0(arc\tan\omega_{x0}/\omega_0 - arc\tan\omega_s/2\omega_0); \tag{7a}$$

   - without anti-aliasing filtering

$$\overline{\varepsilon^2} = 4\int_{\omega_s/2}^{\omega_{x0}} \frac{S_0}{1+(\omega/\omega_0)^2} d\omega = 4S_0\omega_0(arc\tan\omega_{x0}/\omega_0 - arc\tan\omega_s/2\omega_0); \tag{7b}$$

b) with signal processing $G(j\omega) = 1 + j\omega/\omega_0$
   (PD-algorithm)
   - with anti-aliasing filtering

$$\overline{\varepsilon^2} = 2S_0(\omega_{x0} - \omega_s/2) \tag{8a}$$

   - without anti-aliasing filtering

$$\overline{\varepsilon^2} = 4S_0(\omega_{x0} - \omega_s/2) \tag{8b}$$

c) with low-pass filtering ω_{LP} after sampling
   (I-algorithm)
   - with anti-aliasing filtering

$$\overline{\varepsilon^2} = 0 if\, \omega_{LP} < \omega_s/2; \tag{9a}$$

   - without anti-aliasing filtering aliasing errors only appear if $\omega_{xo} > \omega_{S} + \omega_{LP}$

$$\overline{\varepsilon_{al}{}^2} = 2\int_{\omega_s-\omega_{LP}}^{\omega_{x0}} \frac{S_0}{1+(\omega/\omega_0)^2} d\omega = 2S_0\omega_0\left(arc\tan\omega_{x0}/\omega_0 - arc\tan\frac{\omega_s-\omega_{LP}}{\omega_0}\right). \tag{9b}$$

Of special interest in practice is the case of mean operation after sampling, as for instance used in surface measurement. As Fig. 2 shows that means a low-pass-filtering ω_{LP}.

In this case no aliasing error occurs as long as the limiting frequency of the signal ω_C is less than $(\omega_S - \omega_{LP})$. That means in this case the sampling frequency has to be only

$$\omega_S \geq \omega_C + \omega_{LP} \tag{10}$$

Instead of the value due to the sampling theorem now the sampling frequency is allowed to be only half of this value for the extreme case $\omega_{LP} \to 0$ without errors!

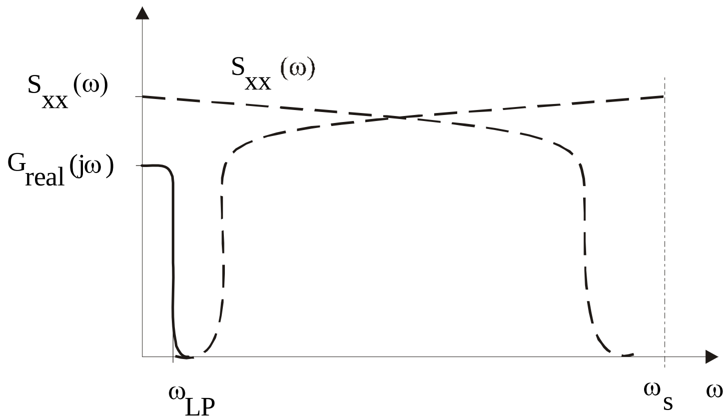


Fig. 2. Mean-operation after sampling

In general we gain the statement
- an algorithm with an integral (I-)part leads to decreasing errors,
- an algorithm with a differential (D-)part leads to increasing errors

compared with the case of not a signal processing after sampling, as normally taken into consideration when dealing with these problems.

4. Conclusions

Today and the more in future in connection with "embedded sensors" sensors with digital output will be used. In these cases an anti-aliasing filtering often is not possible: The sensors are direct-connected to the signal source of the process to be measured and if the signal is e.g. a nonelectric quantity as for instance a force an anti-aliasing filtering (low-pass- filtering) is not possible.

The paper therefore deals with the problem of both cut-off and aliasing errors with and especially without anti-aliasing filtering. Of special interest with respect to the field of intelligent measuring instruments is the fact, that these errors both are depending on the algorithms of signal processing after sampling. It is shown, that an algorithm with an integral (I-)part leads to decreasing errors, while an algorithm with a differential (D-)part leads to increasing errors. In the extreme case $\omega_{LP} \rightarrow 0$ instead of the value due to the sampling theorem the sampling frequency is allowed to be only half of this value.

In general the solution to avoid aliasing errors is oversampling. Due to the generation of micro-electronics it is well-known, that every 7 years the sampling frequencies will increase one order in magnitude [4], [5] and [6], so oversampling will be easily possible in future.

References

[1] Woschni, E.-G.: Informationstechnik, 3. ed., Hüthig Verlag, Heidelberg, 1989.
[2] Woschni, E.-G.: Signal Theory, Systems Theory and Errors in Signal Systems, in Handbook of Measurement Design, Vol. 1. Sydenham PH and Thorn R Editors. JohnWiley & Sons, Chichester 2005.
[3] Woschni, E.-G.: Cut-Off and Aliasing Errors - Exact Solutions and Approximations.In proceedings of the XXIVth General Assembly of URSI, Sektion C2, 1993.
[4] Moore, G.: Gramming more Components onto Integrated Circuits. *Electronics,* April 1965.
[5] Diebold: Research Program-Europa, Report Nr. E 146, 1976.
[6] Woschni, E.-G.: Ein Jahrhundert Informationstechnik wird besichtigt, Hirzel, Leipzig, 1980.

Self-Correction of ADC Error Using Additive Iterative Method and Averaging of Dithered Samples

M. Kamenský, K. Kováč

Department of Measurement, Faculty of Electrical Engineering and Information Technology, Slovak University of Technology, Bratislava, Slovak Republic,
Email: miroslav.kamensky@stuba.sk

***Abstract**. In the paper the combination of additive iterative algorithm with averaging of dithered samples is designed for self-correction of ADC linearity error. Iterative method is known as the aid for automated error correction and dithering uses to be applied for quantiser performance enhancement. Dither theory for Gaussian noise has been used for exhibition of new method abilities in ADC characteristic improvement. Experimental ENOB value improvement is more than 2 bits.*

Keywords: ADC Error Correction, Iterative Method, Nonsubtractive Dithering

1. Introduction

Self-correction functions become important part of modern measurement devices. Analog-to-digital converter (ADC) could be used for direct voltage measurements and it is also the basic part of a general digital measurement channel. It is often not difficult to make correction of offset and gain error of measurement transducer (MT) such as ADC. But correction of nonlinearities of the static transfer characteristic is problematic especially if they vary in time.

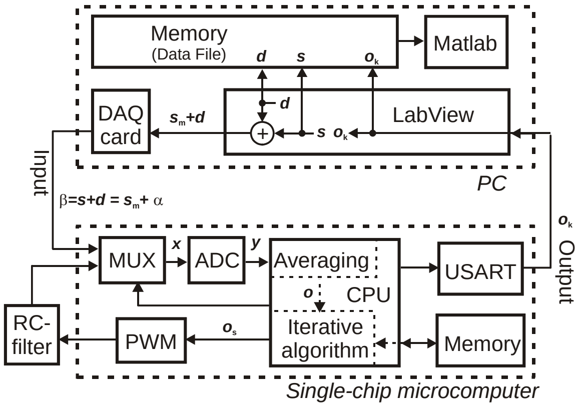


Fig. 1 Block diagram of the workplace.

Methods for automatic correction of ADC have been employed as discussed below. Proposed correction is focused on ADC nonlinearities. Additive iterative method (AIM) is suitable for nonlinear error correction in the case of analog MT. But the ideal characteristic of ADC is fundamentally nonlinear reflecting quantisation error. Quantisation error limits efficiency of AIM, therefore in designed measurement system AIM is combined with nonsubtractive dithering (ND). Block diagram in the Fig.1 shows the experimental system. Diagram of tested measurement unit (MU) is located in the lower part of the figure. Designed measurement system consists only of single-chip microcomputer (with some basic peripherals like power supply etc.) and low-pass RC-filter. The correction method is implemented there. The upper side is devoted to the main PC components used for testing and experiments.

2. Aditive iterative method

For this method four main blocks of the system are needed: MT – in the Fig. 1 it is represented by ADC; block of processing (BP) – CPU (processor) with memory; inverse element (IE) – Pulse Width Modulation (PWM) with RC-filter; switch (SW) – multiplexer (MUX). The correction is performed iteratively in several steps. The BP controls the whole process. It receives input value from the MT and according to the implemented algorithm and data in the memory it calculates next input o_s to the IE. Evaluation of o_s is assigned by correction formula [1]

$$o_{s,i} = o_{s,i-1} + \left(o_{s,0} - h\left[h_{IE}\left(o_{s,i-1}\right)\right]\right) \quad (1)$$

Function $h(x)$ is transfer characteristic of MT and it determines initial value $o_{s,0} = h(s_m)$. Argument s_m represents measured value, which is in these measurements the mean value of input β (Fig. 1). After initial step SW controlled by BP switches the input of ADC from measured signal $x = \beta$ to signal from IE $x = h_{IE}(o_{s,i})$. With each next step of iterations $o_{s,i}$ should become more accurate representation of measured value s_m. After appropriate number of steps actual $o_{s,i}$ could be sent to the output of the MU as a result o_k of correction.

The iterative process is convergent if condition of convergence is satisfied [2]. Then it tends to value given by characteristics of IE $h_{IE}^{-1}(s_m)$. Therefore the aim is to have an ideal IE with inverse characteristics equal to ideal characteristics of MT $h_I(x) = h_{IE}^{-1}(x)$.

Inverse element with averaging

Transfer characteristic of IE determines the best resp. theoretically reachable accuracy of measurement output corrected with the iterative method. Therefore this element must be designed thoroughly. Inverse element for ADC is digital-to-analog converter (DAC) and it has been built by means of pulse width modulation output of microprocessor. PWM circuits are naturally precise but to get the mean $a_{PWM,0}$ ($a_{PWM,k}$ denotes k-th frequency component) of PWM output corresponding to precise DAC result, low-pass filter should be added. Simple RC-filter has been used with frequency characteristics $A_{RCF}(\omega)$ [2], which is influenced by the time constant τ_{RC}. The filter slows down the correction process because after every step the process should wait until settling of filter output.

To speed up the process we proposed to use combination of analog and digital filter. Output of analog RC-filter oscillates in the range of several LSB. As digital filtering a sampling and averaging of N samples in each step of the iterative correction is used. The best way is to use synchronous sampling here but there could be no possibility to synchronize ADC and PWM circuits. Generally sampling gets samples from rectangular window $T_{RW}=N.T_s$ wide (T_s is sampling period) with frequency characteristic $A_{RW}(\omega)$. Denoting $\omega_{RW} = 2\pi/T_{RW}$ and $\omega_{PWM} = 2\pi/T_{PWM}$ (T_{PWM} is period of PWM output), mathematical model of IE error caused by non-synchronous sampling is [2]

$$\Delta h_{IE}\left(o_s\right) = \left[\frac{1}{T_{RW}} \sum_{k=-\infty}^{\infty} a_{PWM,k}\left(\{o_s\}\right) A\left(k\omega_{PWM}, \{\tau_{RC}\}\right) A_{RW}\left(-k\omega_{PWM}, \{t_{s1}, T_{RW}\}\right)\right] - a_{PWM,0} \quad (2)$$

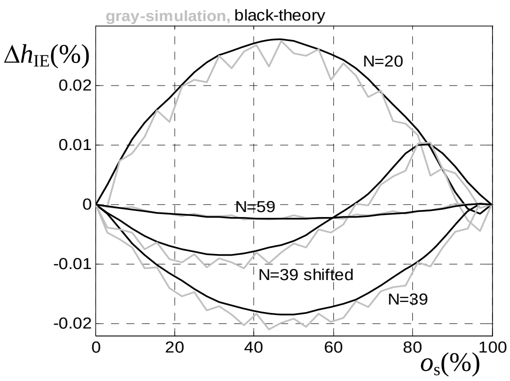


Fig.2 Theoretical error of IE mean evaluated through averaging of samples.

By {} brackets dependency on parameters of system is expressed. According to simulation the RC-filter time constant was set to τ_{RC}=0,1 s. If choosing appropriate number of sample N resp. value $T_{RW}=N.T_s$, special case of quasi-synchronous sampling could be achieved. Suitable values are N=20, 39, 59 (etc.) for which IE error dependencies are shown in the Fig. 2. As could be seen, also time shift t_{s1} – shift from the start of PWM counter to the first sample – influences the accuracy. Expecting error of measurement after correction near to 0,01% of scale, the value N=59 has been chosen.

3. Gaussian noise and averaging

ADC represents real quantiser and its quantisation error limits measurement accuracy of an ideal converter or of real ADC corrected with AIM. Averaging of samples is the way to overcome this limitation employing natural noise present in measured signal. In many cases intentionally added noise (dither) d could help. A process of resolution improvement is called nonsubtractive dithering (ND), if the noise is not subtracted from the signal after quantization.

Noise is present in real applications and usually it is of Gaussian nature, but its dispersion may be too small to get significant resolution improvement. For evaluation of noise influence on accuracy of our system, where dithering with averaging is implemented, the appropriate error parameter must be chosen. Theory [3][4] yields for mean error ε of noisy samples $m_{\varepsilon|s}$

$$m_{\varepsilon|s} = q\sum_{k=1}^{\infty}\frac{(-1)^k}{\pi k}\exp\left[-2\pi^2k^2\left(\frac{\sigma_\alpha}{q}\right)^2\right]\sin\left(\frac{2\pi k s_{\mathrm{m}}}{q}\right) \tag{3}$$

Fig. 3 exposes dependency of the mean error from measured value s_{m} depicted in the range of one quantization step q. Theoretical curves show that with increasing standard deviation σ_α of total input noise α – composed from natural noise and dither together – the mean error decreases. Gray lines are obtained from simulation as a mean from P=20 results of averaging of N=59 samples. The mean is estimated through averaging. The bigger the input noise dispersion is the noisier the curves are. It suggests that an optimal noise variance exists in system with ND and averaging. The formula of theoretical mean error (3) does not include this fact.

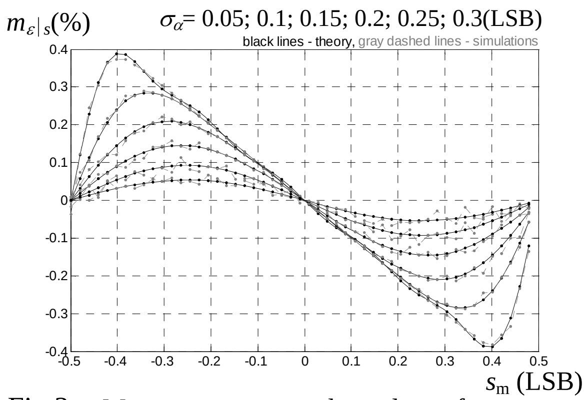


Fig.3 Mean error curves - dependency from standard deviation of Gaussian noise.

The mean-square error (MSE) was chosen as suitable parameter of dithering and averaging performance rating for finding an optimal noise dispersion. It is theoretically evaluated as mean for one whole quantization step [3]. Using (3) and theory from [3] for Gaussian noise it holds

$$\mu_{\mathrm{a}}^2(\sigma_\alpha, N) \cong \frac{\frac{q^2}{12}+\sigma_\alpha^2}{N} + \left(1-\frac{1}{N}\right)\frac{q^2}{2\pi^2}e^{-4\pi^2\left(\frac{\sigma_\alpha}{q}\right)^2} \tag{4}$$

This formula embodies influence of both the mean error (3) and the occurrence of noise in measurement results. For Gaussian dither and N=59 according to theoretical relation from [3] the optimal σ_d is 0,347q, if natural noise is not present in the input signal.

4. Experimental results and discusion

Experiments were performed with designed measurement system, where AIM and ND with averaging of N=59 samples was implemented. In 51 levels of input voltage P=20 correction processes were accomplished. Fig. 4 shows the mean error obtained from these 20 measurement results. As could be seen the additive iterative method used for averaged samples corrects error considerably under the 1 LSB level. In our case appropriate dither helps to suppress nonlinearities involved by quantization. Although implemented method can significantly correct gain error and offset, only nonlinear error component is investigated.

Therefore linear error component has been subtracted from measurement results. To evaluate also dispersion of results within error analysis, RMSE (Root MSE) resp. μ_a curves have been depicted in the Fig. 5. Dither with standard deviation σ_d < 0.05 % (0.5 LSB) has improved results, but differential nonlinearity (DNL) of ADC has caused notable shift of experimental curve against the theory. AIM corrects DNL and therefore it shifts the curve closer to theoretical values. Optimal dither dispersion (see fig. 5) is lower than theoretical optimum because natural noise is present in the input signal β.

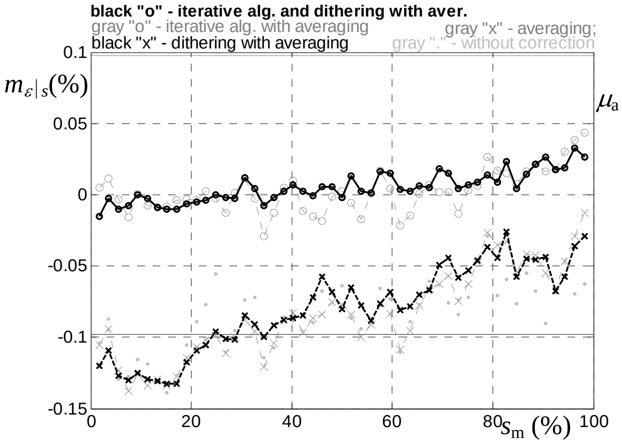


Fig.4 Mean error before and after correction.

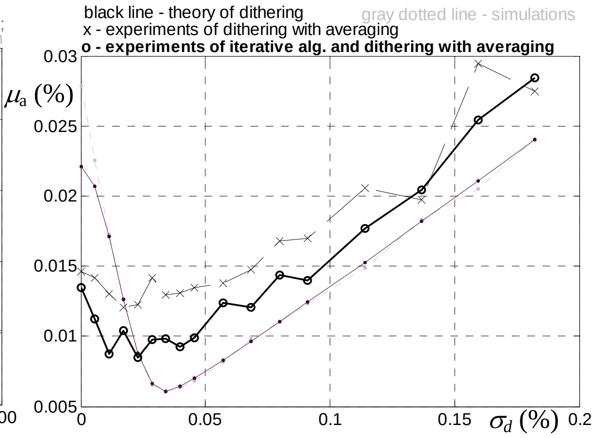


Fig.5 RMSE before and after correction.

5. Conclusions

Combination of additive iterative method and nonsubtractive dithering has been implemented in experimental measuring device. AIM has automatically corrected integral nonlinearity in every process of measurement result evaluation. Averaging has enabled correction under level of 1 LSB of used 10-bit ADC. Analysis of quasi-synchronous sampling of periodic IE output has been performed, leading to negligible error of mean evaluation through averaging. Dispersion of natural noise present in real signal is usually smaller than optimal for dithering. Theoretical dependence of root mean square error (RMSE) upon standard deviation of added noise has been proved through measurements in the whole range and the best dispersion of dither has been found. The RMSE has decreased from 0.036 % to 0.0085 % and adequately ENOB has significantly grown from 9.64 bit to 11.73 bit using proposed correction techniques.

Acknowledgements

Work presented in this paper was supported by the Slovak Ministry of Education under grant No. 2003SP200280802 and by the Slovak Grant Agency VEGA under grant No. 1/0551/09.

References

[1] Muravyov, S. V.: Model of procedure for measurement result correction. In proceedings of the XVI IMEKO World Congress, 2000, published on CD.

[2] Kamenský, M., Kováč, K.: Iterative correction of measurement with averaging of dithered samples. *Advances in Electrical and Electronic Engineering*, 7 (1-2), 358-361, 2008.

[3] Skartlien, R., Øyehaug, L.: Quantization error and resolution in ensemble averaged data with noise. *IEEE Transactions on Instrumentation and Measurement*, 54 (3), pp. 1303-1312, 2005.

[4] Carbone, P., Petri, D.: Effect of additive dither on the resolution of ideal quantizers. *IEEE Transactions on Instrumentation and Measurement*, 43 (3), pp. 389-396, 1994.

Estimation of Exponential ADC Test Signal Using Histogram

[1]M. Sakmár, [2]J. Buša, [1]J. Šaliga, [1]L. Michaeli

[1] Department of Electronics and Telecommunications,
[2]Department of Mathematics, Technical University of Košice, Slovak Republic
Email: michal.sakmar@tuke.sk, jan.busa@tuke.sk,
jan.saliga@tuke.sk, linus.michaeli@tuke.sk

Abstract. *The paper presents a new approach to the ADC INL testing methodology by simple monotonic exponential stimulus which data processing algorithm was suggested with targeting simple implementation in ADC on-board self-testing systems. The novelty of the approach is in determination of stimulus parameters by processing of real measured histogram of record instead of direct record processing in time domain that was used in previous papers. The method was verified by simulations and experimental measurements in comparison with standardised sinewave histogram test that confirmed applicability of the method.*

Keywords: Exponential Stimulus, Histogram Test, On-board ADC Testing

1. Introduction

Analogue-to-digital convertors (ADC) are the main hardware signal processing blocks performing quantization of analogue signals to digital representation (code *k*) within signal digitizing. Quantization levels of ADC (transient levels T_k) are expressed by ADC transfer characteristics. Generally the ideal ADC should have quantization levels uniformly distributed within ADC input range with constant difference between the neighbouring ones. This required difference is usually called nominal quantisation step Q (code bin width W_k). Unfortunately the real ADC have the transient level T_k distributed non-uniformly, i.e. code bin widths are not constant within the ADC full input range. Difference between nominal (ideal) T_k and real T_k or Q and W_k represents errors of real ADC transfer characteristics and usually it is described by so-called integral (INL) or differential nonlinearity (DNL) respectively (IEEE standards). Although the definition of DNL and INL is very simple the measurement of these error parameters is very complex task because of stochastically behaviour or real ADCs. It means that code *k* on the ADC output may stochastically change within a few neighbouring codes even if a DC signal is on ADC input. The most common methods for ADC testing eliminating the ADC stochastic behaviour is histogram method: the real histogram is built from record acquired on ADC output at known in details ADC input stimulus and this is post compared with theoretical histogram calculated for ideal equivalent ADC for the same stimulus. This method is standardised for sinewave stimulus. Because of required extreme quality of sinewave stimulus the generators suitable for testing 16 and more bits ADCs are rather expensive and less common in laboratories. To avoid this constraint other stimulus signals have been proposed for histogram tests [1], [2], [3], et al. Exponential stimulus seems to be one of the best alternative solutions because of its easy generation in on-board ADC self test system (see Fig. 1) ([4]).

The novelty of our approach presented in this paper is determination of stimulus parameters needed for modelling of histogram of ideal equivalent ADC (required for INL estimation) by processing of real measured histogram instead of stimulus record fitting that was used in previous papers ([5], [7]).

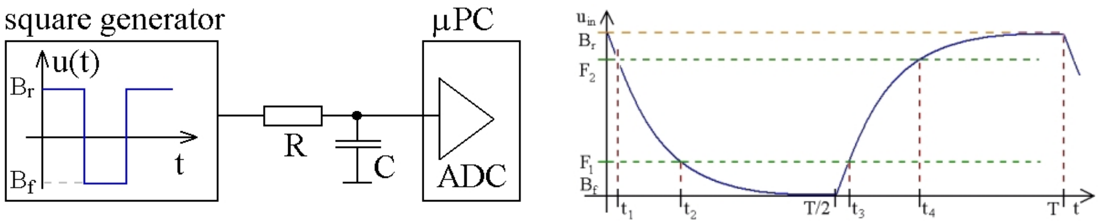


Fig. 1. Application of exponential stimulus in on-board system (left) and shape of exponential signal in time domain (right).

Mathematical representation of such a stimulus is as follows:

$$y(t)=\begin{cases}(F_2-B_f)\cdot e^{-\frac{t-t_1}{\tau_1}}+B_f, & for \quad (t_1<t<t_2)\\ B_r-(B_r-F_1)\cdot e^{-\frac{t-t_3}{\tau_2}}, & for \quad (t_3<t<t_4)\end{cases}. \tag{1}$$

where τ is the time constant of exponential pulse, F_1 and F_2 determines full-scale input range of ADC under test and B_f, B_r are limit values of exponential signal for t → ∞ for each direction (falling and rising) of signal.

2. Exponential stimulus histogram processing

The main goal of processing of measured histogram $H_m(k)$ is to estimate INL of ADC under test. The histogram can be achieved by processing either rising or falling part of stimulus in record. The expressions herein below describe INL estimation for any chosen slope.

To decrease calculation load, the INL can be calculated from recurrent expression:

$$INL(k+1)=INL(k)+DNL(k), \tag{2}$$

Where

$$DNL(k)=\frac{H_C(k)-H_{C_{id}}(k,B)}{H_{id}(k,B)}, \text{ for } k=1, 2, ..., 2^N-2, \tag{3}$$

N is number of bits of ADC under test, INL for the first and last transient level is equal zero, i.e. INL(1) = INL(2^N-1) = 0, $H_C(k)$ is cumulated normalized measured histogram:

$$H_C(k)=\frac{\sum_{i=1}^{k}H_m(i)}{\sum_{i=1}^{2^N-2}H_m(i)}=\frac{H_{mC}(k)}{H_{mC}(2^N-2)}=\frac{H_{mC}(k-1)+H_m(k)}{H_{mC}(2^N-2)}, \text{ for } k=2, 3, ..., 2^N-2, \tag{4}$$

and $H_{id}(k, B)$ is normalised histogram for ideal equivalent ADC. To build $H_{id}(k, B)$ only parameter B has to be estimated. Three formal mathematical expressions of $H_{id}(k, B)$ was published [4] in the past with the goal to simplify determination of B. The recurrent one was chosen to be applied in following because of its mathematical simplicity and using only basic arithmetic operations what is important in on-chip implementation.

The ideal histogram may be expressed ([4]):

$$H_{id}(k,B)=M(B)\cdot\frac{1}{F_1-B+Q\cdot(k-0.5)}, \text{ for } k=1, 2, ..., 2^N-2, \tag{5}$$

where

$$M(B)=\frac{1}{\sum_{i=1}^{2^N-2}\frac{1}{F_1-B+Q\cdot(i-0.5)}}, \tag{6}$$

and Q is nominal code bin width $Q=\frac{F_2-F_1}{2^N-1}$. The value $M(B)$ is independent of k and that is why it can be calculated only once for all values k in $H_{id}(k,B)$.

In general the simple histogram $H_{id}(k,B)$ is very sensitive on superimposed noise and harmonic disturbances in stimulus. That's why the normalized cumulative $H_{Cid}(k,B)$ is preferred for INL estimation:

$$H_{Cid}(k+1,B) = H_{Cid}(k,B) + H_{id}(k+1,B),\ H_{Cid}(1,B) = H_{id}(1,B),\ k=1, 2, \ldots 2^N-3. \quad (7)$$

3. Iteration algorithm for INL estimation

The both histograms $H_{id}(k, B)$ and $H_{Cid}(k, B)$ dependent only on parameter B is required to determine INL of ADC under test. In other words the only parameter that has to be estimated from measured histogram is parameter B. In praxis it can be estimated only numerically because the least mean square fit (LMS) (8) leads to a nonlinear equation that can not be solved analytically.

$$\min(\phi(B)) = \min\left(\sum_{i=1}^{2^N-2}\left(\hat{H}_C(i) - H_{C_{id}}(i,B)\right)^2\right), \quad (8)$$

Where $\phi(B)$ is the cost function of LMS fit. Newton iteration algorithm can be applied to determine B from (8):

$$B_{n+1} = B_n - \frac{\phi'(B_n)}{\phi''(B_n)}. \quad (9)$$

Where $\phi'(B_n)$ and $\phi''(B_n)$ are computed numerically as follows:

$$\phi'(B) = \frac{\phi(B+h) - \phi(B-h)}{2\cdot h},\quad \phi''(B) = \frac{\phi(B+h) - 2\cdot\phi(B) + \phi(B-h)}{h^2}, \quad (10)$$

And h is a constant dependent on number of bits used for representation of numbers in test data processing system, e.g. for double representation (64 bits) the recommended value of h is 10^{-3} - 10^{-4}. The iteration final condition is $\varepsilon > |B_{n+1} - B_n|$, where ε is approximation residual uncertainty. The resulting B is consequentially applied for estimation of INL of ADC under test according the procedure described hereinabove in chapter 2.

4. Experimental results

The proposed algorithm was verified by simulated measurement on simulated ADC with nominal modelled INL. Fig. 2 shows the nominal INL and INL calculated from record on the output of the modelled simulated ADC under test by the suggested algorithm for ideal exponential stimulus.

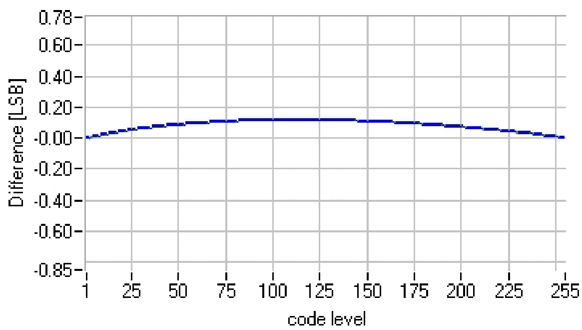


Fig. 2. The modelled (dark) and measured (light) INLs of simulated ADC (the left graph) and difference between them (the right graph).

As it can be seen from the figure there are only little differences between modelled and measured INL. The right graph shows differences of both INLs that can be supposed to be error of measurement.

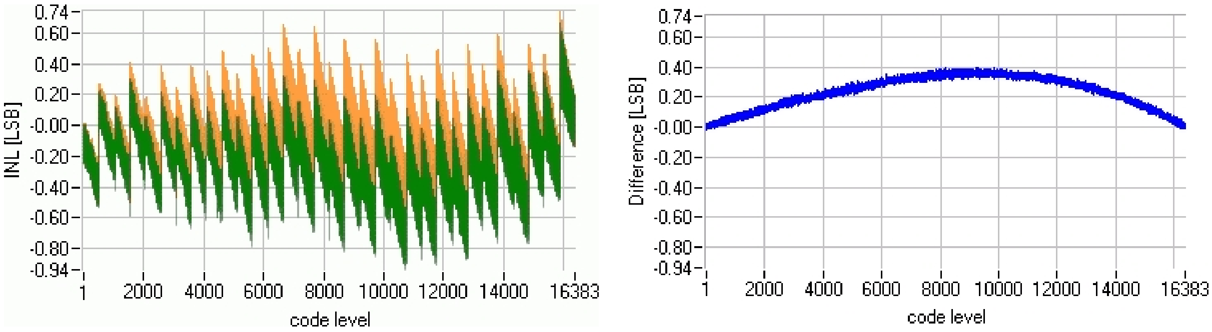


Fig. 3. INL of USB6009 obtained from standardized harmonic stimulus histogram test (dark) and from exponential stimulus histogram test (light) (left) and difference between them (right).

The proposed algorithm was also verified by real experimental measurement. The ADC under test was 14-bit ADC implemented in USB6009 device by National Instruments. The ADC was tested by standardized histogram test method with harmonic stimulus was used (dark curve in Fig. 3 left) as well as by algorithm described hereinabove (light curve in Fig. 3 left). The difference of INLs was calculated and it is shown in Fig. 3 (right).

5. Conclusions

The paper presents a new approach to the ADC INL testing methodology by simple monotonic exponential stimulus which data processing algorithm was suggested with targeting simple implementation in ADC on-board self-testing systems. The method was verified by simulations and experimental measurements in comparison with standardised sinewave histogram test that confirmed applicability of the method. The residual error of measurement can be caused by INL low code frequency component masking [6] if it has nature similar to the shape of input exponential stimulus.

Acknowledgements

The work is a part of project supported by the Science Grant Agency of Slovak republic (No. 1/0103/08).

References

[1] Haasz V., Komárek M., Roztočil J., Slepička D., Suchánek P.: *System for testing middle-resolution digitizers using test signal up to 20 MHz*, IMTC/06 Proceedings of the 23rd IEEE IMT Conference,(2006), p. 266-270. ISBN 0-7803-9360-0.

[2] Fodor B., Kollár I.: *ADC testing with verification* , IEEE Transactions on Instrumentation and Measurement 57 (12), (2008), pp. 2762–2768.

[3] Björsell N., Suchánek P., Händel P., Rönnow D.: *Measuring Volterra kernels of analog-to-digital converters using a stepped three-tone scan*, IEEE Transactions on Instrumentation and Measurement 57 (4), (2008),pp. 666–671.

[4] Michaeli L., Šaliga J., Sakmár M., Buša J.: *New approximation models for processing exponential stimulus histogram test of ADC*, Proc. of 16th IMEKO TC-4 conf.: Florence, Italy, September 22–24, 2008, pp. 1043-1047.

[5] Holcer R., Michaeli L., Šaliga J.: *DNL ADC Testing by the Exponential Shaped Voltage*, IEEE Trans. on Instrumentation and Measurement, Vol. 52, No. 3, June 2003, 946–949.

[6] Michaeli L., Michalko P., Šaliga, J.: *A new ADC fast testing method based on the unified error model*, Measurement. Vol. 41, No. 2 (2008), ISSN 0263–2241.

[7] Vedral, J.: Exponential Fit Test - *Theoretical analysis and practically implementation*, Proc. of IMEKO TC-4 conf., Florence, Italy, September 22–24, 2008, pp. 1033 – 1036.

Evaluation of Calibration of a Thermocouple

P. Benkó, R. Palenčár

Faculty of Mechanical Engineering, Bratislava, Slovakia
Email: peter.benko@stuba.sk

Abstract. *This contribution describes the procedure of evaluating the calibration of thermocouple by means of its comparison with the thermocouple standard. In the process of thermocouple calibration by means of comparison, the resulting uncertainty specified by applying the generalized procedure for evaluating the calibration of measuring devices with continuous scale.*

Keywords: Calibration, Thermocouple, Estimation of Unknown Parameters, Uncertainties, Covariance

1. Introduction

For measuring instrument with continuous scale a generalized procedure for evaluating the calibration uncertainties and covariances has been developed by Palenčár, Wimmer [1,2] and Kubáček [4]. In this paper authors are presenting these procedures for evaluating uncertainties of the calibration of a thermocouple (hereafter TC only) type S by means of comparison.

2. Calibration procedure

Calibration is carried out by comparison of the unit under test TC type S against standard TC type S calibrated in defined fixed points according to ITS-90 (Fig.2.1). The calibration is represented as a curve fitted to the measured values of the deviation E-E_{ref} and generally given as a function of temperature t. This curve is representing deviation function.

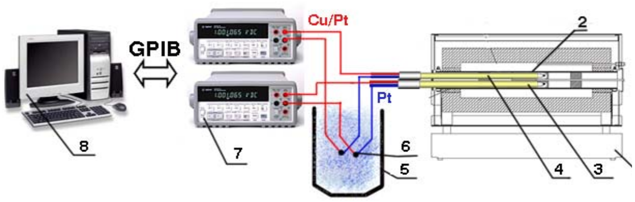


Fig. 2.1: Scheme of calibration

1- Calibration furnace, 2- Isothermal block, 3- Standard TC, 4- Unit under test, 5- Dewar flask, 6- Reference junction of TC's, 7- Voltmeters, 8- Computer with GPIB port

Table 2.1: Measured and computed values i- calibration points (nominal values), $t(\overline{E}_{\mathrm{S}_i})\,(°\mathrm{C})$- values measured by standard TC, $\overline{E}_{\mathrm{K}_i}\,(°\mathrm{C})$- values measured by unit under test TC

Cal. Points i	$t(\overline{E}_{\mathrm{S}_i})\,(°\mathrm{C})$	$\overline{E}_{\mathrm{K}_i}\,(°\mathrm{C})$
100	99,8188	643,85
200	199,7252	1437,82
300	299,7120	2319,79
400	399,7737	3255,32
500	499,8191	4228,52
600	599,6689	5230,63
700	699,7653	6265,91
800	799,8341	7334,86
900	899,7653	8437,38
1000	999,5335	9569,7
1100	1099,4	10736,91

3. Methodology

We consider the case, when number of calibration points r is higher than number of unknown parameters p, $r>p$ the model is overdetermined. Calibration model should be established using following relations

$$W_i = a_0 + a_1 \cdot t_i + a_2 \cdot t^2{}_i + a_3 \cdot t^3{}_i + a_4 \cdot t^4{}_i \qquad i=1, ..., n \tag{3.1}$$

in matrix notation

$$\boldsymbol{W} = \boldsymbol{Ta} \tag{3.2}$$

where $\boldsymbol{T}$ is a matrix, which contains values, arithmetical means of series of measurements in each calibration points measured by standard TC.

Left side of the model (3.1) or (3.2), the observation vector $\boldsymbol{W}$ is presenting the measurement model of unit under test TC

$$\boldsymbol{W} = \boldsymbol{\Delta E} + \boldsymbol{C}_\mathrm{K}\boldsymbol{\Lambda} \tag{3.3}$$

where $\boldsymbol{\Delta E}$ is the vector of deviations from the reference function. Reference function is given by IEC 584.2 standard [3].

$$\boldsymbol{\Delta E} = \overline{\boldsymbol{E}} - \boldsymbol{E}_\mathrm{ref} \tag{3.4}$$

in product of $\boldsymbol{C}_\mathrm{K}\boldsymbol{\Lambda}$ fills every influences of measurement.

Vector of correction $\boldsymbol{\Lambda}$ is given by and matrix $\boldsymbol{C}_\mathrm{K}$ is the known matrix, usually its elements are sensitivity coefficients.

Our aim is to get estimation for unknown parameters of deviation function. This aim could be reached by using least-square method [1,2,3]. Uncertainties are taken into account as well. We apply following expression iteratively because of stochastic character of quantity t [1].

$$\hat{\boldsymbol{a}} = \left(\boldsymbol{T}^\mathrm{T}\boldsymbol{U_W}^{-1}\boldsymbol{T}\right)^{-1}\boldsymbol{T}^\mathrm{T}\boldsymbol{U_W}^{-1}\boldsymbol{W} \tag{3.5}$$

Initial values of unknown parameters $\hat{\boldsymbol{a}}$ of deviation function are determined by zero estimation. Then covariance matrix of input quantities $\boldsymbol{U_W}$ is

$$\boldsymbol{U_W} = \boldsymbol{U}_{\boldsymbol{\Delta E}} + \boldsymbol{C}_\mathrm{K}\boldsymbol{U}_{\boldsymbol{\Lambda}}\boldsymbol{C}_\mathrm{K}{}^\mathrm{T} \tag{3.6}$$

where

$\boldsymbol{U}_{\boldsymbol{\Delta E}}$ - covariance matrix of the vector $\boldsymbol{\Delta E}$ is diagonal matrix, principal-diagonal elements present square of uncertainties estimated by type A method

$\boldsymbol{C}_\mathrm{K}\boldsymbol{U}_{\boldsymbol{\Lambda}}\boldsymbol{C}_\mathrm{K}{}^\mathrm{T}$ - product of these matrix gives diagonal covariance matrix, principal-diagonal elements present square of uncertainties estimated by type B method

$\boldsymbol{U}_{\boldsymbol{\Lambda}}$ - uncertainties of correction measurement by unit under test TC are included in this covariance matrix

Covariance matrix $\boldsymbol{U}_{\hat{\boldsymbol{a}}}$ is represented by matrix of the uncertainties of the estimates

$$\boldsymbol{U}_{\hat{\boldsymbol{a}}} = \left(\boldsymbol{T}^\mathrm{T}\boldsymbol{U_W}^{-1}\boldsymbol{T}\right)^{-1} \tag{3.7}$$

Deviation associated with the reference function is solved by

$$\boldsymbol{\Delta}\hat{\boldsymbol{E}} = \boldsymbol{T}\hat{\boldsymbol{a}} \tag{3.8}$$

uncertainty of the deviation can be achieved by application of law of propagations of uncertainties

$$u_{\Delta \boldsymbol{E}}^{2} = \boldsymbol{T} \cdot \boldsymbol{U}_{\hat{\boldsymbol{a}}} \cdot \boldsymbol{T}^{\mathrm{T}} \tag{3.9}$$

Zero estimation of vector $\hat{\boldsymbol{a}}$ is biased (see Fig.4.1(a)). It is caused by stochastic characters of the quantity t. Therefore the model is nonlinear and requires a solution procedure. It is linearized by application of Taylor series and higher elements of estimated values are neglected. After linearization left side of model vector $\boldsymbol{W}$ will be

$$\boldsymbol{W} = \Delta\boldsymbol{E} + \boldsymbol{C}_{\mathrm{K}}\boldsymbol{\Lambda} + \boldsymbol{D}\left(\boldsymbol{\delta t}_{1} + \boldsymbol{C}_{\mathrm{S}}\boldsymbol{\delta t}_{2}\right) \tag{3.10}$$

$\boldsymbol{D} = \mathrm{diag}\left(d_{100} \quad d_{200} \quad \dots \quad d_{1100}\right)$ - is the known matrix, obtained by application of expansion of Taylor series

After linearization covariance matrix $\boldsymbol{U}_{\boldsymbol{W}}$ has the form

$$\boldsymbol{U}_{\boldsymbol{W}} = \boldsymbol{U}_{\Delta \boldsymbol{E}} + \boldsymbol{C}_{\mathrm{K}}\boldsymbol{U}_{\boldsymbol{\Lambda}}\boldsymbol{C}_{\mathrm{K}}^{\mathrm{T}} + \boldsymbol{D}\left(\boldsymbol{U}_{\boldsymbol{\delta t}_{1}} + \boldsymbol{C}_{\mathrm{S}}\boldsymbol{U}_{\boldsymbol{\delta}t_{2}}\boldsymbol{C}_{\mathrm{S}}^{\mathrm{T}}\right)\boldsymbol{D}^{\mathrm{T}} \tag{3.11}$$

where

$\boldsymbol{U}_{\boldsymbol{\delta t}_{1}}$ - covariance matrix of the vector $\boldsymbol{\delta t}_{1}$ is diagonal matrix, principal-diagonal elements present square of uncertainties estimated by type A method

$\boldsymbol{C}_{\mathrm{S}}\boldsymbol{U}_{\boldsymbol{\delta}t_{2}}\boldsymbol{C}_{\mathrm{S}}^{\mathrm{T}}$ - product of this matrix is given by diagonal covariance matrix, principal-diagonal elements present square of uncertainties estimated by type B method

$\boldsymbol{U}_{\boldsymbol{\delta}t_{2}}$ -uncertainties of correction of measurements by standard are included in this covariance matrix

Now in new iteration we consider the observation vector $\boldsymbol{W}$ (3.10) and covariance matrix $\boldsymbol{U}_{\boldsymbol{W}}$ (3.11) and we use formula for estimation of parameters (3.5).

Numerically, in the most cases design matrix $\boldsymbol{T}$ is badly scaled and its columns are nearly linearly dependent. For this it is reasonable to transform quantities of t to interval $-1 \le t \le 1$

From the viewpoint of the user relevant results are the temperature values and their uncertainties. Temperature value can be obtained by interpolation table which can be edited from deviation function and its uncertainty is determined by application of theorem for implicit function.

$$f\left(E,t,a\right) = E - g\left(t,\boldsymbol{a}\right) = 0 \tag{3.12}$$

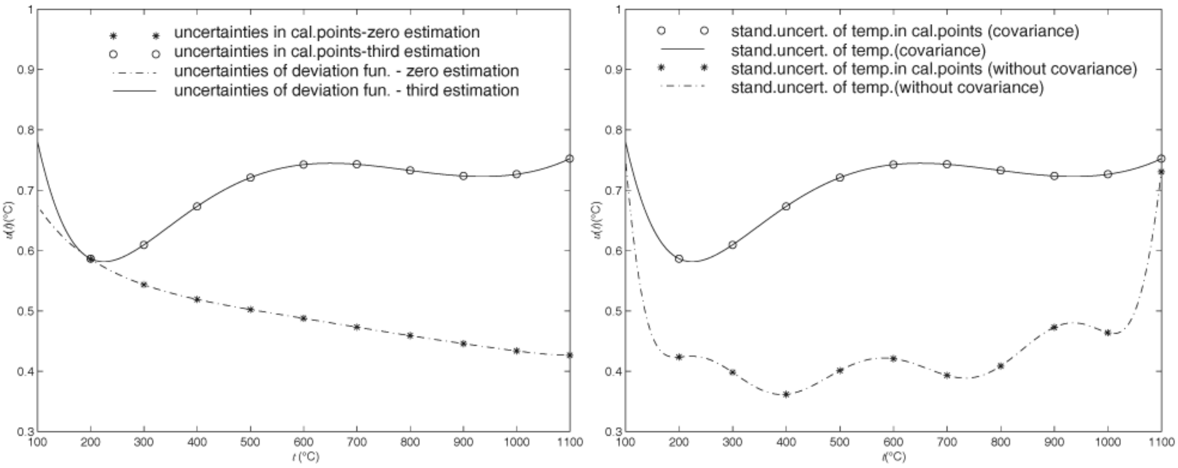
we get it by adding up deviation function and reference function, where variable E is representing the current measured value of emf. Now consider function t=(h,$\boldsymbol{a}$) is defined from the implicit function.

Standard uncertainty is then obtained from the (3.13) relation

$$u^{2}\left(t\right) = \boldsymbol{h}^{\mathrm{T}} \cdot \boldsymbol{U}_{\hat{a}} \cdot \boldsymbol{h} \tag{3.13}$$

4. Conclusion

Procedure for evaluating the calibration of TC was applied to demonstrate whether considering the covariances has an impact on final result of standard uncertainty. For this reason was carried out the evaluation twice. The difference is shown in Fig. 4.1(b). As a conclusion we can claim that covariances had significant effect on final result of a calibration.



(a) (b)

Fig 4.1: Standard uncertainties of deviation function: (a) Difference between zero and third estimation of parameters, (b) Standard uncertainties derived from third estimation when consider covariance and not

Acknowledgements

This work was elaborated by Peter Benkó during his PhD study at the Faculty of Mechanical Engineering in Bratislava. Peter Benkó wishes to thank Prof. Rudolf Palenčár for introducing him the estimation theory of mathematical statistics.

Authors thanks to Faculty of Mechanical Engineering in Bratislava and Grant agency VEGA n. 1/0310/09 for their support.

References

[1] Wimmer G., Witkovský V., Palenčár R.: Spracovanie a vyhodnocovanie meraní, VEDA, Bratislava 2002, 80-224-0734-8.

[2] Palenčár R., Ruiz J., Janiga I., Horníková A.: Štatistické metódy v skúšobných a kalibračných laboratóriách, Grafické štúdio Ing.Peter Juriga, Bratislava 2001, ISBN 80-968449-3-8.

[3] Lira I.: Evaluating measurement uncertainty: Fundamentals and practical guidance, IoP Bristol, ISBN 0-7503-0804-0.

[4] Kubáček L.: Statistika a metrologie, Vydavatelství Univerzita Palackého v Olomouci, 2000.

The Evaluation of Accuracy of Measurement Results in Medical Analytical Laboratory

P. Otomański, A. Odon

Institute of Electrical Engineering and Electronics, Division of Metrology and Optoelectronics, Poznan University of Technology, Piotrowo street 3a, 60-965, Poland

Email: otoman@et.put.poznan.pl, odon@et.put.poznan.pl

Abstract. *Problem of the reliability estimation of measurement methods used to medical diagnostic laboratory, given the example of blood morphology, was presented in the present paper. Measurement method was presented and statistical methods used to evaluation of measurement results were discussed.*

Keywords: Accuracy of Measurement Results, Standard Deviation, Blood Cells Counting

1. Introduction

The result of every measurement is errored, i.e. it contains an error the value of which depends on the accuracy of the applied measurement method. According to [1], the measurement error is a difference of the measurement result and the true value of a measurand. The basic problem in the estimation of measurement error is the fact that we never know the exact true value of a measurand. When we carry out a series of measurements, it is necessary to carry out also a statistical analysis of results obtained with the use of appropriate mathematical procedures. When carrying out a research or test, we should be certain that the applied measuring instrument gives reliable results. The instrument may be checked using a standard, or it may be calibrated by a series of quantities with known yet different values. In terms of medical diagnostic laboratory, a series of control determinations is done every day, and their aim is to check a series of parameters, among them first of all the reproducibility of the results of measuring the same quantity in time. Reproducibility should be understood as the degree of accordance of the results of measuring the same measurand in different measuring conditions. For intra- and interlaboratory control, additional samples with precise value of the examined parameter are used in repeated series of routine determinations. Correct, regular results of this measuring process are significant part of the procedure of laboratory accreditation. This publication is an attempt to take a penetrating look – from the metrological point of view – inside the applied methods and criteria of evaluating the analytical phase accuracy in medical analytical laboratory.

2. The Methods of classifications and counting of blood cells

Measuring equipment used in medical diagnostic lab for blood tests should make it possible to select blood cells and count their number in a unit of 1 mm^3 volume. Blood cells can be divided into three types: erythrocytes, leukocytes and thrombocytes. Regardless of the method applied for measuring, a blood sample is usually diluted in physiological saline, and then a precisely measured volume of the so prepared blood sample is transmitted through a tube with a very small diameter, ca 100 μm; thanks to that, cells in the stream of the diluted sample are arranged one after another (hydrodynamic focusing). The stream of blood sample passes then through the measuring region, in which a detection system detects successively passing blood cells, and transmits information to a data processing and acquisition system. Blood dilution is an indispensable technical operation because, otherwise, more than one cell could pass through the measuring region simultaneously, which causes counting errors.

There are mainly four methods for counting and classification of blood cells. The hemacytometer method - a manual counting method, the photoelectric nephelometric method, can be used only to count normal red blood cells. The other two available methods for blood cell counting and classification use either a Coulter counter or a flow cytometer [2]. They measure single cell flowing through the measuring region based on the electrical impedance, the so-called Coulter principle, or the laser scattering principle respectively.

The resistance method (Coulter's method) is developed to count the blood cells and classify their sizes. It is based on measuring variations in the resistance generated by non-conducting particles, diluted in electrolyte. Fig. 1 shows a simplified diagram illustrating the principle of measuring the number of blood cells with the use of the resistance method. A diluted blood sample moves from a bigger vessel through an aperture with ca 100 μm diameter to a tube in which a hypotension pump is installed, which results in the effect of sucking in the solution from the vessel. There is one electrode in the vessel with the diluted blood sample, whereas the other one is located inside the tube. There is a current generator added to the electrodes, which makes the current pass through the solution. At the moment when a blood cell appears in the aperture region, conductance rapidly decreases and, as a result, a voltage pulse is generated between the electrodes. A detection system with properly set value of actuation threshold allows different types of blood particles to be discriminated. The optic method uses the property that a blood cell placed in liquid medium has a different light absorption coefficient than the solution in which it is placed. Blood cells move in the tube with a very small diameter and pass through the measuring region illuminated by a light source and observed by a photodetector. At the moment when a blood cell crosses the light beam, a voltage pulse is generated at the output of the photodetector [3]. There are a few reasons causing measuring errors during the classification and counting of cells; the most important of them are the following: contamination, incorrect dilution of samples, reduction of the aperture diameter, caused by contamination deposition, incorrect detection threshold. It should be emphasised that even with complete idealization of the measuring apparatus, the analysed biomaterial – because of its heterogeneity – causes a scatter of measuring results, therefore, statistical analysis of the measuring results is a condition sine qua non.

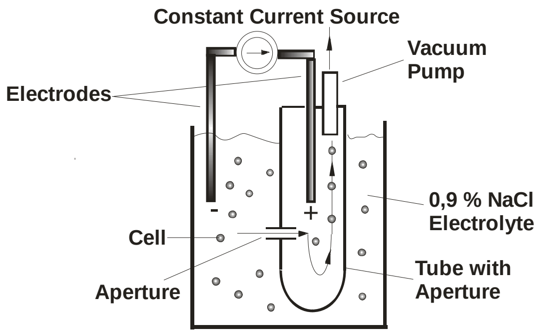


Fig. 1. A scheme illustrating the principle of blood cells measurement using the resistance method

3. Reliability of measurements in medical analytical laboratory

Measurement results obtained in diagnostic lab require an interpretation that makes it possible to determine if the inspected method is accurate and the result reliable. In the classical measuring theory two basic parameters are used (their role is to estimate the distribution quality): an expected value, the estimator of which is most often the mean value of population, being the concentration measure of a random variable, as well as the standard deviation, being the scatter measure. In diagnostic labs, due to a specific character of the examined biomaterial, and also due to a specific character of technical solutions in measuring apparatus, we use any method for estimating the measurement reliability in a modified form – as compared to standard recommendations. In order to determine the quality of a given measurement method, we use the knowledge of the values of three parameters: arbitrarily matched total allowable measuring error, inaccuracy and bias.

Total error allowable – TE_A – is an acceptable difference between the obtained result and the true value of a measurand. This is an idea applied in medical diagnostic labs to determine the requirements that should be met by the used measuring method, so that it could be regarded as reliable [4]. The value of TE_A makes it possible to determine such a confidence interval in which with assumed probability the obtained measuring results are located. The determination of concrete values of this parameter for particular measurands allows us to decide if a given method, with its inaccuracy and bias, is, nevertheless, a reliable method. Because of random errors that accompany every measurement, the results obtained during repeated measuring of the same sample vary. The degree of accordance among independent measurement results in the literature concerning quality analysis in medical diagnostic laboratory, e.g. in [4], is called precision. The quantity coefficient of inaccuracy is the value of standard deviation, or coefficient of variation defined as a ratio of standard deviation and percentage mean value. The correctness of measurements, as a quality factor, is described by means of a number which is the difference between the mean value and assumed reference value. According to the ISO 3534-1 document, quantity information about correctness, presented in the form of that difference, is defined as bias [4]. In practice, as the reference value we most often assume the mean value of a sample of a specified population of measurement results, and determined e.g. in the model sample of control material designed for inter-laboratory control.

One of the most widespread methods of statistical quality control in medical diagnostic lab is the method applying a control chart, on which particular measurement results are plotted; it is known as the Levey-Jennings chart. This method applies control material as a sample which is analysed in order to carry out quality control. The most relevant question is the interpretation of obtained measurement results by the determination of such limits for the allowable error that cannot be exceeded if the method should not be regarded as out-of-control. In this respect, estimations vary in answering the question if the investigated method is within allowable limits or not. We can distinguish here simple rules and complex rules. The most simple way to estimate control results is the method based on a simple interpretative rule, e.g. rules $1_{2.5S}$, where S means standard deviation. In this rule we assume that the method remains out-of-control in the situation when at least one control result obtained in a measurement exceeds the limits ± 2.5S. In the respective literature we can also find rules 1_{2S}, 1_{3S}, $1_{3.5S}$, that differ from one another in the width of an interval in which the measurement results should be contained. The interpretation of the results of control measurement using complex rules consists in applying a few rules simultaneously, which makes it possible to improve the effectiveness of estimating the control results. One of the most well-known and widespread ones is rule $1_{3S}/2_{2S}/R_{4S}/4_{1S}/10_X$, which has taken its name – Westgard rules – from the name of its main author. The algorithm of this method is shown in Fig. 2. In order to evaluate the reliability of measurements, some experiments were carried out with the use of the Levey-Jennings control chart. According to recommendations, the research was conducted for 20 days, with Sysmex XS-1000i analyser, using thrombocythes as the analyte. Examples of research results are presented in the form of control chart in Fig. 3. The mean value as well as the limits of allowable variation interval determined according to rule12.5S was marked in the figure. As it can be noticed, one of the measurements exceeds the limits of the assumed variation range, which allows us to formulate the conclusion that the measurement method applied is out-of-control. Applying for the same series of results the complex Westgard rule, we should conclude that this measurement method is correct. A variety of possible methods of interpreting measurement results evokes a fundamental question: Which option will be the most suitable? Which rule should we choose, a simple or a complex one? In the authors' opinion, there is no unequivocal answer to such question. An extremely significant part in this situation is played by an experienced person who supervises the research and interprets the obtained measurement results.

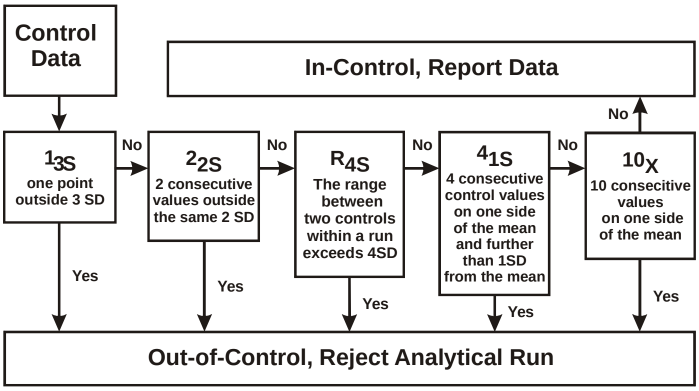


Fig. 2. The modern Westgard Rules $1_{3S}/2_{2S}/R_{4S}/4_{1S}/10_X$

4. Conclusions

The paper presents selected problems of the evaluation of the reliability of measurement methods applied in diagnostic laboratories. The authors discussed the applied measurement methods developed to measure blood cells as well as main sources of measuring errors in these methods. They also pointed to the specific approach to evaluating the accuracy of any measurement method applied in such laboratory, as compared with the classical approach and mathematical formal solutions, known from the measurement theory. The present paper indicates only a modest fragment of a whole area of problems relating to the statistical estimation of the results of measuring difficult objects, such as biological objects. One of the purposes of this work is to bring attention to the fact that measuring practice in diagnostic laboratory develops for its own needs specific principles of the estimation of measurement results. These principles on the one hand prove to be quite suitable in practice, and on the other hand they not necessarily correspond precisely with standards recommended in metrology.

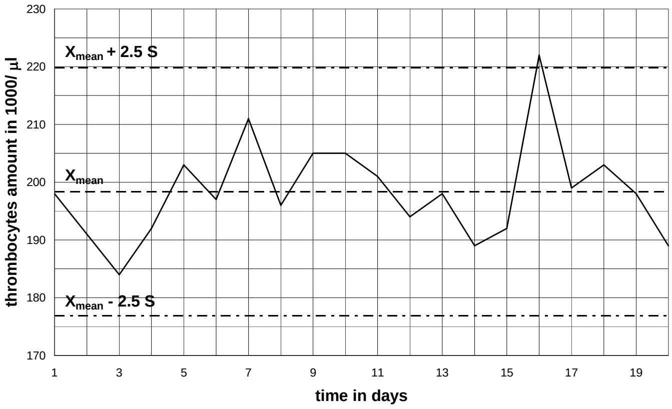


Fig. 3. Levey – Jennings control charts with an example of thrombocytes analysis

References

[1] International Vocabulary of Basic and General Terms in Metrology, International Organization for Standardization, second edition, 1993.

[2] Chiappini D. et al.: Comparability of hematology profile between QBC (R) autoread (M) plus system and all electronic cell counting instrument (Coulter Counter STKS), *Schweiz Med. Wochenschr* 130, (15), 2000, pp. 545–550.

[3] Yang Ye et al.: Blood cell counting and classification by nonflowing laser light scattering method, *Journal of Biomedical Optics* 9(5), September/October 2004, pp. 995–1001.

[4] Gernand W., Fundamentals of quality control of laboratory tests, Centrum Promocji Nauk Medycznych, Lublin, 2000, in Polish.

Image Object Recognition Based on Biologically Inspired Hierarchical Temporal Memory Model and Its Application to the USPS Database

[1]S. Štolc, [2]I. Bajla

[1]Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovakia,
[2]Austrian Research Centers GmbH – ARC, Smart Systems Division, Seibersdorf, Austria
Email: umerstol@savba.sk

Abstract. *In the paper we describe basic functions of a Hierarchical Temporal Memory (HTM) based on a novel biologically inspired network model of the overall large-scale structure of human neocortex. It appeared in a form of research release of the system NuPIC (Numenta Platform for Intelligent Computing) in 2007. In the design of the HTM, hierarchical structure and spatio-temporal relations serve for generation of invariant representations of the outer world (e.g. world of visual patterns), similar to those in human neocortex. There are several open issues for a research into HTM, in particular those applied to pattern recognition tasks. In the paper we report our results of the HTM architecture design and optimization of the network parameters for the task of recognition of the handwritten digits from the well benchmarked USPS database.*

Keywords: Hierarchical Temporal Memory (HTM), Pattern Recognition, USPS

1. Introduction

The HTM is an uncommon, neocortex inspired computational theory [1] that appeared in a form of research release of the software NuPIC 1.6. (Numenta Platform for Intelligent Computing) in 2007 [2]. As an original memory-prediction theory, it has a potential to mature in the future into the state capable to solve also various problems of pattern recognition. In the design of the HTM, hierarchical structure and spatio-temporal relations serve for generation of invariant representations of the world (in our domain restricted to images), similar to those in human neocortex. The functioning of biological regions and subregions is simulated in the nodes (basic units of the HTM) using Bayesian belief revision techniques. The basic difference of the HTM to the neural networks consists in providing a model of the overall large-scale structure of human neocortex. There are several open issues for a research into HTM, especially those applied to specific Pattern Recognition (PR) tasks. In the paper we report our results of HTM architecture design and optimization of the parameters of the individual HTM-nodes for the task of recognition of the handwritten digits of the internationally accepted USPS database. This choice is based on the fact that the recognition accuracy achieved for various classifiers applied to this database is well benchmarked in the literature, and it can be compared to the results of our research into object detection using vector subspace methods, in particular non-negative matrix factorization methods [3].

2. Hierarchical Temporal Memory

In [4,5], the general concept, theory, as well as terminology of the HTM is described. Our interests were focused on research into optimal design of such an HTM that can be used for solving image recognition tasks, therefore we will briefly describe the basic functions of the HTM implementation in relation to modelling the visual world. The HTM is a memory-prediction network that is organized in several layers of elementary units – nodes working in identical mode. The individual layers (levels) are ordered in a hierarchical tree-like structure.

There is a zero sensory level of the HTM which serves as an input to the first level of nodes. In our case zero level (ImageSensor) represents a visual field of image pixels. Since the use of temporal dependences of input spatial patterns is essential characteristic of the HTM, it learns either from natively moving images or sequences of image frames of an artificially generated movie (obtained by applying a limited set of translations, rotations and zoomings to the given training images). It is the explorer plug-ins that generate the temporal sequences within the ImageSensor object. They are responsible for "exploring" the input space of possible images accomplishing two main goals: (i) efficiently select images for presentation to the network, out of a large space of potential images, (ii) generate smooth temporal sequences needed for training TemporalPoolerNode. When the learning process is finished in all levels, the HTM network can classify an unknown pattern into previously defined classes. For every input, the node does three learning operations: (i) memorization of the input vectors, (ii) learning transition probabilities, (iii) temporal grouping.

The memorization of the pattern vectors is carried out in the spatial pooler that actually generates spatial statistical representations of input vectors (patterns). More specifically: at the 1st hierarchy level, the spatial pooler of each node detects clusters of pattern vectors occuring in the field of view of this node in the course of training. Each cluster is memorized by means of a centroid representative. At higher levels of the HTM hierarchy, the spatial clustering algorithm is applied to belief vectors which are input for higher level spatial poolers. Once the memorization process is finished, the spatial pooler can continue with the next step of the learning process. During this stage, for every input pattern its closeness to every vector stored in the node memory [6] is measured by Euclidean distance d_i of the pattern vectors. It is assumed that the probability that the input pattern matches the *i*-th stored pattern vector can be calculated as being proportional to the Gaussian function $e^{-d_i^2/\sigma^2}$ of the distance d_i, where σ is a parameter of the node.

The learning in the temporal pooler is characterized as follows. First, a time adjacency matrix for the pattern vectors is generated, entries of this matrix are numbers of transition events between vectors following each other during image movement in the field of view of the node (the rows and columns of the adjacency matrix represent memorized pattern vectors – coincidences – of the given node. Second, for temporal grouping of the pattern vectors (level 1) or beliefs (other levels), the Agglomerative Hierarchical Clustering is used. In contrast to the clusters generated in the spatial poolers, these clusters reflect temporal dependences and therefore they are called temporal groups. Each temporal group can be seen as an invariant representation of the patterns included in this group.

The topmost HTM level is constituted by the only node – supervised classifier (in our case Zeta1TopNode) [2] – in which a traditional supervised grouping of a network input is carried out based on the corresponding beliefs produced by preceding HTM levels.

3. Application to USPS database

Train and test sets

For the purpose of testing the performance of the HTM model in comparison with other classification approaches, we have decided on the standard USPS (U.S. Post Service) database of handwritten digits collected by CEDAR, Buffalo [7] and later on converted to gray level format by LeCun's research group [8]. The USPS database consists of 9298 digits of 16x16 pixels each which are divided into two non-overlapping groups: 7291 digits for training and 2007 digits for testing. The most pronounced advantage of using this data set is that a vast number of benchmarks has been performed for different classification methods.

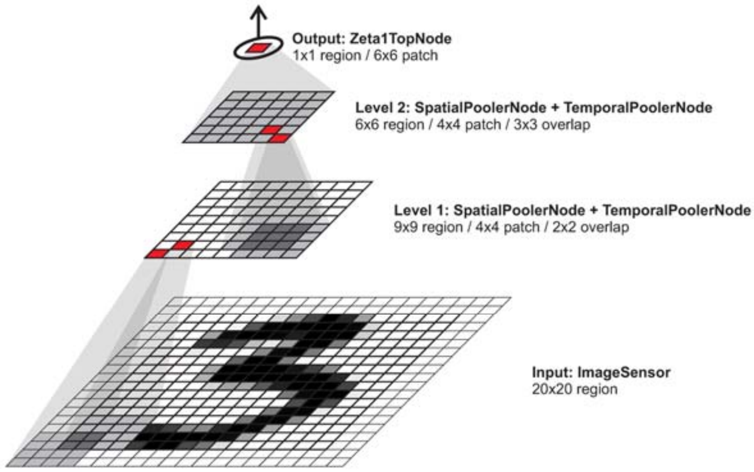


Fig. 1. In the picture 2-level HTM architecture is depicted. The network consists of ImageSensor (input), two levels comprising both SpatialPoolerNode and TemporalPoolerNode regions (level 1 and level 2), and Zeta1TopNode (output).

Optimization of the HTM network architecture

We have experimented with various types of 1, 2 and 3-level architectures where multiple parameters had to be tuned. Let us focus on the specific network architectures which showed the best overall recognition accuracy. The basic assumption originating from the USPS database itself and holding for all our network architectures is that every image passed to the network must have the same size of 16x16 pixels. To avoid undesired border effects, the input images are padded by 2-pixel border while getting the dimension of 20x20 pixels. We have found the following architecture optimal for the given task: a 2-level HTM network that consists of ImageSensor (input), two levels comprising joined SpatialPoolerNode and TemporalPoolerNode regions (level 1 and level 2), and Zeta1TopNode (output). The nodes at level 1 are arranged to a 9x9 array such that they cover full extent of the ImageSensor (20x20 pixels). Each level 1 node receives input from a 4x4 patch while two neighboring patches overlap by 2 nodes. The nodes at level 2 form a 6x6 array while each node receives input again from a 4x4 patch with 3 node overlap. The output Zeta1TopNode receives input from all nodes at level 1 (see Fig. 1).

There are two parameters of each SpatialPoolerNode: *maxDistance* and σ of the Gaussian which are to be optimized at every network level (identical values use all nodes at one level). The TemporalPoolerNodes at every level have only one parameter: *requestedGroupCount* that should be optimized. HTM documentation does not prescribe or include any recommendation or method to be applied for parameter optimization in a particular HTM application. The decision is left on the user. We have decided to use a mixture of two approaches: some of the parameters were estimated using perpendicular-search method [9], whereas other parameters were estimated according to the experience or recommendations from Numenta discussion forum. The perpendicular-search method is based on the principle of searching for the minimum of error function along individual parameters, however, in each iteration the parameter is altered only within the restricted area around the starting point. Best obtained parameter combination is then chosen as the starting point for the next iteration. This means that only one parameter is varied at a time while all other parameters keep the same value.

A key capability of the ImageSensor, which reads data from image files and hands it off to nodes in an HTM network [6], is the ability to generate smoothly-varying patterns forming “virtual” temporal sequences (movies). The ExhaustiveSweep explorer is one of the most common explorers implemented within NuPIC 1.6.1. This explorer performs an exhaustive raster scan through the input space. It can generate rather complex sequences by translating an image side-to-side either horizontally or vertically always by one pixel. The sequences generated are rather long and therefore imply highly memory and time demanding training phase. Since the USPS database contains normalized patterns only (i.e. numbers are always centered and their size is normalized to 16x16 pixels), it is not necessary to build up a positional and dimensional invariance for this data. The ExhaustiveSweep explorer appeared to be inappropriate in this case mainly due to its high computational costs. We have developed an alternative explorer which better meets our expectations involved by the USPS database. Basic idea, on which this explorer is inspired, is a way of how humans are seeing the letters while reading a text. When eyes are moving through the text lines, each single symbol is being seen from different viewing angles. The symbol in the top-left corner of a page looks slightly different than the same symbol (i.e. same character, font and size) in the center or bottom-right corner of a page. The ViewAngleSweep explorer tries to imitate this behavior by smooth alternating the viewing angle in 9 fixed values which form the final temporal sequence presented to the network.

4. Results

The results of the design of suboptimal HTM network for the application to the USPS handwritten digits database can be summarized as follows. The 2-level architecture, characterized in Fig. 1, has been proposed. The following optimum values of the adjustable network parameters have been found:

- Level 1:
    - SpatialPoolerNode: *maxDistance* = 250, σ = 500;
    - TemporalPoolerNode: *requestedGroupCount* = 80;
- Level 2:
    - SpatialPoolerNode: *maxDistance* = 0.3, σ = 1;
    - TemporalPoolerNode: *requestedGroupCount* = 950.

All the nodes in the individual HTM levels have been learned using a special ViewAngleSweep explorer that we developed. We have achieved the overall classification

Table 1. Confusion matrix for 2-level HTM architecture.

True \ Class	0	1	2	3	4	5	6	7	8	9	Accuracy
0	355	0	2	0	0	0	0	1	1	0	98.86 %
1	0	259	0	0	2	0	2	1	0	0	98.11 %
2	1	0	190	0	1	0	0	3	3	0	95.96 %
3	1	0	0	155	0	6	0	1	2	1	93.37 %
4	0	2	0	0	185	1	1	6	0	5	92.50 %
5	1	0	1	1	0	154	0	0	1	2	96.25 %
6	1	0	0	0	0	1	166	0	2	0	97.65 %
7	0	1	1	0	1	0	0	144	0	0	97.96 %
8	5	0	0	2	0	3	0	0	154	2	92.77 %
9	1	0	0	0	1	1	0	0	0	174	98.31 %

accuracy: 96.46 %. For a detailed report on classification accuracy on USPS testing set see table 1.

5. Conclusions

In the paper we have described a suboptimal design of the HTM network when applied to the task of image recognition, in particular, to handwritten digits of the USPS database. Comparison of the obtained results to the results achieved by other classifiers published in [10] showed that the HTM overcomes performance of a number of tested classifiers, and only several of them achieved higher classification accuracy (range between 97-98 %, combination of tangent vector and local representation and SVM-like approaches). Two final conclusions can be drawn: first, in the HTM, the Zeta1TopNode can be replaced and tuned for application of the SVM classifier, second, as commented by D. George, real power of the HTM architecture can be demonstrated in tasks in which a real temporal hierarchy (instead of a virtual movie) occurs.

Acknowledgements

This work has been supported by the Forschung Austria (grant “Application potential of HTM – a new paradigm of intelligence”).

References

[1] D. George and J. Hawkins. Hierarchical bayesian model of invariant pattern recognition in the visual cortex. In *Proceedings of the international joint conference on neural networks*, Montreal, Canada, 2005. International Neural Networl Society.

[2] Numenta. *Zeta1 Algorithms Reference*, March 2007. Document version 1.0.

[3] D. Soukup and I. Bajla. Robust object recognition under partial occlusions using NMF, 2008. http://www.hindawi.com/getarticle.aspx?doi=10.1155/2008/857453

[4] Numenta. *Hierarchical Temporal Memory, Concepts, Theory, and Terminology*, June 2008. Document version 1.8.0.

[5] Dileep George. *How the Brain Might Work: A Hierarchical and Temporal Model for Learning and Recognition*. PhD thesis, Stanford University, 2008.

[6] Numenta. *Numenta Node Algorithms Guide, NuPIC 1.6*, June 2008.

[7] C. H. Wang and S. N. Srihari. A framework for object recognition in a visually complex environment and its application to locating address blocks on mail pieces. *International Journal of Computer Vision*, 2(2):125–151, September 1988.

[8] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural Computing*, 1(4):541–551, 1989.

[9] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear Programming, Theory and Algorithms*. John Wiley & Sons Inc. New York, 1990. ISBN 0-471-59973-5.

[10] J. Dong. HeroSvm 2.1, 2005. http://www.cenparmi.concordia.ca/~jdong/HeroSvm.html

Simultaneous Tolerance Intervals for the Linear Regression Model

V. Witkovský, M. Chvosteková

Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovakia
Email: witkovsky@savba.sk

Abstract. *In this article, we address the problem of constructing the simultaneous tolerance intervals for linear regression. The suggested method is based on inverting the exact likelihood ratio test (LRT) for testing the simple null hypothesis on all parameters of the linear regression model with normally distributed errors, as proposed in Chvosteková and Witkovský (2009).*

Keywords: Linear Regression Model, Exact Likelihood Ratio Test, Simultaneous Tolerance Intervals

1. Introduction

The simultaneous tolerance intervals are important for many measurement procedures. The most common application for simultaneous tolerance intervals is the multiple-use calibration problem; see e.g. Scheffé (1973), Mee *et al.* (1991), and De Gryze et al. (2007). The tolerance intervals has been recognized and considered in various settings by many authors, see e.g. Wilks (1942), Wallis (1951), Wilson (1967), Lieberman and Miller (1963), Lieberman *et al.* (1967), Limam and Thomas (1988), Mee *et al*. (1991), and Krishnamoorthy and Mathew (2009). These simultaneous tolerance intervals are constructed such that with given confidence level $1-\alpha$ at least a specified proportion $1-\gamma$ of the population is contained in the tolerance interval for all possible values of the predictor variates. All known simultaneous tolerance intervals in regression are conservative in that the actual confidence level exceeds the nominal level $1-\alpha$.

Here we suggest a new method for constructing the simultaneous tolerance intervals for linear regression, based on inverting the exact likelihood ratio test (LRT) for testing the simple null hypothesis on all parameters of the linear regression model with normally distributed errors.

2. The New Simultaneous Tolerance Intervals for Liner Regression

Consider the linear regression model $Y = X\beta + \sigma Z$ with normally distributed errors, where Y represents the n-dimensional random vector of response variables, X is the $n \times k$ matrix of non-stochastic explanatory variables (for simplicity, here we assume that X is a full-ranked matrix), β is a k-dimensional vector of regression parameters, Z is n-dimensional vector of standard normal errors, i.e. $Z \sim N(0, I_n)$, and σ is the error's standard deviation, $\sigma > 0$.

Chvosteková and Witkovský (2009) suggested the likelihood ratio test for testing the simple null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ against the alternative $H_1 : (\beta, \sigma) \neq (\beta_0, \sigma_0)$ on all parameters of the linear regression model with normally distributed errors. The LRT rejects the null hypothesis for large values of $\lambda(y)$, the observed value of the test statistic

$$\lambda(Y) = (Y - X\beta_0)'(T - X\beta_0)/\sigma_0^2 - n\log(\hat{\sigma}_{ML}^2/\sigma_0^2) - n, \tag{1}$$

where $\hat{\sigma}_{ML}^2 = (Y - X\hat{\beta})'(Y - X\hat{\beta})/n$ and $\hat{\beta} = (X'X)^{-1}X'Y$.

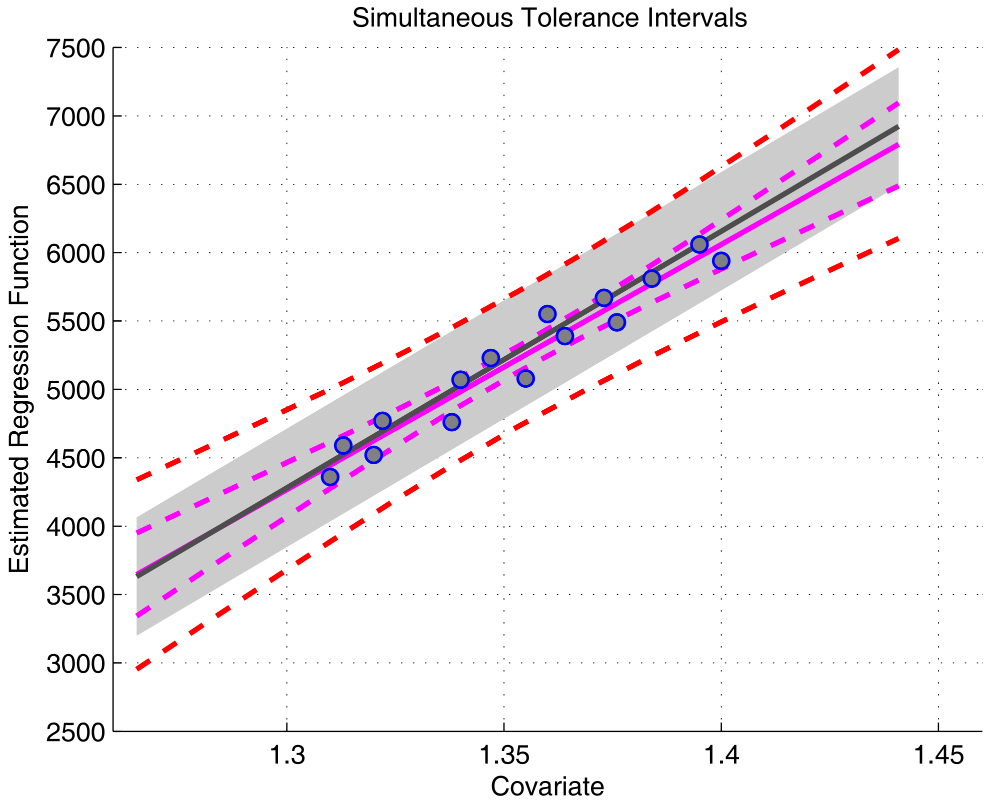


Fig. 1. Illustration of the simultaneous tolerance intervals. The circles represent the observed data. The outer dashed lines represent the simultaneous tolerance interval covering at least the $(1-\gamma)$-content about the true regression function $x'\beta$ with probability at least $(1-\alpha)$, simultaneously for all vectors $x = (x_1, \ldots, x_k)'$ of explanatory variables. The dark solid line represents the true regression function $x'\beta$, together with the $(1-\gamma)$-content region (shaded area) of all possible future observations. The light solid line represents the fitted regression function $x'\hat{\beta}$ together with the simultanoues $(1-\alpha)$-confidence region (light dashed lines) for the true regression line.

So, for the given significance level $\alpha \in (0,1)$ the test rejects the null hypothesis if

$$\lambda(y) > \lambda_{1-\alpha}, \tag{2}$$

where $\lambda_{1-\alpha}$ is the $(1-\alpha)$-quantile of the distribution of the random variable $\lambda(Y)$, where

$$\lambda(Y) \sim Q_k + Q_{n-k} - n\log\left(Q_{n-k}\right) + n\left(\log(n) - 1\right), \tag{3}$$

and $Q_k \sim \chi^2_k$ and $Q_{n-k} \sim \chi^2_{n-k}$ are two independent random variables with chi-square distributions, with k and $n-k$ degrees of freedom, respectively. The critical values of the test could be easily estimated by Monte Carlo simulations, and/or computed exactly by numerical integration. For more details see Chvosteková and Witkovský (2009), where the critical values of the LRT test, at the significance level $\alpha = 0.05$, are given for normal linear regression models with $k = 1, \ldots, 10$ explanatory variables and $n = k+1, \ldots, 100$ observations.

The exact LR test for testing the simple null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ could be directly used to construct the exact confidence region for the parameters of the linear regression model. In particular, the exact $(1-\alpha)$-confidence region for the parameters β and σ is given as

$$\mathcal{C}_{1-\alpha}(Y) = \left\{(\beta, \sigma) : \lambda(Y) \leq \lambda_{1-\alpha}\right\}. \tag{4}$$

Based on that, we define the $(1-\alpha)$-simultaneous tolerance interval covering at least the $(1-\gamma)$-content about the true mean $x'\beta$, for any vector $x = (x_1, \ldots, x_k)'$ of explanatory variables, as

$$\mathcal{T}_{1-\alpha}^{1-\gamma}(x|Y) = \Big[\inf_{(\beta,\sigma)\in\mathcal{C}_{1-\alpha}(Y)} \left\{x'\beta + u_{\gamma_1}\sigma\right\}; \sup_{(\beta,\sigma)\in\mathcal{C}_{1-\alpha}(Y)} \left\{x'\beta + u_{1-\gamma_2}\sigma\right\} \Big], \tag{5}$$

where u_{γ_1} and $u_{1-\gamma_2}$ are pre-specified quantiles of the standard normal distribition such that $\gamma = \gamma_1 + \gamma_2$, with $\gamma \in (0,1)$. As a special case we get the one-sided tolerance intervals, if $\gamma_1 = \gamma$ or $\gamma_2 = \gamma$. However, typically the symmetric tolerance intervals about the estimated regression function are used most frequently, with $\gamma_1 = \gamma_2 = \gamma/2$.
Notice, that directly from the construction of the tolerance intervals $\mathcal{T}_{1-\alpha}^{1-\gamma}(x|Y)$ we get the following probability statement

$$\Pr\Big(\Pr\Big(x'\beta + \sigma Z \in \mathcal{T}_{1-\alpha}^{1-\gamma}(x|Y)\Big) \geq 1-\gamma, \text{ for all } x \text{ and } Z \sim N(0,1), Z \perp Y\Big) \geq 1-\alpha, \tag{6}$$

where $Z \sim N(0,1)$ is a standard normal random variable stochastically independent of the random vector Y.

3. Monte Carlo Method for Approximate Derivation of the Tolerance Bounds

Derivation of the tolerance bounds given by Eq. (4) requires numerical optimization for given x, α and γ (in particular, γ_1 and/or γ_2 such that $\gamma = \gamma_1 + \gamma_2$,), and the observed value y of Y. As an alternative, we suggest a simple algorithm for computing the approximate values of the simultaneous tolerance bounds, based on the Monte Carlo simulations method. The algorithm is based on generating the random sample of size N from the Fisher's fiducial distribution of the regression parameters β and σ, given by

$$\begin{aligned}\tilde{\sigma}^2 &= \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{Q^\star_{n-k}} = \frac{(n-k)S^2}{Q^\star_{n-k}} = \frac{n\hat{\sigma}^2_{ML}}{Q^\star_{n-k}},\\ \tilde{\beta} &= \hat{\beta} - \tilde{\sigma}(X'X)^{-1}X'Z^\star,\end{aligned} \tag{7}$$

with $Q^\star_{n-k} \sim \chi^2_{n-k}$ and $Z^\star \sim N(0, I)$, the stochastically independent random variables. For more details on applications of the fiducial inference see e.g. Fisher (1935), Hannig et al. (2006), and Hannig (2009).
For given observed data y, and the observed fiducial vector of parameters $(\tilde{\beta}, \tilde{\sigma})$ the value of the LR test statistic $\lambda(y|(\tilde{\beta}, \tilde{\sigma}))$ is evaluated

$$\begin{aligned}\lambda(y|(\tilde{\beta}, \tilde{\sigma})) &= \frac{(y - X\tilde{\beta})'(y - X\tilde{\beta})}{\tilde{\sigma}^2} - n\log\left(\frac{\hat{\sigma}^2_{ML}}{\tilde{\sigma}^2}\right) - n\\ &= \frac{(y - X\hat{\beta} + \tilde{\sigma}X(X'X)^{-1}X'Z)'(y - X\hat{\beta} + \tilde{\sigma}X(X'X)^{-1}X'Z)}{\tilde{\sigma}^2} - n\log\left(\frac{\hat{\sigma}^2_{ML}}{\tilde{\sigma}^2}\right) - n\\ &= \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{\tilde{\sigma}^2} + \frac{\tilde{\sigma}^2 Z'X(X'X)^{-1}X'Z}{\tilde{\sigma}^2} - n\log\left(\frac{\hat{\sigma}^2_{ML}}{\tilde{\sigma}^2}\right) - n\\ &= q^\star_{n-k} + q^\star_k - n\log\left(q^\star_{n-k}\right) + n(\log(n) - 1),\end{aligned} \tag{8}$$

where $q^\star_k$ is the observed value of $Q^\star_k \sim \chi^2_k$ and $q^\star_{n-k}$ is the observed value of $Q^\star_{n-k} \sim \chi^2_{n-k}$, the two independent random variables with chi-square distributions, with k and $n-k$ degrees of freedom, respectively. Notice, that the fiducial confidence region

$$\tilde{\mathcal{C}}_{1-\alpha}(y) = \left\{(\tilde{\beta}, \tilde{\sigma}) : \lambda(y|(\tilde{\beta}, \tilde{\sigma})) \leq \lambda_{1-\alpha}\right\}, \tag{9}$$

where $\lambda_{1-\alpha}$ is the $(1-\alpha)$-quantile of the distribution of the random variable $\lambda(Y)$, given by Eq. (3), is equal to the $(1-\alpha)$-confidence region $\mathcal{C}_{1-\alpha}(y)$, defined in (4). Finally, for any vector $x = (x_1, \ldots, x_k)'$, chosen α and γ (in particular, γ_1 and/or γ_2 such that $\gamma = \gamma_1 + \gamma_2$), the $(1-\gamma)$-content $(1-\alpha)$-simultaneous tolerance interval for $x'\beta + \sigma Z$ could be approximately evaluated as

$$\mathcal{T}_{1-\alpha}^{1-\gamma}(x|y) = \Big[\min_{(\tilde{\beta},\tilde{\sigma})\in\tilde{\mathcal{C}}_{1-\alpha}(y)} \Big\{ x'\tilde{\beta} + u_{\gamma_1}\tilde{\sigma} \Big\} \,;\, \max_{(\tilde{\beta},\tilde{\sigma})\in\tilde{\mathcal{C}}_{1-\alpha}(y)} \Big\{ x'\tilde{\beta} + u_{1-\gamma_2}\tilde{\sigma} \Big\} \Big]. \tag{10}$$

The MATLAB algorithm for computing the approximate values of the simultaneous tolerance bounds, based on the Monte Carlo simulations method, is available upon request from the authors.

Acknowledgements

The research was supported by the grants VEGA 1/0077/09, VEGA 2/7087/27, APVV-SK-AT-0003-09, and APVV-RPEU-0008-06.

References

[1] Chvosteková M. and Witkovský V. Exact likelihood ratio test for the parameters of the linear regression model with normal errors. *Measurement Science Review* 9(1): 1–8, 2009.

[2] Fisher R.A. The fiducial argument in statistical inference. Annals of Eugenics, 6:391–398, 1935.

[3] De Gryze S., Langhans I. and Vandebroek M. Using the correct intervals for prediction: A tutorial on tolerance intervals for ordinary least-squares regression. *Chemometrics and Intelligent Laboratory Systems*, 87: 147–154, 2007.

[4] Hannig J., Iyer H.K., Patterson P. Fiducial generalized confidence intervals. *Journal of the American Statistical Association*, 101:484–499, 2006.

[5] Hannig J.: On generalized fiducial inference. *Statistica Sinica*, submitted (2008). See also the working paper at http://www.stat.colostate.edu/research/2006_3.pdf.

[6] Krishnamoorthy K. and Mathew T. *Statistical Tolerance Regions: Theory, Applications, and Computation*. Wiley, ISBN: 978-0-470-38026-0, 512 pages, 2009.

[7] Lieberman G. J. and Miller R. G. Jr. Simultaneous tolerance intervals in regression. *Biometrika*, 50:155-168, 1963.

[8] Lieberman G.J. and Miller R.G. Jr. and Hamilton A. Unlimited simultaneous discrimination intervals in regression. *Biometrika*, 54:133-145. Corrections in Biometrika 58:687, 1967.

[9] Limam M.M.T. and Thomas D.R. Simultaneous tolerance intervals for the linear regression model. *Journal of the American Statistical Association*, 83(403): 801-804, 1988.

[10] Mee R.W., Eberhardt K.R. and Reeve, C.P. Calibration and simultaneous tolerance intervals for regression. *Technometrics*, 33(2): 211-219, 1991.

[11] Scheffe H. A statistical theory of calibration. *Annals of Statistics* 1(1):1-37, 1973.

[12] Wallis, W.A. Tolerance intervals for linear regression. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley. CA: University of California Press. Berkeley, 1951.

[13] Wilks. S.S. Statistical prediction with special reference to the problem of tolerance limits. *The Annals of Mathematical Statistics*, 13: 400-409, 1942.

[14] Wilson A.L. An approach to simultaneous tolerance intervals in regression. *The Annals of Mathematical Statistics*, 38(5): 1536-1540, 1967.

A Table of Critical Values of a Rank Statistic Intended for Testing a Location-Scale Hypothesis

F. Rublík

Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovakia
Email: umerrubl@savba.sk

Abstract. *The exact critical points for selected sample sizes and significance levels are tabulated for the two-sample test statistic which is a combination of the Wilcoxon and the Mood test statistic. This statistic serves for testing the null hypothesis that two sampled populations have the same location and scale parameters.*

Keywords: Distribution-Free Rank Test, Combination of Rank Statistics, Small Sample Sizes

1. Introduction

Suppose that the random variables $X = \sigma_X \varepsilon + \mu_X$, $Y = \sigma_Y \varepsilon + \mu_Y$, where the real numbers μ_X, μ_Y denote the location parameters, $\sigma X > 0$, $\sigma Y > 0$ are scale parameters and the random variable ε has a distribution function continuous on the real line. Assume that $X_1, \ldots, X_m$ is a random sample from the distribution of X, $Y_1, \ldots, Y_n$ is a random sample from the distribution of Y and these random samples are independent. Let $(R_1, \ldots, R_N)$, $N=m+n$, denote the ranks of the pooled sample $X_1, \ldots, X_m, Y_1, \ldots, Y_n$ .The null hypothesis

$$H_0: \quad \mu_X = \mu_Y \ , \qquad \sigma_X = \sigma_Y \tag{1}$$

is sometimes called also the location-scale hypothesis. In practice the change in the location is often accompanied by the change in scale, and in such a case the statistics constructed for testing the location-scale null hypothesis (1) usually yield better results than the statistics constructed especially for the one type change of location or constructed especially for the one type change of scale. Readers interested in further discussion of the need of testing this hypothesis can found further arguments in Section 1 of [6] or in [3].

The null hypothesis (1) is against the alternative H_1 that at least one of the equalities (1) does not hold, tested usually by means of the Lepage test from [4]. This test is included also into the monograph [2]. The Lepage test statistic is given by the formula

$$T = T_K + T_B, \quad T_K = \frac{(S_W - E(S_W \mid H_0))^2}{Var(S_W \mid H_0)} \ , \quad T_B = \frac{(S_B - E(S_B \mid H_0))^2}{Var(S_B \mid H_0)} \ , \tag{2}$$

where $S_W = \sum_{i=1}^{m} R_i$ is the Wilcoxon rank test statistic and $S_B = \sum_{i=1}^{m} a_N(R_i)$ is the Ansari-Bradley rank tests statistic (i.e., the vector of scores $a_N=(1,2,...,k,k,\ldots,2,1)$ if $N=2k$ and $a_N=(1,2,...,k,k+1,k,\ldots,2,1)$ if $N=2k+1$). An analogous statistic has been formulated in the multisample setting in [7], another test statistics has been for this problem studied in [8], where also their non-centrality parameters for testing the null location-scale hypothesis are for some situations computed. The statistics in [8] are defined in the general multisample setting, but they can be used also in the two-sample setting, when the tables of the critical constants can be computed. However, for testing the location-scale hypothesis the only available tables of critical constants are those published in [5].

2. Basic Formulas

The topic of the paper is the computation of the critical values for the two-sample test statistic (3). In accordance with [8] label the combination of the Wilcoxon and the Mood statistic by T_{SQ}. Thus in the notation from the previous section

$$T_{SQ} = T_K + Q \,, \tag{3}$$

$$T_K = \frac{12}{mn(N+1)} \left(S_W - \frac{m(N+1)}{2} \right)^2 , \tag{4}$$

$$Q = \frac{180}{mn(N+1)(N^2-4)} \left(\tilde{S} - \frac{m(N^2-1)}{12} \right)^2 , \quad \tilde{S} = \sum_{i=1}^{m} \left(R_i - \frac{N+1}{2}\right)^2 . \tag{5}$$

Since according to the assumptions the distribution function of ε is continuous, the statistic T_{SQ} is distribution-free whenever the null hypothesis (1) holds. The null hypothesis (1) is rejected whenever $T_{SQ} \geq w_\alpha$. Values of $w_\alpha = w_\alpha(m,n)$ can be found from the table presented in the next section, for the sample sizes not included into the table instead of w_α use the (1-α)th quantile of the chi-square distribution with 2 degrees of freedom.

In the computation of the tables of this paper the following lemma is useful.

Lemma 1. *Let $J(m,n)$ denote the set of all m-tuples $(i_1, \dots ,i_m)$ consisting of integers such that $1 \leq i_1 < \dots < i_m \leq m+n$. Suppose that*

$$D(m,n,k_1,k_2) = \left\{ (i_1, \dots ,i_m) \in J(m,n) \,;\ \sum_{j=1}^{m} i_j = k_1 \,,\ \sum_{j=1}^{m} i_j^2 = k_2 \right\},$$

and $B(m,n,k_1,k_2)$ denotes the number of elements of $D(m,n,k_1,k_2)$.

(I) Let $s > 1$. If at least one of the inequalities $k_1 \leq (r+s)$, $k_2 \leq (r+s)^2$ holds, then $B(s,r,k_1,k_2) = B(s,r-1,k_1,k_2)$.

(II) If $k_1 > (s+r)$, $k_2 > (s+r)^2$, then

$$B(s,r,k_1,k_2) = B(s,r-1,k_1,k_2) + B(s-1, r, k_1 - (r+s) , k_2 - (r+s)^2) .$$

If $R=(R_1, \dots ,R_N)$ is a random vector which is uniformly distributed over the set of all permutations of the set $\{1, \dots ,N\}$, then according to Theorem 1 on p. 167 of [1] for any set $A \subset \{1, \dots ,N\}$ consisting of *m distinct* integers

$$P(\{R_1, \dots ,R_m\}=A) = 1 \Big/ \binom{n}{m} .$$

Combining this equality with Lemma 1 one can construct a program for computation of critical values of the statistic (3).

3. Tables of Critical Values

In this section we present the table 1 of the exact critical values of the statistic (3). First we describe the output of the table.

Since the set $V=V(m,\mathrm{n})$ of possible values of the statistic T_{SQ} is finite for every sample sizes m, n, in general one cannot find exact critical values for arbitrary prescribed probability α of the type I error. In the following table the number on the intersection of the column for $\underline{w}$ with the row for significance level α denote the quantity $\underline{w} = min\{\ t \in V;\ P(\ T_{SQ} \geq t) \geq \alpha$ }, the entry corresponding to α and $\overline{w}$ is the quantity $\overline{w} = max\{t \in V;\ P(T_{SQ} \geq \mathrm{t}) > \alpha\ \}$. Further, for given α, denote $\underline{p} = P(\ T_{SQ} \geq \underline{w}\ )$, $\overline{p} = P(\ T_{SQ} \geq \overline{w}\ )$ the corresponding probabilities of the type I error. Thus $\underline{p}$ is the largest available significance level not exceeding α and $\overline{p}$ is the smallest available significance level greater than α The value of critical constant yielding the significance level closer to the nominal level α is printed in boldface letter. If the difference in computed values exceeds the number of decimal places used to describe the result of computation, then the boldface symbol is used for the value corresponding to $\underline{p}$. For the space reasons only several combinations of sample sizes are given in the following table, a more detailed table will be presented in another paper.

Table 1 : Critical values of the test statistic T_{SQ} from (3).

A	$\underline{w}$	$\underline{p}$	$\overline{w}$	$\overline{p}$	$\underline{w}$	$\underline{p}$	$\overline{w}$	$\overline{p}$	$\underline{w}$	$\underline{p}$	$\overline{w}$	$\overline{p}$
		m=3	n=6			m=3	n=7			m=3	n=8	
0.200	3.5520	.179	**3.2796**	.202	**3.1298**	.200	3.0519	.217	3.3803	.176	**3.2435**	.212
0.100	4.2077	.083	**3.6121**	.107	**4.2532**	.100	4.0454	.117	4.1880	.091	**4.0213**	.103
0.050	**6.3272**	.048	4.7342	.071	**5.5000**	.050	4.2980	.067	6.2435	.036	**4.9359**	.061
0.020			**6.4519**	.024	**7.3181**	.017	7.0000	.033	8.1367	.012	**7.5897**	.024
0.010			**6.4519**	.024			**7.3181**	.017			**8.1367**	.012
0.005			**6.4519**	.024			**7.3181**	.017			**8.1367**	.012
		m=6	n=10			m=6	n=ll			m=6	n=12	
0.200	**3.3170**	.200	3.3086	.201	**3.2812**	.200	3.2748	.202	**3.2998**	.200	3.2998	.200
0.100	**4.4683**	.100	4.4487	.101	**4.4572**	.100	4.4550	.100	**4.4561**	.100	4.4473	.100
0.050	5.4700	.049	**5.4683**	.050	**5.4694**	.050	5.4630	.051	**5.5087**	.050	5.5065	.050
0.020	**6.8313**	.020	6.8016	.020	**6.8474**	.020	6.8410	.020	**6.8837**	.020	6.8771	.020
0.010	**7.7563**	.010	7.6605	.010	**7.8872**	.010	7.8745	.010	**7.9539**	.010	7.9100	.010
0.005	**8.5019**	.0049	8.4733	.0054	8.9718	.0048	**8.8718**	.0050	**9.0235**	.005	8.9183	.0051
		m=7	n=7			m=7	n=8			m=7	n=9	
0.200	**3.3020**	.200	3.2857	.203	**3.3223**	.200	3.3189	.200	3.3445	.198	**3.3328**	.200
0.100	**4.4653**	.100	4.4500	.101	**4.4931**	.100	4.4921	.100	**4.5012**	.100	4.4756	.100
0.050	5.5602	.049	**5.5551**	.050	**5.5029**	.050	5.5029	.050	**5.5694**	.050	5.5480	.050
0.020	**6.7183**	.020	6.6163	.021	**6.7162**	.020	6.6980	.020	**6.7635**	.020	6.7581	.020
0.010	**7.2500**	.010	7.0499	.011	**7.4921**	.010	7.4686	.010	**7.5478**	.010	7.5424	.010
0.005	**8.0500**	.0047	7.8622	.0058	8.0855	.0048	**7.9873**	.0051	8.4740	.004	**8.4441**	.0051
		m=7	n=10			m=7	n=11			m=7	n=12	
0.200	**3.3223**	.199	3.2982	.201	**3.2836**	.200	3.2831	.200	**3.2905**	.200	3.2871	.201
0.100	**4.5132**	.100	4.4972	.100	**4.5140**	.100	4.5140	.100	**4.4804**	.100	4.4789	.101
0.050	**5.5739**	.050	5.5719	.050	**5.5428**	.050	5.5428	.050	**5.6012**	.050	5.6000	.050
0.020	**6.8912**	.020	6.8711	.020	**6.8373**	.020	6.8373	.020	**6.9289**	.020	6.9159	.020
0.010	**7.7453**	.010	7.7052	.010	**7.7821**	.010	7.7805	.010	**7.9493**	.010	7.9373	.010
0.005	**8.5894**	.0050	8.5714	.0051	**8.7792**	.0050	8.7648	.0050	**8.9306**	.005	8.9087	.0050

4. Some Simulation Results

The aim of the following simulation is to obtain a picture of the power of tests based on the statistics (2) and (3) for small sample sizes covered by the Table 1. To consider power of the concerned tests in the case of distributions with various tail behavior, the sampling from normal, logistic and Cauchy distribution is employed, in each case $\mu_X = 0$, $\sigma_X = 1$. The simulations are based in each case on 10000 trials. The critical constants w_α of the statistic T_{SQ} defined in (3) are taken from the previous table, the critical constants t_α of the statistic (2) are those computed in [5]. The power better of the two considered cases is emphasized by the boldface type.

Table 2: Simulation estimates of the power.

	$m = 3$, $n = 7$					
	$\mu_Y = 1.5$, $\sigma_Y = 3$		$\mu_Y = 3.5$, $\sigma_Y = 4$		$\mu_Y = 7$, $\sigma_Y = 5$	
	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.10$
$P(T_{SO} \geq w_\alpha \mid Normal)$	**0.057**	0.07	**0.202**	0.224	**0.533**	**0.547**
$P(T \geq t\alpha \mid Normal)$	0.032	**0.14**	0.136	**0.251**	0.435	0.539
$P(T_{SO} \geq w_\alpha \mid Logistic)$	**0.032**	0.04	**0.073**	0.086	**0.191**	0.208
$P(T \geq t\alpha \mid Logistic)$	0.017	**0.13**	0.047	**0.170**	0.137	**0.247**
$P(T_{SO} > w_a \mid Cauchy)$	**0.041**	0.06	**0.084**	0.104	**0.174**	0.197
$P(T > ta \mid Cauchy)$	0.025	**0.114**	0.060	**0.165**	0.138	**0.233**
	$m = 7$, $n = 8$					
	$\mu_Y = 1.5$, $\sigma_Y = 3$		$\mu_Y = 3.5$, $\sigma_Y = 4$		$\mu_2 = 7$, $\sigma_2 = 5$	
	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.10$
$P(T_{SO} \geq w_\alpha \mid Normal)$	**0.448**	**0.640**	**0.745**	**0.886**	0.950	**0.987**
$P(T \geq t\alpha \mid Normal)$	0.415	0.61	0.730	0.868	**0.952**	0.984
$P(T_{SO} \geq w_\alpha \mid Logistic)$	**0.363**	**0.540**	**0.586**	**0.765**	**0.796**	**0.915**
$P(T \geq t\alpha \mid Logistic)$	0.326	0.51	0.547	0.732	0.774	0.900
$P(T_{SO} > w_a \mid Cauchy)$	**0.228**	0.32	**0.431**	0.566	**0.659**	**0.759**
$P(T > ta \mid Cauchy)$	0.219	**0.36**	0.419	**0.568**	0.643	0.757

5. Discussion and conclusions

As shown on p. 283 of [8], the bounds for the asymptotic efficiency of the test statistics (2) and (3) in the case of there considered sampled distributions do not depend on the number of sampled populations, i.e., they are the same in the two-sample case and in the multisample case. As concluded in [8], taking into account computed values of the asymptotic efficiencies, one sees that a combination of the multisample Kruskal-Wallis statistic (in the two-sample case the Wilcoxon test statistic) with the Mood test statistic appears to be a good choice when one considers symmetric distributions whose type of tail weight is unknown. However, these considerations are related to the asymptotic case, when both *m* and *n* tend to infinity. The simulation results, given in the previous section do not contradict the mentioned asymptotic results. After inspecting the Table 2 it can be said that for small sample sizes and $\alpha = 0.05$ testing based on (3) is preferable to (2), but for $\alpha = 0.1$ it is advisable to use the Lepage test. This suggests that the test based on (3) can be considered as a useful competitor to the Lepage test.

Acknowledgements

The research was supported by the grants VEGA 1/0077/09, APVV-SK-AT-0003-09, and APVV-RPEU-0008-06.

References

[1] Hájek, J, Šidák, Z. and Sen, P. K. *Theory of rank tests*. Academic Press, San Diego 1999.

[2] Hollander, M. and Wolfe, D. A. *Nonparametric statistical methods*. Wiley, New York 1999.

[3] Hutchison, T. P. Should we routinely test for simultaneous location and scale changes? *Ergonomics*, 45(2002), 248-251.

[4] Lepage, Y. A combination of Wilcoxon's and Ansari-Bradley's statistics. *Biometrika* 58(1971), 213-217.

[5] Lepage, Y. A table for a combined Wilcoxon Ansari-Bradley statistic. *Biometrika* 60(1973), 113-116.

[6] Podgor, M. J. and Gastwirth, J. L. On non-parametric and generalized tests for the two-sample problem with location and scale change alternatives. *Statistics in Medicine* 13(1994), 747-758.

[7] Rublík, F. The multisample version of the Lepage test. *Kybernetika* 41(2005), 713-733.

[8] Rublík, F. On the asymptotic efficiency of the multisample location-scale rank tests and their adjustment for ties. *Kybernetika* 43(2007), 279-306.

Confidence Region in Linear Mixed Model for Longitudinal Data

G. Wimmer

Mathematical Institute, Slovak Academy of Sciences, Bratislava, Slovak Republic
Email: wimmerg@mat.savba.sk

Abstract. *In many situations, the response variable is observed on subjects at several time points. Such a data are often referred to as longitudinal data. A widespread model for analysing this type of data is a linear mixed model. We use a score algorithm to estimate a parameters of this model with AR(1) errors to achieve confidence regions for regression parameters. In simulation study we discus qualities of constructed confidence regions and point out some of their deficiencies.*

Keywords: Confidence Region, Linear Mixed Model, Longitudinal Data

1. Introduction

The main characteristic of longitudinal studies is that subjects are measured in some time points or time intervals. With this repeated observations on several subjects we try to describe a common feature, which defines the behaviour of all subjects in time. It is natural to assume that these vectors of outcomes are independent between subjects, but the repeated measurements, done on the single subject exhibit some form of correlation. This is due to the fact that every subject has except these joint attributes also his own individual effects, which affects the final outcome of his repeated measurements. For analyzing such type of data it appears to be advantageous to use the linear mixed model as in [8], which reflects both (common and individual) effects of each subject on his repeated measurements.

It is advisable to note, that individual effects can differ from subject to subject and so the suggested model is able to distinguish between them. Moreover, such model can be used to make statistical inferences not only about the common effects, but also in a broader problem to estimating MSE of the prediction error of the individual's effects, since we are able to estimate with linear mixed model these individual effects and evaluate a deviation of each subject from a common mean. However, this is not the problem discussed in this article.

Apart from that, there is one more advantage to linear mixed model. It permits a certain form of dependence connected with random errors.

2. Model construction

Let us consider a linear mixed model as in [1] with AR(1) errors. The response vector for i-th subject ($i = 1, 2, \ldots, I$) can be written as

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\alpha} + \mathbf{Z}_i \mathbf{b}_i + \boldsymbol{\varepsilon}_i \tag{1}$$

where $\mathbf{X}_i$ is $(n_i \times p)$-dimensional known matrix for i-th subject, $\boldsymbol{\alpha}$ is p-dimensional vector of the unknown regression parameters. These are identical for all subjects. $\mathbf{Z}_i$ are $(n_i \times q)$-dimensional known design matrices for individual effects $\mathbf{b}_i$, where $\mathbf{b}_i$'s are q-dimensional random vectors from $N(\mathbf{0}, \mathbf{D})$ mutually independent. $\boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \mathbf{R}_i)$ are n_i-dimensional error vectors independent of $\mathbf{b}_i$. Here $\mathbf{D}$ and $\mathbf{R}_i$ are some covariance matrices. In light of the above mentioned assumptions we require, that for i-th subject ($i = 1, 2, \ldots, I$) at given time point j ($j = 1, 2, \ldots, n_i$) is

$$\varepsilon_{i,j} = \rho \varepsilon_{i,j-1} + \tau_{i,j} \tag{2}$$

where $\tau_{i,j} \sim N(0, \sigma^2)$. ρ is coefficient of autoregression and σ^2 is some positive scalar.

Covariance matrix for random vector $\mathbf{y}_i$ is then

$$Var(\mathbf{y}_i) = \mathbf{V}_i = \mathbf{Z}_i \mathbf{D} \mathbf{Z}_i^{'} + \sigma^2 \mathbf{R}_i \tag{3}$$

where

$$\sigma^2 \mathbf{R}_i = \frac{\sigma^2}{1-\rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \dots & \dots & \dots & \dots & \dots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{bmatrix}.$$

Let us denote $\boldsymbol{\nu} = (d_{11}, d_{12}, \dots, d_{22}, \dots, d_{rr}, \sigma^2, \rho)'$ as a vector of all variance-covariance parameters in model (1) (i.e. we can write $Var(\mathbf{y}_i) = \mathbf{V}_i(\boldsymbol{\nu})$). Now it is clear that all the parameters of the proposed model (1) are $(\boldsymbol{\alpha}', \boldsymbol{\nu}')'$.

3. Parameter estimation and its properties

With assumption form the section 2 we can use a score algorithm to estimate the parameters of model (1) directly from the likelihood function. Despite the fact, that in our primary interest is to estimate unknown regression parameter $\boldsymbol{\alpha}$, it is necessary to estimate also the variance-covariance parameters $\boldsymbol{\nu}$, since these are usually also unknown. To obtain an estimator of these unknown parameters $(\boldsymbol{\alpha}', \boldsymbol{\nu}')'$, we can use a logarithm of likelihood function, which is proportional to

$$l = -\frac{1}{2} \sum_{i=1}^{I} ln \left| \mathbf{V}_i(\boldsymbol{\nu}) \right| - \frac{1}{2} \sum_{i=1}^{I} \left[(\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\alpha})^{'} \mathbf{V}_i^{-1}(\boldsymbol{\nu}) (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\alpha}) \right] \tag{4}$$

With a given maximum likelihood estimate of $\boldsymbol{\nu}$, $\hat{\boldsymbol{\nu}}$, maximizing (4) we get the following maximum likelihood estimate of $\boldsymbol{\alpha}$

$$\hat{\boldsymbol{\alpha}} = \left(\sum_{i=1}^{I} \mathbf{X}_i^{'} \mathbf{V}_i^{-1}(\hat{\boldsymbol{\nu}}) \mathbf{X}_i \right)^{-1} \left(\sum_{i=1}^{I} \mathbf{X}_i^{'} \mathbf{V}_i^{-1}(\hat{\boldsymbol{\nu}}) \mathbf{y}_i \right) \tag{5}$$

As we can see, it is necessary at first to estimate unknown variance-covariance parameters and then we can calculate an estimate for regression parameter. Successive iteration between these two steps yields maximum likelihood estimates of unknown parameters, such as $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\nu}}$. For more details see [1].

With a given estimate $\hat{\boldsymbol{\alpha}}$ we can consider a construction of confidence regions for some linear combination $\mathbf{L}\boldsymbol{\alpha}$, where $\mathbf{L}$ is known $(r \times p)$-dimensional matrix. If a covariance matrix of vectors $\mathbf{y}_1, \dots, \mathbf{y}_I$ were known, then $\hat{\boldsymbol{\alpha}}$ is asymptotically normally distributed with mean

$$E\left(\hat{\boldsymbol{\alpha}}\right) = \boldsymbol{\alpha} \tag{6}$$

and covariance matrix

$$Var(\hat{\boldsymbol{\alpha}}) = \left(\sum_{i=1}^{I} \mathbf{X}_i^{'} \mathbf{V}_i^{-1} \mathbf{X}_i \right)^{-1} \tag{7}$$

In this case is

$$X^2 = (\mathbf{L}\hat{\boldsymbol{\alpha}} - \mathbf{L}\boldsymbol{\alpha})^{'} (\mathbf{L}\, Var(\hat{\boldsymbol{\alpha}}) \mathbf{L}^{'})^{-1} (\mathbf{L}\hat{\boldsymbol{\alpha}} - \mathbf{L}\boldsymbol{\alpha}) \tag{8}$$

χ^2 distributed with r degrees of freedom.

If, and it is in the majority of practical application, must be estimated also the variance-covariance parameters of the model, we can replace (7) with its maximum likelihood estimate

$$\hat{Var}(\hat{\boldsymbol{\alpha}}) = \left(\sum_{i=1}^{I} \mathbf{X}_i' \hat{\mathbf{V}}_i^{-1} \mathbf{X}_i\right)^{-1},$$

where $\hat{\mathbf{V}}_i = \mathbf{V}_i(\hat{\boldsymbol{\nu}})$ is maximum likelihood estimate of the covariance matrix. Then we can "naïve" assume (and it is that, what many authors do, see for example [2] or [3]) that

$$X^2 = (\mathbf{L}\hat{\boldsymbol{\alpha}} - \mathbf{L}\boldsymbol{\alpha})'(\mathbf{L}\,\hat{Var}(\hat{\boldsymbol{\alpha}})\mathbf{L}')^{-1}(\mathbf{L}\hat{\boldsymbol{\alpha}} - \mathbf{L}\boldsymbol{\alpha}) \tag{9}$$

has also χ^2 distribution with r degrees of freedom.

It was created a MATLAB algorithm "CONFZON" which evaluates the 95% confidence region from (9) for different numbers of subjects and different ranges of measurements for each subject. For some combination of these two parameters, we counted the empirical probability of coverage of the real value $\boldsymbol{\alpha}$ from 10000 simulations taken by this confidence region. We considered model (1) with 2-dimensional regression parameter $\boldsymbol{\alpha} = (1,2)'$, 2-dimensional vector of individual effects $\boldsymbol{b} = (b_1, b_2)'$ with covariance matrix $\mathbf{D} = (1,0;0,1)$, AR(1) parameter $\rho = 0.5$ and errors variance $\sigma^2 = 1$. Results are shown in the Tables 1-2.

Table 1. Simulated probability of coverage of 95% confidence region evaluated from (9) for different numbers of subjects with the same range of the repeated measurements on each subject.

Number of subjects	Range of the repeated measurements on each subject	Probability of coverage
5	5	0.8297
10	5	0.8797
30	5	0.9394
50	5	0.9384
100	5	0.9413
500	5	0.9510
1000	5	0.9586

Table 2. Simulated probability of coverage of 95% confidence region evaluated from (9) for small number of subjects with a different range of the repeated measurements on each subject.

Number of subjects	Size of the repeated measurements on each subject	Probability of coverage
5	5	0.8297
5	30	0.8687
5	50	0.8805
5	80	0.8933
5	100	0.9040
5	150	0.9082

4. Discussion and conclusions

As it turns out, that our "naive" concept of the confidence region is particularly suitable for large numbers of subjects, which shows the Table 1. It is also appropriate to note, that for small numbers of subjects (5 or 10) is this conference region unsuitable or (30 and 50) liberal. However, for a sufficient number of subjects (from 100 subjects) also for small numbers of the repeated measurements on each subject proposed confidence region is approaching the theoretical value. Moreover, from the Table 2 can be concluded that despite the increasing number of repeated measurements on each subject for a small sample of the subjects approaching simulated probability of coverage is very slow to the theoretical value. It is caused because the proposed confidence region does not take into account the uncertainty inherent in estimating the variance-covariance parameters. This can be removed using the F distribution instead of χ^2 distribution, but there arise practical problems with the numbers of degrees of freedom for this F distribution, where are used different approximation, see e.g. [4] and [6], [7]. Unfortunately, these confidence regions were not yet studied in detail for the analysis of longitudinal data. Therefore we think that it would be appropriate, on the basis of additional simulations, to verify their properties, or to propose their improvement.

Acknowledgements

The research was supported by the grants VEGA 1/0077/09, APVV-SK-AT-0003-09 and APVV-RPEU-0008-06.

References

[1] Chi EM, Reinsel GC. Model for Longitudinal Data with Random Effects and AR(1) Errors. *Journal of the American Statistical Association,* 84 (406): 452-459, 1989.

[2] Diggle PJ, Heagerty P, Liang KY, Zeger SL. Analysis of Longitudinal Data. Oxford University Press, 2002

[3] Fitzmaurice GM, Laird NM, Ware JH. Applied Longitudinal Analysis. John Wiley & Sons, New Jersey, 2004.

[4] Harville D. Accounting for Estimation of Variance and Covariance in Prediction Under general Linear Model: An Overview. *Tatra Mountains Mathematical Publication,* 39: 1-15

[5] Hedeker D, Gibbons RD. Longitudinal Data Analysis. John Wiley & Sons, New Jersey, 2006.

[6] Kenward MG, Roger JH. Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood. *Biometrics,* 53 (3): 983-997, 1997

[7] Kenward MG, Roger JH. An improved approximation to the precision of fixed effects from restricted maximum likelihood. *Computational Statistics and Data Analysis,* doi:10.1016/j.csda.2008.12.013, 2009

[8] Laird NM, Ware JH. Random-Effects Models for Longitudinal Data. *Biometrics,* 38 (4): 963-974, 1982.

The ROC Analysis for Classification of Smokers and Non-Smokers Based on Various Prior Probabilities of Groups

K. Cimermanová

Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovakia
Email: katarina.cimermanova@gmail.com

Abstract. *This paper addresses the influence of the prior probabilities of diseases for diagnostic reasoning. For various prior probabilities of smokers and non-smokers used in discriminant analysis we construct the ROC curve and the Youden Index with related asymptotic pointwise confidence intervals. We show how the prior probabilities change probability of diagnostic results.*

Keywords: Breath Analysis, ROC Analysis, Confidence Intervals, Discriminant Analysis, Prior Probability

1. Introduction

The prior probabilities are independent of measured data and known before we have taken any observations. The change of this information changes the probability of a correct output of a diagnostic test [5]. In this paper we show how radical these changes are in classification of measured concentrations of volatile organic compounds of smokers and non-smokers.

The ROC (receiver operating characteristic) curve is a metric for comparing predicted and actual target values in a classification model. The ROC curve plots sensitivity and 1-specificity of the diagnostic test. The sensitivity measures the proportion of actual positives which are correctly identified as such (i.e. the percentage of sick people who are identified as having the condition); and the specificity measures the proportion of negatives which are correctly identified (i.e. the percentage of well people who are identified as not having the condition).

Different classification algorithms use different techniques for finding relationships between the measured values of subjects (e.g. concentrations of selected volatile organic compounds, VOCs, of breath profile) and the known targets (association with groups, e.g. smokers or non-smokers). We use the discriminant function $g(X)$ with a threshold (the decision point used by the model for classification) dependent on prior probabilities of groups, [5]. The ROC curve measures the impact of changes in threshold. For the ROC curve related with changes of prior probabilities we construct asymptotic pointwise confidence interval, [4] (CI describes range where the true ROC curve lies with some specific probability, e.g. 95% CI).

To evaluate effectiveness of classification based on different prior probabilities of discriminated classes we use the Youden Index [3]. This index ranges between 0 and 1, with a value close to 1 indicating that the effectiveness of algorithm is relatively large and a value close to 0 indicating limited effectiveness. For the Youden Index we construct asymptotic pointwise confidence interval, too.

We apply the classification on breath analysis data. Breath analysis as a non-invasive technique is very attractive because it can be easily applied to sick patients, including children and elderly people. It offers potential for detection of some diseases, e.g. diabetes, lung and esophageal cancer etc. In our study we consider measured values of breath profile of smokers and non-smokers measured by proton transfer reaction mass spectrometry PTR-MS, for more details see e.g. [1]. The molecular masses detectable by the PTR-MS range from m/z 21 to m/z 230. The selected compounds (m/z values) for our analysis are m/z 28 (tentatively identified as hydrogen cyanide), m/z 42 (acetonitrile), m/z 53 (1-buten-3-alkyne), m/z 59 (acetone), m/z 79 (benzene), m/z 93 (toluene) and further m/z 97, m/z 109 and m/z 123, for more details see [6]. The measured quantities (counts) are transformed [6] to concentrations of volatile organic compounds in ppb levels.

2. ROC analysis

Let us have random vectors $X_i = (x_{i1}, \dots x_{in})$ where $i = 1, \dots N$, N is the number of all observed subjects, x_{ij} represents measured concentration of the j-th volatile organic compound (VOC) of subject i and n is the number of selected VOCs. For each subject X_i we have categorization to a population, i.e. the target. For the population of each group (smokers and non-smokers) we assume n-dimensional normal distribution.
For classification we can use the discriminant function

$$g(X) = -\frac{1}{2}(X-\mu_1)'\Sigma_1^{-1}(X-\mu_1) + \frac{1}{2}(X-\mu_2)'\Sigma_2^{-1}(X-\mu_2) + \ln|\Sigma_2|/|\Sigma_1|, \qquad (1)$$

where X is a vector of observed values of a subject, μ_1 and μ_2 are mean values estimated from training data and Σ_1 and Σ_2 are covariance matrices estimated from training data. (The database is divided into training and a testing set in some ratio, e.g. 3:2).

For the new observation X (from the testing set) we evaluate the value of the discriminant function $g(X)$. This value is compared with a threshold value k, $-\infty < k < \infty$. When $g(X) > k$ subject is classify to the group of positives ($X \in \omega_1$) and otherwise when $g(X) < k$, the subject is classify as negative ($X \in \omega_2$). In our case the threshold value k is defined as

$$k = \ln\frac{P(\omega_2)}{P(\omega_1)} \qquad (2)$$

where $P(\omega_1)$ and $P(\omega_2)$ are prior probabilities, more in [5].

From results of classification of testing data we can evaluate sensitivity Se and specificity Sp as

$$Se = \frac{TP}{TP+FN} \quad \text{and} \quad Sp = \frac{TN}{TN+FP} \qquad (3)$$

where TP is true positive (positive subject is classified as positive), TN is true negative (negative subject as negative), FP is false positive (negative as positive) and FN is false negative (positive subject as negative).

The sensitivity can be expressed as $Se(k) = P(g(X) > k) = 1 - P(g(X) \leq k) = 1 - G(k)$ and specificity $Sp(k) = P(g(X) \leq k) = F(k)$, where $G(k)$ and $F(k)$ can be interpreted as cumulative distribution functions (cdfs) of discriminant function $g(X)$ for positive group ω_1 and negative group ω_2. The alternative definition of the ROC curve is

$$R(1-t) = 1 - G\{F^{-1}(t)\} \qquad (4)$$

for $0 \leq t \leq 1$ where $F^{-1}(t) = \inf\{X: F(k) \geq t\}$ denotes the generalized inverse function of F. However, since empirical cdfs are discontinuous, the estimate of $R(1 - k)$ might have a very erratic appearance [4]. For this reason, it can be advantageous to use smooth empirical cdfs for calculating the estimator of $R(1 - t)$. From cumulative distribution functions we construct, based on normalized histograms, probability density functions (pdfs) f and g. We smooth the functions f and g, too. Next we assume that f and g are continuous and f/g is bounded on any subinterval (a,b) of (0,1), and $n/m \to \lambda$ as $\min(n,m) \to \infty$, where n and m are sample size of training sets of populations. The asymptotic pointwise estimate of a CI for $R(1 - k)$ is defined as

$$R(1-t) \pm z(\alpha / 2)\sigma(t) \tag{5}$$

where $z(\alpha/2)$ is the $\alpha/2$-quantile of a standard normal distribution, α is a chosen level of significance and σ is the standard deviation of the ROC curve defined later. In [4] it is shown that a probability space exists on which one can define two independent Brownian bridges such that

$$\sqrt{n}\hat{G}\{\hat{F}^{-1}(t) - G\{F^{-1}(t)\}\} = \sqrt{\lambda} B_1^{(n)}(G\{F^{-1}(t)\}) + \frac{g(F^{-1}(t))}{f(F^{-1}(t))} B_2^{(n)}(t) + o(n^{-1/2}(\log n)^2) \tag{6}$$

for this processes we have $E(B^{(n)}(t)) = 0$ and $E(B^{(n)}(t)B^{(n)}(s)) = t\,(1-s)$, for more details see [2]. It can be shown that $\hat{G}\{\hat{F}^{-1}(t)\} - G\{F^{-1}(t)\}$ is asymptotically normally distributed as

$$\hat{G}\{\hat{F}^{-1}(t)\} - G\{F^{-1}(t)\} \approx \hat{G}\{F^{-1}(t)\} - G\{F^{-1}(t)\} - \frac{g\{F^{-1}(t)\}}{f\{F^{-1}(t)\}}[\hat{F}\{F^{-1}(t)\} - t] \tag{7}$$

with zero mean and variance

$$\sigma^2(t) = \frac{1}{n} G\{F^{-1}(t)\}[1 - G\{F^{-1}(t)\}] + \frac{1}{m}\frac{g^2\{F^{-1}(t)\}}{f^2\{F^{-1}(t)\}} t(1-t)\,, \tag{8}$$

where after replacing F, G, f and g by the respective estimators we obtain an estimator of σ^2 for $R(1 - t)$. The Youden Index is defined as

$$J(t) = Se(k) + Sp(k) - 1 \tag{9}$$

for all possible threshold values k [3]. It is the maximum vertical distance between the ROC curve and the diagonal or the chance line, Fig. (1), left. The Youden Index can be rewritten as $J(t) = Se(k) + Sp(k) – 1 = F(k) – G(k) = R(1 - k) + F(k) – 1$, where $F(k) = t$ is regarded as a constant. So for the Youden Index we can write an asymptotic pointwise CI

$$J(t) \pm z(\alpha / 2)\sigma(t)\,, \tag{10}$$

where σ is estimated by Eq. (1) defined for σ estimator of the ROC curve $R(1-t)$. The optimal classification is at the point where the Youden Index is maximal.

3. Results

Recent results suggest that breath-concentrations could be expected to be log-normally distributed and that the logarithmic transformation of the data could be profitable, e.g. [6]. Therefore we analyze logarithmic transformed data. In the database, we have measured concentrations of selected VOCs for 44 smokers and 173 non-smokers.

The sensitivity Se and specificity Sp was estimated by Eq (3), where TN, TP, FP, FN values were computed as arithmetic means based on 1000 times divided database in ratio 3:2 for different k defined by the prior probabilities $P(\omega_1)$ = 1:0.01:1 and $P(\omega_2) = 1 - P(\omega_1)$.

From proportion Se and Sp ecdfs of discriminant function $g(X)$ were evaluated for the positive group G and the negative group F. The functions f and g, pdfs of discriminant function were computed from normalized histogram from ecdf. For computing of the ROC curve by Eq. (4) and the standard deviation of the ROC curve by Eq. (8), we smoothed G, F, f and g functions with Gaussian window. The Youden Index was evaluated by Eq. (9) with 95% confidence interval by Eq. (10).

The results of classification of smokers and non-smokers are plotted in Fig 1. The most effective classification is for prior probability of smokers $P(\omega_1) = 0.3$. We also see that the effectiveness of classification is different for different prior probabilities.

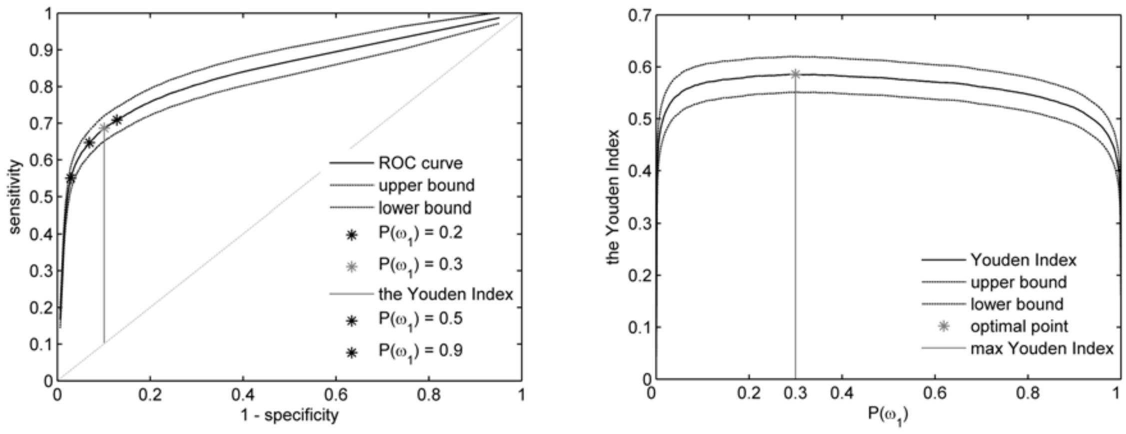


Fig.1. (left) The ROC curve with 95% confidence interval for discriminant function for two groups with threshold dependent on prior probabilities of groups, optimal threshold point with related Youden Index and other threshold points characterized by prior probability of positive group.
(right) The Youden Index with 95% confidence interval for discriminant function for two classes with threshold dependent on prior probabilities of groups and optimal threshold point with related Youden index.

4. Discussion and Conclusions

The ROC analysis is an important tool to summarize the performance of a medical diagnostic test. By the Youden index we see effectiveness of classification.
The confidence bands are a useful graphical tool for visualizing the statistical variability of the ROC curve and the Youden Index estimated from diagnostic test of clinical data.

Acknowledgements

The research was supported partly by the Slovak Research and Development Agency (APVV), grant SK-AT-0003-08 and RPEU-0008-06, by the Scientific Grant Agency of the Slovak Republic (VEGA), grant 2/7087/27 and 1/0077/09.

References

[1] Amann, A., Smith, D. Breath analysis for clinical diagnosis and therapeutic monitoring, World Scientific, Singapore, 2005

[2] Billingley, P. Convergence of Probability Measure, John Wiley & Sons, Inc., New York, 1968, 64-65

[3] Fluss, R., Faraggi, D., Reiser, B. Estimation of the Youden Index and its Associated Cutoff Point, *Biometric Journal*, Vol. 47, 2005, 45-72

[4] Hall, P.G., Hyndman, R.J., Fan, Y. Nonparametric Confidence Intervals for Receiver Operating Characteristic Curve, *Biometrika*, Vol. 91, No. 3, 2004, 743-750

[5] Hand, D.J. Discrimination and classification, Wiley series in probability and mathematical statistics, J Wiley, 1981

[6] Kushch, I., et al. Compounds Enhanced in a Mass Spectrometric Profile of Smokers' Exhaled Breath Versus Non-smokers as Determined in a Pilot Study Using PTR-MS, *Journal of Breath Research*, Vol.2, 2008, 1-26

Determining the Confidence Interval for the Center and Width of a Structure in Fitting Measured Data by the Regression Line

[1] G. Wimmer, [2]K. Karovič

[1]IMS and MI SAS, Bratislava, FNS UMB, Banská Bystrica, Slovakia
and FNS MU Brno, Czech Republic,
[2]Institute of Measurement Science SAS, Bratislava, Slovakia
Email: wimmer@mat.savba.sk

Abstract. *The two dimensional cross section of interest of a structure (e.g. a grating line) is schematically seen in Fig.2. Solid lines are structure's edges. The solid bold arrow at the horizontal axis is the structure's width w for given value F. Here the proper confidence intervals for width and center of the structure are derived.*

Keywords: Confidence Interval, Measurement Uncertainty

1. Introduction

In metrology in case of assumed linear dependence between two quantities *x*, *y* (e.g. length and electric signal, respectively) the functional dependence $y = f(x)$ is fitted using the regression line $y=a+bx$. It is assumed that measurement of the quantity *y* in (exact) point *x* is normally distributed, and the measurements are independent with equal standard deviations. The estimators of the regression coefficients $\hat{a}$, $\hat{b}$ and their standard deviations $s_{\hat{a}}, s_{\hat{b}}$ are determined from pairs of measured values $\{x_i, y_i\}_{i=1}^n$ using standard procedures (see e.g. [3]). A common estimator of the *x* value for assigned level *F* of the quantity *y* (e.g. electric signal) is

$$\hat{x} = \frac{\hat{F} - \hat{a}}{\hat{b}},$$

where $\hat{F}$ is a proper estimator of *F*, independent of $\hat{a}$, $\hat{b}$. Let *f* be the estimate of *F* and s_F be the estimate of the standard deviation of *f* . By using the Law of Propagation of Uncertainties (see [2]) we obtain the estimate of the standard deviation of the estimator $\hat{x}$ as

$$s_{\hat{x}} = \frac{1}{\hat{b}} \sqrt{\left[\frac{s^2}{n\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2} \left(\sum_{i=1}^n x_i^2 - 2\frac{f-\hat{a}}{\hat{b}} \sum_{i=1}^n x_i + n\left(\frac{f-\hat{a}}{\hat{b}}\right)^2 \right) \right] + s_F^2}\ ,$$

where

$$s^2 = \frac{1}{n-2}\left(\sum_{i=1}^n y_i^2 - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n} - \hat{b}\left(\sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n} \right)\right). \tag{1}$$

In deriving the (1-α)-confidence interval for x difficulties arise caused by the generally nonsymmetrical distribution of $\hat{x}$ and so this confidence interval cannot be determined in usual way as $\left(\hat{x} - ks_{\hat{x}},\, \hat{x} + ks_{\hat{x}}\right)$.

The desired (1-α)-confidence interval for x can be determined by the below described procedure.

2. Subject and Methods

The (1-α)-confidence interval for x

It is obvious (see e.g. [3]) that for an arbitrary x the (1-α)-confidence interval for (nonrandom value) $a + bx$ ($= y_x$) is $({}_l y_x,\ {}_u y_x)$, where

$$ {}_l y_x = \hat{a} + \hat{b}\, x - s d_x \left(t_{n-2}\left(1 - \frac{\alpha}{2}\right)\right), \qquad {}_u y_x = \hat{a} + \hat{b}\, x + s d_x \left(t_{n-2}\left(1 - \frac{\alpha}{2}\right)\right), $$

with

$$ d_x = \sqrt{\frac{1}{n}\left(1 + \frac{\left(nx - \sum_{i=1}^{n} x_i\right)^2}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}\right)}, $$

and $t_{n-2}\left(1 - \frac{\alpha}{2}\right)$ is the $\left(1 - \frac{\alpha}{2}\right)$-quantile of the Student t-distribution with n-2 degrees of freedom, where s is given by (1), for illustration see Fig. 1. According to [1], pp. 509-512, in case of sufficiently steep edges of the structure and small value of s (what is assumed here), the (1-α)-confidence interval for x, dented by $({}_l x,\ {}_u x)$, can be constructed by the approach illustrated in Fig. 1. Given the errorless (nonrandom) value F, the following relations hold true for the boundaries ${}_l x$, ${}_u x$ of the (1-α)-confidence interval for x

$$ F = \hat{a} + \hat{b}\, {}_l x - s d_x \left[t_{n-2}\left(1 - \frac{\alpha}{2}\right)\right], $$

$$ F = \hat{a} + \hat{b}\, {}_u x + s d_x \left[t_{n-2}\left(1 - \frac{\alpha}{2}\right)\right]. $$

Solving both preceding equations the bounds ${}_l x$ and ${}_u x$ are given by

$$ {}_l x = \frac{-B}{2A} - \frac{\sqrt{B^2 - 4AC}}{2A}, \quad {}_u x = \frac{-B}{2A} + \frac{\sqrt{B^2 - 4AC}}{2A} \tag{2} $$

for
$$ A = \hat{b}^2 - \frac{ns^2\left[t_{n-2}^2\left(1 - \frac{\alpha}{2}\right)\right]}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}, $$

$$B = 2\left(\frac{s^2\left[t_{n-2}^2\left(1-\frac{\alpha}{2}\right)\right]\sum_{i=1}^{n}x_i}{n\sum_{i=1}^{n}x_i^2-\left(\sum_{i=1}^{n}x_i\right)^2} - \hat{b}\left(F-\hat{a}\right)\right),$$

$$C = \left(F-\hat{a}\right)^2 - \frac{s^2}{n}\left[t_{n-2}^2\left(1-\frac{\alpha}{2}\right)\right]\cdot\left(1+\frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n\sum_{i=1}^{n}x_i^2-\left(\sum_{i=1}^{n}x_i\right)^2}\right).$$

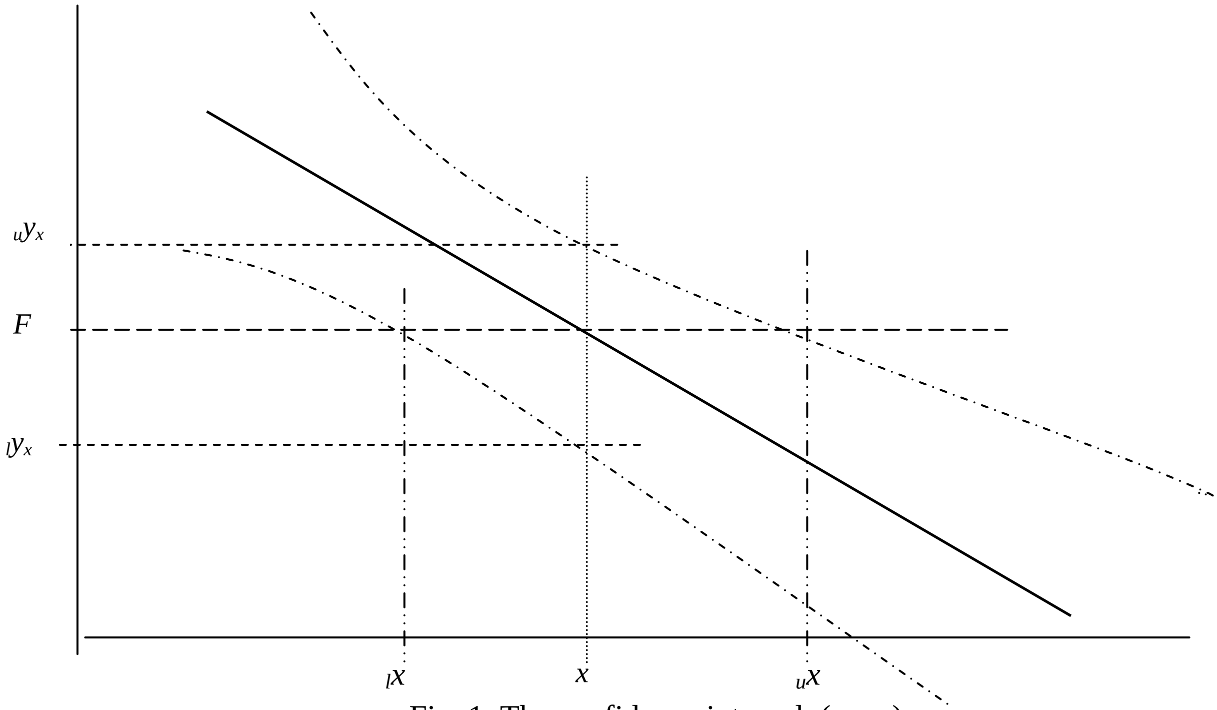


Fig. 1. The confidence interval $({}_lx, {}_ux)$

Determining some parameters of the structure

In order to determine the structure's center and width, for a given (errorless, nonrandom) F in analyzed cross section, it is necessary to use the x values from two confidence intervals $\left({}^{(1)}_{l}x, {}^{(1)}_{u}x\right)$ and $\left({}^{(2)}_{l}x, {}^{(2)}_{u}x\right)$ that correspond to the structure's edges (borders), see Fig. 2.

Using Bonferroni's inequality (see e.g. in [3]) for the structure's center x_s

$$P\left\{x_S \in \left(\frac{{}^{(1)}_{l}x+{}^{(2)}_{l}x}{2}, \frac{{}^{(1)}_{u}x+{}^{(2)}_{u}x}{2}\right)\right\} \geq 1-2\alpha\,,$$

i.e.

$$\left(x_{S_1} = \frac{{}^{(1)}_{l}x+{}^{(2)}_{l}x}{2}, \quad x_{S_2} = \frac{{}^{(1)}_{u}x+{}^{(2)}_{u}x}{2}\right) \qquad (3)$$

is at least (1-2α)-confidence interval for the structure's center.

Similarly, for the structure's width w ,

$$P\left\{w \in \left({}^{(2)}_{l}x - {}^{(1)}_{u}x,\ {}^{(2)}_{u}x - {}^{(1)}_{l}x\right)\right\} \geq 1 - 2\alpha ,$$

and from this,

$$\left(w_1 = {}^{(2)}_{l}x - {}^{(1)}_{u}x,\ w_2 = {}^{(2)}_{u}x - {}^{(1)}_{l}x\right) \tag{4}$$

is the at least (1-2α)-confidence interval for the structure's width w.

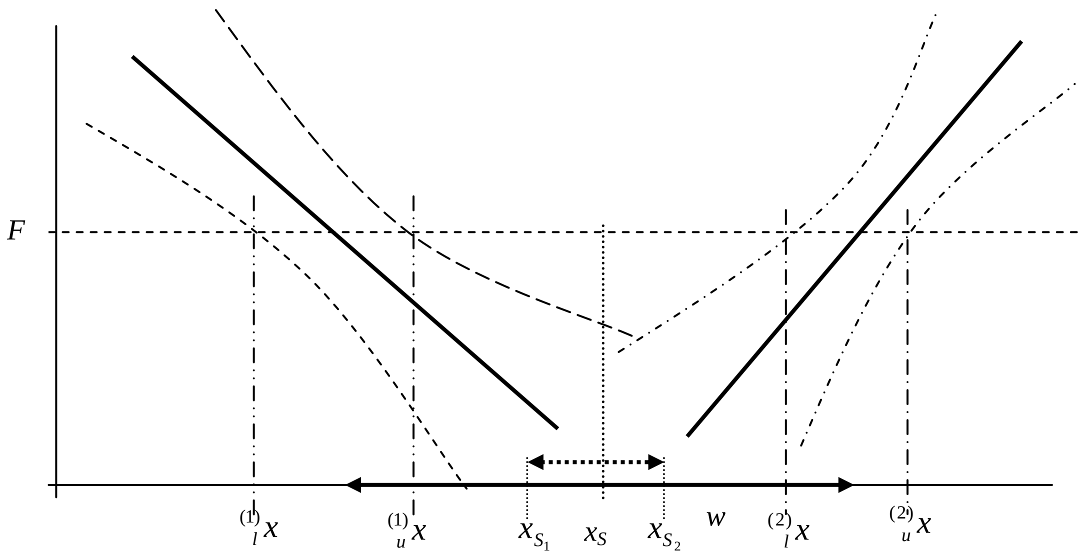


Fig. 2. Determining the sizes of the structure

3. Conclusion

The above achieved assertions are applicable to measurement of the geometry of two-dimensional structures (or cross sections of three-dimensional structures) in the following manner:

For a chosen (errorless) value (level) F of quantity y (e.g. electric signal of a length gauge) and $\alpha \in (0,1)$, the (1-α)-confidence interval for the structure's border (bound) is $({}_{l}x,\ {}_{u}x)$ where ${}_{l}x$, ${}_{u}x$ are given in (2), the (1-2α)-confidence interval for the structure's center is given in (3), and the (1-2α)-confidence interval for the structure's width is given in (4).

Acknowledgment

Contribution was supported by Alexander von Humboldt foundation and grants APVV RPEU-0008-06, VEGA 1/0077/09, VEGA 2/7082/27 and MŠMT ČR LC06024.

References

[1] Cramér H. Mathematical Methods of Statistics. Princeton University Press, Princeton, 1946.

[2] Guide to the Expression of Uncertainty of Measurement (GUM) (1995), ISO, ISBN 91-67-10188-9, Geneve, Switzerland, 101 pages.

[3] Mood R.M., Graybill F.A., Boes D. Introduction to the Theory of Statistics. (3rd edn.). International Student Edition McGraw-Hill, Auckland, 1974.

Comparison of Outliers Elimination Algorithms

P. Mostarac, R. Malarić, H. Hegedušić

Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
Email: {petar.mostarac; roman.malaric; hrvoje.hegedusic}@fer.hr

Abstract. *In this paper one general approach to elimination the single outlier is described, and shown that Chauvenet's criterion is a one particular solution. Peirce's criterion is also described and compared for the rejection of suspicious data with Chauvenet's criterion, their sensitivity to outliers and their characteristics. All expressions needed for calculation the borders for rejection suspicious observations and their solutions are given.*

Keywords: Outliers Elimination, Comparison of Peirce's Criterion vs. Chauvenet's Criterion

1. Introduction

When modeling error of measurement, which is defined as difference between real and measured value, insignificant part of an error is made by impulse error. Impulse errors are common fluctuation of significant deviation and can increase result of measurement interpreted by mean and standard deviation statistics.

If probability of having suspicious measurement, which can be described as impulse error, with assumption of Gauss distribution and estimation of parameters (mean and variance), given by the all measurements, less than the real number of suspicious data, it can be discuss on their rejection. That decision is explained with probability that rejected data are a result of impulse error. By setting boundaries for outlier elimination, strictness of criterion is defined and arbitration over suspicious data is done.

The rest of this paper the focus will be on setting the criterion for one suspicious data, and Peirce's criterion which is able for multiple outlier elimination.

2. Criterions

Criterion for one suspicious data

Criterion for eliminating outliers can be defined by the amount of allowed deviation comparing to standard deviation σ. For N measurements, with standard deviation σ, mean μ and suspicious data *x* defined as:

$$n = \frac{\max\{|x-\mu|\}}{\sigma} \tag{1}$$

It can be defined [1] that probability in N measurements must be greater or equal to *p* so it can be kept:

$$p \geq N \times (1 - P(\mathrm{n}\sigma)), \tag{2}$$

P is a function defined like integral of probability density function (*pdf*) on interval ±nσ. Function *P* can be substituted with Cumulative distribution function (*cdf*) *Φ*,

$$P(\mathrm{n}\sigma)=(1-2\Phi(\mathrm{n}\sigma)), \tag{3}$$

and introducing (3) in (2), with defining *cdf* function through Error function the solution is:

$$n \leq \sqrt{2}erf^{-1}(\frac{p}{N}-1)\ . \tag{4}$$

For Chauvenet's criterion p=0,5. And p can be chosen such that our criterion be less or more rigorous.

Peirce's criterion

Peirce's criterion is a more rigorous than Chauvenet's criterion. Peirce's criterion is also able to remove several suspicious data. It is an iterative method based on theory of probability. In [2] it is described and also required conditions for rejection are presented. We will explain the crucial parts of Peirce's criterion. The principle is that the k suspicious data *should be rejected when the probability* P, with all data including the suspicious data, *of the system of the errors is less then* a probability without suspicious data multiplied by *the probability of making* those suspicious observations P_1.

$$P<P_1 \tag{5}$$

Whence [3],

$$\lambda^{N-k}R^k < Q^N . \tag{6}$$

$$\lambda^2 = \frac{N-m-kn^2}{N-m-k} \tag{7}$$

$$R = e^{\frac{(n^2-1)}{2}} erf\,(\frac{n}{\sqrt{2}}) \tag{8}$$

$$Q^N = \frac{k^k (N-k)^{N-k}}{N^N}, \tag{9}$$

where

k	number of suspicious data
m	the number of unknown quantities contained in the observations.

Equation (6) doesn't have explicit solution. We can solve it on several ways. With some numerical method like Newton-Raphson method or we can use Gould's proposal that he gave in [3]. First to calculate the value for Q^N, it is a constant, then R with arbitrary n_a, and then recalculate new n_n from (7). For large N use logarithm. After that we can repeats all process with updated $n_a = n_n$, until the error $|n_n - n_a|$ is enough small, $n_a \cong n_n = n$. Final n is the number of σ. If suspicious data excide ±nσ, they can be removed. The number of suspicious data $k \in [1, \lfloor N/2 \rfloor]$, there is no point to calculate n for grater k because, if $k > \lfloor N/2 \rfloor$ then we can consider that all data are suspicious and it will be smart to repeat measurements.

3. Results

Results of equation (4) for $p \in (0,1]$ and $N \in [3,50]$ are shown in Fig. 1. Above the function is area of outliers and under the function is area of valid data.

Results for calculations of n with equation (6) for $N \in [3,50]$ and $k \in [1, \lfloor N/2 \rfloor]$ are shown in Fig. 1. The value of n for k greater then $\lfloor N/2 \rfloor$ is zero only for easier representation, but the real value isn't defined. For small number of data and only one suspicious observation the Peirce's criterion is more rigorous as it is shown in Fig. 3.

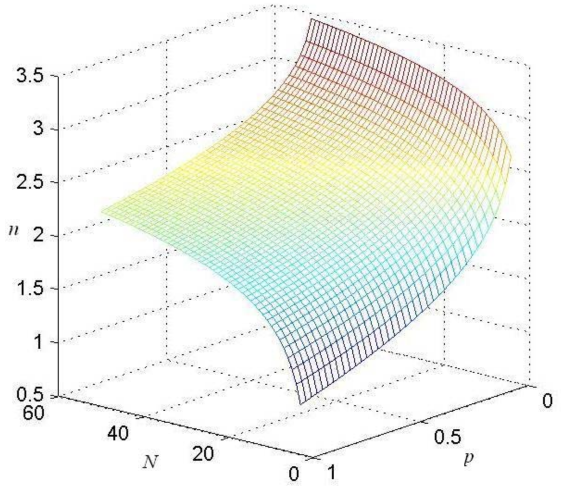


Fig. 1. Parameter *n*, from function (4), for $p \in (0,1]$.

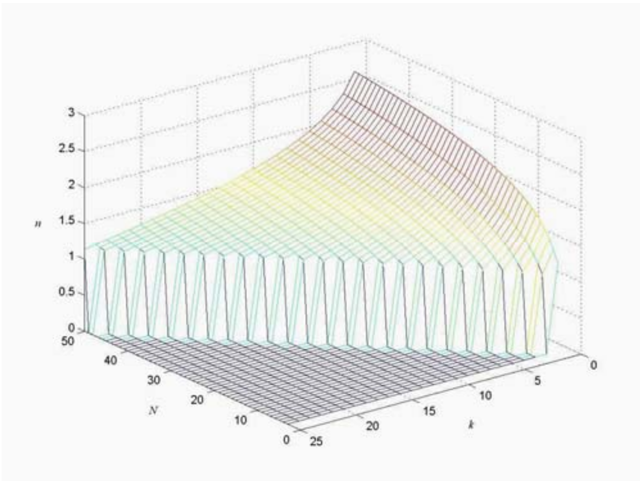


Fig. 2. Parameter *n* calculated from function (6) for $k \in [1, \lfloor N/2 \rfloor]$.

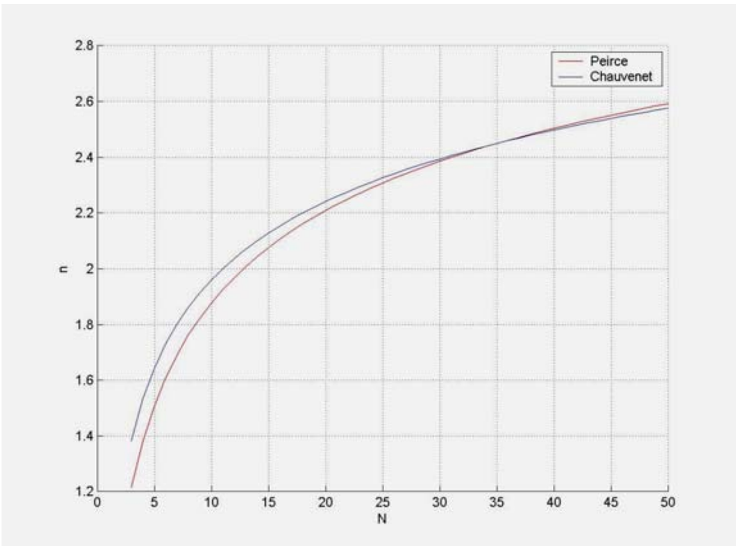


Fig. 3. Parameter *n* calculated from function (4) with *p*=0,5 and from function (6).

4. Conclusions

Chauvenet's criterion is frequently used for removing suspicious data, and also for removing several suspicious data without exact justification. For those data, where are possible more then one suspicious data we must use Peirce’s criterion, or some other iterative algorithm. Chauvenet’s criterion create fix borders independent from number of suspicious data, so second and other removed data aren’t removed because their probability to be outliers, but due to the probability of one data to be outlier.

References

[1] Taylor J. R.: An introduction to Error Analysis, University Science Books, Sausalito, 1997.

[2] Benjamin P., Criterion for the rejection of doubtful observations, *Astronomical Journal*, 45, 161-163, 1852.

[3] Gould B. A., On Peirce's criterion for the rejection of doubtful observations, with tables for facilitating its application, *Astronomical Journal*, 83, 81-83, 1855.

Exact Likelihood Ratio Test for the Parameters of the Linear Regression Model with Normal Errors

M. Chvosteková, V. Witkovský

Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovakia
Email: chvosta@gmail.com

Abstract. *We present an exact likelihood ratio (LR) based test for testing the simple null hypothesis on all parameters of the linear regression model with normally distributed errors. In particular, we consider simultaneous test for the regression parameters (beta) and the error standard deviation (sigma). The critical values of the LR test are presented for small sample sizes and small number of explanatory variables with standard significance level, alpha = 0.05.*

Keywords: Exact Likelihood Ratio Test, Linear Regression Model, Simultaneous Tolerance Intervals

1. Introduction

In the paper we present an exact likelihood ratio test (LRT) for testing the simple null hypothesis, $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ against the alternative $H_1 : (\beta, \sigma) \neq (\beta_0, \sigma_0)$, on the parameters β and σ of the linear regression model $Y = X\beta + \sigma Z$ with normally distributed errors, $Z \sim N(0, I_n)$. Although the derivation of the exact distribution of the likelihood-ratio based test statistic under the null hypothesis H_0 is straightforward, it seems that the result is not available in the standard statistical literature on linear regression models. The critical values of the LR test are presented for small sample sizes $n = k + 1, \ldots, 100$ with different number of explanatory variables, $k = 1, \ldots, 10$, and significance level 0.05.

2. Likelihood RatioTest of the Hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$

Consider the linear regression model $Y = X\beta + \sigma Z$ with normally distributed errors, where Y represents the n-dimensional random vector of response variables, X is the $n \times k$ matrix of non-stochastic explanatory variables (for simplicity, here we assume that X is a full-rank matrix), β is a k-dimensional vector of regression parameters, Z is an n-dimensional vector of standard normal errors, i.e. $Z \sim N(0, I_n)$, and σ is the error standard deviation, $\sigma > 0$.

Here we consider likelihood-ratio (LR) based test for testing the simple null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ against the alternative $H_1 : (\beta, \sigma) \neq (\beta_0, \sigma_0)$. Based on the above assumptions the log-likelihood function, denoted as $\ell(\beta, \sigma | Y = y)$, is given by

$$\ell(\beta, \sigma | y) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta). \quad (1)$$

The (-2)-multiple of the likelihood ratio test (LRT) statistic, say $\lambda(y)$ for observed value y of Y, for testing the null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ is given by

$$\begin{aligned}\lambda(y) &= -2\left(\sup_{(\beta,\sigma)\in H_0} \ell(\beta, \sigma|y) - \sup_{(\beta,\sigma)} \ell(\beta, \sigma|y)\right) = -2\left(\ell(\beta_0, \sigma_0|y) - \ell(\hat{\beta}_{ML}, \hat{\sigma}_{ML}|y)\right) \\ &= \frac{1}{\sigma_0^2}(y - X\beta_0)'(y - X\beta_0) - n\log\left(\frac{\hat{\sigma}^2_{ML}}{\sigma_0^2}\right) - n,\end{aligned} \quad (2)$$

where $\hat{\beta}_{ML} = \hat{\beta} = (X'X)^{-1}X'y$ is the standard least squares estimate (LSE) of β (which is also the MLE of β) and $\hat{\sigma}_{ML}$ is the maximum likelihood estimate (MLE) of the standard deviation σ, i.e. $\hat{\sigma}_{ML} = \sqrt{\frac{1}{n}(y - X\hat{\beta})'(y - X\hat{\beta})}$. Under given model assumptions, and under the null hypothesis H_0, it is straightforward to derive the distribution of the test statistic $\lambda(Y)$:

$$\begin{aligned}
\lambda(Y) &\sim \frac{1}{\sigma_0^2}(Y - X\beta_0)'(Y - X\beta_0) - n\log\left(\frac{(Y - X\beta_0)'M_X(Y - X\beta_0)}{n\sigma_0^2}\right) - n \\
&\sim Z'Z - n\log\left(Z'M_XZ\right) + n\left(\log(n) - 1\right) \\
&\sim Z'(P_X + M_X)Z - n\log\left(Z'M_XZ\right) + n\left(\log(n) - 1\right) \\
&\sim Q_k + Q_{n-k} - n\log\left(Q_{n-k}\right) + n\left(\log(n) - 1\right),
\end{aligned} \tag{3}$$

where $P_X = X(X'X)^{-1}X'$, $M_X = I_n - P_X$, $Z \sim N(0, I_n)$, $Q_k \sim \chi^2_k$ and $Q_{n-k} \sim \chi^2_{n-k}$ are two independent random variables with chi-square distributions, with k and $n-k$ degrees of freedom, respectively. This LRT rejects the null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ for large values of the observed test statistic $\lambda(y)$, i.e. for the given significance level $\alpha \in (0, 1)$ the test rejects the null hypothesis if

$$\lambda(y) > \lambda_{1-\alpha}, \tag{4}$$

where $\lambda_{1-\alpha}$ is the $(1-\alpha)$-quantile of the distribution of the random variable $\lambda(Y)$, given by Eq. (3). The quantiles $\lambda_{1-\alpha}$ could be evaluated numerically, by inverting the cumulative distribution function, of the random variable $\lambda(Y)$, denoted by $\mathcal{F}_{LR}(\cdot)$:

$$\begin{aligned}
\mathcal{F}_{LR}(x) &= \Pr(\lambda(Y) \leq x) \\
&= \Pr(Q_k \leq x - Q_{n-k} + n\log(Q_{n-k}) - n\left(\log(n) - 1\right)) \\
&= \int_0^\infty \mathcal{F}_{\chi^2_k}\left(x - q_{n-k} + n\log(q_{n-k}) - n(\log(n) - 1)\right) f_{\chi^2_{n-k}}(q_{n-k})\,\mathrm{d}q_{n-k},
\end{aligned} \tag{5}$$

where $\mathcal{F}_{\chi^2_k}(\cdot)$ denotes the cumulative distribution function of the chi-square distribution with k degrees of freedom, and $f_{\chi^2_{n-k}}(\cdot)$ denotes the probability density function of the chi-square distribution with $n-k$ degrees of freedom. For illustration, the critical values of the LR test are presented in Table 1 for different number of explanatory variables, $k = 1, \ldots, 10$, selected small sample sizes, $n = k + 1, \ldots, 100$, and the significance level $\alpha = 0.05$. Notice that since the family of normal distributions meets regularity conditions, from standard asymptotic result about the distribution of the LRT we get $\lambda_{1-\alpha} \to \chi^2_{k+1,1-\alpha}$as $n \to \infty$, where by $\chi^2_{k+1,1-\alpha}$ we denote the $(1-\alpha)$-quantile of chi-square distribution with $k+1$ degrees of freedom.

The LRT could be equivalently based on the test statistic $F^\star$ defined as $F^\star = \lambda(Y)/(kS^2/\sigma_0^2)$, where $S^2 = (Y - X\hat{\beta})'(Y - X\hat{\beta})/(n-k)$ and $\hat{\beta} = (X'X)^{-1}X'Y$:

$$F^\star = \frac{1}{k}\frac{(\hat{\beta} - \beta_0)'X'X(\hat{\beta} - \beta_0)}{S^2} + \frac{n-k}{k} - \frac{n}{k}\frac{\log\left((n-k)S^2/n\sigma_0^2\right) + 1}{S^2/\sigma_0^2}. \tag{6}$$

Note, that the leading term in $F^\star$ is the standard F-statistic for testing the hypothesis on regression parameters $H_0 : \beta = \beta_0$ against the alternative $H_1 : \beta \neq \beta_0$. Under null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ we directly get

$$F^\star \sim \frac{Q_k/k}{Q_{n-k}/n-k} + \frac{n-k}{k} - \frac{n}{k}\frac{\log(Q_{n-k}/n) + 1}{Q_{n-k}/n-k}. \tag{7}$$

Then, the test rejects the null hypothesis if

$$F^{\star}_{obs} > F^{\star}_{1-\alpha}, \tag{8}$$

where $F^{\star}_{obs}$ denotes the observed value of the statistic $F^{\star}$ and $F^{\star}_{1-\alpha}$ is the $(1-\alpha)$-quantile of the distribution of the random variable $F^{\star}$. The quantiles $F^{\star}_{1-\alpha}$ could be evaluated by inverting the cumulative distribution function of the random variable $F^{\star}$, denoted by $\mathcal{F}_{F^{\star}}(x)$:

$$\begin{aligned}\mathcal{F}_{F^{\star}}(x) &= \Pr(F^{\star} \leq x)\\ &= \Pr\left(Q_k \leq \frac{xkQ_{n-k}}{n-k} - Q_{n-k} + n\left(\log\left(\frac{Q_{n-k}}{n}\right)+1\right)\right)\\ &= \int_0^{\infty} \mathcal{F}_{\chi^2_k}\left(\frac{xkq_{n-k}}{n-k} - q_{n-k} + n\left(\log\left(\frac{q_{n-k}}{n}\right)+1\right)\right) f_{\chi^2_{n-k}}(q_{n-k})\,\mathrm{d}q_{n-k}.\end{aligned} \tag{9}$$

The MATLAB function for computing the quantiles $\lambda_{1-\alpha}$ and $F^{\star}_{1-\alpha}$ is available upon request from the authors. More details on the numerical algorithm, as well as on its possible application for construction of the simultaneous tolerance intervals, could be found in the extended version of the paper, in Chvosteková and Witkovský (2009).

3. Discussion

The exact LR test for testing the simple null hypothesis $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ could be directly used to construct the exact confidence region for the parameters of the linear regression model. In particular, the exact $(1-\alpha)$-confidence region for the parameters β and σ is given as $\mathcal{C}_{1-\alpha}(Y) = \{(\beta, \sigma) : \lambda(Y) \leq \lambda_{1-\alpha}\}$. Moreover, this could be directly used for constructing the simultaneous tolerance intervals in linear regression model with normal errors, as suggested in Witkovský and Chvosteková (2009). These intervals are constructed such that, with confidence coefficient $1-\alpha$, we can claim that at least a specified proportion, say $1-\gamma$ of the population is contained in the tolerance interval, for all possible values of the predictor variates, see e.g. Lieberman and Miller (1963), Limam and Thomas (1988), De Gryze et al (2007), and Krishnamoorthy and Mathew (2009). For further details see also Chvosteková and Witkovský (2009).

Acknowledgements

The research was supported by the grants VEGA 1/0077/09, VEGA 2/7087/27, APVV-SK-AT-0003-09, and APVV-RPEU-0008-06.

References

[1] Chvosteková M. and Witkovský V. Exact likelihood ratio test for the parameters of the linear regression model with normal errors. *Measurement Science Review*, 9(1): 1–8, 2009.

[2] De Gryze S., Langhans I. and Vandebroek M. Using the correct intervals for prediction: A tutorial on tolerance intervals for ordinary least-squares regression. *Chemometrics and Intelligent Laboratory Systems*, 87: 147–154, 2007.

[3] Krishnamoorthy K. and Mathew T. *Statistical Tolerance Regions: Theory, Applications, and Computation*. Wiley, ISBN: 978-0-470-38026-0, 512 pages, 2009.

[4] Lieberman G. J. and Miller R. G. Jr. "Simultaneous Tolerance Intervals in Regression. *Biometrika,* 50: 155–168, 1963.

[5] Limam M.M.T. and Thomas D.R. Simultaneous Tolerance Intervals for the Linear Regression Model. *Journal of the American Statistical Association*, 83(403): 801–804, 1988.

[6] Witkovský V. and Chvosteková M. Simultaneous Tolerance Intervals for the Linear Regression Model. In *MEASUREMENT 2009. Proceedings of the International Conference on Measurement*, Smolenice, May 20–23, 2009.

Table 1. Critical values of the likelihood ratio test (LRT) for testing the null hypothesis on parameters of the normal linear regression model with $k = 1, \ldots, 10$ explanatory variables, $H_0 : (\beta, \sigma) = (\beta_0, \sigma_0)$ against the alternative $H_1 : (\beta, \sigma) \neq (\beta_0, \sigma_0)$, for selected small sample sizes $n = k + 1, \ldots, 100$ and the significance level $\alpha = 0.05$.

n/k	1	2	3	4	5	6	7	8	9	10
2	11.8545	-	-	-	-	-	-	-	-	-
3	8.8706	19.3470	-	-	-	-	-	-	-	-
4	7.8893	13.4989	27.1540	-	-	-	-	-	-	-
5	7.4046	11.5844	18.3296	35.2052	-	-	-	-	-	-
6	7.1164	10.6358	15.4138	23.3410	43.4538	-	-	-	-	-
7	6.9257	10.0694	13.9556	19.3806	28.5094	51.8679	-	-	-	-
8	6.7901	9.6927	13.0778	17.3814	23.4748	33.8153	60.4240	-	-	-
9	6.6888	9.4241	12.4904	16.1687	20.9120	27.6847	39.2429	69.1047	-	-
10	6.6103	9.2228	12.0691	15.3518	19.3464	24.5412	31.9998	44.7793	77.8962	-
11	6.5477	9.0663	11.7521	14.7630	18.2858	22.6089	28.2622	36.4110	50.4142	86.7873
12	6.4966	8.9412	11.5047	14.3179	17.5176	21.2931	25.9522	32.0684	40.9101	56.1388
13	6.4541	8.8388	11.3063	13.9693	16.9345	20.3360	24.3719	29.3717	35.9540	45.4906
14	6.4182	8.7536	11.1435	13.6888	16.4763	19.6068	23.2179	27.5192	32.8629	39.9134
15	6.3874	8.6813	11.0075	13.4581	16.1065	19.0319	22.3357	26.1616	30.7317	36.4216
16	6.3608	8.6195	10.8923	13.2649	15.8015	18.5667	21.6383	25.1207	29.1648	34.0061
17	6.3375	8.5659	10.7933	13.1008	15.5456	18.1821	21.0724	24.2955	27.9601	32.2250
18	6.3170	8.5190	10.7074	12.9596	15.3278	17.8587	20.6035	23.6244	27.0028	30.8522
19	6.2988	8.4776	10.6321	12.8369	15.1400	17.5829	20.2085	23.0673	26.2225	29.7589
20	6.2825	8.4408	10.5656	12.7292	14.9765	17.3448	19.8711	22.5971	25.5736	28.8659
21	6.2679	8.4079	10.5064	12.6339	14.8329	17.1371	19.5793	22.1946	25.0249	28.1220
22	6.2546	8.3783	10.4533	12.5490	14.7056	16.9544	19.3244	21.8462	24.5546	27.4919
23	6.2426	8.3515	10.4056	12.4728	14.5920	16.7923	19.0998	21.5414	24.1468	26.9512
24	6.2316	8.3271	10.3623	12.4041	14.4901	16.6475	18.9004	21.2725	23.7897	26.4816
25	6.2216	8.3048	10.3230	12.3419	14.3981	16.5175	18.7221	21.0335	23.4743	26.0700
26	6.2123	8.2844	10.2870	12.2852	14.3146	16.3999	18.5618	20.8196	23.1936	25.7060
27	6.2038	8.2657	10.2540	12.2334	14.2386	16.2932	18.4167	20.6270	22.9421	25.3817
28	6.1959	8.2483	10.2236	12.1858	14.1689	16.1959	18.2849	20.4526	22.7155	25.0910
29	6.1885	8.2323	10.1955	12.1420	14.1050	16.1067	18.1646	20.2941	22.5102	24.8288
30	6.1817	8.2174	10.1695	12.1014	14.0460	16.0247	18.0543	20.1492	22.3234	24.5911
40	6.1328	8.1115	9.9864	11.8187	13.6384	15.4640	17.3081	19.1807	21.0900	23.0435
50	6.1038	8.0497	9.8809	11.6577	13.4094	15.1533	16.9006	18.6599	20.4377	22.2394
60	6.0847	8.0092	9.8122	11.5537	13.2627	14.9557	16.6437	18.3343	20.0335	21.7459
70	6.0712	7.9806	9.7640	11.4811	13.1606	14.8190	16.4668	18.1115	19.7584	21.4120
80	6.0611	7.9594	9.7282	11.4274	13.0855	14.7187	16.3376	17.9493	19.5590	21.1709
90	6.0533	7.9429	9.7006	11.3861	13.0279	14.6421	16.2391	17.8259	19.4078	20.9886
100	6.0470	7.9299	9.6788	11.3534	12.9823	14.5816	16.1615	17.7290	19.2892	20.8460
∞	5.9915	7.8147	9.4877	11.0705	12.5916	14.0671	15.5073	16.9190	18.3070	19.6751

On Two-Sided Statistical Tolerance Intervals for Normal Distributions with Unknown Parameters

I. Janiga, I. Garaj

Slovak University of Technology, Bratislava, Slovak Republic
Email: ivan.janiga@stuba.sk

Abstract. *The equation for computing factors for exact two-sided tolerance limits for a normal distribution with unknown mean and variability was developed in [1] and realised in [2]. Later it was discovered that the equation can be used also in computing factors for more than one normal distribution with unknown means and unknown common variability. The results are presented in [3]. In the paper is given unified approach to two-sided statistical tolerance intervals for normal distributions with unknown parameters.*

Keywords: Statistical Tolerance Intervals, Tolerance Factor, Estimate of Common Variability

1. Introduction

The paper deals with a unified approach to exact two-sided statistical tolerance limits for one normal distribution with unknown parameters (see [2]) and for $m \geq 2$ normal distributions with possible different unknown means and unknown common variability (see [3]).

2. Statistical tolerance intervals for m distributions with common variability

Let measurements $(x_{i1}, x_{i2}, \cdots, x_{in})$ be values of m random samples of the same sizes n drawn from m populations. We assume that measured values x_{ij} are realizations of independent normally distributed random variables X_{ij} with mean value μ_i and common variability σ^2, that is $X_{ij} \sim N(\mu_i, \sigma^2)$, $i = 1, 2, ..., m$; $j = 1, 2, ..., n$. The parameters μ_i and σ^2 are supposed to be unknown.

We are looking for two-sided intervals, which with confidence $1-\alpha$ $(0 < \alpha < 1)$ cover at least the fraction p $(0 < p < 1)$ of values of the distributions $N(\mu_i, \sigma^2)$, $i = 1, 2, ..., m$. These intervals are named 100 p % statistical tolerance intervals.

If $X_i \sim N(\mu_i, \sigma^2)$, $i = 1, 2, ..., m$ then two-sided statistical tolerance intervals for distributions $N(\mu_i, \sigma^2)$, $i = 1, 2, ..., m$ are intervals

$$\left(\bar{x}_i - ks_P,\ \bar{x}_i + ks_P\right),\ i = 1, 2, ..., m \tag{1}$$

where

$\bar{x}_i$ value of a sample mean of the *i*th random sample

k tolerance factor

s_P^2 estimate of the common variability (sometimes called “pooled”).

The value $\bar{x}_i$ is an unbiased estimate of unknown parameter μ_i computed by the formula

$$\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij} \tag{2}$$

and the value s_P^2 is an unbiased estimate of the unknown common parameters σ^2

$$s_P^2 = \frac{1}{m(n-1)} \sum_{i=1}^{m} \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)^2 \tag{3}$$

The intervals defined in (1) have to fulfil the conditions

$$P\left[P\left(\bar{x}_i - ks_P < X_i < \bar{x}_i + ks_P\right) \geq p\right] = 1 - \alpha\,, \quad i = 1, 2, ..., m \tag{4}$$

from which the exact formula for computation of tolerance factors was derived (see [1]).

3. Exact formula for computation of statistical tolerance factors

The value of a tolerance factor is the solution of the following equation (derived in [1])

$$\sqrt{\frac{n}{2\pi}} \int_{-\infty}^{\infty} F(x,k)\, e^{-\frac{nx^2}{2}}\, \mathrm{d}x - 1 + \alpha = 0 \tag{5}$$

where

$$F(x,k) = \int_{\frac{\nu R^2(x)}{k^2}}^{\infty} \frac{t^{\frac{\nu}{2}-1} e^{-\frac{t}{2}}}{2^{\frac{\nu}{2}}\, \Gamma\left(\frac{\nu}{2}\right)} \mathrm{d}t$$

$R(x)$ solution of the equation $\Phi(x+R) - \Phi(x-R) - p = 0$.

The values of statistical tolerance factors $k = k(n, \nu, p, 1-\alpha)$ are computed from (5) for $\nu = m(n-1)$ and selected $n, p, 1-\alpha$. Let us suppose two cases.

Case 1 ($m = 1$ **)**: It means that we consider only one sample. Then we obtain from the equation (5) a value of statistical tolerance factor $k = k(n, \nu, p, 1-\alpha)$, where $\nu = n - 1$.

Case 2 ($m > 1$ **)**: We consider m samples and the equation (5) give us the value of $k = k(n, \nu, p, 1-\alpha)$, where $\nu = m(n-1) \neq n - 1$.

4. Comparison of both cases

The comparison was performed by using data. Suppose the percentage of solids in each of four batches of wet brewer’s yeast (A, B, C and D), each from a different supplier, was to be determined. The researcher wants to determine whether the suppliers differ so that decisions can be made regarding future orders.

The random samples from each batch were collected and data are presented in Table 1.

Table 1. Percentages of total solids in four batches of brewner´s yeast

n	1	2	3	4	5	6	7	8	9	10
Batch A	20	18	16	21	19	17	20	16	19	18
Batch B	19	14	17	13	10	16	14	12	15	11
Batch C	11	12	14	10	8	10	13	9	12	8
Batch D	10	7	11	9	6	11	8	12	13	14

For comparing the suppliers there was decided to use 95 % two-sided statistical tolerance intervals with confidence equals $1-\alpha = 0{,}95$.

A good fit with normal distributions with unknown parameters μ_i and σ_i^2 $i = \mathrm{A, B, C, D}$ was confirmed by Shapiro-Wilks W statistics. P-values of all batches are greater than or equal to 0,10 so we can conclude that batches come from normal distributions with 90% or higher confidence (see Table 2).

Table 2. P-values of W tests for normality of batches

Batch	A	B	C	D
P value	0,61	0,99	0,69	0,95

On the basis of the data the values of sample means and standard deviations were computed, that are $\bar{x}_\mathrm{A} = 18{,}4$; $s_\mathrm{A} = 1{,}7127$; $\bar{x}_\mathrm{B} = 14{,}1$; $s_\mathrm{B} = 2{,}76687$; $\bar{x}_\mathrm{C} = 10{,}7$; $s_\mathrm{C} = 2{,}05751$; $\bar{x}_\mathrm{D} = 10{,}1$ and $s_\mathrm{D} = 2{,}60128$.

Table 3. Two-sided statistical tolerance factors for unknown common variability σ^2

$1-\alpha = 0{,}95$; $p = 0{,}95$; $\nu = m(n-1)$					
n \ m	1	2	3	4	5
8	3,7456	3,0609	2,8357	2,7201	2,6488
9	3,5459	2,9541	2,7548	2,6515	2,5873
10	**3,3935**	2,8700	2,6904	**2,5964**	2,5377

Case 1 ($m = 1$): It is required to compute the 95 % two-sided tolerance interval with the confidence level 95 %. The value of $k = k(n, \nu, p, 1-\alpha) = k(10,\, 9,\, 0{,}95,\, 0{,}95) = 3{,}3935$ can be found in Table 3. Then statistical tolerance intervals for batches A, B, C, D are as follows

$$\begin{aligned}
&\text{A: } 18{,}40 \mp 3{,}3935 \times 1{,}7127 \;\Rightarrow\; (12{,}59;\, 24{,}21)\\
&\text{B: } 14{,}10 \mp 3{,}3935 \times 2{,}76687 \;\Rightarrow\; (4{,}71;\, 23{,}49)\\
&\text{C: } 10{,}70 \mp 3{,}3935 \times 2{,}05751 \;\Rightarrow\; (3{,}72;\, 17{,}68)\\
&\text{D: } 10{,}10 \mp 3{,}3935 \times 2{,}60128 \;\Rightarrow\; (1{,}27;\, 18{,}93)
\end{aligned}$$

Case 2 ($m > 1$):It is also required to compute the 95 % two-sided tolerance interval with the confidence level 95 %. First of all it is needed to check variances of the batches A, B, C, D. In testing the null hypothesis H_0: $\sigma_\mathrm{A}^2 = \sigma_\mathrm{B}^2 = \sigma_\mathrm{C}^2 = \sigma_\mathrm{D}^2$ were used Cochran's C test (P-Value = 0,605823), Bartlett's test (P-Value = 0,496912) and Levene's test (P-Value = 0,559791). Since the smallest of the P-values is greater than or equal to 0,10 it can be concluded that batches have the common variability σ^2. Then estimates of the unknown common variability and standard deviation computed from data are $\sigma_\mathrm{est}^2 = s_P^2 = 4{,}6463846$ (see (3)) and $\sigma_\mathrm{est} = s_P = 2{,}3231923$.

Now the value of $k = k(n, \nu, p, 1-\alpha) = k(10,\, 36,\, 0{,}95,\, 0{,}95) = 2{,}5964$ can be found in Table 3. Then statistical tolerance intervals for batches A, B, C, D are as follows

$$\begin{aligned}
&\text{A: } 18{,}40 \mp 2{,}5964 \times 2{,}3231923 \;\Rightarrow\; (12{,}36;\, 24{,}43)\\
&\text{B: } 14{,}10 \mp 2{,}5964 \times 2{,}3231923 \;\Rightarrow\; (8{,}07;\, 20{,}13)
\end{aligned}$$

$$\text{C:}\ \ 10{,}70 \mp 2{,}5964 \times 2{,}3231923 \ \Rightarrow\ (4{,}67; 16{,}73)$$
$$\text{D:}\ \ 10{,}10 \mp 2{,}5964 \times 2{,}3231923 \ \Rightarrow\ (4{,}07; 16{,}13)$$

5. Discussion and Conclusions

The unified approach to two-sided statistical tolerance intervals for normal distribution with unknown parameters was developed in chapters 2 and 3. The equation (5) was used for computation of the values of statistical tolerance factors $k = k(n, v, p, 1-\alpha)$ for the selected $n = 2$ (1) 10; $m = 1$ (1) 4; $v = m(n-1)$; $p = 0{,}95$ and $1-\alpha = 0{,}95$ (see Table 3).

In the first case ($m = 1,\ \ n = 10$) the 95 % two-sided tolerance intervals with the confidence level 95 % were computed for all batches. In the computation there were used the value of tolerance factor 3,3935 given in Table 3 and the values of standard deviation $s_{\mathrm{A}} = 1{,}7127$; $s_{\mathrm{B}} = 2{,}76687$; $s_{\mathrm{C}} = 2{,}05751$; $s_{\mathrm{D}} = 2{,}60128$.

Then in the second case ($m = 4,\ \ n = 10$) the 95 % two-sided tolerance intervals with the confidence level 95 % were also estimated for all batches. But in this case the value of tolerance factors 2,5964 from Table 3 and the estimate of the unknown common standard deviation $s_{\mathrm{P}} = 2{,}3231923$ were used.

When comparing the result of the both cases it can be declared that the statistical tolerance intervals for batches B, C, D are significantly much smaller in the second case than in the first one. But the statistical tolerance interval for batch A is significantly a little larger in the first case.

We can conclude that the tolerance intervals computed simultaneously for several populations can yield intervals shorter than the tolerance intervals computed for each random sample separately, provided that the underlying normal populations have the same variance. This nice property follows from the fact that on the average the estimate of the variance computed from several random samples is “better” than the estimate computed from one random sample, because this is based on smaller number of observations.

Acknowledgements

This contribution was granted by the scientific projects VEGA 1/0437/08 and VEGA 1/4091/07 as well as VEGA 1/0374/08.

References

[1] Janiga I., Miklóš R. Statistical tolerance intervals for a normal distribution. n: Measurement Science Review. ISSN 1335-8871. - Vol. 1, Sec. 1, No. 1 (2001), s. 29-32.

[2] Garaj I., Janiga, I. Dvojstranné tolerančné medze pre neznámu strednú hodnotu a rozptyl normálneho rozdelenia. Vydavateľstvo STU, Bratislava, 2002, ISBN 80-227-1779-7.

[3] Garaj I., Janiga, I. Two sided tolerance limits of normal distributions with unknown means and unknown common variability. Vydavateľstvo STU, Bratislava, 2004, ISBN 80-227-2019-4.

[4] Terek M., Hrnčiarová Ľ. Štatistické riadenie kvality. IURA EDITION, Bratislava, 2004, ISBN 80-89047-97-1.

Minimum Detectable Value and Limit of Detection in Linear Calibration with Standard Deviation Linearly Dependent on Net State Variable

I. Janiga, J. Mocák, P. Cisko

Slovak University of Technology, Bratislava, Slovak Republic
Email: ivan.janiga@stuba.sk

Abstract. *The minimum detectable value obtained from a particular calibration is the smallest value of the net state variable, which can be detected with a probability of* $1-\beta$ *as different from zero. In this work, the procedure for the computation of the critical value and the minimum detectable value was derived from the linear calibration function following the ISO concept, and compared to the limit of detection recommended to chemists by IUPAC. It is assumed that all measurements of the response variable are independent and normally distributed with standard deviation linearly dependent on the net state variable*

Keywords: Linear Calibration, Critical Value, Minimum Detectable Value, Limit of Detection

1. Introduction

The minimum detectable value of the net state variable is an important characteristic of the measurement method and enables its optimization or selection. In order to characterize a measurement process, the minimum detectable value should be stated using appropriate data.

In this work, the procedures for the computation of the critical value of the response variable, Y, and the minimum detectable value of the net state variable, x, are based on the following basic assumptions:

a) The calibration function $y = \mathrm{f}(x)$ is linear.
b) The measurements of the response variable of all specimens J in reference states are assumed to be independent and normally distributed, with standard deviation not constant at different values of the net state variable.
c) The residual standard deviation is linearly dependent on the values of the net state variable.

It should be noted that in chemistry y is usually instrumental signal, x is concentration of the analyte, I – number of calibration standards (including zero concentration), J – number of replicate standard measurements.

2. Estimation of minimum detectable value

The following model is based on assumptions that the calibration function is linear and standard deviation linearly depends on x and is given by:

$$y_{ij} = a + bx_i + \varepsilon_{ij},\ i = 1, 2, \ldots, I;\ j = 1, 2, \ldots, J \tag{1}$$

where

x_i net state variable in the state i

a, b model parameters

y_{ij} response variable for the state i and specimen j

ε_{ij} independent random errors normally distributed with expectation $E(\varepsilon_{ij}) = 0$ and variance $V(\varepsilon_{ij}) = \sigma^2(x_i) = (c + d\,x_i)^2$

The parameters of the model, a, b, c and d were estimated in two steps (see [1]). The first step comprises iterative estimation of the linear relationship between the residual standard deviation and the net state variable. It can be easily performed e.g. in MS Excel by using

Solver option from the Tools menu and setting 3 as the number of iterations. Empirically obtained standard deviations are used as the starting estimates of $\hat{\sigma}_0$. The third iteration usually gives the final result that is

$$\hat{\sigma}(x) = \hat{c}_3 + \hat{d}_3 x = \hat{\sigma}_0 + \hat{d}x \tag{2}$$

The second step includes estimation of the calibration function parameters:

$$\hat{a} = \frac{J\sum_{i=1}^{I} w_i x_i^2 \sum_{i=1}^{I}\sum_{j=1}^{J} w_i \bar{y}_{ij} - J\sum_{i=1}^{I} w_i x_i \sum_{i=1}^{I}\sum_{j=1}^{J} w_i x_i \bar{y}_{ij}}{J\sum_{i=1}^{I} w_i J\sum_{i=1}^{I} w_i x_i^2 - \left(J\sum_{i=1}^{I} w_i x_i\right)^2} \tag{3}$$

$$\hat{b} = \frac{J\sum_{i=1}^{I} w_i \sum_{i=1}^{I}\sum_{j=1}^{J} w_i x_i \bar{y}_{ij} - J\sum_{i=1}^{I} w_i x_i \sum_{i=1}^{I}\sum_{j=1}^{J} w_i \bar{y}_{ij}}{J\sum_{i=1}^{I} w_i J\sum_{i=1}^{I} w_i x_i^2 - \left(J\sum_{i=1}^{I} w_i x_i\right)^2} \tag{4}$$

where w_i denotes the i-th weight

$$w_i = \frac{1}{\hat{\sigma}^2(x_i)} = \frac{1}{\left(\hat{\sigma}_0 + \hat{d}x_i\right)^2} .$$

The critical value of the response variable is derived by expressing the variance $V(\bar{y} - \hat{a})$ between the average $\bar{y}$ at the basic state and the estimated intercept as:

$$y_c = \hat{a} + t(v, 0.95)\sqrt{\hat{\sigma}_0^2 + \left(\frac{1}{J\sum_{i=1}^{I} w_i} + \frac{\bar{x}_w^2}{S_{xxw}}\right)\hat{\sigma}_{yxw}{}^2} \tag{5}$$

and the critical value of the net state variable is

$$x_c = \frac{t(v, 0.95)}{\hat{b}}\sqrt{\hat{\sigma}_0^2 + \left(\frac{1}{J\sum_{i=1}^{I} w_i} + \frac{\bar{x}_w^2}{S_{xxw}}\right)\hat{\sigma}_{yxw}{}^2} \tag{6}$$

where

$t(v, 0.95)$ - critical t- value for the number of degrees of freedom ν and quantil $(1-\alpha)$

$$\bar{x}_w = \sum_{i=1}^{I} w_i x_i \Big/ \sum_{i=1}^{I} w_i$$

$$S_{xxw} = J\sum_{i=1}^{I} w_i (x_i - \bar{x}_w)^2$$

$$\hat{\sigma}_{yxw}{}^2 = \frac{1}{I \cdot J - 2}\sum_{i=1}^{I}\sum_{j=1}^{J} w_i (\bar{y}_{ij} - \hat{a} - \hat{b}x_i)^2 .$$

Finally, the minimum detectable value of the net state variable is given by

$$x_d = \frac{\delta(\nu,\alpha,\beta)}{\hat{b}} \sqrt{\hat{\sigma}^2(x_d) + \left(\frac{1}{J\sum_{i=1}^{I} w_i} + \frac{\overline{x}_w^2}{S_{xxw}} \right) \hat{\sigma}_{yxw}{}^2} \tag{7}$$

where

$\delta = \delta(\nu, \alpha, \beta)$ - the non-centrality parameter of the non-central t distribution [1,2]

$\nu = I \cdot J - 2$

Since $\hat{\sigma}^2(x_d)$ depends on the value of x_d yet to be calculated, x_d has to be calculated iteratively; three iterations are usually sufficient.

3. Results and Discussion

Measurement results obtained in calibration procedure used for quantitative analysis 4-amino-naphtalene-1-sulfonic acid (NSA) are summarized in Table 1. The second column contains the NSA concentrations, the third column involves the mean signal values calculated from three replicate signal measurements (J=3), the fourth column comprises the values of empirical standard deviation calculated from three replicates. The results of iterative calculation of the linear relationship between the residual standard deviation and NSA concentration are located in the right upper part of the table. Results of final statistical analysis are collected in the right bottom part of the table.

It happens frequently in chemical measurements that very small amounts of the component of interest are concerned. Then it is important to distinguish small concentration values from the zero concentration, given by the blank (containing all accompanying sample components except the main determined analyte). ISO 11843 standard procedures, characterizing the capability of detection by determining the critical value of the net state variable and, mainly, the minimum detectable value of the net state variable, are not widespread among the chemists who, instead of the mentioned characteristics, employ the limits of detection, LoD, in the signal (y) domain and, above all, in the concentration (x) domain. The recommendation of IUPAC for the LoD calculation [3] is relatively similar to the ISO calculation of x_c but it was hitherto given only for the homoscedastic case (constant standard deviation). The mentioned IUPAC approach [3] assumes a constant standard deviation and uses the one-sided upper confidence limit of the calibration line so that the LoD in the concentration domain is defined by

$$\text{LoD} = [t(\nu,1\text{-}\alpha)\,\hat{\sigma}_{yx}/\hat{b}\,]\,[1 + 1/N + \overline{\overline{x}}^{\,2}/\sum_{n=1}^{N}(x_n - \overline{x}\,)^2]^{1/2} \tag{8}$$

where $\hat{\sigma}_{yx}$ means the residual standard deviation obtained in linear regression y vs. x, and N = $I.J$ (I - the number of standard solutions, J - the number of parallel measurements). For the sake of better compatibility with older IUPAC LoD definitions, the value α = 0.01 was recommended [4].

If the standard deviation is linearly proportional to concentration, the following equation can be derived, in which all used symbols are the same as defined in part 2:

Table 1. Measurement results and results of statistical analysis

i	x_i	$\overline{y}_i$	s_i	Iterative computation of $\hat{\sigma}(x) = \hat{c}_3 + \hat{d}_3 x$			
1	0.022	4.26	0.39509		c	d	
2	0.044	6.35	0.77019	*Iteration 1*	0.100532	13.98287	
3	0.059	9.06	0.99000	*Iteration 2*	0.100532	13.98287	
4	0.073	11.33	1.13530	*Iteration 3*	0.100532	13.98287	
5	0.088	12.89	1.28508				
6	0.100	14.17	1.41500	*Results of statistical analysis*			
7	0.130	18.45	1.87502	$a =$	1.013391	$t(\upsilon,0.95) =$	1.687094
8	0.150	21.38	2.04035	$b =$	137.185145	$\hat{\sigma}_0^2 =$	0.0104346
9	0.160	23.21	2.42160	$\overline{x}_w =$	0.0514275	$\hat{\sigma}_{yxw}^2 =$	0.8619128
10	0.180	26.70	2.66010	$S_{xxw} =$	0.0763458	$y_c =$	1.442377
11	0.240	35.20	3.63113	$\upsilon =$	37	$x_c =$	0.0031271
12	0.290	40.21	4.17129	$\alpha =$	0.05	$x_d =$	0.0076365
13	0.350	50.61	5.01099	$\beta =$	0.05	LoD =	0.0045067

$$\text{LoD} = [t(\nu,1\text{-}\alpha)\,\hat{\sigma}_{yxw}/\hat{b}\,]\,[\frac{\hat{\sigma}_0^2}{\hat{\sigma}_{yxw}^2} + \frac{1}{J\sum_{i=1}^{I} w_i} + \frac{\overline{x}_w^2}{S_{xxw}}]^{1/2} \tag{9}$$

The LoD value calculated by weighted linear regression is in the last entry of Table 1.

4. Conclusions

Derived set of statistical equations was applied to linear calibration under assumption of standard deviation linearly proportional to the net state variable. Compared to the case with constant standard deviation the weights reflecting the variances at particular states of the net state variable and iterative solution of equations were applied. Comparison of the ISO approach and the IUPAC recommendation applied to the same chemical calibration revealed that the calculation procedures are similar and differ in the way of application of the t-distribution as well as in assumed probabilities.

Acknowledgements

This contribution was granted by the scientific projects VEGA 1/0437/08 and VEGA 1/4091/07 as well as VVCE-0004-07.

References

[1] Janiga I, Garaj I, Cisko P. Estimation of Minimum Detectable Value by Using Linear Calibration with Standard Deviation Linearly Dependent on Net State Variable. *Forum metricum slovacum*, 2004, Tom VIII: 60-67.

[2] Janiga I, Garaj I, Cisko P. Estimation of Minimum Detectable Value by Using Linear Calibration with Constant Standard Deviation *Forum metricum slovacum*, 2004, Tom VIII: 68-73.

[3] Mocak J., Bond A.M., Mitchell S., Schollary G.: *Pure & Applied Chemistry*, 1997, 69: 297-328.

Measurement in Biomedicine

The Electromyogram and Mechanomyogram in Monitoring Neuromuscular Fatigue: Techniques, Results, Potential Use within the Dynamic Effort

M. T. Tarata

University of Medicine and Pharmacy of Craiova, Romania
Email: mihaitarata@yahoo.com

Abstract. *The paper provides a functional, practically oriented overview of the concept and technical possibilities of monitoring the neuromuscular fatigue via the electromyographic and the mechanomyographic signals, as essentially related to the muscle contraction and intimately mirroring muscle activation and contraction mechanisms. The Fast Fourier Transform-based and Wavelet Transform-based techniques are critically analysed, illustrated and discussed, on the international arena and through original work. An original technique - the Area/Amplitude Ratio (Raa) - is also comparatively discussed, as a practical alternative instrument in monitoring the neuromuscular fatigue in a dynamic contraction - important issue in real life activity - , e.g. work in difficult environments, pilots in mission, difficult work in normal environments.*

Keywords: Area/Amplitude Ratio, Neuromuscular Fatigue, Electromyography, SEMG, Mechanomyography, MMG, Wavelet

1. Introduction

Muscle in voluntary contraction produces force, based on two mechanisms: (i) the firing frequency and (ii) the recruitment of the motor units. A motor unit is a functional entity consisting of a motor neuron and the whole set of muscular fibers it innervates. To increase the muscle force, either the firing frequency or the number of recruited motor units has to be increased. The actual muscle force is built up from individual muscle fiber twitches, the smoothness of the force output being enhanced by the firing rate / recruitment interaction within the motor unit pool [1] and the mechanical filtering effect of the tissue, compensating for the discrete nature of the process, The variation of the firing rates of the motor units occurs simultaneously for all the motor units within a muscle, and even in different muscles acting on the same joint, according to the phenomenon of 'common drive' [2], therefore small variations of the output force may occur. The control loops involving Ia, Ib afferents are also responsible for the occurrence of certain peaks within the output force in the isometric exercise, and others are due to the activation from higher levels [3].

The myoelectric activity detected with surface electrodes, surface electromyogram (SEMG), may be considered as the summation of the electrical signals generated by a number of motor units, active within the same motor territory in the proximity of the electrodes. The SEMG signal is a convenient mean to study the muscle behavior under fatiguing exercise, as it proves time-dependent changes, provided care is taken to prevent cross talk from adjacent muscles.

Sustained muscular contractions externally associated with not being able to maintain a certain force lead to physiological fatigue, tremor or pain, localized in the specific muscle (localized muscular fatigue). Fatigue is defined as 'any reduction in the force generating capacity, measured by maximum voluntary contraction (MVC), regardless of the task performed' [4]. The failure in maintaining the motor task defines the endurance limit; endurance time is the total duration of the task up to the endurance limit. Fatigue is associated with a compression of the power spectral density of the SEMG toward lower frequencies, from the very beginning of the voluntary contraction [5]. This is due to the reduction in the conduction velocity in direct relation with the muscular fiber membrane excitability and with neural adaptations, resulting in an increase of the lower frequency content of the signal. Fatigue is also associated with higher amplitudes of the SEMG signal toward the end of the exercise. It has been shown that, in sustained motor tasks, changes at different levels, including motoneural discharge behavior, develop before an endurance limit is reached - phenomenon called 'central fatigue' [6, 7]. Central and peripheral fatigue develop together, and have to be seen not as a result, but as complementary elements of a complex strategy striving to insure the optimality of the motor behavior within the framework of available resources.

Under muscular contraction, mechanical vibrations occur, due to three main processes: (i) the inner muscular vibrations, which are the intrinsic components of the muscle contraction [8], (ii) oscillations of the human motor system, e.g. tremor and clonus [9], and (iii) artefacts. They are located in specific frequency ranges, with a certain overlapping: the artefacts - due to large movements - show the lowest frequencies, possible tremor contributions are below 10 Hz, in healthy subjects usually between 5.85 and 8.8 Hz [10], and the mechanical inner vibrations affect the range between 10 and 40 Hz. As a signal complementary to SEMG, the mechanomyogram (MMG) reflects the mechanical muscle vibrations generated by the spatio-temporal summation of the individual muscle fiber twitches, evoked through motor unit (MU) activation by the motor neurons. The MMG may be considered to reflect the mechanical muscle vibrations generated by the spatio-temporal summation of the individual muscle fiber twitches which are evoked through MU activation by the motor neurons.

SEMG and MMG, recorded simultaneously from the same muscles under steady contraction, show a compression of the spectra toward lower frequencies since the beginning of the contraction [11, 12]. After an initial approach based on Fast Fourier Transform (FFT) techniques, in order to study transitional phenomena in muscle contraction and to monitor neuro muscular fatigue in a dynamic contraction, the use of the Wavelet Transform (WT) has been investigated, via Instantaneous Mean Frequency and Instantaneous Median Frequency [13]. The use of WT was shown on a rather limited scale until now, only for epochs where FFT can also be consistently used, i.e. windows of signal where no acceleration or deceleration occurs, situation which may occasionally happen in a steady contraction or in some isokinetic exercise. As an alternative, the author explored the use of the Raa parameter (Area/Amplitude Ratio) [14, 15], together with the Instantaneous Mean Scale (IMS) and Instantaneous Median Scale (IMedS) with the purpose to validate its use and to assess its computational efficiency in terms of speed and required memory space.

2. Recording SEMG and MMG

The SEMG signal is usually recorded via two disposable self adhesive surface EMG conductive gel electrodes, (e. g. 22.5 x 22.5 mm H59P, MVAP, USA), with their centres 25 mm apart from each other, placed on abrased, clean skin, longitudinally, immediately under the thickest point of the muscle or close to the innervation point.

The SEMG signals were amplified (x 2000, 100 MΩ input impedance, 100 dB CMRR, 500 Hz antialias filter, Beckman R611, USA) and acquired together with the MMG signals via a computerized acquisition system (DAP1200 Microstar Laboratories USA), at 1000 Hz sampling rate on all the channels simultaneously.

The MMG was recorded with acceleration or sound transducers and recently with laser distance sensors. The technique of recording MMG was refined using piezoresistive silicon accelerometers in a surface mount package [16] to provide a reliable acquisition of MMG. An accelerometer (+-2g, ICS Sensors, model 3031,USA) and author-made original amplifier (x 50000, 10-250 Hz band pass filter, 250 Hz antialias filter) picks up the MMG. The accelerometer was placed between the SEMG electrodes to pick up the maximal MMG, orthogonal to the muscle from the same motor territory.

The SEMG and MMG recorded simultaneously from the same muscle (Figure 1) have similar behavior [11, 17], i.e. the median frequencies of SEMG and MMG decrease from the very beginning of the contraction, and their RMS values increase.

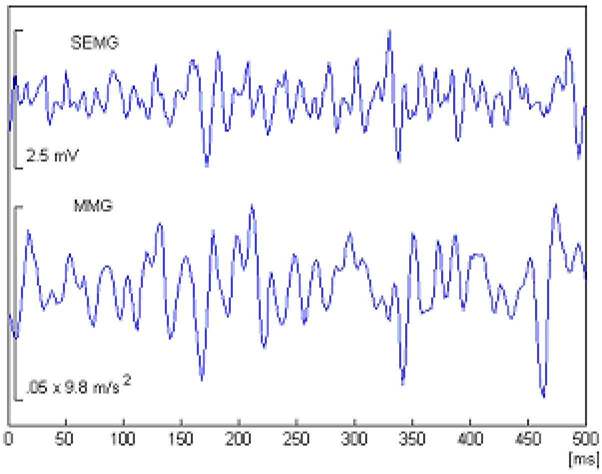


Fig. 1. SEMG and MMG recorded simultaneously from the same location of the Biceps muscle (Courtesy [11]).

Recent research showed the MMG to be a reliable signal in studying the development of fusion [18] and the changes in muscle contractile properties during repetitive unfused contractions [19], as well. Following preliminary work with refined techniques using piezoresistive silicon accelerometers in a surface mount package [15, 16], while the use of accelerometers has been further validated [20] for recording MMG, dedicated work, exploring the isometric steady contraction of different muscles, showed that the SEMG and MMG, recorded simultaneously from the same muscle, have similar behavior [11].

3. Techniques and Results

Until now, the power spectral density obtained via FFT was consistently used to compute mean (MNF) or median (MDF) frequencies to monitor activation in isometric steady contraction [21, 11, 22, 23, 24, 25]. A time window of 500 – 1000 ms was currently used, either for SEMG or for MMG in this type of contraction, thus fulfilling one major restriction in accurately using FFT to compute power spectral density from these signals, to ensure consistency across the data, according to the wide sense stationarity [26, 27, 11].

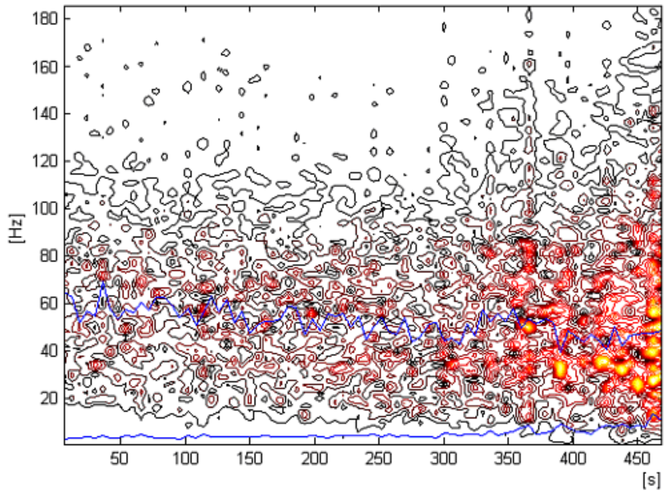


Fig. 2. SEMG power spectral density evolution in time (Biceps muscle) – contour plot of 15 levels isolines of the matrix containing the succession of the spectra, from the beginning to the end of the contraction (black – zero level; red, yellow - higher power spectral density peaks). The power spectral density compression towards lower frequency and higher peaks by the end of contraction, can be seen. The middle trace (blue) shows the SEMG MDF evolution; MDF decreases with increasing fatigue, thus proving the power spectral density compression. The lower trace (blue) shows the SEMG RMS evolution; RMS increases with increasing fatigue (Courtesy [11]).

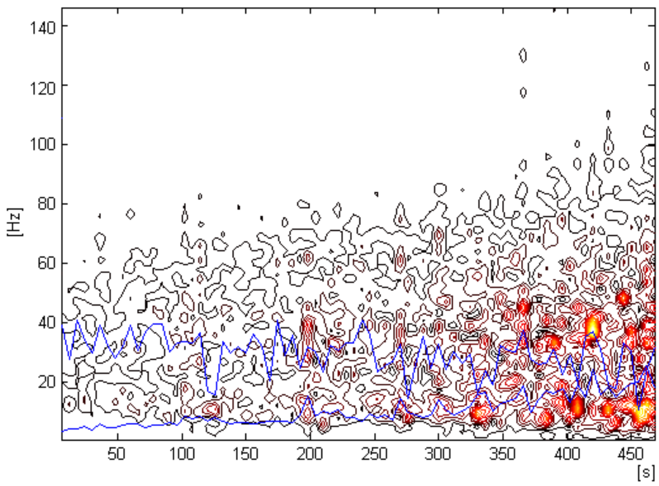


Fig. 3. MMG power spectral density evolution in time (Biceps muscle) – contour plot of 15 levels isolines of the matrix containing the succession of the spectra, from the beginning to the end of the contraction (black – zero level; red, yellow - higher power spectral density peaks). The power spectral density compression towards lower frequency and higher PSD peaks by the end of contraction, can be seen. The middle trace (blue) shows the MMG MDF evolution; MDF decreases with increasing fatigue, thus proving the power spectral density compression. The lower trace (blue) shows the MMG RMS evolution; RMS increases with increasing fatigue (Courtesy [11]).

Because of this demand, FFT has to be cautiously used to compute power spectral density from SEMG or MMG. Therefore, to avoid the necessity to satisfy the wide sense stationarity condition, the use of WT has been investigated to monitor the dynamic contraction via Instantaneous Mean Frequency and Instantaneous Median Frequency [28, 29, 30, 31]. Equivalence of this approach with the use of FFT was shown on a limited scale until now, only for epochs where FFT can also be consistently used, i.e. windows of signal where no acceleration or deceleration occurs, which may happen in some isokinetic exercise [13].

Previous work [11] showed a compression of the spectra toward lower frequencies, with advancing fatigue of the spectra, both for SEMG and MMG (Figures 2, 3). MNF, MDF, computed on a time window of 500 ms via the FFT from SEMG and MMG acquired from Biceps Brachii and Brachioradialis muscles under voluntary, steady contraction, decreased in all the subjects, from the very beginning of the contraction to its very end - when the task could not be sustained any more -, thus proving the consistence of SEMG and MMG signals in monitoring central fatigue (Figures 2, 3, 4).

Even when the wide sense stationarity is satisfied during an isometric steady contraction for time windows of 500 – 1000 ms, while the FFT approach stands in this case, other techniques are necessary to allow the exploration of the signals within time windows shorter than 500 ms, e.g. in order to study local variations in activation, with not speaking of the dynamic contraction associated with real life work, when the FFT approach fails.

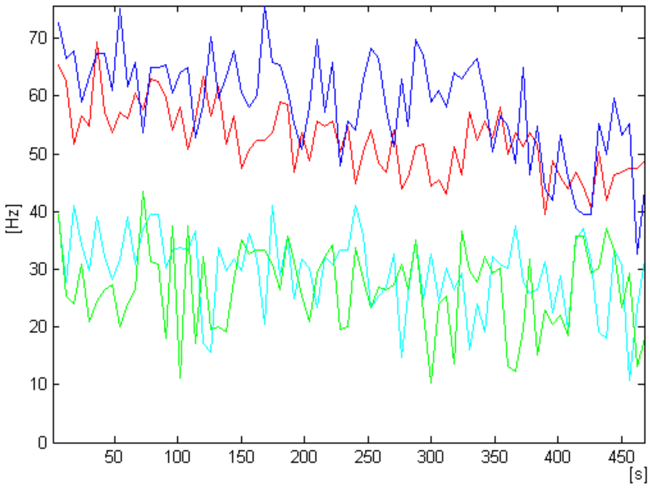


Fig. 4. SEMG MDF and MMG MDF evolutions with increasing fatigue, both for the Biceps and Brachioradialis muscles (red – Biceps SEMG MDF, blue – Brachioradialis SEMG MDF, cian – Biceps MMG MDF, green – Brachioradialis MMG MDF). An overall decrease can be noticed starting at the very beginning of the contraction, which proves the central component of the fatigue (Courtesy [11]).

The WT approach could successfully demonstrate the compression of the spectra toward lower frequencies in neuromuscular fatigue, on epochs of enough time length to allow the computation of FFT based parameters, too.

To explore shorter epochs we performed preliminary work on exploring the behavior of the IMS (Instantaneous Mean Scale) and IMedS (Instantaneous Median Scale) computed via the WT on epochs as short as 100 ms, from the SEMG and MMG signals acquired from the same muscle (Biceps Brachii) under voluntary contraction. This allowed a comparison with the evolution of Raa (Area /Amplitude Ratio), an original parameter computed from the time domain, with a dimension of time [14,15], which increases with increasing fatigue, from the very beginning of the contraction.

Raa – Average Area /Amplitude Ratio, with a dimension of time [ms], is computed from the signal in the time domain, as an average of the Area/Amplitude ratios over the considered epochs, calculated between consecutive transversals of the isoelectric line, called 'phases' [14, 15]:

$$Raa = \frac{1}{n}\sum_{i=1}^{n}\frac{S_i}{A_i} \qquad (1)$$

with
n - the number of phases within the current epoch,
S_i - current phase area, the integral of the i^{th} phase of the signal within the current signal segment,
A_i - the maximal amplitude of the i^{th} phase of the signal within the current signal segment, selected on all m samples within the current phase.

The advantage of Raa is its computational efficiency, comparing to using FFT or Wavelet techniques. As an example, figure 5 shows the evolution of the Area/Amplitude Ratio (Raa), IMS and IMedS with advancing fatigue, for both the SEMG and MMG – computed on 100 ms epochs -. All three parameters increase with advancing NMF.

Raa, IMS and IMedS show positive slopes in all subjects (SEMG and MMG), from the beginning of the contraction, for all the epoch lengths. This proves the central component of the fatigue.

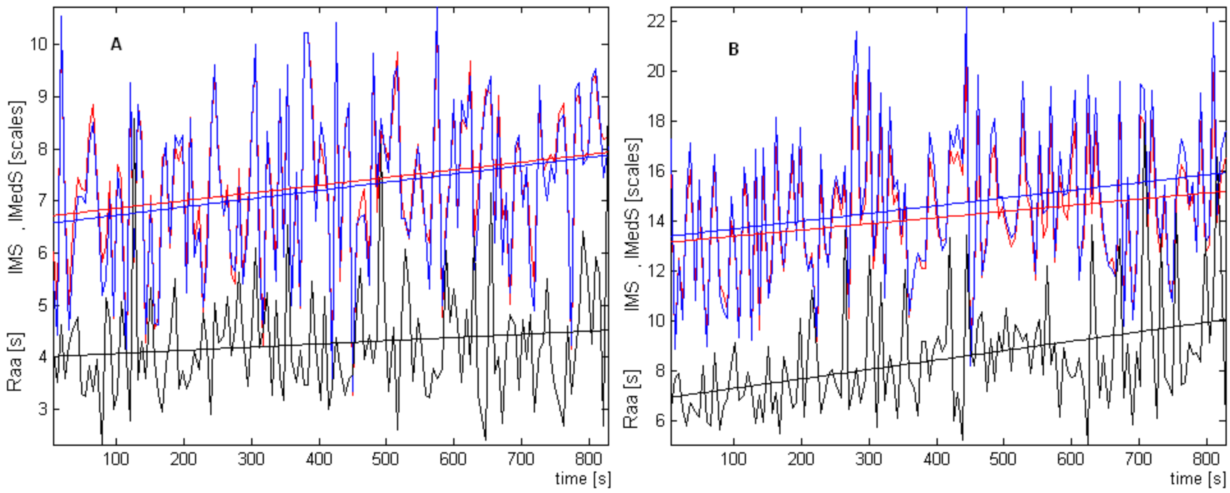


Fig. 5. The evolution of Area/Amplitude Ratio (Raa), Instantaneous Mean Scale (IMS – red), Instantaneous Median Scale (IMedS - blue) with advancing fatigue, for the SEMG (A) and MMG (B), on PTW 100ms.

The two-way analysis of variance (ANOVA), performed separately on the SEMG and MMG, shows (i) no significant differences between the slope of the linear interpolation of the evolution means either among the epochs or between sexes and (ii) no interaction between any epoch and any of the subject sexes - the probability values were greater than 0.05 -, for Raa, IMS, and IMedS as well. Raa approach was 5.52 ±0.97 times quicker than the WT. For the given number of scales considered (30) the memory required for IMS, IMedS is 285±57 times larger than for Raa.

4. Discussion and conclusions

The Raa, IMS and IMedS increase with advancing fatigue may be the effect of the centrally generated progressive alteration of the activation of individual MNs [33], possibly explained by the withdrawal of the tonic fusimotor driven spindle-support via the fusimotor loop [34], mechanism responsible for up to 30 % intervention, as the muscle afferents provide up to 30% excitation to the MNs. Other contributors may be (i) the presynaptic inhibition, which affects the Ia and Ib afferents, (ii) the group II non-spindle muscle afferents which have increased discharge rates with increasing fatigue, further contributing to the inhibitory effect, (iii) group III and IV muscle afferents, activated by the accumulation of metabolites in the fatiguing muscle, thus centrally inhibiting the alpha MNs. Within this context, the critical role of somatosensory feedback from working muscles on the centrally mediated determination of central motor drive and power output has been emphasized [35], so that the development of peripheral muscle fatigue is confined to a certain level, by having shown that attenuated afferent feedback from exercising locomotor muscles results in an overshoot in central motor drive and power output normally chosen by the subject, thereby causing a greater rate of accumulation of muscle metabolites and excessive development of peripheral muscle fatigue.

Together with our findings, other reports concluded that (i) during dynamic muscle actions, the MMG signal provides valid information regarding muscle function, (ii) SEMG and MMG provide complementary information about the electrical and mechanical activity of the muscle [36, 12], (iii) MMG MNF (IMS, IMedS in our case) can vary during fatiguing sustained isometric muscle action, despite a constant torque level, thus providing different information from the torque signal, and decreases in MMG mean or median frequency may reflect reductions in the global motor unit firing rate, rather than decreases in the firing rates of individual motor units [37], (iv) fatigue induced decreases in MMG MNF may reflect reductions in MU firing rates and/or derecruitment of fast twitch MU [38].

All these findings advocate the use of MMG alone or together with SEMG, further underlining the intimate link between the SEMG and MMG as functionally related signals, both witnessing the neural activation of the muscular fibers and its mechanical effect, respectively. While the potential for cross-talk in surface MMG is relatively small even for muscles close to each other and having a common innervation [39] and the MU activation strategy might be estimated in more detail by the MMG than by the SEMG [40], the use of MMG is furthermore encouraged and justified.

It was recognized that SEMG and 'MMG signals recorded during dynamic muscle actions could require different signal processing methodologies when compared to isometric muscle actions' [32]. It was to the purpose of finding other processing techniques to cope with the non-stationarity inherent to the both signals, mainly in the dynamic contraction, that the FFT approach has been progressively replaced by the WT approach.

Characterizing the dynamic contraction is only possible via parameters computed on very short epochs, thus being able to appropriately explore transitional episodes of muscle contraction or relaxation, which naturally alternate within the muscle function during normal work. WT is able to meet such challenges, yet some comments are needed regarding whether to use the Discrete (DWT) or the Continuous Wavelet Transform (CWT) and what mother wavelet to use.

Different authors used the DWT [29, 30] with different mother wavelets: Daubechies db2, db3 [29], db4 [13], db5 [30], db10 [32], Morlet [31], Sym 4, Sym5 [29], as alternatives to FFT, to be able to deal with stronger nonstationarities which were observed in the SEMG and MMG signals from a dynamic contraction.

We chose to use CWT to compute IMS and IMedS because it works at every scale and preserves all the information within the signal [41]. Provided the exact selection of the mother wavelet is not critical in CWT given that each MW gives approximately the same qualitative results [42], our choice was the 'Mexican Hat', chosen from a set of wavelets (coif5, db3, db4, gaus5, mexhat, meyr, morl, rbio3.5) after a selection based on sensitivity, by computing the ratio of variation of IMS and IMedS, over their maximal value. The 'Mexican Hat' mother wavelet gave an average ratio of 15%±4% comparing to 9%±4% for the others, therefore showing a higher sensitivity, while satisfying the admissibility condition and showing excellent localization in time and frequency [43].

Concluding on the use of WT, one of the problems - specific to using this approach - is the complexity of the mathematical instrument, associated with the need to use appropriate criteria to choose a proper mother wavelet, making such methods over-complete (redundant) and therefore inefficient in terms of computational time or storage space [28].

The Raa parameter overcomes such problems by its simple definition [14, 15] and proves a similar behavior in monitoring NMF, with IMS and IMedS computed via WT. Our experiments demonstrate that Raa has a similar behavior with IMS and IMedS (Figure 5), with increasing fatigue, with no significant differences between slopes, neither among the PTWs, nor between sexes , while being quicker than the WT approach and requiring less memory. Therefore it is a convenient alternative to monitor the development of neuromuscular fatigue, mostly in a dynamic contraction, due to the possibility to be computed on short epochs (100 ms).

This opens way to monitoring neuromuscular fatigue in any type of exercise, in difficult environments and activities.

Acknowledgements

This work was performed under research grant 82108/2008/CNMP/Romania and aimed a review, analysis and proper choice of techniques to address monitoring of neuromuscular fatigue of pilots in dynamic effort.

References

[1] De Luca CJ. Control properties of motor units. J. of Experimental Biology, 115:125-136, 1985.

[2] De Luca CJ et al. Control Scheme Governing concurrently Active Human Motor Units During Voluntary Contractions. Journal of Physiology, 329:129-142, 1982.

[3] Tarata MT. Sensorimotor interactions within the context of muscle fatigue. In: "Sensorimotor Control" (Dengler R., Kossev A., eds.), NATO Science Series, Series 1: Life and Behavioural Sciences 2001, 326: 84-91.

[4] Bigland-Ritchie B, Woods JJ. Changes in muscle contractile properties and neural control during human muscular fatigue. Muscle & Nerve, 7: 691-699, 1984.

[5] Stulen FB, De Luca CJ. Frequency parameters of the myoelectric signal as a measure of muscle conduction velocity. IEEE Trans. Biomed. Eng. , BME-28: 515-522, 1981.

[6] Bigland-Ritchie B et al: Central and peripheral fatigue in sustained maximum voluntary contractions of human quadriceps muscle. Clinical Science and Molecular Medicine, 54: 609-614, 1978.

[7] Gandevia SC et al. Central fatigue: critical issues, quantification and practical implications. Adv. Exp. Med. Biol., 384: 281-294, 1995.

[8] Barry DT. Vibrations and sounds from evoked muscle twitches. *Electromyogr. &Clin. Neurophysiol.*, 32: 35-40, 1992.

[9] Iaizzo PA, Pozos RS. Analysis of multiple SEMG and acceleration signals of various record lengths as a means to study pathological and physiological oscillations. *Electromyogr. & Clin. Neurophysiol.*, 32: 359-367, 1992.

[10] Comby B, Chevalier G, Bouchoucha M. A new method for the measurement of tremor at rest. Archives Internationales de Physiologie, de Biochimie et de Biophysique, 100: 73-78, 1992.

[11] Tarata MT. Mechanomyography versus Electromyography, in monitoring the muscular fatigue. Biomed Eng Online, 2:3, 2003.

[12] Madeleine P, Jørgensen LV, Søgaard K, Arendt-Nielsen L, Sjøgaard G. Development of muscle fatigue as assessed by electromyography and mechanomyography during continuous and intermittent low-force contractions: effects of the feedback mode. Eur J Appl Physiol, 87:28–37, 2002.

[13] Beck TW, Housh TJ, Johnson GO, Weir JP, Cramer JT, Coburn JW, Malek MH. Comparison of Fourier and wavelet transform procedures for examining the mechanomyographic and electromyographic frequency domain responses during fatiguing isokinetic muscle actions of the biceps brachii. J Electromyogr Kinesiol, 15(2):190-199, 2005.

[14] Tarata MT. The Average Area/Amplitude Ratio (Raa), A Consistent Parameter in the Quantitative Analysis of the Electromyogram. In Proceedings of the International Conference on Medical Physics & Biomedical Engineering MPBE'94 May 5-7. Nikos Milonas, Nicosia, 1994, 53-57.

[15] Tarata MT. Monitoring the Evolution of the Muscular Fatigue, Via New Parameters developed from the SEMG Signal. In Proceedings of the ECSAP'97 The First European Conference on Signal Analysis and Prediction 1997 June 24-27. ICT Press, Prague 431-434.

[16] Tarata MT, Spaepen A, Puers R. The Accelerometer MMG Measurement Approach in Monitoring the Muscular Fatigue, MEASUREMENT SCIENCE REVIEW, ISSN 1335-8871, http://www.measurement.sk, 1: 47-50, 2001.

[17] Tarata MT. Sensorimotor interactions within the context of muscle fatigue. In "Sensorimotor Control" (Dengler R., Kossev A., eds.), NATO Science Series, Series 1: Lifeand Behavioural Sciences 2001, 326: 84-91.

[18] Gobbo M, Ce` E, Diemont B, Esposito F, Orizio C. Torque and surface mechanomyogram parallel reduction during fatiguing stimulation in human muscles, Eur J Appl Physiol, 97: 9–15, 2006.

[19] Miyamoto N, Oda S. Effect of joint angle on mechanomyographic amplitude during unfused and fused tetani in the human biceps brachii muscle. *Eur J Appl Physiol,* 95(2-3): 221-8, 2005.

[20] Beck TW, Housh TJ, Johnson GO, Weir JP, Cramer JT, Coburn JW, Malek MH. Comparison of a piezoelectric contact sensor and an accelerometer for examining mechanomyographic amplitude and mean power frequency versus torque relationships during isokinetic and isometric muscle actions of the biceps brachii. *J Electromyogr Kinesiol,* 16(4): 324-35, 2006.

[21] Madeleine P, Bajaj P, Søgaard K, Arendt-Nielsen L. Mechanomyography and electromyography force relationships during concentric, isometric and eccentric contractions. *Journal of Electromyography and Kinesiology,* 11: 113–121, 2001.

[22] Esposito F, Ce` E, Gobbo M, Veicsteinas A, Orizio C. Surface EMG and mechanomyogram disclose isokinetic training effects on quadriceps muscle in elderly people. *Eur J Appl Physiol,* 94: 549–557, 2005.

[23] Coburn JW, Housh TJ, Cramer JT, Weir JP, Miller JM, Beck TW, Malek MH, Johnson GO. Mechanomyographic and electromyographic responses of the vastus medialis muscle during isometric and concentric muscle actions. *J Strength Cond Res.* 19(2):412-20, 2005.

[24] Cescon C, Sguazzi E, Merletti R, Farina D. Non-invasive characterization of single motor unit electromyographic and mechanomyographic activities in the biceps brachii muscle. *J Electromyogr Kinesiol,* 16(1):17-24, 2006.

[25] Kawczyński A, Nie H, Jaskólska A, Jaskólski A, Arendt-Nielsen L, Madeleine P. Mechanomyography and electromyography during and after fatiguing shoulder eccentric contractions in males and females. *Scandinavian Journal of Medicine & Science in Sports,* 17(2): 172-179, 2007.

[26] Rabiner RL, Gold B. Theory and Application of Digital Signal Processing. *Prentice-Hall, Englewood Cliffs,* 1975.

[27] Merletti , Lo Conte LR. Advances in processing of surface myoelectric signals Part 1. *Medical & Biological Engineering & Computing.* 5: 361-372, 1995.

[28] Karlsson S, Gerdle B, Mean frequency and signal amplitude of the surface EMG of the quadriceps muscles increase with increasing torque — a study using the continuous wavelet transform. *Journal of Electromyography and Kinesiology*, 11:131–140, 2001.

[29] Kumar DK, Pah ND, Bradley A. Wavelet Analysis of Surface Electromyography to Determine Muscle Fatigue. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(4):400-406, 2003.

[30] Kilby J, Gholam H. Wavelet analysis of surface electromyography signals. Conf Proc IEEE Eng Med Biol Soc. 2004;1:384-7.

[31] Lauer RT, Stackhousea C, Shewokisa PA, Smith BT, Orlina M, McCarthya JJ. Assessment of wavelet analysis of gait in children with typical development and cerebral palsy. *Journal of Biomechanics,* 38:1351–1357, 2005.

[32] Beck TW, Housh TJ, Cramer JT, Weir JP, Johnson GO, Coburn JW, Malek MH, Mielke M. Mechanomyographic amplitude and frequency responses during dynamic muscle actions: a comprehensive review. *Biomed. Eng. Online,* 4(1):67, 2005.

[33] Orizio C, Gobbo M, Diemont B, Esposito F, Veicsteinas A. The surface mechanomyogram as a tool to describe the influence of fatigue on biceps brachii motor unit activation strategy. Historical basis and novel evidence. *Eur. J. Appl. Physiol.*, 90:326–336, 2003.

[34] Hagbarth KE, Macefield VG. The fusimotor system. Its role in fatigue. In Gandevia SC editor. Fatigue. Plenum Press: New York, 1995; 259-70.

[35] Amann M, Proctor LT, Sebranek JJ, Pegelow DF, Dempsey JA. Opioid-mediated muscle afferents inhibit central motor drive and limit peripheral. *J. Physiol., 0:* jphysiol.2008.163303v1, 2008.

[36] Madeleine P, Bajaj P, Søgaard K, Arendt-Nielsen L. Mechanomyography and electromyography force relationships during concentric, isometric and eccentric contractions. *J. Electromyogr.Kinesiol.,* 11:113–21, 2001.

[37] Beck TW, Housh TJ, Johnson GO, Cramer JT, Weir JP, Coburn JW, Malek MH. Does the frequency content of the surface mechanomyographic signal reflect motor unit firing rates? A brief review. J. Electromyogr. Kinesiol., 2007; 17(1):1-13.

[38] Perry-Rana SR, Housh TJ, Johnson GO, Bull AJ, Cramer JT. MMG and EMG responses during 25 maximal, eccentric, isokinetic muscle actions. *Med. Sci. Sports Exerc.,* 35:2048-54, 2003.

[39] Cramer JT, Housh TJ, Weir JP, Johnson GO, Ebersole KT, Perry SR, Bull AJ. Power output, mechanomyographic, and electromyographic responses to maximal, concentric, isokinetic muscle actions in men and women. *J. Strength Cond. Res*., 16:399-408, 2002.

[40] Akataki K, Mita K, Watakabe M. Electromyographic and mechanomyographic estimation of motor unit activation strategy in voluntary force production. *Electromyogr Clin Neurophysiol*., 44(8):489-96, 2004.

[41] Percival DB, Walden AT. Wavelet Methods for Time Series Analysis. Cambridge University Press: New York, 2000; 5-19.

[42] Torrence C, Compo G. A practical guide to wavelet analysis. *Bulletin of the American Meterological Society*, 79 (1): 61–78, 1997.

[43] Debnath L. Wavelet Transforms and Their Applications. Birkhauser: Boston, 2002: 361-378.

Selected Biosensors for Neurotoxicity Testing

S. Ďaďo

Faculty of Electrical Engineering, CTU Prague, Czech Republic
Email: dado@fel.cvut.cz

Abstract. *In vitro neurotoxicity testing and toxicity effect quantification plays an important role in many disciplines of biomedicine as an alternative to in vivo methods. The principle of the majority of in vitro methods corresponds to the basic concept of biosensors. i.e. measured quantity is by means of biological sensing element transformed to physical quantity easily measurable by electrical methods of measurements. Two types of biosensors suitable for neurotoxicity measurements are described in the paper. A common feature for both types is an application of the living cell as biological sensing element. In first type of biosensor the morphology of cell is evaluated using image processing methods known as videometry. In the second type of biosensors the electrical impedance of cells using an improved version of an ECIS (Electric Cell-substrate Impedance Sensing) method is a measure of toxicity effects. The results of experiments with biosensors using videometry and proposal for improvements of ECIS based biosensors are included in the paper.*

Keywords: In vitro neurotoxocity testing, Biosensors, Impedance of cells, ECIS method

1. Introduction

Neurotoxicity is defined as any adverse effect on the chemistry, structure and function of the nervous system during development or at the maturity induced by chemical or physical influences. For both economic and humane considerations there has been growing interest in alternatives to the use of animals in toxicity testing of chemical agents. Tissue culture has the potential to replace animal testing, but for the success of *in vitro* approaches new and sensitive methods to detect cellular activities are required.

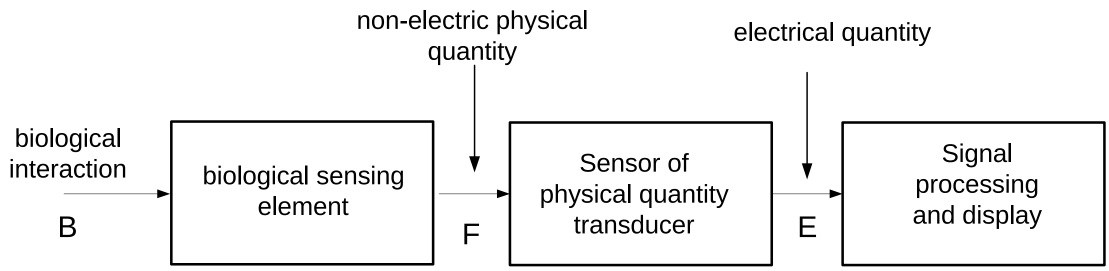


Fig.1 Block diagram of the typical biosensor

For the detection of cellular activities the concept of biosensors can be used. As it is well known *biosensors* (Fig.1) are devices incorporating a biological sensing element coupled to a variety of transducers (sensors) which convert a biological interaction into an easily measurable electrical signal. For the neurotoxity testing by *videometric* methods biological sensing elements are spinal ganglions. Spinal ganglions are clusters of nerve cells – neurons with nerve fibres sprouts called neuritis (Fig. 2). The change of neuritis morphology is very sensitive indicator of toxicity of the tested substance.

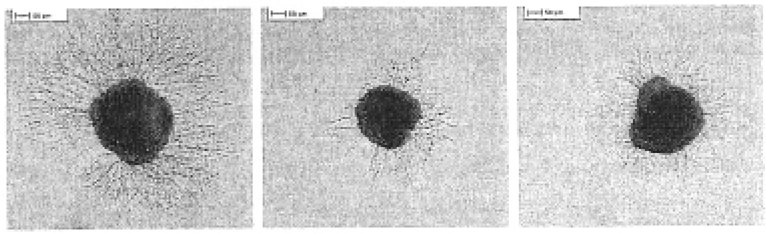


Fig. 2 The morphological changes of ganglion caused by toxic substances. (On the left - reference ganglion, in the middle - ganglion damaged by the effect of *TPPS*$_4$, on the right - the effect of *Photosan*).

Typical biological sensing elements for testing toxicity by *impedance methods* (ECIS) are fibroblastic V79 cells. The measure of toxicity is time response function of impedance when cells are exposed to toxicants.

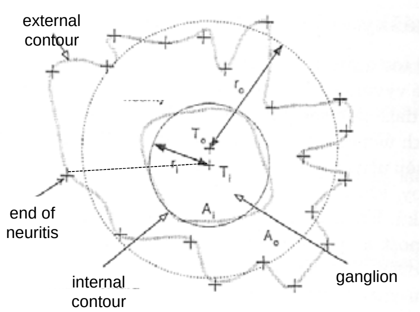


Fig. 3 Internal and external contours

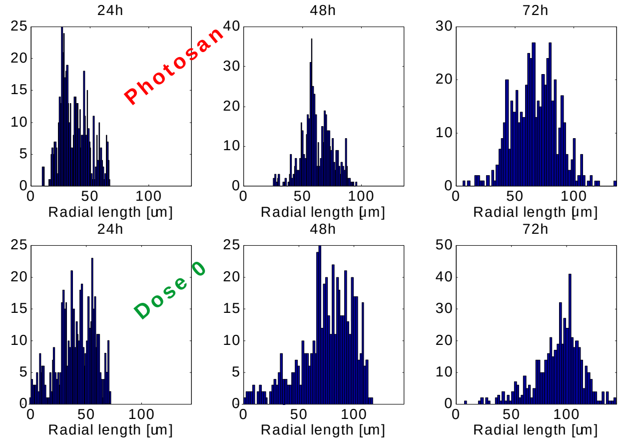


Fig. 4 Histograms of radial length of neuritis. Dose 0-reference, upper row- ganglions affected by Photosan

2. Biosensors based on image processing (videometry)

The degree of toxicity is estimated from morphological changes of neuritis exposed to toxic substance. The goal is to find quantitative parameters of ganglion geometry satisfactorily expressing the degree of toxicity of tested substance. The changes of the radial length of neuritis (Fig. 4) can be used as a measure of toxicity [1]. Another possibility is an analysis of *contours*. In case of ganglion the *external* contour is an outline connecting the endpoints of neuritis and internal contour is a circle approximating ganglion shape (Fig. 3).

Methods based on analysis of contours

Describing the contour by equation in polar coordinates analysis the discrete Fourier transform (DFT) or wavelet transform (DWT) can be used. By interpolation between endpoints we obtain continuous function c(φ) in polar coordinates. Then continuous wavelet transform of function *f(t)* defined by Equ.1 can be used.

$$Wf(\tau,s) = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{s}}\overline{\psi}(\frac{t-\tau}{s})dt \qquad (1)$$

where s is scale (amplitude) parameter and τ determines position of mother wavelet $\psi(t)$

WT with properly chosen mother wavelet allows description of contour by minimum number of coefficients [2].

Descriptor CWT

In this approach the contour is divided to N segments having length *k* and for each segment transformation (1) is applied. The *descriptor CWT* is then defined by relation

$$CDc(\tau,s)=\sum_{k=1}^{N}\int_{k}^{k+1} c(k)\frac{1}{\sqrt{s}}\psi(\frac{t-\tau}{s})dt \tag{2}$$

The maximum values of coefficients are then used for the characterization of contour.

Skeleton

By plotting local maximums of absolute values of descriptors in coordinate τ for *each* particular scale *s* so called *skeleton* (Fig. 5, on the right) is created.

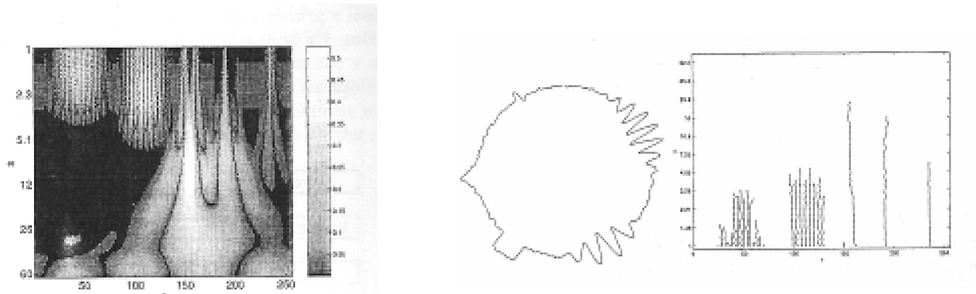


Fig. 5, left: 3-D graph of WT coefficients (wavelet Mexican hat). middle: simulation of contour radial noise. right: skeleton lines – ridges of maxima.

The advantage of skeleton description is simple elimination of noise (stochastic variation of radial length). The noise causes the appearance of skeleton line for small values of s which are then easily separated from long lines – attributes of contour [3].

3. Biosensors based on measurement of cells impedance

In this case the biological sensing element is a monolayer of cells between electrodes located in wells (Fig. 6). The presence of cells increases impedance due to isolation properties of cell membranes. Ideally, the effect of individual cell on the impedance should be observable (spatial resolution). The arrangement of electrodes using “spread resistance” principle (known from semiconductor resistivity measurement or from mercury drop polarography) can fulfill this task. Two (gold) electrodes; miniature active electrode (diameter 250 μm) and large reference one (area approx. 300 times larger) are used. The current density through active electrode is much higher thus only processes in vicinity of it affect the impedance (Fig. 6).

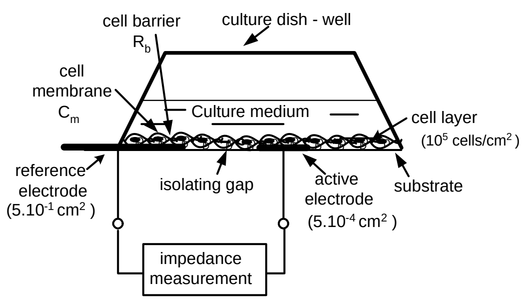


Fig.6 The principle of ECIS [4]
(Patent US 7,399,631 B2)

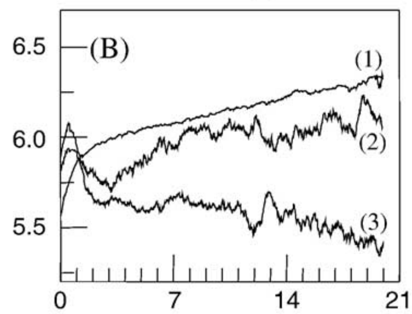


Fig.7 The change of C_m [nF] with time [h] (f = 4 kHz)
(line 1-no cells, line 2,3- cell culture V79)

Impedance is measured with a weak AC signal (1 μA) in frequency range from 10 to 10^5 Hz. When cells attach and spread on and between electrodes, their insulating membranes constrain the current, forcing it to flow beneath and between the cells that must be anchored and spread upon substratum. A limited population of cells (1 to 1,000) is measured at the time. This results in impedance changes that can be readily measured and used to quantify cell behavior.

The perspectives of ECIS method and problems to be solved

Biosensors using ECIS concept is perspective but not yet as widely used as Surface Plasmon Resonance biosensors. Further improvement of the ECIS method should solve following tasks:

1. Asymmetrical geometry of electrodes might lead to the dependence of impedance on polarity and in consequence to the presence of DC component of electrode current (“rectifier effect”). The remedy: new symmetrical arrangement of electrodes.
2. The research of feasibility of replacement expensive instrumentation by modern impedance measuring IC (e.g. impedance converter AD 5933) is in progress. At the same time the parasitic impedances of leads will be eliminated.
3. Information capacity of impedance of cell is huge and can be fully exploited by modern methods of signal processing and then widely applicable (besides toxicity testing also e.g. in cells wounding research, cell motility and spreading, etc.).
4. Finding the quantitative factors for evaluation of toxicity from the measured impedance is a difficult task requiring a lot of experimental work and the consensus of experts.

4. Results and conclusions

Biosensors for toxicity testing using morphological changes of cells are most widespread methods of in vitro toxicity testing. The transduction of cell morphology to image processing (videometry) offers more information, but requires complicated instrumentation and signal processing. The biosensors transforming cell morphology changes to impedance are much simpler and less demanding. They are just on the beginning on the era of their application in toxicology and represent the perspective orientation of in vitro toxicity testing.

But in both approaches finding of “universal” toxicity quantification parameter and method is extremely difficult, mainly due to the problems with time demanding and even dangerous experimental activities necessary for verification of the proposed toxicity criterion. At the present situation the choice of toxicity testing procedure depends to large extent on the type of toxic agent.

Acknowledgements

This work was supported by the research project MSM ČR 6840770012 “Transdisciplinary research in the domain of biomedical engineering”. The author would like to thank Prof. V. Mandys of the 1st Department of Pathology, Charles University, Czech Republic, for valuable information and experimental data.

References

[1] Ďaďo,S., Mandys V, Šmíd R. Biomedical shape measurement for neurotoxicity testing. *Measurement Science Review*, 1 (1): 55-58, 2001.

[2] Ďaďo,S.,Šmíd, R.: Charakterizace tvaru objektu v obraze pomocí spojité vlnkové transformace. In: Nové smery spracovania signálov, Liptovský Mikuláš, 1998

[3] Ďaďo,S.,Šmíd, R.: Wavelet Correlation for Biomedical Shape Evaluation, Measurement 2003, Smolenice 2003, Poster Section

[4] Keese, CR. At al., Cell–Substratum Interactions as a Predictor of Cytotoxicity. *In Vitro &Molecular Toxicology*, vol.11, No. 2, pp. 183-192, 1998.

Experimental Analysis of Gait Parameters Variability Using VideocameraRecords

D. Šimšík, A. Galajdová, J. Majerník, Z. Dolná

Technical University of Košice, Faculty of Mechanical Engineering,
Department of Biomedical Engineering, Automation and Measurement,
Košice, Slovakia
Email: Dusan.Simsik@tuke.sk

Abstract. *The human gait analysis can provide important data for gait patterns classification of individuals as input for identification of persons according to gait parameters. The set of gait parameters that are relevant for identification purposes is not defined yet. Therefore our recent research target is to distinguish similarities and differences, evaluate the variability of gait parameters of the relevant group of people and conclude the usability for identification purposes. The gait parameters of ten subjects were recorded, calculated and evaluated in our Human motion analysis laboratory at the Technical University of Košice. We use SMART system to capture motion data – dynamic parameters describing trajectories of 25 markers. Their positions, velocities and accelerations provide input for further calculation and linear analysis of human gait as joint angles, rotation angles, step frequency, length and with of step, gait cycle phases.*

Keywords: Human Motion Analysis, SMART System, Gait Parameters, Variability, Identification of People

1. Introduction

The main task of identification of persons is to distinguish person according to the physical appearance, social behaviour, role of interaction in the environment (family, society, friends, work) and identify person according to characteristics, knowledge, occupation and biometrical features (physical, biological, genetic). In the criminology, the identification process can be seen from different point of view and utilize several different methods and algorithms. [1, 2]

Some researchers already showed that gait style and its quantitative description are suitable sources for identification even in the cases when the observation of face, ear, and eye is impossible. In the last years there were made several attempts to identify persons using their motion traces. In order to capture human body motion traces the video capturing systems are applied. It is difficult task to use the video records taken in a real situation in public environment, because usually there is only a less quality record made by a simple camera with low data density. However, cameras improve their parameters, and there are already methods that are able from a simple record of human body movement can provide enough information for identification of individuals. Persons committing crime are often using different camouflage, hidden faces, clothes that can cover body, so the standard methods for identification fail. On the other hand there are laboratory methods with promising results but human motion analysis requires a lot of data, several cameras, special markers for isolating of important points on the human body, and other techniques, to provide reliable results. Video record from the public places is usually in captured in a dynamic, non structures space with possible artefacts caused by other persons and environment.

Identification based on gait analysis has several advantages comparing with other biometric data. It is not invasive, can be remotely captured that is not possible for other biometric features.

Identification of a person is a particular case of a general identification process. It can be classified into [2]:

- external identification (recognition of physical, biological characteristics of person),
- internal identification (perception of the psychological, philosophical and social self-identity).

Gait analysis has many specific individual features for movement identification purposes but on the other hand, there are limitations for use of gait in biometric applications as there are many other factors, which influence the movement patterns and gait styles – intentional change of movement performance, its speed/acceleration, overall behaviour, or external factors like various diseases and injuries influencing posture, short-term and long-term physical and psychical overloading of the body, drugs and medical treatment, changes in body dimensions, walking surface, environment conditions and many others. Also the conditions of imaging can have influence on the gait pattern – calibration, light, shadows, background, and colour of clothing. Some methods will be suited for usage in the laboratory conditions, some of them, on contrary, in the common terrain and environment (streets, public places, corridors, financial institutions, etc.).

2. Subject and Methods

Natural variability in human gait patterns is the precondition for differentiations between individuals when we expect to identify the only one individual between numbers of persons inside of an evaluated group. Therefore we started with an analysis of the gait parameters of a selected homogenous group of subjects. Ten subjects were recorded, calculated and evaluated to analyse variability of obtained data for identification purposes of subjects. In our Human motion analysis laboratory at the Technical University of Košice we use SMART system to record 3D data of moving subjects (6 infrared cameras, 50Hz).

Measurements offer values of dynamic parameters and their time characteristics - trajectories of selected points attached on human body according to the marker set for the whole gait cycle. A model consists of 25 reflex markers placed on head, spine, torso, pelvis, upper and lower extremities. The data available immediately after 3D motion reconstruction are markers positions, velocities and accelerations in time. They provide input for further calculation and linear data analysis of human gait as joint angles, rotation angles, step frequency, length and width of step, gait cycle phases. The subjects were walking during trials in two different speeds - standard walking speed and moderately faster speed. We used 3 types of situations – for each speed we captured gait in tight clothes, in indoor clothes and outdoor outfit with coat and shoes.

Literature analysis shows that there are 2 basic approaches in the gait analysis – identification methods oriented on modelling and recognition of human movement and methods focused on the human body silhouette dynamic changes describing by integrated values [3].

3. Results

The identification process is based on the interpersonal and intrapersonal comparison and individual variance determination. We gathered data from 25 markers because after evaluation of the data variability we will be looking for significant gait parameters that could help us to distinguish between subjects with the highest probability.

In order to detect significant differences the simple ANOVA test will be used. The best parameters with significant differences between subjects can be used for identification procedure. Parameters with very low score shall prove poor contribution for identification purposes and we plan to release them from the further research.

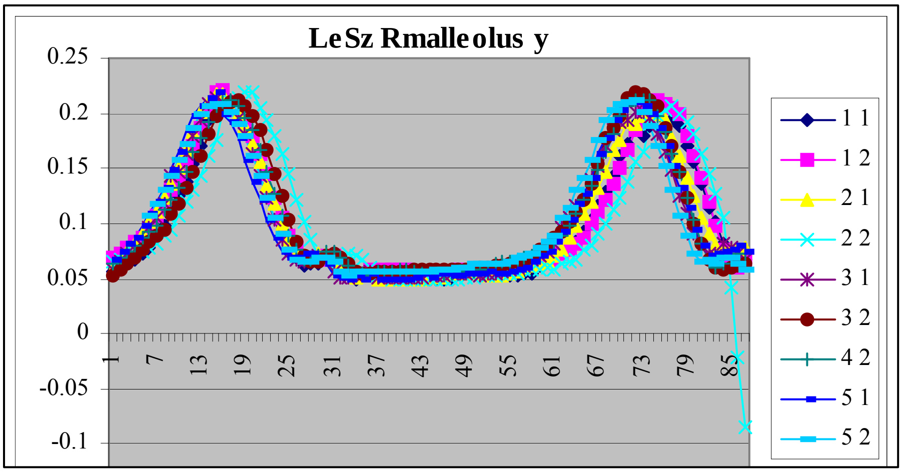
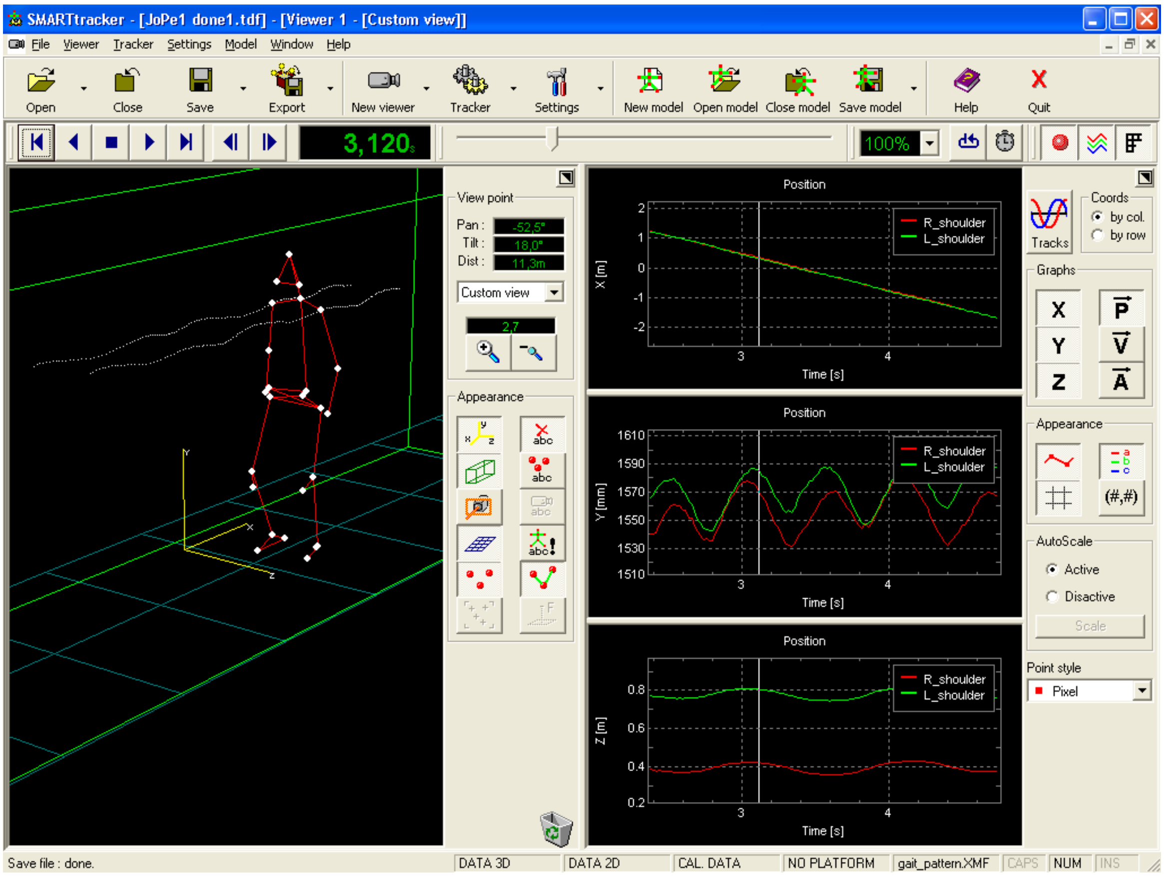


Fig. 1 Variability of the lateral ankle marker position in time

Variability (Fig. 1) of all position data of 25 markers are currently being evaluated.

We use SMART tools for data evaluating in combination with MATLAB and Excell (Fig.2).



4. Discussion

We are currently in the early stage of our research of identification procedures of individuals based on the dynamic stereotype patterns of the human gait.

Variability of gait linear gait parameters is well known fact. We work on proving that fact on the selected group of probants.

The further step will be the factorial analysis of gait parameters looking for a minimal set of significant parameters that are critical for an efficient analysis of an individuals gait record and proving a person identity.

5. Conclusions

Early results of the research obtained ten years ago in different research centres confirmed that gait has a large potential for identification and verification tasks. Only further intensive research in the area of newly developed and applied methods using computational technologies and engineering approach will confirm whether gait will be equally effective, powerful and sufficient biometric alternative to other biometric methods used nowadays.

Gait biometry is associated to high expectations, especially in relation to physical safety and potential use of the already introduced technical resources for safety purposes. There is an option to use already installed industrial cameras, monitoring systems upgraded for automatic evaluation of human identity (real-time systems or retrospectively to identify person from the archive, databases) based on the face and gait recognition. The interest is strongly supported and motivated by security systems due to growing incidences and terror attacks.

Acknowledgements

Presented research was funded by the grant kept on the College of Karlovy Vary, Institute of Criminalistics and Forensic Science, Karlovy Vary, and by the Slovak grant VEGA 1/0820/08 “Dynamic Analysis of Human Motion in Clinical Praxis, Rehabilitation and Sports Medicine”. The obtained results will be used for the Handbook writing supported by KEGA grant 3/6441/08 Rehabilitation engineering – new handbook and research concept of the new study programme.

References

[1] Rak, R., Porada, V., Mesároš, M.: Biometrical Identification. In: Proceeding of Specialized Digital Forensic Forum&Meeting. Digital Forensic Forum Prague 2007. Prague: College of Karlovy Vary. pp.183-214. ISBN 978-80-254-1536-8.

[2] Porada, V.: Kriminalistika. Brno: Ústav soudního inženýrství, VUT, 1993.

[3] Šimšík, D., Porada, V., Rak, R., Majerník, J., Galajdová, A.: Analýza pohybu človeka pri identifikácii osôb v kriminalistike. 1. vyd. Košice : TU, SjF, 2008. 272 s. ISBN 978-80-553-0023-8.

[4] Majerník, J. et al.: Some clinical studies of human movement at the rehabilitation clinic. In: Gait and Posture. vol. 18, no. 2 (2003), p. s94-s95.

Selection of Measures for Sleep Stages Classification

K. Šušmáková, A. Krakovská

Institute of Measurement Science, Slovak Academy of Sciences
Dúbravská cesta 9, 842 19 Bratislava, Slovak Republic
E-mail: kristina.susmakova@savba.sk

Abstract: *In this work a large amount of features from polysomnographic recordings was tested to find the best set of variables for sleep stages classification. Discriminant analysis was done with Fisher quadratic classifier and forward selection procedure. Resulting set contains 14 measures from EEG, EOG, EMG and ECG signals, some of them are used in this context for the first time (e. g. fractal exponent and entropy of EMG).*

Keywords: Sleep stages classification, Forward selection procedure, rules of Rechtschaffen and Kales, nonlinear, spectral measures

1. Introduction

The evaluation of sleep stages is done after broadly appreciated Rechtschaffen-Kales manual [1], which involves parameters, techniques and wave patterns of three physiological signals: electroencephalogram (EEG), electrooculogram (EOG) and electromyogram (EMG) needed for definitive assignation of sleep stages. In clinical practice also other physiological signals, e. g. electrocardiogram (ECG), blood pressure etc., may be measured.

The main states of vigilance are wakefulness (W), REM sleep and non-REM sleep (NREM). NREM sleep is further divided into four stages from the lightest Stage 1 (S1) to the deepest Stage 4 (S4). Stages 3 and 4 are referred to as slow wave sleep (SWS).
During W there is a low voltage (10−30 μV) and mixed frequency EEG, substantial alpha activity in EEG and relatively high tonic EMG.
S1 is characterized by low voltage, mixed frequency EEG with the highest amplitude in 2-7 Hz range. Alpha activity may be present but it must not take more than 50% of an epoch. Vertex sharp waves may occur; their amplitude can reach the value of about 200 μV. In S1 after wakefulness slow eye movements can be present. The EMG level is lower than in the wakefulness.
S2 is characterized by sleep spindles and K-complexes on a relatively low voltage, mixed frequency background activity and the absence of slow waves. Sleep spindles are bursts of brain waves of 12-16 Hz. A K-complex is a sharp negative wave (the amplitude demand is at least 75 μV) followed by a slower positive one.
SWS is detected if more then 20% of the epoch of EEG record contains delta waves which are characterized with 2 Hz or slower and with the amplitudes above 75 μV. Sleep spindles and K-complexes may also be present.
REM sleep shows low voltage and mixed frequency (similarly to S1) of EEG, sawtooth wave pattern is often present. EMG reaches the lowest level and episodic rapid eye movements occur.
The convenience of developing a computerized system for automated analysis and classification of sleep stages has been recognized by several authors [2-4]. A few commercial systems are also available; however they showed substantial differences from the visually scored polysomnographs in the distribution of the sleep stages. Most of these works were concentrated on the choice of the best kind of classifier, however there are only a few works dealing with selection of proper set of discriminative features [5].

2. Subject and Methods

Data

Data with all-night polysomnographic records were kindly provided by Prof. G. Dorffner, received by The Siesta Group Schlafanalyse GmbH. The records were obtained from 20 healthy subjects, 10 men and 10 women. Ages ranged from 23 to 82 years old with an average 50 ± 21.5 years. Subjects slept at their usual sleeping time, typically from 11 pm, the average recording time was 7.5 hours with sleep efficiency 87.1%. Sleep stages were scored by two experts on 30 s non-overlapping segments, if there was an ambiguity, the third independent expert made the decision. All measures were computed on these 30 s long windows, for 1 channel of EMG, 2 channels of EOG, 6 EEG channels (derivations: Fp1-M2, C3-M2, O1-M2, Fp2-M1, C4-M1, O2- M1, where M1, M2 are the left and right mastoids) and 1 channel of ECG. Following numbers of sleep stages were analyzed: 2069 stages of waking, 1452 stages of S1, 7860 stages of S2, 1586 stages of S3, 1865 stages of S4, and 3226 stages of REM sleep.

Computed measures

Following measures were computed for all 10 channels: average frequency, average amplitude, variance, skewness, kurtosis, normality test [6], spectral moments [7], spectral edge [8], spectral exponent [9], spectral entropy [8], fractal dimension [10], exponent of detrended fluctuation analysis [11, 12], entropy, absolute spectral powers [14], relative spectral powers [14], relative power ratios [14]. Coherency [14], phase angle [14] and mutual information were computed for all 15 combinations of EEG channels and between EOG channels.

Powers, coherences, and phase angles were computed for EEG channels in following frequency bands: delta 1: 0.5 - 2 Hz, delta 2: 2 - 4 Hz, theta 1: 4 - 6 Hz, theta 2: 6 - 8 Hz, alpha 1: 8 - 10 Hz, alpha 2: 10 - 12 Hz, sigma 1: 12 - 14 Hz, sigma 2: 14 - 16 Hz, beta: 16 - 30 Hz, and gamma: 30 - 40 Hz and total power: 0.5 Hz - 40 Hz. Power ratios were computed between the relative spectral powers in the main frequency bands: delta/theta, delta/alpha, delta/sigma, delta/beta, delta/gamma, theta/alpha, theta/sigma, theta/beta, theta/gamma, alpha/sigma, alpha/beta, alpha/gamma, sigma/beta, sigma/gamma, and gamma/beta. For EMG and ECG spectral measures were computed only for frequency bands from 10 Hz due to the high-pass filter used on these signals.

Discriminant analysis

Discriminant analysis was done by Fisher quadratic classifier with forward selection procedure (FSP) [12]. In FSP the choice of best features is realized in iterative way - in the first step the best variable in discrimination of sleep stages is selected. At each step a new variable is added to previous set of variables so that new set maximizes (minimizes) some specific criterion. At each step the significance of change of the value of the criterion is tested with t-test. Null hypothesis is that the value of the criterion is not changed with the new set of variables in comparison with the set in previous step. After the p-value of the t-test exceeds the limit 0.05 the null hypothesis is confirmed and FSP is finished. All measures entered in FSP and the criterion to be minimized was the mean error of classification into all sleep stages. Classification procedure was as follows:

1. With the aim to eliminate effect of various numerosity of sleep stages on the classification error from the whole database the same amount of epochs of each sleep stage was randomly chosen; this amount was equal to the less numerous stage (S1).
2. From this representative data set of all sleep stages 90% of randomly determined values created a training set on which the discrimination function was derived.
3. Testing was done on the rest of the data. The total error rate - the percentage of incorrect classified epochs - and the errors of the respective states were computed.

This procedure was repeated 100 times and the mean values and standard deviations of errors were calculated.

3. Results and conclusions

The result of SFP can be seen in Table 1, where selected measures, p – values of significance of change between two steps, mean and standard deviation of classification errors and also mean errors of respective sleep stage are listed. The mean error of classification was diminished from 42.5% by best single-performing variable (delta/beta C3 derivation) to 19.3% by 14 selected measures. On the first three positions, there are EEG, EOG and EMG measures placed and in the last 14th position, the zero-crossing rate of ECG are placed, which indicates the benefit of measuring all polysomnographic signals.

Table 1: Results of forward selection procedure. Selected measures, p-values of hypotheses testing, means and standard deviations of classification errors in [%] and also mean errors of particular sleep stages are listed. Abbreviations: f. fractal; a. absolute, r, relative; c. coherency, zero-cross. zero-crossing rate.

Measure	Channel	p	mean error[%]	Std error[%]	error W[%]	error S1[%]	error S2[%]	error SWS[%]	error REM[%]
delta/beta	C3	0	42.5	1.1	32	80.4	40.3	11.4	48.4
f. exponent	EMG	0	35.3	0.9	29.4	60.5	46.1	11.6	28.5
variance	EOG2	0	29.9	0.9	25.7	47.1	35.1	9.2	32.4
a.sigma1	C3	0	27	0.9	25.1	43.8	30.9	9.6	25.8
r.delta2	O2	0	24.2	0.9	18.6	39.2	29.8	9.4	24.2
theta/gamma	C3	0	23.2	0.9	17.9	39.1	28.6	8.6	21.4
theta/alpha	O1	0	22.4	0.9	16.4	39.9	26.7	8.1	20.8
sigma/gamma	C4	0	21.7	0.8	15.9	39.2	26.5	7.6	19.1
c.delta1	O1-O2	0.002	21.3	0.8	15.2	38.7	26.7	7.9	18.3
entropy	EMG	0.001	20.9	0.8	15.2	39.7	27.2	7.7	15.1
a.theta2	Fp1	0	20.4	0.9	14.7	38.4	26.2	7.7	14.9
theta/alpha	Fp1	0.009	20.1	0.8	13.7	38.3	26	7.7	14.5
c.sigma2	O1-C3	0	19.6	0.8	13.4	38.1	25.3	7.4	13.8
zero-cross.	ECG	0.016	19.3	0.9	13	37.6	24.8	7.5	13.7
alpha/sigma	O1	0.555	19.3	0.8	12.6	36.8	25.5	7.5	13.9

Table 2: Confusion matrix using results from FSP (first 14 measures from Table 1): labels in the first column denote sleep stages determined by experts and in the row there are percentages of sleep stages how this particular stage was determined by classifier.

	W	**S1**	**S2**	**SWS**	**REM**
W	86.36	12.12	0.48	0.02	1.01
S1	10.68	61.20	11.89	0.29	15.95
S2	1.12	10.42	74.66	9.40	4.40
SWS	0.50	0.39	6.14	92.89	0.08
REM	1.81	8.58	3.35	0.11	86.15

Confusion matrix of these 14 selected measures can be found in Table 2. The best determined stage was SWS with 92.89% classification occurrence. Two less predicable stages were S1 with 61.20% of successful classifications and S2 with 74.66%. S1 is a transient state between wakefulness and first “real sleep“ stage S2. From the confusion matrix it can be seen that about 11% is misclassified as wakefulness and about 12% as S2. The high error of S1 detection is not so serious problem with regard to automatic sleep stages classification and its application in sleep medicine because S1 amounts only about 5% of all sleep stages of normal healthy subject. The most numerous stage, S2 (about 50% of all stages), was misclassified with about 10% as neigbouring NREM stages S1 or SWS. The error of S2 could be improved with combination of resulted measures and algorithms of detection of K-complexes and sleep spindles which occur typically during S2 but could be present in smaller amount also during SWS.

Acknowledgements

This work was supported by Slovak Grant Agency for Science (grant No 2/7087/27).

References

[1] Rechtschaffen A. and Kales A. (Eds.): A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subject, US Government Printing Office, National Institute of Health Publication, Washington DC, 1968.

[2] Park, H.-J., Oh, J.-S., Jeong, D.-U. and Park, K. - S.: Automated sleep stage scoring using hybrid rule- and case-based reasoning. *Computers and Biomedical Research*, 33(5):330349, 2000.

[3] Principe, J. C., Gala, S. K. and Chang, T. G.: Sleep staging automaton based on the theory of evidence. *IEEE Transactions on Biomedical Engineering*, 36(5):503–509, 1989.

[4] Anderer, P., Gruber, G., Parapatics, S., Woertz, M., Miazhynskaia, T., Klösch, G., Saletu, B., Zeitlhofer, J., Barbanoj, J., Danker-Hopfe, M., H. Himanen, S.-L., Kemp, B., Penzel, T., Grözinger, M., Kunz, D., Rappelsberger, P., Schlögl, A. and Dorffner, G.: An E-Health Solution for Automatic Sleep Classification according to Rechtschaffen and Kales: Validation Study of the Somnolyzer 24 x 7 Utilizing the Siesta Database, *Neuropsychobiology*, 51(3):115–133, 2005.

[5] Zoubek, L., Charbonnier, S., Lesecq, S., Buguet, A. and Chapotot, F.: Feature selection for sleep/wake stages classification using data driven methods, *Biomedical Signal Processing and Control*, 2(3): 171–179, 2007.

[6] Galbraith, C. G. and Wong E.: Moment analysis of EEG amplitude histograms and spectral analysis: relative classification of several behavioral tasks, *Perceptual and Motor Skills*, 76, 1993, p. 859-866.

[7] Vogel F., Holm S., and Lingjærde O. C.: Spectral moments and time domain representation of photoacoustic signals used for detection of crude oil in produced water, In Proc. Nordic Symp. on Physical Acoustics, Ustaoset, Norway, 2001.

[8] Fell J., Röschke J., Mann K., and Schäffner, C: Discrimination of sleep stages: a comparison between spectral and nonlinear EEG measures, *Electroencephalography and clinical Neurophysiology*, 98, 1996, p. 401-410.

[9] Pereda E., Gamundi A., Rial R., and González, J.: Nonlinear behaviour of human EEG: fractal exponent versus correlation dimension in awake and sleep stages, *Neurosci. Lett.*, 250, 1998, p. 91–94.

[10] Higuchi T.: Approach to an irregular time series on the basis of the fractal theory, *Physica D*, 31, 1988, p. 277 – 283.

[11] Little M., McSharry P., et al: Nonlinear, biophysically-informed speech pathology detection, In Proceedings of ICASSP, Toulouse, France, 2006, EEE Publishers.

[12] Peng C.-K., Havlin S., Stanley H. E., and Goldberger A. L.: Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series, *Chaos*, 5, 1995, p. 82 – 87.

[13] Thatcher W., North R.D., and Biver C.: EEG and intelligence: Relations between EEG coherence, EEG phase delay and power, *Clinical Neurophysiology*, 116, 2005, p. 2129-2141

[14] Meloun M., Militký, J.: Statistická analýza experimentálních dat (Statistical Analysis of Experimental Data, in Czech). ACADEMIA, Praha, 2004.

[15] Larose, D.: Data Mining Methods and Models. Wiley InterScience, 2006.

Two Decades of Search for Chaos in Brain.

A. Krakovská

Inst. of Measurement Science, Slovak Academy of Sciences,
Bratislava, Slovak Republic,
Email: krakovska@savba.sk

Abstract. *A short review of applications of methods of chaos theory to investigation of brain dynamics represented by EEG is given.*

Keywords: EEG, Brain, Chaos, Fractals, Self-organized criticality

1. Introduction to measures of chaos

Advances in computer technology allowed studying the behaviour of nonlinear systems of differential equations, when there are no solutions for the equations. In 1963 the meteorologist Edward Lorenz, using numerical integration of a simple nonlinear model of convection in the atmosphere called attention to the unpredictable complicated behaviour of the system and published the first picture of a strange attractor (attracting subset of non-integer dimension), the so called Lorenz attractor [1].

The three non-linear Lorenz equations belong to the chaotic systems. Chaos is a non-linear deterministic process, which looks random. The distinguishing feature of chaos is sensitive dependence on initial conditions, meaning that nearby states will rapidly evolve toward very different positions. This makes long-term prediction of the course of chaotic orbits impossible, except in the short run. Moreover, change of system parameters may dramatically change the nature of the system behaviour.

In 1980 Packard et al. showed how a real time series could be represented in a multi-dimensional state space [2]. The process of reconstruction is called phase space embedding. Then Takens mathematically established that, if we can measure any single variable of a dynamical system with sufficient accuracy, then it is possible to reconstruct a state portrait, topologically equivalent to the attractor of the original system [3]. Complexity of the reconstructed attractor may provide important information about the system. The most popular tool to assess the complexity is the correlation dimension (D_2), computed by algorithm of Grassberger and Procaccia (GP) published in 1983 [4]. The correlation dimension is an example of measures that can capture the fractal character of the strange attractors. Moreover, the dimension is related to the minimum number of variables needed to model the system behaviour.

Algorithm of Grassberger and Procaccia made it possible to apply the element of chaos theory to various observations, and led in 1985 to the first applications to electroencephalograms (EEG), when Rapp et al. described their results regarding chaotic analysis of neural activity in the motor cortex of a monkey [5].

At the same time Babloyantz and her co-workers reported the observations on the correlation dimension of human sleep EEG [6]. This and some following studies have concluded that the deeper the sleep, the lower the brain dynamics complexity. Dimension has been reported to be the highest in REM sleep, the smallest in slow wave sleep, and significantly higher in the second half of the night than in the first half.

The early years of nonlinear analysis of brain, roughly between 1985 and 1990, were characterized by enthusiastic search for low-dimensional chaos in EEG signals. Around 1990 some of the limitations of algorithms for chaos analysis became clear, and the previous findings of chaos in the brain were critically reexamined.

First, the understanding of non-integer values of the correlation dimension as a sign of deterministic chaos has been questioned. Initially, fractality was considered to be a manifestation of chaos, but in fact we have chaotic attractors that are not fractal and strange attractors that are not chaotic.

Moreover, Theiler has shown that for data sets with long autocorrelation time the application of GP algorithm leads to spuriously low estimates of dimension due to an anomalous shoulder in the graph of correlation integral [7]. To reduce the effects of linear correlations, the author recommended taking much more data than the characteristic autocorrelation time □, and omitting pairs of points closer in time than □. In 1996, after application of the suggested correction to brain signals, a set of EEGs, previously reported to exhibit low dimensions, was reexamined [8]. The scaling regions disappeared and the authors switched to position that in the case of EEG there is no convincing argument for preference low-dimensional representation over modelling by linearly filtered noise.

Another remarkable paper questioning the view that stochastic time series lead to a non-convergence of the correlation dimension is that of Osborne and Provenzale [9]. They showed that noise exhibiting power-law spectra, may result in low values of the correlation dimension. These noises in fact are self-affine signals, generating a fractal curves whose noninteger dimensions are educible from their power spectra. The GP-method cannot distinguish between fractal attractor of deterministic system and fractal random curve if their dimensions equal.

The above findings have also been tested in the case of EEG [10, 11]. The authors have found negative linear correlation between D_2 estimation and the decay of EEG spectra. As a consequence, low values of D_2 computed for EEG must be reinterpreted and the hypothesis of presence of scale-invariant fractal like structures, rather than the suggestion of deterministic dynamics, is to be preferred.

While in some specific cases, e.g., epileptic seizure, the EEG does appear to exhibit low complexity, in general the brain is continually interacting with many other complex systems and EEG seems to be a mixture of noise, certain cyclic processes and possibly some random fractal signals. Each part of such a composition itself has been frequently reported to fool the algorithms used to detect chaotic dynamics. Therefore, the estimates of correlation dimension and other chaotic measures of brain activity should be interpreted with extreme caution.

2. Self-organized criticality

Scaling behaviour (or scale-free behaviour) means that no characteristic scales dominate the dynamics of the underlying process. The long-range correlations build up until they extend throughout the entire system. Then the dynamics of the system exhibit power-law scaling behaviour, and the underlying process works in a so called critical state.

This alternative type of nonlinear process, discovered in 1987 by Bak et al. [12], is known as self-organized criticality (SOC). Unlike chaos, however, this is a probabilistic process. Bak et al. have shown that, under some very general conditions, systems that consist of a large number of interacting non-linear elements self-organize as energy and matter flow through the system and evolve to a critical state with fluctuations at different spatial and temporal scales. Critical states are characterized by power law (fractal) event size distributions and $1/f$ noise.

Noise with a $1/f$ power spectrum is emitted from a huge variety of sources, including electric current passing through a vacuum tube, quasars, sunspots, economic and communication systems, annual amount of rainfall, or rate of traffic flow.

Interestingly, $1/f$ noise and power law scaling have also been found in electroencephalographic recordings. This proposal is consistent with the observation of the background activity pattern - neurons constantly emit pulses. A large cluster of synchronized neurons seems to attract further neurons and causes the oscillation amplitude to increase. Thus, smaller concentration of neurons gives rise to low amplitude signals of higher frequency while large clusters are associated with slow, high amplitude oscillations. As a result, most of the power of the EEG signals is concentrated in the low frequency spectrum. Such a non-trivial long-range correlation within the signal is often present in fractal or scale-invariant processes.

Advocates of the above hypothesis argue that, in a critical state, the brain could be both stable and variable, and may be optimally suited for information processing. Some findings suggest that power law scaling EEG is characteristic of healthy cerebral activity and the breakdown of the scaling may lead to incapability of quick reorganization during processing demands and may be useful in identification of some brain diseases.

3. Discrimination ability of nonlinear measures

In [11] we looked for the presence of exponential or power-law decay in the power spectra of EEG as the next statements are generally accepted: While chaotic behaviour has power spectrum that falls exponentially at high frequencies, stochastic behaviour has power spectrum that decreases as $1/f^{\gamma}$ with increasing frequency. γ came to be called spectral decay, fractal exponent, or power-law exponent. Thus, examination of the power spectrum can help us to answer the question, whether the observed erratic behaviour is essentially deterministic or stochastic. Our result was clear: Power-law model proved to be preferable over exponential model in 99% of frequency ranges both before alpha activity and following alpha activity.

Our next goal was to verify declarations about the relation of correlation dimension to power-law decay. The average value of γ established from the whole EEG spectrum (in our case from 5 to 250 Hz) was about 2.28. As regards correlation dimension, relatively low values of estimates (between 3 and 6) with the mean of 4.35 were found. We found strong negative correlation between the evolutions of the two measures indicating that, in this case they reflect the same information - the dimension estimate by GP-algorithm only mirrors the spectral features of signal.

While it seems that chaotic brain in general sense is no longer an issue, measures used for identifying low dimensional chaotic systems, such as the correlation dimension, continue to be used for studying the EEG signals. There is a range of new EEG measures successful in the monitoring of sleep, anesthesia and seizures and in distinguishing between normal and pathological or otherwise differing states.

Beside correlation dimension, let us mention Lyapunov exponent (reflects the exponential rate of divergence of nearby orbits) and Higuchi's fractal dimension. Higuchi's dimension does not refer to the reconstructed attractor but to the EEG signal itself, which is considered a geometric figure. This dimension yields values between 1 and 2, since a simple curve has dimension 1 and a plane has dimension of 2.

In [13, 14] we tested a lot of traditional and novel scoring measures on the same data to produce a quite systematic overview, missing in the literature, of how varying audio–visual input influences the brain signals and of how single measures allow the sleep stages classification. We confirmed the remarkable efficiency of some novel measures. Our concurrent testing of 73 measures showed that a lot of traditionally used characteristics had poor classification ability as compared with the small number of the best spectral and nonlinear measures. For instance, fractal exponent and the closely related fractal dimension were overcoming the most of traditional spectral measures in discrimination between the individual states of sleep. Presumably the new measures capture the fundamental properties of the underlying system. This indicates potential advantage of nonlinear methods over standard spectral methods in any task associated with classification, modelling or prediction of the discussed systems.

4. Conclusion

The complex system of the brain ranges from neurons to large neuronal networks. It seems to evolve continuously and in real time in conjunction with changes in the surrounding world. Although a convincing demonstration of chaos has only been obtained at the level of neurons, acting as coupled oscillators, some scientists still believe that there could be considerable benefits for the brain to operate in chaotic regimes due to rich range of behaviours. Similar ideas were presented by Skarda and Freeman in 1987 already [15]. They hypothesized that brain self-organizes to generate relevant (possibly chaotic) activity patterns that serve as the essential ground states of the brain activity. Even today, the discussion of this topic is far from finished.

On the other hand, regardless of the presence of chaos in brain activity, it became more and more obvious that the neuroscience should benefit from methods developed for the analysis of nonlinear and chaotic behaviour. So far, nonlinear methods are used mainly in research as they are not yet rooted in everyday clinical practice. But this will certainly change in the near future.

Acknowledgements

This work was supported by Slovak Grant Agency for Science (Grant No. 2/7087/27).

References

[1] Lorenz EN. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences,* 20 (2): 130 - 141, 1963.

[2] Packard NH et al. Geometry from a Time Series. *Phys. Rev. Lett.* 45, 712 - 716, 1980.

[3] Takens F. Detecting Strange Attractors in Turbulence. *Lecture Notes in Mathematics,* 898, 366-381, 1981.

[4] Grassberger P, Procaccia I. Measuring the strangeness of strange attractors. *Phys. Rev. Lett.* 50 (5), 346 - 349, 1983.

[5] Rapp PE et al. Dynamics of spontaneous neural activity in the Simian motor cortex: the dimension of chaotic neurons. *Physics Letters A*, 110 (6), 334-338, 1985.

[6] Babloyantz A, Salazar JM, Nicolis C. Evidence of chaotic dynamics of brain activity during the sleep cycle. *Physics Letters A*, 111 (3), 152 - 156, 1985.

[7] Theiler J. Spurious dimension from correlation algorithms applied to limited time-series data. *Phys. Rev. A,* 34, 2427 – 2432, 1986.

[8] Theiler J, Rapp PE. Re-examination of the evidence for lowdimensional, nonlinear structure in the human electroencephalogram, *Electroencephalogr.Clin. Neurophysiol,* 98, 213 - 222, 1996.

[9] Osborne AR, Provenzale A. Finite correlation dimension for stochastic systems with power-law spectra, *Physica D*, 35, 357–381, 1989.

[10] Pereda E, Gamundi A, Rial R, González J. Nonlinear behaviour of human EEG: fractal exponent versus correlation dimension in awake and sleep stages. *Neuroscience Letters,* 250(2), 91 - 94, 1998.

[11] Krakovská A, Štolc S. Spectral decay vs. correlation dimension of EEG. *Neurocomputing*, 71, 2978 - 2985, 2008.

[12] Bak P, Tang C, Wiesenfeld K. Self-organized criticality: an explanation of 1/*f* noise. *Phys Rev Lett,* 59, 381 – 384, 1987.

[13] Teplan M, Krakovská A, Štolc S. EEG responses to long-term audio–visual stimulation. *International Journal of Psychophysiology,* 59, 81 – 90, 2006.

[14] Šušmáková K, Krakovská A. Discrimination ability of individual measures used in sleep stages classification. *Artificial Intelligence in Medicine,* 44, 261 – 277, 2008.

[15] Skarda C, Freeman WJ. How brain make chaos in order to make sense of the world. *Behav. Brain Sci.,* 10, 161 – 195, 1987.

Advantages of the Area/Amplitude Ratio over Wavelet-derived Parameters, in Monitoring Neuromuscular Fatigue

M. T. Tarata

University of Medicine and Pharmacy of Craiova, Romania
Email: mihaitarata@yahoo.com

Abstract. *The work compares the Raa (Area/Amplitude Ratio) with Instantaneous Mean Scale (IMS) and Instantaneous Median Scale (IMedS) - both computed via Wavelet Transform (WT) - in a fatiguing contraction (Biceps Brachii, 50% MVC, (six females, six males), computed on different epochs (100, 200, 500 ms) from the surface electromyogram (SEMG) and mechanomyogram (MMG) from same motor territory. All three parameters increase from the beginning of the contraction, demonstrating the central component of neuromuscular fatigue. The ANOVA test shows no significant differences (P=0.05) among epochs or between sexes and no interaction between epochs and sexes, which proves Raa as a practical alternative to WT. Raa is 5.52±.97 (MEAN±STDEV) times quicker and requires 285±57 (MEAN±STDEV) times less buffer memory space than the WT. The simple definition and computational efficiency of Raa are major advantages over the WT approach, thus allowing the efficient monitoring of fatigue on very short epochs, also in a dynamic contraction.*

Keywords: Area/Amplitude Ratio, Neuromuscular Fatigue, Electromyography, Mechanomyography, Wavelet.

1. Introduction

A noninvasive method of monitoring the development of NMF is based on the compression of the Power Spectral Density of the SEMG toward lower frequencies, from the beginning of the voluntary contraction [1], due to the reduction of the conduction velocity in direct relation with the muscular fiber membrane excitability and with neural adaptation [2]. As a complementary signal, MMG reflects the mechanical muscle vibrations generated by the spatio-temporal summation of the individual muscle fiber twitches, evoked through motor unit activation by the motor neurons. SEMG and MMG, recorded simultaneously from the same muscles under steady contraction, show a compression of the spectra toward lower frequencies since the beginning of the contraction [3, 4]. In order to study transitional phenomena in muscle contraction and to monitor fatigue in a dynamic contraction, the use of WT has been investigated, via Instantaneous Mean Frequency and Instantaneous Median Frequency [5]. The use of WT was shown on a rather limited scale only for epochs where Fast Fourier Transform can also be consistently used, i.e. windows of signal where no acceleration or deceleration occurs, situation which may just occasionally happen in a steady contraction or in some isokinetic exercise. As an alternative, we explored the use of the Raa parameter (Area/Amplitude Ratio) [6, 7] together with IMS and IMedS, with the purpose to assess its computational efficiency in terms of speed and required memory space.

2. Subjects and Methods

A pair of SEMG electrodes (22.5 x 22.5 mm H59P, MVAP, USA) were placed on the Biceps (six males, six females). The subject sitting on a chair, with a 90^{o} anteflexion between the forearm and the arm had to carry a weight hanging from the wrist via a soft belt. The 100% MVC was estimated, as the maximum weight that could be sustained for two seconds. Tests were performed for 50% MVC up to exhaustion, on different days. An accelerometer (±2g, ICS Sensors, model 3031, USA) and author-made original amplifier (x 50000, 10-250 Hz band pass filter, 250 Hz antialias filter) was used to pick up the MMG. The accelerometer was placed between the SEMG electrodes to pick up the maximal MMG, orthogonal to the muscle from the same motor territory. The SEMG signals were amplified (x 2000, 100 MΩ input impedance, 100 dB CMRR, 250 Hz antialias filter, Beckman R611, USA) and acquired together with the MMG signals via a computerized acquisition system (DAP1200 Microstar Laboratories USA), at 500 Hz sampling rate on all the channels simultaneously.

Parameters

The Raa, IMS and IMedS were computed from the original signals (SEMG, MMG) on successive epochs, for all the subjects, and all epoch widths, using a rectangular window.

Raa – Average Area /Amplitude Ratio, with a dimension of time [ms], is computed from the signal in the time domain, as an average of the Area/Amplitude ratios over the considered epoch, calculated between consecutive transversals of the izoelectric line, called 'phases' [6, 7]*:*

$$Raa = \frac{1}{n}\sum_{i=1}^{n}\frac{S_i}{A_i} \qquad (1)$$

with

n - the number of phases within the current epoch,

S_i - current phase area, the integral of the i^{th} phase of the signal within the current signal segment,

A_i - the maximal amplitude of the i^{th} phase of the signal within the current signal segment, selected on all m samples within the current phase.

IMS and IMedS are computed (MATLAB, Mathworks, USA) via the Continuous Wavelet Transform over 30 scales, using the 'Mexican Hat' mother wavelet, chosen from a set of mother wavelets (coif5, db3, db4, gaus5, mexhat, meyr, morl, rbio3.5), after the sensitivity of each wavelet was estimated over the same set of recordings, by computing the ratio of variation of IMS and IMedS, over their maximal values. The 'Mexican Hat' wavelet gave an average ratio of 15.8± 4.2% comparing to 8.4±3.3% for the others, therefore showing a higher sensitivity, satisfies the admissibility condition and shows excellent localization in time and frequency [8].

Data processing

To provide greater measurement reliability, maximal voluntary intra-subject amplitude normalization was applied to the SEMG and MMG, as a percentage of the maximal value recorded for MVC. Raa, IMS and IMedS were calculated for all the subjects, for each time within the epoch, for all the epochs; their average on each epoch was computed. In each case the processing time and necessary storage were measured, to allow a comparison of the computational efficiency. Statistical processing consisted in applying the two-way analysis of variance (ANOVA) to the data, separately for each specific signal (SEMG or MMG). The existence of any significant difference between the means of the slopes and intercept points respectively, among epochs and between sexes was investigated.

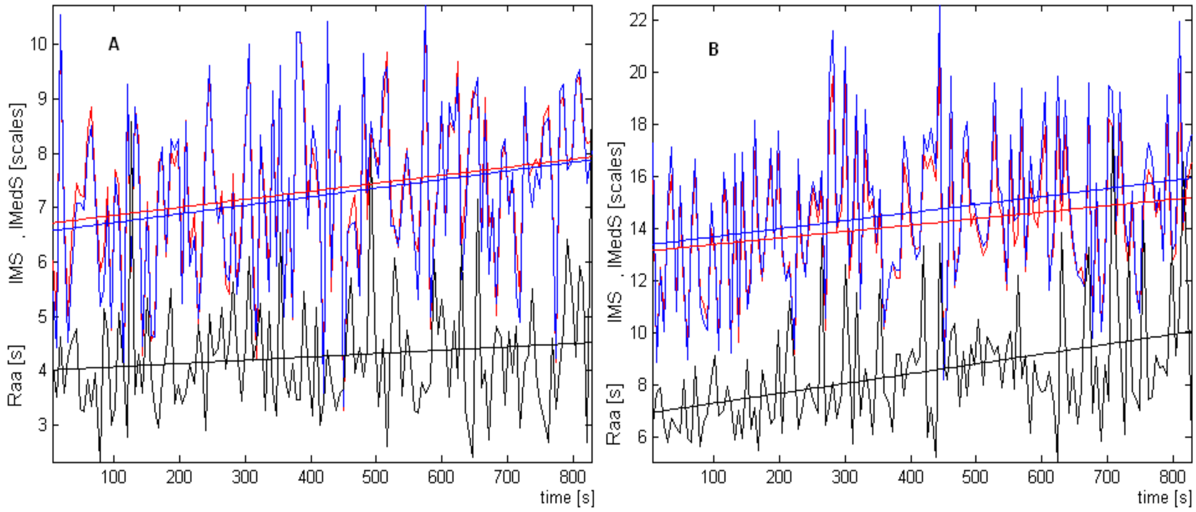


Figure1. The evolution of Area/Amplitude Ratio (Raa), Instantaneous Mean Scale (IMS - red), Instantaneous MedianScale (IMedS - blue) with advancing fatigue, for the SEMG (A) and MMG (B), computed on epochs of 100ms.

3. Results

Raa, IMS and IMedS showed positive slopes (Figure 1) in all subjects (SEMG and MMG), from the beginning of the contraction, for all the epochs. This proved the central component of the fatigue. The two-way analysis of variance (ANOVA) - performed separately on the SEMG and MMG - showed (i) no significant differences between the means either among the epochs or between sexes and (ii) no interaction between any epoch and any of the subject sexes - the probability values were greater than 0.05 -, for Raa, IMS, and IMedS as well. Computationally, the Raa was 5.52±.97 times quicker than the WT. For the given number of scales considered (30) IMS, IMedS – computed from WT – required 285±57 times larger memory than needed for Raa.

4. Discussion and conclusions

As hypothesized, the experiments showed an increase of Raa, IMS and IMedS during the development of neuromuscular fatigue from the beginning of the contraction both for the SEMG and MMG, due to a central intervention in modulating the muscle activation with increasing fatigue.

WT methods are redundant and inefficient in terms of computational time or storage space [9] - due to a complex mathematical instrument - associated with the use of appropriate criteria to choose a proper mother wavelet. The Raa parameter overcame such problems. The work demonstrated that Raa has a similar behavior as IMS and IMedS and requires less memory.

This work performed a functional comparison between scale-oriented parameters computed via the WT and author's original parameter Raa and demonstrated that they have a similar behavior.

Computationally, Raa is 5.52±.97 times quicker than the WT. For the given number of scales considered (30) the memory required for IMS, IMedS is 285±57 times larger than for Raa.

Acknowledgements

This work intended to assess the memory usage and computational efficiency of Raa over the WT approach in monitoring the neuromuscular fatigue. It was performed under research grant 82108/2008/CNMP/Romania and aimed a proper choice of techniques to address monitoring of neuromuscular fatigue of pilots in a dynamic effort. I am thanking the human subjects; without their commitment, these preliminary results could not have been possible.

References

[1] Stulen FB, De Luca CJ. Frequency parameters of the myoelectric signal as a measure of muscle conduction velocity. IEEE Trans Biomed Eng, 28:515-522, 1981.

[2] Gandevia SC, Allen GM, McKenzie DK. Central fatigue: critical issues, quantification and practical implications. Adv Exp Med Biol, 384:281-294, 1995.

[3] Tarata MT. Mechanomyography versus Electromyography, in monitoring the muscular fatigue. Biomed Eng Online, 2:3, 2003.

[4] Madeleine P, Jørgensen LV, Søgaard K, Arendt-Nielsen L, Sjøgaard G. Development of muscle fatigue as assessed by electromyography and mechanomyography during continuous and intermittent low-force contractions: effects of the feedback mode. Eur J Appl Physiol, 87:28–37, 2002.

[5] Beck TW, Housh TJ, Johnson GO, Weir JP, Cramer JT, Coburn JW, Malek MH. Comparison of Fourier and wavelet transform procedures for examining the mechanomyographic and electromyographic frequency domain responses during fatiguing isokinetic muscle actions of the biceps brachii. J Electromyogr Kinesiol, 15(2):190-199, 2005.

[6] Tarata M. The Average Area/Amplitude Ratio (Raa), A Consistent Parameter in the Quantitative Analysis of the Electromyogram. In Proceedings of the International Conference on Medical Physics & Biomedical Engineering MPBE'94 May 5-7. Nikos Milonas, Nicosia, 1994, 53-57.

[7] Tarata M. Monitoring the Evolution of the Muscular Fatigue, Via New Parameters developed from the SEMG Signal. In Proceedings of the ECSAP'97 The First European Conference on Signal Analysis and Prediction 1997 June 24-27. ICT Press, Prague 431-434

[8] Debnath L. Wavelet Transforms and Their Applications. Birkhauser: Boston, 2002, 361-378.

[9] Karlsson S, Gerdle B. Mean frequency and signal amplitude of the surface EMG of the quadriceps muscles increase with increasing torque — a study using the continuous wavelet transform. J Electromyogr Kinesiol 11:131–140, 2001.

Measuring the T Wave of the Electrocardiogram; Why and How

A. van Oosterom

Department Cardiomet, Faculty of Biology and Medicine; University of Lausanne, Lausanne, Switzerland

Email: adriaan.vanoosterom@epfl.ch

Abstract. *The genesis of the T wave of the electrocardiogram is discussed on the basis of a biophysical model, derived from the underlying electrophysiology. This model is used to scrutinize the commonly assumed significance of the individual features of this waveform. As is shown, the last word has not been said about this topic.*

Keywords: T wave morphology; QT level abnormalities, Timing Apex, T wave duration.

1. Introduction

The T wave is one of the prominent features of the electrocardiogram (ECG). An example is presented in Fig. 1, where it is shown next to the other wavelets commonly used for describing the major features of the ECG: P,Q,R,S and U. The nomenclature was introduced by Einthoven in 1900 [1]. The labels PQRST were chosen by Einthoven for no other reason than that a labelling starting with A, B, etc. was already in use at the time for other types of components of signals related to the heart, significantly those of heart sounds. After some initial try-outs, the significance of most of the individual wavelets was soon recognized.

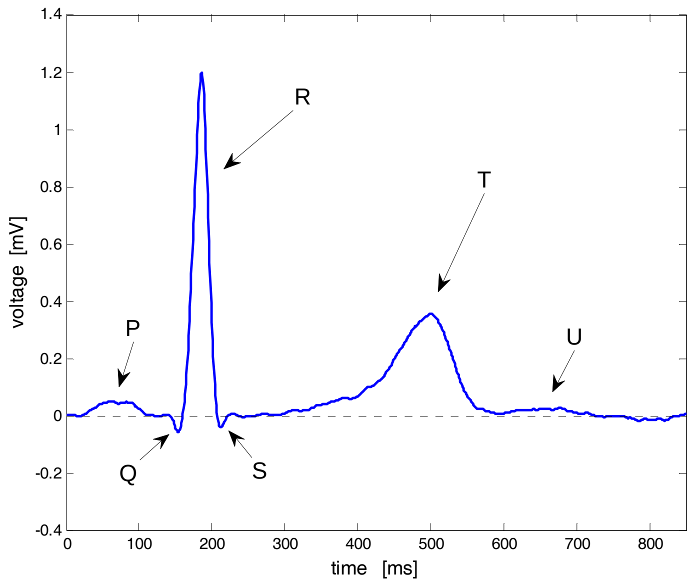


Fig. 1. Nomenclature of wavelets of the ECG as recorded on a healthy subject (lead V4).

The QRS wavelets, elements of the QRS complex during the so-called QRS interval, are attributed to the depolarization phase of the heart’s ventricles, initiating and accompanying their mechanical activity. The subsequent T wave signals are associated with the subsequent return of the transmembrane potentials of the ventricular myocytes to their resting state. The P wave, preceding the QRS complex, is generated by the electric activity of the atria that support the filling of the ventricular cavities prior to the contraction of the ventricular walls.

Curiously, the significance of the U wave is currently, more than a century after Einthoven's work, still being debated [2, 3].

During the twentieth century, the ECG has developed into one of the most important non-invasive tools of electrocardiology [4]. Contributions came from fields such as basic electrophysiology, clinical cardiology, biophysical modelling, improved recording technology and, during its final decades, the field of signal processing supported by computer science.

The widespread application of the ECG has brought about an ever growing-demand on the diagnostic accuracy of the technique. Again, curiously enough, more then a century later, its full potential has not yet been reached, and several basic problems still remain to be solved.

In the diagnostic application of the ECG, two different categories can be distinguished: the analysis of heart rhythm and that of the waveforms observed in the ECG. This invited, didactic paper relates to the latter problem, and in particular so to the significance of features observed in T wave morphology.

2. Measuring the T wave; how?

By now, the technology for measuring bio-electric signals has attained a high quality [5, 6]. The basic principle involved is straightforward. All it amounts to is the placing of two electrodes at different locations on the thorax, which sense the potential difference between these two locations; in the case of the ECG resulting in a signal generated by current sources inside the myocardium.

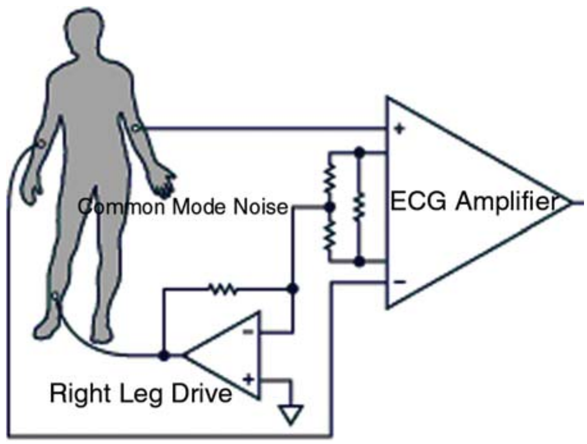


Fig. 2. Basic principle used for recording a single ECG, sensing the potential difference generated by the heart's electric activity between the left and the right arm. The third electrode shown, the one on the right leg serves the very important purpose of minimizing the contribution of electromagnetic fields external to the body on the observed signal [7].

The potential differences on the body surface are transferred to the amplifier by placing electrodes on the skin. At the interface between the skin and the –normally- metal electrodes an exchange of charge takes place between the electrons inside the metal and ions inside the tissues. As a consequence, voltage differences build up at both interfaces, having magnitudes that may be much greater than those of the ECG itself. Moreover, the polarity and magnitude may change slowly over time in response to changes in skin temperature and moisture, as well as pressure exerted. To reduce the effect of such artefacts, the DC type of components in the ECG, components having frequencies below 0.1 Hz, are invariably suppressed by analogue or digital filtering in the first stages of the recording system. A direct consequence of this is that in the signals available for diagnostic purposes, the DC component, being the mean value of the signal over a single beat, is irreparably lost. This is highly relevant, in particular for the interpretation of the ECG during the QT interval, as will be evident later in this paper.

3. Measuring the T wave; but which one?

Measuring bioelectric potentials requires the use of at least two electrodes, since the physics involved indicates that electric potential as such is specified uniquely up to a constant only, and only potential *differences* between different locations are those that matter. In ECG terminology, the signal observed between two terminals of any electrode configuration is referred to as a "lead". The configuration shown in Fig. 2 relates to the recording of a single lead. The third electrode serves to suppress any interferences stemming from electromagnetic fields in the environment, in particular when using a mains-powered amplifier. This third electrode does not, or at least should not, influence the waveform of the recorded signal. Such electrodes are irrelevant for the subsequent discussion.

Multilead signals

In most applications the ECG is recorded by placing more than two electrodes on the thorax. The most frequently used electrode montage is depicted in Fig. 3. It is the configuration of the so-called standard 12-lead system. As can be seen, this lead system involves nine electrodes only. The signals observed over the time interval of a single beat are shown in Fig. 4. The columns 2, 3 and 4 represent the signals observed at the electrode locations shown in Fig. 3 acting as one of the terminals of the lead signals, the other terminal being the mean of the potentials present at electrodes VR, VL and VL, the so-called Wilson Central Terminal. As a consequence, the sum of the potentials of the corresponding leads, VR, VL and VF, is zero and at most eight independent signal components are contained in signals stemming from the standard 12-lead system.

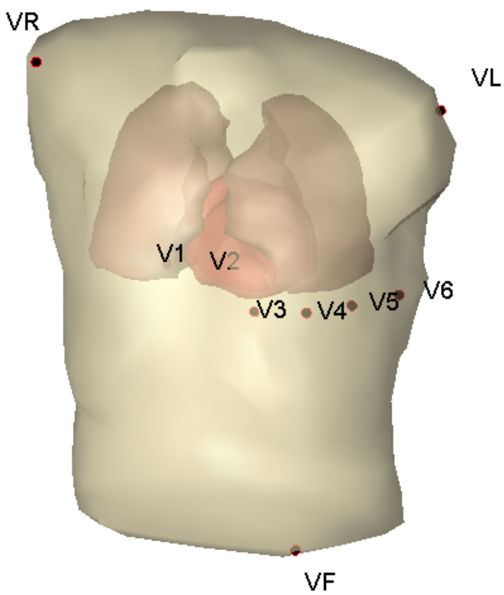


Fig. 3. Electrode montage of the standard 12-lead system. Note that this only involves 9 electrodes. The thorax is that of a 54-year-old male, a numerical model based on his magnetic resonance (MR) images. It includes ventricles, atria and lungs, as used in the computation of the volume conductor effects on the observed ECG waveforms. The lead signals for a single beat, observed at these locations are shown in Fig. 4.

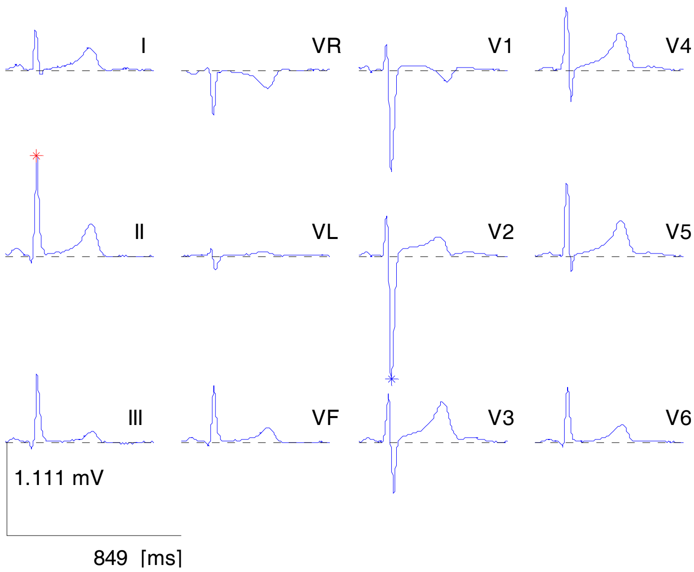


Fig. 4. Waveforms recorded by means of the standard 12-lead system during one complete heart beat. Electrode positions as shown in Fig. 3.

The signals in the first column of Fig. 4 are the three linear combinations that exist on the basis of the potentials at VR, VL and VF, e.g. II=VL-VF; as such these so-called bipolar leads (a misnomer) do not carry extra information. The redundancy involved in the 12 lead signals has proved to be valuable for diagnostic applications of the ECG on the basis of visual inspection of its waveforms. Various other electrode montages are used in specific applications, involving either fewer or more electrode locations for sampling the spatial aspect of the electric field on the thorax [8-10].

Signal features

Even for the ECGs of normal subjects, the ECG waveforms of individual leads exhibit clear differences. In a hybrid type of convention, the nomenclature of the Q, R and S wavelets of the QRS complex are strictly coupled to their polarity; early negativity: negative S, subsequent positivity: R and subsequent negativity: S. In contrast, the polarity of the other wavelets: P, T and U may be either positive or negative and biphasic variants can also be observed. In the example shown (Fig. 4), it can be seen that in an individual lead, not all of these wavelets may be present.

Waveform features used in diagnostic applications of the ECG are timing, magnitude and duration of the individual wavelets. For each of these, in all of the leads of the standard 12-lead system, extensive sets of normal values have been collected [11]. Unfortunately, all these normal values were found to have a large standard deviation, which hampers the classification of distinct types of cardiac disease on the basis of such features. Even when using multivariate types of analysis and applying stratifications on the basis of gender, age or other constitutional variates, there remains a clear need to improve the diagnostic performance. In particular, for the T wave, no indication is given about the individual lead in which the T wave should preferably be analyzed.

4. Measuring the T wave; why?

The statistical analysis of large ECG data bases, in combination with clinical observations, has revealed that the T wave, and in particular the signals during the interval leading up to its apex, the ST segment, contains useful diagnostic information, in particular with respect to episodes of acute ischemia that might lead to myocardial infarction (MI). In extreme cases, when effecting particular regions of the ventricles, the departures from the normal range are indicative. Other interest in T wave morphology stems from the more recently identified hazards arising from repolarization abnormalities, expressed in the more extreme departures of the waveform normality in the variants of the Long-QT syndrome [12], the Brugada syndrome [13] and the Short-QT syndrome [14].

In all of the above mentioned cases, strict diagnostic criteria that can be applied to the ECG are either lacking or of insufficient quality. Improvement will probably be slow if derived from statistical analysis only: the incidence of these syndromes is, fortunately, low. Another approach is to try to understand the basic mechanisms involved, research that is currently being carried out in many research centres. This concerns in particular work at the cellular and membrane level, at which an impressive collection of distinct, genetically related defects have been identified. On a larger scale, the link between cellular activity during repolarization and the resulting ECG waveforms, much less activity takes place. Widely differing opinions about the optimal features of the T wave to be used, and their significance in terms of the underlying electrophysiology can be seen in the literature. It is here that the biophysical modelling of the genesis of the T wave comes into play, an approach that may lead to the improved use of T wave features and to understand their limitations.

One may question if the heading of this section is in accord with its contents so far. In fact, without always being aware of the underlying model-based assumptions, invariably all measurements taken, and in particular the interpretations of their numerical values, are linked to a model [15].

5. Modeling the T wave; how?

Theory

The electrocardiogram arises from the electric activity at the membranes of the cardio-myocytes. In the following, the myocardium is modelled as an ensemble of myocytes, with their intracellular domains coupled by sets of intercalated discs. On a macroscopic scale, the membrane processes at a location $\vec{r}$ at time instant t feed an electric current into the extracellular domain, with source surface density $\vec{J}^{\mathrm{i}}(\vec{r},t)$ (with unit A/m^2). As implied in the work of Wilson[16], and later derived explicitly by Plonsey [17], the impressed density is related to the local transmembrane potential by

$$\vec{J}_{\mathrm{m}}(\vec{r},t) = \nabla(\sigma_{\mathrm{i}} V_{\mathrm{m}}(\vec{r},t)) \ , \tag{1}$$

in which ∇ denotes the gradient operator (unit:1/m), σ_{i} the electric conductivity of the intracellular domain (unit: S/m), and Vm is the local transmembrane potential (unit: V).

In the passive extracellular domain, based on the conservation of charge, the impressed current density $\vec{J}^{\mathrm{i}}(\vec{r},t)$ sets up an electric field $\vec{E}(\vec{r},t)$ (unit: V/m) (a vector field) that should satisfy the differential equation

$$\nabla \bullet \sigma_{\mathrm{e}} \vec{E}(\vec{r},t) = -\nabla \bullet \vec{J}_{\mathrm{m}}(\vec{r},t) , \tag{2}$$

in which $\nabla\bullet$ denotes the divergence operator, (unit:1/m), and σ_e the electric conductivity of the external medium. As one of the basic results of potential theory, we find that, in an external medium of infinite extent, the electric field $\vec{E}(\vec{r},t)$ is the negative gradient of an “auxillary”, scalar function $\Phi(\vec{r}\,',t)$, called the electric potential at field point $\vec{r}\,'$. The potential is found by integration over all of space, justified by the superposition principle that applies to a linear medium:

$$\Phi(\vec{r}\,',t) = -\frac{1}{4\pi\sigma_\mathrm{e}}\int \frac{\nabla\bullet\vec{J}_\mathrm{m}(\vec{r},t)}{R}\,\mathrm{d}V\,, \tag{3}$$

in which R denotes the length of the vector $\vec{R} = \vec{r}\,' - \vec{r}$ pointing from source location $\vec{r}$ to observation point (field location) $\vec{r}\,'$.

Next, by combining (1) and (3) we have

$$\Phi(\vec{r}\,',t) = -\frac{1}{4\pi\sigma_\mathrm{e}}\int_\mathrm{V} \frac{\nabla\bullet\nabla(\sigma_\mathrm{i}V_\mathrm{m}(\vec{r},t))}{R}\,\mathrm{d}V\,. \tag{4}$$

In the final step we recall that we are dealing with cardiac sources located within a volume V (the myocardium), bounded by a closed surface S (endocardium+ epicardium) and apply Green’s second theorem from potential theory [18] (dating from 1828) to the integral in (4), for field points outside the myocardium. Moreover, we take σ_i to be a uniform scalar function within the source region. This permits σ_i to be moved to the front of the integral. Green’s second theorem applied to the scalar fields V_m and $1/R$ reads:

$$\int_\mathrm{V}\frac{1}{R}\nabla^2 V_\mathrm{m}\mathrm{d}V - \int_\mathrm{V} V_\mathrm{m}\,\nabla^2\frac{1}{R}\,\mathrm{d}V = \int_\mathrm{S}\frac{1}{R}\nabla V_\mathrm{m}\bullet\mathrm{d}\vec{S} - \int_\mathrm{S} V_\mathrm{m}\,\nabla\frac{1}{R}\bullet\mathrm{d}\vec{S}\,. \tag{5}$$

The integral to the left of the equality sign is zero, since $\nabla^2\frac{1}{R} = 0$ for field points outside S [19]. The integral to the right of the equality sign is clearly zero since we take *S*, while close to the myocardium, keeping free of it. In the remaining integral on the right, the term

$\nabla^2\frac{1}{R}\bullet\mathrm{d}\vec{S}$ is, $\mathrm{d}\,\omega = \mathrm{d}\omega(\vec{r}\,',\vec{r})$, a scalar function signifying the solid angle subtended by an infinitesimally small element of S as viewed from the observation point. After combining these results with (4), we have

$$\Phi(\vec{r}\,',t) = -\frac{\sigma_\mathrm{i}}{4\pi\sigma_\mathrm{e}}\int_S V_\mathrm{m}(\vec{r},t)\,\mathrm{d}\,\omega\,. \tag{6}$$

This expression shows that, under the assumed condition of a homogeneous intracellular conductivity, the potential in the region exterior to the active sources, i.e. outside the myocardium, depends exclusively on the transmembrane potentials at the surface bounding it.

The term $\sigma_{\mathrm{i}} V_{\mathrm{m}}$ expresses the strength of an equivalent current dipole layer surface density (unit: A m/m^2). Multiplied by the surface element $\mathrm{d}\vec{S}$, a vector directed along the outward normal of S, it takes on the nature of a current dipole (unit: A m). It represents a double layer of electric sources and sinks, and is, accordingly, referred to as an equivalent double layer (EDL), the instantaneous strength of which is proportional to V_{m}.

This highly significant result was first described by Geselowitz in 1989 [20]. As shown, it follows directly from the classic theory of potential theory. In Chapter 7 of the second edition of Comprehensive Electrocardiology [21], it is shown that it forms a direct extension of the classic solid angle theory, originally used for modelling the depolarization phase of the myocytes (QRS complex), to that of the subsequent repolarization phase. In addition, it discusses various implications of the situations in which the assumed uniformity and isotropy of the conductivity of the myocardium can not be taken to be realistic (see also [22]) as well as its relationship with the more general bi-domain theory.

Numerical Implementation

The use of (6) for simulating body surface potentials requires the specification of the surface S and of the distribution of the transmembrane voltage $V_{\mathrm{m}}(\vec{r},t)$. A numerical variant of (6) reads

$$\mathbf{\Phi} = \mathbf{\Omega V}_{\mathrm{m}}\,, \tag{7}$$

expressing the multiplication of a matrix Vm of transmembrane potentials at T discrete time instants at a dense grid of N nodes on the ventricular surface S (matrix size N×T) by matrix $\mathbf{\Omega}$ (matrix size K×N), the elements of which are (scaled) variants of the solid angles subtended by the sources around the N nodes and K observation points in the infinite homogeneous medium. This yields matrix $\mathbf{\Phi}$ (matrix size K×T) : the potentials at the corresponding time instants.

In the period preceding the onset of ventricular electric activity the observed potential (differences) at the thorax surface is minimal, and generally assumed to be zero. During this period the transmembrane potential of all myocytes are taken to be at their resting potential level (about -80 mV). Because of this, matrix $\mathbf{\Omega}$ must have a unit eigenvalue, with corresponding unit vector e:

$$\mathbf{\Omega e} = \mathbf{o}\,. \tag{8}$$

This agrees with the well-known property of a closed surface: it subtends a zero solid angle at an exterior observation point. This, in turn, corresponds to the well-known fact that a completely uniformly polarized cell does not generate an external electric field [16].

In practical implementations of this source description, the cardiac sources are obviously not situated in an infinite medium, but rather in a bounded medium: no electric current passes normally through the skin into the non-conductive external medium: air. The effect of this major inhomogeneity can be accounted for by application of the boundary element method (BEM) [23]. This method generates a linear transfer function between the potentials at all interfaces while assuming the medium to be infinite and homogeneous, and the potentials in an inhomogeneous, bounded state at the same locations [24]. In this way, major inhomogeneities, like the one at the thorax surface and those resulting from the low electric conductivity of lung tissue and the high conductivity of blood inside the ventricles, can be taken into account. These are the compartments of the numerical thorax model illustrated in

Fig. 3. Its numerical expression is that of a matrix transfer matrix B (matrix size L×K), with L the number of electrode locations on the thorax.

The final, general expression for finding the simulated potentials $\mathbf{V}_b$ in the bounded, inhomogeneous medium then reads [25]

$$\mathbf{V}_b = \mathbf{B}\boldsymbol{\Omega}\mathbf{V}_m = \mathbf{A}\mathbf{V}_m , \tag{9}$$

with matrix **A** expressing the general transfer between the transmembrane potentials on S and the potentials on the body surface. Note that $\mathbf{Ae} = \mathbf{o}$, a property “inherited” from (8).

The effectiveness of the EDL source model has been demonstrated in several publications, some related to the forward problem as such: the simulation of the potentials based on assumed EDL source parameters, others to inverse application: the computation of its parameters on the basis of observed body surface potentials [26-29]. This source model is the one used in ECGSIM [30], an interactive computer program for simulating QRST waveforms, freely downloadable from www.ecgsim.org.

EDL-based modeling of the T wave.

We now look at some examples of T waves, modelled on the basis of the EDL source model. As indicated above this demands the specification of individual transmembrane potentials V_m at all of a set of dense nodes on the ventricular surface S. In the examples shown these in turn are modelled as products of a number of logistic functions (hyperbolic tangents), functions of the type,

$$\frac{1}{1+e^{-\beta(t-\tau)}} . \tag{10}$$

These sigmoid functions having a maximum slope $\beta/4$ at $t=\tau$ [31]. In the examples shown in Fig. 5, three such functions are involved, one for modelling the upstroke of the transmembrane potential V_m, the other two for modelling its down slope: an initial slow phase,

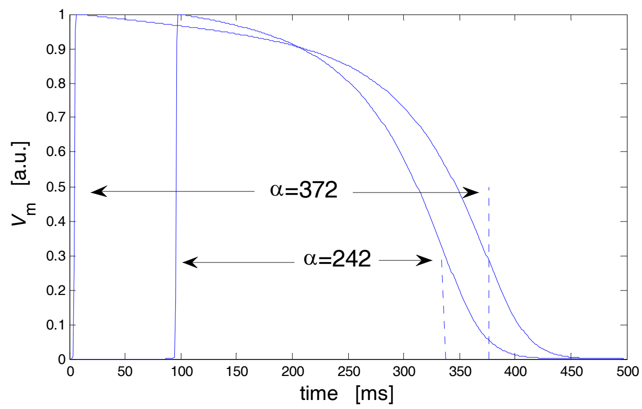


Fig. 5. Examples of two waveforms modelling the transmembrane potential V_m as a product of three logistic functions. The intervals between the arrows denote the activation recovery intervals.

phase 2 of the transmembrane potential, followed by its much faster phase 3. For each individual V_m, just two parameters are used, $\tau=\delta$ setting the timing of local depolarization, and $\tau=\rho$ specifying the timing of the inflection point of the downslope, taken as a marker for local repolarization. The shape parameters β are taken to be identical for all. The parameter shown in Fig. 5 is the activation recovery interval: $\alpha = \rho - \delta$, an indicator of the local absolute refractory period of the local myocytes [32]. The transmembrane potential is specified in arbitrary units. Its zero level is irrelevant because of (8), its magnitude is incorporated in the transfer matrix **A**.

The normal T wave

The ECG waveforms of a healthy subject were simulated on the the basis of the EDL source model. The values of local depolarization δ_n and repolarization ρ_n at each of 317 nodes on S were derived from a dedicated inverse procedure, involving the thorax model shown in Fig. 3 and ECG signals of a 64-lead system involving 65 electrodes "strategically" placed on the thorax. This procedure is similar to the one used for activation imaging published in [33], now extended in order to include the repolarization phase. The δ_n and ρ_n values, mapped on the ventricular surface S are shown in Fig. 6. The waveforms of the transmembrane potential V_m were of the type specified in the previous sub-section. The validation of these kinds of results is impossible. The only way to estimate its realism is by comparing the sequences with what is known from invasive electrophysiology. For the depolarization sequence, the left part of Fig. 6, the outcome is in full qualitative agreement with the unique data set published by Durrer and colleagues in Amsterdam [34]. For the repolarization sequence, shown in the right panel of Fig. 6, no such data is available, which is in fact the reason for the ongoing discussion on this topic mentioned in the introduction. However, the elements of the results were in agreement with the sparse available invasive studies [35, 36]. Moreover, the results are in full qualitative agreement with those observed, by means of essentially the same method, when applied to several healthy individuals. The ECGs simulated on the basis of the activation and recovery sequences depicted in Fig. 6 are represented in Fig. 7, superimposed on the corresponding measured data.

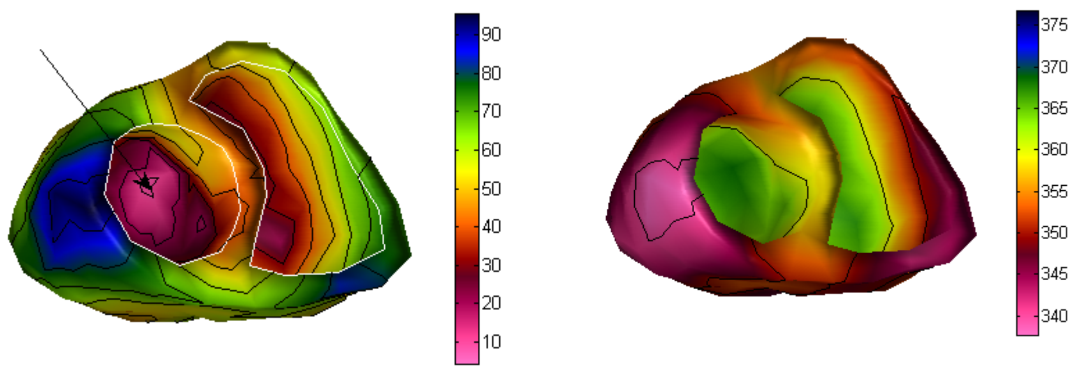


Fig. 6. Posterior view of the ventricles in their natural orientation. Left panel: Timing of depolarization δ; right panel: timing of repolarization ρ. Isochrones are drawn at 10 ms intervals. White contours show the transition between the endocardial and epicardial aspect of the ventricles. Note the small range of repolarization values relative to those of depolarization and its overall shift in time. The sequence of repolarization is largely, but never entirely, the reverse of that of depolarization.

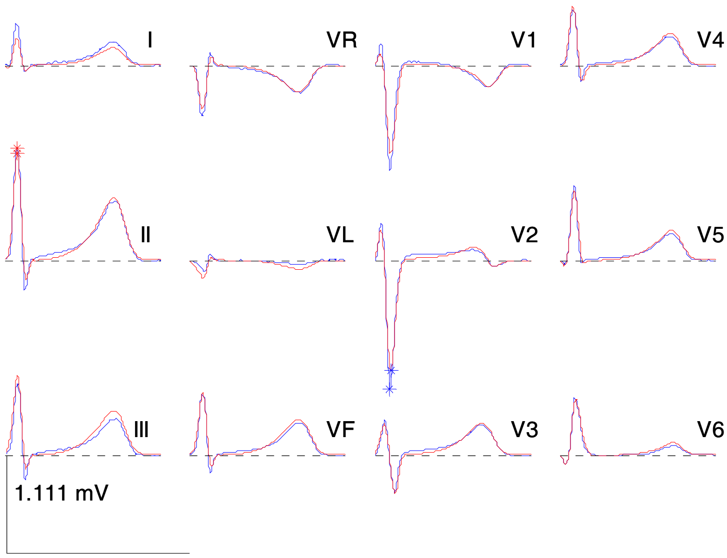


Fig. 7. Superposition of the measured standard 12-lead signals and those simulated on the basis of the EDL source model.

The EDL source model is now discussed in two applications: finding the significance of major T wave morphology and trying to understand ST segment changes during periods of acute ischemia.

T wave morphology

The single parameter model for the timing of repolarization, ρ permits a study of its effect on T wave morphology. Throughout, it should be stressed that while it is a single parameter, in the EDL model it needs to be specified for all of the *N* small elements of the surface carrying the source, ρ_n. Even if the distribution of ρ_n, like the one shown in the right panel of Fig. 6, does not represent the full reality, its subsequent, global variation provides an interesting view on the involved mechanisms.

The non-uniformity of the timing of repolarization is commonly referred to as its dispersion, both for the entire myocardium as well as for transmural (endo-epicardial) differences. Based on an available set of ρ_n values it can be quantified more precisely. Its full significance can only be appreciated by mapping it on the ventricular surface. A more crude documentation of dispersion is to state its basic parameters of descriptive statistics, such as the mean value $\bar{\rho}$, standard deviation $\tilde{\rho}$ and range $\vec{\rho}$. For the distribution of ρ shown in the left panel of Fig. 6, these values were: $\bar{\rho}$ = 355 ms; $\tilde{\rho}$ =7.82 ms; $\vec{\rho}$ =338—377 ms (width: 39ms).

The most commonly used T wave features are the extreme absolute value of the T wave (its apex_T value), and markers taken to be signalling the completion of ventricular repolarization The timing markers commonly used are illustrated Fig. 8, on the recorded signals of leads V1 and V4, also shown in Fig. 4. The marker *a* is the timing of the apex, *b* the timing of the inflection point of the segment following apex_T and *c* is the timing of the intersection of the tangent at time *t*=*b* with the baseline.

The statistics of the timing markers *a*, *b*, and *c* are listed in Table 1. All these markers, are taken as indicating the end of the repolarization phase are frequently referred to as “the end”

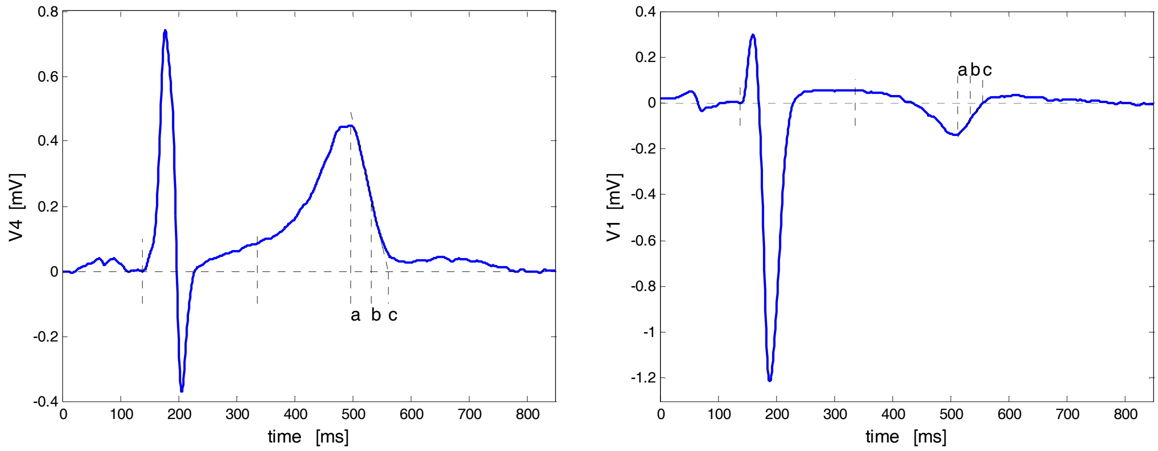


Fig. 8. Markers of timing indicated on measured leads V1 (left panel) and V4 (right panel). From left to right: onset of QRS, mid QT interval; a: the timing of the apex, b: that of the subsequent inflection point, c: the intersection of the tangent at t=b and the baseline. Linear baseline correction was applied from data at the onset QRS to the end of the signal segment show: the beginning of the subsequent P wave (not shown).

of the T wave. The values observed on the individual signals of the 64-lead system were strongly correlated, the highest value of the linear correlation coefficient being 0.96 (*b* and *c*), the lowest 0.88 (*a* and *c*). This suggests that, apart from a shift in their mean value in a direction that is obvious from their definition (Fig. 8), these markers *b* and *c* provide the same estimate of end-repolarization as is observable in individual signals.

In clinical data, the signal quality and that of the subsequently applied signal processing (baseline correction) is not always sufficient to make a robust estimation of markers b and c. Frequently, their values in individual leads are ignored in reports about their potential diagnostic value. Moreover, the presence of the U wave (V1, Fig. 8) tends to be ignored, or severely masked by taking the end of the baseline correction to be at an early stage of the U wave. The end of the T wave, estimated by either of the two markers, clearly does not signal the end of repolarization, whatever their effectiveness may have seemed to be from statistical studies. Moreover, normal values like the length of the QT interval need to be specified in a lead-specific manner.

Table 1. Statistics of the markers of end-repolarization a, b, and c as observed on different leads; for the 64-lead system showing: minimum, mean and maximum values. Time t=0 is taken at the onset of the QRS complex.

lead	*a* (ms)	*b* (ms)	*c* (ms)
V1	374	396	416
V4	359	394	423
64-leads	316 356 385	351 391 420	367 417 467
STD(*t*)	358	393	419

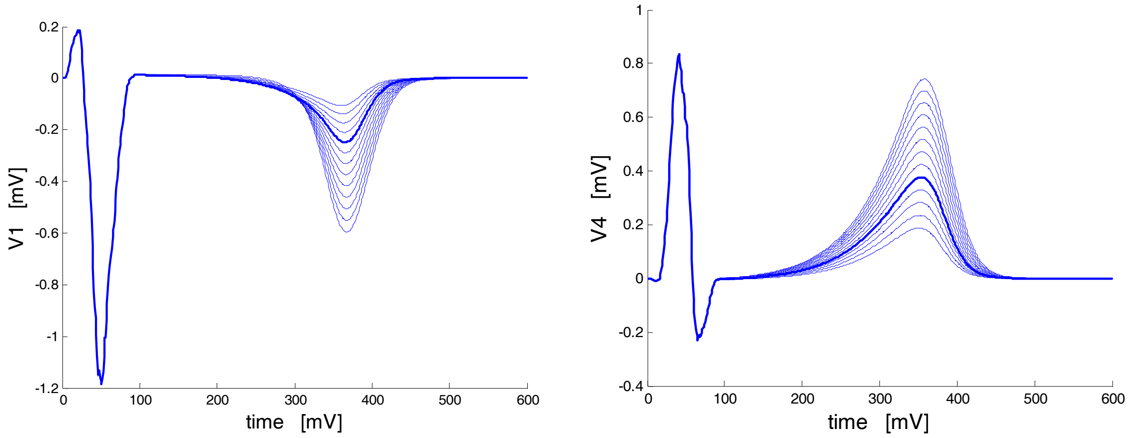


Fig. 9. The effect of the dispersion of the timing of repolarization on T wave magnitude as observed in simulations based on the EDL source model. Left panel: lead V1, right panel: lead V4. With reference to the dispersion involved in the simulation most closely resembling the measured data (heavy trace), the other traces result from scaling the dispersion by factors ranging from 0.5 to 2, in 12 steps of 0.25, while keeping the spatial distribution (pattern) of the ρ values the same. Note that the major effect of this is a scaling of T wave magnitude.

The mean timing of the T wave, marker *a*, has a special significance. A comparison with the mean value of the repolarization parameter found in the inverse procedure ($\bar{\rho}$ = 355 ms) shows that both are equal. For the normal ECG, the mean timing of the T waves is the mean value of the timing of the inflection point of the downslope of the transmembrane potentials. The correspondence is in particular close if the range of the ρ values is small. Moreover, in that situation the shape of a "typical" T wave, such as the one shown for lead V4 (Fig. 7), resembles the derivatives of any of the transmembrane potential waveforms, which under these conditions have very similar shapes. They are almost replicas of ono another, approximately being just shifted and slightly stretched in time according to the local values of δ_n and ρ_n. A mathematical analysis of this property, essentially based directly on (9), can be found in [29].

The significance of the amplitude of the T wave can also be interpreted on the basis of the EDL source model. An example of this is illustrated in Fig. 9. It demonstrates the effect of the dispersion of the timing of repolarization on T wave magnitude as observed in simulations based on the EDL source model. Left panel: lead V1, right panel: lead V4. With reference to the dispersion involved in the simulation most closely resembling the measured data (thick trace), the other traces result from scaling the dispersion by factors ranging from 0.5 to 2, in 12 steps of 0.25, while keeping the spatial distribution (pattern) of the ρ values the same. The effect of using a scaling factor *f* resulted in the scaling of the T wave magnitudes of all leads by the same factor. Significantly, for a scaling factor of zero, associated with a hypothetical uniform timing of repolarization, the magnitude of the T wave is zero in all leads. Over the wide range of scaling factors shown, the effect of the global scaling of dispersion on T wave morphology is only marginal. Note that increasing the scaling factor is accompanied by an apparent increase in the width of the T wave. This is a secondary effect: the basic shape of the T wave remains –in close approximation- the same.

The ST segment

As is mentioned in the Introduction, the signals during the interval leading up to its apex, the ST segments, contain useful diagnostic information, in particular with respect to episodes of acute ischemia that might lead to myocardial infarction (MI). In healthy subjects the majority of the 12-lead signals show relatively small values (Figs 4 and 8). During the various stages of acute ischemia, extreme changes in the ST level may be observed. The ST level is used to monitor the ongoing process. A direct interpretation of the ST level on the basis of the underlying electrophysiology is not always straightforward. A part of this problem relates to basic factors involved in the recording of bioelectric signals as discussed in Sect. 2. Below, ST changes as observed in the ECG are discussed, while incorporating these factors and supporting the analysis on the basis of the EDL source model.

As is discussed in Sect.2, bioelectric potentials observed by means of electrodes are contaminated by contact potentials at the electode-tissue interface. The magnitudes of these potentials are much larger than those of the ECG, unknown, and may change slowly in time: the so-called baseline ‘wander’ or ‘drift’. This problem is treated by including a highpass filter in the first stage of the amplifier. The resulting signal is referred to as an AC-recording (AC=alternating current). This produces an output signal, which, when computed over a long time, has zero mean value: the DC component of the signal is zero. If rapid perturbations of the contact potential occur, the baseline of the signal, i.e. the values in the observed signal that should be assigned a zero value, must be specified on a beat-to-beat basis. Even if the dynamic range of the input stage were wide enough to encompass the full range of sensor potentials (allowing a DC recording in spite of the presence of the contact potentials) the subseqent shift to a zero level remains to be *defined*.

The treatment of the problem is generally referred to as baseline *correction*. However, this is a somewhat misleading term since it assumes that the signal level that should be specified as zero is self evident, which is, in fact, not the case. The problem involved is illustrated in the following example.

The ECG records express differences in the potential field generated at different locations on the thorax. This potential field is uniform over the thorax (thus resulting in zero ECG potential differences) if all myocytes have the same transmembrane potential. In healthy myocardial tissue, the time instant at which this situation is most closely approximated is just before the beginning of atrial depolarization: the onset of the P wave. The signal of lead V3 shown on the upper trace of the left panel of Fig. 10 is the (unprocessed) output of an AC-coupled amplifier. For this healthy subject and this particular beat, the onsets of the P-waves were close to zero, and so only a minor shift was required to refer this signal to zero at the onset of the subsequent P waves.

The problem of baseline definition is more pressing for recordings taken during periods of acute ischemia. During ischemia, the transmembrane potential (TMP) of the myocytes within the ischemic region change: the resting potential decreases (tends to zero), the upstroke of the TMP is reduced. The size of these changes varies during the various stages of ischemia [37]. The intercellular intra-cellular coupling between myocytes I in the ischemic region and myocytes H in the healthy region and the differences in the TMPs of the myocytes within these regions result in a current flow. In the intracellular myocardium, there is a current flow from cells I to H during the time interval between the end of the T wave and the onset QRS interval. In the extracellular domain the direction of the current flow has to be in the opposite

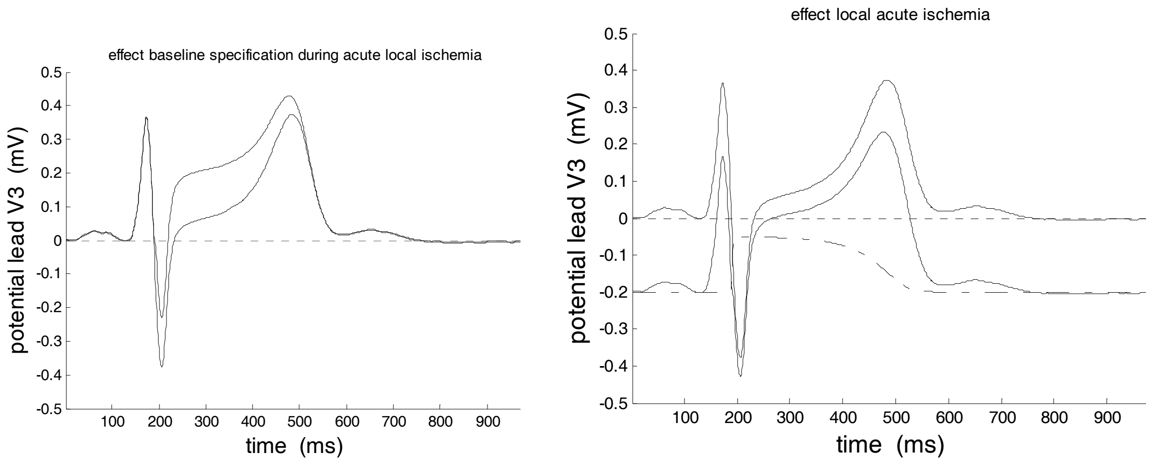


Fig. 10. Lead V3 signals. Left panel, upper trace: healthy myocardium; lower trace: true, i.e. DC variant, of lead V3 during ischemia; dashed line: the difference. Right panel, lower trace: once more the healthy situation, upper trace: the waveform as observed in the ECG after baseline "correction". Note that the true ST depression that takes place (left panel) is observed as an elevation in the ECG.

direction, since charge is conserved. This can be expressed by an equivalent current sink at the ischemic region, which lowers the nearby extracellular potentials [37].

In the lower traces of Fig. 10 the effect of this current sink on the potential of lead V3 is simulated by a downward DC shift of 0.2 mV. During the activation of the tissue surrounding the ischemic zone the potential difference between the regions is smaller and the loading effect is smaller. In this simulation, the contribution from the ischemic zone to the potential of lead V3 was set to -0.05 mV. During repolarization the contribution of the sink slowly returns to the -0.2 mV level. The time course of the contribution of the current sink is shown by the dash-dot line. The addition of this contribution to the signal of the non-ischemic situation (the upper trace) results in the type of waveform for the V3 signal that can be expected during the ischemic period (the lower solid trace on the left panel of Fig.10.). Note that between the end of the T wave and the moment of depolarization reaches the ischemic-zone (R-S segment) the signal is lowered by 0.2 mV and the remaining part by 0.05 m V, then followed by a slow return to the (depressed) value of -0.2 mV at the end of the repolarizing phase.

Standard ECG recordings do not reveal the baseline depression shown in the lower solid trace in the right panel of Fig.10. The presence of the highpass filter results in a signal that has a zero mean value. A (subsequent) baseline 'correction' to the onset of the P wave then results in the upper trace of the same figure. Note that as a result of the current sink within the ischemic zone, the actual baseline depression shows up as an elevation of the ST-segment, merely as a result of the *defining* the zero baseline at the onset of either the P wave or QRS complex. This phenomenon hinders a direct interpretation of the ECG features in terms of the underlying source-sink configurations associated with electrophysiology.

6. Discussion

As shown in this paper, the measurement of the T wave as such is a straightforward procedure. The major concern in the correct interpretation of the observed waveform features originates from the nature of the currents generating the T wave. During the depolarization phase, the dominant source distribution is restricted to the depolarizing wave front only. In contrast, strength of the sources during repolarization is much smaller and their distribution is present throughout the myocardium. The lack of a complete set of invasive data on the repolarization of the entire myocardium clearly hampers the development of a validated source model. The EDL source model used in the analysis of this paper seems to be a workable model, which, like all models, may need to be replaced by an alternative should more complete data become available. The following "conclusions" are inspired by the analysis based on this model.

The analysis shows that the major significance of the timing of apex_T is that of the mean value of the timing of the downstroke of the (mean) transmembrane potential of the myocytes. This is in fact a property that follows directly from the completely general expression (4), a property related to the fundamentals of volume conduction theory, and applies independently of the limiting assumptions involved in the derivation leading up to the EDL source model. This statement may seem to be in contrast to those based on measured data that state that the timing of apex-T signals the recovery time of the epicardial surface. However, the latter statements are usually based on *in vitro* observations of small tissue segments. These do not take into account the full complexity of the diffuse nature of repolarization currents, which are present continuously within the entire ventricular myocardium.

The analysis of the results shown in Fig. 9 indicated that the magnitude of the T wave is the proper indicator of the magnitude of the dispersion of the timing of repolarization. The width of the observed T wave is a secondary factor, related to magnitude. Note that the scaling factor f is unknown in clinical situations. However, the scaling property can be used while monitoring ST changes in individual patients. As in all other ECG wavelets, the magnitude of apex_T in the individual leads of a subject depends on heart position relative to the electrode positions involved as well as on the orientation of the heart. This causes voltage related ECG parameters to be less robust than desirable. The use of the EDL model in the analysis of ST changes during acute ischemia has emphasized the current confusion in the interpreting such observed data and yielded an explanation consistent with observations and interpretations of results from electrophysiological studies [37] for the corresponding effects on body surface potentials. It is a problem that is inevitably linked to the polarization effect at the electrode-tissue interface, which causes the observation of the involved DC shifts to be impossible.

In conclusion, the proper interpretation of repolarization processes on the basis observed T wave morphology needs to be supported by adequate modeling of the electrophysiological properties of the tissue, while fully taking into account the physics involved.

References

[1] W. Einthoven and K. de Lint. Ueber das normale menschliche Elektrokardiogram und Uber die capillar-elektrometrische Untersuchung einiger Herzkranken, *Pflugers Arch ges Physiol*, vol. 80, pp. 139-160, 1900.

[2] H. Ritsema van Eck, J. A. Kors, and G. van Herpen. The U wave in the electrocardiogram: A solution for a 100-year-old riddle.. *Cardiovasc Res*, vol. 67, pp. 256-262, 2005.

[3] D. di Bernardo and A. Murray. Origin on the electrocardiogram of U-waves and abnormal U-wave inversion. *Cardiovasc Res*, vol. 53, pp. 202-208, 2002.

[4] Einthoven 2002; 100 Yeart of Electrocardiography. M. J. Schalij, M. J. Janse, A. van Oosterom, E. E. van der Walle, and H. J. J. Wellens, Eds. Leiden: The Einthoven Foundation, 2002, pp. 616.

[5] J. Webster. Medical Instrumentation, Application and Design. 3 ed. New York: John Wiley and Son, 1998.

[6] A. C. MettingVanRijn, A. P. Kuiper, A. C. Linnenbank, and C. A. Grimbergen. Patient isolation in multichannel bioelectric recordings by digital transmission through a single optical fiber. *IEEE Trans Biomed Eng*, vol. 40, pp. 302-8, 1993.

[7] X. Winter and e. al. Driven-Right-Leg Circuit Design. *IEEE Trans Biomedical Eng*, vol. BME-30, pp. 62-66, 1983.

[8] P. W. Macfarlane. Lead Systems. In: *Comprehensive electrocardiology: Theory and Practice in Health and Disease*, vol. 1, P. W. Marfarlane and T. D. Veitch, Eds. Oxford: Pergamon Press, 1989, pp. 315.

[9] R. Hoekema, G. J. H. Uijen, and A. van Oosterom. The Number of Independent Signals in Body Surface Maps. *Meth. Inform. Med.*, vol. 38/2, pp. 119-124, 1999.

[10] B. Taccardi. Distrbution of Heart Potentials on the Thoracic Surface of Normal Human Subjects. *Circulation Research*, pp. 341-352, 1963.

[11] P. W. Macfarlane. *Normal Limits*, vol. volume III. Oxford: Pergamon Press, 1989.

[12] W. M. Jackman, K. J. Friday, J. L. Anderson, and E. M. E. M. A. C. R. Lazzara. The long QT sybdromes: a critical review, new clinical observations and a unifying hypothesis. *Prog Cardiovasc Dis*, vol. 31, pp. 115-172, 1988.

[13] C. Antzelevitch, P. Brugada, J. Brugada, R. Brugada, J. A. Towbin, and K. Nademanee. Brugada Syndrome: 1992-2002. *Journal of the American College of Cardiology*, vol. 41, pp. 1665-1671, 2003.

[14] I. Gussak and C. Antzelevitch. Early repolarization syndrome: Clinical Characteristics and possible cellular and ionic mechanisms. *J Electrocardiol*, vol. 33, pp. 299-309, 2000.

[15] T. J. C. Faes. Assessment 0f Cardiovascular Autonomic Function; An Inquiry into measurement. In: *Faculty of medicine*, vol. PhD. Amsterdam: Vrije Universiteit Amsterdam, 1992.

[16] F. N. Wilson, A. G. Macleod, and P. S. Barker. The Distribution of Action Currents produced by the Heart Muscle and Other Excitable Tissues immersed in Conducting Media. *J Gen Physiol*, vol. 16, pp. 423-456, 1933.

[17] R. Plonsey. An Extension of the Solid Angle Formulation for an Active Cell. *Biophysical J*, vol. 5, pp. 663-666, 1965.

[18] W. R. Smythe. *Static and Dynamic Electricity*. New York: McGraw-Hill, 1968.

[19] W. K. H. Panofski and M. Phillips. *Classical Electricity and Magnetism.* London: Addison-Wesley, 1962.

[20] D. B. Geselowitz. On the Theory of the Electrocardiogram. *Proc IEEE*, vol. 77/6, pp. 857-876, 1989.

[21] A. van Oosterom. The Equivalent Double Layer; source model for repolarization. In: *Comprehensive Electrocardiology*, vol. I, P. W. Macfarlane, A. v. Oosterom, M. J. Janse, P. Kligfield, J. Camm, and O. Palm, Eds. New York: Springer-Verlag, 2009.

[22] D. B. Geselowitz. Description of cardiac sources in anisotropic cardiac muscle. Application of bidomain model. *J Electrocardiol*, vol. 25 Sup., pp. 65-67, 1992.

[23] R. M. Gulrajani. *Bioelectricity and Biomagnetism*. New York: John Wiley & Sons, 1998.

[24] R. C. Barr, M. Ramsey, and M. S. Spach. Relating Epicardial to Body Surface Potentials by means of Transfer Coefficients based on Geometry Measurements. *IEEE Trans. Biomed. Eng.*, vol. BME-24, pp. 1-11, 1977.

[25] T. F. Oostendorp and A. van Oosterom. Source parameter estimation in inhomogeneous volume conductors of arbitrary shape. *IEEE Trans. Biomed. Eng.*, vol. BME-36, pp. 382-391, 1989.

[26] P. M. van Dam and A. van Oosterom. Volume conductor effects involved in the genesis of the P wave. *Europace*, vol. 7, pp. S30-S38, 2005.

[27] G. J. M. Huiskamp. Simulation of Depolarization and Repolarization in a Membrane Equations based Model of the Anisotropic Ventricle. *IEEE Trans Biomed Eng,* vol. BME-45, pp. 847-855, 1998.

[28] G. Fischer, F. Hanser, B. Pfeifer, M. Seger, C. Hintermuller, R. Modre, B. Tilg, T. Trieb, T. Berger, F. X. Roithinger, and F. Hintringer. A Signal Processing Pipeline for Noninvasive Imaging of Ventricular Preexcitation. *Methods of Informetics in Medicine,* vol. 44, pp. 588-515, 2005.

[29] A. van Oosterom. The Singular Value Decomposition of the T wave: Its Link with a Biophysical Model of Repolarization. *Int. J. Bioelectromagnetism*, vol. 4, pp. 59-60, 2002.

[30] A. van Oosterom and T. Oostendorp. ECGSIM; an interactive tool for studying the genesis of QRST waveforms. *Heart,* vol. 90, pp. 165-168 2004.

[31] A. van Oosterom and V. Jacquemet. A Parameterized Description of Transmembrane Potentials used in Forward and Inverse Procedures. presented at Int Conf Electrocardiol, Gdansk; Poland, 2005.

[32] C. W. Haws and R. L. Lux. Correlation Between In Vivo Transmambrane Action Potential Durations and Activation-Recovery Intervals From Electrograms. *Circulation,* vol. 81/1, pp. 281-288, 1990.

[33] G. J. H. Huiskamp and A. van Oosterom. The depolarization sequence of the human heart surface computed from measured body surface potentials. *IEEE Trans Biomed Eng,* vol. 35, pp. 1047-1058, 1989.

[34] D. Durrer, R. T. van Dam, G. E. Freud, M. J. Janse, F. L. Meijler, and R. C. Arzbaecher. Total excitation of the isolated human heart. *Circulation*, vol. 41, pp. 899-912, 1970.

[35] M. R. Franz, K. Bargheer, W. Rafflenbeul, A. Haverich, and P. R. Lichtlen. Monophasic action potential mapping in a human subject with normal electrograms: direct evidence for the genesis of the T wave. *Circulation,* vol. 75/2, pp. 379-386, 1987.

[36] J. C. Cowan, C. J. Hilton, C. J. Griffiths, S. Tansuphaswadikul, J. P. Bourke, A. Murray, and R. W. F. Campbell. Sequence of epicardial repolarization and configuration of the T wave. *Br Heart J,* vol. 60, pp. 424-433, 1988.

[37] M. J. Janse, F. J. L. v. Capelle, H. Morsink, A. G. Kleber, F. Wilms-Schopman, R. Cardinal, C. N. d'Alnoncourt, and D. Durrer. Flow of injury current and patterns of excitation during early ventricular arrhythmias in acute regional myocardial ischemia in isolated porcine and canaine hearts; evidence for two different arrhythmogenic mechanisms. *Circ Res*, vol. 47, pp. 151-165, 1980.

Biophysical Model for Beat-to-Beat Variations of Vectorcardiogram

E. A. I. Aidu, V. G. Trunov, L. I. Titomir

Institute for Information Transmission Problems,
Russian Academy of Sciences, Moscow, Russia
Email: aidu@iitp.ru

Abstract. *For a realistic biophysical model, a linear deformation of the myocardium lead to a linear transformation of the orthogonal electrocardiosignals under the condition of topological invariability of the depolarization and repolarization processes in the myocardium. Certain relationships have been found between the linear transformations in the physical and functional spaces. Several practically important conclusions follow as a logical consequence of these statements.*

Keywords: Vectorcardiogram, Biophysical Model, Beat-to-Beat Alteration

1. Introduction

Comparison of single beats within one ECG recording or consecutive ECGs from the same patient during cardiologic tests and in long term monitoring is an essential part of the contemporary noninvasive electrocardiologic diagnosis.

Variations of the ECG signals from one heartbeat to another are affected by a number of factors, which do not reflect the electrophysiological state of the myocardium itself and complicate the analysis of the ECG signals. These factors are usually treated as distortion or extracardiac factors [1]. Considering comparison of consecutive ECG cycles from the same person, the main distortion factors are the patient's breathing movements, possible changes in position, and haemodynamic alterations resulting in rotation and deformation of the miocardium.

A simple biophysical model of excitable media is proposed to describe the dipole moment changes after deformation of the excitable media (myocardium). The first approximation of the deformation is linear transformation.

2. Subject and Model

Suppose, that the excitable media (region M, Fig. 1) was deformed by linear transformation **T** (region M_T, Fig. 1):

$$\rho = \mathbf{T}r, \tag{1}$$

where $\rho = (\xi, \eta, \zeta)' \in \mathsf{M}_\mathrm{T}$ is the new position of the myocardium point $r = (x, y, z)' \in \mathsf{M}$ after transformation **T** ($'$, transpose symbol). Let $d, \mathrm{J}(r)$ and $d_\mathrm{T}, \mathrm{J}_\mathrm{T}(r)$ be the dipole moments and current density before and after transformation **T,** respectively. Dipole moment vectors are expressed as integrals of current density over the excitable media regions:

$$d = \int_{\mathsf{M}} \mathrm{J}(\rho)\, \mathrm{d}v_\rho \,,\; d_\mathrm{T} = \int_{\mathsf{M}_\mathrm{T}} \mathrm{J}_\mathrm{T}(\rho)\, \mathrm{d}v_\rho \,. \tag{2}$$

For bidomain model of the excitable media, the current density is determined by gradient of the transmembrane potential [2], then before and after transformation **T** the current densities are:

$$J(r) = -\sigma_i \nabla_r U(r), \quad J_T(\rho) = -\sigma_i \nabla_\rho U_T(\rho), \tag{3}$$

where σ_i – intracellular conductivity, $U(r)$, $U_T(\rho)$ – transmembrane potentials before and after transformation **T**, $\nabla_r = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$; $\nabla_\rho = \left(\frac{\partial}{\partial \xi}, \frac{\partial}{\partial \eta}, \frac{\partial}{\partial \zeta}\right)$.

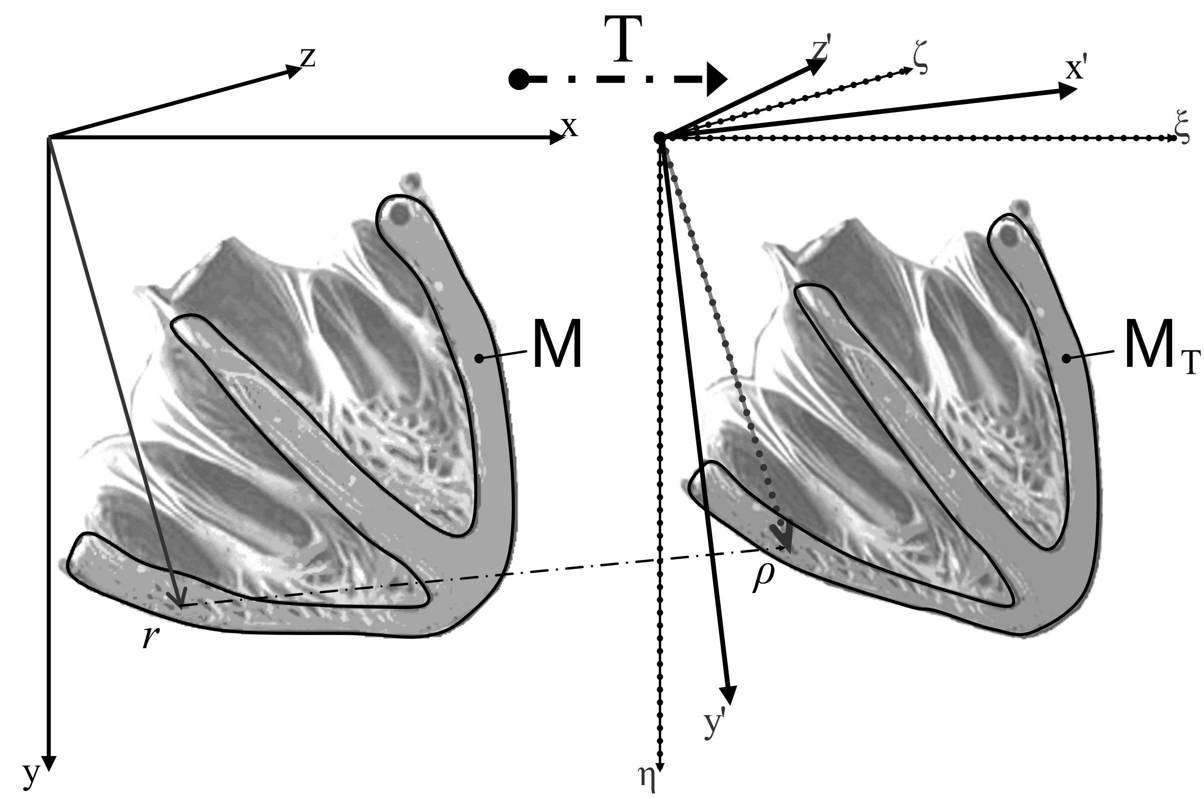


Fig. 1. Deformation of myocardium. The point $\rho = (\xi, \eta, \zeta)' \in \mathsf{M}_T$ is the new position of the myocardium point $r = (x, y, z)' \in \mathsf{M}$ after transformation **T**.

Assume further that the state of myocardial cells does not change after transformation **T**, and transmembrane potential remains the same, so

$$U_T(\rho) = U_T(\mathbf{T}r) = U(r). \tag{4}$$

The main aim of the present reasoning is to obtain the relation between dipole moments before and after the deformation of the excitable media. Using Eqs. 1 – 4, taking into account that gradient of a function is a contravariant tensor of rank 1, after several simple manipulations and change of variables $r = \mathbf{T}^{-1}\rho$, the relation is obtained:

$$\begin{aligned} d_T &= \int_{\mathsf{M}_T} J_T(\rho)\, \mathrm{d}v_\rho = -\sigma_i \int_{\mathsf{M}_T} \nabla_\rho U_T(\rho)\, \mathrm{d}v_\rho = -\sigma_i \int_{\mathsf{M}} (\mathbf{T}')^{-1} \nabla_r U(r) \frac{\partial v_\rho}{\partial v_r}\, \mathrm{d}v_r \\ &= -\sigma_i (\mathbf{T}')^{-1} |\mathbf{T}| \int_{\mathsf{M}} \nabla_r U(r)\, \mathrm{d}v_r = (\mathbf{T}')^{-1} |\mathbf{T}|\, d, \quad |\mathbf{T}| = \det(\mathbf{T}). \end{aligned} \tag{5}$$

3. Results and Corollaries

For the bidomain model, a linear deformation **T** of the excitable media leads to linear transformation **G** of the dipole moment under the condition of the topological invariability of the activation propagation. The relation between these two linear transformations is:

$$d_{\mathrm{T}} = \mathbf{G}\, d\,, \quad \mathbf{G} = \left(\mathbf{T}'\right)^{-1}\left|\mathbf{T}\right|. \tag{6}$$

Assume further that vectorcardiogram (VCG) reflects the heart vector evolution during depolarization process in the heart. Let $\{\mathbf{f}_0(t), \mathbf{f}_1(t), \ldots, \mathbf{f}_n(t)\}$ be a sequence of n QRS loops, and $\mathbf{G}_i$ be the linear transformation that maps $\mathbf{f}_0$ (reference loop) to approximate $\mathbf{f}_i$ with the minimum root mean square error ε_i:

$$\mathbf{f}_i = \mathbf{G}_i\,(\mathbf{f}_0)+\mathbf{e}_i \;(i=1,..n),\; \varepsilon_i = \|\, \mathbf{e}_i \,\| \tag{7}$$

The transformations $\mathbf{G}_i$ of the reference QRS loop are represented by the rotations $\mathbf{O}_i$ and dilatations $\mathbf{S}_i$: $\mathbf{G}_i = \mathbf{O}_i\,\mathbf{S}_i$. Any linear transformation in the Euclidean space may be obtained by sequential executing rotation and dilatation (in any sequence). This provides separate analysis of these two kinds of transformations. The rotation of the reference QRS loop is characterized by the eigenvector of this transformation and the angle of turn around this vector.

The dilatation of the reference QRS loop is characterized by the coefficients of dilatation along three orthogonal directions, or, in other words, by the eigenvalues (σ_{i1}, σ_{i2}, σ_{i3}) and three orthogonal eigenvectors of this transformation. The product of these three eigenvalues is equal to the factor of the volume expansion for any three dimensional body after the transformation.

Looking back to the corresponding myocardium transformations $\mathbf{T}_i$ through Eq. 6, it is possible to calculate the eigenvalues (λ_{i1}, λ_{i2}, λ_{i3}) and eigenvectors that characterize the myocardium expansion (or contraction) for the i-th beat against reference beat. The product of eigenvalues $\lambda_i = \left|\mathbf{T}_{\mathrm{i}}\right| = \lambda_{i1}\,\lambda_{i2}\,\lambda_{i3}$ may be treated as the end-diastolic heart volume dilatation (or contraction) relative to the reference beat.

The simple relations between the eigenvalues of linear transformations in the physical and functional spaces are obtained using Eq. 6:

$$\sigma_{i1} = \lambda_{i2}\,\lambda_{i3}\,,\; \sigma_{i2} = \lambda_{i1}\lambda_{i3}\,,\; \sigma_{i3} = \lambda_{i1}\,\lambda_{i2}\,;\; \sigma_i = \lambda_i^2 \;\text{ or }\; \left|\mathbf{G}_i\right| = \left|\mathbf{T}_i\right|^2. \tag{8}$$

The same results may be obtained while analyzing the double layer activation front deformation (Fig. 2), and for more general cases.

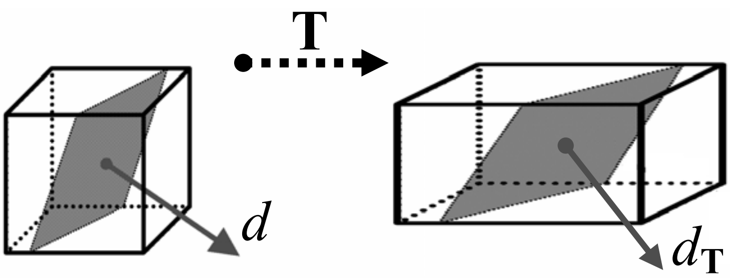


Fig. 2. Elementary cube and double layer dilatation along three eigenvectors of transformation **T**. Calculation of the relation between dipole moments before and after transformation and integration over the whole surface of excitation gives the Eqs. 6, 8.

4. Discussion

The influence of the ventricular volume on the electric heart vector is usually explained by Brody effect [1] without mentioning the changes of myocardium that contain the sources of the electric field. Brody effect means that the intracardiac blood surface plays a role of an imperfectly reflecting curved mirror. Changes in VCG are caused both by myocardial deformation and its reflection on intracardiac blood. It is important to find out their relative contribution. The same changes of VCG after changes of the left ventricular end-diastolic volume may be easily explained by Brody effect [1] and also by linear deformation of myocardium.

The presented model was used in [3], where a set of vectorcardiograms recorded during the parabolic flights of a laboratory aircraft was analyzed. In all normal cases, the QRS loop for each person, as a curve in the three dimensional (3D) vectorcardiographic space, remains virtually unchanged after the proper 3D linear transformation (the relative error of approximation was less than 0.05). The distance between the superimposed QRS loops served as an indicator of changes in the heart depolarization process. Changes of the QRS volume factor $|\mathbf{G}_i|$ were in accordance with haemodynamic changes due to gravitation acceleration.

5. Conclusions

For the bidomain model, linear deformation of the excitable media leads to linear transformation of the dipole moment under the condition of the topological invariability of the repolarization and depolarization processes.

The parameters of linear VCG transformations (eigenvectors, eigenvalues, rotation angles, determinants) may be used as indices of the heart position and haemodinamic changes.

The real interrelation of VCG transformations and myocardium deformations is much more complex. Further theoretical and experimental investigations are needed to assess the validity of the proposed biophysical model and the results obtained.

Acknowledgements

This work was supported by Presidium of the Russian Academy of Sciences, program “Basic Sciences for Medicine”.

References

[1] Malmivuo J, Plonsey R. Bioelectromagnetism: Principles and Application of Bioelectric and Biomagnetic Fields. Oxford University Press, New York, 1995.

[2] Miller W. T., Geselovitz D. B. Simulation studies of the electrocardiogram. 1. The normal heart. *Circulation Research*, 43:301-315, 1978.

[3] Aidu E.A.I., Trunov V.G., Titomir L.I., Capderou A., Vaïda P. Transformation of vectorcardiogram due to gravitation alteration. In: Measurement 2003. Bratislava: Inst. Meas. Sci. SAS - VEDA, 2003, 169-172.

Evolutionary Algorithms for Cardiovascular Decision Support

B. Gaál , A. Endrődi, K. Fülöp, A. Király, I. Vassanyi, G. Kozmann

Department of Electrical Engineering and Information Systems, Veszprém, Hungary,
Email: bgaal@almos.vein.hu

Abstract. *The extraction of information from cardiovascular models may improve the diagnosis and treatment of cardiovascular diseases. Retrieving information on the dynamics of the cardiovascular system is done by identifying the structure, order and parameters of various type of models, traditionally lumped parameter and ARMA models, which provide a way to extract knowledge on the overall state of the cardiovascular system. There are specific system identification techniques developed for specific cardiovascular models. Our method is general, easily applicable to various model variations. The identification of structure, order and parameters is considered as a multi-objective optimization problem, which we solve with Genetic Algorithms (GAs). The paper deals with the comparison of variants of GAs for the task, and introduces a novel multi-level GA approach (muleGA), which, in this particular application, is preferable over the more traditional Pareto-based and aggregating fitness techniques.*

Keywords: Cardiovascular Modelling, Parameter Estimation, Genetic Algorithms

1. Introduction

Non-invasive continuous-time beat-to-beat measurement and estimation of highly informative cardiovascular parameters such as arterial blood pressure, cardiac output and stroke volume is becoming more available since the introduction of devices such as Portapres or Finapres based on the Penaz method, making the signal acquisition available outside the intensive care unit (ICU). Invasive beat-to-beat measurement of arterial blood pressure is carried out in the intensive care or high-dependency unit as a key monitoring technique of real-time hemodynamic monitoring systems. Both invasive and non-invasive measurements are used mainly for monitoring purposes, no sophisticated data-mining and time-series analyzing and forecasting algorithms are exploited in present day medical practice.

Advanced techniques for modeling the dynamics of the cardiovascular system however exist [1, 2 ,3], although these are mostly used for research purposes [4, 5]. By the means of high performance generic methods, employing the search and optimization capabilities of Genetic Algorithms [6], real-time estimation of hard-to-measure cardiovascular parameters is possible to some extent. Encouraging results were presented on the application of evolutionary search methods for model identification [7] and for cardiovascular modeling [8]. Zaid et al. reports basic results on simulating the cardiovascular state with a model whose parameters are adjusted according to the patient's easily measureable vital signals [9, 10]. This topic was further researched by Heldt et al. [11]. The latter system not only executes model identification but runs pattern recognition and physiologic reasoning tasks to comprehensively form a modern ICU device. As Zaid et al [9], we also use the 9-compartment CVSIM model [12] both for the simulation of cardiovascular signals (virtual patient), and for adjusting the model's configuration to have its outputs match the outputs of the virtual patient's (see Fig.1).

2. Subject and Methods

We used and extended the GALib [13] GA-library with multi-objective [14] and multi-level functionality. Four type of GAs were tested against each other in our single-level test configurations, Steady-State-GA [13], micro-GA [15], NSGA-II[16] and SPEA-II [17]. For the multi-level experiments, we used our published multi-level algorithm [18] fine-tuned for the current application. In short, the multi-level GA (muleGA) works the same way as any other GA, with the main difference being the individuals of the populations, as in our multi-level case, the information represented by individuals are functions of so-called lower level GAs, which the individuals encompass. So, each individual runs few GAs in itself. When these lower-level GAs are evolved is detailed in our previous research paper [18]. Each GA in the multi-level structure can be of any type. In our multi-level tests, at each level, we use the Simple-GA, provided by the GALib library [13]

Our problem statement was: for the given output (virtual patient) signals find the input parameters that would be difficult to measure otherwise. The simulator at our disposal (CVSIM) did an excellent job at the forward calculation, and we tried to find the most suitable input vector to the simulator by using single- and multi-level genetic algorithm structures. The search space was 28-dimensional (of which 10 were held constant) and we considered four output signals (x0 - Left Ventricle Pressure, q1 - Arterial Flow, q3 - Right Ventricle Flow, v5 - Pulmonary Venous Volume). Each time a candidate input vector evaluated we programmed the simulator with it and ran the simulator until we got 700 samples for the output signals our experiment was concerned with. Then for each signal we calculated the absolute difference from the reference over all samples and normalized the result. The objective of the optimization was to minimize the average of the differences.

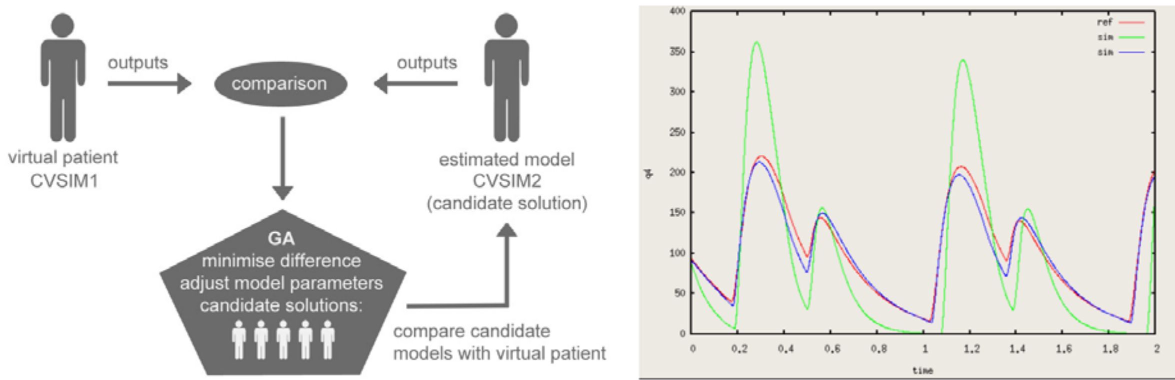


Fig. 1. The figure on the left shows our test configuration. The parameter sets evolved by the Genetic Algorithm(GA) are run in candidate models which are compared to the virtual patient's output. The absolute difference of the output signals are taken into account by the GA's fitness function. The figure on the right shows one output from the virtual patient (red) and one output from the best candidate solution plotted 2 times, at the first (green) and at the last iteration (blue). Note, that in the last iteration, the output of the virtual patient and the output of the estimated model are virtually the same.

Input vectors were represented by 28-gene genomes. The initial value of the genes was chosen randomly from a specific range. The genome mutated by randomly shifting all of its genes which weren't kept constant, while ensuring they don't fall out of bounds.

In the multi-level case, bottom-level GAs encoded full input vectors in order to be able to feed the simulator, but only aimed to optimize for two output signals out of four. Genomes of the top-level GA were constructed such that each of them had two subordinate GAs, one of them optimizing for one pair of the output signals, and the other for the rest. During the evaluation, a top-level individual will fed the simulator (CVSIM2 on Fig.1) with the average of the input vectors represented by the current best solutions of its subordinate GAs. On the top-level, all output signals of the simulator is considered for evaluation.

3. Results and Discussion

First we tested the traditional single level GAs. Results show (see Fig.2) that there is no significant difference in the goodness of solutions regardless of GA type, however the Steady-State GA running an aggregating fitness function had significantly better running time (10-20 times faster) than the Pareto-based approaches.

	MicroGA	SPEA-II	SState	NSGA-II
MicroGA		TT: $0 \rightarrow p = 0.20$ KS: $0 \rightarrow p = 0.55$	TT: $0 \rightarrow p = 0.54$ KS: $0 \rightarrow p = 0.79$	TT: $0 \rightarrow p = 0.98$ KS: $0 \rightarrow p = 0.79$
SPEA-II			TT: $0 \rightarrow p = 0.55$ KS: $0 \rightarrow p = 0.55$	TT: $0 \rightarrow p = 0.23$ KS: $0 \rightarrow p = 0.55$
SState				TT: $0 \rightarrow p = 0.53$ KS: $0 \rightarrow p = 0.67$
NSGA-II				
Mean	-3.914	-3.723	-3.818	-3.917
Div	1.094	1.115	1.057	1.227

Fig. 2. The results of comparison of four GAs. Mean values are close to the theoretically optimal -4 (problems specific) value. Paired T (TT) and Kolmogorov-Smirnov (KS) tests fail to reject the null-hypothesis (that the expected values of the compared results are identical) at 5% significance level. We can conclude, that no single-level GA is performing better than the others in this particular problem, as far as the quality of the solution is concerned.

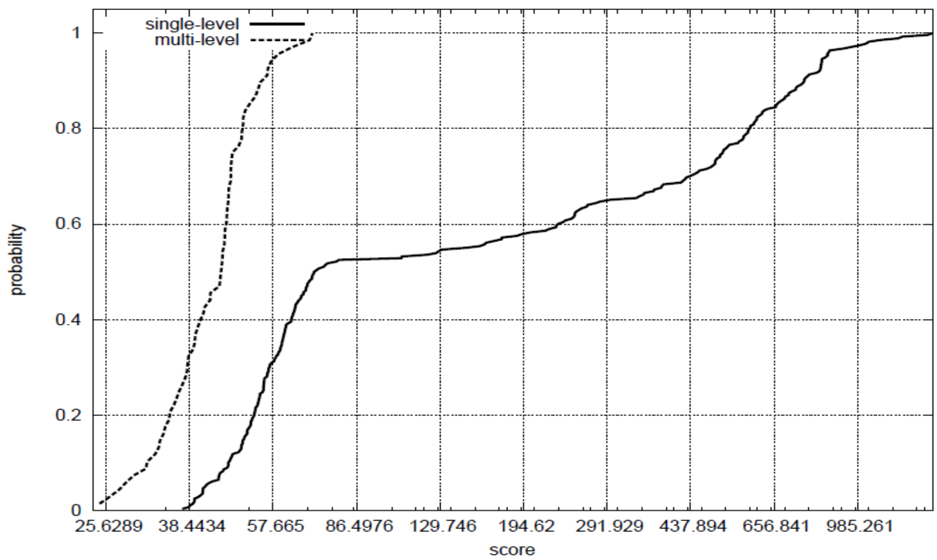


Fig. 3. The probability of solutions found by single-level and multi-level GAs are lower than a specific score (low-is-best) is shown on the figure. For this problem domain, the multi-level GA approach is clearly more powerful considering the quality of the solutions.

We repeated the CVSIM experiment a several times and examined the performance and the objective score distribution of the algorithms. Single-level and multi-level configurations were run more than 300 times to calculate average running time for the different methods. The multi-level technique is 2 to 5 times slower, but still faster than the Pareto-based algorithms. Considering the quality of the solutions, the best results of the single-level algorithm achieved worse scores much frequently. Computing the Kolmogorov-Smirnov test (KS=5.052) confirms the two distributions indeed different (see Fig.3)

4. Conclusion and Future Work

We showed that GAs are capable of estimating the parameters of a lumped parameter cardiovascular model such that it mimics the outputs of the reference signals. As our reference signals were synthesized with the same CVSIM model, future work is needed to test the method with real signals. Our first tests show that for real signals, the 9-compartment CVSIM model is insufficient, so the 21-compartment counterpart needs to be used. Because of this, the efficiency of our multi-level GA is emphasized, as it is capable to perform better in high dimensional search spaces than any of the single-level GA variants tested.

Acknowledgments

The work presented was funded by the Hungarian National Research projects “AALAMSRK” and INNO-6-2007-0021 OMFB-01013/2007.

References

[1] Korhonen. Multivariate closed-loop model for analysis of cardiovascular dynamics. *Methods of Information in Medicine*, Vol 36, pp 264-267, 1997.

[2] R. Mukkamala. A Cardiovascular Simulator for Research-User's Manual and Software Guide. Harvard-MIT Division of Health Sciences and Technology Massachusetts Institute of Technology, Cambridge, November 27, 2001.

[3] R. Mukkamala and R. J. Cohen. A forward model-based validation of cardiovascular system identification. *Am. J. Physiol. Heart Circ Physiol*, Vol 281, pp H2714-H2730, 2001.

[4] T. J. Mullen, M. L. Appel, R. Mukkamala, J. M. Mathias, and R. J. Cohen. System identification of close-loop cardiovascular control: effects of posture and autonomic blockade. *Am. J. Physiol. Heart Circ Physiol*, Vol 272, pp H448-H461, 1997.

[5] R. Mukkamala, J. M. Mathias, T. J. Mullen, R. J. Cohen, and R. Freeman. System identification of closed-loop cardiovascular control mechanisms: diabetic autonomic neuropathy. *Am. J. Physiol. Heart Circ Physiol*, Vol 276, pp R905-R912, 1999.

[6] T. Back. Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms, Oxford University Press, Oxford, UK, 1996.

[7] Susanne Rolf, Joachim Sprave, Wolfgang Urfer (1997). Model Identification and Parameter Estimation of ARMA Models by Means of Evolutionary Algorithms.

[8] Jingyu Liu (2004). Human cardiovascular dynamics identification, simulation, and application using a novel hybrid method. PhD Thesis.

[9] Cardiovascular Parameter Estimation using a Computational Model, Zaid Samar, MSc Thesis, Massachusetts Institute of Technology, May 2005

[10] Model Model-Based Approaches to Signal Analysis, Patient Parameter and State Estimation, Alarm Generation and Hypothesis Generation, presentation. http://mimic.mit.edu/Archive/Presentations/ModelingProjects.pdf (Verified as of 28th February 2009)

[11] Integrating Data, Models, and Reasoning in Critical Care, Proceedings of the 28th IEEE EMBS Annual International Conference New York City, USA, Aug 30-Sept 3, 2006.

[12] http://www.physionet.org/physiotools/cvsim/ (Verified as of 28th February 2009)

[13] The M.I.T. GALib C++ Library of Genetic Algorithm Components at http://lancet.mit.edu/ga/ [Verified 28th Feb, 2009]

[14] C. A. Coello Coello, D. A. Van Veldhuizen, and G. B. Lamont, Evolutionary Algorithms for Solving Multi-Objective Problems, Kluwer Academic Publishers, New York, May 2002, ISBN 0-3064-6762-3.

[15] Carlos A. Coello Coello and Gregorio Toscano Pulido. Multiobjective Optimizationusing a Micro-Genetic Algorithm. In Lee Spector et al., editor, Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'2001), pages 274-282, San Francisco, California, 2001. Morgan Kaufmann Publishers.

[16] K. Deb, S. Agrawal, A. Pratap and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation* 6 (2) (2002). pp. 182-197

[17] N. Srinivas and K. Deb. Multiobjective function optimization using nondominated sorting genetic algorithms. *Evol. Comput*., vol. 2, no. 3, pp. 221-248, Fall 1995.

[18] B. Gaál, I. Vassányi, G. Kozmann. A Novel Artificial Intelligence Method for Weekly Dietary Menu Planning. *Methods Inf Med* 2005; 44: 655–64. ISSN: 0026-1270

Comparison of Several ECG Intervals Used for Identification of Ischemic Lesions Based on Difference Integral Maps

[1]J. Švehlíková, [2]M. Kania, [1]M. Turzová, [1]E. Hebláková, [1]M. Tyšler, [2]R. Maniewski

[1]Institute of Measurement Science, SAS, Bratislava, Slovakia,
[2]Institute of Biocybernetics and Biomedical Engineering, PAS,Warsaw, Poland
Email: umersveh@savba.sk

Abstract. *Repolarization changes in small areas of myocardium can be detected from difference integral maps computed from body surface potentials measured on the same subject in situations with and without manifestation of ischemia. Detection was made by inverse solution with 2 dipoles. On 10 patients and 3 healthy subjects surface potentials were recorded at rest and during stress. Difference integral maps for 4 intervals of integration (QRST, QRSU, STT and STU) of electrocardiographic signal were computed and their properties and applicability as input data for inverse identification of ischemic lesions were compared. The results showed that better (more reliable) inverse solutions can be obtained from difference integral maps computed either from QRST or from STT interval of integration. The average correlation between such maps is 97%. The use of the end of U wave instead of the end of T wave didn't improve the results.*

Keywords: Integral body surface potential maps, Ischemic lesions, Inverse solution to 2 dipoles

1. Introduction

Myocardial ischemia is accompanied by changed shape of the repolarization phase of action potentials (AP) of myocytes in the ischemic area. Previously we reported a method for identification of local ischemic lesions by computing an inverse solution with two dipoles using difference integral maps (DIMs) of the QRST interval [1]. DIMs were computed by subtraction of QRST integral maps obtained under normal conditions from QRST integral maps obtained in situations with manifestation of ischemia.

Use of the whole QRST that represents both, depolarization and repolarization of the myocardium was based on the assumption that ischemic changes could affect also depolarization phase of myocytes AP by reduced AP amplitude and rate of rise [2]. The other possibility would be to evaluate only the STT interval in electrocardiogram (ECG) reflecting only repolarization phase of the myocardium.

Another issue when selecting the ECG interval for evaluation of repolarization changes is determination of the end of the repolarization. In many real signals small U wave appears after the T wave in ECG. According to some studies [3] T and U waves together represent the repolarization period. In such a case inclusion of U wave into the evaluated interval would be desirable.

In presented study the differences between DIMs computed from QRST and STT intervals of measured data were analyzed and obtained results of identification of small ischemic lesions were compared. Also influence of inclusion of U wave into the evaluated interval representing the myocardium repolarization was investigated.

2. Methods and Material

To reveal ischemic lesions with changed repolarization body surface potential maps from the same person recorded in situations with and without manifestation of possible ischemia were measured and used to compute DIMs. Integration intervals QT, QU, STT and STU including the myocardium repolarization phase were alternatively used in the computation supposing that these intervals include information on the changed repolarization. Equivalent cardiac generator with two dipoles was computed from the DIMs by an inverse solution. To find the best pair of dipoles, inverse solution was calculated for all pairs of dipoles located in predefined points evenly distributed within the modeled ventricular myocardium and minimum of rms difference (RMSDIF) between the original DIM and the DIM generated by the inversely estimated pair of dipoles was used. Not only dipoles from the best result but dipoles from all results with RMSDIF within 1% difference from the best solution were analyzed. Modified K-means clustering method was applied on all analyzed dipoles to divide them into 2 clusters [4]. If the dipoles from one dipole pair represented different ischemic lesions, it should be possible to assign them to different clusters. Any pair for which this procedure was not successful was excluded from the evaluation. The final gravity center of each cluster was then calculated and the mean dipole moment computed from all dipoles in the cluster was used to represent the lesion. The number of excluded pairs and mutual distance of cluster centers were used as the criteria to decide whether these clusters represent two separate lesions or only one lesion (i.e. the type of the lesion).

Body surface potential maps of 10 patients (p1-p10) with coronary artery disease and 3 healthy subjects (h1-h3) were computed from 64 ECG leads measured at rest and during an exercise test on supine ergometer at the load of 75 W. Fiducial points Q, S, T and U in averaged ECG signals were determined manually by visual observation of the computed rms signal from all measured leads in each time instant (Fig.1). Then, for each patient, integral maps for the time intervals QT, QU, STT and STU were computed at rest and during exercise. To subtract the integral maps measured at rest and during stress correctly, changes of heart rate were compensated by recalculation of the time integral values to the same interval length. DIMs for these intervals were then used as input data for the inverse procedure. The differences between the data obtained using the four time intervals and their influence on the results of the inverse solution were studied.

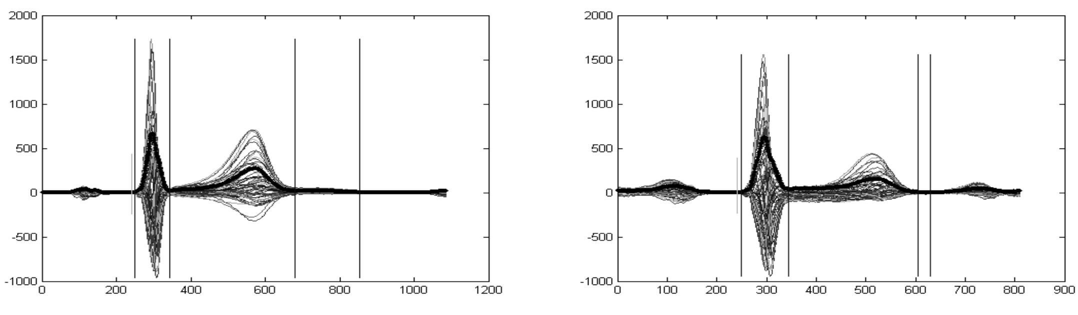


Fig.1. Averaged signals from 64 leads measured on patient p3 at rest (left) and during exercise (right). Black thick line represents rms value of signals. Vertical lines represent estimated fiducial points, from left to right: Q, S, T, U.

3. Results

To study properties of QRST, QRS and STT intervals of integration QRS integral maps were computed for all measured subjects at rest and during the stress. Also DIMs for QRST and

STT intervals of integration were computed. The correlation coefficients between the maps and their rms signals were evaluated.

The correlation of QRS integral maps at rest and at stress was 93-99%, average 97%. RMS values of maps at stress differed from the RMS values of maps at rest from 0% (p6) to 25% (p1), average 10%. Analogically, evaluation of DIMs showed that the correlation between DIMs computed for QRST and STT intervals of integration was high (94–100%, average 97%) except for patient p6 (55%). RMS values of maps differed from -1 to 26 %, average 14%.

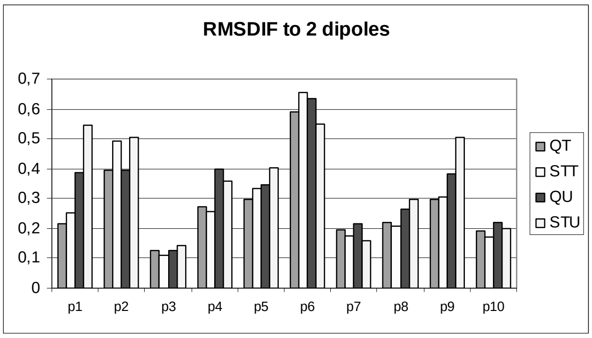
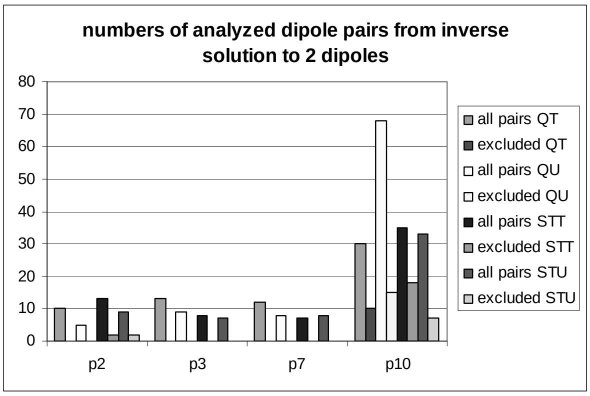


Fig.2. Relative rms differences (RMSDIF) between original DIMs and maps generated by the inversely estimated pairs of dipoles computed in all 10 patients.

To study the influence of U wave inclusion in each patient data DIMs were computed for 4 possible intervals of integration: QT, QU, STT and STU and the differences in inverse solutions with 2 dipoles were observed. The main observed feature of the inverse solution was always the value of the RMSDIF difference between the original DIM and the map generated by the inversely estimated dipole or pair of dipoles (Fig. 2). This value carries information on dipolarity of the input DIM and applicability of the inverse solution for each particular case. According to the RMSDIF values from inverse solutions to 2 dipoles patients were divided into 2 groups. Patients with similar values of RMSDIF regardless the interval of integration (p2, p3, p7, p10) were assigned to the first group. Patient p6 was excluded from further examination because neither 1 dipole nor 2 dipoles could satisfactorily represent his DIM (RMSDIF was 65% or 59% resp.). Rest of patients was assigned to the second group.

In the first group of patients following selected features of inverse solution characterizing the type of the lesion were similar regardless the interval of integration (Fig.3): no excluded pairs from clustering method and a large mutual distance between cluster centers characterizing two simultaneous lesions (p3, p7); small mutual distance determining 1 lesion (p2); large number of excluded pairs and also a large number of solutions indicating that DIM does not represent neither 1 nor 2 small local lesions (p10).



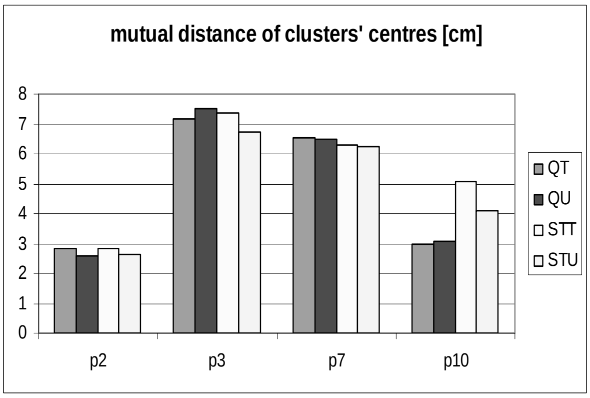
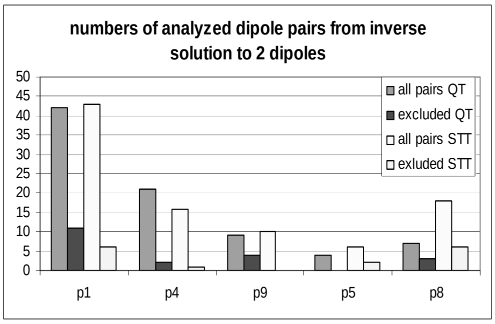


Fig.3. Properties of inverse solutions to 2 dipoles for the first group of patients. Left: The numbers of all pairs of dipoles used in the clustering method and the numbers of pairs that could not be divided uniquely to different clusters. Right: Mutual distance of inversely determined centers of clusters for 4 intervals of integration.

In the second group of patients the RMSDIF for DIMs computed from intervals defined by the end of U wave were greater than for DIMs computed using the end of T wave. Therefore we decided to consider the QT and STT intervals to be more reliable. Further evaluations were done on these 2 intervals.

The nonzero number of excluded pairs together with a large standard deviation of positions of dipoles in the clusters (Fig.4) indicated that these results didn't represent a local lesion but most probably large ischemic areas.



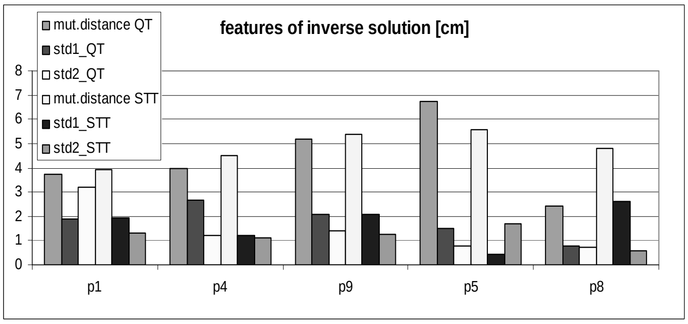


Fig.4. Properties of inverse solutions to 2 dipoles for the second group of patients for QRST and STT intervals of integration. Left: The numbers of all pairs of dipoles used in the clustering method and the numbers of pairs that could not be divided uniquely to different clusters. Right: Mutual distance of inversely determined centers of clusters and standard deviations of positions of dipoles in each cluster.

4. Discussion and Conclusions

Although it is assumed that U wave represents late repolarization activities of myocytes, the presented study indicates that the substantial information on repolarization changes for our inverse method is included in the QRST or STT intervals. The results for 13 measured subjects showed that DIMs computed from QRST or STT interval of integration are in very good correlation (97%). The properties of inverse solutions to 2 dipoles for identification of local ischemic lesions computed either from QRST or from STT interval of integration are comparable and give similar results.

Acknowledgements

This work was supported by grants 2/7092/27 from the VEGA grant agency and APVV-51-059005 from the APVV agency in Slovakia and partly supported by research project No. NN518 328035 financed by the Polish Ministry of Science and Higher Education.

References

[1] Švehlíková J, Tyšler M, Turzová M, Hebláková E. Identification of local repolarization changes in the heart by an inverse solution with two dipoles. In: XXXVth International Congress on Electrocardiology, Abstracts, St. Petersburg, Russia, 2008, 109, ISBN 5-85574-257-5.

[2] Mirvis D. Electrocardiography. A Physiologic Approach; Mosby, 1993.

[3] Ritsema van Eck H J, Kors J A, van Herpen G. The U wave in the electrocardiogram: A solution for a 100-year-old riddle. *Cardiovascular Research,* 67:256-262, 2005.

[4] Bishop Ch M. Pattern Recognition and Machine Learning; Springer 2006, 424-428.

On Measuring the Absolute Ventricular Volumes for the Estimation of End Systolic Pressure-Volume Relation in Rabits

[1]I. Rocha, [2]M. T. Tarata, [1]L. Silva -Carvalho†

[1]Universidade de Lisboa, Faculdade de Medicina, Lisboa, Portugal
[2]University of Medicine and Pharmacy of Craiova, Romania
Email: isabelrocha@yahoo.com

Abstract. *The paper deals with a new approach to calibrate the measurement of absolute left intraventricular volumes, necessary to compute the elastance curve and End-Systolic Pressure-Volume Relation (ESPVR) on single cardiac cycles. The method is based on using the conductance catheter technique, while injecting known absolute volumes of blood in the isolated left ventricle and has been tested on rabbits. Thus the influence of the parallel conductance does not affect the volume measurement anymore.*

Keywords: End Systolic Pressure-Volume Relation, Elastance, Left Ventricle Volume, Conductance Catheter, Rabbit.

1. Introduction

To characterize left ventricular contractile status more methods and indexes have been developed (1-6). ESPVR is a valuable tool; it designates a maximal elastance (maximal pressure/volume ratio). A traditional way to estimate ESPVR is to alter the heart load and use multiple cardiac cycles (Figure 1). Inducing such a transitional status is time consuming and increases the risk of heart failure. It was assumed the ESPVR stay on a line (see red line in Figure 1, left). Whether this linear relation stands it is still debated, evidence suggesting a nonlinear evolution of ESPVR or, in other words, only local linearity.

Such computations have been largely based on arbitrary mathematical curve fitting of left ventricular pressure data measured during isovolumetric contraction and relaxation of an ejecting beat. Different studies showed the linearity is just an assumption [1], raising questions on the accuracy of the ESPVR thus computed. Attempts have been made to compute the ESPVR on single cardiac cycles using algorithms, not always convergent [7].

We proposed [7] an original algorithm which is always convergent and provides accurate computation of the elastance curve and ESPVR on single cardiac cycles.

One of the problems is to determine absolute ventricular volumes by using the conductance catheter technique, where an estimate of the conductivity of structures surrounding the ventricular blood pool on parallel conductance is required. Usually, the contribution of the parallel conductance ranges from 50 to 70% of the total volume signal [8, 9]. Thus it is important to accurately quantify this volume offset, or to accurately measure the volume. Parallel conductance volume is normally estimated by injection of hypertonic saline, which transiently changes the conductivity of the blood in the ventricle [10] with disadvantages: (i) inappropriate saline loading with multiple

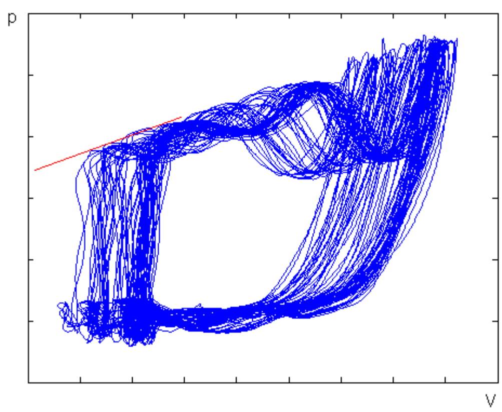


Figure 1. Cardiac cycles during a transition – an approximate ESPVR curve (relative units).

measurements, while being difficult to perform in small animals and (ii) the characteristics of the injectate itself may alter the calculated values [11] and hemodynamics [10]. Our work shows a technique of calibrating the measurement of absolute left intraventricular volumes, alleviating the above mentioned disadvantages of the mentioned methods.

2. Subjects and Methods

We performed preliminary experiments in 6 New Zealand rabbits (2.5 – 4.2 Kg) of either sex, anaesthetized with sodium pentobarbitone (40 mg/Kg, I.V.) supplemented as necessary with IV bolus injections. The descending aorta was cannulated retrogradely through the femoral artery with a Swann-Ganz catheter (Edwards size 4F, Baxter, UK) and the arterial blood pressure was measured by connecting the lumen of the catheter to a pressure transducer (SensoNor 840). The left carotid bifurcation was identified and the tip of a plethysmography catheter with an open central lumen was placed inside the left ventricle by retrograde cannulation of the common carotid artery to allow the simultaneous recording of left ventricular pressure (SensoNor 840) and volume (conductance plethysmography Dias Amado).

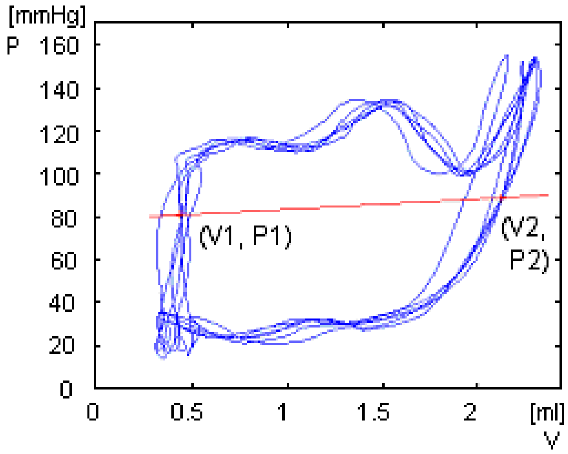


Figure 2. Two points of equal elastance are shown (red crosses) on one of the cardiac cycles, to illustrate the method.

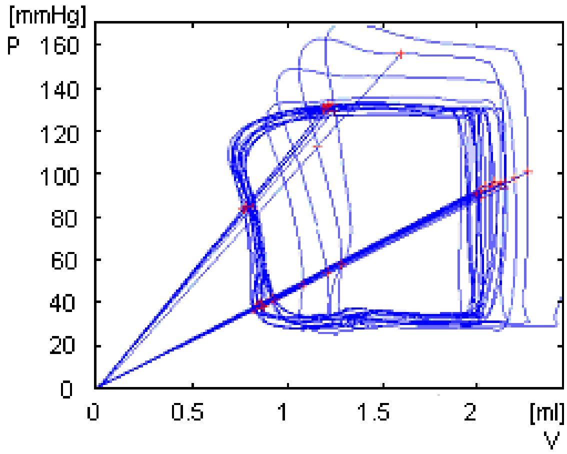


Figure 3. V0 is at the intersection of the equal elastance lines with the V axes (relative units).

The electrocardiogram (ECG) was recorded (Lectromed) from four electrodes applied to the limbs. Pressure, volume and the ECG were recorded. At the end of the experiment the animal was killed with an overdose of anaesthetic and the magnitude of infarction was confirmed in the isolated heart. Estimates derived from the normalized elastance curve are physiologically based and have the potential for noninvasive assessment. To compute the elastance (E) at any instance of the cardiac cycle, the current ventricular pressure (P) and volume (V) are needed together with a residual computational volume (V_0):

$$E(t) = \frac{P}{(V - V0)} \qquad (1)$$

To estimate V_0, our method [7] takes into consideration two conditions of equal elastance, one in the diastolic (V_1, P_1) and the other the systolic phase (V_2, P_2) of the same cardiac cycle (Figure 2). The algorithm computes an initial estimate of V_0 using these states of equal elastance:

$$V_0=P_2(V_2-V_1)/(P_2-P_1) \qquad (2)$$

Iteratively refined, V_0 (Figure 3) allows the computation of the elastance curve on all the current cardiac cycle. ESPVR is built by the points of maximal elastance and corresponds to the instances of end-systolic maximal pressure.

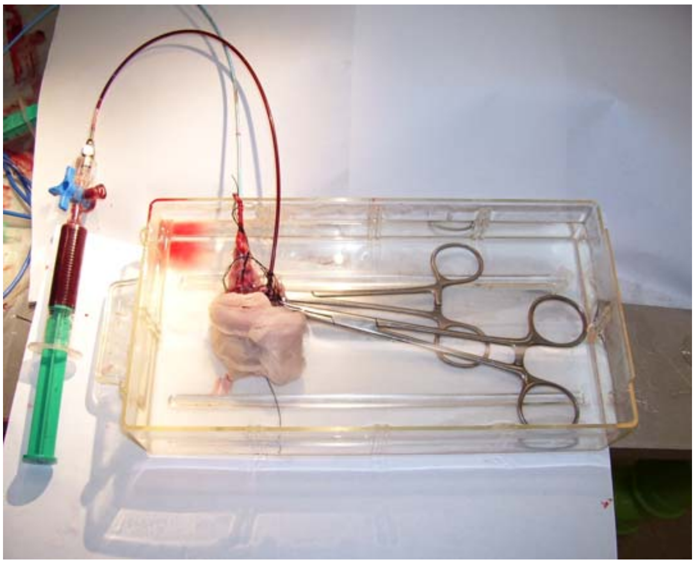


Figure 4. Isolated heart prepared for volume calibration

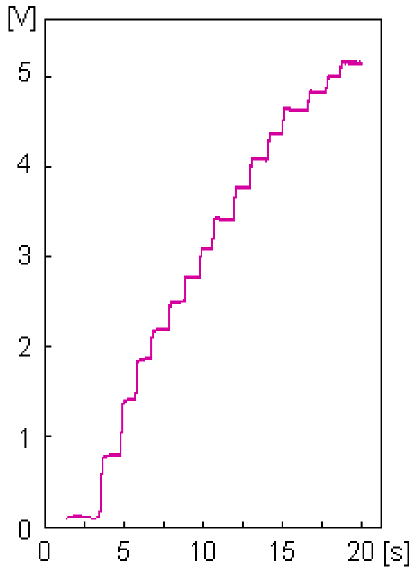


Figure 5. Calibration curve – steps of 0.2 ml, starting from 0.

The volume calibration is performed on the isolated heart, progressively injecting (Figure 4) in the left ventricle heparinated blood, in increments of 0.2 ml. Figure 5 shows the calibration curve. The resulting voltage, corresponding to known absolute volumes, generates a function through cubic interpolation, which is subsequently used to read actual real absolute volumes from the previously recorded data.

3. Results

Using this procedure and the method [6] to compute the cardiac elastance in single cardiac cycles, the elastance for an 'average' animal model (rabbit) could be computed, together with important cardiac parameters: Ejection Fraction (0.4221 ±.1238), End Diastolic Volume (1.4845e-6 ±4.1403e-7 m^3), End Systolic Volume (8.2268e-7 ±1.4823e-7 m^3), Emin (dP/dt) (1.1505e+10 ±4.1974e+9 Pa/ m^3), Emax (dP/dt) (1.1505e+10 ±4.1974e+9 Pa/m^3). The averages were taken on 2000 individual cardiac cycles, equally distributed on the animal population the experiments wee performed on.

4. Discussion and conclusions

To ignore V_0 and report the elastance as p/V would be inadequate as there are marked differences in V_0 in cases of infarction or dilated cardiomyopathy i.e. V_0 is sensitive to the pathologic condition of the heart. In situations of high contractility V_0 may have negative values. The similarity between V_0 computed on single cardiac cycles and V_0 computed on multiple cardiac cycles suggests that the evolution of elastance within the physiological loading range can locally predict linear ESPVR. This also supports the idea that nonlinear physiology rather than volume-catheter signal artifacts during Inferior Vena Cava obstruction explains negative values for V0. Thus, the accurate computation of the elastance, based on absolute ventricular volumes and accurate V0, is essential. The algorithm and the calibration method we developed, provide the automatic, dynamic computation of the elastance curve on single cardiac cycles and therefore of the ESPVR. The algorithm offers an adequate quantitative tool to refine physiological studies on the heart and circulatory system condition. The preliminary results are encouraging us to further refine the work.

References

[1] Noda T, Cheng CP, DeTombe PP, Little WC. Curvilinearity of LV end-systolic pressure-volume and dP/DTmax – end-diastolic volume relations. Am. *J. Physiology* 265:H910-H917, 1993.

[2] DeTombe PP, Jones S, Burkhoff D, Hunter WC, Kass DA. Ventricular stroke work and efficiency both remain nearly optimal despite altered vascular loading. *Am. J. Physiol*. 264 (Heart Circ. Physiol. 33) H1817-H1824, 1993.

[3] Asanoi H, Kameyama T, Ishizaka S, Nozawa T, Inoue H. Energetically Optimal Left Ventricular Pressure for the Failing Human Heart, *Circulation,* Vol.93, No.1, January 1 67-73, 1996

[4] Burkhoff D, DeTombe PP, Hunter WC, Kass DA. Contractile strength and mechanical efficiency of left ventricle are enhanced by physiological afterload, *Am. J. Physiol*. 260 (Heart Circ. Physiol. 29) H569-H578, 1991.

[5] Karunanithi MK, Feneley MP. Single-Beat Determination of Preload Recruitable Stroke Work Relationship: Derivation and Evaluation in Conscious Dogs, *Journal of the American College of Cardiology*, Vol.35, No.2:502-512, 2000

[6] Tarata M., Silva-Carvalho L., Rocha Isabel, Rosario L. A New Approach to Single-Beat Estimation of the Elastance Curve of the Cardiac Muscle, Proceedings of The 1st MEDINF Int. Conference on Medical Informatics & Eng. MEDINF' 2003, October 9 - 11, 2003, Craiova, Romania, Craiova Medicala Journal, ISSN 1454-6876, Vol. 5, Sup. 3, (2003): 218-221.

[7] Senzaki H, Chen C-H, Kass DA. Single_Beat Estimation of End_Systolic Pressure_Volume Relation in Humans – A New Method With the Potential for Noninvasive Application, *Circulation,* Vol.94, No.10, November 15:2497-2506, 1996.

[8] Steendijk P, Lardenoye JW, Van Der Velde ET, Sehalij MJ, Baan J. Evaluation of a new transcardiac conductance method for continuous on-line measurement of left ventricular volume. *Crit Care Med* 28: 1599–1606, 2000.

[9] White, P.A., Brookes, C.I.O., Ravn, H., Hjortdal, V., Chaturvedi, R.R., Redington, A.N. Validation and utility of novel volume reduction technique for determination of parallel conductance. *American Journal of Physiology* - Heart and Circulatory Physiology, Volume 280, Issue 1 49-1, H475-H482, 2001.

[10] Baan J, Van der Velde ET, Hein G, Debruin HG, Smeenk GJ, Koops J, Van Dijk AD, Temmerman D, Senden J, Buis B. Continuous measurements of left ventricular volume in animals and humans by conductance catheter. *Circulation* 70:812–823, 1984.

[11] Herrera MC, Olivera JM, and Valentinuzzi ME, Parallel conductance determination in cardiac volumetry using dilution manoeuvres: theoretical analysis and practical implications.*Med Biol Eng Comput* 37:169–174, 1999.

Atherosclerosis: Nonlinear Optical Measurements Related to Disease Progression

[1]M.G. Sowa, [1]L.B. Mostaco-Guidolin, [1]M.S.D. Smith, [1]E.K. Kohlenberg, [2]M. Shiomi, [3]A. Ridsdale, [3]A. Stolow, [1]A.C.T. Ko

[1]Institute for Biodiagnostics, National Research Council Canada, Winnipeg, Canada, R3B 1Y6
[2]Experimental Animals, Kobe University School of Medicine, Japan
[3]Steacie Institute for Molecular Sciences, National Research Council Canada, Ottawa, Canada, K1A 0R6
Email: mike.sowa@nrc-cnrc.gc.ca

Abstract. *Nonlinear optical (NLO) microscopy is used to follow key structural and biochemical changes associated with the progression of atherosclerosis. Arteries from WHHL-MI rabbits are examined using a 3 channel NLO microscope that can simultaneously monitor the coherent anti-stokes Raman scattered light (CARS), the two-photon excited fluorescence (TPEF) and the second harmonic generation (SHG) from a sample. Distinct differences in the nonlinear optical signals are observed that correlate with the age of the vessel and the presence of atherosclerotic plaque. These differences are attributed to the changing extracellular matrix and the increased lipid deposition associated with plaque development. The capability of NLO to perform 3D sectioning in thick highly scattering vessels in order to visualize structural details of the artery wall and highlight vessel pathology is demonstrated. These features make NLO a potentially valuable tool to help understand the progression of atherosclerosis.*

Keywords: Nonlinear optical microscopy, coherent anti-Stokes Raman scattering, two photon excited fluorescence, second harmonic generation, atherosclerosis

1. Introduction

Atherosclerosis once thought of as a passive disease characterized by the progressive accumulation of fatty deposits in arteries is now recognized as a complex disease where plaque development is mediated by a range of factors.[1] Early consequences of the disease include a thickening of the arterial wall and the appearance of fatty streaks on the vessel. In more advanced disease, vascular smooth muscle cells and a fiborous collagen-elastin network form to encapsulate lipid pools. Superimposed on these pathological changes is the normal aging process that affects vessels. For example, vessels become progressively stiffer with age as a result of changes to the elastin – collagen network within the vessel wall. NLO where elastin, collagen and lipid structures can be simultaneously imaged through the monitoring of two-photon excited fluorescence (TPEF), second harmonic generation (SHG) and coherent antistrokes Raman scattering (CARS) is ideally suited to following these changes in the wall of the artery.[2,3] The technique requires no stains to be added to the tissue and the imaging can be perform in an epi-configuration, obviating the need to prepare the tissue in thin sections. These additional features makes NLO microscopy are very convenient tool to use to study atherogenesis.

2. Subject and Methods

Nonlinear Optical Microscopy

As discussed in detail in Pegoraro et al, [4] a multiphoton laser scanning microscope system based on the schematic in Figure 1 was built to study bulk arterial tissue samples. A Ti:Sapphire oscillator (Spectra-Physics, Tsunami) centered at 800nm with a pulse width of 100fs, average output power of 1W when pumped with 7.25 W green laser at 532nm (Spectra-Physics, Millennium Pro) was used as the laser source. The output femto-second laser pulses were first passed through a Faraday isolator (Newport), re-compressed using a pair of GTI laser mirrors (Layertec GmbH, Germany) and then split into reflected pulses (pump) and the transmitted pulses (Stokes), using a 50/50 beam splitter. An objective lens focussed the transmitted pulses into a photonic crystal fiber (FemtoWhite 800, Crystal Fibre, Denmark) to generate broadband emission where the NIR portion of the broadband emission was retained as the Stokes pulses for generating the CARS signal. The pump and Stokes pulses were combined and sent into the microscope assembly where three non-descanned modular type PMT detectors (Hamamatsu) were used for the simultaneous detection of TPEF, SHG and CARS signals in the epi-direction. The laser pulses were focused onto sample through a 20x, 0.75 NA infinity corrected air objective lens (Olympus) with the emitted TPEF and epi-SHG/CARS signals being collected through the same objective lens. ScanImage (Cold Spring Harbour, NY) was used for laser scanning control and image acquisition and ImageJ ver 1.42b was used for post image processing and image viewing.

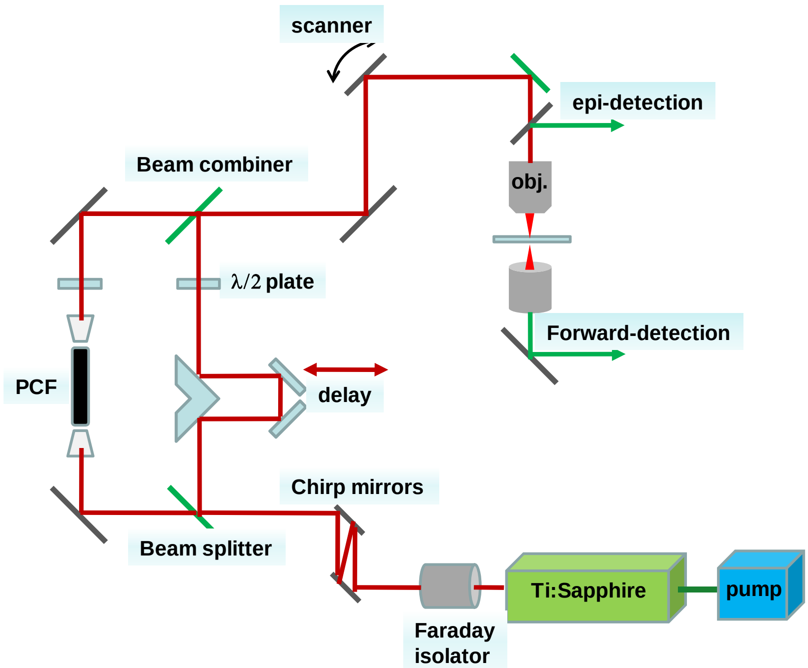


Fig. 1. Schematic of the nonlinear optical microscope with the Stokes pulses being generated in PCF which is pumped by the same Ti:sapphire laser that provides the pump pulse.

Arterial tissue samples and sample preparation

All animal experiments conformed to the guidelines set out by the Canadian Council on Animal Care regarding the care and use of experimental animals and were approved by the local Animal Care Committee of the National Research Council of Canada. The myocardial infarction prone Watanabe heritable hyperlipidemic (WHHL-MI) rabbits spontaneously develop atherosclerotic plaques due to a hereditary defect in LDL processing.[5,6] This model was used to study plaque development where the aorta was dissected from the ascending aorta to the external iliac artery and then rinsed in heparinized saline. The exterior aorta was subdivided into ~20-30 mm sections resulting in 5-7 pieces that were cut open longitudinally exposing the luminal surface. The samples were placed in petri dishes with the luminal surface facing up on a moist surface and hydration was maintained throughout the measurements by applying PBS solution periodically. Digital photos of the luminal surface were acquired and regions of interest were identified prior to nonlinear optical measurements.

3. Results & Discussion

Representative NLO measurements from the luminal surface of a healthy aorta are presented in Figure 2 and compared to images in Figure 3 taken from a region of older aorta with atherosclerotic plaque.

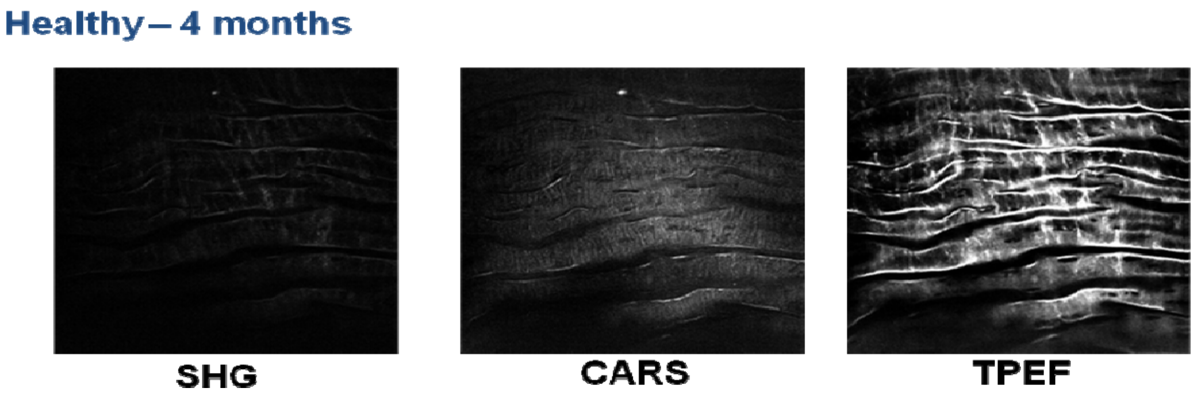


Fig. 2. Second harmonic generation (SHG), coherent antistokes Raman scattered (CARS) and two photon excited fluorescence (TPEF) images of healthy vessel.

Distinct differences can be observed in the TPEF, SHG and CARS images of young and old vessels. The TPEF images of the surface of young, healthy vessels show an elastin membrane structure with relatively weak SHG and CARS signal indicating a limited presence of type I collagen and lipids at the surface of the vessel. With age, the texture of the TPEF image changes revealing a dense interconnected elastin network but also with much high SHG indicating a greater contribution from collagen. In addition, from regions with plaque an intense CARS lipid signal is observed. Simple parameterizations of these images in terms of texture and intensity can be used to quantify these differences.

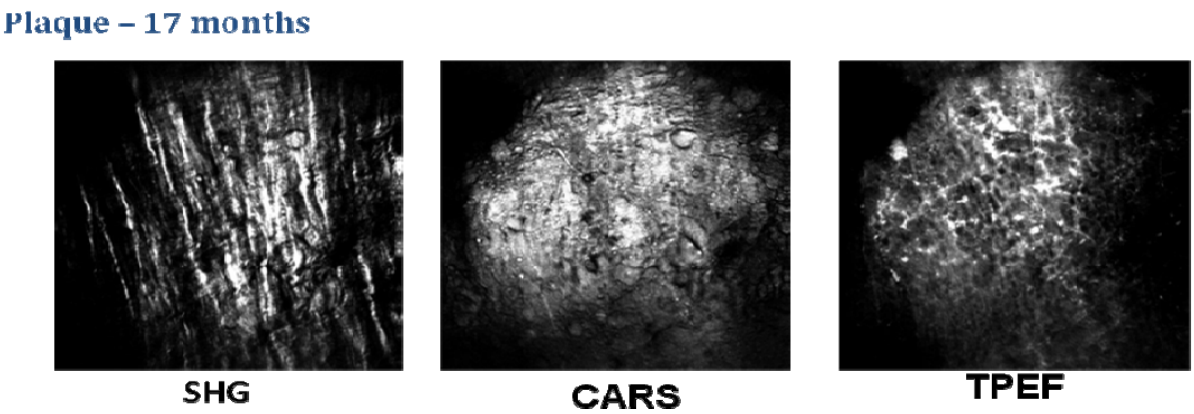


Fig. 3. Second harmonic generation (SHG), coherent antistokes Raman scattered (CARS) and two photon excited fluorescence (TPEF) images of older vessel with plaque.

4. Conclusions

Nonlinear optical imaging of the two-photon excited fluorescence and second harmonic generation in arterial tissue is demonstrated to be a useful tool for understanding the elastin-collagen network of vessels. The organization of the network is shown to change with age and with the advancement of atherosclerotic disease. In addition, coherent antistrokes Raman scattering can be used to follow lipid deposition in the wall of the artery.

Acknowledgements

The authors acknowledge Vijay Iyer of Cold Spring Harbour Laboratory (NY,USA) for assistance with the ScanImage software. This work was supported by National Research Council Canada, Genomics and Health Initiative.

References

[1] Libby P. Atherosclerosis: Disease Biology Affecting the Coronary Vasculature. *Am. J. Cardiol 98[suppl]*: 3Q-9Q (2006).

[2] Le TT, Langohr IM, Locker MJ, Sturek M, Cheng J-X. Label-free molecular imaging of atherosclerotic lesions using multimodal nonlinear optical microscopy. *J. Biomed. Opt.* 12(5): 054007 (2007).

[3] Ko ACT, Ridsdale A, Pegoraro AF, Smith MSD, Mostaço-Guidolin LB, Hewko MD, Kohlenberg EM, Schattka BJ, Shiomi M, Stolow A, Sowa MG. Label-free imaging of arterial tissues using photonic crystal fiber (PCF) based nonlinear optical microscopic system. In: *Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues VII.* Proc. SPIE Vol. 7182: 718204 *(2009).*

[4] Pegoraro AF, Ridsdale A, Moffatt DJ, Jia Y, Pezacki JP, Stolow A. Optimally chirped multimodal CARS microscopy based on a single Ti:sapphire oscillator. *Optics Express* 17(4): 2984-96 (2009).

[5] Shiomi M, Ito T, Yamada S, Kawashima S, Fan J. Correlation of Vulnerable Coronary Plaques to Sudden Cardiac Events. Lessions from a Myocardial Infarction –prone Animal Model (the WHHLMI Rabbit). *J. Atheroscler Thromb*, 11: 184-9 (2004)

[6] Shiomi M, Ito T, Yamada S, Kawashima S, Fan J. Development of an Animal Model for Spontaneous Myocardial Infarction (WHHLMI Rabbit). *Arteriosclerosis, Thrombosis, and Vascular Biology*. 23: 1239 (2003).

Robust Classification of Endocardial Electrograms Fractionation in Human using Nearest Mean Classifier

V. Křemen, L. Lhotská

Department of Cybernetics, Faculty of Electrical Engineering,
Czech Technical University, Prague, Czech Republic
Email: xkremen@fel.cvut.cz

Abstract. *Complex fractionated atrial electrograms (CFAEs) may represent the electrophysiological substrate for atrial fibrillation (AF). Progress in signal processing algorithms to identify CFAEs sites is crucial for the development of AF ablation strategies. A novel algorithm for automated description of atrial electrograms (A-EGMs) fractionation based on wavelet transform and several statistical pattern recognition methods was proposed and new methodology of A-EGM processing was designed and tested in such a comprehensive form than ever before. The algorithms for signal processing, description and classification were developed and validated using a representative set of 1:5 s A-EGMs (n = 113) ranked by 3 independent experts into 4 classes of fractionation: 1 – organized atrial activity; 2 – mild; 3 – intermediate; 4 – high degree of fractionation. New feature extraction and classification algorithms used and tested here showed mean classification error over all classes ~ 5.9%, and classification error of highly fractionated A-EGMs of ~ 9%. These operator-independent and fully automatic algorithms for A-EGMs complexity description is the first usage of such novel approaches in A-EGMs processing and analysis and may be easily incorporated into mapping systems to facilitate CFAEs identification and help to guide AF substrate ablation.*

Keywords: Atrial fibrillation, catheter ablation, complex fractionated atrial electrograms, wavelet transform, classification.

1. Introduction

Significant progress has been achieved in the field of curative ablation of atrial fibrillation (AF) in recent years. While empirical isolation of pulmonary veins is usually an effective strategy in paroxysmal AF, targeting extrapulmonary substrate within left (right) atrium is often necessary in the case of persistent/permanent AF [1]. Both areas with high dominant frequency of atrial electrograms (A-EGMs) [2] and areas with complex fractionated atrial electrograms (CFAEs) [3] were shown to play a role in the maintenance of the arrhythmia. In order to identify those sites, great effort has been made to describe the patterns of activation in AF [4] and to quantify general characteristics of A-EGMs either in time- or frequency-domain [5, 6]. Recently, two software algorithms for time-domain analysis of CFAEs were implemented in commercially available mapping systems - CARTO (Biosense-Webster) and EnSite NavX system (St. Jude Medical). Both methods require initial setting of specific input parameters making them, at least to some extent, operator-dependent. We used algorithms for automatic classification (pattern recognition), based on description of signal, using features extracted from recorded and preprocessed signals. This approach is based on the idea that there are signal complexes [7] in every A-EGM signal, which are related to electrical activation of electropathological substrate during AF. These signal complexes – fractionated segments of A-EGM (FSs) can be found automatically and then used for several features extraction (degrees of freedom of the signal), which could be used for automatic evaluation of electrogram complexity (or level of fractionation) in next stages. In this paper we focus on evaluation of A-EGM signal complexity of A-EGMs recorded during AF. In this paper we bring the results of a novel robust method for A-EGM processing, based on the wavelet transform signal analysis, several feature extraction followed by classifier.

2. Subject and Methods

Experimental dataset of A-EGMs

Atrial bipolar electrograms were collected during left-atrial endocardial mapping using 4-mm irrigated-tip ablation catheter (NaviStar, Biosense-Webster) in 12 patients (9 males, aged 56±8 years) with persistent AF. The A-EGMs acquired before the ablation procedure were band-pass filtered (30-400Hz) and sampled at frequency of 977Hz by CardioLab 7000 (Prucka Inc.). Discontinuous recordings from distal catheter bipole during left-atrial mapping outside the pulmonary veins and their tubular ostia (in order to treat only the signals from sites that are usually targeted during extrapulmonary substrate modification) were exported in digital format and reviewed by independent expert. The fragments with inadequate endocardial contact, relatively high signal-to-noise ratio or artifacts were excluded. Remaining parts of recordings were split into not necessarily contiguous 1500 ms segments with stable signal pattern. The A-EGMs very close to the mitral annulus were discarded to prevent the interference of relatively sharp ventricular signals with atrial signal analysis. This yielded approximately 250 segments with high-quality A-EGMs. This set of A-EGMs was further scrutinized. Finally, selection of 113 such segments represented wide spectrum of A-EGMs including those very organized, extremely fractionated, and all intermediate forms.

Expert classification of A-EGMs

Although the degree of fractionation of the A-EGM signals in the experimental dataset was a continuous variable by nature, expert classification into categories was chosen for the purpose of our study. Three experts, who perform AF ablation procedures on regular basis, independently ranked raw A-EGMs into those 4 classes of fractionation (1 - organized activity; 2 - mild degree of fractionation; 3 - intermediate degree of fractionation; 4 - high degree of fractionation) according to the subjective perception of signals. This procedure was facilitated by the purpose-written software for displaying A-EGMs in the same aspect ratio as on real-time screen during the left atrium mapping. This software also allowed experts scrolling through all A-EGMs with the possibility to reorder them repetitively according to assigned classification until the final ranking were reached. No specific criteria for signal assessment (e.g. dominant frequency or percentage of continuous electrical activity) were given. The experts were asked to classify the A-EGMs by their subjective judgment according to how the ablation at particular site would be valuable for atrial debulking. They were only instructed to keep approximately equal percent occurrence for each A-EGM category.

A-EGMs processing, feature extraction and classification

A-EGM preprocessing algorithm described in [8] was used to filter and prepare signals for the phase of feature extraction. The algorithms for A-EGM feature extraction [9] were used to describe A-EGM complexity in a new way. Based on the Automated Fractionated segments Search (AFS) preprocessing algorithm [8], the algorithms automatically search for areas of the A-EGM signal, where local electrical activity is found (FSs), also described by Faes et al. as local activation waves (LAWs) [6]. Several features of A-EGM are then defined based on FSs description. Following features are therefore derived from the characteristics of the automatically observed FSs or LAWs:

1) A number of fractionated segments found by AFS in particular A-EGM signal in dataset.

2) Minimum of Inflection Points in found FSs in a particular A-EGM signal.
3) Maximum of Inflection Points in found FSs in a particular A-EGM signal. Arithmetical Mean value of Inflection Points added together in automatically found FSs.
4) Sum of width of all FSs found in particular A-EGM.
5) Minimal width of found FSs in particular A-EGM signal.
6) Maximal width of found FSs in particular A-EGM signal.
7) Arithmetical Mean value of inter-segment distance of automatically found FSs.

To determine the quality of the feature selection and to evaluate the automatic classification of the individual level of fractionation of A-EGM, we used Nearest mean classifier, where classes are represented by their mean (prototype) and which associate new sample with the class of nearest mean. Training data were used in 1:1 fashion (training:testing).

3. Results

The mean classification error across all classes obtained for the best performing Nearest mean classifier was achieved at ~ 5.9 %. The classification errors of this classifier for individual classes are shown in Table 1. This classifier was able to clearly discriminate classes 1 and 2 with 0% error, while classes 3 and 4 were more difficult to approach.

Table 1. Unweighted classification error for Nearest mean classifier using described 7 features.

Classes of fractionation	1	2	3	4
Error [%]	0	0	14.3	9.1

4. Conclusions

While the current methods are focused on dominant frequency classification or evaluation of electrogram fractionation, the above described method primarily eliminates segments of electrograms, where evidently no local electric activity is present and then future techniques will extract more features from the signal and describe complexity of electrogram based on found electrogram segments. In conclusion, we proposed a novel robust algorithm based on wavelet transform for automated and operator-independent assessment of A-EGMs fractionation to facilitate CFAEs identification and to guide AF substrate ablation. Because of the low computational costs it can be easily incorporated into real-time mapping systems provided it will be first validated off-line in larger and independent A-EGMs sample and compared with currently available algorithms. By now, its clinical value is unknown and warrants further investigation.

Acknowledgements

This work was supported by the research projects #MSM 6840770012 "Transdisciplinary Biomedical Engineering Research II from Ministry of Education, Youth and Sports of the Czech Republic.

References

[1] Haïssaguerre M, Wright M, Hocini M and Jaïs P. The substrate maintaining persistent atrial fibrillation *Circ. Arrhythmia Electrophysiol.* **1,** 2-5, 2008.

[2] Sanders P, Berenfeld O, Hocini M, Jais P, Vaidyanathan R, Hsu L F, Garrigue S, Takahashi Y, Rotter M, Sacher F, Scavee C, Ploutz-Synder R, Jalife J and Haissaguerre M. Spectral analysis identifies sites of high-frequency activity maintaining atrial fibrillation in humans *Circulation* **112,** 789–797, 2005.

[3] Nademanee K, McKenzie J, Kosar E, Schwab M, Sunsaneewitayakul B, Vasavakul T, Khunnawat C and Ngarmukos T. A new approach for catheter ablation of atrial fibrillation: mapping of the electrophysiologic substrate *J. Am. Coll. Cardiol.* **43,** 2044–2053, 2004.

[4] Kumagai K. Patterns of activation in human atrial fibrillation *Heart Rhythm* **4,** (3)Suppl. S7–S12, 2007.

[5] Mainardi L T, Corino V D, Lombardi L, Tondo C, Mantica M, Lombardi F and Cerutti S. Assessment of the dynamics of atrial signals and local atrial period series during atrial fibrillation: effects of isoproterenol administration *Biomed. Eng. Online* **3,** 37, 2004.

[6] Ravelli F, Faes L, Sandrini L, Gaita F, Antolini R, Scaglione M and Nollo G. Wave similarity mapping shows the spatiotemporal distribution of fibrillatory wave complexity in the human right atrium during paroxysmal and chronic atrial fibrillation *J. Cardiovasc. Electrophysiol.* **16,** 1071–1076, 2005.

[7] V. Křemen and L. Lhotská. Novel approach to search for individual signal complexes in Complex Fractionated Atrial Electrograms using wavelet transform. *In 6th International Special Topic Conference on ITAB - Proceedings*, pages 83-86, IEEE, Piscataway, 2007.

[8] Křemen, V., Lhotská, L., Macaš, M., Čihák, R., Kautzner, J., et al. A New Approach to Automated Assessment of Fractionation of Endocardial Electrograms During Atrial Fibrillation. *Physiological Measurement,* vol. 29, no. 12, p. 1371-1381. ISSN 0967-3334, 2008.

[9] Křemen, V. Automated Assessment of Endocardial Electrograms Fractionation in Human. PhD Thesis. *Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague*. 2008.

Effect of Air-Ion Concentration on Microflora in Living Spaces

[1]Z. Szabó, [2]K. Bartusek,

[1]Department of Theoretical and Experimental Engineering,
Faculty of Electrical Engineering and Communication, Brno University of Technology,
Kolejní 2906/4,612 00 Brno, Czech Republic
[2]Institute of Scientific Instruments, Academy of Sciences of the Czech Republic, v.v.i.
Kralovopolska 147, 612 00 Brno, Czech Republic
Email: bar@isibrno.cz

Abstract. *Using the aeroscopic method and aspiration capacitor the effect of the density of ion concentration on the number of bacterial colonies in selected spaces was examined. The experiment was carried out in a Faraday cage and in a laboratory of the Department of Theoretical and Experimental Electrical Engineering (DTEEE). During the measurement the current value of the density of air ions was recorded. In the paper the methodology is described for the measurement of ion fields and the results are given of the change in the number of bacterial colonies due to short-term ionization of a closed space. A comparison of the measurement results obtained for different ion concentrations indicates the possibility of reducing pathogenic bacteria and fungi by increasing the concentration of air ions in closed spaces.*

Keywords: aspiration capacitor, measurement of ion concentration.

1. Introduction

Air ions are extremely important for human health. Air ions affect the metabolic functions of cells in the lungs, the blood supply in the organism, and also the psychic functions of man. Through their electric charge the air ions also affect the earth’s atmosphere and the environment we live in. The metrology of air ions has been known for a long time. One of its methods is based on measurement using the aspiration capacitor. This method is applied in the measurement described in this paper. The aim of the study described is to verify the hypothesis of reducing the number of bacteria and fungi in the air of living spaces.

2. Measuring method

An aspiration capacitor with variable electric field and an electrometer for measuring small currents were used to measure the concentration of various kinds of ions (light-weight, light heavy-weight, heavy-weight) and ions of different polarities. The principle of the measuring method is obvious from Fig. 1; it has been taken over from [1] and [2]. A known amount of air under examination $M = 1.021 \,.\, 10^{-3}$ m/s (volume rate of flow) flows through a cylindrical capacitor, which has polarization voltage U ($U = 0 – 100$ V).

Electrostatic forces attract the air ions to the electrodes. The number of ions captured by the electrode create a small electric current I. The concentration of ions of one polarity is proportional to the magnitude of this current according to the relation

$$n = \frac{I}{M \cdot e} \qquad (1)$$

where $e = 1.6 \,.\, 10^{-19}$ C is the electron charge.

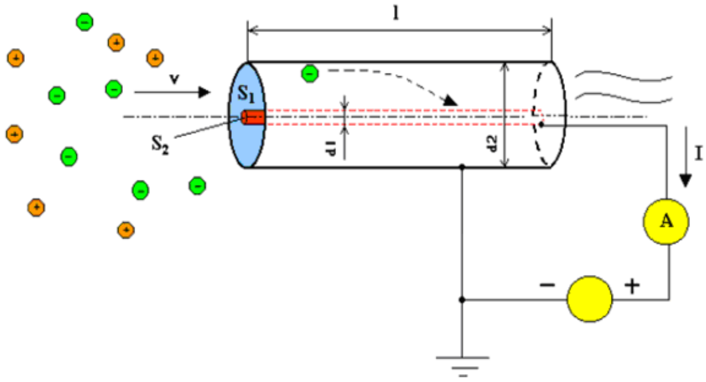


Fig. 1. Principle of measuring air ions, using aspiration capacitor.

The capacitor collector captures all ions whose mobility is lower than the minimum mobility given by the relation

$$k_m = \frac{M}{4 \cdot \pi \cdot C \cdot U} = \frac{M \cdot \ln \frac{d_2}{d_1}}{2 \cdot \pi \cdot l \cdot U}, \tag{2}$$

where M is the volume rate of flow of air through the capacitor, C is the capacitor capacitance, U is the voltage across the electrodes, and d_1, d_2 and l are the dimensions of the aspiration capacitor. The methodology of measuring and the calculations are described in more detail in [1, 2 and 3].

Aeroscopic measurement of air was carried out in cooperation with the Department of Preventive Medicine of Masaryk University in Brno. Samples of air were taken using the BIOTEST HYCON RCS Plus measuring instrument. Used as the culture medium was the strip for total capture of bacteria. The volume of air taken was 200 l and the results were converted to values per 1 m^3.

3. Experimental results

Two different spaces were chosen for the measurement. The first was a Faraday cage with a minimum magnitude of electric field. We assumed that the negative air ions would be concentrated in the centre of the cage and that the number of bacteria and fungi would be comparable with the laboratory space. In the course of measuring the quality of air by the aeroscopic method the density of ions in the space being measured was recorded during and after ionization. The space was ionized by a BIV-06 ionizer.

The other space was an infrequently used laboratory. This space approaches working and living spaces in which people will usually stay for quite some time. The measurement proceeded in the same way as in the case of Faraday cage.

Measurement in the Faraday cage

The BIV-06 ionizer was axially positioned 2.5 m from the aspiration capacitor opening. In the course of the measurement of ions, samples of air were taken simultaneously. The first two measurements were performed without ionization and the results were averaged. Another two measurements took place after 10 minutes' ionization and the results were again averaged. The results of measuring the ion fields are given in Fig. 2. On this figure the ion concentration increases after 60 and then it decays. This effect could have been caused by air turbulences what was started up after closing the Faraday cage door. With a natural ion concentration in the Faraday cage of ca. 700 ions/cm^3 (Fig. 2, left) an average of 198 bacterial colonies per 1000 l of the air being measured and 2 colonies of fungi were measured. The results show the evident effect of ionization on the number of bacteria in the space. After 10 minutes' ionization, the number of bacterial colonies dropped by 7.5%.

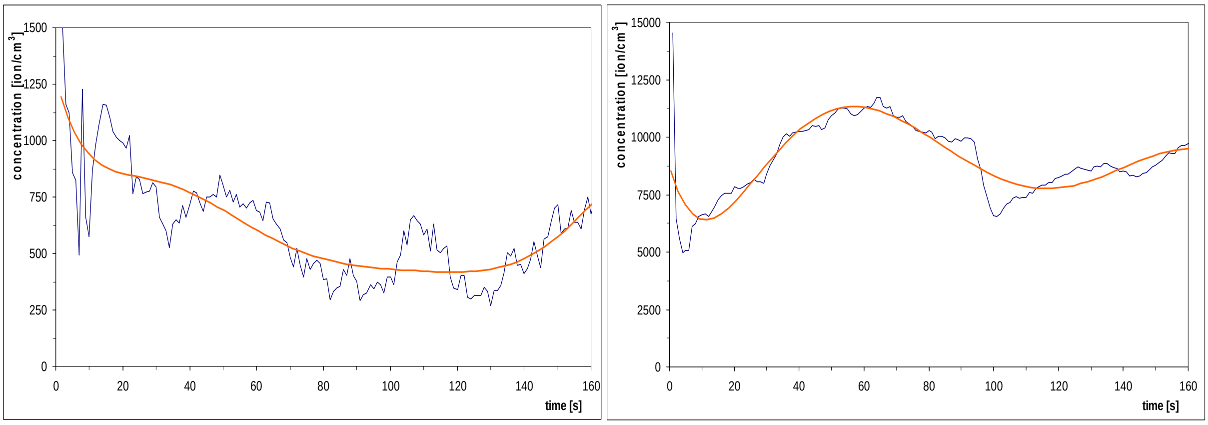


Fig. 2. Natural concentration of air ions in Faraday cage, left – before ionization, right – after ionization.

Measurement in the living space

The BIV-06 ionizer was used for the measurement in the living space. The ionizer was positioned in the axis of aspiration capacitor at a distance of 2.3 m from its opening. Prior to the ionization, 125 ions/cm^3 (Fig. 3), 300 bacterial colonies and 1 colony of fungi were measured in the space.

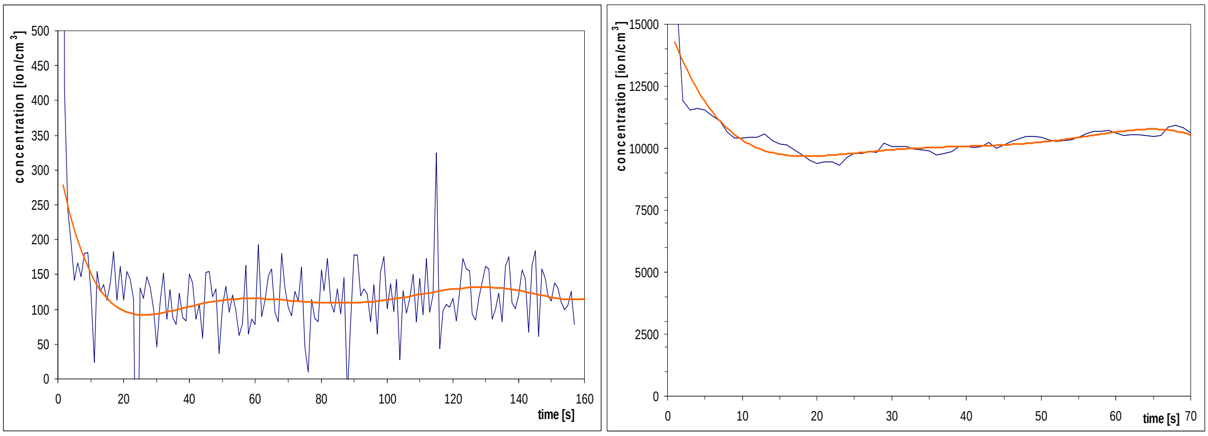


Fig. 3. Natural concentration of air ion in the laboratory, left – before ionization, right – after ionization.

After 25 minutes' ionization in the laboratory an average of 10 500 ions/cm^3 (Fig. 3), 125 bacterial colonies and no fungi were measured. The drop in the number of bacterial colonies was 58 %.

4. Conclusion

In cooperation with the Department of Preventive Medicine of Masaryk University in Brno the effect of the concentration of air ions on the number of bacteria was measured in two spaces. An electromagnetically shielded Faraday cage and the space of a DTEEE laboratory were chosen as spaces for the experiment. In the two spaces the natural concentration of air ions was measured using an aspiration capacitor while natural microflora was measured by the aeroscopic method. In the Faraday cage an average air ion density of 900 ions/cm^3 was measured while the value for the living space was 150 ions/cm^3. The high number of air ions in the Faraday cage corresponds with the reduced recombination of ions in a space with markedly reduced electric field. For the above density of air ions and average of 200 bacterial colonies per 1000 l of air were measured in the Faraday cage and an average of 275 bacterial colonies per 1000 l of air in the living space. After the measurement of natural concentration of air ions, the air in the Faraday cage was ionized by a BIV-06 ionizer for a period of 10 minutes. After the 10-minute ionization the average ion density measured for the cage was 9000 ions/cm^3 and the average number of bacterial colonies was 180 per 1000 l of air. The same measurement was carried out in the living space. After 25 minutes of ionization by the BIV-06 ionizer, 10 500 ions/cm^3 were measured while the number of bacterial colonies dropped to an average of 110 per 1000 l. It follows from the results measured that the density of air ions and the time of ionization affect the number of bacterial colonies in ionized spaces. With an initial ion density of 10 500 ions/cm^3 and an ionization duration of 25 minutes, a 60 % drop in bacterial colonies was established in the living space. In the Faraday cage after 10 minutes of ionization, 9000 ions/cm^3 and a 7.5 % drop in bacterial colonies were established. The higher drop in the number of bacterial colonies in the living space could have been caused by a longer time of ionization. The measurement results show that ionization reduces the number of bacterial colonies and fungi, which is very favourable to the health of people staying in such an environment.

Acknowledgements

The paper was prepared within the framework of BUT Brno of the Czech Republic and with the support of the research plan MSM 0021630516 and grant GAČR 102/09/0314.

References

[1] Bartušek K. Měření spektrálních charakteristik iontových polí. http://www.elektrorevue.cz/clanky/01038/index.html, 2001.

[2] Steinbauer M., Fiala P., Bartušek K., Szabó Z. Experiments of Accuracy Air Ion Field Measurement, Progress In Electromagnetics Research Symposium 2007, http://piers.mit.edu/piers/, 2007.

[3] Bartušek K. Měření s aspiračním kondenzátorem, http://www.elektrorevue.cz/

[4] Alpin K. L. Instrumentation for atmospheric ion measurements, PhD. Thesis, University of Reading, August 2000.

In-situ Response Measurement of Degenerated Articular Cartilage to the Loading in 3-T MRI

[1,2]V. Juráš, [1,2]P. Szomolányi, [3]S. Gäbler, [2]I. Frollo, [1]S. Trattnig

[1]MR Center, Highfield MR, Department of Radiology, Medical University of Vienna, Lazarettgasse 14, 1090 Vienna, Austria,
[2]Department of Imaging Methods, Institute of Measurement Science, Slovak Academy of Sciences, Dúbravska cesta 9, 84104 Bratislava, Slovakia,
[3]TU Wien, Institute of Materials Science and Technology, Vienna, Austria
Email: umervjur@savba.sk

Abstract. *The aim of this study was to assess the changes in MRI parameters during applied load directly in MR scanner and correlate these changes with biomechanical parameters of human articular cartilage. Cartilage explants from patients who underwent total knee replacement were examined in the micro-imaging system in 3-T scanner. Concerning MRI parameters, T_1 without- and T_1 with contrast agent as a marker of proteoglycan content, T_2 as a marker of collagen network anisotropy and ADC as a measure of diffusivity were calculated in pre- and during compression state. Subsequently, these parameters were compared to the biomechanical properties of articular cartilage, instantaneous modulus (I), equilibrium modulus (Eq) and time of tissue relaxation (τ). Significant load-induced changes of T_2 and ADC were recorded. High correlation between T_{1Gd} and I was found (r = 0.6324), and between ADC and Eq (r = -0.4884). Multi-parametric MRI may have great potential in analysing static and dynamic biomechanical behaviour of articular cartilage.*

Keywords: Articular cartilage, MRI, biomechanics, loading

1. Introduction

Only limited information is currently available on the relationship between cartilage compression and signal changes in T_1, T_2 and diffusion weighted images as a consequence of biochemical and biomechanical alterations during and after compression. However, in order to interpret medical images from MRI correctly, one should know several circumstances that precede measurement itself. Cartilage compression can be considered as one of these factors. Cartilage tissue consists of several macromolecules that provide basic functions: to distribute the load within a joint and provide a smooth surface for articulation. Load distribution in the joint is influenced by cartilage anatomy and mechanical properties, the presence of menisci and ligaments, bone stiffness and anatomy, and loading direction and kinetics. MR imaging has become the method of choice in the evaluation of normal [1] and damaged [2] cartilage due to the improved soft tissue contrast and multi planar capability without radiation exposure. Several studies were performed in the field of mechanical testing of cartilage during compression [3]. Biochemical and biomechanical changes were observed: fluid-flow and internal deformation, intrinsic viscoelasticity, changes in the water content or catabolism and loss of proteoglycans [4]. Investigators attempted to use MRI to evaluate mechanical properties of cartilage after applying load. In some cases, an MR-compatible device was built for controlled loading of cartilage explants and intact joints [5]. In another study, the rate and degree of deformation were increased after trypsin degradation [6]. A recent in vivo study demonstrated the ability to measure changes in cartilage volume as a function of mechanical stress (i.e. exercise) [7]. These studies imply that MRI is a modality with sufficient sensitivity to evaluate load-induced changes in cartilage tissue.

The goal of this study was to evaluate common MRI parameters of human articular cartilage and their changes as a consequence of static compression by a special designed non-magnetic device for indentation tests of cartilage tissue and correlate these with selected biomechanical parameters.

2. Subject and Methods

Cartilage samples were prepared from joints of 10 patients, who underwent a total knee joint replacement. The samples were cuboid-shape, with 10x10x6 mm in dimension. Study was performed on a Bruker 3T Medspec whole-body scanner (Bruker, Ettlingen, Germany) using BGA-12 micro-gradients (capable of delivering 200mT/m) with a special designed compression device built for this gradient system. Cartilage sample was compressed in the way that 15% of thickness decrease was accomplished. T_1 mapping was performed after filling the water-proof chamber of the compression device with a 2 mmol solution of gadopentetate dimeglumine (Gd-DTPA, Schering, Berlin, Germany) and soaking cartilage sample in solution for 24 hours. Then inversion recovery spin echo pulse sequence with TI times were 15, 30, 60, 160, 400 and 2000 ms. For T_2 mapping a multi-echo multi-slice spin echo sequence with TE times 15, 30, 45, 60, 75 and 90 ms was applied. ADCs were calculated from data from pulsed gradient spin echo (PGSE) with 6 different b-values (10.472, 220. 627, 452.8, 724.5 and 957.7). Each of parameters was calculated by fitting to appropriate exponential function. Fitting routines were written in IDL (Interactive Data Language, Research Systems, Inc.) using *mpcurvefit* routine. Regarding biomechanical parameters, *Eq* (equilibrium modulus), *I* (instantaneous modulus) and τ (time of tissue relaxation) were measured by indentation tests on a Zwick Z050 universal testing device with a 20N-load cell of 1mN resolution. OA status was determined by histogical evaluation using hematoxylin-eosine staining. Biomechanical and MR parameters were compared using Pearson correlation coefficient and statistical significance was determined by pair T-test (p-value lower than 0.05 was considered as statistical significant).

3. Results

Calculated MR parameters in pre- and during compression states are summarized in Table 1. The summary of Pearson correlation coefficients between MR and biomechanical parameters can be found in the Table 2. The example of MR appearance of articular cartilage sample is depicted in the Fig. 1.

Table 1. Values of T_1, T_2 and ADC, comparison of pre- and during-compression states

	ADC [10^{-3}mm^2/s]	T_1 [ms]	T_2 [ms]
number of pixels in ROI	166	110	108
without compression	**0.96 ± 0.40**	180 ± 30	**30 ± 9**
with compression	**0.85 ± 0.39**	230 ± 30	**27 ± 8**
change (%, p value)	**11% (<0.05)**	22% (0.238)	**10% (<0.05)**

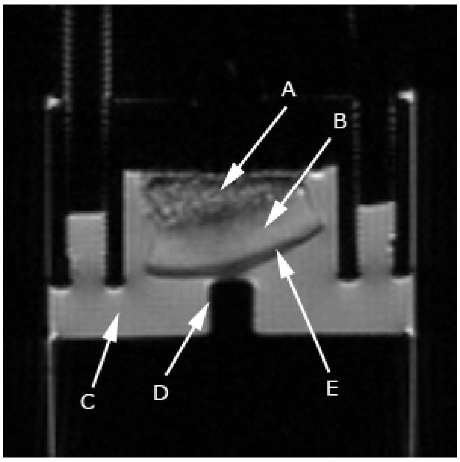


Fig. 1. Cartilage specimen in micro-imaging system (IR, TI = 60ms, TR = 4s, TE = 15ms): A - bone tissue, B – deep zone of cartilage, C - PBS + Gd-DTPA^{2-} solution, D - compressive piston, E - superficial zone of cartilage

Table 2. Pearson correlation coefficients between MR parameters during loading and biomechanical parameters measured by indentation tests

	Eq [MPa]	I [MPa]	I/Eq	T [s]
T1c/T1w	-0.1884	**0.6324**	**-0.5275**	**0.6092**
T2c/T1w	0.3620	0.1118	-0.0860	0.0257
ADCc/ADCw	**-0.4884**	0.1276	-0.2612	**0.5039**

The rows are expressed as a ratio of the MR parameters during compression (c-index) and without compression (w-index).

4. Discussion

It is generally known that the response of cartilage to compressive load is not uniform [8]. Therefore, the ability to spatially localize the response of cartilage to compression is necessary. While interpreting data in comparison with in-vivo conditions there are two important factors affecting resulting deformation: first one is the dynamics of loading and second one is the distribution of loading. Dynamics of loading strongly affects load displacement relationship. T_2 values decreased after compression in 10%. These observations agree well with published results and support the hypothesis that cartilage compression results in greater anisotropy of superficial collagen fibers, because region of interest was defined very close to piston head, i.e. mostly in superficial part of cartilage. Eckstein et al demonstrated a 5 to 6% decrease in patellar cartilage volume after compressive loading induced by performing 50 deep knee bends [9]. Besides tissue consolidation and the resulting decrease in water content due to efflux of water as a possible factor responsible for the observed change in cartilage T_2 prior ex vivo studies designed for the evaluation of changes in collagen fiber orientation with loaded conditions suggest that changes in fiber orientation is the dominant factor for T_2 shortening [10]. T_1 values of cartilage after penetration of Gd-DTPA^{2-} allow assessment of the proteoglycan (GAG) component of articular cartilage [11]. Applying compression to cartilage sample invokes changes in water content in cartilage tissue. Amount of proteoglycans stays unaltered, but amount of liquid part decreases. Due to squeezing liquid content from cartilage, concentration of Gd-DTPA^{2-} decreases as well. In contrast to T_{1Gd}, which reflects changes in chemical composition, diffusion constants may reflect microstructural degradation of the cartilage matrix. ADC correlates with structural changes rather than with changes in chemical composition [12]. Compression may induce cartilage matrix impairment which leads to decreasing of ADC. Since *Eq* and *I* reflect dynamic mechanics of cartilage, they correlate mostly with proteoglycans content (T_{1Gd}) and diffusivity of water molecules (ADC)

5. Conclusions

It was shown that multi-parametric MRI has a great potential in analysis the static and dynamic biomechanical behaviour of articular cartilage.

Acknowledgements

Austrian Science Fund (FWF) FWF-Project P-18110-B15 and Slovak Scientific Grant Agency VEGA 2/0142/08 is gratefully acknowledged.

References

[1] V. Mlynarik, and S. Trattnig. Physicochemical properties of normal articular cartilage and its MR appearance. *Investigative Radiology* 35 (2000) 589-594.

[2] D. Burstein, A. Bashir, and M.L. Gray, MRI techniques in early stages of cartilage disease. *Investigative Radiology* 35 (2000) 622-638.

[3] C. Herberhold, T. Stammberger, S. Faber, R. Putz, K.H. Englmeier, M. Reiser, and F. Eckstein, An MR-based technique for quantifying the deformation of articular cartilage during mechanical loading in an intact cadaver joint. *Magnetic Resonance in Medicine* 39 (1998) 843-850.

[4] R.L.Y. Sah, Y.J. Kim, J.Y.H. Doong, A.J. Grodzinsky, A.H.K. Plaas, and J.D. Sandy, Biosynthetic Response of Cartilage Explants to Dynamic Compression. *Journal of Orthopaedic Research* 7 (1989) 619-636.

[5] T. Stammberger, C. Herberhold, S. Faber, K.H. Englmeier, M. Reiser, and F. Eckstein, A method for quantifying time dependent changes in MR signal intensity of articular cartilage as a function of tissue deformation in intact joints. *Medical Engineering & Physics* 20 (1998) 741-749.

[6] J.H. Kaufman, R.R. Regatte, L. Bolinger, J.B. Kneeland, R. Reddy, and J.S. Leigh, A novel approach to observing articular cartilage deformation in vitro via magnetic resonance imaging. *Journal of Magnetic Resonance Imaging* 9 (1999) 653-662.

[7] F. Eckstein, M. Tieschky, S. Faber, K.H. Englmeier, and M. Reiser, Functional analysis of articular cartilage deformation, recovery, and fluid flow following dynamic exercise in vivo. *Anatomy and Embryology* 200 (1999) 419-424.

[8] M. Wong, and D.R. Carter, Articular cartilage functional histomorphology and mechanobiology: a research perspective. *Bone* 33 (2003) 1-13.

[9] F. Eckstein, M. Tieschky, S.C. Faber, M. Haubner, H. Kolem, K.H. Englmeier, and M. Reiser, Effect of physical exercise on cartilage volume and thickness in vivo: MR imaging study. *Radiology* 207 (1998) 243-248.

[10] R. Grunder, M. Kanowski, M. Wagner, and A. Werner, Visualization of pressure distribution within loaded joint cartilage by application of angle-sensitive NMR microscopy. *Magnetic Resonance in Medicine* 43 (2000) 884-891.

[11] A. Williams, A. Gillis, C. McKenzie, B. Po, L. Sharma, L. Micheli, B. McKeon, and D. Burstein, Glycosaminoglycan distribution in cartilage as determined by delayed gadolinium-enhanced MRI of cartilage (dGEMRIC): Potential clinical applications. *American Journal of Roentgenology* 182 (2004) 167-172.

[12] V. Mlynarik, I. Sulzbacher, M. Bittsansky, R. Fuiko, and S. Trattnig, Investigation of apparent diffusion constant as an indicator of early degenerative disease in articular cartilage. *Journal of Magnetic Resonance Imaging* 17 (2003) 440-444.

MRI Depiction of Blood-Spinal-Cord-Barrier Permeability after Spinal Cord Contusion Injury in Animal Model – in Vitro Study

[1,3]Z. Holúbeková , [2]N. Walder, [2]E. Sezemsky, [1,3] P. Szomolányi, [2]M. Brejnikow, [2]R. Schmidhammer, [1]I. Frollo, [3]S. Trattnig, [2]H. Redl

[1]Institute of Measurement Science, Slovak Academy of Sciences, Dúbravská cesta 9, 84104 Bratislava, Slovakia
[2]Ludwig Boltzmann Inst. for Exp. and Clin. Traumatology, AUVA Res. Center, Vienna, Austrian Cluster for Tissue Regeneration
[3]MR Center of Excellence, Dep. of Radiology, Med. Univ. Vienna, Vienna, Austrian Cluster for Tissue Regeneration

Email: zuzana.majdisova@gmail.com

Abstract. *The blood-spinal cord barrier (BSCB) controls the exchange of substances between the blood and the central nervous system. Spinal cord injury (SCI) causes a BSCB breakdown, which results in increased capillary-permeability for plasma molecules that increase neuronal damage. Contrast enhanced MRI is a technique to determine evolution and duration of BSCB-breakdown after standardized SCI and define the "therapeutic window" for this injury model.*

Keywords: MRI, blood-spinal cord barrier, "therapeutic window"after spinal cord injury

1. Introduction

The Blood-Spinal-Cord-Barrier represents a selective physiologic barrier that provides a stable microenviroment within the neuronal tissue. It was shown that blood–spinal cord barrier permeability can be measured using quantitative autoradiography following contusion injury to the rat spinal cord [1]. MR technique was used with advantage for high resolution in-vivo mouse models [2]. Here we used contrast enhanced MRI to determine evolution and duration of BSCB-breakdown after standardized SCI and define the "therapeutic window" for this injury model.

2. Subject and Methods

In male Sprague-Dawley rats, a laminectomy was carried out at TH11 and a contusion injury was inflicted, using the IH® Impactor with a force of 150kdyn. Rats were divided into groups with different observation times: 0h (n=8); 24h, 72h, 4d, 5d, 6d, 10d (n=5 each).

At the end of the observation time each rat received an i.v. injection of 0.8ml/kg gadopentetate dimeglumine (Magnevist®, Bayer HealthCare). For definition of MRI timepoint, three rats were euthanized 10 minutes (0h group) after contrast agent (CA) administration and dynamic MRI measurement of excised spines was performed every 2 hours up to 18 hours. All other rats were euthanized 1 hour after CA administration, for evaluation of BSCB permeability. Subsequently, imaging was performed on the 3T MEDSPEC whole body scanner (Bruker, Ettlingen, Germany) equipped with a 200 mT/m microimaging gradient system and a 35 mm inner diameter micro-imaging coil. T1 weighted multislice SE sequence with the setup: FOV = 6 cm, slice thickness = 1,5 mm, matrix dimensions = {256, 256}, TR = 850 ms, TE = 17,6 ms. In case of high resolution sagitally oriented measurements (as shown on Figure 1.), matrix dimensions = {512, 512}.

Figure 1. High resolution sagitally oriented MR image of a rat spinal cord – in vitro, after the spinal cord contusion injury.

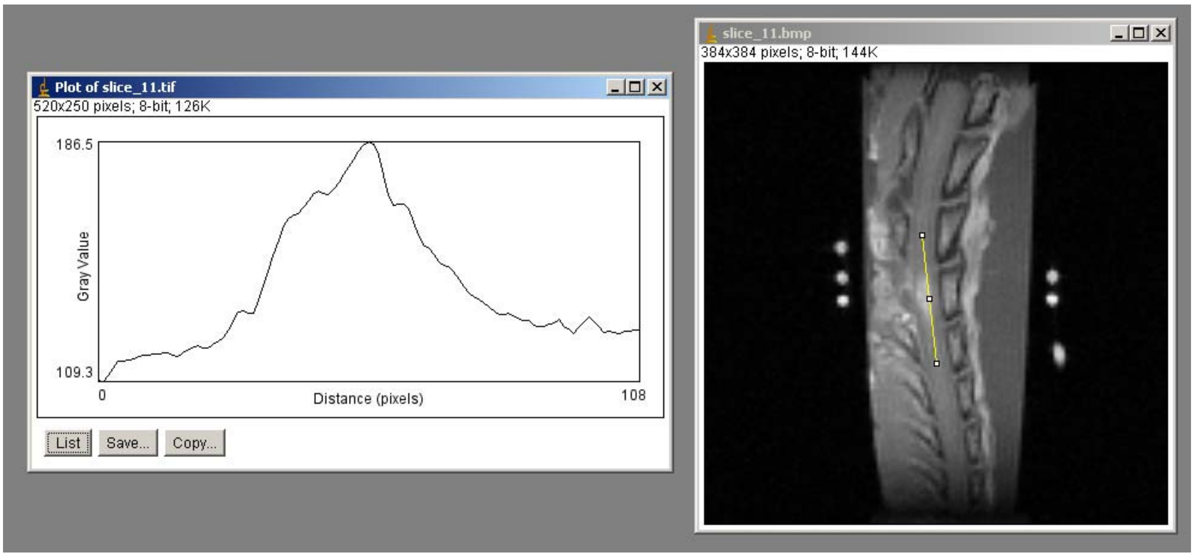


Figure 2: image intensity profile drawn along the spinal cord in the lesion site.

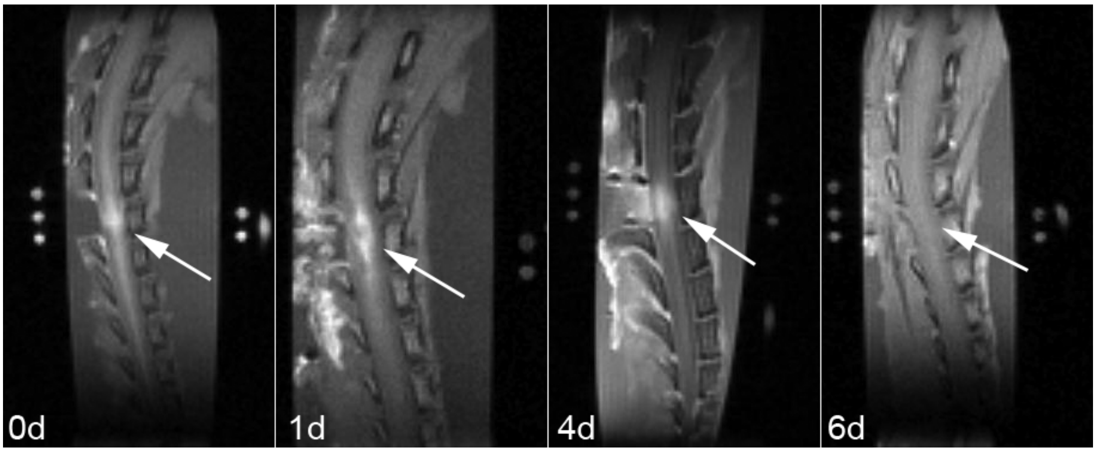


Figure 3. Series of time dependent MR images of the contrast enhancement rat spinal cord, after the injury – in vitro study 0, 1, 4 and 6 days after CSI

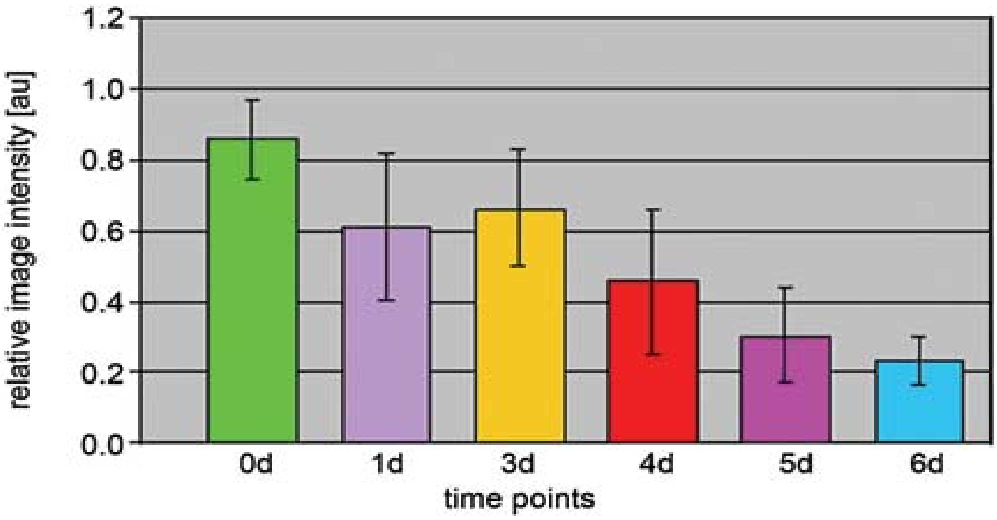


Figure 4: Extent of BSCB permeability up to 6 days

3. Results

Figure 1 shows high-resolution MR image of a rat spinal cord without CA application. Lesion site is apparent due to tissue damage on the right side. Without CA application lesion site has only limited contrast in comparison to the healthy tissue.

Figure 2 shows intensity profile drawn along the spinal cord. Line shows the profile trajectory. Line plot on figure 3 shows image intensity along the profile.

Definition of MRI timepoint

When euthanization and MRI were performed 10 minutes after CA application, signal increase was mainly detectable in the spinal cord arteries, the dynamic post-mortem measurement over 18h yielded increasing enhancement of the inury site over time. Adequate distribution of the CA within the neuronal tissue could be observed when euthanization and MRI were performed 1h after CA application.

Evaluation of BSCB permeability

Signal enhancement at the injury epicenter was measured after observation periods up to 6 days, gradually decreasing with time as well as with distance from the injury site. (Fig.3.) After 6 days or later, little or no signal enhancement was visible. Figure 4 shows decrease of the relative contrast intensity in image.

4. Discussion and conclusion

The BSCB controls the exchange of substances between the blood and the central nervous system. These barriers, formed by cells lining the blood vessels in the brain and the spinal cord, protect nerve cells by restricting entry of potentially harmful substances and cells of the immune system. Impairment in cellular machinery of the BSCB may lead to a barrier breakdown in many brain and spinal cord diseases or injuries.

Our study shows that the inflicted SCI distinctly increases BSCB permeability for 5 days. Delayed dispersion of the contrast agent within the neuronal tissue has to be considered. We show on the animal model, that BSCB is open for 5 days after the injury. These 5 days could be used for diagnostics and treatment as well. This is critical period, during which neuroregenerative medication can be applied and efficiently used. After this period, further treatment has only a limited efficiency, due to re-organized BSCB, which protects spinal cord from all kind of metabolites transported to the lesion site after intravenous treatment administration.

Acknowledgements

Funding for this study was provided by Austrian Science Fund (FWF) FWF-TRP-Projekt L243-B15 and FWF project: P18110 B15

References

[1] P.G. Popovich et al. Experimental Neurology 142: 258-275, 1996

[2] M. Bilgen et al. Magn. Reson. Imaging 20(4): 337-41, 2002

Detection of Variations in Biomedical Signals Based on Continuous Wavelet Transform Modulus Maxima

S. Karas, M. Teplan, M. Tysler

Institute of Measurement Science, Bratislava, Slovakia,
Slovak Academy of Sciences
Email: slavomir.karas@savba.sk

Abstract. *This paper presents an approach to identification of small variations in sampled signals based on continuous wavelet transform and nonlinear solution of the least-mean squares problem. Proposed technique is able to distinguish changes which are not visible by direct observation. The study was performed on signals from an ECG simulator recorded by the ProCardio 8 system. The method could be used in a beat-to-beat morphological analysis of ECG signals.*

Keywords: ECG, continuous wavelet transform, modulus maxima point, Gaussian wavelet

1. Introduction

The continuous wavelet transform (CWT) represents a tool for analysis of signals which can be the best described as aperiodic, noisy, intermittent, transient and so on [1]. CWT transform coefficients (TC) contain detailed information concerning the signal dynamics across the time-scale-amplitude (T-S-A) space. However this representation is redundant and not easy to interpret. Analysis of such TC can be focused on variations of frequency content with respect of time or on discontinuities, edges and sharp transitions. The latter are important in ECG signals that are quasi-periodic type of sequences with repeating fiducial points in P, QRS, T and U waves [2]. These points can be regarded in the time domain as singular points [3]. We analysed these singularities by measurement of the decay of TC computed by the CWT.

2. Methods

It was shown [3] that each sample in a discrete sampled signal can be considered as a point ν with some neighbourhood $[\nu - h, \nu + h]$ and the signal $f(t)$ can be approximated by the function:

$$p_\nu = \sum_{k=0}^{m-1} \frac{f^{(k)}(\nu)}{k!}(t-\nu)^k \tag{1}$$

with an approximation error $\varepsilon_\nu(t) = f(t) - p_\nu(t)$ (2). Signal $f(t)$ can be expressed as:

$$f(t) = p_\nu(t) + \varepsilon_\nu(t) \tag{3}$$

where the error of approximation is $|\varepsilon_\nu(t)| \le K|t-\nu|^\alpha$ (4), α is Lipschitz exponent and K is constant that is independent of ν. If exists the wavelet $\psi(t)$ with $n > \alpha$ vanishing moments $\int_{-\infty}^{\infty} t^k \psi(t)dt = 0$ (5) for $0 \le k < n$, for the wavelet transform (WT) is valid:

$$Wp_\nu(u,s) = \int_{-\infty}^{+\infty} p_\nu(t)\frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right)dt = 0 \tag{6}$$

With respect to Eq.(3), WT can be written as $Wf(u,s) = Wp_v(u,s) + W\varepsilon_v(u,s)$ (7) and thus: $Wf(u,s) = W\varepsilon_v(u,s)$ (8), where $Wf(u,s)$ are amplitudes of TC in u time and s scale. From Eq.(8) it is clear that the TC of WT express the error of approximation of function $f(t)$ by Eq.(1). It is also obvious that the amplitudes of TC are related with Lipschitz exponent. We used this dependence to measure the decay of amplitude of TC to characterize singularities that correspond to fiducial points in analyzed ECG signal. For this purpose we computed the modulus maxima points (MMPs) across the T-S-A space by using the equation:

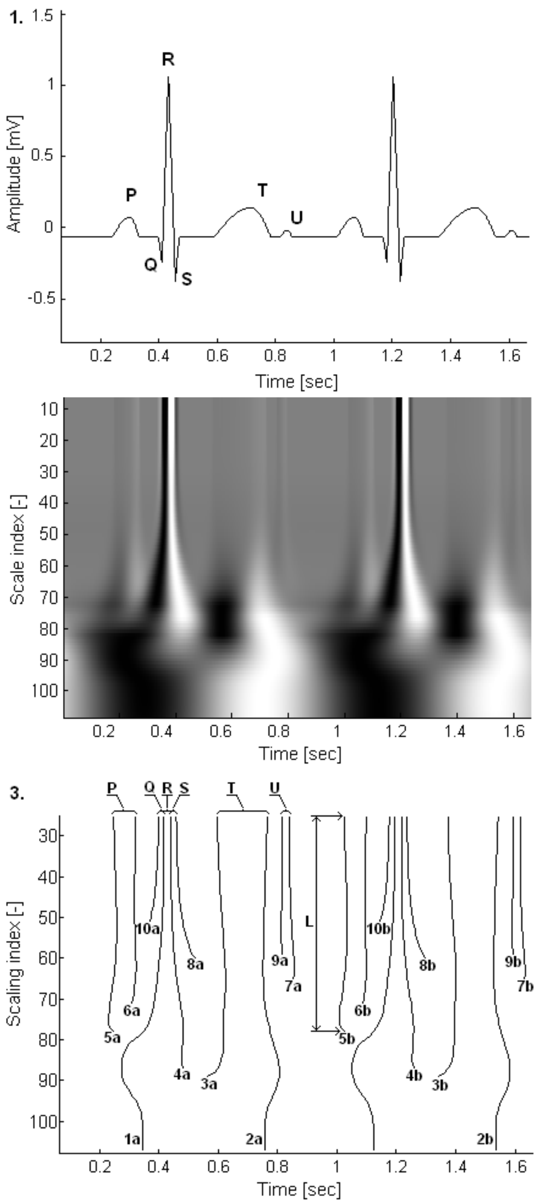


Fig. 1.1. Two similar cycles of recorded ECG signal;
Fig. 1.2. TC computed from signal depicted in Fig. 1.1;
Fig.1.3. MMLs (marked as 1a - 10a according to the length L in T-S-A space) that correspond to the fiducial points in ECG signal depicted in Fig. 1.1.

$$\frac{\partial W f(u_k, s_l)}{\partial u} = 0 \tag{9}$$

where u_k is k-th value of time and s_l is l-th value of scale. By connecting all MMPs across the scale s in the T-S-A space we created modulus maxima lines (MMLs). CWT of an ECG signal (Fig. 1.1) acquired by the ProCardio-8 system [4] was computed using the first derivation of Gaussian wavelet. We searched TC (Fig. 1.2) in T-S-A space and localized all points where Eq.(9) is valid. From these points we constructed the MMLs. Not only MMLs belonging to the fiducial points but also those belonging to the singularities that represent the noise, were created. MMLs that correspond to the fiducial points were separated by tresholding (Fig. 1.3). From Fig. 1.1 and Fig. 1.3 it is obvious that each MML corresponds to one specific fiducial point (1a to 10a represent the first and 1b to 10b the second activation sequence in the ECG signal). Each MML can be then analyzed in 3D space. This option gives the complete characterization of all fiducial points in the T-S-A space (Fig. 2.1). We analyzed each MML in a 2D projection, namely in the S-A (scale-amplitude) projection (Fig.2.2) and limited the analysis only to the exponential segment of the MML. The analysis was performed with an exponential model:

$$\mathbf{y} = A.e^{(\mathbf{x}-D).B} + C \tag{10}$$

where $\mathbf{x}$ and $\mathbf{y}$ are values for input segment of MML in S-A projection (Fig. 2.2) and A, B, C, D are approximation coefficients. We solved Eq.(10) for all MMLs associated to the fiducial points in selected segment of the ECG signal, and obtained a set of parameters that are related to behavior of these points in the T-S-A space. Parameters for MMLs in the first and in the second activation sequence of the ECG signal were computed.

3. Results and discussion

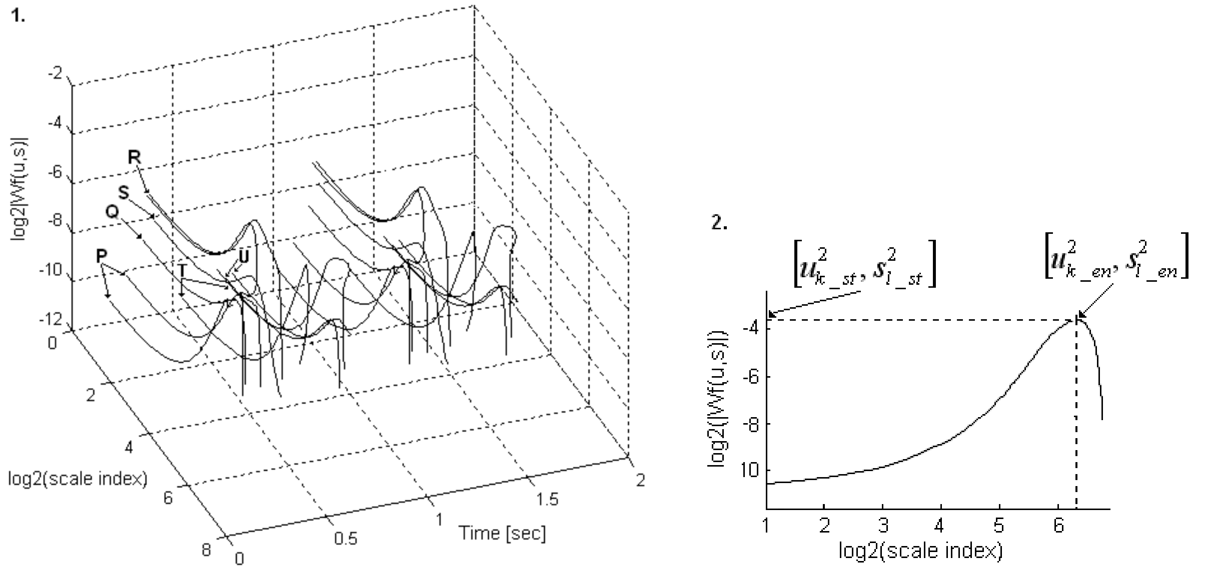


Fig. 2.1. MMLs of ECG signal in T-S-A space. Fig.2.2. MML of ECG signal pertaining to T wave in S-A projection (coordinates $u^2_{k\_st}$, $s^2_{l\_st}$ denote the start and coordinates $[u^2_{k\_en}, s^2_{l\_en}]$ denote the end point of analyzed exponential segment of the maxima line in S-A projection).

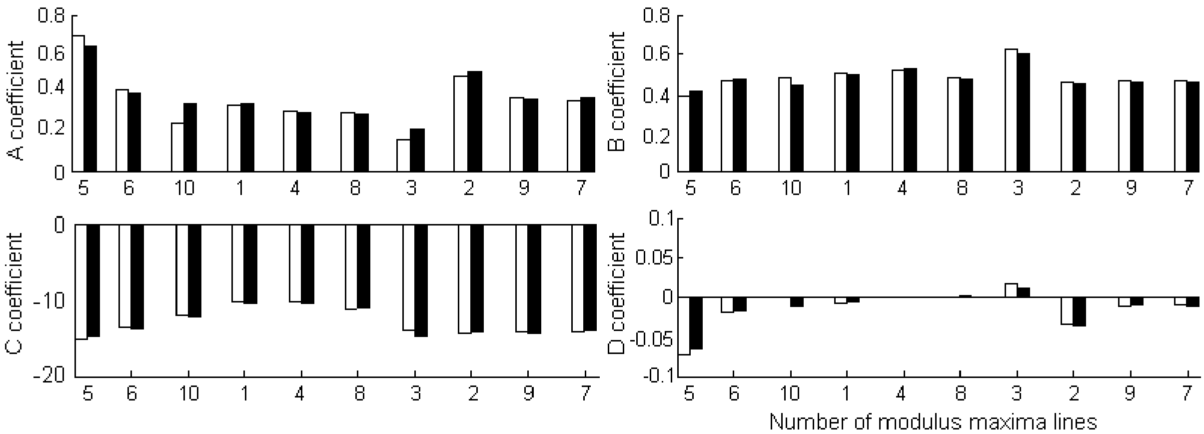


Fig. 3. Coefficients computed from MMLs (Fig. 1.3) for ECG signals depicted in Fig. 1.1 for each line number. Two sets of values of particular coefficient are depicted in each figure that corresponds to the specific MMLs (1a-10a for the first and 1b-10b for the second set depicted in Fig. 1.3). The first set is for coefficients that correspond to the first (white color) and the second set is for coefficients that correspond to the second activation sequence (black color) in the ECG signal.

While no obvious differences between signals were visible in the time domain, noticeable differences were detected in proposed coefficients. The maximal differences in coefficient values (Fig. 3) were localized at positions that correspond to the 5th, 10th and 3th MML. Computed values shown that the maximal difference occurred at 10th MML in all studied coefficients except of C. It corresponds to the onset of the Q wave in ECG signal (compare Fig. 1.1 with Fig. 1.3). In C the maximal difference occurred at 3th MML. The absolute maximal difference among all coefficients, which corresponds to 10th MML, was in coefficient A that refers to a scaling of exponent in the model. Absolute minimal difference that corresponds to that MML was detected in coefficient D that is related to the linear translation along x axis (this axis expresses log2 (scale index) and it is inversely proportional to the frequency [1]) as it is apparent from Eq.(10) and Fig. 2.2. It is also worth to note that absolute maximal difference from all coefficients (except C that represents the linear translation along y axis – this axis expresses log2 (|Wf (u, s)|)) was detected at position that correspond to the 10^{th} MML and represents the onset of the Q wave. Each coefficient represents different property of particular MML that corresponds to the specific fiducial point. Coefficient D determines frequency offset of each particular MML in the T-S-A space. Coefficient C expresses translation along the axis of amplitude of transform coefficients and determines the amplitude offset of each MML. Another important parameter is length of MML in the T-S-A space. This parameter determines a range of each MML along frequency axis. From Fig. 1.3 it is apparent that the sharpest features of ECG signal (such as R wave and T wave) have the longest corresponding MMLs. This means that frequency content of these features tends from the highest to the lowest frequencies and even very low frequencies are included unlike in the other features of the analysed ECG signal. We can speculate that specific change in activation sequence of ECG signal is related to corresponding change in some property of particular MML in the T-S-A space and thus with the change of coefficient values of the exponential model (with more or less sensitivity).

4. Conclusions

From the presented results we can conclude that even small changes in characteristic features of ECG signals are projected into coefficients of the proposed exponential model. We expect that the measure of these changes is different for each coefficient what means that different type of change in signal is coupled with change of different coefficient and thus it can be detected and evaluated. Each coefficient thus carries different information about the fiducial point in the analyzed signal and is connected with the corresponding singularity. A complete characterization of each activation sequence can be computed what offers new possibilities of the beat-to-beat morphological characterization of ECG signals. Future work will be focused on examination of applicability of the presented approach on real ECG signals [5].

Acknowledgements

This work was supported by project ESF 2004/4-016, APVV grant number 51-059005 and Vega grant number 2/7092/27.

References

[1] Paul S Addison. The illustrated wavelet transform handbook. 2002.

[2] Rangaraj M. Rangayyan. Biomedical Signal Analysis. IEEE Press. 2002.

[3] Stéphane Mallat. A wavelet tour of signal processing. Academic press. 1999.

[4] Rosík V., Karas S., Hebláková E. Tysler M.Portable Device for High Resolution ECG Mapping.Measurement Science Review, Vol.7, Sec.2, 4, 57-61, 2007. ISSN 1335 – 8871.

[5] Karas S., et. al. Software for Analysis Of Biosignals Measured From Isolated Perfused Animal Hearts In Langendorff Setup.Biosignal 2008. Brno. ISSN 1211- 412X.

Compensation of Heart Rate Variation during Noninvasive Identification of Repolarization Changes

[1]M. Turzová, [1]M. Tyšler, [1]E. Hebláková, [1]J. Švehlíková, [2]M. Kania

[1]Institute of Measurement Science SAS, Bratislava, Slovakia
[2]Institute of Biocybernetics and Biomedical Engineering PAS, Warsaw
Email: umerturz@savba.sk

Abstract. *Local repolarization changes in the myocardium can be noninvasively identified by an inverse solution using multichannel measurements of ECG potentials and QT integral maps in situation with and without changed repolarization. In this study, possible error of the inverse solution due to heart rate changes between the two measurements was analyzed on a model. Minimization of the error by compensation of QT interval changes was proposed. Four commonly used formulas for prediction of QT intervals from heart rates were compared on real measured data.*

Keywords: Body Surface ECG Potentials, Heart Rate Changes, QT Interval, Inverse Solution

1. Introduction

The ischemic heart disease arises as a consequence of decreased blood perfusion in the main coronary arteries. Early detection of starting ischemia allows its more effective medical treatment. Electrical activity of cardiac myocytes can be characterized by their action potentials (AP). Repolarization of a myocyte in the ventricular myocardium is practically independent of the activation sequence and can be described by an elementary integral generator (IG) over the whole cycle of ventricular activity (QT interval). During ischemia, AP amplitude and duration decrease and these changes can be represented by differences between IG's obtained from myocardium with and without ischemia. If the region with ischemia is reasonably small, these differences can be represented by single dipole generator. For known chest geometry and electrical properties, such generator can be assessed by inverse solution [1] from the surface map of differences between QT integrals:

$$DIQT = IQT_{ISCH} - IQT_{NORM} \qquad (1)$$

where IQT_{ISCH} and IQT_{NORM} are surface maps of QT integrals with and without ischemia, DIQT is surface map of differences between these two integrals.

In initial state, cardiac ischemia is not recognizable at rest conditions, but it can be induced by mental or pharmacological load or by exercise that is usually connected with increased heart rate (HR). As a consequence of changed HR, QT interval is also changed and values of IQT_{ISCH} and DIQT maps are influenced what can cause an error of the inverse solution. Therefore QT interval changes should be compensated. To avoid the difficulty with direct QT measurements [2] and compensation, a dependence of QT on HR can be used. In this study, several formulas describing this relation were tested on real ECG measurements.

2. Methods and material

Simulation of ischemia and its inverse identification

Body surface potentials (BSP) were simulated for a normal heart and for several cases of ischemia located subepicardially and subendocardially in three regions supplied by main coronary vessels. Cellular automaton and experimentally observed shapes and durations of

APs were employed to simulate the cardiac depolarization-repolarization. Boundary element method was used to compute BSP in 62 leads. For each lesion, DIQT maps representing the difference between normal and particular ischemic case were computed and together with information on torso geometry and conductivities were used for inverse identification of the best dipole representing the lesion. [1]. Error of the inverse solution was evaluated relatively to a representative dipole (RD) of the lesion obtained as the sum of dipolar moments of model elements in the ischemic region and placed at the centroid of the lesion.

Simulation of changed heart rate between two measurements and its compensation

In simulated IQT maps with and without ischemia manifestation, the same HR in signals was assumed. To analyze the influence of uncompensated HR changes on the inverse solution, changes of the QT_{ISCH} interval from 70% to 130% of the QT_{NORM} were simulated, QT_{NORM}=377 ms and HR_{NORM} = 75 heart beats per minute (BPM) were set for the normal case as the reference values. Changes of HR_{ISCH} from 38 to 165 BPM corresponded to these changes of QT_{ISCH}.

If no change of the QT duration due to ischemia is assumed, the simplest way to compensate the changed HR between the two measurements is to recalculate both *IQT* integrals in (1) to the same QT, e.g. for the QT_{NORM} during normal activation, using a compensation coefficient c_f and the compensated difference integral maps $DIQT_f$:

$$DIQT_f = c_f \, . \, IQT_{ISCH} - IQT_{NORM} \qquad (2)$$

$$c_f = QT_{NORM} / QT_{ISCH} \qquad (3)$$

Measured data and prediction of QT interval

Body surface potentials at 64 leads were measured in 9 patients suffering from heart diseases and in 3 healthy volunteers. All subjects underwent exercise test on supine ergometer with stepwise increasing load from 25 to 125 W. DIQT maps were computed from averaged ECGs at rest (before the exercise) and at load of 75W.

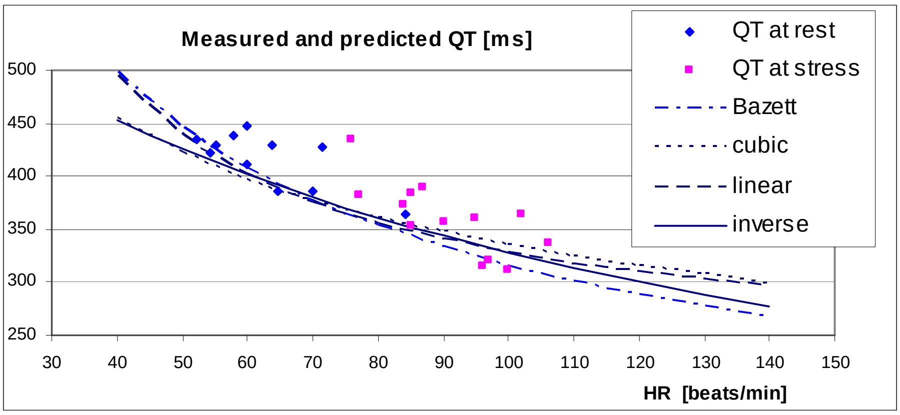


Fig.1.Relations between HR and QT intervals. Dots: QT intervals measured in 12 subjects at rest and during stress. Lines: Tested formulas for estimation of QT from HR or RR with parameters by [3].

ECG signals before the stress test were averaged during about 3 minutes and mean RR intervals were computed. During the load of 75 W, only 12 – 18 heart beats were averaged and HR was traced from the graphic protocol of load and HR. For individual subjects, HR before test varied from 44 to 84 BPM and during the load from 76 to 106 BPM; differences between HR before and during the stress were from 23 to 32 BPM (excluding two healthy persons with differences 16 and 46 BPM). Instants of Q-onset and T-end were set manually, using graphs of averaged ECG signals in all leads and rms signal.

Because of the difficulties with direct QT measurements we proposed to use besides measured values of QT durations also their values predicted from the HR. Four formulas describing the dependence of QT duration on HR or on interval RR (RR=60/HR) were tested. The first formula is the well known formula QT = a1 / RR1/2 proposed by Bazett. Later on, alternative formulas were suggested and claimed to be more accurate: a cubic formula QT = a2 / RR1/3 (Fridericia), a linear formula QT = a3 + b3 RR and an inverse formula QT=1/(a4+ b4 HR) by Rautaharju [3]. Values of parameters obtained from large data sets [3] were implemented. Mean differences between measured and predicted QT were about 20-30 ms (up to 70 ms).

3. Results

Simulation of influence of QT interval changes on inverse identification of local ischemia
Uncompensated changes of QT_{ISCH} caused some errors of location and orientation of the inverse dipole. Mean and standard deviations (SD) of these errors are shown in Fig. 2.

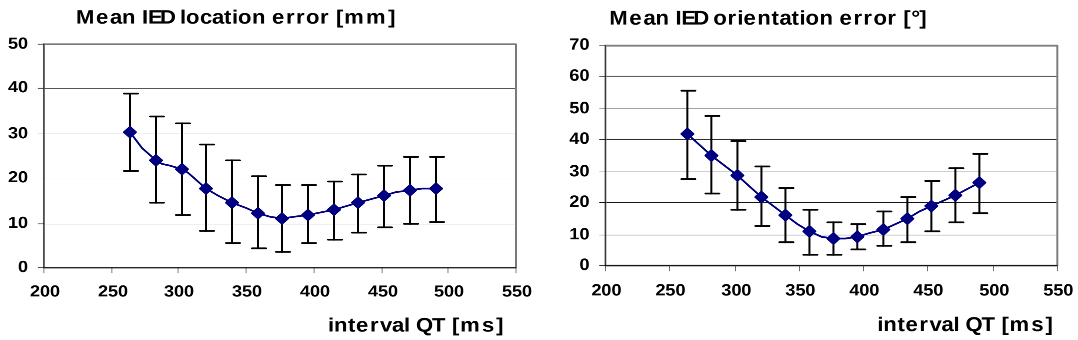


Fig.2. Simulated influence of uncompensated QT changes. The mean and standard deviation of the error of dipole location (left) and its orientation (right) for all simulated lesions.

Minimal mean location and orientation errors represented the error of the inverse method itself were observed for QT_{ISCH}=QT_{REF} = 377 ms (11mm and 9° respectively). Little changes of the interval QT_{ISCH} of about 5% (corresponding to HR changes between 64 and 87 BPM) caused very small increase of the mean dipole location error up to 12 mm and the mean dipole orientation error up to 12°. Changes of QT_{ISCH} to 264 ms (0.7 of QT_{NORM}) caused a considerable increase of the mean error of dipole location of up to 30 mm and error of orientation up to 41°. For some individual lesions the errors were far higher.

Testing of formulas for prediction of QT intervals from measured data

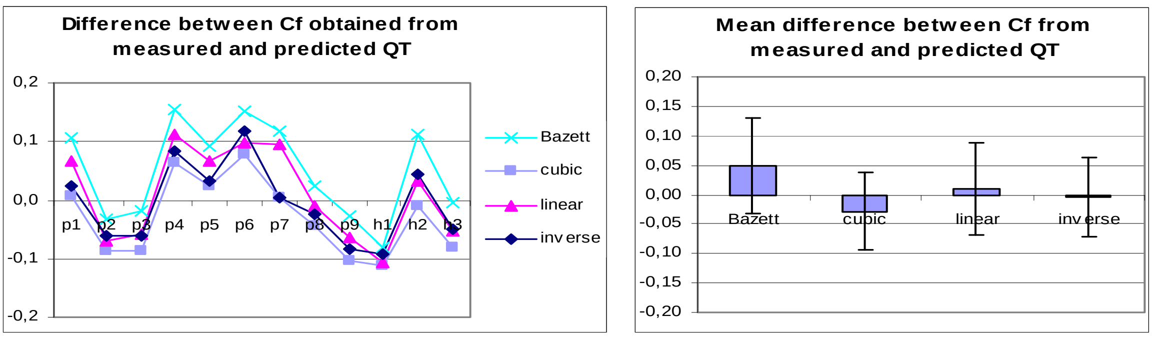


Fig. 3. Differences of c_f coefficients obtained from predicted and measured QT intervals for 4 tested formulas. Left: Individual values for 12 subjects. Right: means and standard deviations of values.

All 4 formulas slightly underestimated measured QT durations (Fig. 1). But for $DIQT_f$ maps and for the inverse solution precision of the c_f coefficient is more important than the QT value itself. Differences between c_f coefficients from measured and predicted values of QT for tested 12 subjects and 4 formulas are shown in Fig. 3 left; mean and SD for 4 formulas are shown in Fig.3 right. The minimum mean value was observed for inverse formula; the minimum SD for cubic and inverse formulas (Fig.3 left).

4. Discussion and Conclusions

Influence of HR variations and connected QT interval changes on the inverse identification of an ischemic lesion was simulated on a model. Simulation results revealed that it is desirable to compensate QT changes above 5%. It corresponds to HR changes of about 12 BPM. High changes of HR significantly influenced the inverse solution.

For real data, HR changes of up to 46 BPM were observed between situations at rest and during exercise. Simple compensation of related changes of QT duration was proposed. In previous study, selection of the formula appeared as more important than selection of its parameters for c_f computation. [4]. Four formulas for prediction of QT interval were tested on 12 subjects and gave moderate differences between predicted and measured QT intervals. When measured and predicted compensation coefficients c_f were compared, inverse formula gave the best results; the worst results were obtained from the formula of Bazett (in agreement with [3]).

However, differences between measured and predicted QT intervals in several individual cases exceeded 5% and corresponding differences of c_f were out of the 0.05 limits (Fig. 2 left) that ensure acceptable error of the inverse identification. To improve the QT prediction, HR stability during each measurement should be checked and interval of averaging should be selected so that HR before and during the period is relatively invariable (for example at the end of the load). Particular heart cycle can be included into the averaged data only if the whole QT pattern falls into prescribed boundaries.

Acknowledgement

This work was supported by grants 2/7092/27 from the VEGA grant agency and APVV-51-059005 from the APVV grant agency in Slovakia and partly supported by research project No. NN518 328035 financed by the Polish Ministry of Science and Higher Education.

References

[1] Tyšler M, Kneppo P, Turzová M, Švehlíková J, Karas S, Hebláková E, Hána K, Filipová S. Non-invasive Assessment of Local Myocardium Repolarization Changes using HR Surface ECG Mapping. *Physiological Research,* 56 (Suppl 1): S133-S141, 2007.

[2] Murray A, McLaughlin NB, Bourke JP, Doing JC, Furniss SS, Cambell RWF. Errors in manual measurements of QT intervals. *British Heart Journal,* 71: 386-390, 1994.

[3] Rautaharju PM, Zhou SH, Wong S, Prineas R, Berenson GS. Functional characteristics of QT prediction Formulas. The concept of QTmax and QT rate sensitivity. *Computers and Biomedical Research,* 26: 188-204, 1993.

[4] Turzová M, Tyšler M, Hebláková E, Švehlíková J. Influence Of Heart Rate Changes On Noninvasive Identification Of Local Ischemia And Its Compensation. In: Analysis of Biomedical Signals and Images. Proceedings of Biosignal 2008. 220 (7 p.).

EEG Characterization of Psycho-Physiological Rest and Relaxation

M. Teplan, A. Krakovská, S. Štolc

Institute of Measurement Science, Slovak Academy of Sciences,
Dúbravská cesta 9, 841 01 Bratislava, Slovakia
Email: michal.teplan@savba.sk

Abstract. *EEG correlates of psycho-physiological rest are addressed. Experiment consisted of 88 relaxation sessions of 8 subjects. 6-channel EEG data of 3-minute duration were examined. Firstly EEG characteristics of rest were revealed in a form of linear regression trends. On the contrary to general expectations, during resting conditions of 3-minute session in darkened room in lying position with eyes closed, relative alpha-1 powers were decreasing, followed by increase in theta-1 and decrease in total powers over the whole cortex. EEG features derived from linear regression model were selected according their ability to discern between subjectively more and less successful relaxation. Lower overall power in backward areas as a distinctive characteristic of more successful relaxation was accompanied by increased alpha and beta-1 features in some cortex areas. Potential applications involve clinical, pharmacological, and stress management areas.*

Keywords: EEG, relaxation, signal analysis

1. Introduction

The relaxation response is an integrated body reaction, which has been found to have such benefits as increased mental and physical health and improved ability to deal with tension and stress. In the literature we had found no direct characterization of EEG features of relaxation. However need for characterization of resting status of patients may be relevant for many areas, e.g. stress reduction, sleep deprivation, and testing of pharmacological substances related to hypnotic and sedative drugs. Poor relaxation response is usually connected to problems of stress. Actually, stress is acknowledged as one of the major problems of modern society. While calling for stress reduction, need for monitoring tools for stress may grow.

It might be difficult to define relaxation properly. First of all, it is related to subjectivity and may vary across individuals. Implicitly, under the term relaxation it is usually meant certain positive and beneficial phenomenon. Regular relaxation practice should have more side effects on different physiological and psychical parameters of individuals (related e.g. to aging, digesting, general peace, psycho-somatic deceases). There exist no generally accepted references in a form of other physiological parameters. According to Travis five different categories of physiological variables might be sensitive to level of relaxation [1] including breath and heart rates, and skin conductance. Similar parameters for characterization of stress level are reported by Vavrinský [2]. Physiological components of relaxation response introduced in [3] were decrease in oxygen consumption, decrease in respiratory rate, decrease in heart rate, and increased in alpha brain wave production. Researchers usually consider increased alpha production as a sign occurring during mental and physical rest [4, 5, 3]. According to Ossebaard relaxation is made up of several biological, psychological, and social components [6].

In our previous study we interpreted increased alpha and theta powers during long term audio-visual stimulation as signs of relaxation [7]. The aim of this study was to examine basic EEG characteristics during resting conditions. In particular, two consequent tasks were addressed: 1) Finding changes in EEG measures during sensori-motorical rest reflecting functioning of the central nervous system. 2) Selecting the most appropriate objective EEG features that are able to distinguish between more and less successful relaxation determined by subjective assessment.

2. Subject and Methods

Eight right-handed healthy subjects (3 females and 5 males, mean age 27 years, st.dev. 7 years) volunteered in the experiment. Subjects were instructed to keep their eyes closed and relax both physically and mentally. The lying position during 3-minute EEG measurements was comfortable enough to avoid unwanted activities and to diminish the occurrence of some artefacts caused by feeble motion. Monopolar EEG montage was realized by EEG cap and comprised eight channels with electrodes placed on F3, F4, C3, C4, P3, P4, O1, and O2 locations according the International 10-20 system. The reference electrode was located at Cz and the ground electrode at Fpz point. A standard cap system (Electro Cap Inc.) with Ag-AgCl electrodes was employed. In order to prevent signal distortions, impedances at each electrode contact with the scalp were kept below 5 kΩ, and balanced within 1 kΩ of each other. EEG recording was realized by unit with the following parameters: no. of channels: 8, amplifying gain: 402, sampling frequency: 500 Hz, A/D converter resolution: 16 bits, input resolution: 0.46 µV, noise: max 4.1 µV pp.(0.07 to 234 Hz), low pass filter: 234 Hz (-3dB), high pass filter: 0.07 Hz (-3dB).

From the 8-channel signal between active electrodes and reference electrode six difference signals F3C3, F4C4, C3P3, C4P4, P3O1, and P4O2 were derived by off-line transformation in order to avoid undesirable effects of common reference electrode. The electroencephalograms were analyzed first by online visual control of the ongoing EEG in 8 channels and later by off-line analysis. Sequences contaminated by either subject-related or technical artefact and obvious sleep occurrences were excluded by eye inspection and according to the subject's assessment. Digital high pass FIR filter with cut-off at 0.75 Hz and with 3000 point Blackman window was utilized.

To uncover objective changes from obtained EEG data, we computed the following signal characteristics: total power, relative frequency powers, and magnitude squared coherences in nine bands: delta-1 (0.5-2 Hz), delta-2 (2-4 Hz), theta-1 (4-6 Hz), theta-2 (6-8 Hz), alpha-1 (8-10 Hz), alpha-2 (10-12 Hz), beta-1 (12-16 Hz), beta-2 (16-30 Hz), and gamma (30-45 Hz). Powers and coherences were computed in Matlab environment. Time subwindows were 2 seconds long with 0.75 overlap.

From each of 8 subjects we obtained 24 recording sessions. For the first task we chose 11 artefact-free recordings from each participant. From them, in the second task we chose 4 least and 4 most successful relaxation sessions for each subject. Choice was done according to answers to the following task: "Assess a level of your relief accomplished during 3-minute resting period." Subjective measure was rated on 7-point bipolar scale.

3. Results

We identified evolutions of various measures during 3-minute intervals. Most of them could be reasonably interpolated by linear regression. For evaluation of trend magnitudes we constructed index out of absolute difference obtained by measure’s trend according to linear model (Δ_{abs}) divided by standard deviation of residuals in respect to linear regression:

$$\Delta^{res} = \frac{\Delta_{abs}}{std(res)}$$

This index is reasonable alternative to p-value from testing whether regression slope differs from flat one by F-test and they are in mutual unambiguous nonlinear correspondence.

In order to discover general features of the resting process, we focused on strong nonzero trends from average (n = 88) course of each measure, that should reflect certain changes in physiological functioning. In Fig. 1 samples of averaged behaviour of measures with cortex location identification during resting conditions are presented.

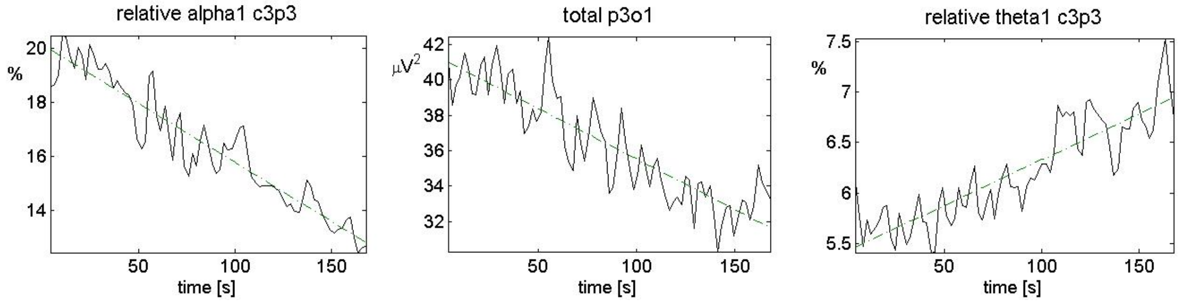


Fig. 1: Strongest trends of averaged behaviour of EEG measures in particular cortex locations during resting conditions.

Sensori-motorical rest can be characterized by trends of the following EEG measures. The strongest changes were observed across all six cortex regions by decrease of relative alpha-1 power (cortex average $\Delta^{res} = -8.9$), further by increase in theta-1 relative power ($\Delta^{res} = -6.1$) and by global decrease in total power ($\Delta^{res} = -4.8$). The strongest combination of alpha decrease and theta increase was observed in C3P3 area (Fig. 1), where alpha-1 relative power was diminished by 40 % from its initial values (typical st.dev. 15 % for average points in graph).

Further goal was to find out special EEG features bound to 'successful' relaxation expressed on the contrast of 'less successful' relaxation. Recordings were divided into two groups according their subjective evaluation of general well-being during the relaxation period. For each subject groups were formed separately; four artefact-free sessions were taken with the lowest scoring and four with the highest one.

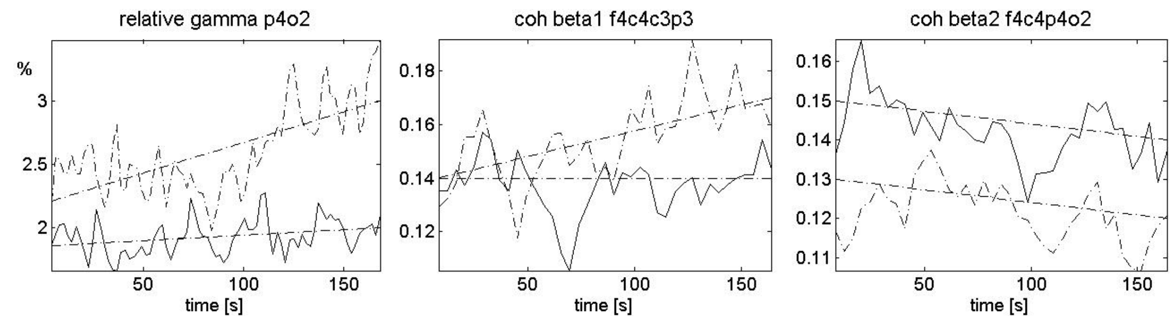


Fig. 2: Samples of average curves for more and less (dashed) successful relaxation differing by either initial values, slopes or terminal values.

For each individual data (32 cases from 8 persons in each group) linear regression model was counted. Two relaxation categories might differ by three types of features related to linear regression: initial values, trends (Δ^{res}), and terminal values. In order to find distinct features of two relaxation categories, two-sample heteroscedastic student's t-test was utilized for data that did not violate normality in Shapiro-Wilk test ($p > 0.05$) and nonparametric Kruskall-Wallis test for those data where normality was refused (both $p < 0.002$ after involving Bonferroni correction). More successful relaxation was characterized by lower final levels of total power in parieto-occipital areas, overall higher relative alpha powers, and higher terminal relative beta-1 powers in backward regions. Further significant changes were observed in beta coherence (in both relation directions) mostly interhemispherically through all three types of features.

4. Discussion

Decrease of total power over the whole cortex implies that overall brain activity was gradually diminished during the resting process, with the strongest contribution on this change realized by alpha-1 band. In spite of the fact that during our resting conditions we detected no increase in appearance of alpha waves during 3 minutes, consistent increase of alpha waves in our data could be found: Focus on shorter time windows shows that relative power in alpha-2 band in average increased during the first 30 seconds in all cortex regions. In the literature commonly accepted fact on alpha wave increase during rest and relaxation [3] should be rather related to eye closing phenomenon, by which resting is usually initiated, and to kind of lighter rest - rest with shorter time span and perhaps to sitting position. Moreover we have shown that in principle discrimination of two relaxation categories from EEG data is possible. In further work referencing to more physiological outcomes (e.g. cardio-vascular, respiratory, muscular, blood secretion) is suggested.

Acknowledgements

This work was supported by Slovak Grant Agency for Science (grant No 2/7087/27).

References

[1] Travis, F. Autonomic and EEG patterns distinguish transcending from other experiences during Transcendental meditation, *Int. J. Psychophysiol.,* 42: 1–9, 2001.

[2] Vavrinský, E. Tenkovrstvové mikroelektródy pre elektrochemické vodivostné senzory aplikované v biomedicínskom monitorovaní stresu, PhD thesis, Slovak Technical University in Bratislava, Faculty of Electrical Engineering and Informatics, Department of microelectronics, 2005.

[3] Foster, D. EEG and subjective correlates of alpha-frequency binaural-beat stimulation combined with alpha biofeedback. Memphis State University, 1990.

[4] Lindsley, D. Psychological phenomena and the electroencephalogram, *Electroenceph. Clin. Neurophysiol.,* 4: 443, 1952.

[5] Brown, B. Recognition of aspects of consciousness through association with EEG Alpha activity represented by a light signal, *Psychophysiology*, 6: 442–452, 1970.

[6] Ossebaard, H. Stress reduction by technology? An experimental study into the effects of brainmachines on burnout and state anxiety, *Applied psychophysiology and biofeedback,* 25: 93–101, 2000.

[7] Teplan M. et al. EEG responses to long-term audio-visual stimulation. *Int. J. Psychophysiol.*, 59: 81-90, 2006.

Specific Measuring Systems

Quality Control and Nanometrology for Micro/Nano Surface Modification of Orthopaedic/Dental Implants

P.H. Osanna, M.N. Durakbasa, K. Yaghmaei, L. Kräuter

Institut für Fertigungstechnik - Technische Universität Wien
Department for Interchangeable Manufacturing and Industrial Metrology
Vienna University of Technology, A-1040 Wien, Austria

Abstract: *Nowadays implant surfaces are created by a large variety of manufacturing processes and techniques such as: CNC turning and milling, broaching, casting, grinding, polishing, honing, electro chemical etching, welding, brazing, stamping, bending, etc. The manufacturing accuracy range and allowable tolerances for implants lie between 1 µm and 100µm.*

The most important parameters in determining the suitability of an implant are its biocompatibility, functionality, performance, and corrosion resistance. In this paper the aspects of implant surface modification methods influencing biocompatibility, corrosion resistance and mechanical property nature of implants are discussed and how these depend on geometrical dimension and size, shape, size of roughness and surface topography.

Natural bone surface contains features that are about 100 nm. The interaction of bone tissue and total hip joint and precise assessment of wear, friction and miniaturisation demand for creating of nanometer scaled implant structure, implant surfaces thin film deposition and ultra precision surface treatment require the application of state of the art of new manufacturing and measurement instrumentation and techniques. These include micro and nanofabrication of surface patterns and topographies by the use of laser machining, photolithographic techniques, and electron beam and colloidal lithographies to produce controlled structures on implant surfaces in the range of size from l0 nm to 100µm.

Keywords: Structured Surfaces, Atomic Scale Machining, Laser Machining, Nanometrology, Surface Modification, Implants, Medical Devices

1. Introduction

Nowadays implant surfaces are created by a large variety of manufacturing processes and techniques such as: CNC turning, milling, broaching, casting, grinding, polishing, honing, electro chemical etching, welding, brazing, stamping, bending, etc. The range of manufacturing accuracy and tolerances for implants lies between 1µm and 200 µm [1].

Natural bone surface contains features that are about 100 nm. The interaction of bone tissue and total hip joint and precise assessment of wear, friction and miniaturisation demand for creating of nanometer scaled implant structure, implant surfaces thin film deposition and ultra precision surface treatment require the application of state of the art of new manufacturing and measurement instrumentation and techniques. At present these developed techniques and instruments are missing in manufacturing processes in the implant industry. The scalability benefit of implant design and production from pushing the micro into nanoscale implant engineering techniques needs multidisciplinary approaches of ultra-precision engineering techniques, micro/nanoscale surface metrology, material science and medicine.

The past decades and current research and development in the area of biomaterials and medical implants show some general trends. One major trend is an increased degree of functionalisation of the implant surface, better to meet the demands of the biological host system. While the biomaterials of the past and those in current use are essentially bulk materials (metals, ceramics, polymers) or special compounds (bioglasses), possibly with some additional coating (e.g. Hydroxyapatite), at present the current research on surface modifications points toward much more complex and multifunctional surfaces for the future. Such surface modifications can be divided into three classes, one aiming toward an optimized three-dimensional physical microarchitecture of the surface (pore size distributions, "roughness", etc.), the second one focusing on the (bio) chemical properties of surface coatings and impregnations, and the third one dealing with the viscoelastic properties (or more generally the micromechanical properties) of material surfaces.

These properties are expected to affect the interfacial processes cooperatively, i.e., there are likely synergistic effects between and among them: The surface is recognized by the biological system through the combined chemical and topographic pattern of the surface, and the viscoelastic properties [1].

2. Biocompatibility

Biocompatibility consideration of metallic, polymeric and ceramic materials has led to their use as standard materials for implants and medical devices. However, increasing demands on implant materials, and the trend towards device miniaturisation due to the enhanced functionality and performance has put the spotlight on the implant surface and its properties. Implants all possess inherent morphological, chemical, and electrical surface qualities which elicit reactionary responses from the surrounding biological environment. In fact, biocompatibility can be described as multi-factorial in that simultaneous stimuli from any of these material properties can affect the host response. There are many factors which influence implant biocompatibility such as implant size, shape, material composition, surface quality and roughness.

3. Corrosion of Implants

Implant Corrosion is one of the major degradation processes that might occur in vivo, and should be considered for evaluating new biomaterials and new designs of medical devices [3]. The bio-environment may be described as "aggressive and angry", and is associated with a variety of conditions. Metals and alloys, which are extensively used in medical devices, might corrode severely in this bio-environment in accordance with both thermodynamic and kinetic considerations. This degradation process is undesirable primarily because it limits the functionality and lifetime of medical devices, and secondly because it releases corrosion products that may elicit an adverse biological reaction in the host. Corrosion is a complex phenomenon that depends on geometric, metallurgical, mechanical, etc. parameters. Implant manufacturing understanding of these parameters and their synergistic effects is required in order to control implants corrosion. Fretting corrosion occurs to some extent on all load bearing orthopaedic implants.

4. Mechanical Properties

The major problem associated with the currently used implants is due to inadequate implant-tissue interface properties. The integration of load-bearing implants, such as hip and knee prosthesis and dental implants, into surrounding tissue is important [4]. Both in vivo and in vitro studies have shown that implant surface topography may affect epithelial and connective tissue behaviour. Based on these observations, a mathematical theory of ideal surface pit morphology, dimensions and densities of biomechanical significance was formulated. Another theoretical analysis concluded that if the geometric form of surface roughness is held constant, then the peak elastic stress depends on the form rather than the size of roughness. There are several in vitro studies that show how the physical, chemical and mechanical properties of implant surface can be improved. Unfortunately, they did not measure the

topography of these differently treated surfaces. It would be interesting to compare surface topographies to the above results.
Fretting of total hip and knee replacements occurs at such places as the bone-stem interface, the cement interface, and on the interfaces of modular connections between implant components. Fretting also occurs at areas of cyclic load bearing contact on metallic bone-plates and screws, such as screw-plate connections, bone-plate interfaces and screw/bone interfaces.

5. Ultra Precision Manufacturing and Nanoscale Surface Metrology for Implants

Implant surfaces are created by a large variety of manufacturing processes. Implant manufacturing has been helped in the last few years with CAD-CAM systems applying casting or machining (milling, grinding, lapping, polishing).

The indirect method for characterisation of dimensional and functional properties of engineering surfaces measures the integrity of a surface by means of micro/nano scaled surface metrology [5]. The development of surface texture, form and 3D surface characterisation methodologies, topography and roughness parameters considering wear and friction has lead the micro/nano scaled surface metrology research to become involved in characterization of implant surfaces. At the time being 3D surface measurement is already proving more and more to be an important tool in several areas of surface analysis including wear, indentation, topography, contact problems and functional behaviour of surfaces.

The measurement process leads to “control by measurement” of the manufacturing process. With the recent developments in new implant surface treatments implant surface layers are treated, or surface engineered. In new and developing technology, the implant surface features such as roughness (10nm to 100μm) are often aimed at producing the implant surface functionality which leads to enhanced implant biocompatibility and performance.

It is obvious that a full understanding between surface topography and functional performance of an implant can only be realised if a 3-dimentional approach to surface characterisation is utilised. 3D-surface measurement is a very valuable tool in implant surface analysis including wear, indentation, topography and contact problems of implants (see Figure 1).

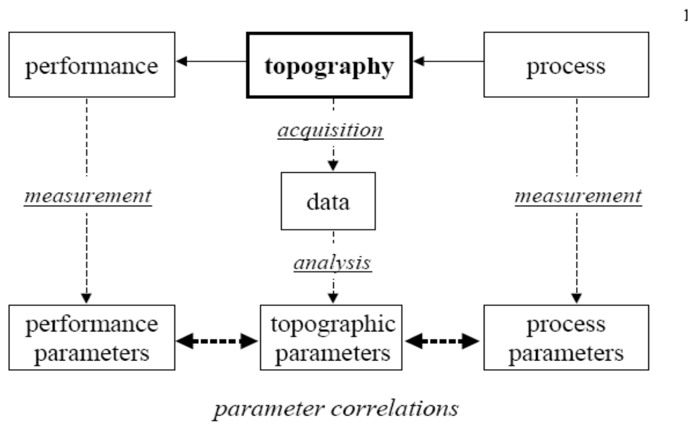


Fig. 1. Parameter correlation according to process, topography and performance.

Due to the topography assessment of implant surfaces:
- Optical microscopy,
- Scanning electron microscopy (SEM) and
- Atomic force microscopy (AFM)

are applicable for metrology.

6. Structure at the Bone-Implant-Interface

Natural bone surface contains features that are about 100 nm across. If the surface of an artificial bone implant were left smooth, the body would try to reject it. Because of that smooth surface is likely to cause production of a fibrous tissue covering the surface of the implant. This layer reduces the bone-implant contact, which may result in loosening of the implant and further inflammation [6]. It was demonstrated that by creating nano-sized features on the surface of the hip or knee prosthesis one could reduce the chances of rejection as well as to stimulate the production of osteoblasts. The osteoblasts are the cells responsible for the growth of the bone matrix and are found on the advancing surface of the developing bone.

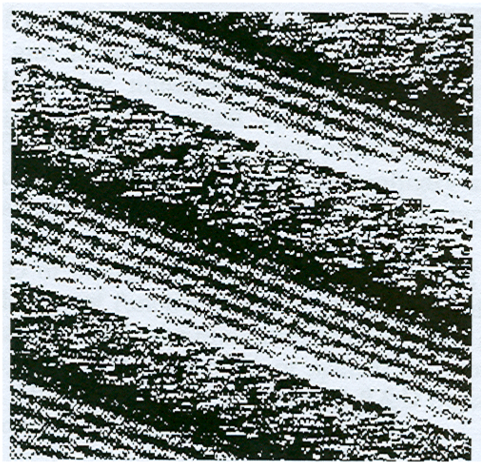


Fig. 2. Laser machined surface

In the last few years, there has been a recent development in the functional significance of surface finish. This is the use of periodic pattern on the surface to enhance a specific use. Structured surfaces are surfaces with a deterministic pattern of usually high aspect ratio, geometric features designed to give a specific function [7,8]. In general structured surfaces are made using conventional processes like cutting with single point diamond turning. Some other ways of making structured surfaces are:

- Moving/removal material (machining, burnishing),
- etching (plating, evaporation),
- replication (hot embossing, casting),
- material modification (Laser texturing) (see Figure 2),
- solidification of liquids),
- photolithographic techniques, and
- electron beam lithographies are attractive ways of making structured surfaces.

Using these techniques enables structures to be imparted on many different and difficult materials.

These techniques for structuring of implant surfaces have made surface structure with pore diameters up to 10 nm to 100µm possible. This permits the development of defined, reproducible and economical surface topographies which reduce wear and friction and will contribute to corrosion resistance and improving biocompatibility, functionality and performance of implants.

7. Biocompatible Coatings

Developing thin film technology for different industrial applications including machining and bio-medicals industries with regard of enhanced orthopaedic implant performance is one of the fastest-growing areas in the field of biomaterials and many developments are anticipated over the next decade.

Biocompatible coating deposited on implants facilitates implant fixation and bone ingrowths. Modifying of implant surfaces by ion implantation or physical vapour deposition exhibit superior hardness and wear resistance. Such technology will also be applicable to the rapidly expanding field of dental implantology where osseo-integration depends on the surface roughness of implants.

Coating is applied to orthopaedic components and other medical devices for a variety of reasons. Porous metal and ceramic coatings deposited on implants facilitate implant fixation and bone ingrowth [9]. Implant surfaces modified by ion implantation or physical vapour deposition exhibit superior hardness and wear resistance. Polymeric coating formulations are used to enhance biocompatibility and biostability, thromboresistance, antimicrobial action, dielectric strength, and lubricity; make medical devices used within the body more visible to ultrasound.

One challenge therefore appears to be to design biocompatible coatings that can resist both the chemical and mechanical environments, which are often poorly defined for both orthopaedic and dental applications, without degrading the very property that is required in the advance material. Hydroxyapatite (HA), is the material for biologically compatible coatings on metal substrates.

8. Further Development and Future Research

8.1 Basic Research:

- Nano scaled surface metrology.
- Micro/Nanotribology and material characterization.
- Development of new parameters and tolerances for micro and nanoscale surfaces.
- Characterization of ultraprecise manufactured implant surfaces and devices.
- New approaches to a comprehensive Quality assurance for implants.
- Biomaterials: Biocompatibility test methods, standards.
- Medical Informatics: Image analysis and reconstruction, knowledge based systems.
- System Analysis: Modelling, Computer aided simulation, functional analysis.
- Collection of all received data in a database.
- Biomechanics.

8.2 Applied Research:

- Application of new measurement instrumentation and methods for implant production.
- Application of PVD, CVD, Hydroxyapatite porous coatings in implant manufacturing.
- Application of nanoscale surface modification techniques in the implant production process.
- Micro/nanotribology assessment and material testing for implants.
- Approaches to development of new measurement instruments.

9. Synopsis and Concluding Remarks

This needs the integration of nanometrology in implant manufacturing processes for precise manufacturing and biocompatibility of implants. The novel nanoscale surface metrology philosophy introducing the new attempt of industrial nanometrology in implant science and technology will lead to industrial breakthrough in manufacturing and measurement of orthopaedic and dental implants and devices, and to point out specific points of special interest [2]:

- Ultra-precision machining of orthopaedic and dental implants and devices.
- Application of nano scaled surface metrology in the implant production process and characterization of ultra-precise manufactured implant surfaces and devices.
- Application of new measurement instrumentation and methods for implant production.
- Development of new parameters and tolerances for micro and nanoscale implant surfaces due to the development of standards at the nanometer level.

The other related activities to the objectives are:

- Investigation of implant size, shape, roughness influencing biocompatibility.
- Investigation of implant size, shape, roughness influencing corrosion resistance.
- Improving the implant-tissue interface properties for load-bearing implants through nano-scale 3D surface topography.
- Determining of the implant mechanical properties (elasticity modules, Poisson's ratio, and stress), plastic deformation, density, adhesion, friction, and tribological behaviour.
- Tribology of hard bearing surfaces in hip prosthesis
- Surface modification of implants due to enhanced tribological performance through reducing the wear of artificial joints.
- Biocompatibility testing on metallic, ceramic polymer implants.
- Measurement methods, standards for the scientific understanding at the interface between the materials and biological sciences.
- Work towards providing traceable nanometrology resources and development of standards at the nanometer level.

References

[1] Kasemo, B., Gold, J.: Implant Surfaces and Interface Processes. Adv. Dent. Res. 1999, June 13, 8/20.

[2] Yaghmaei, K.: Direct and Indirect Methods for Micro/NanoCharacterisation of Dimensional and Functional Properties of Engineering Surfaces. P.H.D. Dissertation, Thesis, 2005.

[3] Eliaz, N., Mudali, U.K.: Biomaterials Corrosion Reviews. Special Issue Vol. 21, 2003, N.2-3.

[4] Kapanen, A.: Biocompatibility of Orthopaedic Implants on Bone Forming Cells. ISBN 951-42-6606-4.

[5] Stout, K.J.., Blunt L.: Application of 3-D Topography to Bio-Engineering, 6th International Conf. on Metrology and Properties of Engineering Surfaces. 1994, Birmingham, UK.

[6] Salata, O.V.: Applications of Nanoparticles in Biology and Medicine. Journal of Nanobiotechnology,2004, 2:3 doi:10.1186/1477-3155-2-3.

[7] Whitehouse, D.J. : Handbook of Surface and Nanometrology. 2000.

[8] Toenshoff, J., Von Alvensieben, H.K., Temme T., Willrnann G.: Review of Laser Micro Structuring. Proc. 1st Conference of Euspen. 1999, Bremen, Germany.

[9] Handbook of Materials for Medical Devices. ASM International, Materials Park, Ohio, USA.

Metrology and Nanotechnology in Interrelation with Intellectual Property Rights

L. Kräuter, M.N. Durakbasa, P.H Osanna

Department for Interchangeable Manufacturing and Industrial Metrology
Vienna University of Technology, Vienna, Austria,
Email: osanna@mail.ift.tuwien.ac.at

Abstract. *Production and measurement systems reached in the past years resolution of nanometer. In this ultra small size of three dimensional structures, there can be discovered unknown physical effects for any kind of technology (electricity, friction, adhesive force,…). But it is possible that not all of this effects or manmade technologies will find protection under the existing Intellectual Property (IP) rights (IPR).*

Keywords: Nanotechnology, Semiconductor, Design of Layers and Surfaces, Metrology, Topographical Structure of Materials, Intellectual Property

1. Introduction

The history of intellectual property right (IPR) on three dimensional structures was found in the USA due to their research and industrial leadership on microelectronic and semiconductor devices in the 80's decade, like on Integrated Circuits (IC) or Central Processing Units (CPU) - or simply called Microchips or only Chips. This movement of the semiconductor industry found its result in the "Semi-conductor Chip Protection Act – SCPA" in 1984, which was called in technical terms "mask work".

The USA forced other industrial states to establish also such new IP rights to their IPR, so that treaty states mutual guarantee protection of those electronic devices, if it was applied for.

In Europe, Germany established a protection of the topography of semiconductor chips in the »Halbleiterschutzgesetz – HalbSchG« (= Semiconductor) law in 1987, in Austria the right was established as »Halbleiter-schutzgesetz – HlSchG« (= Semiconductor) law in 1988 and in Switzerland it was established as »Topographiengesetz« (= Topography) law in 1992.

To receive a protection of the semiconductor technology, the topography of the chip unit must have an own, unique character (sometimes called idiosyncrasy or secret of nature), which means that the design of it must be based on an intellectual work and not be a copy of another registered or widely

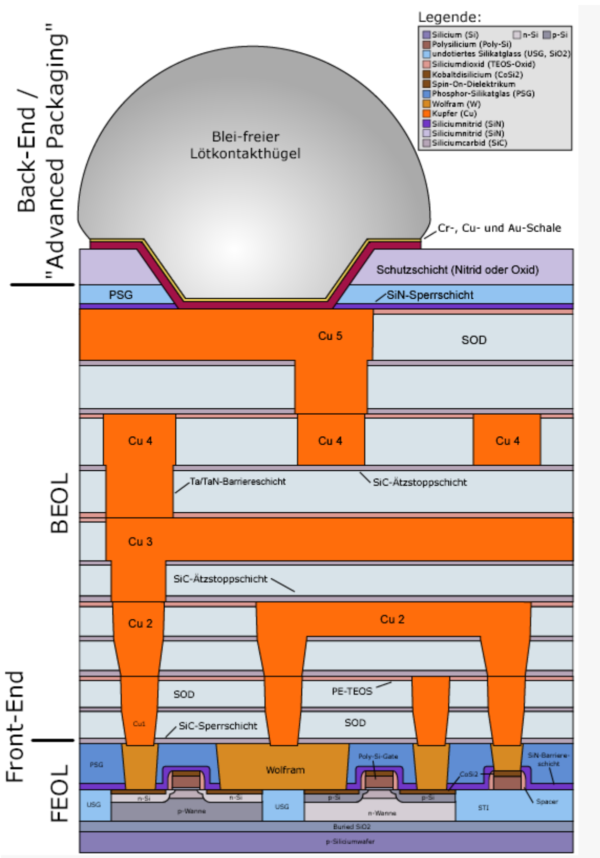


Fig. 2. Structure of a modern IC [1]

well known semiconductor topography. During the registration process in the patent offices, the application is only executed on formalities.

The maximum duration time of this industrial property right is limited up to 10 years, starting on the date of filling. The numbers of yearly applications for this industrial property right is at the moment not high, for example in Germany it is lower than 50 during the past years.

2. Metrology in Interrelation with Nanotechnology and Bionic

With the development of the Scanning Tunnelling Microscopy (STM) and the Atomic Force Microscopy (AFM) in the early 80's, a new area of metrology started. It was the first time that the resolution of microscopy reached the level of atomic structure layers or even further, the position of a single atom unit had become detectable. The STM can be operated as a measurement instrument in two modes:

- "Constant Height Mode", which allows a faster operation, but it allows no safe operation (stylus tip crashes on the surface if the profile is higher than expected),
- "Constant Gap width Mode" (CGM), also known as "Constant Current Mode" (CCM), which is slower in operation, but aware of stylus tip crashes.

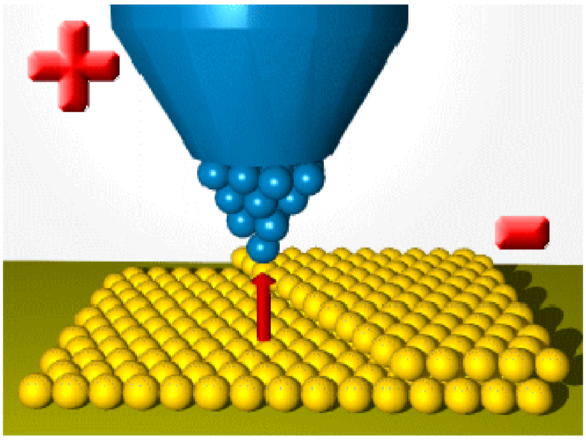


Fig. 1. STM tip functionality [2]

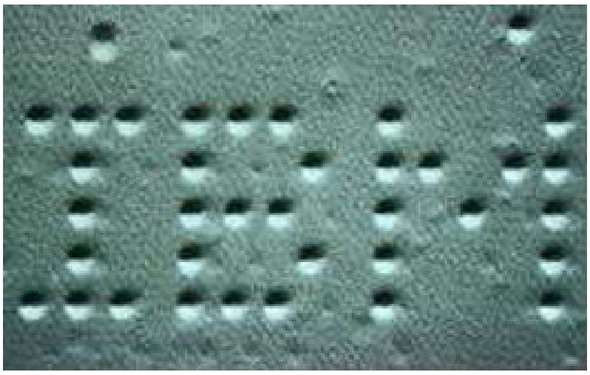


Fig. 3. Xenon atoms on a nickel surface [3]

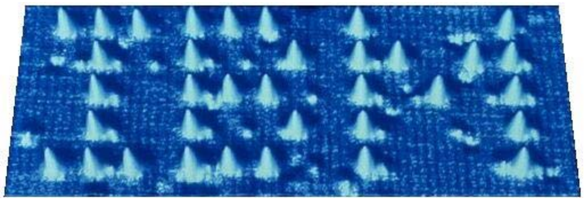


Fig. 4. Xenon atoms on a nickel surface, STM measurement result [4]

A little bit later after the development of the STM microscope, another effect of this measurement instrument had been discovered: it became an atomic machine tool. Only with a slightly change of the polarity of the stylus tip, the same measurement machine became an atomic machine tool by attaching / absorbing atoms, moving them to a new position and finally emitting / dispensing those atoms on a atomic layer surface (mostly called substrate) – and this whole process is repeatable controlled by human mind. On of the first and famous images had been created by IBM researchers, who positioned single Xenon atoms on a nickel substrate in letter form, like a dot matrix printer, in 1990. The created famous first word was simply: "IBM" – as shown in Figure 2 and 3. Since that, the technology was developed much further. Nowadays, a single electron can be surrounded by a corral of Fe-Atoms. Quantum effects can be measured by the latest version of STM's in the same way like distances.

But not only in atom level the structure surface has a main impact on the physical properties, also in the nanometer level in the living nature it has effects. One of those effects, which had been scientifically investigated, was the selfcleaning effect of the lotus flower surface where water doesn't wet the surface. This effect is nowadays common well known as the "lotus-effect". In the early 70's, a German biologist – Wilhelm Bartholtt – investigated the surface structure with a scanning electron microscope. This was also the time, when bionic – scientific research of nature effects and implementation on or in machines in the industrial gods – was developed, even if it was not named so in that way at these days.

A simple explanation of the effect is: water trop has a high surface tension and follows the tendency to minimize it's surface to a sphere shape. The plant uses two effects to avoid wetting:

1. it segregates wax on the leave surface and
2. the segregate has the shape of stalagmites (tips, grains or small hills, about 10 - 20 µm high and with a distance distribution of 10 - 15 µm from each other. See Figure 5)

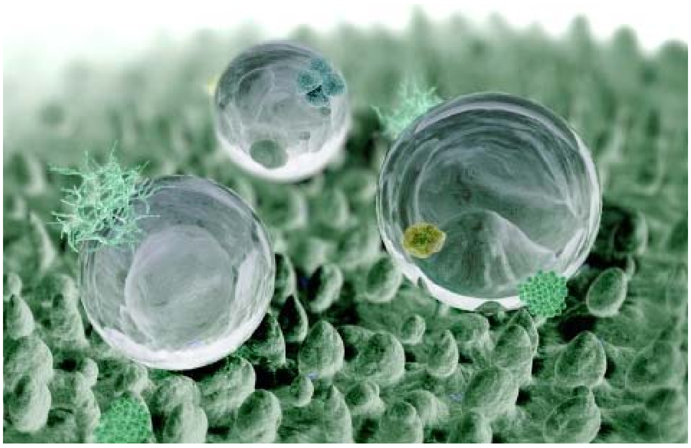


Fig. 5. Lotus-effect, virtual graphic (VG) [5]

So a water trop will be balance on the top of these tips so that less than 5% of the trop surface is in contact with the plant. Due to the surface tension of the water trop and the water resistant layer of the plant, the water trop doesn't coat the surface. With the industrial application of this effect, it is possible to reduce the amount of cleaning work of building surfaces, like facade, glass windows etc.

3. Topological Surface Structures and Intellectual Property

To fulfil positive the criteria's of the standard patent application procedure, a product, a process, a chemical substance, etc. must:

1. have a technological usability,
2. be novel,
3. have an inventive step,
4. is not based on human spirit,
5. be repeatable and finally,
6. the complete disclosure of the technology.

In the case of the Lotus-effect, this means that the effect won't be patentable, because it existed already in the nature. And if someone would find an algorithm with whom it would be possible to improve the Lotus-effect, it wouldn't also be patentable, because it is a work of human spirit. What would be patentable, would be for example, a coated film or foil which imitates the Lotus-effect, a machine which produces this kind of film or foil.

Also it seems to be possible to get protection of the product under the umbrella of the industrial design right. The industrial design right belongs also to the IPR and protects designs, if the design has a important impact on the creation of the shape, the configuration or composition of pattern or colour or combinations of both and has a product inherent three dimensional form containing special and distinguishing aesthetic values.

The important legal questions in the decision making procedure would be:

- Is it a design feature that water is not wetting the surface?
- And would be design of the surface with these countless, invisible tips for the standard average consumer and market participant identifiable?

Since the effect has a main impact from the topographical structure, so why not apply for protection under the Topography /Semiconductor law?

The application would be rejected, because it has no feature of a semiconductor device (transistors, conductive pathways, no use of electricity …). So there exists not only a "metrological gap" in this area of resolution, it exists also an "IP gap" of rights.

It would be possible for patent attorneys to reformulate the technology subject to reach an IP right under the existing patent law. But would this reformulation process really protect the invention in that way like it had been in the mind of the inventor?

True is, that nanotechnology is still a young science and only a few effects have been found nowadays and even less had been scientifically completely understood. Even more the possible technology applications are still rare in everyday consumer products. Also laws will only be established, if new society questions, cases, technologies, etc. must be regulated with new rules, so that everybody of the society can follow them. Of course, sometimes this will take decades of years, till the legal organisations will recognize these topics and pick up those demands to find practicable solutions for these questions. A good example for the "signal response time" of the legal system is the above stated "Semiconductor Chip Protection Act – SCPA" law. From the first semiconductor transistor in the laboratory in 1945 till to the protection act of 1984 it took almost 40 years for the legal system to establish (according to the demand of the industry – not on their own motivation) a practicable right [6].

If one takes this circumstance as a scale which can also be applied to this topic, it will still take some years, till new kinds or different powerful sorts of IP rights will be established.

4. Concluding Remarks

Science and technology is an on going discovering and developing process, everyday and worldwide. As in the past history, technologies had been invented which had a major impact on the society and with a delay also on the legal system.

At the moment, the legal system of any government is trying to get the last "technological revolution" – not in a scientific way, but on the economic and society impact – the internet (also called: word wide web, short: www) into useful regulations.

From the point of view of the immaterial property rights, the copyright of content in the www belongs to creator of the content – the author, exists no needs to work on that field.

For the two next technological revolutions – the genetic technology and nanotechnology – only the genetic technology is in the focus of the legal system, maybe because of the direct influence of our traditional ethic and moral values.

Nanotechnology is known most times for the semiconductor industry or for some simple visual effects, like the Lotus-effect. But IPRs would mainly be provided for electronic technologies – and this should be extended also on further technical fields.

References

[1] http://de.wikipedia.org/wiki/Integrierte_Schaltung download on: 2009-02-10.

[2] http://www.tf.uni-kiel.de/matwis/amat/mw1_ge/kap_2/illustr/i2_1_1.html download on: 2009-02-10.

[3] http://www.deutsches-museum.de/en/sammlungen/ausgewaehlte-objekte/meisterwerke-ii/mikroskop/ download on: 2009-02-10, Copyright © IBM.

[4] http://www.zdnet.de/itmanager/gallery/0,39030617,39134149-2,00.htm?PAGE_WIDTH=660 download on: 2009-02-10, Copyright © IBM.

[5] http://de.wikipedia.org/wiki/Lotuseffekt download on: 2009-02-10.

[6] Leon Radomsky: Sixteen Years After the Passage of the U.S. Semiconductor Chip Protection Act: Is International Protection Working? Berkeley Technology Law Journal, Bd. 15, 2000.

Measurement for Intelligent Control in Greenhouses

P. Eredics

Department of Measurement and Information Systems
Budapest University of Technology and Economics, Budapest, Hungary
Email: eredics@mit.bme.hu

Abstract. *The paper describes the design of an environment monitoring system supporting the development of black-box model based greenhouse climate control. Despite many currently available control schemes, only black-box modelling can be carried out efficiently for small or medium size greenhouses. For this large amount of measurements is needed, calling for a new distributed measurement system architecture. A novel system described in this paper collects measurements from a greenhouse in many measurement points, and provides an interface for the development and testing of model based control solutions.*

Keywords: Distributed Measurement System, Greenhouse Environment Monitoring

1. Introduction

Greenhouses have transparent walls and roof and are widely used for vegetable production and growing flowers. Sun radiation is essential for photosynthesis of the plants, and also to keep the inner temperature within an acceptable range. In the cold season a heating system may also be necessary. Contrary in hot weather other actuators, like roof vents, shading systems, exhaust fans or evaporative cooling may be used to avoid overheating. Almost every greenhouse today has a control system using some kind of environment control computer (ECC). Yet despite the computing power control methods did not change much along with the technology. Even today almost all greenhouse control systems work with independent, set-point based PID controllers [1], which suffer from the missing synchronization of the actuators, and their dependence on the user to find the appropriate set-points.

Model based control

The efficiency of the control can be improved using more advanced models, see e.g. IntelliGrow [2]. There is much of research on complex control models in greenhouse environment, e.g. model predictive control [1]. Such models could provide better control, but they require accurate estimates of plenty of parameters to work. To this end a large amount of measurements is needed to find the correct model of the house. Regrettably they are also expensive in the measurements and calculations. Besides, these methods cannot be used in smaller size industrial greenhouses. Such houses have usually low budget, are mostly uniquely designed to best fit the grower's needs, making it hard to find houses with identical structures. The houses typically grow several types, and often changing plants, which makes the model of even the similar houses different, and always varying. In low utilization periods, some houses can even be split into used and unused compartments, also changing the model.

All this makes it difficult to create a general yet useful analytic model. The alternative is a black-box model, necessarily a learning system. A learning system can identify the model on its own, and is also able to follow model changes. The price is a large number of measurements needed for learning and operation. Traditional control and ECC-s can be used to obtain measurements from a particular house. BipsArch system provides generalized interface for computer access to different ECC-s [3], making it possible to develop and run sophisticate control on separate PC-s, but it cannot overcome main limitations of the ECC. An

ECC requires set-points for the operation (does not accept direct commands for the actuators), and models the house as a single thermal entity. Closer examination of plenty of industrial greenhouses shows that the inner temperature is not homogenous and specific thermal zones appear related to the greenhouse structure and operation.

It is widely accepted, that improvement of the quality of greenhouse control systems can be achieved using complex models. However the mentioned limitations of current systems make it necessary to consider a new system design, described in this paper. The requirements of such a system are reliability, multipoint measurement in the house, remote access capabilities, openness, incremental functionality, low cost and a user friendly interface. It also should be mentioned that as the traditional ECC based control is able to operate the greenhouse alone if the model based control fails. This component should be also present in the system.

2. Pilot Study and Measurement System Architecture

The direct goal of the design was a data acquisition system for a 100 m^2 area greenhouse producing extremely sensible young plants in the western region of Hungary.

The Zones
Fig. 1 shows the simplified structure of the greenhouse, with thermal zones labelled. Traditional control system would make measurements in only one or some of these zones. To help accurate modelling of the house, data should be collected from every indicated zone.

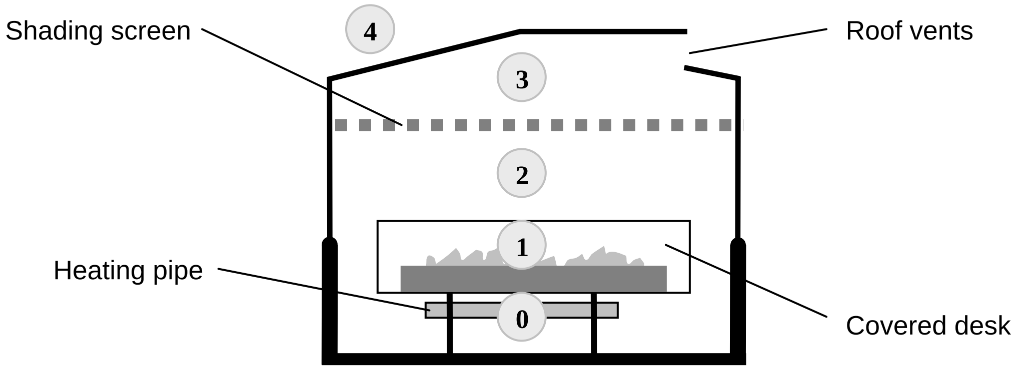


Fig. 1. Thermal zones (identified by their numbers) in a simplified greenhouse environment.

The current goal is to create the thermal model of the house, so we focus on the temperature and closely related radiation measurements. Zone 4 (external environment) is special because the control system can not change it. Zone 3 & 2 are separated only by an active shading screen. Zone 1 is the closest environment of the plants, in the current house usually inside veiled desks. Zone 0 (heating) is not a true thermal zone, but it can be best modelled as one.

Temperature measurements are collected in the current house from Zone 0 and 4 at a single point, from Zone 2 and 3 at two points, and in Zone 1 from all the desks. Two point measurements in Zones 2 and 3 are necessary, because the house is split sometimes into two sections. For larger, industrial size greenhouses more sections can exist, making more measurement points necessary. In the current house thus 21 temperature values, and two radiation values (from above and below of the shading system) are recorded.

Functional Architecture
The analysis of the requirements led to a layered system design, shown in Fig 2, corresponding to the gradual fusion of the measurement information. The 1st layer is the physical environment of the greenhouse. The measuring instrumentation and the actuators

form the 2nd layer. The 3rd layer contains sensor and actuator control implemented on local microcontrollers. This layer is responsible for collecting the measurements and directly controlling the actuators on site. 4th layer is implemented on a PC for recording and storing measurement data. This layer communicates with the microcontrollers and also serves higher layers, thus playing the role of a communication channel with storage capabilities. 5th layer contains the user interface, and here is the place of the model based high level control. Incorporating traditional ECC based control gives increased reliability to the whole system, and is implemented at the 3rd layer, running on a master microcontroller.

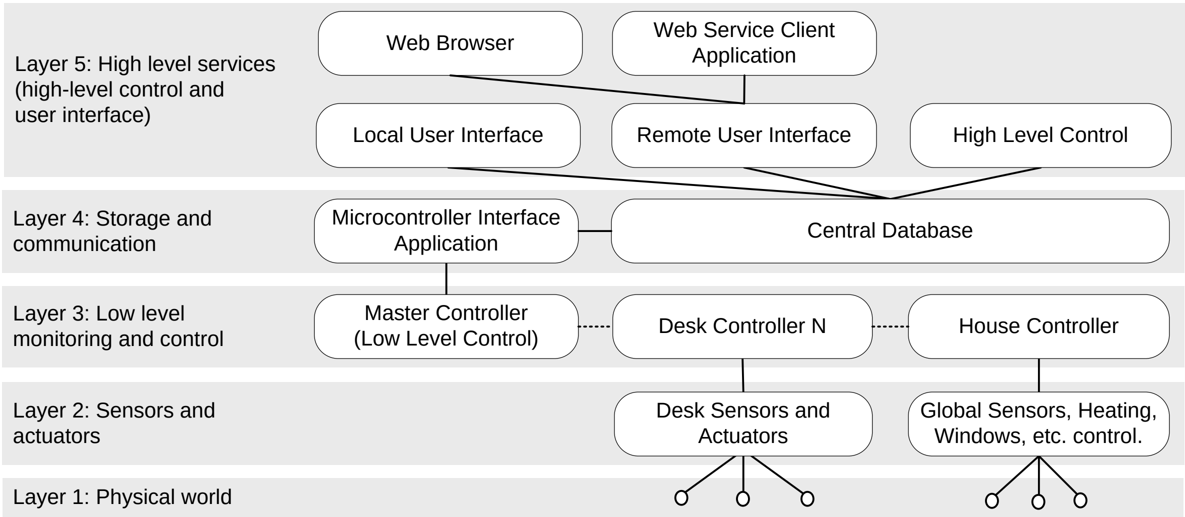


Fig. 2. System layers and the relation of the components implemented in the functional layers

Software and Structure

In the current implementation of the 3rd layer, we used ATmega microcontrollers: The master controller is running the low level control and communicates with a PC and the other parts of the controller network. Every four desks have a dedicated controller, for collecting local measurements, and controlling desk based actuators. (All desks have built in evaporating, and some desks have ventilating fans too.) There is another special microcontroller for controlling the global actuators of the house. The controllers communicate on an EIA-485 bus, while the master controller and the PC have an EIA-232 connection.

System Integration and User Interfaces

The central component of the higher layers resident on a PC is a MySQL database. The microcontroller interface application inserts periodically new measurements, and checks for configuration changes or commands for the lower layers. If the low level control's settings change or a command is issued, the master controller is informed. The central database can be directly accessed though a network connection from anywhere on the internet. It makes possible to develop and run the high level control on a dedicated computer at a distant location, e.g. in the owner's home environment. In such a setup the reliability of the high level control depends on the network conditions, but the reliability of the whole system is guaranteed in all cases by the low level control in the 3rd layer.

A modern greenhouse control and management system cannot exist without a user interface. The present system has multiple user interfaces to facilitate the interaction between the owner and the system. It has a simple display to show the current measurements in the greenhouse. To exercise remote management capabilities the system has a web based management interface. It contains a web server and a PHP web application. This design eliminates the need for a client application on the management computers. Measurement data is visualized with a Java applet in the browser window. Simple web services can be offered through the web server, making it possible to develop desktop applications for instant monitoring or alerting purposes. Fig. 3 shows a sample diagram from the implemented remote management system.

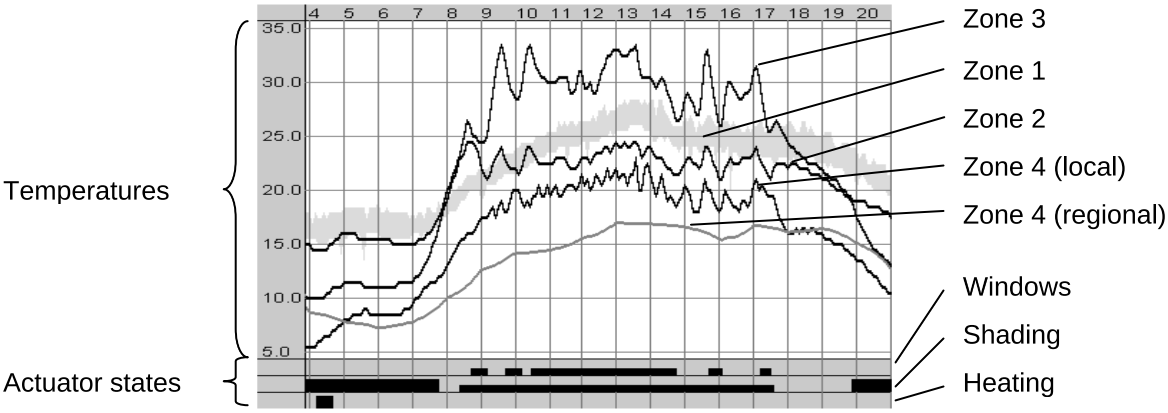


Fig. 3. Temperature data, recorded on 04-05-2008, visualized in the remote management webpage.

Data Processing and Intelligent Control

Measurements are continuously recorded from March 2008. During the winter 2009 the greenhouse held no plants for a month. In this period the goal of the control was to maintain the highest possible temperature with minimal use of costly heating. To utilize the most of the natural radiation a simple intelligent control application was developed. It could only control the shading screens. Control commands were computed by a decision tree with Zone 2 and 3 temperatures, and the trend of Zone 3 as inputs. This simple control application provided optimal control of the shading, and also verified the high level control interface.

3. Results

The system described above is aimed to serve as a supporting environment to develop (black-box) model based, high level control methods. It has three major benefits compared to other available solutions: it collects measurements in all important thermal zones of the greenhouse, it serves as an interface for direct access to the measured data and to control the actuators, and it can be simply accessed from remotely developed applications. The modular design of the system makes it possible to easily extend it with more sensors and microcontrollers, making it possible to monitor a much larger greenhouse without architectural modifications.

References

[1] M.Y. El Ghoumari, H.-J. Tantau, J. Serrano. Non-linear constrained MPC: Real-time implementation of greenhouse air temperature control. *Computers and Electronics in Agriculture,* 49: 345-356, 2005.

[2] J.M. Aaslyng, J.B. Lund, N. Ehler, E. Rosenqvist. IntelliGrow: a greenhouse component-based climate control system. *Environment Modelling & Software,* 18: 657-666, 2003.

[3] J.M. Aaslyng, N. Ehler, L. Jakobsen. Climate control software integration with a greenhouse environmental control computer. *Environment. Modelling & Software,* 20: 521-527, 2005.

Measurements of Performance in Laboratory IEEE 802.11 b, g Point-to-Point Links

[1,2] J. Pacheco de Carvalho, [1,2] P. Gomes, [1,2] H. Veiga, [2] C. Pacheco, [1] N. Marques, [2,3] A. Reis

[1]Centro de Informática, [2]U. de Detecção Remota-Departamento de Física,
Universidade da Beira Interior, Rua Marquês d'Ávila e Bolama,
6201-001 Covilhã, Portugal
[3]Departamento de Electrónica e Telecomunicações / Instituto de Telecomunicações,
Universidade de Aveiro, 3810 Aveiro, Portugal
Email: pacheco@ubi.pt

Abstract. *The importance of wireless communications has been growing. Performance is a very relevant issue, leading to more reliable and efficient communications. Laboratory measurements are made about several performance aspects of Wi-Fi (IEEE 802.11 b, g) point-to-point links. A contribution is given to performance evaluation of this technology, using available access points from Linksys. Detailed results are presented and discussed, namely at OSI levels 4 and 7, from TCP, UDP and FTP experiments: TCP throughput, jitter, percentage datagram loss and FTP transfer rate.*

Keywords: WLAN, Wi-Fi, Point-to-Point Links, Wireless Network Laboratory Performance Measurements

1. Introduction

Wireless communications are increasingly important for their versatility, mobility and favourable prices. It is the case of microwave based technologies, e.g. Wi-Fi.

The importance and utilization of Wi-Fi have been growing for complementing traditional wired networks. Wi-Fi has been used both in ad hoc mode and infrastructure mode. In this case an access point, AP, is used to permit communications of Wi-Fi devices with a wired based LAN through a switch/router. In this way a WLAN, based on the AP, is formed. Wi-Fi has reached the personal home, forming a WPAN, allowing personal devices to communicate. Point-to-point and point-to-multipoint configurations are used both indoors and outdoors, requiring specific directional and omnidirectional antennas. Wi-Fi uses microwaves in the 2.4 and 5 GHz frequency bands and IEEE 802.11a, 802.11b and 802.11g standards [1]. Nominal transfer rates up to 11 (802.11b) and 54 Mbps (802.11 a, g) are permitted. CSMA/CA is the medium access control. Wireless communications, wave propagation [2,3] and WLAN practical implementations [4] have been studied. Detailed information is available about the 802.11 architecture, including performance analysis of the effective transfer rate. An optimum factor of 0.42 was presented for 11 Mbps point-to-point links [5]. Wi-Fi (802.11b) performance measurements are available for crowded indoor environments [6].

Performance has been a very important issue, resulting in more reliable and efficient communications. Telematic applications have specific performance requirements, depending on application. New telematic applications present special sensitivities to performances, when compared to traditional applications. Application characterization and requirements have been discussed for cases such as voice, Hi Fi audio, video on demand, moving images, HDTV images, virtual reality, interactive data, static images, intensive data, supercomputation, electronic mail, and file transfer [7]. Several measurements have been made for 2.4 GHz Wi-

Fi [8], as well as WiMAX and high speed FSO [9,10]. In the present work further Wi-Fi (IEEE 802.11 b,g) results arise, through OSI levels 4 and 7. Performance is evaluated in laboratory measurements of point-to-point links using available equipments.

The rest of the paper is structured as follows: Chapter 2 presents the experimental details i.e. the measurement setup and procedure. Results and discussion are presented in Chapter 3. Conclusions are drawn in Chapter 4.

2. Subject and Method

The measurements used Linksys WRT54GL wireless routers [11], with a Broadcom BCM5352 chip rev0, firmware DD-WRT v24-sp1-10011 [12], and 100-Base-TX/10-Base-T Allied Telesis AT-8000S/16 level 2 switches [13]. The wireless mode settings included bridged access point, point-to-point, and default minimum transmitted power. Interference free communication channels were used in the communications. WEP encryption was not activated. No power levels above the minimum were required as the access points were very close (30 cm).

The laboratory setup is shown in Fig. 1. TCP and UDP experiments at OSI level 4, were as mentioned in [10], permitting network performance results to be recorded. For a TCP connection, TCP throughput was obtained. For a UDP connection with a given bandwidth parameter, UDP throughput, jitter and percentage loss of datagrams were obtained. TCP packets and UDP datagrams of 1470 bytes size were used. A window size of 8 kbytes and a buffer size of the same value were used for TCP and UDP, respectively. One PC, with IP 192.168.0.2 was the Iperf server and the other, with IP 192.168.0.6, was the Iperf client. Jitter, which represents the smooth mean of differences between consecutive transit times, was continuously computed by the server, as specified by RTP in RFC 1889 [14]. This scheme was also used for FTP measurements, where FTP server and client applications were installed in the PCs with IPs 192.168.0.2 and 192.168.0.6, respectively.

Batch command files were written to enable the TCP, UDP and FTP tests. The results were obtained in batch mode and written as data files to the client PC disk.

3. Results and Discussion

The access points were configured, for each standard IEEE 802.11 b, g, with typical fixed transfer rates. For every fixed transfer rate, data were obtained for comparison of the laboratory performance of the links, measured namely at OSI levels 4 and 7 using the setup of Fig. 1. At OSI level 1, SNR values and noise levels N were recorded, which are shown in Fig. 2. For each standard and every nominal fixed transfer rate, an average TCP throughput was determined. This value was used as the bandwidth parameter for every corresponding UDP test, giving average jitter and average percentage datagram loss. The main results, which did not significantly vary versus time of the day, are shown in Figs. 2-3. In Fig. 2, polynomial fits were made to the TCP throughput data. It is seen that the best TCP throughputs are for 802.11g. The average values are 13.9 and 2.9 Mbps for 802.11g and 802.11b, respectively. In Figs. 2 and 3, the data points representing jitter and percentage datagram loss were joined by smoothed lines. In Fig. 2, the jitter data show some fluctuations, being on average higher for IEEE 802.11b (2.1 ms) than for 802.11g (1.2 ms). Fig. 3 shows that, generally, the percentage datagram loss data (1.4 % on average) agree reasonably well for all standards.

At OSI level 7 we measured FTP transfer rates versus nominal transfer rates configured in the access points for the IEEE 802.11b, g standards. Every measurement was the average for a single FTP transfer, using a binary file size of 100 Mbytes. The results thus obtained are represented in Fig. 3. Polynomial fits to data were made for the implementation of each standard. It was found that the best performance was for 802.11g.

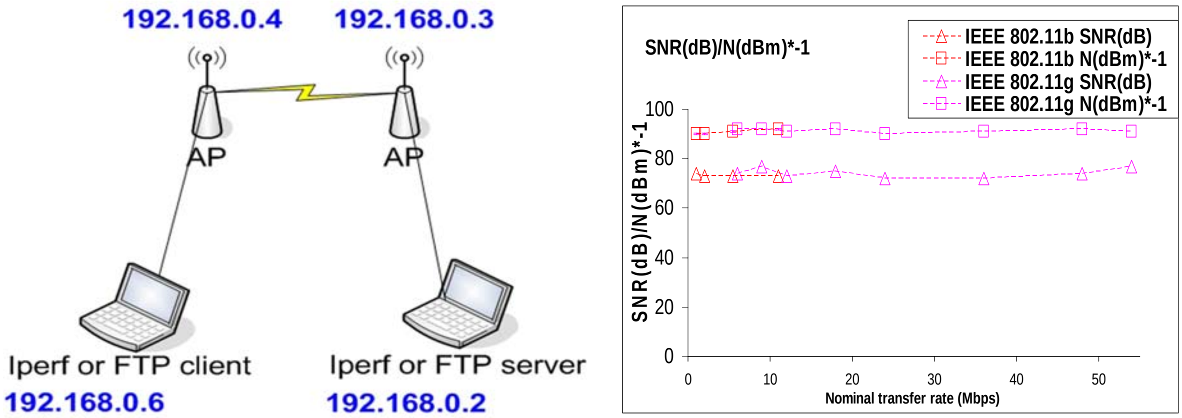


Fig. 1. Wi-Fi laboratory setup scheme and typical SNR (dB) and N (dBm).

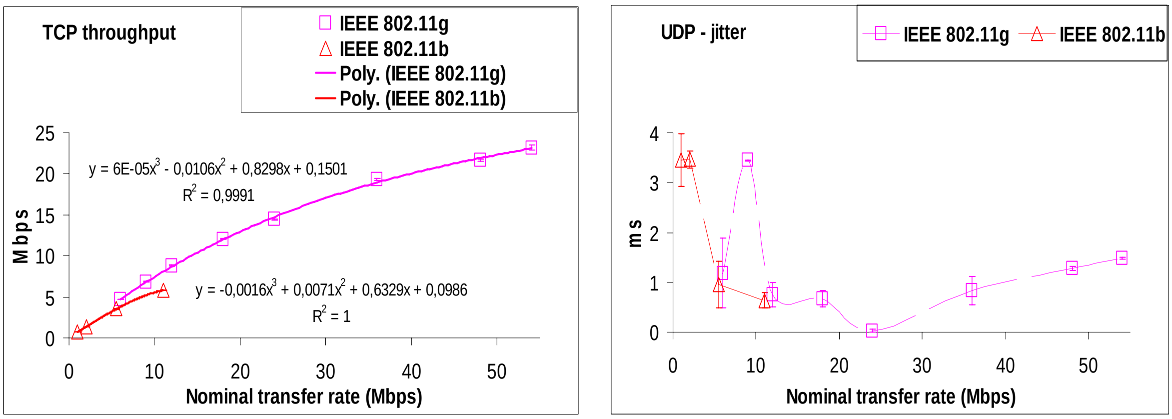


Fig. 2. TCP throughput and UDP jitter results, versus technology (IEEE 802.11 b,g).

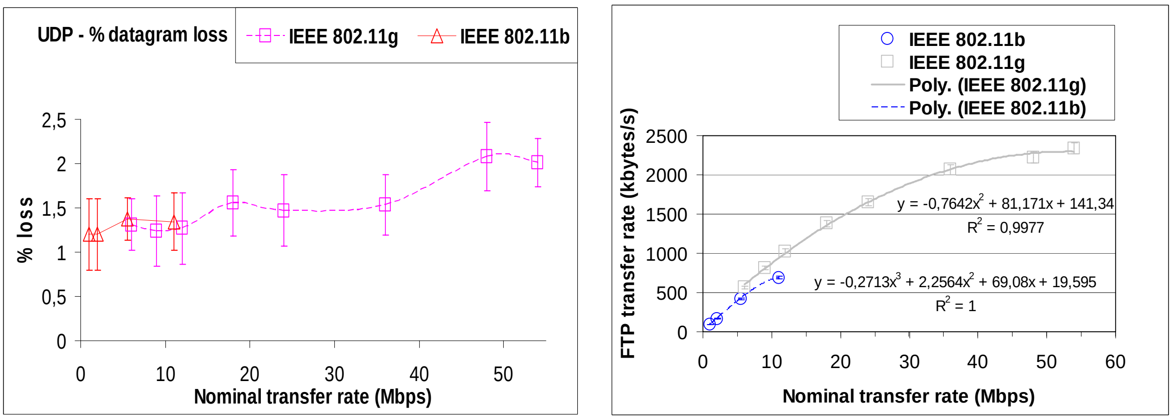


Fig. 3. Results of UDP percentage datagram loss and FTP transfer rate, versus technology (IEEE 802.11 b,g).

4. Conclusions

In the present work a simple laboratory setup was planned that permitted systematic performance measurements of available access point equipments (WRT54GL from Linksys) for Wi-Fi (IEEE 802.11b, g) in point-to-point links. Through OSI level 4, TCP throughput, jitter and percentage datagram loss were measured and compared for each standard. The best TCP throughput and average jitter were found for 802.11g. For the percentage datagram loss, a reasonably good agreement was found for both standards. At OSI level 7, the best FTP performance was for 802.11g. Additional performance measurements either started or are planned using several equipments, not only in laboratory but also in outdoor environments involving, mainly, medium range links.

Acknowledgements

Supports from Universidade da Beira Interior and FCT (Fundação para a Ciência e a Tecnologia)/POCI2010 (Programa Operacional Ciência e Inovação) are acknowledged.

References

[1] Web site http://standards.ieee.org

[2] Mark JW, Zhuang W. Wireless Communications and Networking. Prentice-Hall, Inc., Upper Saddle River, NJ, 2003.

[3] Rappaport TS. Wireless Communications Principles and Practice, 2nd ed., Prentice-Hall, Inc., Upper Saddle River, NJ, 2002.

[4] Bruce III WR, Gilster R. Wireless LANs End to End, Hungry Minds, Inc., NY, 2002.

[5] Schwartz M. Mobile Wireless Communications, Cambridge University Press, 2005.

[6] Sarkar N, Sowerby K. High Performance Measurements in the Crowded Office Environment: a Case Study, In Proc. ICCT'06-International Conference on Communication Technology, Guilin, China, 27-30 November 2006, 1-4.

[7] Monteiro E, Boavida F. Engineering of Informatics Networks, 4th ed., FCA-Editor of Informatics Ld., Lisbon, 2000.

[8] Pacheco de Carvalho JAR, Gomes P, Veiga H, Reis AD. Development of a University Networking Project, in Encyclopedia of Networked and Virtual Organizations, Putnik GD, Cunha MM, Editors. IGI Global, Hershey, PA (Pennsylvania), 2008, 409-422.

[9] Pacheco de Carvalho JAR, Veiga H, Costa R, Gomes P, Reis AD. A Contribution to Experimental Performance Evaluation of Point-to-Point WiMAX Links. In Proc. ISSPIT 2008-8th IEEE International Symposium on Signal Processing and Information Technology, Sarajevo, Bosnia and Herzegovina, December 16-19, 2008, 150-153.

[10] Pacheco de Carvalho JAR, Veiga H, Gomes PAJ, Ribeiro Pacheco CFFP, Reis AD. Experimental Performance Study of a Very High Speed Free Space Optics Link at the University of Beira Interior Campus: a Case Study. In Proc. ISSPIT 2008-8th IEEE International Symposium on Signal Processing and Information Technology, Sarajevo, Bosnia and Herzegovina, December 16-19, 2008, 154-157.

[11] Web site http://www.linksys.com

[12] Web site http://www.dd-wrt.com

[13] Web site http://www.alliedtelesis.com

[14] Network Working Group, "RFC 1889-RTP: A Transport Protocol for Real Time Applications", http://www.rfc-archive.org

Workstation for Automated Measurement of Electron Emission Characteristic of Samples Covered by Carbon Nanotubes

[1,3]J. Janík, [2]O. Čičáková, [1]J. Špakovský, [3]V.Csiba, [4]M. Čaplovičová

[1]Dpt. of Microelectronics, FEI STU, Ilkovičova 3, 812 19 Bratislava, SR.
[2]Dpt. of Measurement, FEI STU, Ilkovičova 3, 812 19 Bratislava, SR. [5]PF UK,
[3]BIONT, a. s. Karloveská 63, 842 29 Bratislava, SR
[4]PF UK, Dpt. of Geology and Mineral deposits, 842 15 Bratislava, SR
Email: olga.cicakova@stuba.sk

Abstract. *The paper deals with an automated current-voltage measurement of the cold electron emission characteristic. A measured sample was NI wire covered by a layer consisted of carbon nanotubes (CNT). The CNT were fabricated at Department of Microelectronics. The technology of preparation is briefly described. A compact high voltage converter (source of high voltage) controlled by a commercial digital low voltage power supply was used. Both the digital power supply and the multimeter were controlled by a PC through a USB/GPIB interface.*

Keywords: Automated Measurement, Field Emission, Carbon Nanotubes, Hot Filament CVD

1. Introduction

Carbon nanotubes (CNT) are one of known allotrope forms of carbon. The CNT were first described by Iijima [1]. Since that time a number of other kinds of nanotubes were studied and different properties of the CNT were found. Metalic CNT, semiconducting CNT, single walled or multiwalled CNT were studied. The CNT properties are chirality dependent [2]. Due to high length to diameter ratio an electric potential applied between cathode and anode creates great electric field in a vicinity of an individual CNT and cold electron emission occurs (cca 10^8 V/m). This effect can be exploated for cold electron emitters. Semiconducting CNT can be used for FET transistor production. Metalic CNT can be added to the polymers or sheets of papers to avoid their electrostatic adhesion properties. Also CNT are known as a hydrogen reservoar for vehicles, which (maybe in the future) replaces oil/petrol/alcohol/gas by the hydrogen. A heat conductivity of nanotubes is extremely anisotropic. It can be used for special coolers. The mentioned electron emission properties has a wide field of application. Because of the analogy of electron tubes known in the thirties, a micro-electron-tubes with the CNT cathodes, named as „vacutrons“, can be fabricated. They can be used up to terahertz frequencies. The displays known as a „CNT FED“ (Field Effect Displays) are very promissing. The CNT can be used also for an space propulsion vehicles [3].

2. Subject and Methods

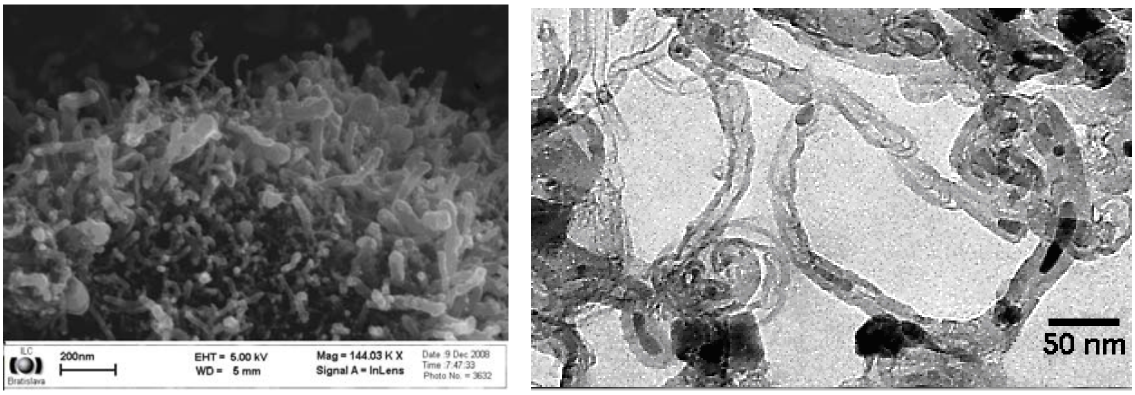
Apparatus for CNT preparation

Many kinds of preparations the CNT are known (arc, laser ablation, chemical vapour deposition -CVD-, electrolyse in organic liquids, etc.). We used a modified CVD technology. The apparatus (CVD reactor) was made after S. Bederka‘s design.

Nanotubes identification

For nanotube identification a few analyzing techniques are used. Most promissing and often used is an Scanning Electron Microscopy (SEM). Also a Transmision Electron Microscopy

(TEM) is convenient. The TEM is rather expensive, but a resolving power of the TEM is excellent. Atomic Force Microscopy (AFM) was made, but result is not shown here.



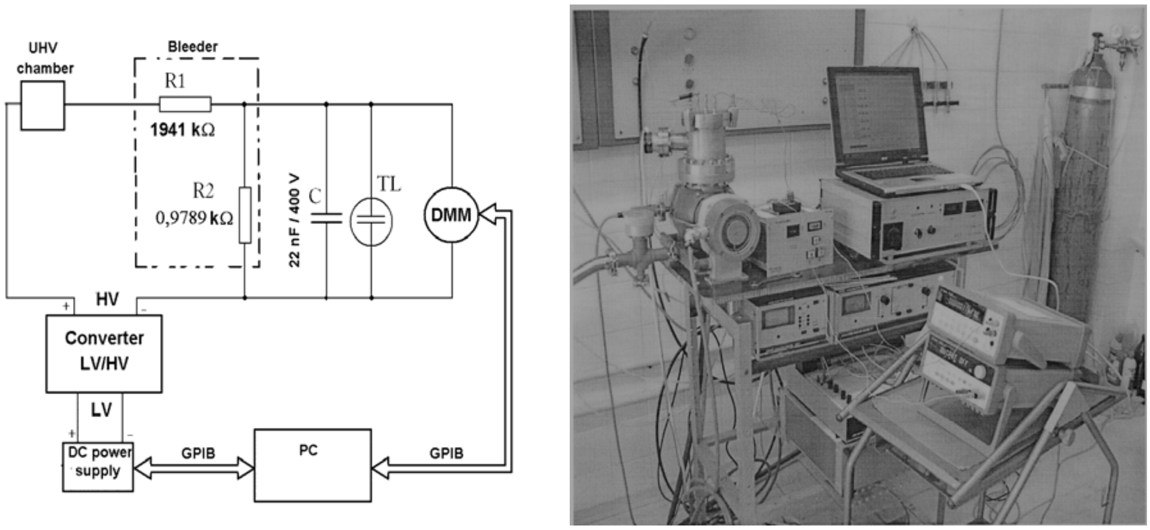
a) b)

Fig. 1a,b. SEM and TEM images of CNT.

Due to a special preparation of sample one can see (Fig. 1b) a structure of an individual nanotube [4]. For CNT identification the Raman spectroscopy is used very often. The Raman spectroscopy was used also in our case, but the results are not reported here.

Experimental measurement

The block diagram of the measurement setup is shown in the Fig. 2a. The view of the workstation is in Fig 2b. The workstation consists of ultra-vacuum chamber evacuated by a turbomolecular pump, low voltage/high voltage converter (LV/HV), digital multimeter, digital power supply, USB/GPIB interface and PC with MS EXCEL software.



a) b)

Fig. 2a,b. Block diagram and view of workstation for cold electron emission measurement

List of used devices:

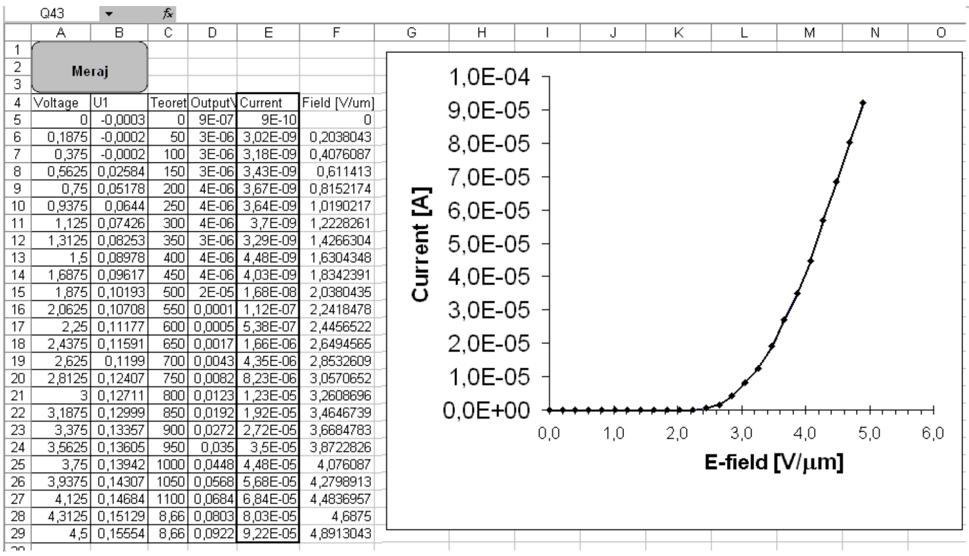
DMM - digital multimeter Agilent technologies 34401A
DC power supply - DC Power Supply Agilent technologies E3641A
GPIB - Agilent 82357B USB/GPIB Interface High-Speed USB 2.0
Converter LV/HV- converter Emco (Hivolt.de) F40 Series
TL - glow-discharge tube

The properties of AGILENT devices are available in [6]. The converter LV/HV has some properties published, but not its linearity. In spite of possibility of this converter to work up to input voltage 15 V, for our purposes the maximum voltage was only 4 V. To evaluate the linearity of the converter we measured its transfer function. Then the substitution line for input voltage 1.6 - 4.0 V was calculated. By this procedure we found out that the linearity error in this range was max. ±0.05 %.

Measurement Procedure

The measuring of electron emission is a macroscopic process. On the surface of Ni wire there are many emitting nanotubes. Some of them are partially damaged in the process of measuring. The long CNT are damaged primary because they are settled and burned in the presence of high electric field. Due to this fact, the measurement should be made as quickly as possible. It is not realizable by the manual measurement. The damage is great at the beginning of the measurement. After some time the current has character of saturation. After our experience the saturation current after some hours reaches about 70 percent of the beginning values. The automation of the measurement gives us a possibility to quantify better the saturation process due to the shorter sampling time compare with the manual measurement. Due to the accuracy of used measuring devices the estimated precision of the measurement is better then 0.1 %. Because of the CNT damaging process the change of the emission current is much worse than mentioned precision.

3. Results



	Voltage	U1	Teoret	OutputV	Current	Field [V/um]
5	0	-0,0003	0	9E-07	9E-10	0
6	0,1875	-0,0002	50	3E-06	3,02E-09	0,2038043
7	0,375	-0,0002	100	3E-06	3,18E-09	0,4076087
8	0,5625	0,02584	150	3E-06	3,43E-09	0,611413
9	0,75	0,05178	200	4E-06	3,67E-09	0,8152174
10	0,9375	0,0644	250	4E-06	3,64E-09	1,0190217
11	1,125	0,07426	300	4E-06	3,7E-09	1,2228261
12	1,3125	0,08253	350	3E-06	3,29E-09	1,4266304
13	1,5	0,08978	400	4E-06	4,48E-09	1,6304348
14	1,6875	0,09617	450	4E-06	4,03E-09	1,8342391
15	1,875	0,10193	500	2E-05	1,68E-08	2,0380435
16	2,0625	0,10708	550	0,0001	1,12E-07	2,2418478
17	2,25	0,11177	600	0,0005	5,38E-07	2,4456522
18	2,4375	0,11591	650	0,0017	1,66E-06	2,6494565
19	2,625	0,1199	700	0,0043	4,35E-06	2,8532609
20	2,8125	0,12407	750	0,0082	8,23E-06	3,0570652
21	3	0,12711	800	0,0123	1,23E-05	3,2608696
22	3,1875	0,12999	850	0,0192	1,92E-05	3,4646739
23	3,375	0,13357	900	0,0272	2,72E-05	3,6684783
24	3,5625	0,13605	950	0,035	3,5E-05	3,8722826
25	3,75	0,13942	1000	0,0448	4,48E-05	4,076087
26	3,9375	0,14307	1050	0,0568	5,68E-05	4,2798913
27	4,125	0,14684	1100	0,0684	6,84E-05	4,4836957
28	4,3125	0,15129	8,66	0,0803	8,03E-05	4,6875
29	4,5	0,15554	8,66	0,0922	9,22E-05	4,8913043

Fig. 3. Measured and calculated values and calculated graph.

The I-V characteristic of Ni wire (cathode) covered by CNT was measured. The sample was prepared after [5]. As an example of automated measured data there is shown a sheet „xls“ of EXCEL in Fig. 3. There are the measured values and drawn graph of emitted current versus electric field. In the chamber anode-cathode electrodes have a cylindrical symmetry. The mean value of the electric field on the cathode was calculated by (1)

$$E = \frac{V}{r_1 \ln \frac{r_2}{r_1}} \tag{1}$$

where

r_1, r_2 radius of cathode and anode

V anode potential.

4. Conclusions

An automated workstation for measuring the I-V characteristics of described CNT was realized. The emission current was involved by an field emission in the strong electric field. The measured dependance of current versus voltage was recalculated to the dependance of current versus electric field. The graph of this dependance was immediately drawn. From the measured values other interesting data (treshold voltage, work function, etc.) can be extracted. Next advantage of automated measurement setup is that no special cards for PC are needed. So the notebook can be used (as it was in our case). A big asset of the automated measurement is shorter measuring time. More, the measuring process avoids subjective mistakes. Compared with the measurements made manually some quantification of damaging process of CNT can be better estimated. The whole measuring process is more effective than it was before. Also there is no need of permanent presence of operator during long time measurements (some days). An optimal sampling time remains an open question, but this will be a task for further work.

Acknowledgements

The CNT preparation were made at Centre Excelencie „CENAMOST“ (Agency for supporting of science an research, Contract No. VVCE-0049-07) and projects VEGA1/0807/08, VEGA 1/0337/08 and VEGA 1/0551/09. The authors are indebted to Mr. M. Michalka from ILC for SEM images.

References

[1] Iijima S. Helical Microtubes of Graphitic Carbon, Nature 354 56-58 1991.

[2] Dresselhaus M. S, Dresselhaus G, Avouris Ph. Carbon Nanotubes: Synthesis, Structure, Properties and Application. Springer – Verlag Berlin, Heidelberg 2001.

[3] Technical Digest of the 18th International Vacuum Nanoelectronisc Conference. Editors Huq SE, Wilshaw PR. St Catherine's College 10 – 14 July 2005, Oxford, UK.

[4] Caplovicova M, et al.:An Alternative Aproach to Carbon Nanotube Samp. Preparation for TEM Investigation, Ultramicroscopy 107, 8, 2007, 692-697.

[5] Spakovsky J., Diploma thesis 2008 (Superwiser Janik J.), KME FEI Bratislava SR.

[6] http://www.home.agilent.com.

Monitoring of Detector Structures at the CERN ALICE Experiment

F. Lackner[1], W. Riegler[1], H. Kopetz[1], P. H. Osanna[2], M. N. Durakbasa[2]

[1]CERN, European Organisation for Nuclear Research,
CH-1211 Geneva-23, Switzerland
[2]Department of Interchangeable Manufacturing and Industrial Metrology,
Vienna Univ. of Technology (TUW), Karlspl. 13/3113, A-1040 Vienna, Austria
Email: Friedrich.Lackner@cern.ch

Abstract. *One of the four CERN LHC (Large Hadron Collider) experiments is ALICE, short for "A Large Ion-Colliding-Experiment", a detector consisting of multiple layers of sub detectors around the collision point to detect different types and properties of particles created in the collision [1] [2] [3]. The sub detectors are held in place by a structure called 'Spaceframe'. It has a cylindrical 18 - sided geometry with a length and diameter of 8 m. The weight of all the sub detectors, about 80 tons on the whole, will deform the Spaceframe in radial direction. For obvious reasons these deformations have to be monitored. This paper presents results from the deformation analysis during a magnetic field test and the installation of detector modules.*

Keywords: Laser Alignment, Detector Monitoring, BCAM, Structure Alignment

1. Introduction optical deformation monitor

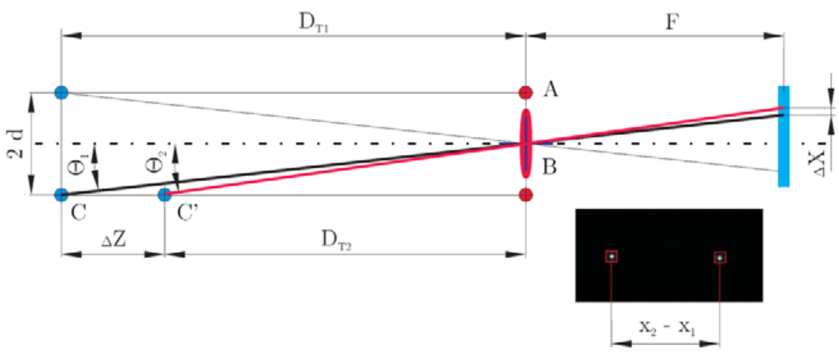


Fig. 1. Principle of the BCAM Operation

The BCAM is a simple optical device which has been developed by Brandeis University for the LHC ATLAS experiment End - Cap Alignment system [10]. The principle of Measurement is presented in Figure 1. A BCAM consists of an electronic camera and a pair of light sources (see Fig. 1. A, B), all integrated into a single enclosure kinematically mounted on three steel balls. The camera contains a CCD (Charge Coupled Device) image sensor and a lens with a focal length F of 75mm. Its field of view is 40mrad horizontally and 30mrad vertically to its mounting plane. The CCD provides an array of 344 by 244 pixels, serving as a two - dimensional coordinate system. A pixel measures 10µm square. The light sources of the BCAM are red laser diodes, treated as being point - like. Each laser transmits at $\lambda = 650$nm in a rectangular cone that measures 75mm by 25mm on a screen at a distance of 100mm from the BCAM. Lasers and CCDs can be controlled via an RJ - 45 sockets. The centers of the steel balls on which the BCAM is mounted in a local BCAM - coordinate system. All relevant BCAM - parameters, e.g. the center of the lens or CCD rotation, can be related to this local coordinate system by a calibration procedure. Thus one only needs to know the position of the centers of the steel balls in a global coordinate system to know the position and orientation of the BCAM - coordinate system. Figure 1 shows the principle of a standard BCAM alignment based on the example of a longitudinal movement. As can be seen, a small displacement ΔZ is limited by the angular BCAM resolution of 5 µrad defined by $\varphi_2 - \varphi_1$ according to the following equation (1), [6]:

$$Z = \frac{D_T^2\,\varphi}{d + D_T\,\varphi} \approx \frac{D_T^2\,\varphi}{d} \qquad (1)$$

DT	Distance between lens pivot point of the BCAM A and the light source of BCAM B
φ	Angular resolution of the BCAM system in µrad
d	Distance between the two light sources
Z	Accuracy

2. The ALICE Spaceframe

The monitoring system of the ALICE spaceframe structure has to determine all the corner positions to an accuracy of better than 500µm. The BCAM system is implemented to the Spaceframe by fixing a mounting plate holding two BCAMs on each corner. The 18 corners of the Spaceframe show relative movements within a few millimeters for the entire detector load. Measuring the relative angles of all BCAMs the 18 internal angles of the space frame are monitored. In summer 2005 a load test was performed to verify the expected deformations [7]. Different load conditions were measured by the SMS and with external survey done by the CERN survey group [4]. Both results were compared and showed that the SMS could reconstruct the positions of the 36 vertices on both sides of the Spaceframe for every load condition within the expected 500µm of their actual position.

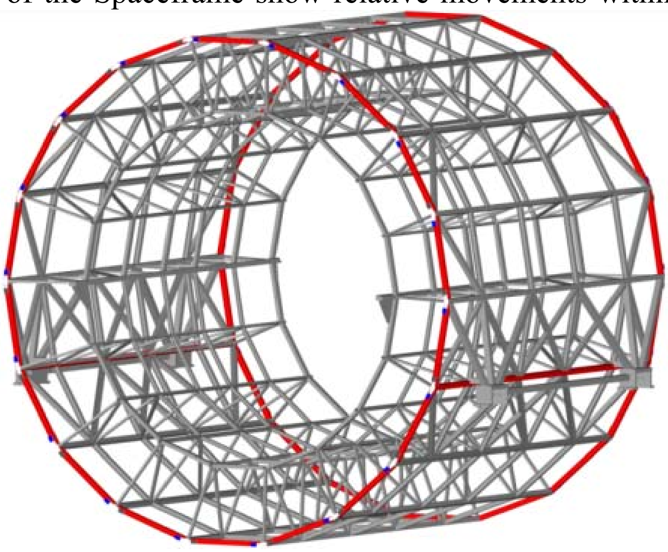


Fig. 1. ALICE Spaceframe - diameter: 8.5 m, length: 7 m, weight: 10.5 tons; Monitored surfaces are highlighted in red

3. Magnetic Field Mapping

After the primary installation and load test verification in the installation hall, the frame was transferred to the cavern and installed into the L3 magnet in early 2006 [6]. After this installation a magnetic field tests was performed in order to verify the field stability [11]. During this period it was decided to operate the spaceframe monitoring system verifying deformation levels due the field influence. In order to calculate and reconstruct the deformed shape of the spaceframe, a least square adjustment was implemented [12]. The Least Square Adjustment method was implemented in a program based on C++ and Root [13]. All calibration constants of the BCAMs and zero measurements of coordinates are stored in external files accessed by the program during runtime.

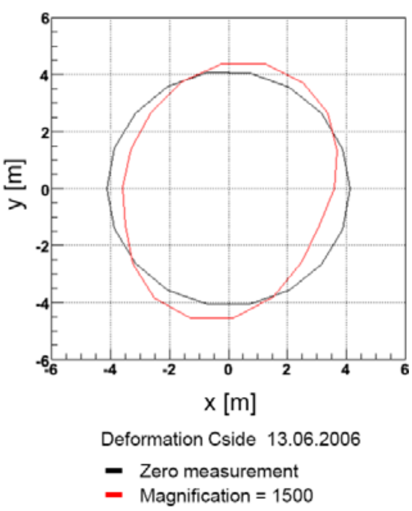


Fig. 2. Spaceframe deformation due to magnetic field influence (exaggerated)

Figure 2 presents an exaggerated plot for reconstructing the spaceframe deformation during the described magnetic field test. As can be seen, the fit represents one face of the spaceframe. The start up of the ALICE magnet to the nominal current of 30000A (0.5T) showed a maximum deformation of 0.9mm vertically and 0.86mm horizontally. The shutdown of the magnet results in a residual deformation of 0.3mm in the vertical and 0.35mm in the horizontal direction. A first analysis of the resulting deformation of the spaceframe due to magnetic field influence has shown that the iron joke deformation is transmitting deformations to the spaceframe support rails. In order to monitor the spaceframe rail deformation an additional monitoring system based on a BCAM – retroreflector application will be installed [5].

Table 1. Vertical and horizontal elastic deformation due to the magnetic field influence

	Solenoid current [A]	ΔV [µm]	ΔH [µm]
02.06.2006	0	0	0
08.06.2006	0	1.7	3
13.06.2006	30000	902.8	860.2
14.06.2006	0	307.4	343.3
20.06.2006	0	303.5	350.6

4. Detector Installation

In early 2009 one module of the Transition Radiation Detector (TRD) was installed into the Spaceframe (Fig. 3.). In parallel, the entire data acquisition was implemented to the detector control system which allows to the read measurement system remotely. The analysis of the raw data taken before and after the detector installation showed that the spaceframe was deformed according to the exaggerated blue curve in Fig. 3.

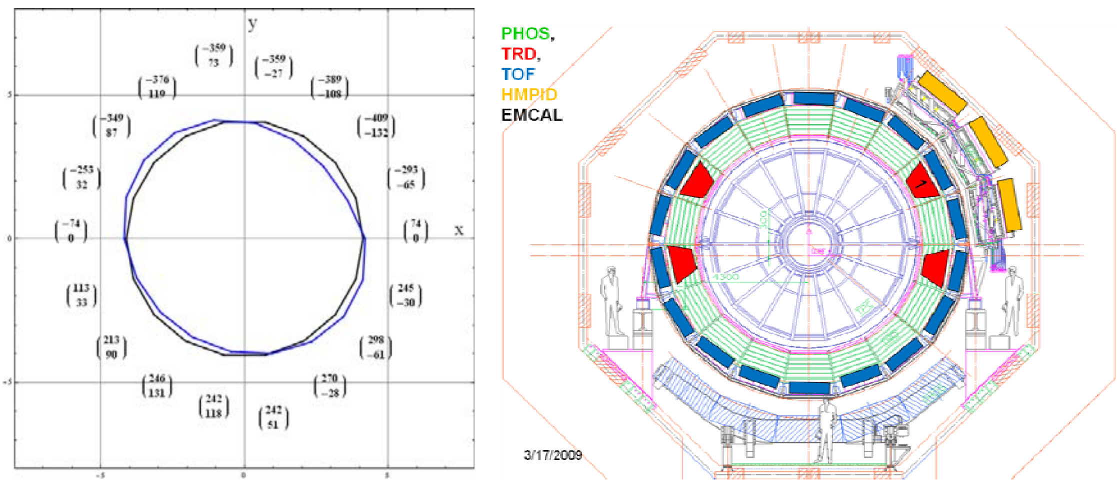


Fig. 3. Left: Deformation due to the TRD module installation. The black curve represents the zero measurement whilst the blue curve shows the deformation after the installation. The curve is drawn with an exaggeration factor of 1000. Right: ALICE Detector layout; (1) installed detector module.

As can be seen, the installation of the 1300kg detector module deforms the spaceframe structure locally by a maximum value of 132µm in vertical and 400µm in the horizontal plane. The implementation and automatic analysis of the monitoring system via the detector control system will allow encountering critical deformations during delicate detector operations in shutdown periods. Furthermore the system will give additional information during the entire life cycle of the ALICE experiment.

5. Discussion and Conclusion

The spaceframe detector structure in the CERN LHC ALICE experiment has to be monitored during the entire lifecycle. This is partly given by the required high precision and the tight clearances in order to install and handle detector structures, but also because the status and long term mechanical stability of the entire experiment is depended on the spaceframe structure. The paper showed first results of the analyzed raw data taken during a magnetic field test and the installation of detector modules. The successful measurements during the magnetic field test showed a deformation of 0.9mm vertically and 0.86mm horizontally on one spaceframe surface. The installation of one TRD module deformed the spaceframe elastically by 132µm in vertical and 409 µm in the horizontal plane. Currently the integration of the monitoring system into the ALICE detector control system is near completion. The implementation of further monitoring systems based on the optical device is foreseen in order to get more detailed information of deformations and small movements of the spaceframe with respect to an external reference.

References

[1] CERN ALICE Collaboration, Technical Proposal for A Large Ion Collider Experiment at the CERN/ LHC, CERN/LHCC/95-71LHHC/P3, 15 December 1995

[2] CERN, Large Hadron Collider Project General Information. http://lhc.web.cern.ch/lhc/general/gen_info.htm

[3] ALICE Collaboration, Physics Performance Report, Volume I, October 2004

[4] Behrens A., Lasseur C., Positioning Strategy, Metrology and Survey in ALICE, November 2007, CERN TS-SU-EM – EDMS 884850

[5] Lackner, F., Riegler, W., Kopetz, H., Osanna, P. H.: A Novel Solution for Various Monitoring Applications at CERN, May 2005

[6] Lackner F., Design and High Precision Monitoring of Detector Structure at CERN;

[7] PhD Thesis, Vienna University of Technology, 2007

[8] Kopetz H., A novel method of precision surveying of large particle detectors.

[9] Master thesis, University of Technology, Graz, 2005.

[10] K. Hashemi, "BCAM User Manual (Version 11)", Brandeis University, Boston (October 2004), Open Source Instruments, www.opensourceinstruments.com, BCAM drawings, reports, circuit diagrams and manuals are distributed for free under the GNU Public License, copyright Kevan Hashemi.

[11] K. Hashemi, "LWDAQ User Manual (Version 3.2.2)", Brandeis University, Boston (December 2004)

[12] C. Amelung, J.R. Bensinger, F. Cerutti, C.W. Fabjan, K. Hashemi, S.Palestini, J. Rothberg, A. Schricker and I. Trigger. Reference Bars for the Alignment of the Atlas Muon Spectrometer. http://Sciencedirect.com, 2005.

[13] J.N. Joux A. Behrens V. Prasad, N. Mietakhov. ALICE - Magnet Test, Measurement on the Spaceframe, Measurement on the Muon area 15th June 2006. Technical report, June 2006.

[14] G. Schlager. Least Square Adjustment for the ALICE Spaceframe Monitoring System. CERN, 2006.

[15] ROOT, An Object-Oriented Data Analysis Framework. Uder Guide 5.21, December, 2008 http://root.cern.ch

[16] CERN ALICE Collaboration, Technical design report TRD Detector , 2004-009 v.1

A Novell Setup of Electronic Techniques for Solving Asynchronous High Speed Illumination in High Speed Measuring

A. Göpfert, S. Lerm, M. Rosenberger, M. Rückwardt, M. Schellhorn, G. Linß

Technical University of Ilmenau, Ilmenau, Germany,
Email andre.goepfert@tu-ilmenau.de

Abstract. *With the demand in shorter process cycles in the production, the requirements for light pulses increase. This applies in particular to build stable and high intensity pulses with good reproducibility. With the help of the freezing the picture, the moving object can be almost independent of disturbances recorded. The aim of the project was a digitally controlled lighting control with regard to design specific criteria.*

On the market available flash and lighting controls were investigate the suitability of a particular application. Given the low fitness, theoretical and practical methods were used for providing a regulated current source. On this basis, a circuit design and board layout prepared. The prototype was made in preview to the use of programming. The programming took initialization routines between several microcontrollers and electronic components. After programming, practical tests were carried out and the desired properties investigated. The tests were positive and a manifold useful.

Keywords: Lighting Control, Flash Control, Strobe Controller

1. Introduction

Machine Vision is a wide area of parameters which can be affected. But also there are problem specifications, which are given and only can be optimized with other components. One parameter can be controlled is illumination.

Illumination is an inherent part in machine vision, because there is no image you can grab or analyze without it. Finding right adjustments of illumination for a device under test is not easy. Solving the measuring problem gets more difficult if the sample is in motion.

The resulting motion can be an effect from free fall, transportation mechanics (conveyer) or actuators, which generate vibrations to the system.

High precession measuring of geometrical variables is affect of these motions. To generate a snapshot, the camera-integration-time must be very low in dimensions of microseconds. The snapshot from a short moment is necessary, because the distance by time will be affect motion blur effects. A new addicting result problem of short integration-times is lower sensitivity of the camera-sensor. To get although object-information it is necessary that the light intensity must be high in a short time.

A sample of necessity is real time measurement in the production cycle of screws. The measurement at own is under high vibrations and parasitic vibrations. In the moment of measure the screw is on the fly. Summarized in this special measurement situation the camera is under vibration and the screw is in the air. Those are two parameters that provide a fail of a standard measurement.

2. Requirements at the device

The light will be generated by LED. There are two methods to control LED. One method is to control voltage and another is to control current. It came to the decision of current-control, because light intensity can be different at the same voltage and different temperatures.

After the background to generate an electronic system to control, adjust and regulate the illumination in machine vision. The Requirement specifications are the following. This section describes the format for paragraphs, figures, tables, equations and references.

Table 1: List of specifications

Parameter	Specification
Channels	2
Voltage IN & Out	24V
Current pulsed	Adjustable from 0 to 10 A
Current continuous	Adjustable from 0 to 2A
Step size	1mA
Timer	0 to 1s
Step size	1µs
Delay	<20µs
Parameter	Specification

To reach the specifications an elaborate electronic design must be created. The basic structure contains different integrated circuits – each for a separate job. For example there are microcontrollers, high current control circuits, USB-Controller, decoupled inputs, lots of filter circuits, but no current-measure-parts. To reach this short operation times a measuring system will be inert. There are different ways to measure the current with voltage above a shunt resistor or a contactless hall sensor. But this method is time inefficient. It would be an addition about test times of measuring voltage, converting into voltage, calculate new and settling the parameters of DAC.

3. Creating a structure of components

At the beginning there were ideas of a control-structure. The ideas have a range from one microcontroller for all to one for each channel or substitute microcontrollers for CPLD or FPGA. The founded concept is based on short reaction-times, after a trigger-signal. The concept is a combination of one master microcontroller and one microcontroller for each part. The matters for microcontrollers are low prices and a lot of functions. Fig.1 shows the simplified concept.

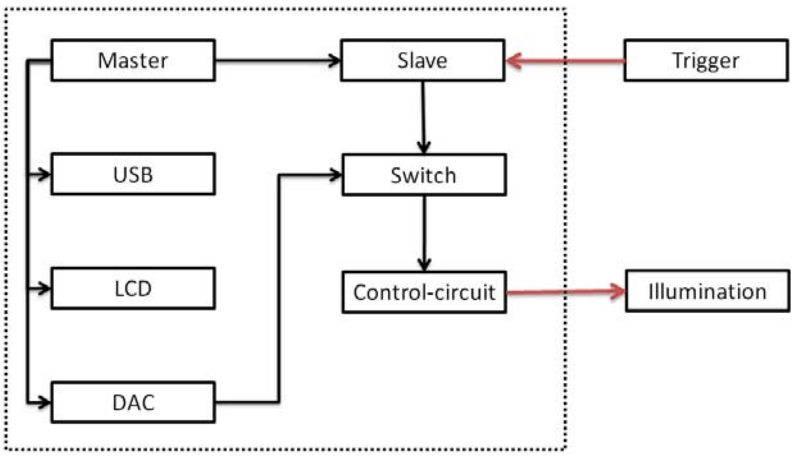


Figure 1: Simplified structure-design of global components [1]

Theoretical founded control circuit

The control circuit with his elements is special in its self. There are hands of methods to control and adjust a current. As example there are control circuits with fixed voltage regulator, transistor in a lot of combinations like current mirror, FET, OPA and combination of FET or transistor with an OPA.

After careful consideration of requirements and circuit possibilities, at the end the decision was falling at a combination of OPA with FET or transistor. The structure is visualized in Fig. 2.

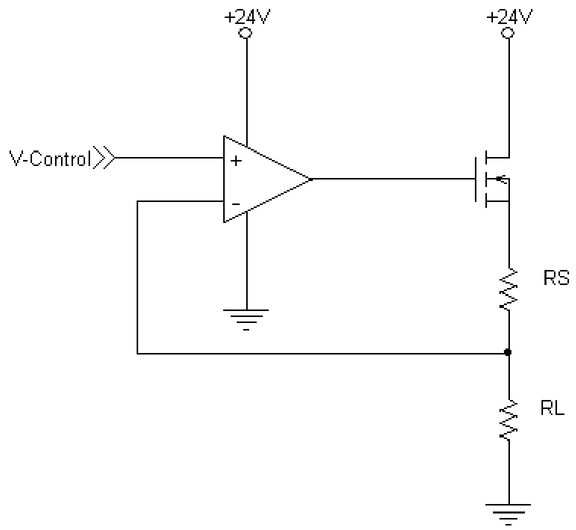


Figure 2. Simplified current-control-design of OPA and FET [2]

In Fig. 2 the resistor named RL is the light source and RS is a resistor for sense, also named shunt. Simplified the output can be described by developing formula (1) to (2) [3].

$$R = \frac{U}{I} \tag{1}$$

$$I = \frac{V_{Control}}{R_S} \tag{2}$$

where

R	resistance
U	voltage
I	current

Build-Up and practical tests

Building Up was made on a double-sided circuit board, illustrated in Fig. 3. Simple failures can be detected and solved. After placing and soldering there was the part of programming. Programming was made in a development environment in programming language C.

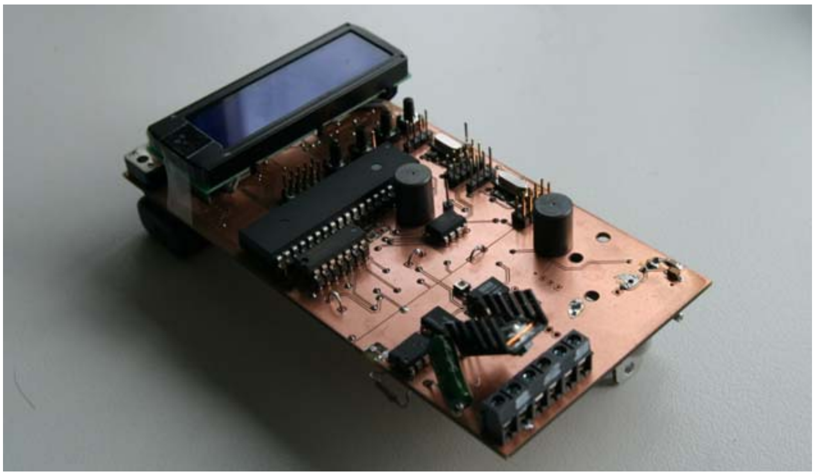


Figure. 3. Simplified structure-design of global components [4]

In aspect of practical tests there were tests about aspects like reaction time, time accuracy, time resolution, current accuracy, current resolution and specially slew rate.

4. Conclusions

Actually the prototype is well operating for high speed illumination. There are high precision adjustable timer and precision adjustable currents. All this parameters are tested at industrial test conditions.

High speed illumination can be used in machine vision systems. Because all parameters are adjustable, it can be adapted to every setup.

In the future are higher voltages or a higher number of channels able. The number of channels can be up to 255, because every channel is independent to the other.

References

[1] Göpfert A, Untersuchungen zum Einsatz verschiedener Hochstromquellen bei synchronem Bildeinzug per Hard- und Softwaretriggerung. Ilmenau, 2008.

[2] Tiezte U, Schenk Ch, Springer, Berlin, 2002.

[3] Hoffmann W, Hochfrequenzstechnik, Springer, Berlin, 1997

[4] Michel T, Optimierung eines Prototypen zur Blitzbeleuchtung in Hochfrequenzbild-aufnahmesystemen, Ilmenau, 2008.

Test Bench for Electrical Drives

V. Hubík

UNIS a.s., Mechatronic System, Brno, Czech Republic,
Email: vhubik@unis.cz

Abstract. *Testing equipment is an integral part of every development department. This article presents the bench for testing and measurement of electrical drives parameters and their control algorithms. With its design and versatility, it is different from known commercial products such as various dynamometers, mechanical brakes, etc. Control of the whole measurement process is realized by the EVA processor system (designed by UNIS, a.s. as well), which provides detailed display on the integrated graphical LCD unit. Maximal parameters of tested electrical drives are up to 40 000 rpm and up to 2 N·m torque.*

Keywords: BLDC Motor Test Bench, Electrical Brake, Mechanical Brake, EVA Processor System, CAN Communication Bus

1. Introduction

The described measuring system, designed by UNIS, a.s. [1], is mainly intended for the testing of the electrical drives' mechanical characteristics and their control algorithms. The prime area of interest in the Mechatronic systems division is the development of control electronics and software algorithms for driving sensor and sensor-less electronic commutated motors (BLDC motors) in the power range up to 1.5 kW and in the voltage range up to 48 V DC. Usage of these electrical drives is mainly in the aerospace industry and other safety critical applications.

The new test bench has been created on the basis of these requirements. Its block schematic is in Fig. 1 and graphical model in Fig. 2. For easier orientation it can be separated into several parts – mechanical brake, electrical brake, processor control system, graphical and operating panel. Individual parts are connected to a common CAN bus, which controls whole test bench, data collects and displays the actual physical values on an LCD panel.

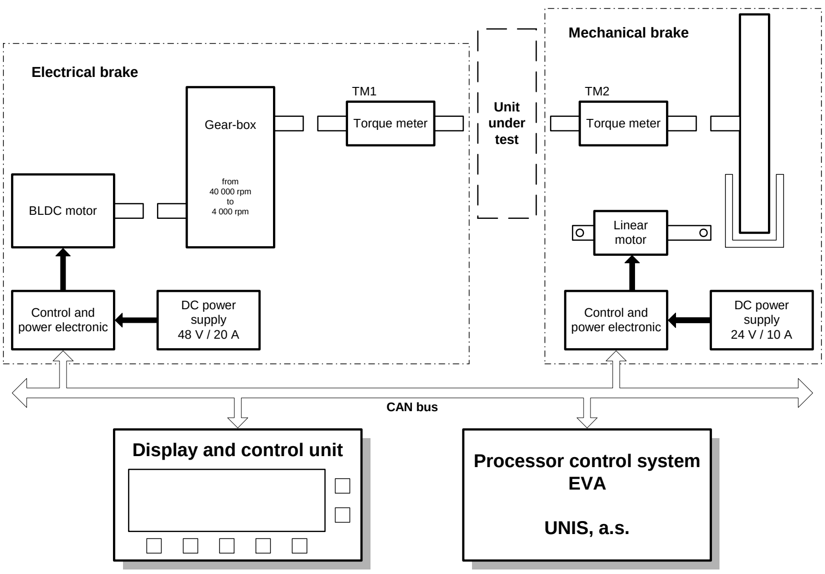


Fig. 1. Block schematic of the test bench.

The following chapters contain a more detailed description of particular part of the test bench, their parameters, and possibilities. Graphical model of several parts, e.g. Fig. 2, created in AutoDesk Inventor environment is also enclosed.

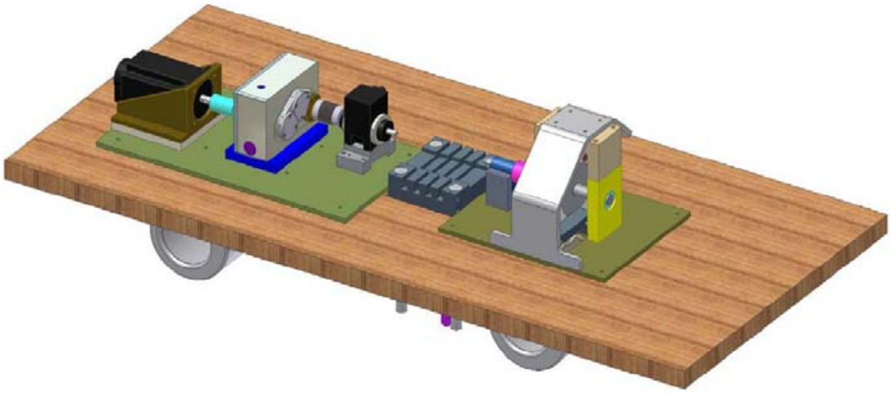


Fig. 2. Graphical model of the whole test bench.

2. Electrical Brake

The electrical brake, Fig. 3, consists firstly of the torque meter (the first black box in Fig. 3), which is a product of Magtrol, with maximal parameter range up to 40 000 rpm and up to 2 N·m torque. The next component is the two-stage gear-box, each stage having the same gear ratio. The gearbox consists of helical gears immersed in oil for permanent lubrication and cooling. The gear ratio of the whole gearbox is 1 to 10 and maximal transferred power up to 1.5 kW. The gearbox output is connected to the BLDC motor (black one, Fig. 3) with a nominal power of 1 kW. Its principle is similar to that of the active electric brake and produces the negative torque to tested device. The braking torque, thanks to smart control, is continuously variable over a whole range according to requirements. Information about torque and angular speed are available from the torque meter in analog values that are converted to digital form and processed in the processor system.

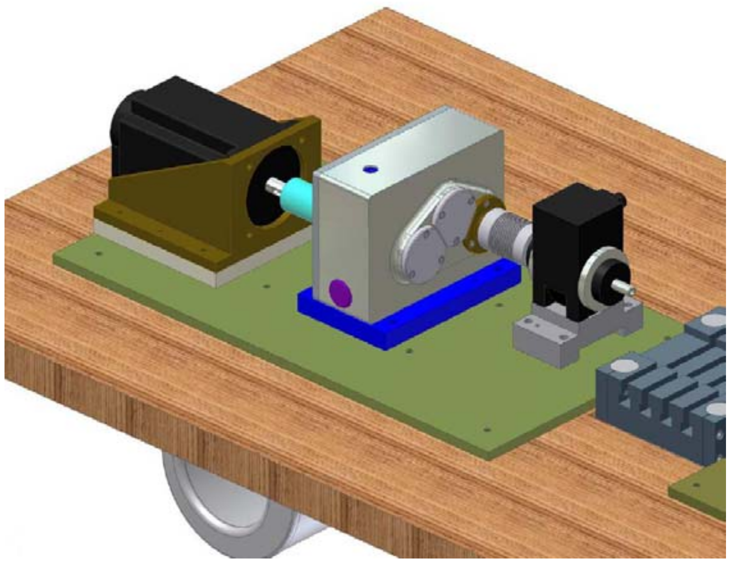


Fig. 3. Model of electrical brake, part of the test bench.

3. Mechanical Brake

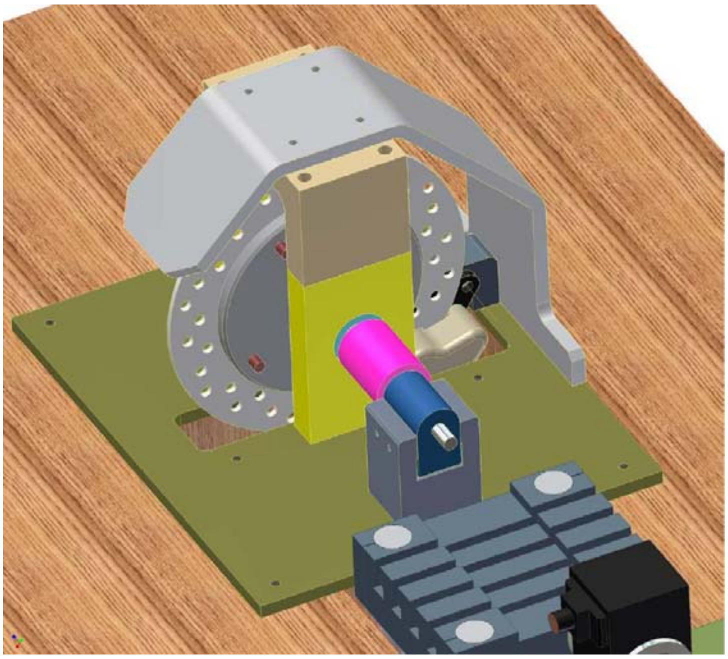


Fig. 4. Model of mechanical brake, part of the test bench.

The input stage of the mechanical part, Fig. 4, of the test bench is also includes an independent torque meter with maximal parameter range up to 10 000 rpm and up to 10 N·m torque. Its mechanical brake consists of a wheel-brake disk that is usually used in the automotive industry. To enable continuous variable control of braking torque, the brake caliper is actuated by a special linear electric motor. The temperature of the wheel-brake disk is measured by a contactless infrared thermometer (pyrometer). Angular speed of the disk is sensed by an inductive sensor, which responds to movement of the disk's clamping screws. The processor unit again processes analog information from the sensors.

Several parts of the test bench are connected together by elastic couplings, which if necessary can be used to couple shafts of different diameters. Alignment of the tested motor and torque meter in the mechanical and electrical part is by means of a height-adjustable central fixture, to which the tested motor is fastened.

4. Control and Display Unit

The control and display unit consists of the graphical monochromatic LCD display and the set of control buttons. The unit is connected to the central CAN bus and receives physical values from the individual parts for display. According to user requirements, the unit sends control commands to the processor unit, which makes appropriate changes. The basic displayed values are – torques from both torque meters, angular speed, temperature of the wheel-disk brake, voltage and current generated by the electric motor in the electrical part of the test bench, etc. The basic parameters for setting are – percent of braking effect for brake caliper, percent of braking effect for electric brake and e.g. settings of the internal protective bonding circuits against overload of individual parts and sensors.

5. Conclusions

The test bench described above is used to perform testing of the mechanical parameters of electrical drives, mainly low and middle power electrically commutated (BLDC) motors. The use of the processor measuring system enables presentation of obtained results in digital form on personal computer or directly on the graphical unit integrated in the test bench. Through the use of the height-adjustable central fixture it is possible to adapt almost any electric drive with compatible dimensions.

During the development work on the control algorithms, several electrically commutated motors by reputable producers, such as Maxon Motor [2], Anaheim Automation [3], Dunkermotor [4], were tested on this stand in UNIS a.s.

Acknowledgements

Published results were acquired using the subsidization of the European Union, FP6 research project No.: 30 888, "**CESAR – Cost Effective Small Aircraft**". Hardware and software development and physical modules were realized and project was supported by UNIS, a.s.

This work was also supported by project No. MSM0021630518 "Simulation modelling of mechatronics systems" solved on the Faculty of Mechanical Engineering, Institute of Production Machines, Systems and Robotics, Brno Technical University.

References

[1] Company web pages, division of Mechatronic System. *UNIS, a.s.* http://www.mechsys.cz

[2] Electrical drives producer web pages. *Maxon Motor*. http://www.maxonmotor.com

[3] Electrical drives producer. *Anaheim Automation.* http://www.anaheimautomatiom.com

[4] Electrical drives producer web pages. *Dunkermotoren*. http://www.dunkermotoren.de

Distributed Measurements and Control Systems for Rapid Prototyping of Artificial Intelligence Controller

[1,2]R. Dindorf, [1]P. Laski, [1]J. Takosoglu

[1]Kielce University of Technology, Kielce, Poland,
[2]AGH University of Science and Technology, Krakow, Poland
Email: dindorf@tu.kielce.pl

Abstract. *The paper deals with distributed measurements and control systems for rapid prototyping of artificial intelligence controller of fluid power driver. The proposed distributed system of measurements and control consists of two PC computers: Target and Host where the first computer directly controls the fluid power driver and is connected to controlled systems, while the second functions as the operator towards the direct control layer. PC computers run Windows with Matlab-Simulink software package. On the basis of the distributed system of measurements and control the fuzzy logic controller of electro-pneumatic servo-drive and adaptive controller of electro-hydraulic servo-drive were designed.*

Keywords: Distributed Measurements and Control Systems, Rapid Prototyping, Fluid Power Control Systems

1. Introduction

In the extended procedures of data acquisition and processing measurement when the advanced control algorithms in fluid power control drives (electro-hydraulic servo-drive and electro-pneumatic servo-drive) are applied both the processing capacity and PLC controllers communications are considerably limited. On the basis of PC computers running Windows XP environment with Matlab-Simulink software package advanced configurations of distributed measurement, control and adjustment systems were set up. To facilitate data acquisition and control fluid power drives a distributed environment running on two PC computers was created. The first computer is connected to controlled systems – fluid power driver while the second functions as a supervisory control layer and the operator towards the direct control layer. In the supervisory control layer complex controlled processing is carried out, the state of controlled system is analyzed and the parameters of control procedures are adjusted to obtain the optimal control conditions. The supervisory control layer, in addition to identification and optimization procedures, may contain a model of controlled system with control algorithm. In the direct control layer the processing, measurement and filtration procedures are conducted. The supervisory control system generates executable files and sends them to the direct control system. Data transmission between PC computers is carried out by TCP/IP protocol (LAN, Ethernet) or by means of serial ports of RS232 type. Industrial distributed measurement and control systems are based on Ethernet networks. The PC computers used in real time systems generate sampling frequency up to 100 kHz. Sampling frequency depends upon processing speed and controlled systems parameters. Distributed measurements and control systems support many input/output formats while additional modules: *xPC Target Explorer* lub *xPCrctool* are used for data processing and acquisition.

The proposed distributed measurements and control systems based upon two PC computers Host and Target are shown in the Fig. 1. The distributed system was used for rapid prototyping of fluid power servo-drives (electro hydraulic and pneumatic servo-drives) in real time. On PC computes Matlab-Simulink and *xPC Target* were installed. In Matlab-Simulink package it is possible to create processing procedures for both conventional and artificial intelligence controllers and to execute own control and visualization applications. PC has the card of analog input/output and *Real-Time xPC Target* system which is used for measurement data acquisition and fluid power drives control. Target PC can simulate the flow of control and measurement signals in real time by means of HIL *(Hardware-in-the-Loop)* method. Applications run by Simulink model use a real time kernel of the PC computer. Host PC and Target PC communicate with each other by TCP/I protocol. The communication of supervisory control layer in Host PC with direct control layer in Target PC may occur continuously, periodically or at operator's specified time intervals. The software suite used in Host PC and Target PC and connections between the two computers are shown in Fig. 2. Working with the rapid prototyping suite consists in building a model of the control algorithm in Simulink. Next the model is compiled and sent to the Target PC which serves as the controller of fluid power drives together with the input/output card and the *Real-Time xPC Target* system. Measurement sensors and transducers are attached to the Target PC through the measurement card. Thanks to *xPC Traget Spy* software the visualization of the processed data and the analysis of the process of controlling fluid power drives is possible.

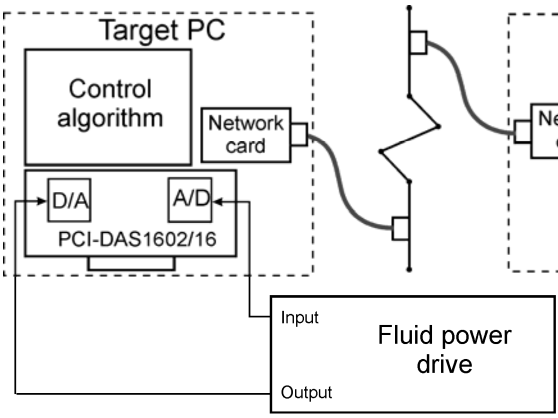


Fig. 1. Schematic diagram of distributed measurements and control systems for fluid power drives.

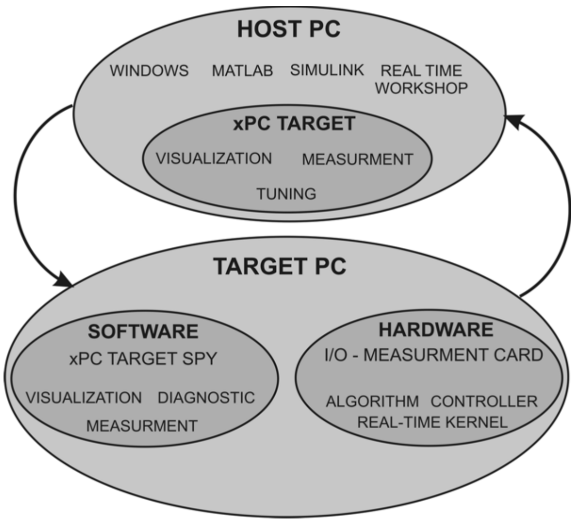


Fig. 2. Diagram of Host PC and Target PC connections according to [1].

2. Distributed Control System for Rapid Prototyping of Fuzzy Logic Controller

On the basis of the proposed distributed measurement and control system a test stand for fast prototyping of fuzzy logic controller of electro-pneumatic servo-drives was set up. The distributed control system with FLC (*Fuzzy Logic Controller*) of PD type controlling pneumatic servo-drive is schematically presented in Fig. 3 [2]. Fuzzy PD controller constructed in Fuzzy Logic Tollbox of Matlab-Simulink package was suggested for the purpose of controlling pneumatic servo-drive. The pneumatic servo-drive together with fuzzy PD controller constitute a system of MISO type with two inputs: position error $e(t) = y_o - y(t)$ and change of position error $\Delta e(t)$ and one output: proportional valve coil voltage $u(t)$. Output and input signals underwent fuzzification process with regular distribution of 7 fuzzy sets of triangular and trapezoid membership functions. The database rules of fuzzy controller are 49 Mac Vicar-Whelen rules described in the table entered to Fuzzy Logic Toolbox. In the inference process the firing degree was determined by means of MIN operator, implication operator and all the inputs of particular rules were aggregated by MAX operator. In the defuzzyfication process the center-of-gravity-method (COG) was applied. The dialogue window "Rule Viewer" of Fuzzy Logic Toolbox is a kind of diagnostic device which enables to trace which fuzzy rules were activated on particular states of input. It also enables observation of fuzzy system output value. The fuzzy logic controller of PD type was tuned by means of Simulink Response Optimization Toolbox of Matlab-Simulink package.

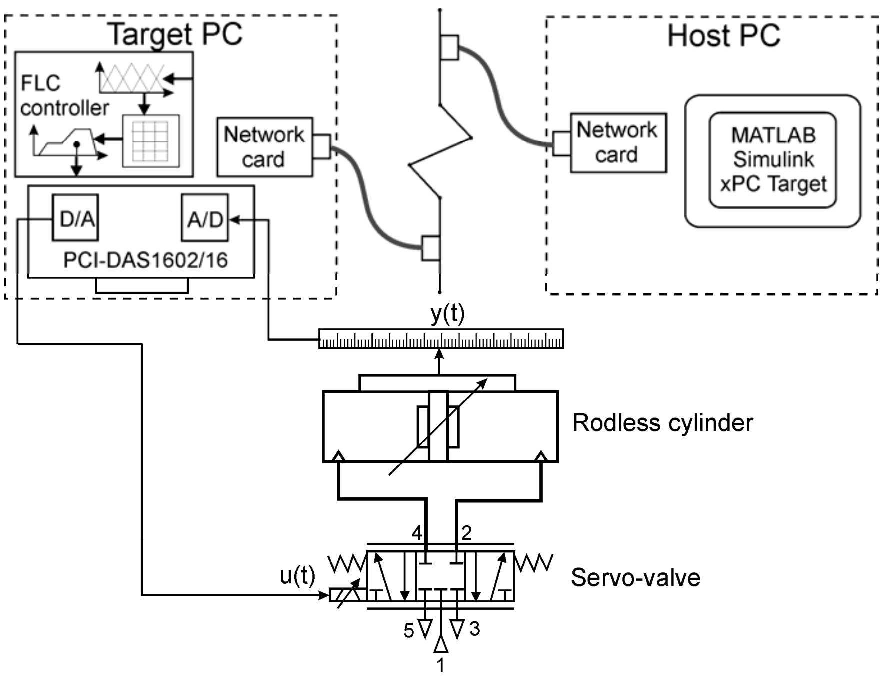


Fig. 3. Schematic diagram of electro-pneumatic servo-drive control system.

3. Distributed Control System for Rapid Prototyping of Adaptive Controller

On the basis of the distributed measurement and control system a test stand for rapid prototyping of adaptive controller of electro-hydraulic servo-drive was set up [3]. To test electro-hydraulic servo-system control algorithms the experimental test stand presented in Fig. 4 was constructed. The experimental test stand consists of two separately controlled control objects composed of hydraulic cylinders controlled by servo-valves (proportional 4/3 valves). The load in the analysed electro-hydraulic servo-drive resulted from slide and resistance movement caused by load cylinder. Piston displacement *y(t)* of hydraulic servo-cylinder was conducted by means of optical transducer. To measure pressure values in cylinder chambers $p_1(t)$ i $p_2(t)$ and pressure in supply line p_o and F force exerted by load cylinder tensometric transducers were used. The distributed control system enables rapid prototyping of adaptive controllers resistant to random interferences resulting from sudden changes of masses and load forces electro-hydraulic servo-drive. In hydrostatic servo-drives the attempts are made to work out such control algorithms which would be insensitive to outer disturbances caused by load mass and external forces and would ensure high precision of positioning and good dynamic properties.

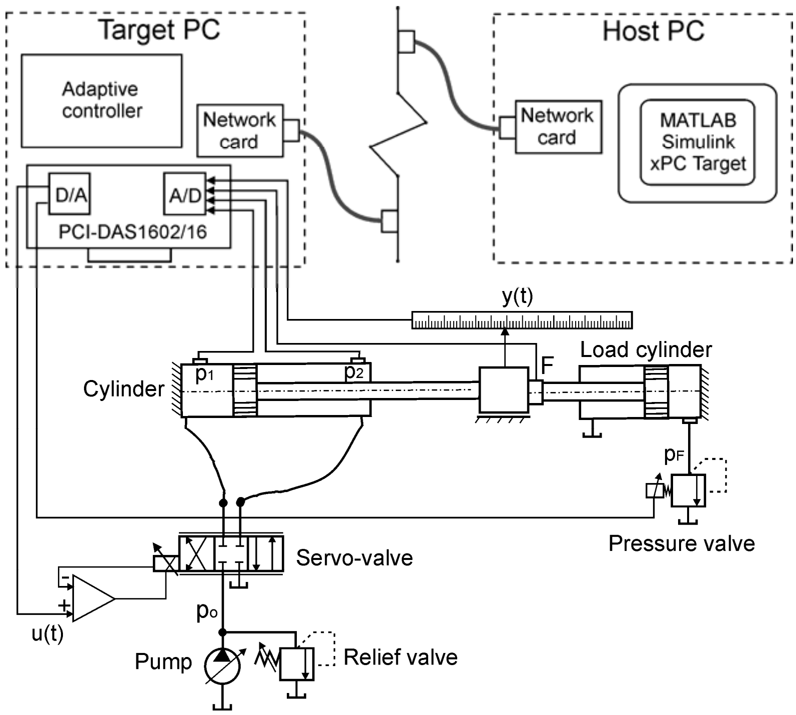


Fig. 4. Schematic diagram of electro-hydraulic servo-drive control system.

4. Conclusions

The article presents the concept of a distributed system for acquiring measurement data and a control system for the rapid prototyping of artificial intelligence controllers of fluid power drives. The proposed measurement and control system consists of two computers: Host PC and Target PC. The Target PC makes up the direct control layer and is connected to the controlled systems (hydraulic and pneumatic servo-drive), while the Host PC makes up the supervisory control layer and serves as an operator of the direct control layer.

References

[1] Users Manuel: xPC TargetFor Use with Real-Time Workshop. The MathWorks, 2005.

[2] Takosoglu J.E., Dindorf R.F., Laski P.A.: Rapid prototyping of fuzzy controller pneumatic servo-system. *International Journal of Advanced Manufacturing Technology*, Vol. 9, 2008. Springer Verlag, London.

[3] Woś P., Dindorf R.: Problems of the adaptive control for electro-hydraulic servo-system. 5th Fluid Power Net International – PHD Symposium, Cracow (Poland), 1-5 July 2008.

High Resolution Audio Codec Test

P. Fexa, J. Svatoš, J. Vedral

Czech Technical University, Technická 2, 166 27 Prague 6, Czech Republic
Email: fexap1@fel.cvut.cz, svatoj1@fel.cvut.cz, vedral@fel.cvut.cz

Abstract. *This paper is to show the methods of dynamic parameters testing of high quality audio codec with 24 bits nominal resolution. In the paper are briefly the best Sine Wave Fit Test, Spectral Purity Test and Noise Histogram Test presented. The above mentioned testing methods are then used for testing of these ADCs and obtained data are discussed..*

Keywords: Audio Codec Test, Sine Wave Fit Test, Spectral Purity Test, Noise Histogram Test

1. Introduction

High resolution audio analog-to-digital converters are one of the most important parts in different electronic applications in this time. The typical example is spectral analyzing in measurement applications and processing from sensors of physical (non-electric) quantities (magnetometery, resistive bridges, and thermometry). The degradation of signal rising on the beginning of measuring chain is very difficult to correct. Therefore the knowledge of parameters of ADC used in such systems is necessary.

2. Used Methods of Testing

Main parameters of tested ADC from the point of view of dynamical testing audio converters are the ratio of *Signal Noise and Distortion SINAD, Effective Number of Bits ENOB, Total Harmonic Distortion THD, Spurious Free Dynamic Range SFDR* and *Effective Resolution ER*. The principles of these methods are well described in the IEEE standards [1], [2]. The dynamic testing methods are in principle divided into two main parts.

In the *Time Domain* is the S*ine Wave Fit Test* witch uses pure harmonic signal as an input signal. The principle of method is fitting of sin wave into measured record of data by the least minimum squares algorithm. Basic algorithms are possible to use *four parameter* least squares fit to sine wave data. The parameter *ENOB* is possible to re-count from this result as

$$ENOB = n - \log_2 \frac{\varepsilon}{RMS_q} \tag{1}$$

where n is number of *nominal bits*, ε is *minimum square error* of Fit Test and $RMS_q = 2^{-n}/\sqrt{12}$ is quantized noise of ideal AD converter.

In the *Spectral Purity Test method is* data recorded not absolutely coherent with used sampling frequency in ideal case. In this case is not necessary to use any window function because the required periodicity of sampled time interval is conformed a priori. Often used window is *Blackman - Harris's Window*. The parameter *SINAD* is calculated as the ratio of the RMS_{Sin} value of basic harmonic component to the RMS_{Tn} value of the total noise including harmonic distortion, spurious components and various types of noise

$$SINAD = 20\log_{10} \frac{RMS_{Sin}}{RMS_{Tn}} (dB) \tag{2}$$

Parameters *SINAD* and *ENOB* are mutually re-countable by means of well-known formula

$$SINAD = 6.02ENOB + 1.76(dB) \tag{3}$$

It should be stated the method of determination of this parameter if any dynamic parameter of ADC is stated.

Total Harmonic Distortion THD represents the nonlinearity of the A/D converter transfer characteristic. For a pure input sine wave is the ration of the root sum of squares of all the harmonic distortion components, including their aliases in the spectral output of the converter.

$$THD = 20\log_{10}\frac{\sqrt{\sum_{i=2}^{m}U_i^2}}{U_i}(dB) \tag{4}$$

Spurious Free Dynamic Range SFDR is defined for a pure sin wave input as the ration of the root sum of squares of all the lowest harmonic distortion components, including their aliases to the *RMS* value of the main harmonic amplitude

$$SFDR = 20\log_{10}\frac{U_{\max}}{U_1}(dB) \tag{5}$$

Effective Resolution is defined for grounded inputs of converter

$$ER = \log_2\frac{FS}{RMS_n}(bit) \tag{6}$$

where *FS* is nominal range of converter and RMS_n is effective value of noise signal.

3. Test Conditions

Both above described converters have been tested by means of the boards supplemented by Creative Sound Blaster Live. The audio codec under test is 24 bit codec Terratec Phase 26 USB with maximal sample rate 96 kSa/sec. The part contains a *sigma-delta modulator* 8 order, a *finite impulse response FIR digital filter*. The digital filter frequency response can be programmed to be either low pass or band pass. The on-chip filtering combined with an *oversampling ratio* 256 reduces the external antialiasing requirements to first order in most cases. The board is connected to the PC by USB2 interface. Measured data are processed by soft in MATLAB.

The ultra-low distortion function audio generator DS360 from Stanford Research has been used as a source of reference harmonic signal. The manufacturer guarantees the total harmonic distortion *THD* ≤ 120 dB in the frequency range from 20 Hz to 50 kHz. The frequency of generated signal should not be an integer multiple of line frequency. With such arranged measuring conditions all next measurements have been performed.

4. Results

In Fig.1 and Fig.2 is presented *FFT spectral plot* and *Noise Histogram Test* grounded input of audio codec. The *Effective Resolution* of codec is 15,6 bits. In FFT spectral plot are evident parasitic spectral components with USB supply frequency 200 kHz. For FFT test is 64 kbyte samples and Blackmann-Harris Windows 4 order used.

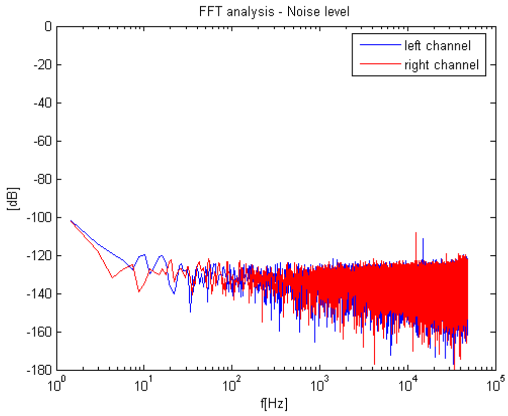


Fig. 1 FFT spectral plot of Noise Histogram Test

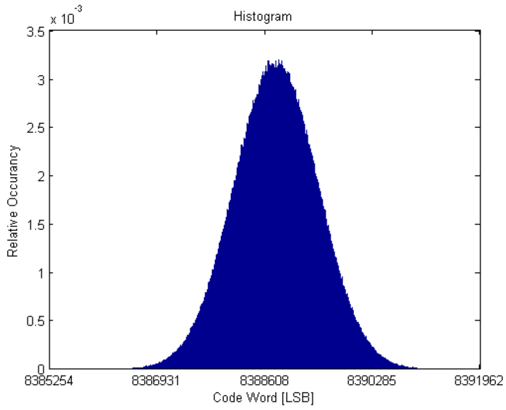


Fig. 2 Noise Histogram Test

In Fig.3 and Fig.4 is presented FFT plot of reconstructed sinus signal 1 kHz with USB and battery supply. In spectral plots of Fig. 3 is evident parasitic spectral components in amplitude under - 100 dB versus major spectral component 1 kHz.

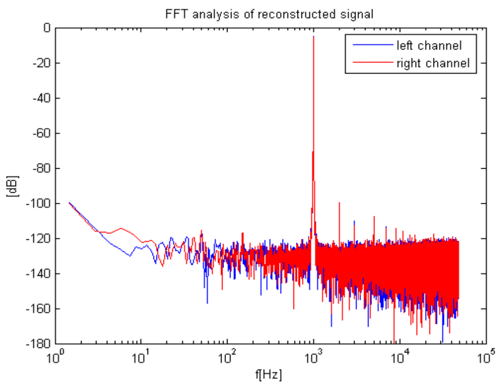


Fig. 3 FFT spectral plot 1 kHz with USB supply

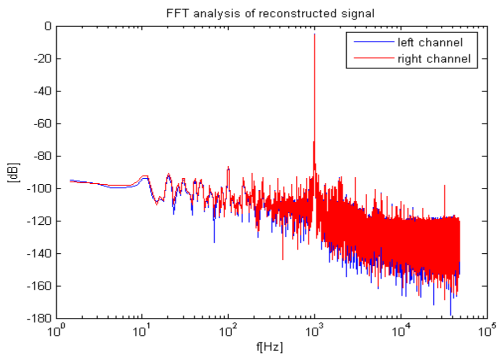


Fig. 4 FFT spectral plot 1 kHz with battery supply

In Fig.5 and Fig.6 is presented plot of *Effective Number of Bits* and *Total Harmonic Distortion* in frequency range 20 Hz to 20 kHz.

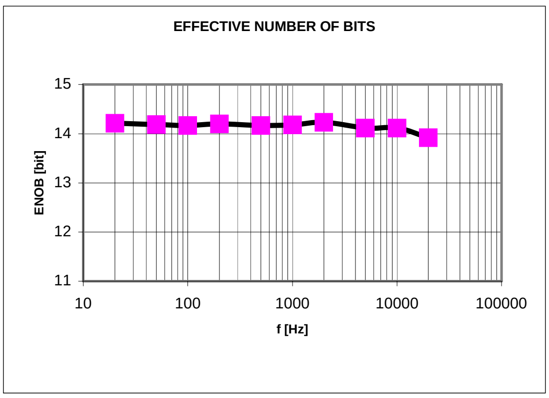


Fig. 5 Effective number of bits frequency plot

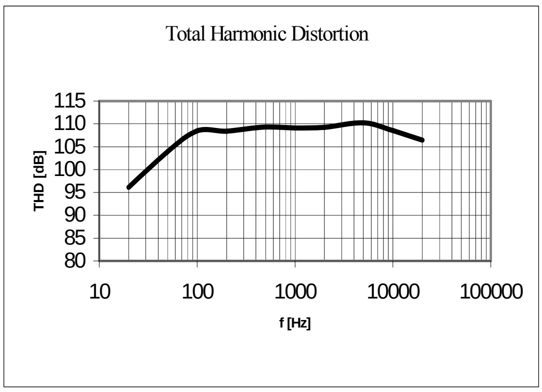


Fig. 6 Total harmonic Distortion frequency plot

5. Discussion

From tested audio codec in frequency range 20 Hz to 20 kHz *ENOB* = 14,5 bits and *THD* up to 110 dB in frequency range 100 Hz to 10 kHz. Frequency range corresponding to drop-off of transfer function – 3 dB is 20 Hz to 40 kHz. In this frequency range is stereo crosstalk 48 dB, and in the frequency range 1 kHz to 10 kHz is this stereo crosstalk up to 85 dB. Characteristic noise of tested sound card reaches to level of 120 dB. Disruptive spectral elements in both channels in frequency area close to 10 kHz reduce this distance to 110 dB. The official parameters of audio codec Terratec Phase 26B is certified: *THD* = 105 dB and *SINAD* = 98 dB from input signal frequency 1 kHz [9].

6. Conclusions

Under testing it's very important to separate alternating supply source from inner analog to digital converters in tested sound card. External accumulator was used as power source and decrease of disruptive signals more than 20 dB was achieved. We can say, that tested audio sound card match common 16 bit digital converters except the distortion done by high frequency signals and signal to noise ratio. Those parameters are close to nominal resolution 24 bits.

Acknowledgements

This project is supported by the Czech ministry of Education: MSM 6840770015.

References

[1] IEEE Standard 1057-1994, *IEEE Standard for Digitizing Waveform Recorders.*

[2] IEEE Standard 1241 - 2000, *IEEE Standard and Teminology and Test Methods for Analog-to-Digital Converters.*

[3] Jeng., Y.C., Crosby, P.B.: Sinewave parameter estimation algorithm with application to waveform digitizer effective bits measurement. IEEE *Trans. Instrum. Meas.,* vol. IM-37, pp. 529-532, Dec. 1988.

[4] Jeng., Y.C.: Measuring harmonic distortion and noise floor of A/D converter using spectral average, *IEEE Trans. Instrum. Meas*., vol. IM-37, pp. 525-528, Dec. 1988.

[5] Harris, F.J.: On the use od windows for the harmonic analysis with the DFT. *Proc. IEEE*, vol.66, pp. 51-83, Jan. 1978.

[6] Vedral, J.: Holub, J.: Measurements od ADCs effective number of bits. ADDA 2005, *International conference on advanced A/D and D/A Conversion Techniques and their Applications.* Limerick, 26005, pp. 255-257, ISBN 0-86341-542-3.

[7] Podlipny, J.: High resolution ADC Testing. Diploma work, CTU FEE, Prague, 2008.

[8] Vedral, J.: Aspects of ADC Verification . In *Testmethoden und Zuverlässigkeit von Schaltungen und Systemen*. Erlangen: Universität Erlangen-Nürnberg, 2007, p. 23-25.

[9] http//www.teratec.com.

Lower Layer Design of Narrowband PLC Node for AMR and AMM Applications and Network Layer Protocol Verification

J. Dvořák

Czech Technical University - FEE – Department of Measurement,
Prague, Czech Republic,
Email: dvoraj6@fel.cvut.cz

Abstract. *This paper describes design of a network node, which uses narrowband powerline communication for data exchange in AMR and AMM networks. Also an implementation of first two ISO/OSI layers is described in this paper. Data reception synchronization and some indicators for upper layer functions as the capabilities of physical layer are detailed. A passage about data link layer is focused on the frame structure and implementation of FEC and ARQ methods for reliable data transfers.*

Keywords: Narrowband Powerline Communication, Media Layer Protocols, AMM and AMR Systems

1. Introduction

Automatic Meter Reading (AMR) and Advanced Metering Management (AMM) are relatively new ways for data collection from electricity, gas and other meters. These systems provide accurate view on consumer demand thus allowing media providers to control their networks more easily and effectively. AMM systems can also support direct control of media delivery points. This can be helpful, when company wants to disconnect media flow to delivery point or just to reduce it. Energy and media providers often tend to narrowband powerline communication (PLC) at the last mile of these data collecting systems. PLC is preferred for its simplicity and cheapness compared to extra data buses or wireless solutions. Main advantages in comparison with a broadband PLC communication are generally lower signal attenuation at low frequencies and the regulations allowing only energy providers to transmit in 9 – 95 kHz bandwidth in power lines. Fully working network of PLC nodes with effective lower (media) layer protocols would be capable to collect data from the covered area easily and with lover costs then other mentioned systems.

2. Powerline Communication and Design of a PLC Node

Channel Characterisation

Power lines are primarily designed for electrical energy delivery with no regards for data transmissions. Data transfers are obtained by the use of some digital modulation technique with carrier frequency in range of tens of kHz for narrowband PLC and MHz to tens of MHz for broadband PLC. Power line parameters are frequency dependent on connected loads. It means that the impedance of connected loads can rapidly change the signal attenuation at measured distance. Generally the attenuation is higher with increasing frequency and therefore the narrowband PLC can work at longer distances than the broadband PLC. Another effect, which makes problems in PLC, is the noise produced by connected loads. It can be a background continuous noise from universal motors, short strong impulses from switching electrical circuits, synchronous and asynchronous noise from switching with mains frequency or from switched power supplies. All described characteristics of the power lines are not stable. They are time varying, therefore the power lines are very hostile and hardly predictable for the data transfers. Powerline channel between two communicating nodes is often asymmetric due to the different position of noise sources. This has to be considered within the development of PLC node and protocols for communication.

Physical Layer

The first version of hardware design of my PLC node consists of two boards (Fig. 1). The first is a ST7540 narrowband PLC modem based board with coupling circuits and communication interface. The second board is based on 40 MHz ARM7 core microcontroller equipped with serial interfaces and interface for PLC board communication. Communication protocol stack implementation is written in C language with possibility to insert upper layers in future.

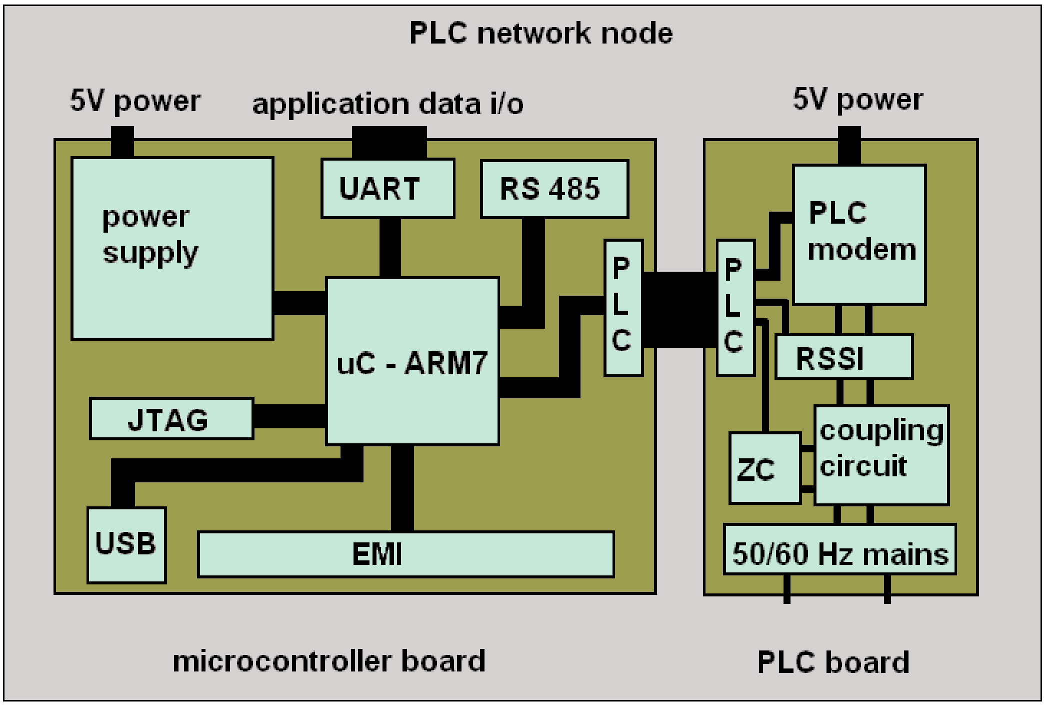


Fig. 1. Block scheme of PLC network node

The PLC modem uses binary FSK modulation for data transfers. It supports 4 bit rates, carrier and preamble detection, automatic output level control, input digital filtering, input sensitivity setting and 6 different carrier frequencies in A and C bands of the CENELEC EN 50065. The coupling circuit is a band-pass filter designed for 72,05 kHz centre frequency. Bit rate is set to 4800 b/s and the host communication utilizes a synchronous SPI interface.

Physical layer software protocol stack implementation contains buffers and functions for physical layer data frame transmission and reception. The structure of the physical layer frame is shown in Fig. 2.

Fig. 2. Structure of a physical layer frame

As in other similar communication methods the data are correctly stored after reception of the synchronization sequence. It is indicated by proper synchronization sequence recognition, which follows the preamble sequence of zeros and ones used for the modem bit synchronization. Because of the strong disturbances in power lines it is necessary to ensure the proper reception even with some degree of errors in synchronization sequence (3 errors in first 16 bits and 8 in following 32 bits are accepted in my case). This synchronization sequence is relatively long, as it was designed to capture frames more correctly for an accurate BER measurement compared with simple 16-bit synchronization. Future development will probably lead to shorten the synchronization sequence. The physical layer provides data link layer information (for MAC protocol), whether node is busy or communication channel is occupied.

Another important capability of physical layer is a Received Signal Strength Indicator – RSSI. In-band signal strength can be provided to data link layer to support the MAC protocol and the received signal strength during the frame reception can be provided to the data link layer or to upper layers. This indicator is especially relevant for network layer routing protocol.

Data Link Layer

The data link layer takes care of data exchange between two nodes, which are close to each other and able to communicate directly. Addressing of nodes was designed with regards to the first proposed version of the network layer protocol. A structure of the typical data link layer frame for data transfer is at Fig. 3.

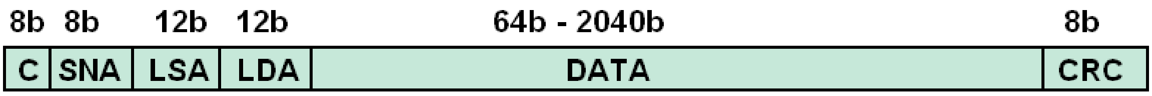


Fig. 3. Structure of a data link layer frame

The first byte defines the frame type. Next byte represents a subnet address, which allows coexistence of more than one network of these nodes under the control of different data concentrators. Following three bytes are link layer source and destination addresses. Next follow the upper layer data and 8-bit CRC fields.

The data link layer supports frame acknowledging and ARQ method for data exchange. The acknowledge frame has the same structure as a normal data link frame and in the data block it carries information about the quality of received frame reception. The protocol allows up to four retransmissions. This method improves data transfer reliability when long burst errors are present in powerline communication channel.

Power lines are very hostile and noisy environment for PLC, so it was necessary to implement a forward error correction (FEC). For the future optimization of the transmission power efficiency 3 levels of forward error correction methods were implemented. The data link layer then supports 4 types of data processing based on the applied BCH codes and interleaving. Approximate relation of measured BER and signal to noise quotient for all implemented FEC methods is shown in Fig. 4. The best performance in reliable area for data transfers (under 0,5 % error rate) has concatenation of BCH codes combined with interleaving for better burst error correction. Disadvantage of this relatively strong FEC is 3,4 code rate, but it’s the only simple way to transfer data without repeaters, when the S/N ratio is decreasing closer to zero.

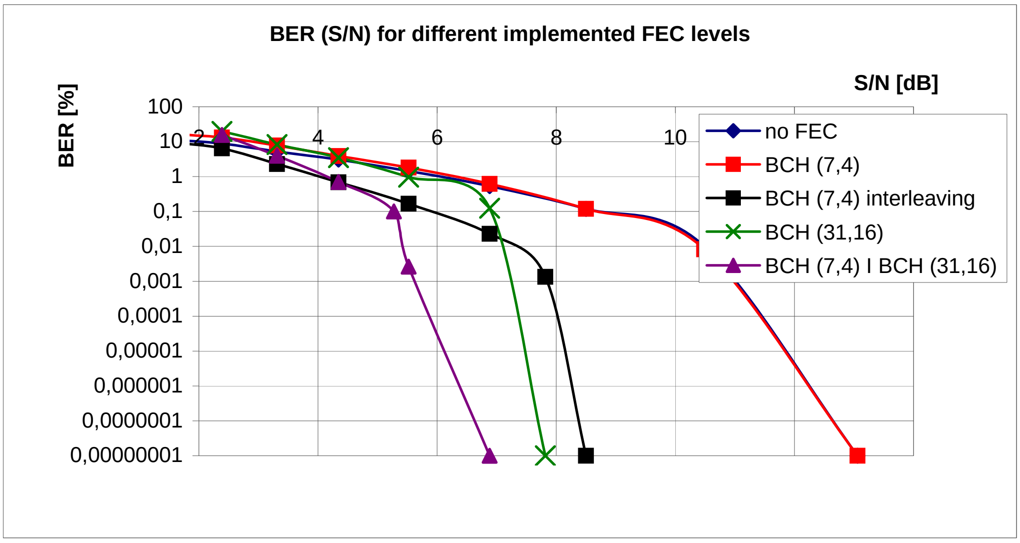


Fig. 4. Comparison of BER in relation with S/N of implemented error correcting coding methods

3. Conclusions

This robust lower layer protocols implementation ensures reliable data transfer over the low voltage power lines. All described methods and procedures are essential for communication in such a noisy and unpredictable channel. The BER measurement shows that strongest implemented level of FEC improves reliability of data transfers, when S/N ratio is getting lower due to the distance between nodes or higher level of noise in power lines. The coding gain between data transfers with no FEC and strong FEC with concatenation of two BCH codes and interleaving is 5dB. This is of course further improved by the use of ARQ method. The described protocols will be improved and optimized for the transmission power efficiency in future. Next important level dealing with narrowband PLC for AMM and AMR is network layer protocol. Network protocol is important, because it makes it possible to send data at longer distances and in more noisy networks, when physical and data link layers are unable to reach all nodes in network. The physical and link layer implementation will further be used for verification of the adaptive network layer protocol focused on minimized power consumption of the overall network communication.

Acknowledgement

This research is supported by The Czech Science Foundation (PN: 102/09/H081).

References

[1] Andreou. G. T., Labridis D. P., Modeling of Low Voltage Distribution Cables for Powerline Communication, IEEE Bologna Power/Tech Conference, Bologna, Italy, 2003

[2] Gault S., Ciblat P., Hachem W., An OFDMA Based Modem for PoweLine Communications over the Low Voltage Distribution Network, 2005 International Symposium on Powerline Communications and Its Applications, 42, 2005

[3] Adámek J., Kódování, SNTL, Prague, 1989

[4] Proakis J. G., Digital Communications, McGraw-Hill, New York, 2001,312, 416-470

Remote Measurement of NNTP Internet Traffic

T. Surmacz

Wroclaw University of Technology,
Institute of Computer Engineering, Control and Robotics
Email: tsurmacz@ict.pwr.wroc.pl

Abstract. *Measuring internet traffic related to NNTP protocol (Usenet News) provides vital information needed for optimum news server configuration. However, gathering data is usually restricted to the "local view", even though in networked environment an overview of the neighbourhood traffic would be more appropriate. This article describes method of gathering such data from remote points and analysing it to provide the required traffic flow information.*

Keywords: Measuring Internet Traffic, NNTP Protocol, Usenet News

1. Introduction

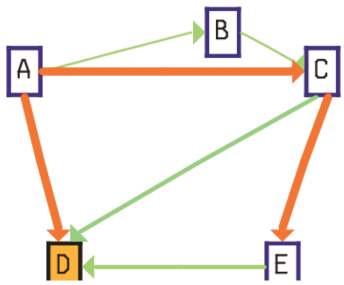
Usenet is a large network of servers exchanging messages (articles) sent by users throughout the world in various discussion groups. Message sent on any server in a particular discussion group gets propagated to all other servers carrying this group through a flooding algorithm [1] that provides a good tradeoff in achieving two main purposes: delivering the messages to all servers and saving the required bandwidth by not transmitting the message to servers already containing it. However, as full usenet news traffic often exceeds 200-300 GigaBytes of data per day on a single server-to-server link, and a server may have several such links open to other servers, more sophisticated tools are required to properly engineer such a traffic, especially that there are usually many redundant links between servers set up to increase the reliability of the overall system.

Measurement of this traffic is the first step for optimizing network bandwidth usage of a news server. The typical setup, however, is to gather locally generated statistical data, which gives only a "local view" of the news server operation. This may be used to spot irregularities or trends in day-to-day operation, but does not provide a "bigger picture" of the overall news traffic in the neighbourhood of servers. In order to get this bigger picture, acquisition of traffic data from remote servers is required.

2. Subject of Measurement

Let's consider a simplified network of just 5 servers located in two different cities (Fig. 1). Servers A, B and C are connected by a wide bandwidth metropolitan area network, as well as servers D and E sharing a fast connection, but the link between these two groups is much slower. If a message is posted to server A (either by a local user, or received from some external link), it will be propagated to all neighbouring servers (B, C and D). The same happens all over again on every consecutive server, so server B will try to propagate it to C, while C will try to push it through to D and E. Servers D and E will eventually exchange the message too, depending on which of them receives it first, but there is a good chance, that the message will traverse the slow link twice (from A to D and from C to E) instead of being distributed mostly by fast links. A way of news bandwidth control by message diverting has been proposed in [2], however a proper measurement method to evaluate the results is still required.

Fig. 1. Sample news servers network topology



Considering the message propagation shown in Fig. 1, server D can calculate the incoming traffic from all its neighbours, but cannot determine the bandwidth used between them (e.g. the traffic between C and E or B and C). Even if data collection is possible from nodes other than D, it will not be possible for all nodes, so some information must be derived from the data collected at other points of the network.

The aim of a measurement method described in this article is to estimate the bandwidth used between all the servers being considered, by analysing the data which is already available locally in headers of received messages, and possibly retrieving the missing data from other servers.

3. Measurement Method

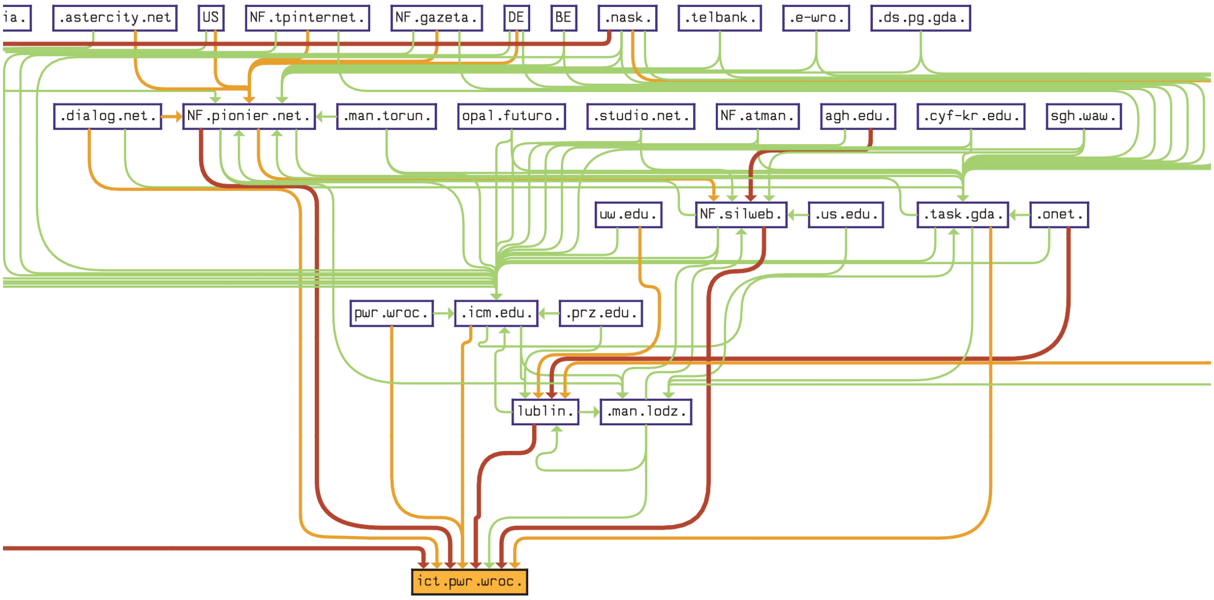
Messages sent through the Usenet News system contain two main parts: a header and a body. The header contains all the information about the message origin (author, date posted, server used for posting, etc.), as well as the “Path:” field, which accumulates the information about the servers the message has been sent through. So if a message is sent from server A to B, then from B to C, the Path: field will contain “Path: C!B!A”, as it shows all server names in reverse order, split by the exclamation mark. This data is normally used by NNTP servers to reduce the required bandwidth by not offering the message to servers that have already seen it (i.e. server C will not offer this article to B nor A), but may also be used to calculate the NNTP-related bandwidth of remote links.

The method for calculating the bandwidth first uses all the locally contained messages and analyses them by splitting the “Path:” field. The directed graph is being built, in which nodes represent servers and edges represent server links. Each message coming from server A through servers B and C to server D (i.e. “Path: D!C!B!A”) adds to edges AB, BC and CD the size of the message (called its *weight*) which is known, as the message is stored locally. However, if the Path: header contains only vectors BA and AD, then it is not known how the message was delivered to server B, or how server C has been offered this, if at all, and to which other edges should the appropriate weight be added. Another target of measurement may be to count just the number of messages sent, but this is a trivial simplification done by ignoring the actual size of messages.

The next step is to analyse the data that can be obtained from remote servers. If other servers offer public access to their repositories, it is possible to connect to them using the NNTP protocol as a client and retrieve message headers for analysis. This must be done in several steps:

1. connect to a remote server as a client;
2. retrieve the list of accessible groups;
3. retrieve the numbers of the first and the last message in each observed group;
4. compare these numbers to the results of a previous session with this server;
5. retrieve headers of all messages that have not been seen in previous run and store them locally.

Fig. 2. Visualisation of Path analysis results for server ict.pwr.wroc.pl, based on its local data only. For picture clarity, the leading ”*news*” and trailing ”*pl*” parts have been omitted in server names and such names are left with a starting or ending dot. The ”*newsfeed*” prefix in real names has been replaced with ”*NF*”.



Steps 3-5 have to be repeated for each discussion group and steps 1-5 for each server that can be used for getting reading access. It is possible to retrieve only some of the existing header fields using XHDR and XOVER NNTP protocol extensions described in [3]. The headers used in this method are “Path:”, “Xref:”, “Message-ID:”, “NNTP-Posting-Date:”, “Date:” and “Bytes:”. The “Xref:” header contains the list of groups where the message has been posted and may be used to speed up data processing by simplifying message exclusion algorithms (so it does not get counted twice if sent to two different groups). The two posting dates (one set by the client, the other by the server) are currently not used, but may be included in propagation delay analysis or for selecting messages sent in particular period (for now, this is done by collecting data in regular intervals and analysing them in batches). The most important headers are the remaining ones – “Path:”, “Bytes:” and “Message-ID:”. The latter one contains a unique identifier of the message which can be used to correctly identify all messages whose weights (the “Bytes:” header) have been already added to particular graph edges and should not be counted again on these links.

Consider the last example – the message that comes from server B to A, and then from server A is sent independently to servers C and D. When it is analysed on D as a locally stored message, the Path: header contains “Path: D!A!B”, so its weight is added to vectors BA and AD. The header of the same message retrieved from server C would contain “Path: C!A!B”, so its weight would get counted twice on edge BA. To prevent that, it is necessary not only to add message weights, but also to keep track of their Message-IDs in a sorted list associated with each edge being considered.

4. Results

The results of the counting algorithm are twofold – first, it allows to find NNTP connections between remote servers and draw a graph showing them, second – it finds the bandwidth used in these connections. Analysis performed on data gathered lately (beginning of February 2009) showed surprisingly large number of redundant links between servers to the extent which made creating any visual graphs impossible, as they were not readable at all, with links spanning in all directions. Fig. 2 shows the results of stage 1 "Path:" analysis from local data, gathered on server ict.pwr.wroc.pl. The weights calculated for particular links (represented by line thickness) had to be used to exclude all links below some threshold values in order to make this graph readable. However, for actual calculations of server "importance" and the bandwidth it utilizes, all the edge weights should be used. Line thickness/color represents the importance of the link, i.e. the number of messages being sent. Fig. 3 shows the bandwidth analysis results of remotely gathered data, presented in graph form, and restricted to traffic in "pl.*" subset of discussion groups only. For graph clarity, all edges and nodes with less then 10 MB of traffic have been omitted and the numbers represent kilobytes of data sent. Highlighted nodes are the servers used for data collection. The analysed data was collected between 7.02.1009 and 14.02.2009 and consisted of 322477 article headers (71236 unique articles) with 258 different data paths discovered before the graph reduction.

5. Discussion

The counting algorithm presented in this paper helps system administrators of NNTP servers to visualize how messages are being sent between the servers and calculate the bandwidth used for NNTP service between local and remote servers. This data may be used to optimise network usage by modifying server setup and divert traffic between existing server links helping to achieve both the required reliability from redundant links, and efficiency. The nature of data collection implies that the counted numbers cannot be 100% accurate, as this would require connecting to every single server on the network and including its data in overall analysis, which would be impractical, if possible at all. However, by Message-ID: analysis performed in parallel to weight counting, it is possible to find what messages have been omitted in calculations for particular node, and thus – estimate the corrections for edge weights related to this node. This will be marginal for nodes for which the data was collected using the NNTP protocol, but may be significant for those implied from Path: analysis and graph building. However, for "important" servers which are large message exchange hubs these uncertainty numbers are also low, as the messages sent through them are widely distributed through other servers, and thus – are included in the overall analysis.

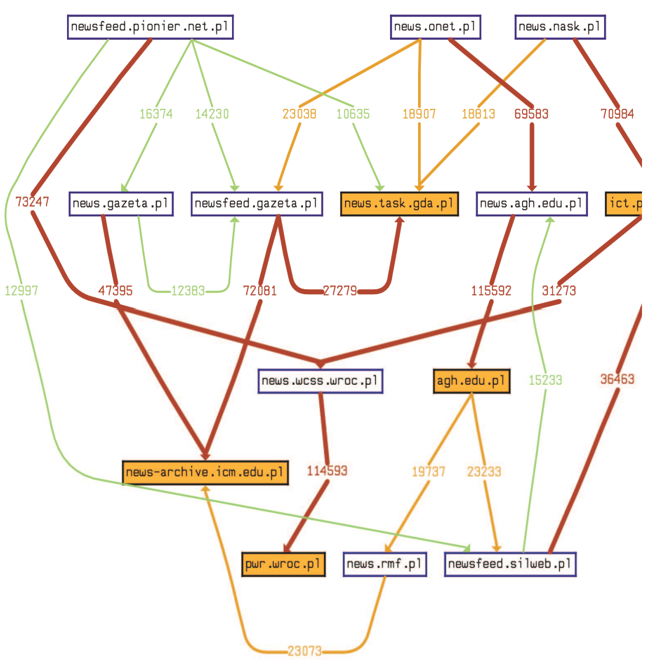


Fig. 3. Full analysis results – kilobytes of messages sent between servers

6. Conclusions

The method described in this article allows measurement of NNTP traffic in IP networks based on just few probe points. The results of these measurements may be used by Usenet news system administrators to monitor daily bandwidth usage and assist in traffic engineering, or calculate various characteristics of news flow between news servers.

References

[1] Horton M, Adams R. Standard for interchange of USENET messages. *RFC 1036*. Dec 1987.

[2] Surmacz T. Bandwidth control in redundant news server links. In proceedings of International Conference on Dependability of Computer Systems (DepCoS-RELCOMEX 2006), Szklarska Poręba, Poland, IEEE Computer Society 2006, 143-149.

[3] Barber S. Common NNTP Extensions. *RFC 2980*, Oct 2000

Evaluation of Parameter Impact on Final Briquettes Quality

[1]M. Svátek, [2]P. Križan, [2]Ľ. Šooš, [1]E. Kureková

[1]Slovak University of Technology, Faculty of Mechanical Engineering,
Institute of Automation Measurement and Applied Informatics, Bratislava, Slovakia,
[2]Slovak University of Technology, Faculty of Mechanical Engineering,
Institute of Manufacturing Systems, Environmental Technology and
Quality Management, Bratislava, Slovakia,
Michal.Svatek@Officedepot.com; peter.krizan@stuba.sk

Abstract. *This paper describes shortly evaluation process of parameter impact on final briquettes quality. At the beginning we present compacting technologies of the solid organic waste as briquetting or pelleting and also products of these technologies. In the middle of the paper we describe parameters that have great influence on briquette quality and design experimental pressing stend. Last part of this paper deals with experiment and evaluation of measured values.*

Keywords: Briquetting, Pelleting, Biomass, Briquette Density, Quality of Extrusion

1. Introduction

Biomass forms a big part of unutilized waste. Production of high-grade bio fuels - briquettes - is a convenient way how to energetically efficiently utilize the waste. Waste has to modify during the process of energetic upgrade. It is very important to know the influence of all parameters on the final quality of briquettes, which is evaluated by the density of briquettes. This report deals with evaluation of measured values that influence the final quality of briquettes.

2. Compacting technologies

Compacting technologies have been known for more than 130 years. The feature that all technologies of compacting – pelleting, compacting and briquetting – have in common is pressing of material under high pressure. Essential preconditions for compressed material to stay compact after pressing even without binder are appropriate setting of pressure and temperature inside the pressing chamber. Further precondition appears to be the maximal fraction size as well as allowable humidity of compressed waste. In the case that the mentioned values are exceeded, the waste needs to be crushed and dried out. The final products of briquetting, compacting and pelleting are briquettes of various shapes and sizes. The technology of compacting is not being described in the report, as it is not used to produce briquettes due to absence of endurance phase of applied pressure as well as absence of cooling phase. This fact is very important in biomass compacting.

Pelleting

Peletting is a progressive way how to compact the crushed and dried mass. It is a new and very progressively developing compacting technology, by which the dried wooden mass (biomass) or crushed wooden mass is processed by the compacting press under a very high pressure. Its introduction needs higher investment in technology purchase as well as investments from the customer side. The final products of the mentioned technology are pellets of cylinder-like shape. Pelleting machine is shown in Fig. 1. [1]

Briquetting

Briquetting is the most known and a widespread technology of compacting. The technology uses mechanical and chemical properties of materials which are compressed into the compact shapes (briquettes) without additive binders using the high pressure compacting. When, for example the biomass undergoes the process of briquetting, while high pressure and a temperature simultaneously act upon the mass, the cellular structures within the material release lignin, which binds individual particles into a compact unit. The briquetting can be used to compact the following materials: sawdust particles, wood shavings, bark, wood dust, straw, cotton, fabrics (natural materials, synthetic materials combined with the material containing lignin), paper, waste of raw material origin, etc. Briquetting press is shown in Fig. 2. [1]

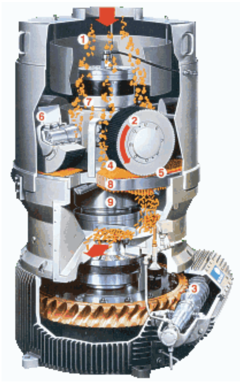


Fig. 1. Pelleting machine   Fig. 2. Mechanic crankshaft briquetting press

3. Parameters influencing the final quality of briquettes

The final quality of briquettes is influenced by many factors. The most significant effect have pressing temperature, compacting pressure, largeness of input fraction, kind of material, size and humidity of fraction. [4]

Compacting pressure is the most important factor having influence particularly upon the strength of briquettes. The strength of briquettes increases with increasing pressure within its strength limit and tendency to absorb atmospheric humidity during the long term storage is decreasing.

Pressing temperature belongs together with compacting pressure to the most considerable parameters, what means - that it has significant effect on the quality and strength of briquettes. It determines the lignin excretion by cellular structures within the wood. Lignin is released under certain pressing temperature, which has to be unconditionally reached to assure undergoing process.

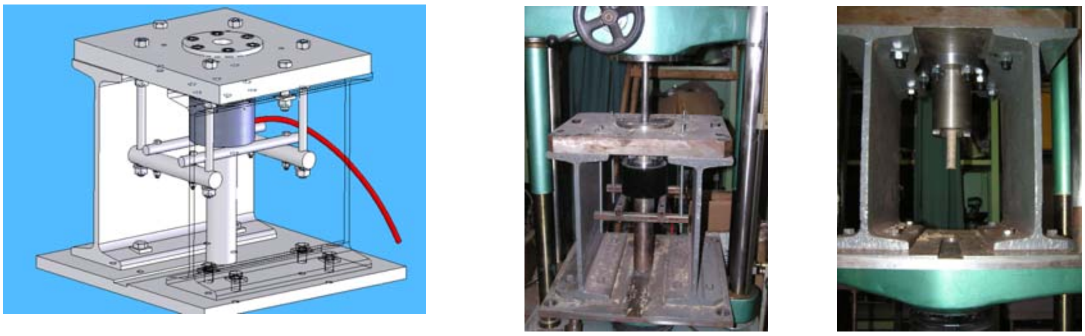
Largeness of input fraction has also effect on compacting process. The bigger is input fraction, the higher is power output required for compacting. In spite of that, the briquette has lower homogeneity and strength. With increasing largeness of fraction, the strength of bonds is decreasing results in fast briquette disintegration in the process of burning (briquette burns down faster, which represents its disadvantage).

Humidity of input material has influence on lignin plastification. Water has positive role in growing tree because it is essential life condition for existence of every plant organism. Presence of water in tree that has been cut is undesirable. All recently known compressing technologies enable to compact material having relative humidity lower than 18%. The humidity of 18 % appears to be the value optimal for compacting, because if the humidity is very small or in contrary very high, particles of material are not compact and briquettes may disintegrate.

4. Experiment and evaluation

At Faculty of Mechanical Engineering of Slovak University of Technology in Bratislava a stend has been designed and constructed and it serves to observe (measure) effects of compactig pressure, pressing temperature, humidity of input fraction and largeness of input fraction on quality of briquettes (see Fig. 3) [2].

The report presents the results of experiments, which were made with the same material, while above mentioned effects were changing. The aim of the experiments was to find out the effect rate of all important parameters on the quality of the briquette.



a) b) c)

Fig. 3. Experimental pressing stend
a) 3D model; b) produced pressing stend; c. pressing stage of a briquette

The measurements on a pressing stend were carried out using different values of individual parameters, which were analysed (see Table 1). The range of the values was obtained on the base of previous measurements. The final and in our opinion critical parameter to consider the quality of a briquette was its density, which ought to lie within the interval from 1 to 1.4 kg/dm^3. The density was evaluated by two steps. The first step was for a density value before stabilization and the second step was represented by density value after stabilization of briquette. The reason is that the density of briquettes is changing with time due to stabilization process and it seems to be important to understand the way that leads to the change as well as to quantify it. [3]

Table 1. Levels of measured parameters

Pressure p [MPa]	Temperature T [°C]	Largeness L [mm]	Humidity w [%]
95 - 159	85 - 115	1 - 4	8 - 12

On the base of measured data, the rate of effect of individual parameters on the final density value was determined. The method of expression of individual parameters effects was applied. Fig. 4 shows horizontal bar chart, which illustrates the influence of individual parameters. As it becomes evident from the chart, the critical parameters that have effect on the final density are the pressing temperature and humidity of compressed material before stabilization as well as after stabilization. [2]

Fig. 5 illustrates the dependence of density change on temperature change.

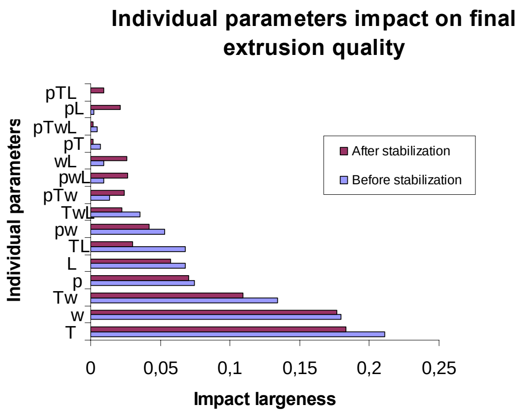


Fig. 4. Rate of individual parameters impact

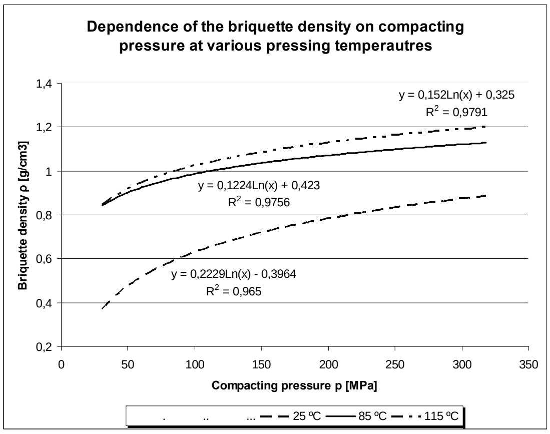


Fig. 5. Dependence of the briquette density on compacting pressure at various pressing tempereatures [2]

5. Conclusion

The main aim of the experiment was to identify the effect rate of observed parameters on the final quality of the briquette, notably the density of a briquette. By the individual steps we discovered that the most critical effect upon the briquette quality has the pressing temperature and then the humidity of material and mutual action of these two parameters. The results testify our hypothesis that compacting pressure, which may seem to be a parameter having the biggest effect on the final quality of a briquette, is minor in analyze of effects on briquettes quality.

Acknowledgements

The research work described in the paper was performed by a financial support of the Slovak Scientific Agency (VEGA), grant No. 1/0265/08 and project SK-SRB-0011-07.

References

[1] Križan, P: Research of factors influencing the quality of wood briquets. *Acta Montanistica Slovaca*. 12 (3): 223-230, 2007.

[2] Križan, P. Svátek, M. Navrhovanie experimentu pre výskum procesu zhutňovania. In proceedings of the International conference of the ERIN, Bratislava, 2008, III-4, (in Slovak).

[3] Jarošová, E. Process variance models for robust design. *Journal of Applied Mathematics*, 1 (1): 251-259, 2008.

[4] Holm, J. Henriksen, U. Hustad, J. Sorensen, L. Toward an understanding of controlling parameters in softwood and hardwood pellets production. published on web 09/09/2006, American Chemical Society

Laboratory Performance Measurements of WiMAX and Wi-Fi Point-to-Point Links

[1,2] J. Pacheco de Carvalho, [1,2] H. Veiga, [3] R. Costa, [1,2] P. Gomes, [2] C. Pacheco, [2]A. Reis

[1]Centro de Informática, [2]U. de Detecção Remota-Departamento de Física,
[3]F. C. Saúde, Universidade da Beira Interior, Rua Marquês d'Ávila e Bolama,
6201-001 Covilhã, Portugal
Email: pacheco@ubi.pt

Abstract. *Wireless communications using microwaves are increasingly important. Performance is a very relevant issue, resulting in more reliable and efficient communications. Laboratory measurements are made about several performance aspects of WiMAX (IEEE 802.16d) and Wi-Fi (IEEE 802.11a, b, g) point-to-point links. An initial evaluation and comparison of these technologies is made, using Alvarion equipments and access points from Enterasys Networks. Through OSI levels 1, 3 and 4, detailed results are presented and discussed, namely: SNR, latency and ICMP packet loss. It is concluded that WiMAX, which is an important emerging and rapidly expanding wireless wide band technology, shows the best performance.*

Keywords: WLAN, Wi-MAX, Wi-Fi, Point-to-Point Links, Wireless Network Laboratory Performance Measurements

1. Introduction

Wireless communications are increasingly important for their versatility, mobility and favourable prices. It is the case of microwave based technologies, e.g. WiMAX and Wi-Fi.

WiMAX permits medium and long distance wireless links, for both point-to-point and point-to-multipoint configurations. It corresponds mainly to the IEEE 802.16 standard and its evolution [1]. There is a fixed version of WiMAX based on IEEE 802.16d or 802.16-2004, allowing for data rates up to 75 Mbps. WiMAX has become interesting for enabling alternatives to traditional telecommunication infrastructures such as phone and coax TV cabling, through technologies such as VoIP and IPTV, while maintaining the capacity for Internet services. IEEE 802.16e/g has provided for mobility, permitting the use of cellular-like phones moving at considerable speeds. WiMAX has, efficiently, enabled to extend medium and long range backbones. Frequency bands of 5.4 and 5.7 GHz are used. Analysis and performance evaluation of IEEE 802.16 has been made [2]. WiMAX performance has been studied [3].

The importance and utilization of Wi-Fi have been growing for complementing traditional wired networks. Wi-Fi has been used both in ad hoc mode and infrastructure mode. In this case an access point, AP, is used to permit communications of Wi-Fi devices with a wired based LAN through a switch/router. In this way a WLAN, based on the AP, is formed. Wi-Fi has reached the personal home, forming a WPAN, allowing personal devices to communicate. Point-to-point and point-to-multipoint configurations are used both indoors and outdoors, requiring specific directional and omnidirectional antennas. Wi-Fi uses microwaves in the 2.4 and 5 GHz frequency bands and IEEE 802.11a, 802.11b and 802.11g standards [1]. Nominal transfer rates up to 11 (802.11b) and 54 Mbps (802.11 a, g) are permitted. CSMA/CA is the medium access control. Wi-Fi performance is available in indoor environments [4].

Performance has been a very important issue, resulting in more reliable and efficient communications. Several measurements have been made for 2.4 GHz Wi-Fi [5], as well as WiMAX and high speed FSO [6,7]. In the present work further results arise, through OSI levels 1, 3 and 4. WiMAX and Wi-Fi (IEEE 802.11 a,b,g) are evaluated and compared for performance in laboratory measurements of point-to-point links.

The rest of the paper is structured as follows: Chapter 2 presents the experimental details i.e. the measurement setup and procedure. Results and discussion are presented in Chapter 3. Conclusions are drawn in Chapter 4.

2. Subject and Methods

WiMAX measurements were performed using the Alvarion BreezeNET B100 model, operating in the 5.4 GHz band. This equipment supports, namely: IEEE 802.3 CSMA/CD; OFDM; BPSK, QPSK, 16-QAM, and 64-QAM modulations; 21 dBi antennas; high RF saturation, with high performance; NLOS; 802.1Q; QoS; a maximum transfer rate of 108 Mbps (turbo mode) [8]. Modulation was set to 8, 64-QAM-3/4, with signal to noise ratios above 23dB. In the tests, 54 Mbps (normal mode) and vertical polarization were used. There were interference free conditions.

Wi-Fi experiments of two types were carried out: EXP1 and EXP2. EXP1 used Enterasys RBTR2 level 2/3/4 access points (AP1) with firmware version 6.08.03 [9], and 100-Base-TX/10-Base-T Allied Telesis AT-8000S/16 level 2 switches [10]. The configuration was for minimum transmitted power, micro cell, point-to-point, LAN to LAN mode, using the radio card antenna. EXP2 used Enterasys RBT-4102 level 2/3/4 access points (AP2) with firmware version 1.1.51 [9], and the same type of level 2 switch. The configuration was similar to EXP1. In both EXP1 and EXP2 interference free communication channels were used. WEP encryption was not activated. No power levels above the minimum were required as the APs were very close (30 cm).

WiMAX and Wi-Fi laboratory tests were made in point-to-point mode, using the setup scheme shown in Fig. 1, where RB and BU represent remote bridge and base unit, respectively. The RB-BU distance was 100 m. For 7-echo UDP traffic injection (OSI level 4) the WAN Killer software was available [11]. Packet size was set to the default of 1500 bytes. The traffic injector was the PC with IP 192.168.0.1, having the PC with IP 192.168.0.5 as destination. Latency was measured as the round trip time of ICMP packets (OSI level 3) involving the PCs having IPs 192.168.0.2 and 192.168.0.6. Percentage packet loss was also measured for different ICMP packet sizes (32 and 2048 bytes) through the same two PCs.

3. Results and Discussion

The main WiMAX results are shown in Fig. 2. A polynomial fit was made to percentage packet loss for 32 byte ICMP packets. The data indicates that, for traffic peaks up to 70 % of the maximum nominal transfer rate (54 Mbps), communications quality is possible as recorded in Table1, as the values of latency and percentage packet loss are within acceptable limits (less than 10 ms and 2%, respectively). For Wi-Fi experiments both AP1 and AP2 were configured, for every standard, with typical fixed transfer rates. For every experiment and fixed transfer rate, measurements were made, with percentages of injected traffic varying from 0% to maximum values. At OSI level 1, typical SNR values are shown in Fig. 1. The best noise levels N were for 802.11a. Measurements through OSI levels 3 and 4 permitted determination, for every standard and fixed transfer rate, of the maximum percentage of network utilization under quality conditions. Some sensitivity to AP type was observed. Table 1 shows the average values obtained from EXP1 and EXP2. It was found that, for every standard, the maximum percentage of network utilization under quality conditions decreases with increasing fixed transfer rate. Some results obtained for 802.11g, at a 54 Mbps are illustrated through Fig. 2. Table 1 shows that WiMAX has the best performance (70%). It is followed by Wi-Fi: 802.11g and 802.11a at 54 Mbps, having 40% each.

Table 1. WiMAX /Wi-Fi average maximum percentages of network utilization under quality conditions versus IEEE802.16d/IEEE 802.11 a, b, g standards and fixed transfer rate (Mbps).

IEEE standard/ Fixed transfer rate	5.5 (Mbps)	11 (Mbps)	12 (Mbps)	36 (Mbps)	54 (Mbps)
802.16d					70%
802.11b	70%	45%			
802.11g			60%	50%	40%
802.11a			65%	55%	40%

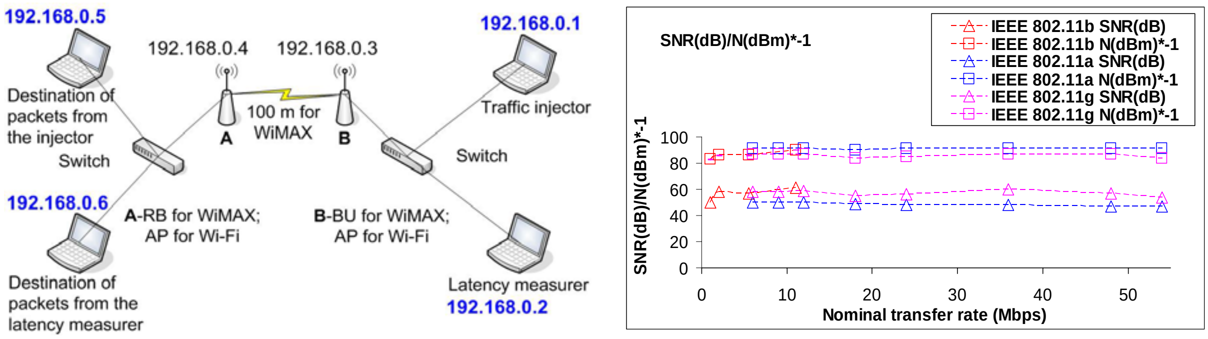


Fig. 1. WiMAX and Wi-Fi laboratory setup scheme, and Wi-Fi typical SNR (dB) and N (dBm).

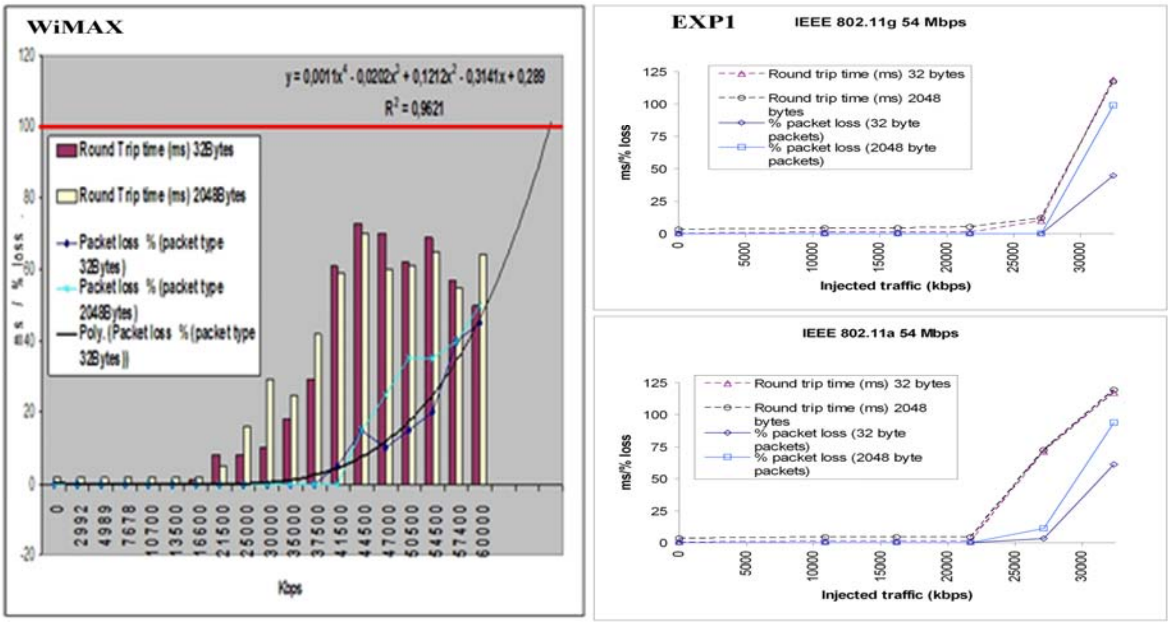


Fig. 2. Synthesis of results showing latency (ms) and ICMP percentage packet loss for WiMAX [6] and Wi-Fi (802.11g and 802.11a at 54 Mbps) experiment 1 (EXP1).

4. Conclusions

In the present work some systematic performance measurements were made using available equipments: both WiMAX (802.16d) and Wi-Fi (IEEE 802.11a, b, g) in point-to-point links. Through OSI levels 3 and 4, the measurements permitted to find maximum percentages of network utilization under conditions of communications quality. WiMAX was found to have the best performance. Further tests are required and planned mainly at OSI levels 4 and 7, using both normal and turbo mode. For Wi-Fi, the maximum percentage of network utilization under quality conditions was found to decrease with increasing fixed transfer rate. Additional measurements either started or are planned using several equipments, both in laboratory and outdoors involving, mainly, medium range links.

Acknowledgements

Supports from Universidade da Beira Interior and FCT (Fundação para a Ciência e a Tecnologia)/POCI2010 (Programa Operacional Ciência e Inovação) are acknowledged. We acknowledge Alvarion and Enterasys Networks for their collaboration.

References

[1] Web site http://standards.ieee.org

[2] Hoymann C. Analysis and performance evaluation of the OFDM-based metropolitan area network IEEE 802.16. *Computer Networks*, 49: 341-363, 2005.

[3] Mach P, Bestak R. WiMAX Performance Evaluation. In Proc. ICN'07- 6th International Conference on Networks, Sainte Luce, Martinique, France, 22-28 April 2007, 17-20.

[4] Sarkar N, Sowerby K. High Performance Measurements in the Crowded Office Environment: a Case Study, In Proc. ICCT'06-International Conference on Communication Technology, Guilin, China, 27-30 November 2006, 1-4.

[5] Pacheco de Carvalho JAR, Gomes PAJ, Veiga H, Reis AD. Development of a University Networking Project, in Encyclopedia of Networked and Virtual Organizations, Putnik GD, Cunha MM, Editors. IGI Global, Hershey, PA (Pennsylvania), 2008, 409-422.

[6] Pacheco de Carvalho JAR, Veiga H, Costa R, Gomes PAJ, Reis AD. A Contribution to Experimental Performance Evaluation of Point-to-Point WiMAX Links. In Proc. ISSPIT 2008-8th IEEE International Symposium on Signal Processing and Information Technology, Sarajevo, Bosnia and Herzegovina, December 16-19, 2008, 150-153.

[7] Pacheco de Carvalho JAR, Veiga H, Gomes PAJ, Ribeiro Pacheco CFFP, Reis AD. Experimental Performance Study of a Very High Speed Free Space Optics Link at the University of Beira Interior Campus: a Case Study. In Proc. ISSPIT 2008-8th IEEE International Symposium on Signal Processing and Information Technology, Sarajevo, Bosnia and Herzegovina, December 16-19, 2008, 154-157.

[8] Web site http://www.alvarion.com

[9] Web site http://www.enterays.com

[10] Web site http://www.alliedtelesis.com

[11] Web site http://www.solarwinds.net

A Bit Error Rate Model of M-ary PSK OFDM Satellite Communication Systems for Educational Purpose

I. Bacic, K. Malaric

University of Zagreb, Faculty of Electrical Engineering and Computing
Department of Wireless Communication, Zagreb, Croatia
Email: iva.bacic@fer.hr

Abstract. *In this work, after a basic theoretical background on orthogonal frequency division multiplexing, our virtual bit error rate mode of M-ary PSK OFDM satellite communication systems, realised in Matlab, is presented. Elementary guide for the program is given: by simply typing the wanted parameters, communication system is defined. BER for this system can be plotted. This program is especially suitable for educational propose. Furthermore, in this work, influence of different PSK modulation, number of subcarriers and length of cyclic prefix on the signal performance and communication quality is analysed. It was shown that the BPSK gained the best signal performance, number of subcarriers has no influence on the transmission quality and by increasing the cyclic prefix BER is smaller.*

Keywords: OFDM, M-ary PSK, Subcarrier, Cyclic Prefix, BER, Matlab

1. Introduction

In satellite communication high transmission quality and availability are essential [1]. For measuring signal performance, a criterion called the bit error rate, BER is used. BER represents the number of incorrectly received bits to the number of transmitted bits and is, commonly plotted against E_b / N_0, where E_b stands for energy per bit and N_0 is the noise power spectral density. To improve communication quality in a communication system, a parallel-data transmission of symbols such as Orthogonal Frequency Multiplexing, where the total transmission bandwidth is split into a number of orthogonal subcarriers is used [2].

In this work, software for analysing the signal performance and communication quality is represented. After defining the elementary parameters for an OFDM system the program results with a BER versus E_b / N_0 graph. Described program was developed in Matlab.

2. Theoretical background

OFDM involves sending several signals at a given time over different frequency channels, or subcarriers which are orthogonal indicating there is a precise mathematical relationship between carrier frequencies ensuring that the modulation symbol can be recovered from the transmitted signal without intersymbol interference, ISI [3] [4]. OFDM divides serial data stream into several parallel streams, one for each subcarrier. Each subcarrier is modulated with a conventional modulation scheme. M-ary PSK scheme used in satellite communication systems [5] [6], is a digital phase modulation of a sinusoidal carrier:

$$s_i(t) = \sqrt{\frac{2E_S}{T}}\cos(2\pi f_C t)\cos(\frac{2\pi i}{M}+\frac{\pi}{4}) - \sqrt{\frac{2E_S}{T}}\sin(2\pi f_C t)\sin(\frac{2\pi i}{M}+\frac{\pi}{4}) \tag{1}$$

where Es is the symbol energy and i is the imaginary number $i = \sqrt{-1}$. Parameter M is chosen as a power of 2 since the data to be converted are usually binary [7]. PSK data bits are grouped into unique pattern to form symbols, represented by a particular phase which is sent across the channel after modulating the carrier. Receiver determines phase of received signal and maps it to the symbol it represents. Usually Gray mapping is used for allocating bits to a symbol providing that adjacent symbols differ by only one bit [8]. This gives the best immunity to corruption. OFDM treats the PSK symbols as though they are in the frequency-domain. These symbols are feed to an Inverse Fast Fourier Transform, IFFT block which brings the signal into the time-domain and is defined with [9]:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N}nk} \qquad n = 0,1,2 \ldots N-1 \tag{2}$$

where N is the IFFT size, chosen as a power of 2 and has to be significantly larger than the number of subcarrieres (like show in Tab.1) to ensure that the edge effects are neglectible at half the sampling frequency [4]. After IFFT is taken, the cyclic prefix, CP can be added. CP is a copy of the last part of the OFDM symbol which is appended to the front of transmitted OFDM symbol and is introduced because of no perfect synchronization between transmitter and receiver (therefore, wrong messages may be received) [10]. Its duration is determined by the expected duration of the multipath channel in the operating environment and is usually set to 1/32, 1/16, 1/8 or 1/4 of duration of whole OFDM symbol [10]. Extended message is transmitted through the channel. At the receiver's side, after removing the cyclic prefix, the Fast Fourier Transform block performs the reveres process on the received signal and brings it back to frequency domain to be demodulated [12].

Table 1. The FFT size and number of subcarriers used in communication systems

IFFT/FFT size	128	256	512	1024	2048
Number of subcarriers	72	192	360	720	1440

3. Virtual bit error rate model program

OFDM system explained in previous section is implemented in a test program for analysing the bit error rate. For deal the problem, Matlab was used.

Program's environment is user friendly: simply by typing the wanted parameters (type of PSK modulation, number of subcarriers and length of cyclic prefix) into the graphical user interface, the communication system is defined. Clicking on button *plot*, the result, a BER versus E_b / N_0 graph with its legend is plotted on the right side of the GUI. After changing parameters, another graph may be plotted at the same figure for comparison the influence of different parameters on the signal performance. Our program provides up to seven curves at the same figure (more than that would be confusing). By clicking the button *clear* the whole figure is deleted, and a new analysis is available. Optionally, grid can be added to the graph (button *grid*).

Thus defined, our test program is especially suitable for educational propose.

4. Results

In the following, the computational results of BER model of M-PSK OFDM satellite communication system are introduced.

First the influence of M is observed. For every pair in Tab. 1 BER versus E_b / N_0 graphs with different modulations are plotted and for all, the same conclusion was made: lower parameter M gives better communication quality. For measuring the same BER level, the BPSK ($M = 2$) allows the highest level (E_b / N_0 is lower) in the channel. For illustration Fig.1 with the results for three modulations; $M = 2, 4, 8$ for the first values of IFFT size and number of subcarriers in Tab. 1 is given.

IFFT size, N is chosen as a power of 2 (first row in Tab.1). Let's observe the influence of number of data subcarriers in an OFDM system with fixed IFFT size and only one modulation type. A representative result is given on Fig. 2. It is seen that the number of subcarriers has no influence on the transmission quality but their amount is directly related with the transmission rate. The more carriers used, transmission rate is higher [4]. Limitations on number of subcarriers are date rate defined by norms and also the highest possible power of the signal (signal power is higher if there are more subcarriers) [4] [12].

On Fig. 3 can be seen that increasing the cyclic prefix duration improves the BER performance of the OFDM system and the best results were gained with length of CP set to 1/4 of the whole OFDM symbol length. Although this length gained lowest data rate, it's mostly used because of providing best protection.

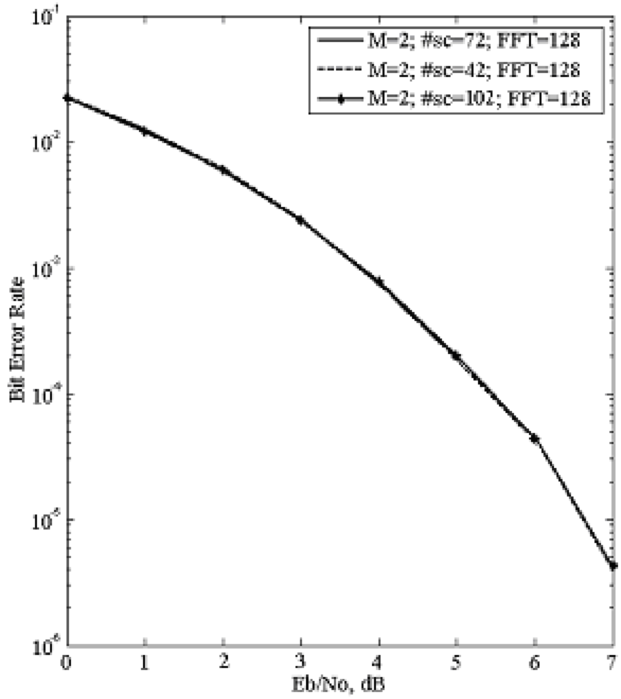


Fig.2. BER vs. E_b / N_0 graph: influence number of subcarriers (right)

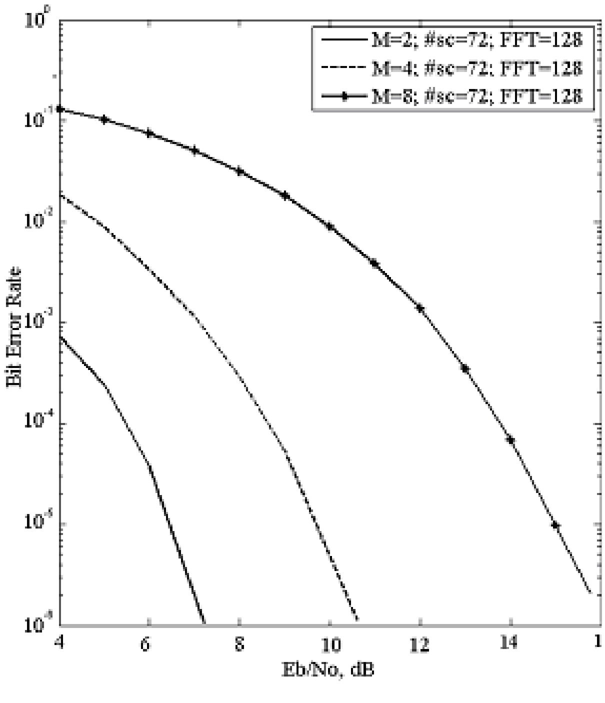


Fig.1. BER vs. E_b / N_0 graph: influence of modulation parameter M

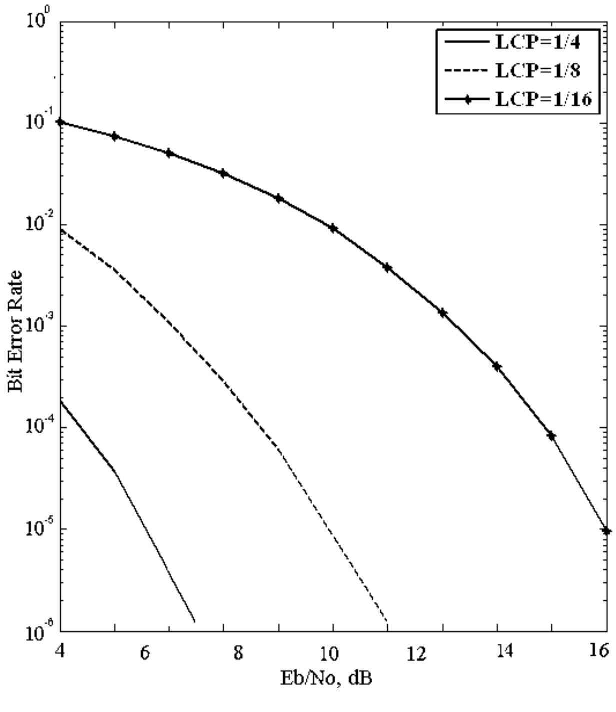


Fig. 3. BER vs. E_b / N_0 graph: influence of cyclic prefix length

5. Conclusions

This work proposes a virtual BER model for M-ary PSK OFDM satellite communication systems. The program was realised in Matlab. Program's interface is providing a user friendly environment. Simply by typing the wanted parameters, the communication system is defined and the result, a BER versus E_b / N_0 graph is plotted. For comparison of influence of different values of parameter on BER, different graphs may be plotted at the same figure. The test program is especially suitable for educational propose. Furthermore, in this work, analysis of the influence of different types of PSK modulation, number of subcarriers and length of cyclic prefix on the signal performance and communication quality is made. It was shown that the BPSK gained the best signal performance. The number of subcarriers has no influence on the transmission quality but as number of subcarriers increases data rate increases. The limitation on number of subcarriers is the date rate defined by norms and also the highest possible power of the signal. It was shown that increasing the cyclic prefix duration improves the transmission quality, so the cyclic prefix with a length 1/4 of whole OFDM symbol provided best protection but had the lowest data rate.

References

[1] The Telecommunication Union, ITU. Handbook on Satellite Communication (HSC). 3rd Edition, John Wiley & Sons, West Sussex, 2002, 58-71.

[2] Patel J.N, Dalal U.N. A Comparative Performance Analysis of OFDM Using MATLAB Simulation with M-PSK and M-QAM mapping. In Proceedings of the International Conference on Computational Intelligence and Multimedia Applications, 2007, 406-410.

[3] Prasad R. OFDM for Wireless Communications Systems. Artech House, Inc., Berlin-London, 2004, 11-17, 123-125.

[4] Schulze H, Lueders Ch. Theory and Applications of OFDM and CDMA, Editor. John Wiley & Sons, West Sussex, 2005, 145-154, 242-261, 153-154.

[5] Xiong F. Digital Modulation Techniques. Artech house, Inc., Norwood, 2000, 9-17, 123-160, 195-210

[6] Jia L, Qingchong L. PSK Communication Systems using Fully Saturated Power Amplifiers. IEEE transactions on aerospace and electronic systems, 2006, 464-477.

[7] Wilson S.G. Digital Modulation and Coding. Prentice Hall, 1996, 141-161, 185-191.

[8] Information Technology. Digital Satellite Communication, Corazza G, Editor. Springer, New York, 2007, 176-183, 202-203.

[9] Weisstein, E.W. Discrete Fourier Transform, unpublished, Retrieved 2009 from MathWorls—a Wolfram Web Resource's Website: http://mathworld.wolfram.com/DiscreteFourierTransform.html

[10] Tse D, Viswanath P. Fundamentals of Wireless communication. Cambridge University Press, 2005 95-102.

[11] Hasan, M.A. Performance Evaluation of WiMAX/IEEE 802.16 OFDM Physical Layer, Master thesis Helsinki University of Technology, 2007, unpublished, Retrived 2009 from Helsinki University of Technology's Website: http://lib.tkk.fi/Dipl/2007/urn009599.pdf

[12] Briggs M, Martinet J, Bare D. Power Measurements of OFDM Signals. International Symposium on Electromagnetic Compatibility, 2004, 485488.

Investigating of Digital to Analog Converter Using Program Designed in LabVIEW Environment

A. Odon, P. Otomański

Poznan University of Technology, Institute of Electrical Engineering and Electronics, Piotrowo 3A, 60-965 Poznan, Poland
Email: odon@et.put.poznan.pl

Abstract. *An application has been designed in the LabVIEW environment for computer simulation of the operation principle and testing of binary weighted digital to analog converter (DAC). The front panel contains graphical objects, permitting interactive communication with the user. The application permits performance virtual experiments thanks to the possibility of the modification of the DAC electronic circuit parameters. The application can be used both for education and for specific research purposes.*

Keywords: LabVIEW, DAC Modelling, DAC Simulation

1. Introduction

Presentation of the operation principle of digital-to-analog converter (DAC) and experimental study of its properties in the traditional way is time consuming and poses difficulties for the lecturers and students because of the need to present considerable amounts of information in the graphical form. To illustrate the problems related to the DAC operation, many schemes and characteristics describing relevant properties of the system are needed and have to be adapted for each new value of the parameters of the converter scheme. One solution of this problem is to use computer based educational tools for lecturers and students to be used either in lectures or laboratory classes. In particular, the LabVIEW environment [1] seems to be commonly used at the university level teaching process for realisation of specific programs for presentation of various didactic problems. Many authors have been interested in these subject areas and many applications dedicated to the simulation both of the analog to digital and digital to analog converters [2-4] and measuring paths employing different sensors have been published [5].

This paper presents the application designed in the LabVIEW environment for computer simulation of the operation principle and testing of DAC based on binary weighted resistors ladder with the summing amplifier.

2. Simulation and Investigation of the Properties of the Binary Weighted Digital-to-Analog Converter

The binary weighted DAC is based on the circuit with a network of parallel branches (the so called resistor ladder), one branch for each bit. Each branch contains a resistance weighted in proportion to the significance of the bit in the digital word. The resistor ladder cooperates with the current summing amplifier. This is one of the fastest conversion methods but for a large number of ADC bits there are problems with accuracy because too much precision is required for the ladder resistors for the most significant bits. This type of converter is usually limited to 8-bit resolution. Because both the architecture and principle operation of this ADC circuit are relatively simple, so this type of ADC is excellent for educational purposes. Fig. 1 presents the front panel of the designed application for the simulation of a binary weighted DAC and illustrates the principle of operation of this converter.

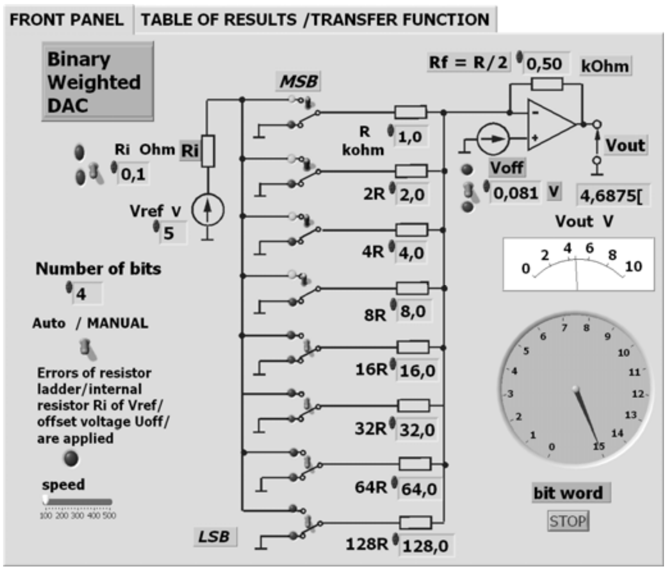


Fig. 1. The front panel of the designed LabVIEW application for the simulation of a binary weighted DAC

The output voltage value for the n-bits DAC is expressed as:

$$V_{out} = -V_{ref}\,\frac{R_F}{R}\left(\frac{1}{2^1}k_1 + \frac{1}{2^2}k_2 + ... + \frac{1}{2^{n-1}}k_{n-1} + \frac{1}{2^n}k_n\right), \tag{1}$$

where V_{ref} is the reference voltage, k_n are individual bits of input word bits, R is value of the resistance in resisor ladder and R_F is a feedback resistor of the current summing amplifier.
The front panel contains a block diagram of the DAC system and controls permitting preset of the desired values of the binary input signal or converter parameters and graphic displays at most important points of the system. The resolution from 2 bits to 6 bits of the DAC may be chosen by the computer operator by presetting a suitable number in the window "number of bits" of the application front panel. The results of simulation are presented in the second panel by using the bookmark "TABLE OF RESULTS/ TRANSFER FUNCTION" in form of plot of transfer function and the data are collected in the table of results." Fig. 2 presents results of simulation 4-bits DAC for the ideal case - that is for the ideal values of DAC circuit parameters.

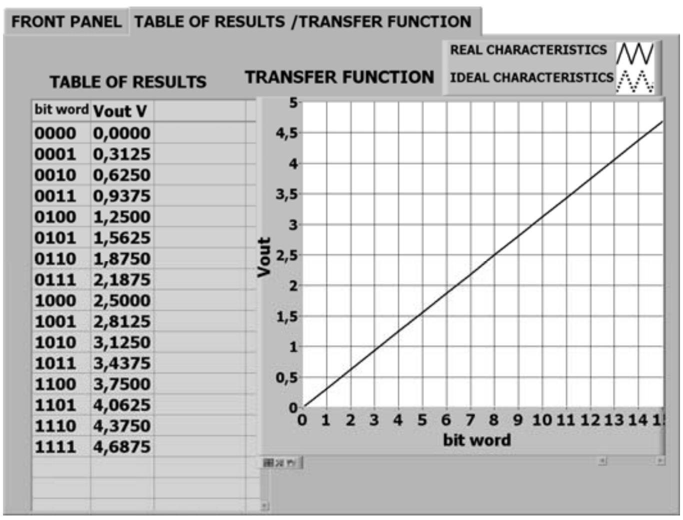


Fig. 2 The results of simulation of 4-bits DAC for the ideal case, all the DAC circuit parameters take the ideal values.

The application permits preset of all important parameters of the D/A circuit including: the reference voltage U_R, value of weighted resistor in each branch, feedback resistor R_F in current summing amplifier and offset voltage simulating the offset voltage of operational amplifier.
Fig. 3 presents the exemplary simulation for non-ideal parameters of DAC. The offset voltage V_{off} = 0.2 V is applied and the other DAC parameters are ideal. Fig. 4 presents the results of this simulation.

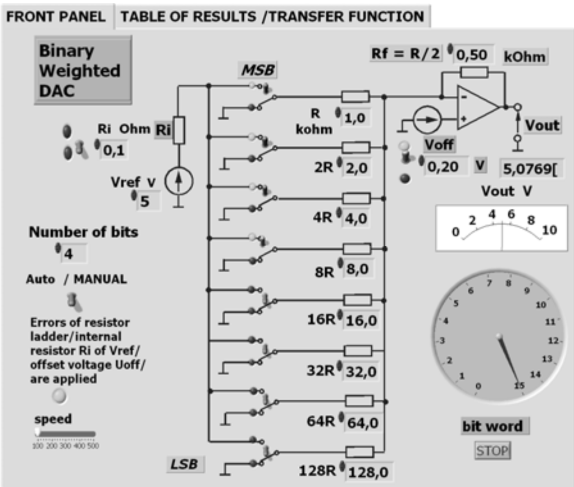


Fig. 3. The front panel for simulation of 4-bits DAC. The offset voltage value V_{off} = 0.2 V. The other DAC parameters have ideal values.

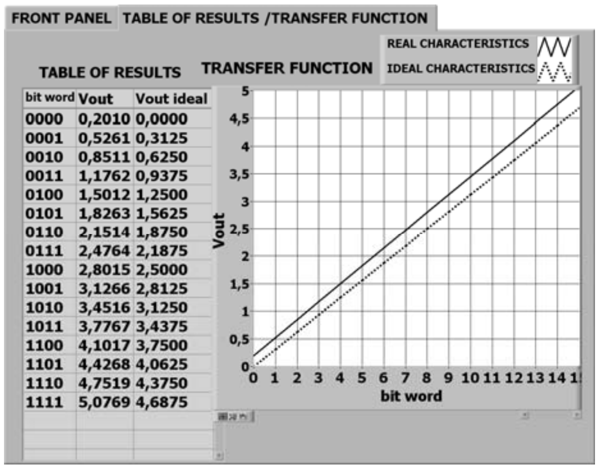


Fig. 4. The results of simulation of 4-bits DAC for the offset voltage value V_{off} = 0.2 V. The other ADC parameters have ideal values.

The results of this investigation presented in Fig.4 show that both an offset error and a gain error of transfer function have occurred. As follows from the results of the simulation presented in Fig.4 (result tables and plot of transfer function), these errors have no effect on the linearity of the transfer function.

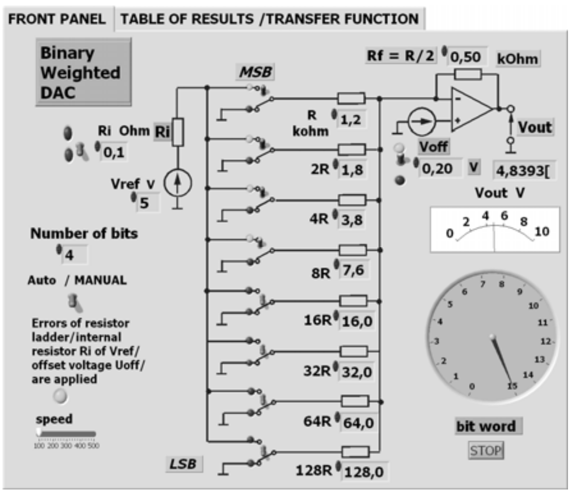


Fig. 5. The front panel for simulation of 4-bits DAC. The offset voltage value V_{off} = 0.2 V. The values of ladder resistors are improper.

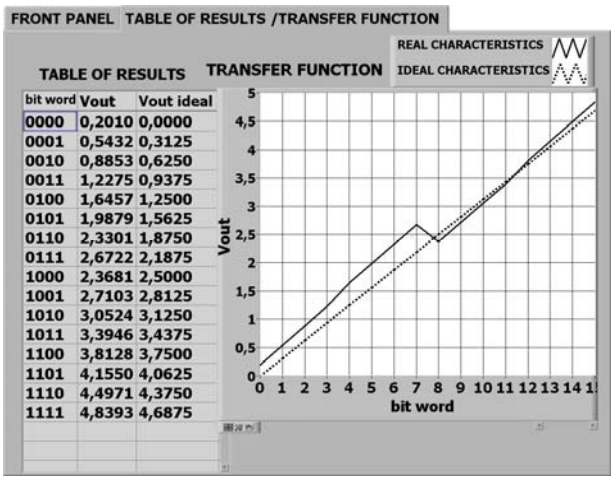


Fig. 6. The results of simulation of 4-bits DAC for the offset voltage value V_{off} = 0.2 V and for improper values of ladder resistors.

Fig. 5 presents results of another simulation for non-ideal parameters of DAC. The offset voltage V_{off} = 0.2 V is applied and the ladder resistors are preset on improper values. Fig. 6 presents the results of this simulation. The results of the simulation presented in Fig. 6 show that improper values of the ladders resistors cause non-linearity of the DAC transfer function and can additionally cause non-monotonicity of the transfer function. All date are collected in the table of results and can be exported to another program (an example Excel) for further data processing such as calculating of linearity errors, gain error, offset error etc.

3. Conclusions

The LabVIEW application written permits presentation of the principles of operation of the binary weighted digital-to-analog converter and performance of simulation experiments thanks to the possibility of preset of the parameters of the digital input signal and those of the elements of the converter system. Students can analyse the effect of changes in the parameters of the DAC system on the plot of transfer function shown in the second window of the front panel which is activated after switch of bookmark. All results of the simulation are collected in table of results and can be exported to another program for further processing. An important advantage of this application over similar ones presented in literature is the possibility of setting a large number of parameters of the system over wide ranges. The application permits performance of virtual experiments thanks to the possibility of modification of the DAC electronic circuit parameters. The application can be used both for education and for specific research purposes. The analysis of the converter work and the effects of different parameters on its performance is conducted in the interactive mode and can be realised with the help of a PC unit only, so the results are achieved very fast.

References

[1] LabVEW TUTORIAL MANUAL, January 1996 Edition, Part Number 320998A-01, National Instruments Corporation, 1993, 1996.

[2] Andrzej Odon, Zbigniew Krawiecki. Using labVIEW for computer modelling of conversion methods in analog to digital converters. In proceedings of the 12th IMEKO TC1 & TC7 Joint Symposium on Man, Science & Measurement September, 3 – 5, 2008, Annecy, France, 2008, 395 - 400.

[3] Tłaczała W., Gorghiu G., García de la Santa A., Kordaki M., Rodriguez J. T. Użycki J. Modeling of the D/A and A/D conversion techniques based on virtual instrumentation. In proceedings of the 4th International Conference on Multimedia and Information and Communication Technologies in Education, Current Developments in Technology-Assisted Education - , Seville, Spain, November 2006, vol. 1, 564 – 568.

[4] L. Referowski, D. Swisulski, Progress in Education of Electrical Measurements. *Electronics and electrical engineering*, 3967, 2006, 25 – 30.

[5] Katranas G.S. et al. Simulation and measurement of bilayer sensor characteristics. *Sensors and Actuators*, A129, 2006, 243-246.

A/D Converter Design Evaluation Engine in MGC Environment

[1]J. Židek, [1,2]O. Šubrt, [1]P. Martinek

[1]Department of Circuit Theory FEE CTU Prague, Technická 2,
166 27 Prague, Czech Republic,
[2]ASICentrum Novodvorská 994, 142 21 Prague, Czech Republic
Email: zidekj1@fel.cvut.cz

Abstract. *Environment for testing static parameters of analog-to-digital converters is presented in this article. It is a novel concept of powerful engine suitable for to educate students in working with modern CAD tools. The source code of each block of the design is written in Verilog-A which offers relatively effortless portability on different design systems (e.g. Cadence). The core of our proposal is based on Servo-Loop with improved search algoritm [1]. The simulation outputs are curves of static INL and DNL. A part of article deals with the example of simple Flash ADC testing.*

Keywords: A/D Converter, Design Verification, CAD Tools, Educational Purpose

1. Introduction

Integral (INL) and differential (DNL) non-linearity are two of basic parameters of A/D converters. The ways of their measurement can be divided in two groups. Algorithms belonging to the so-called open-loop category are advantageous for production test. Best known member of open-loop methods is the histogram methods. This method is, however, hardly applicable for the ADC circuit simulation as it takes too many simulation steps for a given accuracy.

Procedures from the second group (referred to as closed-loop) create a reasonable compromise regarding the simulation requirements; therefore, they are good candidates to be used during circuit-level ADC simulation. The basic method is the standard Servo-Loop algorithm [5], but for much shorter simulation time Improved Servo-Loop Algorithm can be used [1] meeting the same performance specifications.

Recent works in this field are mostly oriented either to measurement level or behavioral model simulations employing mathematical software such as Maple or Matlab. The environment proposed in our article is built up completely in Verilog-A and therefore it can be used in a direct co-operation with analog and mixed-signal circuit simulators (e.g. Eldo, Spectre, Advance MS, etc.) up to full transistor-level complexity without the need of any other computational or post-processing software.

2. Improved Servo-Loop Algorithm

In Fig. 1, block scheme of the proposed Servo-Loop system is outlined. Our Servo-Loop implementation is based on this scheme, suggesting significant improvements against the basic approach [5]: effective usage of the discrete-time integrator, application of the initial condition and refinement of the integrator step.

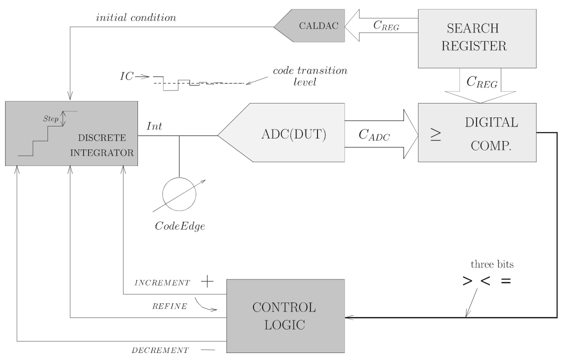


Fig. 1. Block scheme of Improved Servo-Loop

Algorithm Principle and Definitions

The flowchart of the proposed novel algorithm variant is depicted in Fig. 2A). Here, V_{min} and V_{max} are the minimum and maximum ADC input voltages representing the full scale range. BITS is the number of ADC bits, i.e. the output word width. The LastEdge variable stores the value of the previous code transition level, see further explanation below. Finally, V_{lsb} is the code width expressed in term of the input voltage, i.e. the analog input increment corresponding to 1LSB code change of an ideal ADC with the same analog input range as the DUT.

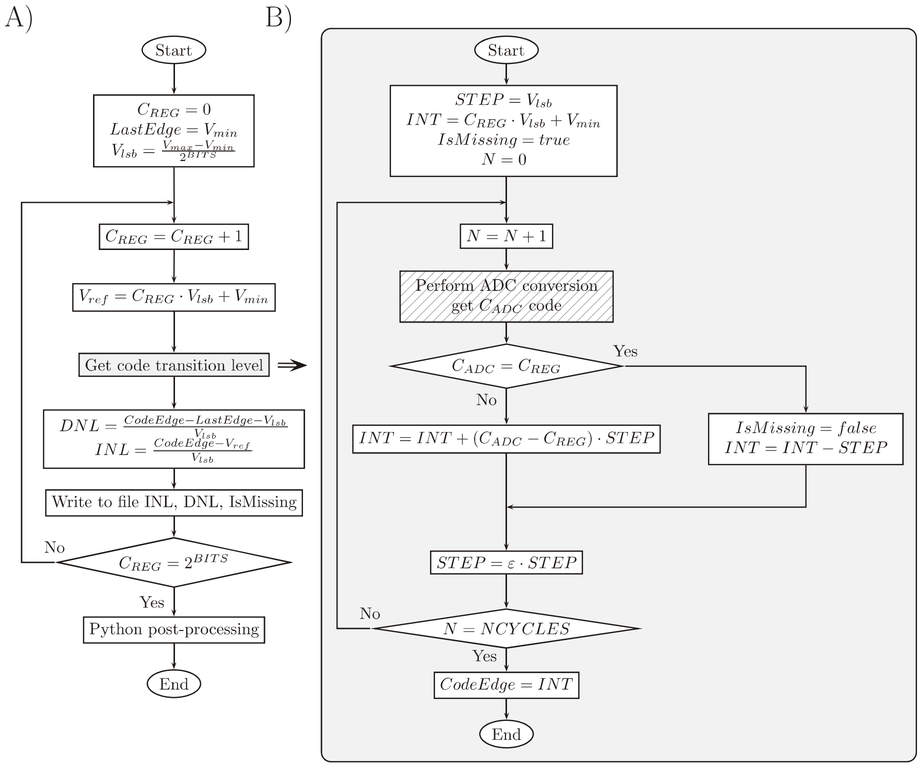


Fig. 2. Algorithm flow chart

In Fig. 2, the initial variables are set immediately after start. The main algorithm cycle is executed for each transition level, i.e. 2^{BITS}-times for the whole set of the ADC codes. The looping statement is ensured by incrementation of CREG variable representing the actual code for which the transition level has to be found. Based on the CREG value, the V_{ref} is calculated which is then used for INL computation. The next step is the most important part of the algorithm formed by the transition level search procedure; it is detailed in grey box in Fig. 2B). First, the initial values of the internal variables are set and after that, single ADC conversion is performed. If the converter output code C_{ADC} is not equal to C_{REG} value the discrete integrator output INT is changed; its incrementation or decrementation dependents on the actual C_{ADC} and C_{REG} relationship. Here, the C_{ADC}-C_{REG} term ensures a quick convergence action in case that the actual C_{ADC} is too far from C_{REG} target. During the first loop cycle, the STEP value equals to 1LSB and it is continually refined to converge to the desired code transition level. At this point, it is important to note that the lower step transition level definition is applied [5]. The STEP size refinement by $\varepsilon<1$ constant is done at the end of each cycle. Once the C_{ADC} code equals to C_{REG} during the whole search procedure, the IsMissing boolean variable is reset; it indicates that the appropriate C_{ADC} code is present on the ADC transfer characteristic. The extracted code transition level value is outputted to the main algorithm cycle (2A) the DNL and INL are then calculated. The INL and DNL data, together with the IsMissing variable are written to separate files for the

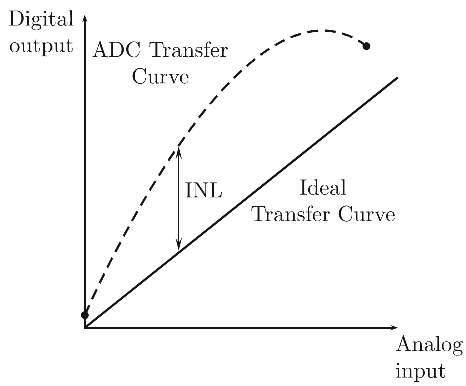


Fig. 3. Basic INL computed by Eldo

next processing in Python script language. The algorithm terminates when the set of code transition levels is complete.

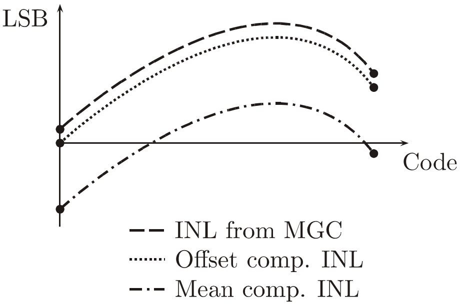


Fig. 4. Offset and Mean comp. INL

Python Extension for Result Postprocessing

Here, it is necessary to notice that the Verilog implementation in MGC software has one specific feature. The file writing subsystem adds unwanted additional lines into the output file together with the useful data. Therefore, it is impossible to format the file in compliance with the EzWave input format. That is why a script in the Python language was used. The proposed implementation can evaluate five types of INL representation (Basic, Best-straight-line, End-Point-Corrected, Offset compensated, Mean compensated).

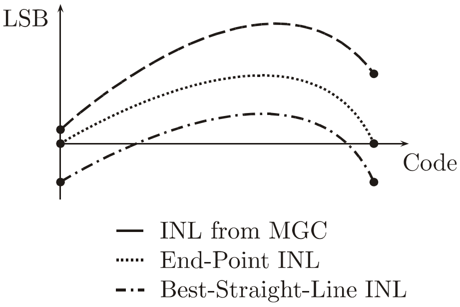


Fig. 5. Best-Straight-Line and End Point Corrected INL

The Fig. 4 and Fig. 5 refer to each INL description. The Basic INL curve is evaluated directly in MGC in coincidence with the equation in Fig. 2A) and with Fig. 3. The meaning of the definitions can be observed directly from Fig. 3 to Fig. 5; detailed explanation can be found e.g. in [3].

System Accuracy

The algorithm resolution is one of the most important parameters. In our implementation, the algorithm can effectively change accuracy by using two parameters. The first parameter is the number of iterations NCYCLES and as it is thoroughly discussed in [1], the algorithm resolution depends also on the ε variable:

$$\Delta_N^{LSB} = \varepsilon^{NCYCLE-1} \tag{1}$$

Where Δ_N^{LSB} is the maximum possible error, ε is the refinement constant and finally, $NCYCLE$ is the number of iteration steps.

3. Environment Implementation

As it was mentioned above, the algorithm was implemented in Verilog-A. Algorithm block diagram is in Fig. 6. Output word from ADC DUT of maximum size of 16 bits is connected to the block labeled as D2A. This block converts the digital signal to a form which can be easily processed by Verilog-A. The next block is the voltage controlled voltage source outputting the difference between the input value (in principle it is C_{ADC}) and the reference value (C_{REG}). This value is led to the input of Step control block computing an appropriate size of the next step. The last block is the discrete integrator with the built-in initial condition; the condition is loaded to the comparator output when reset is at zero level. Each block works only at the time, when its clock signal is active. It is advantageous due to effective usage of simulation time. Clock signal is

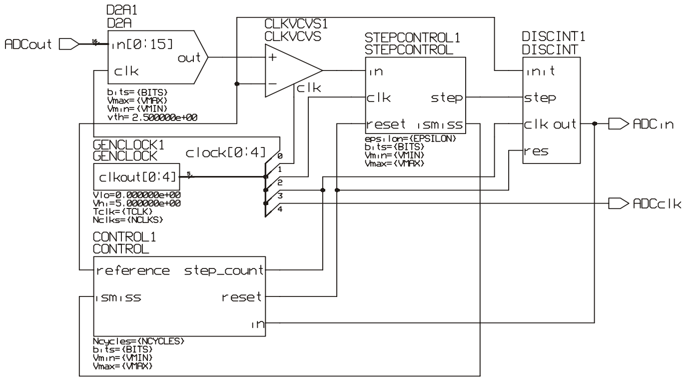


Fig. 6. MGC implementation scheme.

generated by GENCLOCK. The function of CONTROL matches to the A) part of Fig. 3.

4. Simple Flash ADC Testing Example

This section presents simulation result of the Flash ADC in conjunction with the proposed Servo-Loop unit. The ADC is the basic 8-bits realization with resistor chain and ideal comparators. Value of each resistor in chain is 1kΩ execpt R1, R32, R64, ..., R256, which have value 1.5kΩ. In Fig. 7 are displayed DNL and INL for illustration. This shows extrema case of one error source, which can be observed with our implementation.

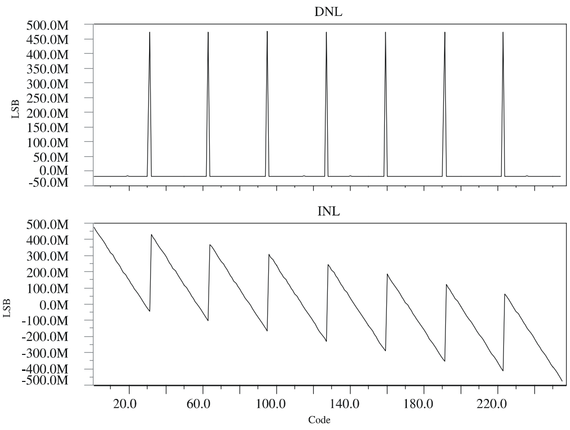


Fig. 7. INL and DNL illustration graph.

5. Conclusion

This work presents an innovative approach to the extraction of ADC performance, suitable for educational purpose. The Servo-Loop unit presented was written as a versatile program module and is suitable for co-operation with any analog simulator supporting behavioral (Verilog-A) device models. The next significant advantage of the Servo-Looper module is the fact that it is capable to extract the static non-linearity of any ADC architecture, described at analog or behavioral simulation level of abstraction. This means that the netlist with extracted parametrs from layout can be used for simulation with our algorithm. This brings nearly the same results as DNL and INL measuring on real chip with no extra time for chip production.

Acknowledgements

The work has been supported by the grant GACR 102/07/1186 of the Grant Agency of the Czech Republic running in co-operation between the Czech Technical University in Prague and ASICentrum Prague. The work has been also supported by the research program MSM6840770014 of the Czech Technical University in Prague. ICstudio®, Advance MS®, Design Architect®, Eldo® and EZwave® are registered trademarks of the Mentor Graphics Corporation. Maple® is a trademark of the Waterloo Maple, Inc., Matlab® is a trademark of the Mathworks, Inc.

References

[1] Šubrt O., Martinek P., Wegener C. A Powerful Extension of Servo-Loop Method for Simulation-based A/D Converter Testing. In proceedings of 15th IMEKO TC4 Symposium on Novelties in Electrical Measurements and Instrumentation, 2007.

[2] Geelen G., A 6b 1.1GSamples CMOS A/D Converter, ISSCC 2001, 128-129, 438.

[3] M. Burns, G. W. Roberts, An Introduction To Mixed-Signal IC Test and Measurement, Oxford University Press, New York, 2001.

[4] Estrada P., Maloberti F., Virtual Test Bench for Design and Simulation of Data Converters, Proc. IEEE Conf. BMAS, Behavioral Modeling and Syst., Santa Rosa, California, 2002, 84- 89.

[5] The Institute of Electrical and Electronics Engineers, Inc., IEEE Standard for Terminology and Test Methods for Analog-to-Digital Converters., IEEE Std. 1241-2000, New York, December 2000.

Measurement of Physical Quantities

Precision Length Measurements by Multi-Wavelength Interferometry

R. Schödel, K. Meiners-Hagen, F. Pollinger, A. Abou-Zeid

Physikalisch-Technische Bundesanstalt (PTB), Bundesallee 100,
D-38116 Braunschweig, Germany
Email: Ahmed.Abou-Zeid@ptb.de

Abstract. *Three different applications of multi-wavelength interferometry are reviewed which have been developed at PTB. They demonstrate the high precision achievable by this technique on multiple scales, from absolute geodetic distances up to 10 m, over prismatic bodies, e.g. gauge blocks, in the sub-meter regime, down to the microscopic roughness of surfaces.*

Keywords: interferometry, diode laser, length measurement, long distance, surface profile, gauge blocks

1. Introduction

Displacement measurements by a laser interferometer are typically performed by moving one reflector of the interferometer along a guideway and counting the periodic interferometer signal, e.g. the interference fringes. Such a counting technique requires a relatively slow continuous movement of the reflector along the entire distance to be measured. When the integer order of interference is lost during the movement, only the fractional order of interference is obtained (phase/2π) resulting in a length unambiguity of $\lambda/2$, only. This small unambiguity can be enlarged by using more than one wavelength, i.e. multi-wavelength interferometry.

An ambiguity of lengths measured by single-wavelength interferometry also occurs when there are length discontinuities at the surfaces of objects to be measured resulting in phase steps. Using two wavelengths $\lambda_{1,2}$, the difference of two interferometer phases acts like a single phase of a "synthetic" wavelength $\Lambda= \lambda_1 \cdot \lambda_2/(\lambda_1 - \lambda_2)$ which is longer than both optical wavelengths. Within half of this synthetic wavelength no counting of interference fringes is necessary. The measurement uncertainty, however, is increased if the measured length L is calculated using the synthetic wavelength. Each uncertainty in both, phase measurements and in the wavelengths is scaled by the ratio of synthetic to optical wavelength. To overcome this problem, it is effective to use the synthetic wavelength only for calculation of the fringe order of the optical wavelengths and benefit from the lower uncertainty of the latter for the length measurement. In this case the measurement uncertainty with the synthetic wavelength must not exceed one quarter of the optical wavelengths to get their correct fringe order. The maximum possible synthetic wavelength is therefore limited. To get a larger range of unambiguity more wavelengths can be used which offer different synthetic wavelengths. Starting from the longest one, each synthetic wavelength is used to get the fringe order of the next shorter one and finally that of an optical wavelength.

In general, multiple lasers are necessary for the multi-wavelength technique whose beams have to be aligned on the same path of the interferometer. These must be detected simultaneously after their separation at the interferometer output for avoiding errors due to the thermal and mechanical drift of the setup.

The effort can be reduced by using wavelength tunable diode lasers. The specific features of such lasers are their tunability by variation of the injection current and the availability of wavelengths from infrared to blue light. This offers a wide range of possible synthetic wavelengths. Depending on the application, the frequency of diode lasers has to be stabilised. By simple parameter stabilisation, i.e. injection current and heat sink temperature, the vacuum wavelength variation is of the order of $\Delta\lambda/\lambda > 10^{-7}$ within days and up to 10^{-5} long term [1]. More sophisticated methods like stabilising the laser on an external cavity or onto atomic or molecular absorption lines, e.g. rubidium, potassium, or iodine, improves the stability by orders of magnitude ($< 10^{-9}$) [2], [3].

This paper gives an overview of three different applications of multi-wavelength interferometry which are well established at PTB: precise measurements of absolute distances (Chap. 2), surface profiles (Chap. 3) and length of prismatic bodies (Chap. 4).

2. Precise measurements of absolute distances

The measurement of distances in the order of ten metres with a relative measurement uncertainty of 10^{-6} is important in a variety of practical applications such as the positioning of components in automotive engineering or aircraft construction, or the inspection of windmill blades. Optical measurements of such distances are typically performed by laser trackers. Recent laser trackers are based on two different length measurement methods: a standard counting interferometer with a HeNe laser and an absolute distance measuring system using a time of flight measurement with an amplitude modulated diode laser. The uncertainty of the absolute length measurement is larger than some 10 µm. Conventional counting interferometers provide a smaller uncertainty but require that the beam follows the reflector movement so that the interference fringes can be counted continuously. In practice, this requires increased efforts and leads to a considerable increase of measurement time. Under well controlled ambient conditions relative measurement uncertainties of 10^{-7} can be achieved.

Absolute distance interferometers (ADI) [4]-[7] achieve a better resolution than systems measuring the time of flight. They could meet the above demand concerning a relative measurement uncertainty of 10^{-6}. Nevertheless, they are not widely used in practice till now, probably due to the complexity of their implementation. Here, an approach to an ADI based on a homodyne interferometer with two diode lasers is presented.

Absolute distance interferometry with a variable synthetic wavelength is performed by use of a laser whose emission frequency ν can be tuned continuously. Diode lasers, with a large mode-hop free tuning range for this purpose, are usually used. In the case of an external cavity diode laser the tuning of the laser frequency $\Delta\nu$ (typically 50…100 GHz) is obtained by changing the diffraction angle by tilting the grating (Littrow type) or by tilting the mirror (Littman type) in the external resonator. In case of a distributed feedback laser diode (DFB) the injection current is modulated. The laser frequency is usually tuned periodically by an oscillator.

In an interferometer the phase Φ is generally given by

$$\Phi = \frac{4\pi}{\lambda_0} nL = \frac{4\pi\nu}{c} nL \, . \tag{1}$$

Here, λ_0 is the laser wavelength in vacuum, ν the laser frequency, c is the velocity of light in vacuum, L is the arm length difference in the interferometer, and n is the refractive index of air. The latter requires the measurement of the ambient parameters (temperature, air pressure, humidity) and can be derived from an Edlén-type empirical equation [8].

In an absolute interferometer the length L_{ADI}, i.e. the length difference between the two interferometer arms, can be calculated from the phase change $\Delta\Phi$ caused by the frequency tuning $\Delta\nu$ according to

$$L_{ADI} = \frac{c}{4\pi n}\frac{\Delta\Phi}{\Delta\nu}. \tag{2}$$

In this case $\Delta\nu$ may be obtained from the transmission of a Fabry-Pérot resonator. The laser frequency at the transmission peaks is given by $\nu = \nu_0 + i \cdot FSR$, with FSR being the free spectral range of the resonator, i the number of the transmission peaks which are swept during the tuning, and ν_0 the initial frequency. Thus, the factor $\Delta\Phi/\Delta\nu$ in (2) is obtained as the slope of a linear interpolation of the phases $\Delta\Phi(i)$.

Alternatively, the light of the laser can be coupled simultaneously into the ADI and into a reference interferometer of a known length L_{ref} which has to be determined independently. The length L_{ADI} of the ADI can be calculated from the phase changes in the ADI ($\Delta\Phi_{ADI}$) and in the reference interferometer ($\Delta\Phi_{ref}$):

$$L_{ADI} = L_{ref}\frac{n_{ref}}{n_{ADI}}\frac{\Delta\Phi_{ADI}}{\Delta\Phi_{ref}}. \tag{3}$$

The refractive indices have to be considered in the ADI (n_{ADI}) and in the reference interferometer (n_{ref}). If both interferometers are placed closely together, the refractive indices are almost equal, and their influence almost cancels out. During the tuning of the laser frequency several phase values from both interferometers are recorded and then fitted to a linear function which passes the origin. The number of data pairs for the fit is only limited by the measurement rate (speed of the detector electronics and A/D conversion). This is a significant advantage compared to the use of a Fabry-Pérot resonator as reference because in the latter case only a limited number of transmission peaks can be used for data sampling.

Length changes, due to, e.g., reflector vibrations that occur during the tuning of the laser frequency, contribute considerably to the measurement uncertainty of frequency sweeping interferometry. If the length is not constant in time, the measured phase values $\Delta\Phi_{ADI}$ in (2) and (3) correspond to different values of L_{ADI}. Variations of L_{ADI} are scaled up with the ratio $\nu/\Delta\nu$, which is typically in the range of 10^3 to 10^4. In the case of vibrations, the impact on L_{ADI} can be reduced by averaging over an appropriate number of single measurements.

In our approach to absolute distance interferometry, we monitor changes of the length L_{ADI} during tuning by a conventional counting interferometer with a permanently frequency stabilised laser with the vacuum wavelength λ_1, which runs parallel to the ADI. The phase change $\Delta\Phi_{ADI}$ of the ADI originating from frequency tuning is corrected by the result of the stabilised laser:

$$\Delta\Phi_{ADI}(corr.) = \Delta\Phi_{ADI} - \frac{\lambda_1 n_2}{\lambda_2 n_1}\Delta\Phi_1. \tag{4}$$

Here, $\Delta\Phi_1$ is the measured phase change of the conventional interferometer. n_1 and n_2 are the refractive indices for λ_1 and λ_2 in the ADI. They differ from each other only due to the dispersion in air since both beams are on the same optical path. The central vacuum wavelength of the tuned laser is denoted as λ_2. By inserting (4) into (3), the final result L_{ADI} for the coarse measurement is obtained.

The range of unambiguity can be extended by using more than one wavelength in the interferometer. If the interferometer is operated with two optical wavelengths λ_1 and λ_2 the difference $\Delta\Phi_{synth}$ between the phases is given by

$$\Delta\Phi_{synth} = \Phi_2 - \Phi_1 = \left(\frac{4\pi}{\lambda_2} - \frac{4\pi}{\lambda_1}\right) nL = \frac{4\pi}{\Lambda} nL \,,$$

$$\Lambda = \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} \,. \tag{5}$$

The phase difference $\Delta\Phi_{synth}$ corresponds to the phase of a fixed synthetic wavelength Λ. It is larger than the optical wavelengths and determines the range of unambiguousness to $\Lambda/2$. Here the wavelengths are assumed to be closely together so that the refractive index n is approximately identical for both wavelengths. By considering not only the phase difference, but the phases itself the range of unambiguousness can be extended to more than $\Lambda/2$. After moving the reflector by one period of the synthetic wavelength the phase difference $\Delta\Phi_{synth}$ is the same as before, but the single phase values are usually different after the movement. When the phase measurement itself has a low uncertainty, a "phase recovery criterion" can be applied. In principle such criterion corresponds to the "method of exact fractions" as further discussed in Chap. 4.

If several synthetic wavelength are used, the longest one determines the unambiguous measuring range. The measurement uncertainty increases according to the ratio of the synthetic to the optical wavelengths. Therefore, the shortest synthetic wavelength gives the lowest uncertainty. The length result from one synthetic wavelength is typically used to determine the fringe order of the next shorter synthetic wavelength.

Since the absolute distance interferometry with a variable synthetic wavelength uses a frequency stabilised laser 1 and a second frequency swept laser 2, it seems straightforward to expand this setup for a measurement with a fixed synthetic wavelength. For that purpose the frequency sweep from laser 2 has to be stopped, and the laser has to be stabilised. The measurement consists of two steps: First, the integer order of interference for Λ is determined by the measurement with variable synthetic wavelength according to (3) and (4). After switching the modulated laser in the stabilised mode, the fractional order of interference is determined according to (5). Thus, the final result L of the two-stage length measurement is

$$L = \left\lfloor \frac{L_{ADI}}{\Lambda/2} - \frac{\Delta\Phi_{synth}}{2\pi} + \frac{1}{2} \right\rfloor \frac{\Lambda}{2} + \frac{\Delta\Phi_{synth}}{2\pi n} \frac{\Lambda}{2} \,, \tag{6}$$

where $\lfloor x \rfloor$ denotes the floor function which returns the largest integer $\leq x$. The synthetic phase $\Delta\Phi_{synth}$ is measured without counting fringes and is in the range $\Delta\Phi_{synth} \in [0...2\pi]$. The subtraction of the term with $\Delta\Phi_{synth}$ and the addition of ½ in the floor function has a practical reason: without these terms the equation is mathematically correct, but the measurement uncertainties lead to a scatter of L_{ADI}. At the border between two fringe orders even a small scatter can lead to the wrong fringe order. Both terms shift the value in the floor function mathematically to constant levels of 1/2, 3/2, 5/2, …, which lead to the fringe order 0,1,2,…. Therefore, scatter smaller than $\Lambda/4$ has no influence on determining the fringe order. It should be noted that this absolute interferometer can also be applied as laser tracker, i.e. as a standard counting interferometer using the permanently stabilised laser.

The optical setup of our ADI as a homodyne Michelson interferometer is depicted in Fig. 1. Laser 1 is a Littrow type extended cavity diode laser (ECDL) operating at approx. 770.1 nm wavelength, and laser 2 is a Littman type ECDL at approx. 766.7 nm wavelength. Both lasers are equipped with antireflection coated laser diodes. For both ECDLs, a small amount of their intensity is split off by a combination of a half-wave plate and a polarizing beam splitter and is fed to a setup using polarization-optical differential detection for the stabilisation to Doppler-reduced potassium (K) absorption lines as described for Rubidium in [2]. Laser 1 is permanently stabilised to a K-D1 line. The emission frequency of laser 2 is tuned by the use of a programmable function generator. A low pass filtered and smoothed triangle signal with a frequency of 5…10 Hz is used for the modulation.

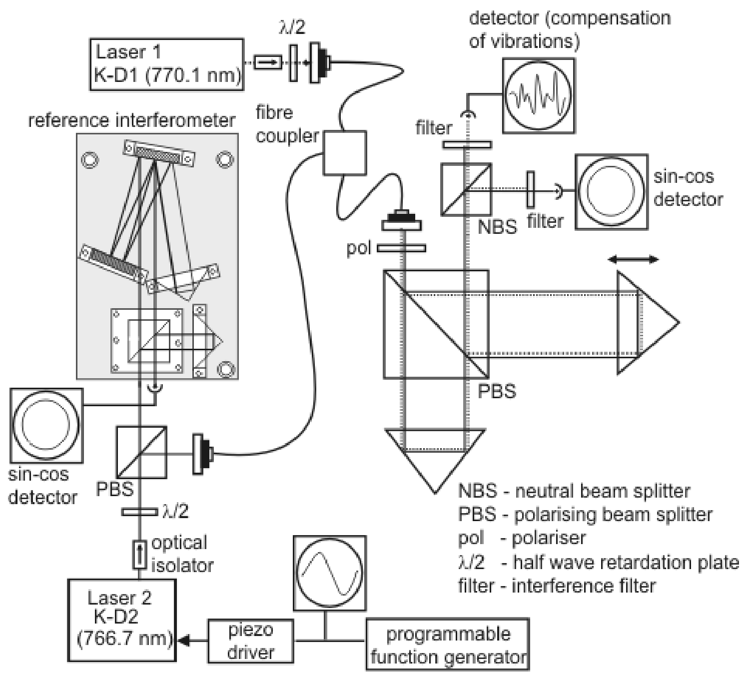


Fig. 1. Optical setup of the absolute distance interferometer with two ECDLs. The light of both diode lasers is coupled into a Michelson interferometer via one polarization maintaining single mode fibre. For clarity, the details for the stabilisation of the ECDLs to potassium absorption lines are not shown.

A portion of the modulated laser's intensity is coupled into the reference interferometer which is placed in a temperature stabilised box. For the ADI, both laser beams are coupled into one polarisation maintaining single mode fibre in order to ensure that they traverse the interferometer on identical paths. By the use of a half-wave plate in front of and a polarizer behind the fibre, matching of the polarization of the lasers light with the fibre axis is achieved. At the interferometer output the two beams are separated by a grating for detection. The interferometer is composed of a polarizing beam splitter and two triple reflectors.

The interference signals are processed within the detectors in such a manner that two signals are obtained which are shifted by 90° to each other. To obtain these quadrature signals for the modulated laser, it has turned out advantageous to use Fresnel rhombi instead of quarter wave plates since for the latter the retardation is more strongly dependent on the wavelength. The quadrature signals for the ADI, the reference interferometer, and the counting interferometer operated by laser 1 are recorded with an acquisition rate of 250000 samples s^{-1} by a 16-bit, 8 channel A/D converter card. First, a data block is sampled with a length depending on the averaging time. The raw sine-cosine values are processed by a Heydemann correction [9]. Since the frequency modulated laser has a slight amplitude modulation the sine-cosine signals do not follow an ellipse but have a slightly spiral shape. The amplitude modulation is ≈ 5% of the total intensity so that for a few 2π periods an ellipse is a good approximation. Therefore several ellipse fits over a few periods instead of one over all data points are calculated.

Subsequently, single phase values are calculated by the arctan function, and are sorted according to the rising and falling edges of the modulation. The phase of the ADI is corrected by the phase of the stabilised laser interferometer according to (4). For each edge, the phases of the ADI and the reference interferometer are linearly interpolated by a least square fit with the axis intercept set to zero. With the slope, the length L_{ADI} is determined according to (3).

For the fixed synthetic wavelength approach, laser 2 is stabilised on a K-D2 absorption line at approx. 776.1 nm giving a synthetic wavelength of ≈ 173 µm. The fractional part of the length can be directly evaluated from (6) after the Heydemann correction. In the case of stabilised lasers, it is sufficient to determine the parameters for the Heydemann correction during a slow movement of the measurement reflector prior to the measurement.

Length measurements with the setup shown in Fig. 1 were performed at the geodetic base of the PTB. It consists of a 50 m long bench with a moving carriage and is equipped with a conventional counting HeNe laser interferometer with a folded beam path, which was used as reference. The refractive index of air is determined by the Edlén equation [8]. The environmental parameters were measured by one air pressure and one relative humidity sensor, and temperature by PT100 sensors placed at an interval of 2.5 m along the bench. Results of measurements for lengths of ~ 2 m were published in [7].

The combined evaluation of ADI and fixed synthetic wavelength measurements of distances up to approximately 10 m leads to results shown in Fig. 2. According to the data, the deviation between the ADI and the HeNe interferometer is below 0,5 µm + 0.5 µm/m except of three points. Thus, the well known approach of an absolute interferometer with one laser with modulated laser frequency and a second, frequency stabilised laser for compensation of vibrations can be improved by frequency stabilising both lasers and using the synthetic wavelength for further interpolation of the result.

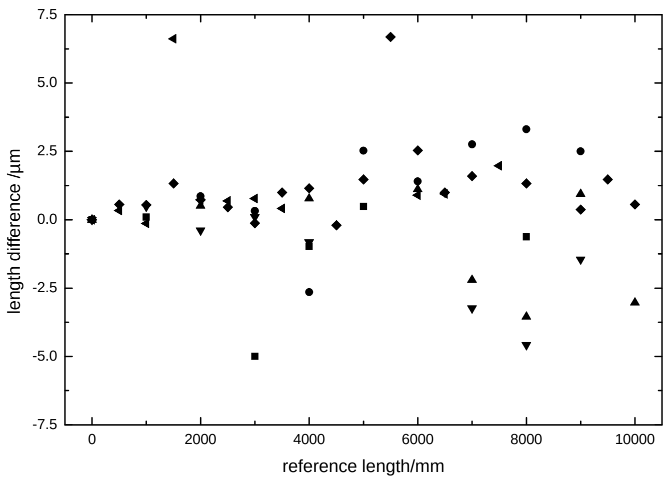


Fig. 2. Difference between the absolute interferometer and a counting HeNe laser interferometer as a function of the reference length after evaluation according to equation (6). The different symbols denote individual measurements.

3. Precise length measurements of surface profiles

High precision measurements of surface topography and properties like roughness are important tasks in quality assurance and control. There are a number of different technologies which differ by the measurement principle, operating complexity, resolution, and measuring range. Interferometric methods have a very high resolution within the nanometre range. Laser interferometers, however, are giving ambiguous results for distance variation of more than half the wavelength λ, e.g. due to steps in the surface. Nevertheless, using the multi-wavelength technique the interferometric measurement of surface profiles is possible.

For the measurement of surface profiles the measuring range of the interferometer was supposed to be in the order of 100 µm. Using parameter stabilised laser diodes three wavelengths were necessary. Wavelengths of 789.5 nm, 822.95 nm, and 825.3 nm were chosen which lead to synthetic wavelengths of approx. 14.0 µm, 14.8 µm, and 289 µm.

A modulation technique allows the separation of the wavelengths at the interferometer output. The injection current of the three laser diodes is modulated with different frequencies ω_m around 1 MHz. This results in a modulation of the laser frequency by $\Delta\nu$ and a modulation of the interferometer phase. At the interferometer output a spectrum appears with harmonics of ω_m. The amplitude of two adjacent harmonics are proportional to the sine and cosine of the interferometer phase. Lock-in amplifiers detecting at the second and third harmonic give a quadrature signal [10].

The different modulation frequencies of the laser diodes allow the simultaneous detection of the three interferometer signals with only one photo detector. The optical setup, sketched in Fig. 6, is also simplified by the modulation technique.

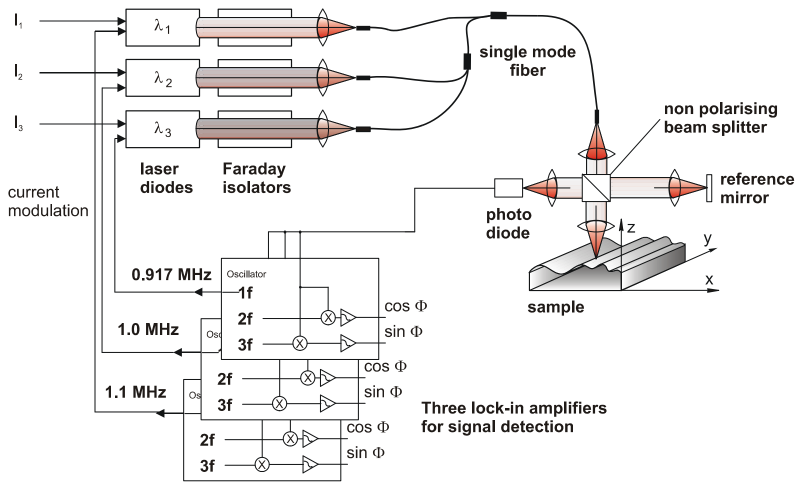


Fig. 6 Optical setup of the surface profilometer with a multiple wavelength diode laser interferometer

The beams of the diode laser are coupled into one single mode fibre. Faraday isolators protect the diodes from feedback light. The fibre is connected to a simple unbalanced Michelson interferometer without polarization optical components. One arm of the interferometer leads to the reference mirror, the light of the other arm is focused onto the sample. Depending on the measurement range in z-direction, i.e. focus depth, and on the lateral resolution, the focal length of the objective lens can be chosen suitable according to the surfaces under investigation. The sample can be moved in both lateral directions with mechanical translation stages by approximately 15 cm.

The developed diode laser profilometer was tested using a PTB depth setting standard chosen as a test surface with a well-known surface topography [11]. This standard is a glass block, 8 mm thick, 20 mm wide, and 40 mm long with six grooves of circular profile, the depth ranging from 190 nm to 9.2 µm. The reference values of this standard have been reproduced within the uncertainties stated for the depth setting standard as shown in table 1.

Table 1. Comparison of the measured groove depths with the reference values of a depth setting standard.

Diode laser profilometer / nm	PTB reference value / nm
196	191±6
445	440±6
945	952±10
2130	2141±32
4520	4543±48
9225	9208±80

The measurement uncertainty of the profilometer is limited by the quality of the translation stages (~ 25 nm). The uncertainty of the interferometer itself is estimated to approximately 20 nm for a 100 µm measuring range and an arm length difference of 10 cm [10]. Moreover, uncertainty contributions arise from the long and short term wavelength stability of diode lasers used (1 nm and 10 nm, respectively). Contributions due to changes of the refractive index of air are negligible. A further uncertainty contribution can arise from the phase measuring, e.g. due to the lock-in amplifiers, or optical quality of the surface. This contribution amounts to approx. 4 nm for smooth, highly reflective surfaces.

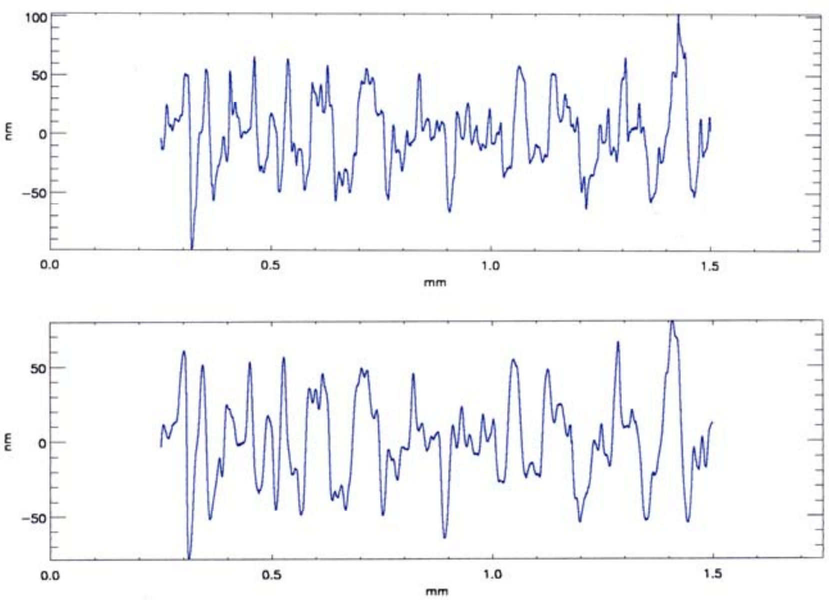


Fig. 4. Surface profile measured with a contact stylus instrument (upper curve) and with the interferometer (lower curve)

On rough surfaces the interference contrast varies over the sample which affects the interferometer signal. Three so-called “superfine roughness standards” with R_z values between 134 nm and 450 nm were measured to investigate the limits of the interferometer on rough surfaces. These standards are cylindrical disks with approx. 50 mm diameter with a 4 mm wide rough zone around the centre. Figure 4 shows a part of the surface profile of a roughness standard with an average maximum height R_z = 134 nm as measured with the diode laser interferometer in comparison with the contact stylus instrument. It turns out that the profile is wll reproduced by the interferometer. However, fine details like the sharp spike near position

1.4 mm and the narrow groove near 0.3 mm are smoothed by the interferometer. This is caused by the focus diameter of approx. 2 µm of the microscope objective used for these measurements so that the measured height values are averaged over this range. This smoothing leads to an apparent decrease of the roughness. All roughness parameters derived from the interferometer data are smaller than their reference values from the calibration with a contact stylus instrument. The roughness average R_a is well reproduced by the interferometer data. The maximum deviation to the reference value is 9%. The other averaged roughness parameters like rms roughness R_q which are not shown here have similar deviations. The average maximum height R_z and the maximum roughness depth R_{max} as peak values are up to 17% smaller than their reference values.

As shown above the use of the three-wavelength diode laser interferometer offers a measurement range of ≈145 µm and allows to measure objects not accessible to one-wavelength interferometers, i.e. surfaces with steps larger than half the optical wavelengths. Also, a temporary loss of the signals, e.g. due to poor reflecting parts of the surface, only leads to missing data for these areas and does not interrupt the scan as it would be the case for a fringe counting method.

Although the interferometer itself has a resolution of about 4 nm and its expected uncertainty is of the order of 20 nm, the resulting overall accuracy is unfortunately limited by the mechanical translation stages used in the experiments to values up to 250 nm for scan lines longer than 4 mm.

The interferometer reaches its limits on rough surfaces. Such surfaces could be measured, although the interferometer signal amplitude fluctuated and dropped partially to 10% of the value on smooth areas. For a calculation of the roughness parameters fringe counting was applied to get equally spaced data without missing points. The roughness parameters derived from these values are in good agreement with the calibration values of the roughness standards.

4. Precise length measurements of prismatic bodies

For many industrial applications precise measurement of prismatic bodies, e.g. gauge blocks, or cylindrical samples, by multi-wavelength interferometry is necessary for dissemination of the length unit metre, i.e. traceability of the length measurement. Precise length changes of such bodies with temperature (thermal expansion), pressure (compressibility), and time (long term stability) can be measured with PTB's Precision Interferometer. Figure 5 shows the interferometer situated in a temperature controlled and vacuum tight environmental chamber. The light provided by the three different lasers alternatively passes a fibre representing the entrance of the interferometer. The reference path of the interferometer can be varied for phase stepping by slightly tilting the compensation plate. The tilt angle is monitored by an auxiliary interferometer and servo controlled. For measurements in air the measuring path contains a 400 mm vacuum cell close to the sample to determine the refractive index of air at the specific environmental conditions.

A 512 x 512 pixel camera system (Photometrics CH 350) provides data frames at 16-bit per pixel. Phase stepping interferometry based on intensity frames at 8 different phase steps is used to obtain the phase map of the sample including the end plate [12]. The centre position of the samples front faces with respect to the camera pixel coordinates is assigned [13]. Interferometer autocollimation was adjusted by retroreflection scanning before a measurement was started [14]. Three stabilised lasers are used subsequently in the measurements. The gauge block lengths resulting from using the two J_2-stabilised lasers at 532 nm and 633 nm were averaged. The Rb-stabilised laser at 780 nm is used for a coincidence check, only.

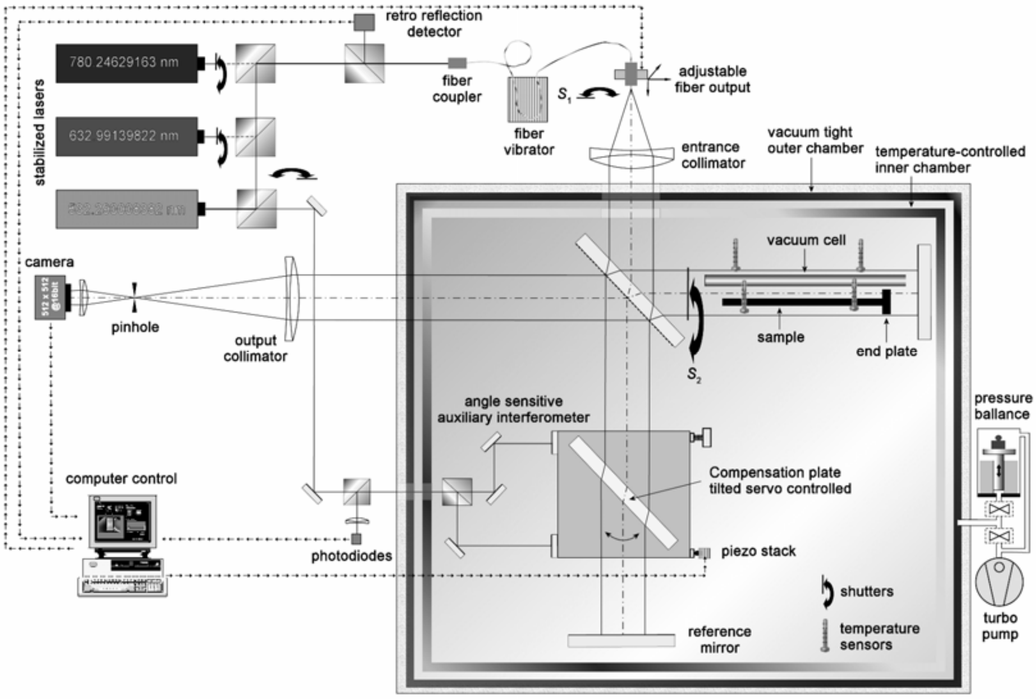


Fig. 5. Scheme of PTB's Precision Interferometer

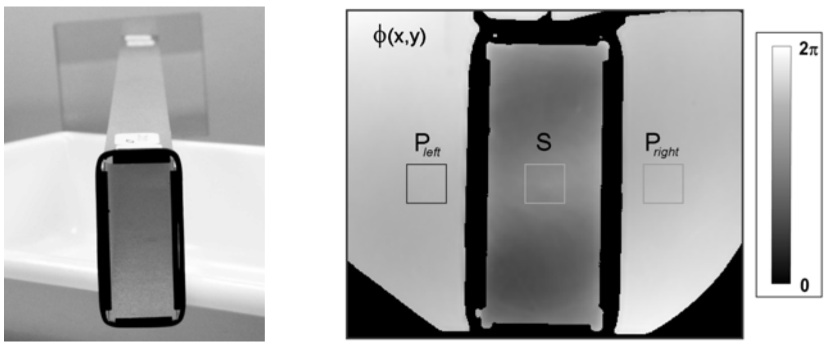


Fig. 6. Left: front view of the sample wrung to the end plate, right: measured interferogram with a region of interest (ROI) at the sample's front face (S) and two symmetrically arranged ROIs at the end plate ($P_{left/right}$)

The fractional order of interference, f, can be extracted from the interferogram. Figure 6, left, shows a photograph of a typical sample. At the right side the measured interferogram of this sample is shown. The rectangles indicate the regions of interest (ROIs) in which the phase values are averaged. This leads to the values ϕ_S, ϕ_P^{left}, and ϕ_P^{right} from which the fractional order of interference is calculated:

$$f = \frac{1}{2\pi}\left[\frac{1}{2}\left(\phi_P^{left} + \phi_P^{right}\right) - \phi_S\right]. \tag{7}$$

Using optical interferometry, the length L of a sample body is expressed as a multiple of the light wavelength used. For one-wavelength interferometry the exact knowledge of the integer interference order (entailed with a precision mechanical pre-measurement) would be required before the exact length can be determined. The use of different wavelengths, however, results in independent length which should coincide. The expected coincidence is predicted by the uncertainties for the individual λ_k and f_k, respectively. A variation technique can be used for the extraction of the integer orders of interference known as "method of exact fractions". Here integer variation numbers δ_k are introduced in the length evaluation:

$$L_k = \frac{1}{2}\lambda_k\left(i_k + \delta_k + f_k\right), \tag{8}$$

in which i_k is evaluated from the estimated length $i_k = L^{est} \Big/ \frac{1}{2}\lambda_k$. At a set of $\{\delta_1, \delta_2, ...\}$ the mean length $\overline{L}$ and the average deviation Δ are calculated according to:

$$\overline{L} = \frac{1}{r}\sum_{k=1}^{r} L_k \ , \ \ \Delta = \frac{1}{r}\sum_{k=1}^{r}\left|\overline{L} - L_k\right|. \tag{9}$$

The amount of Δ can be used as a coincidence criteria and displayed vs. the related values of $\overline{L}$. Fig. 7 shows a measurements example for this approach, in which the estimate length of the sample is 299.15 mm. The measurements were performed under vacuum conditions using three wavelengths λ_k subsequently, each measurement resulting in the fractions as given in Table 2.

Table 2.

λ_k / nm	$u(\lambda_k)/\lambda_k$	f_k (in example)	$u(f_k)$	$u(L_k)$ /nm (in example)
532.290008382	3e-12	0.0271	0.0003	0.08
632.99139822	2.5e-11	0.1723	0.0004	0.13
780.24629163	1e-9	0.6465	0.0006	0.38

Obviously, a large number of values exist below a level of Δ = 5 nm. The x-values of these points are separated by some tenth of micrometer. Certain minima of Δ exist whose separation exceeds very large values of more than 0.6 mm. It is tempting to identify the minimum closest to L^{est} with the length of the sample. However, before such conclusion, it is necessary to review the uncertainty of the length caused by the uncertainties of the individual wavelengths and fractional orders of interference. From (8) follows:

$$u(L_k)^2 = \left(L^{est}u(\lambda_k)/\lambda_k\right)^2 + \left(\frac{1}{2}\lambda_k u(f_k)\right)^2 \tag{10}$$

and, in case of the mentioned example with PTB's Precision Interferometer, the uncertainties $u(L_k)$ are close to 0.1 nm as depicted in the last column in Table 2.

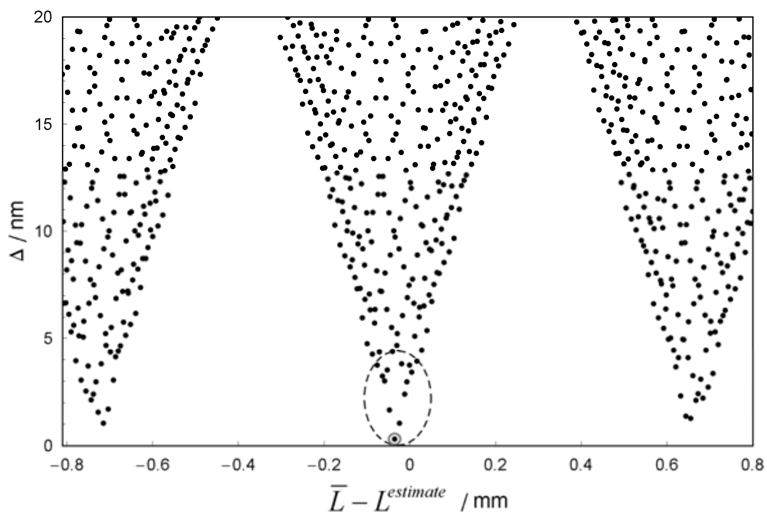


Fig. 7 The average length resulting from variations of the integer orders (x-coordinates of the data points) and the related average deviation.

The actual minimum, Δ_{min}, is about 0.3 nm (see the data point marked by the grey circle in Fig. 7). This value is somewhat larger than the total uncertainty of the mean length, $u(\overline{L})$, resulting to 0.14 nm. The reason for this somewhat larger value of Δ_{min} compared to 0.14 nm can be attributed to small constant influences onto the coincidence between the different lengths L_k as discussed in [15]. The low value of $u(\overline{L})$ together with the fact that the neighbouring points are off by more than 1 nm (see data points within the dashed circle of Fig. 7) justifies the conclusion that the mean length for which Δ_{min} is found can be looked upon the actual length. Thus, the range of unambiguity is in fact about 0.6 mm as assumed in the above example. Therefore, a rough estimate of the length, obtained from a simple measurement, is sufficient for measurements at PTB's Precision Interferometer.

5. Conclusions

The multi-wavelength measurement technique is a proper method to remove ambiguity in interferometric length measurements. In general, the more accurate the fractions of the fringes can be measured, the larger is the range of unambiguity.

Based on this principle, an absolute distance interferometer has been developed for the measurement of long distances of up to approx. 10 m by combination of frequency-sweeping interferometry and two-wavelength interferometry, using two external cavity diode lasers. A comparison of the absolute distance interferometer to a counting HeNe laser interferometer shows a deviation below 0.5 µm + 0.5 µm/m.

Furthermore, a three-wavelength diode laser interferometer has been developed for the measurement of surface profiles. An uncertainty of approx. 20 nm was achieved for a measuring range below 150 µm.

Finally, the example of PTB`s Precision Interferometer demonstrates the benefit of three-wavelength interferometry. The length of bodies with parallel end faces (e.g. gauge blocks) can be measured there with an uncertainty of the averaged length below 0.3 nm, while the range of unambiguity is drastically enlarged. Therefore, an ultra precise length measurement at PTB's Precision Interferometer requires only a rough estimate of the length in advance.

Acknowledgements
The authors are grateful to the Deutsche Forschungsgemeinschaft (DFG) for financial support.

References

[1] Abou-Zeid A. Diode Lasers for Interferometry, *J. of Prec. Eng.,* 11(3), 139, 1988.

[2] Bodermann B, Burgarth V, and Abou-Zeid A. Modulation-free stabilised diode laser for interferometry using Doppler-reduced Rb transitions. In Proc. 2nd Euspen Topical Conf. (Turin), vol 1, ed A Balsamo et al, 2001, 294–297

[3] Zarka A, Abou-Zeid A, et al. Intern. Comparison of Eight Semiconductor Lasers Stabilised on 127I2 at 633 nm, *Metrologia*, 37 (4), 329, 2000.

[4] Bourdet GL and Orszag AG. Absolute distance measurements by CO2 laser multiwavelength interferometry, *Appl. Opt.* 18, 225–227, 1979.

[5] Bechstein KH and Fuchs W. Absolute interferometric distance measurements applying a variable synthetic wavelength, *J. Opt.* 29, 179–182, 1998.

[6] Kinder T and Salewski KD. Absolute distance interferometer with grating stabilised tunable diode laser at 633 nm, *J. Opt. A: Pure Appl. Opt.* 4, 364–368, 2002.

[7] Hartmann L, Meiners-Hagen K, and Abou-Zeid A. An absolute distance interferometer with two external cavity diode lasers, *Meas. Sci. Technol.* 19, 045307 (6pp), 2008.

[8] Bönsch G and Potulski E. Measurement of the refractive index of air and comparison with modified Edlens formulae, *Metrologia* 35, 133–139, 1998.

[9] Heydemann PLM. Determination and correction of quadrature fringe measurement errors in interferometers, *Appl. Opt.* 20, 3382–3384, 1981.

[10] Meiners-Hagen K, Burgarth V, and Abou-Zeid, A. Profilometry with a multi-wavelength diode laser interferometer, *Meas. Sci. Technol.* 15, 741–746, 2004.

[11] Abou-Zeid A, Wolf M. Profilometry: using a diode laser interferometer with three wavelengths. In Proc. of the 1st euspen Topical Conf. on Fabrication and Metrology in Nanotechnology, Copenhagen, Denmark, L. De Chiffre and K. Carneiro (Eds.). Vol. 1, p. 137–140, 2004.

[12] Schödel R, Nicolaus A, Bönsch G. Phase stepping interferometry: Methods to reduce errors caused by camera nonlinearities, *Appl. Opt*. 41, 55-63, (2002).

[13] Schödel R, Decker JE. Methods to recognize the sample position for most precise interferometric length measurements, In Proc. SPIE, Vol. 5532, 237-247, 2004.

[14] Schödel R, Bönsch G. Highest accuracy interferometer alignment by retroreflection scanning, *Appl. Opt.* 43,. 5738-5743, 2004.

[15] Schödel R, Nicolaus A, and Bönsch G. Minimizing interferometer misalignment errors for measurement of sub-nanometer length changes. In Proc. SPIE 2003, Vol. 5190, p. 34-42, 2003.

The Magnetoplasmic Measurements of the Carrier Density and Mobility in Semiconductors

[1]Z. Jankauskas, [1]V. Kvedaras, [1,2]R. Gaivenis

[1]Department of Electrical Engineering, Vilnius Gediminas Technical University,
Sauletekio av. 11, LT 10223, Vilnius, Lithuania
[2]Computer Software Department, Vilnius College in Higher Education,
J.Jasinskio Str.15, LT 2600, Vilnius, Lithuania
Email: zigmantas.jankauskas@el.vgtu.lt

Abstract. *Many semiconductor materials manufactured by the help of nanotechnology have the charge carriers of different type and mobility. Already existing carrier density and mobility measurement methods are not accurate enough for the case of several carrier components. The use of the magnetoplasmic waves provides a simple and most precise way to determine the density and mobility of each type of the carriers (electrons and/or holes).*
Magnetoplasmic Waves may be excited in semiconductors when the strong magnetic field H is applied. The semiconductor sample becomes partially transparent under these conditions. In the case of magnetoplasmic resonance within each of carrier groups the transparence coefficient has maximum. For fixed values H and excitation frequency ω the density and mobility of every carrier type can be found.
Dispersion relation for two types of charge carriers is obtained and resonance curves are calculated.

Keywords: Magnetic measurement, imaging, magnetic susceptibility, helicons, calculation.

1. Introduction

Many semiconductor materials manufactured by the help of nanotechnology have the charge of different type and mobility. Already existing carrier density and mobility measurement methods are not accurate enough in the case of several carrier components. The use of the magnetoplasmic waves (helicons) provides a simple and more precise way to determinate the density and mobility of each type of the carriers (electrons and/or holes).

Magnetoplasmic waves may be excited in semiconductors when the strong magnetic field H is applied and large Hall currents may exist. In the case of magnetoplasmic resonance within each of carrier groups the transparency coefficient has maximum. For fixed values H and excitation frequency ω the density and mobility of every carrier type can be found.

2. Subject and Methods

Let us consider a semiconducting plate with a electric coil placed on its surface. We shall for the sake of simplicity assume the semiconductor to be infinitely large. Let the semiconductor surface be parallel to the xy plane and the z-axis be directed perpendicular to it. Then the electrical current has only one component along the y-axis, and the magnetic has one along the x-axis. We assume, that the field H_x varies with time according to a harmonic law

$$\mathrm{H}_x = H\cos\omega t, |z| = a, \qquad (1)$$

Where a is the thickness of the plate along z-axis [1]. The currents induced in the semiconductor sample are directed in such a way so as to counteract the penetration of the field. As a result the varying magnetic field within the semiconductor will be other than zero only to a certain deph (skin depth). If the semiconductor plate is at the same time placed into the strong magnetic field $H_x = H_0$, the Hall currents j_x appear and helicon magnetoplasmic waves may be excited. Assuming that the conductivity of the plate is provided by electrons with an isotropic mass we have the following equations of motion for the current components j_x and j_y

$$\frac{d}{dt} j_x + \tau^{-1} j_x - \frac{eH_0}{mc} j_y = \frac{Ne^2}{m} E_x, \qquad \frac{d}{dt} j_y + \tau^{-1} j_y + \frac{eH_0}{mc} j_x = \frac{Ne^2}{m} E_y, \tag{2}$$

where N, e, m and τ are the density, charge, mass and collision time of the electrons, E_x and E_y, are the varying electrical field, and H_0 is the static magnetic field along the z-axis.

The equations of motion (2) must be solved along with the Maxwell equations

$$rot\vec{\varepsilon} = -\frac{1}{c}\frac{\partial \vec{\mathrm{H}}}{\partial t} \quad rot\vec{\mathrm{H}} = -\frac{1}{c}\frac{\partial \vec{\varepsilon}}{\partial t} + \frac{4\pi}{c}\vec{j} \tag{3}$$

The solutions of the system of equations (2), (3) were sougth in the form [2]

$$\vec{\varepsilon} = \mathrm{Re}\,\vec{E}\exp(-i\omega t), \ \vec{\mathrm{H}} = \mathrm{Re}\,\vec{H}\exp(-i\omega t), \qquad \vec{j} = \mathrm{Re}\,\vec{I}\exp(-i\omega t), \tag{4}$$

where the complex amplitudes $\vec{E}, \vec{H}$ and $\vec{I}$ depend on only coordinate z and have the components x and y. Taking into account that $\omega << \tau^{-1} = 10^{11} - 10^{13}$ Hz and ignoring the displacement currents in comparison with conductive ones we obtain a system of equations for the complex amplitudes of the varying magnetic field H_x and H_y

$$c^2\frac{\partial^2}{\partial z^2}H_x + i\omega\delta_{11}H_x + i\omega\delta_{12}H_y = 0 \quad c^2\frac{\partial^2}{\partial z^2}H_y + i\omega\delta_{22}H_y - i\omega\delta_{12}H_x = 0, \tag{5}$$

where the components of the conductivity tensor σ are

$$\sigma_{11} = \sigma_{22} = \sigma_0\frac{1}{1+\left(\frac{eH_0\tau}{mc}\right)^2}, \quad \sigma_{12} = \delta_0\frac{\frac{eH_0\tau}{mc}}{1+\left(\frac{eH_0\tau}{mc}\right)^2}, \quad \sigma_0 = \frac{Ne^2\tau}{m}. \tag{6}$$

The boundary conditions have the form

$$H_x = H, H_y = 0 \text{ if } z = \pm a, \qquad \frac{\partial H_x}{\partial z} = \frac{\partial H_y}{\partial z} = 0 \text{ if } z = 0, \tag{7}$$

and the solutions of (5) satisfying (7) are

$$H_x = \frac{1}{2}H\left(\frac{\cos k_- z}{\cos k_- a} + \frac{\cos k_+ z}{\cos k_+ a}\right), \qquad H_y = \frac{1}{2}iH\left(\frac{\cos k_+ z}{\cos k_+ a} + \frac{\cos k_- z}{\cos k_- a}\right), \tag{8}$$

where $k\pm$ can be detected from the characteristic equation

$$c^2 k_\pm^{\ 2} = \frac{\omega_p^{\ 2}\omega}{\pm\omega_H + i\tau^{-1}}, \omega_p^{\ 2} = \frac{4\pi Ne^2}{m}, \omega_H = \frac{eH_0}{mc}. \tag{9}$$

The items with the argument k_- are caused by the helicon magnetoplasmic waves.

3. Two types of charge carriers

If the semiconductor sample contains two types of electrons with different densities and mobilities the characteristic equation for magnetoplasmic wave vector $\underset{\rightarrow}{k}$ can be written as follows

$$k_{\pm}^{2} = \omega\mu_0 \, e N_1 u_1 \left(\frac{-i \pm u_1 B}{1+(u_1 B)^2} \right) + \omega\mu_0 \, e N_2 u_2 \left(\frac{-i \pm u_2 B}{1+(u_2 B)^2} \right), \qquad (10)$$

where N_1, N_2 and u_1, u_2 are the densities and mobilities of the different types of the charge carriers; μ_0 - magnetic permeability of vacuum. The resonant curves in dependence of magnetic field $\underset{\rightarrow}{B}$ were calculated from (17) for various values of N_1, N_2, u_1, u_2, and angular frequencies ω, and are shown on Fig.1, 2. The resonant curves of amplitude and phase for forward wave *K* and reflected wave *P* from InSb plate in dependence of magnetic field *B* are calculated using program [3] for frequency f=300 MHz. Parameters of semiconductor plate: thickness a=5 mm, dielectric constant ε=15, charge carriers density $N = 0{,}15 * 10^{23} m^{-3}$, charge carriers mobility $u = 5 m^2 V^{-1} s^{-1}$.

The first main peak for $B_1 = 29T$ corresponds to the resonance of electrons with higher mobility u_1 (density N_1). The second peak for $B_2 = 7T$ is responsible for the resonance of both carriers type with summary density ($N_1 + N_2$). The heights of both peaks are proportional to the products ($u_1 B$) and ($u_2 B$) respectively. The comparison of experimental and theoretical resonant curves provides the possibility to calculate the parameters N_1, N_2, u_1, and u_2 for both types of charge carriers.

There are observed, that different sign of mobilities u_1 and u_2 increase second and higher resonans maximums in resonant curves.

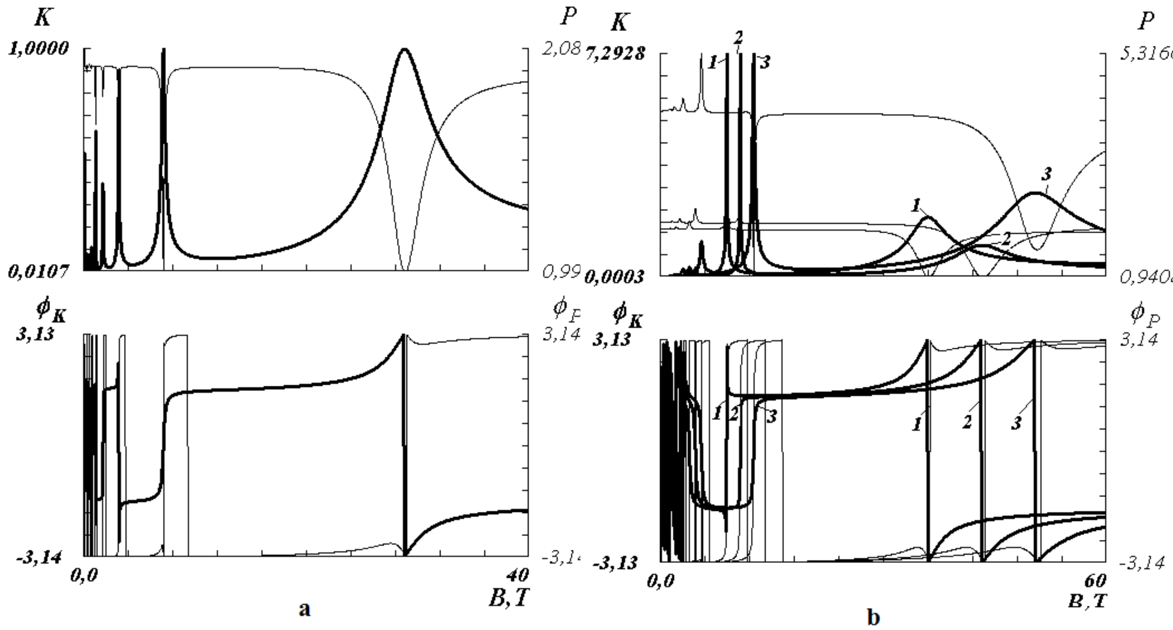


Fig. 1 The resonant curves in dependence of magnetic field $\underset{\rightarrow}{B}$ for:

a - $u_1 = -5 m^2 V^{-1} s^{-1}, u_2 = 5 m^2 V^{-1} s^{-1}$, $N_1 = N_2 = 0{,}15 * 10^{23} m^{-3}$;

b - $u_1 = -2.5 * m^2 V^{-1} s^{-1}, u_2 = 5 m^2 V^{-1} s^{-1}$, $N_1 = 0{,}3 * 10^{23} m^{-3}$, $N_2 = 0{,}15 * 10^{23} m^{-3}$:

1 – f=250 MHz, K_{max}=4,01, 2 - f=300 MHz, K_{max} =7,19, 3 - f=350 MHz, K_{max} =2,74

4. Experimental verification

The theoretical results were confirmed experimentally for the n-Ge semiconductor material. In Fig. 2 the experimental resonance curve is shown for the case f=25 MHz and specimen tickness d=3 mm. The relatively weak signal in detection coil for B=1,5 T pertains to the electrons with a higher mobility 0,65 $m^2V^{-1}s^{-1}$, and the stronger maximum for B=8 T – to the electrone with a smaller mobility 0,125 $m^2V^{-1}s^{-1}$.

The measurements were performed of the National High Magnetic Laboratory, which is supported by MSF Cooperative Agreement No. DMR-0084173, by the state of Florida, and by the DOE.

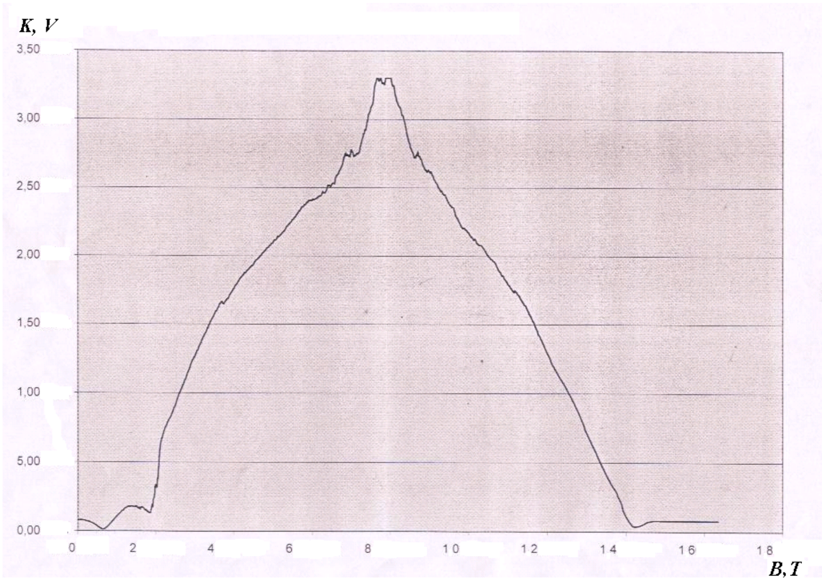


Fig. 2 The experimental resonant curve in dependence of magnetic field $\underset{\rightarrow}{B}$

5. Conclusions

The measurement of the charge densities and mobilities for the charge carriers of various types in semiconductor materials by the help of magnetoplasmic waves can be provided in contactless mode. Many semiconductors thus can be investigated if the high magnetic fields (~30 Tesla) are available. The measurement results are in compliance with the data obtained by the use of already existing methods.

References

[1] Jankauskas Z., Kvedaras V., Laurinavičius L. Experimental realization of Helicon Maser. *Physica B: Condensed Matter*, vol.346-347: 539-542, 2004.

[2] Irvin I.D. Fundamentals of Electrodynamics. Willey, New York , 2002.

[3] Gaivenis R., Jankauskas Z., Laurinavičius L. Electromagnetic Magnetoplasmic Wave in Multi-layer Resonator with Several Semiconductors. *Electronics and Electrical Engineering*, 53 (4): 16-21, 2004.

Magnetic Field Measurement of a Planar Coil Using Magnetic Resonance Imaging Methods

I. Frollo, P. Andris, J. Přibil, L. Vojtíšek, Z. Holúbeková, T. Dermek

Institute of Measurement Science, SAS, Bratislava, Slovakia,
frollo@savba.sk

Abstract. *A new method for mapping and imaging of the magnetic field of planar electromagnetic coil, placed into the homogenous magnetic field of an NMR imager, is proposed. First attempts of electromagnetic planar coil computation and testing on an NMR 0,2 T imager were accomplished. In our experiments a homogeneous phantom (reference medium) was used - a container filled with water - as a medium. The resultant image represents the magnetic field distribution in the homogeneous phantom. A standard gradient echo imaging method, susceptible to magnetic field homogeneity, was used for detection. An image acquired by this method is actually a projection of the sample properties onto the homogeneous phantom. The goal of the paper is to map and image the magnetic field deformation using NMR imaging methods.*

Keywords: magnetic field, magnetic resonance imaging, electrical phantom, meander coil

1. Introduction

Nuclear Magnetic Resonance (NMR) measurement and imaging of proton density of biological and physical structures is a routine investigating procedure. Another case is when an object that generates a weak magnetic field is inserted into a stationary homogeneous magnetic field resulting in deformation of the basic stationary magnetic field. If the space in the vicinity of this object is filled with a water containing substance, we are able to image this object. The acquired image represents a modulation of the basic magnetic field by the weak field of the coil.

First attempt of a direct measurement of the magnetic field created in living tissue by a simple wire fed by a current was reported in [1]. A method utilizing the divergence in gradient strength that occurs in the vicinity of a thin current-carrying copper wire was introduced in [2]. Using pulsed gradient spin-echo NMR sequence, in vitro micro images of a sample, a solution of polyethylene oxide in water, were presented. A simple experiment with thin, pulsed electrical current– carrying wire and imaging of a magnetic field using a plastic sphere filled with agarose gel as phantom, was published in [3]. First attempt of the indirect susceptibility mapping of thin-layer samples using phantoms was described in [4].

In this paper we propose an electromagnetic planar coil supplied by small and stabilised DC current. Computation of the magnetic field of the planar coil serves for comparison of theoretical results with experimental.

2. Subject and Methods

We suppose that the conductor position is on the x-axis in the rectangular complex coordinate system (**i**x, **j**y, **k**z) and the diameter of the conductor is neglected.
According to Fig. 1, we suppose one conductor limited by two points (P, Q), lengths of 2L, with the left - right symmetry. Conductor is fed by currents +I. The basic magnetic field B_0 of the NMR imager is parallel with the z-axis. The task is to calculate the $B_z(x,y,z)$ component of the magnetic field near the conductor.

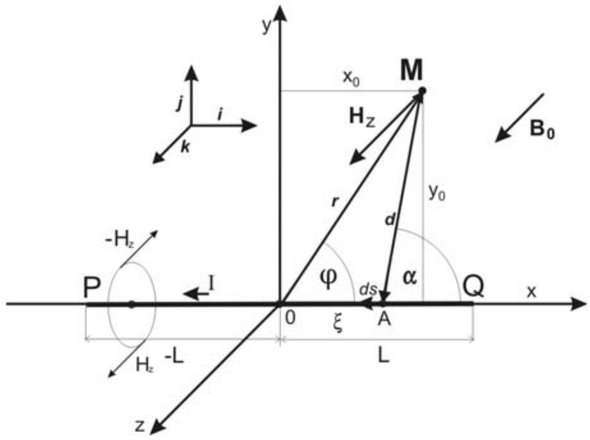


Fig. 1. Basic configuration of the limited length conductor $\overline{PQ}$ placed into the x – axis.

The formula for magnetic field in the point $M(x_0,y_0)$ is supposing that $z_0 = 0$. For computation of the magnetic field $B_z = \mu_0 H_z$ the Biot-Savart law in vector form was used [5].

$$\mathbf{H} = \frac{I}{4\pi} \oint \frac{d\mathbf{s} \times \mathbf{d}}{d^3} \tag{1}$$

From the physical interpretation the magnetic field is line-symmetry in compliance with x -axis, and symmetric according to the yz – plane. The calculation is limited to the first quadrant of the xy – plane. Following the Fig. 1 and formula (1) we can write:

$$d\mathbf{s} \times \mathbf{d} = ds\, d\, \sin(\alpha) = d\xi\, d\, \sin(\alpha) \tag{2}$$

Using formula (1) we get the basic expression as follows

$$\mathbf{H} = \mathbf{k} \frac{I}{4\pi} \int_{-L}^{L} \frac{\sin\alpha}{d^2}\, d\xi \tag{3}$$

Considering Fig.1 : $\sin\alpha = \frac{r}{d}\sin\varphi$, $\quad d^2 = r^2 + \xi^2 - 2r\xi\cos\varphi$, $\quad r\sin\varphi = y$, $\quad r\cos\varphi = x$

In the complex coordinate system (**i**x, **j**y, **k**z) we get the final formula for magnetic field **H** in the integral form

$$\mathbf{H} = \mathbf{k} \frac{Iy}{4\pi} \int_{-L}^{L} \frac{d\xi}{\sqrt{\left(x^2 + y^2 + \xi^2 - 2\xi x\right)^3}} \tag{4}$$

After integration of (4) we get the final general formula for magnetic field calculation

$$H_z = \frac{I}{4\pi y}\left[\frac{L - x}{\sqrt{\left(x^2 + y^2 + L^2 - 2Lx\right)^3}} + \frac{L + x}{\sqrt{\left(x^2 + y^2 + L^2 + 2Lx\right)^3}}\right] \tag{5}$$

As a mathematical and physical model a flat meander rectangular coil was designed and magnetic field using formula (5) for every wire was calculated. The resultant plots of magnetic field relative values are depicted in Fig. 2.

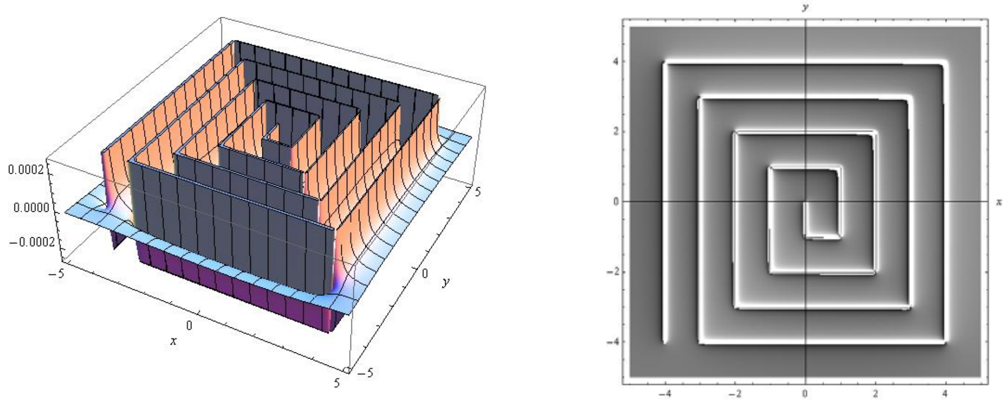


Fig. 2. Calculated magnetic field H_z (x,y) of the planar meander rectangular coil, limited plot-range. Left: 3D-plot of relative values. Right: Density-plot of relative values.

3. Experimental Results

As a physical model, a planar meander rectangular coil was realized on a thin plate. The coil was placed between two rectangular holders – homogeneous phantom filled with 0,1 wt% solution of $CuSO_4$ in distilled water, see Fig.3.

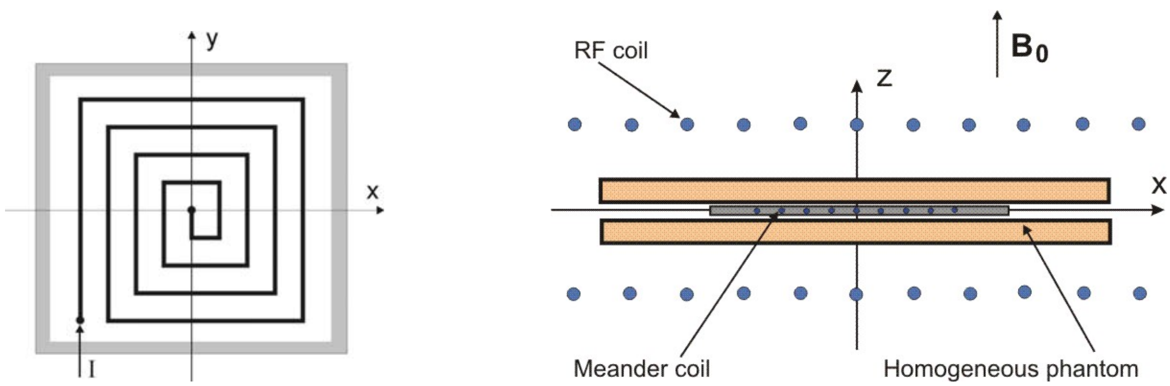


Fig. 3. Left: Meander planar coil design, dimensions 60 x 60 mm, realized on a printed board. Right: RF coil, active measuring volume (homogeneous phantom) and meander planar coil.

An NMR imager (ESAOTE Opera), permanent magnet 0,2 T with vertical orientation of the basic magnetic field B_0 was used. Feeding current I = 30 mA was applied to the meander coil creating a planar source of a weak magnetic field in the shape of a meander. The "gradient-echo" NMR sequence (Fig. 4) was selected for the measurement [6]. A special feature of the sequence is its sensitivity to basic magnetic field inhomogeneities.

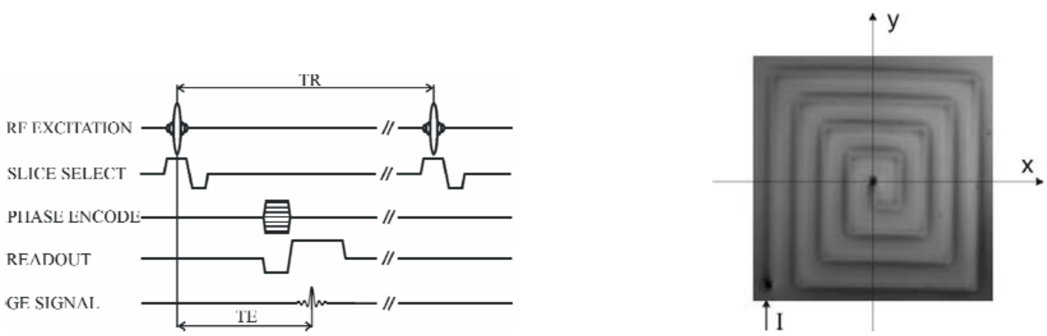


Fig. 4. Left: Time diagram of the NMR sequence for the gradient-echo signal detection. TE - echo time of 18 ms, TR – repetition time of 300 ms. Right: NMR image of the magnetic field distribution of a meander planar coil (60 x 60 mm), number of measured voxels 150 x 150.

4. Results and discussion

The single meander coil served for verification of this methodology, for the adjustment of basic parameters of the imaging sequence: time intervals TE and TR, number of averages, and for testing of a reference environment – $CuSO_4$ doped water in connection with relaxation times of the measuring sequence. It is evident that imaging of the magnetic field of the electromagnetic coil can be performed only if the vector of the static magnetic field $B_0 = B_z$ is perpendicular to the phantom plain. Other orientations caused significant blurring of the image edges in the x- and y-axis directions due to strong basic magnetic field B_0.

5. Conclusions

A new method for mapping and imaging of the magnetic field of planar electromagnetic coil placed into the homogenous magnetic field of an NMR imager is proposed. The method is based on a projection of magnetic field of the electromagnetic coil into a homogeneous planar phantom and subsequent NMR imaging using a gradient-echo sequence. The method was tested using a single planar coil – meander fed by electric currents and generating weak magnetic field.

Electromagnetic coil can serve as a special phantom and additional tool for measurement and imaging quality testing using NMR imaging methods. The advantages of such an electromagnetic phantoms are: universality, stability, repeatability, simple modification of basic parameters and precision. They are suitable for S/N ratio testing and very useful for imaging pulse sequences adjusting and examination. The first results showed the feasibility of the method and some of possibilities offered in this field of research.

Acknowledgements

The financial support by the Grant Agency of the Slovak Academy of Sciences, project no. VEGA 2/0142/08 and by The State program ŠPVV no. 2003SP200280203 is gratefully acknowledged.

References

[1] Joy M., Scott G., Henkelman M. In vivo imaging of applied electric currents by magnetic resonance imaging, *Magnetic Resonance Imaging*, Vol.7, 1989, 89-94.
[2] Callaghan P.T., Stepisnik J. Spatially-distributed pulsed gradient spin echo NMR using single-wire proximity, *Physical Review Letters*, Vol. 75, No. 24, 1995, 4532-4535.
[3] Sekino M., Matsumoto T., Yamaguchi K., Iriguchi N., and Ueno S. A method for NMR imaging of a magnetic field generated by electric current, *IEEE Trans. on Magnetics*, Vol. 40, 2004, 2188-2190.
[4] Frollo I., - Andris P., - Přibil J., - Juráš V. Indirect Susceptibility Mapping of Thin-Layer Samples Using Nuclear Magnetic Resonance Imaging, *IEEE Trans. on Magnetics*, Vol. 43, no. 8, (2007), 3363-3367.
[5] Kraus J. D., Fleisch D. A.: *Electromagnetics with applications*, 5th Edition, McGraw-Hill, 1998.
[6] Liang Zhi-Pei, Lauterbur Paul C., *Principles of Magnetic Resonance Imaging: A signal processing perspective*, 1999, Wiley-IEEE Press.

Magnetic Field of Spiral-shaped Coil

[1]P. Andris, [2]J. Weis, [1]I. Frollo

[1]Institute of Measurement Science, Slovak Academy of Sciences,
Dubravska cesta 9, Sk-841 04 Bratislava, Slovak Republic
Email: peter.andris@savba.sk
[2]Department of Radiology and Department of Medical Informatics and Bioengineering, Uppsala University Hospital, SE-751 85 Uppsala, Sweden
Email: jan.weis@radiol.uu.se

Abstract. *Spiral-shaped coil is one of several basic RF coils used in magnetic resonance instruments. It is appropriate mainly for surface measurements. Magnetic field of spiral-shaped coil was calculated using the Biot-Savart law and numerically computed. A sample coil was made and its magnetic field was measured using a magnetic resonance method. The verification showed very good qualitative and quantitative agreements between the calculated and measured values of the magnetic field.*

Keywords: magnetic field, magnetic resonance, spiral-shaped coil, NMR

1. Introduction

Although Nuclear Magnetic Resonance (NMR) tomographs are manufactured and disposed by many commercial firms a researcher can come across a moment when the supplied equipment is insufficient. Such part is for instance an RF coil. It must excite protons in a sample and/or convert RF magnetic field from the sample into electrical signal. Several parameters of the coil must be considered during a design. One of them and very significant is magnetic field of the coil and its homogeneity. Homogeneity or sensitivity of the coil can be increased by optimisation. Knowledge of the magnetic field is inevitable in the both cases. Magnetic field can be calculated in more manners. Calculation using the Biot-Savart law is rather frequent because it can be applied on a coil with an arbitrary shape. The spiral-shaped coil is very frequently used RF coil for construction of surface RF coils or arrays of NMR coils. It can be used for systems working at low as well as at high magnetic field. The purpose of the article is to present equations for computing the magnetic field of spiral-shaped coil based on the Biot-Savart law. Correctness of the equations was verified by NMR method described in [1].

2. Subject and Methods

Magnetic field of a three-turn spiral-shaped coil was calculated using the Biot-Savart law (see a two-turn spiral-shaped coil in Fig. 1). The calculation was made in a vector form, that is why the calculation is very mistake-proof. Calculations in vector form were performed very successfully using the program package Mathematica (Wolfram Research Inc., Champaign, IL). The computed magnetic field was compared with the magnetic field measured on a sample of the spiral-shaped coil using the NMR method [1]. Measurement was performed on 1.5 T NMR system (Gyroscan NT, Philips, Best, the Netherlands).

3. Results

Look at a two-turn spiral-shaped coil in Fig. 1. The application of the Biot-Savart law yielded the following equations:

$\mathbf{B} = \frac{\mu_o I}{4\pi} \oint \frac{d\mathbf{s} \times (\mathbf{P} - \mathbf{s})}{|\mathbf{P} - \mathbf{s}|^3}$ is magnetic field calculated using the Biot-Savart law [2], where

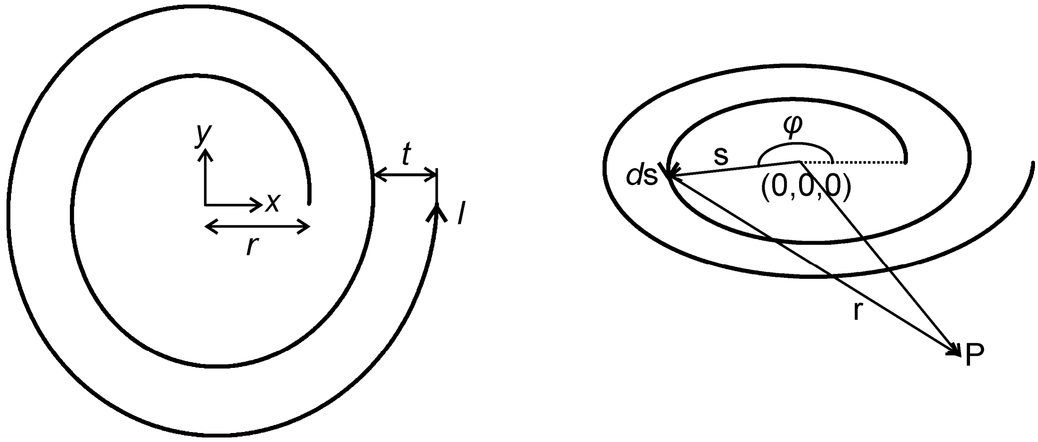


Fig. 1 Two-turn spiral-shaped coil.

I is current through the coil,

P is vector pointing to the observer and

s is vector pointing to centerline element of the coil conductor d**s**.

$\mathbf{P} = \mathbf{i}p_x + \mathbf{j}p_y + \mathbf{k}p_z$ is vector pointing to the observer.

$\mathbf{i}, \mathbf{j}, \mathbf{k}$ are unity vectors in directions of the x,y,z coordinates.

$\mathbf{s} = \mathbf{i}\left(r + \frac{t}{2\pi}\varphi\right)\cos\varphi + \mathbf{j}\left(r + \frac{t}{2\pi}\varphi\right)\sin\varphi$ is vector pointing to the coil.

$$d\mathbf{s} = \left(\mathbf{i}\left(\frac{t}{2\pi}\cos\varphi - \left(r + \frac{t}{2\pi}\varphi\right)\sin\varphi\right) + \mathbf{j}\left(\frac{t}{2\pi}\sin\varphi + \left(r + \frac{t}{2\pi}\varphi\right)\cos\varphi\right)\right)d\varphi$$

is vector element of the coil in the point determined by the vector **s**.

$\mathbf{R} = \mathbf{P} - \mathbf{s}$ is vector between the observer and the coil.

$$\mathbf{B} = \left(I \cdot 10^{-7}\right)\int_0^{2\pi n} \frac{d\mathbf{s} \times \mathbf{R}}{|\mathbf{R}|^3}$$

is resulting magnetic field in the point **P** due to the current I.

The NMR method used for verification is described in [1]. The magnetic field was measured on the three-turn spiral-shaped coil with the parameters:

I = 4.5 mA, r = 0.01 m, t = 0.007 m, n = 3

The magnetic field was calculated and measured [1,3] in the three basic planes, Fig. 2 depicts the both calculated and measured values on centrelines in the plane xy and yz. It is evident very good agreement between the calculated and the measured values.

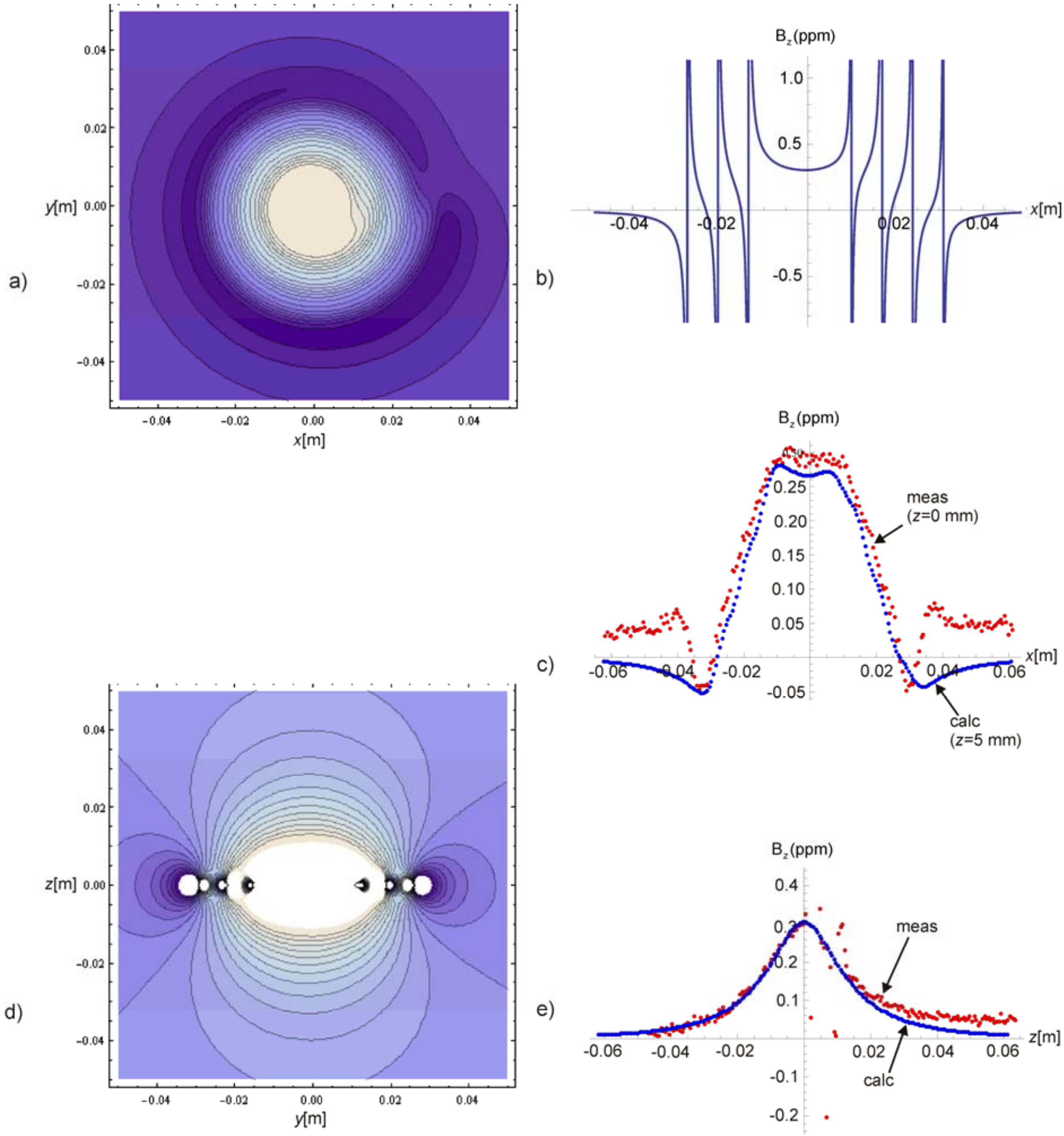


Fig. 2 Comparison between the calculated and the measured magnetic field of the three-turn spiral-shaped coil. The contour plot with the calculated values of the magnetic field in the plane *xy*, *z*=5 mm a); Calculated values at the centreline in the *xy* plane for *z*=0. The discontinuities due to zero distance between the conductor and the observer are obvious. b); A comparison between the magnetic fields at the centrelines in the plane *xy*: the measured for *z*=0, the calculated for *z*=5 mm. Despite of the different *z* coordinates the difference in the magnetic field values has not exceeded 10% in the region of interest c); The contour plot of the calculated values in the plane *yz* for *x*=0 d); A comparison between the calculated and the measured magnetic field at the centrelines in the *yz* plane. The difference in the right part of the graph was caused by a mechanical instability during the measurement but it is beyond the region of interest. e). The noise in the figure e) was caused by a plastic holder at the sample coil.

4. Discussion and Conclusions

Experiments confirmed correctness of the calculated values. Also the numerical integration was completed in relatively short time. The equation so can be used for a subsequent optimisation. Some inaccuracies occurring in the presented results could be caused by some of the following reasons: imperfections in the sample coil (it was rather small), imperfections in the measurement performing, drifting the DC current through the sample coil during the measurement. Whereas the measurement was performed during a stay abroad it can not be repeated simply. The presented equations can be simple used for the magnetic field of a spiral-shaped coil with arbitrary number of turns calculation. Spiral-shaped coil can be used as a surface RF coil for NMR [4,5] or as a device of phase arrays using advantageous signal-to-noise parameters of surface coils to volume measurements. From the presented figures of the magnetic field calculated at z=0 the discontinuity due to zero distance between the observer and the coil conductor is obvious. In the measured figures such discontinuity was not observed. The measured data were processed in different ways to explain the phenomenon e.g. the space close to z=0 was divided into appropriate grid to simulate voxels occurring in NMR measurement and to average the magnetic field within them, but the acquired results were still distant from the measured results. The probable reason is the fact that the measured values are limited by number of excited protons in the sample, but the calculated values can change without limitation. Examination of the phenomenon will continue.

Acknowledgements

Support of the Grant agency of the Slovak Academy of Sciences, grant no 2/0142/08 is gratefully acknowledged.

References

[1] Weis J., Andris P., Frollo I., Ahlström H.: A simple method for mapping the B_1 distribution of linear RF coil. *MAGMA*, 2005, 18: 283-287.

[2] Ilkovic D.: Physics. ALFA, Bratislava, 1972. (in Slovak)

[3] Andris P., Weis J., Frollo I., Ericsson A.: Magnetic field distribution of RF coil measurement using nuclear magnetic resonance. *Journal of Electrical Engineering*, 2002, 53: 32-34.

[4] Andris P., Frollo I.: Varicaps for NMR receiving coil matching and sensitivity changes. *Journal of Electrical Engineering*, 2007, 58: 291-293.

[5] Andris P.: Matching and tuning RF coils for NMR tomograph. *Measurement Science Review*, 2001, 1: 115-118.

Structural and Magnetization Properties of Ce doped $YBa_2Cu_3O_{7-\delta}$

[1]J. Polovková, [1]A. Cigáň, [2]T. Eslanloo Pereira, [1]J. Maňka, [1]P. Jurdák, [1]M. Kopčok

[1]Institute of Measurement Science, Slovak Academy of Sciences,
Dúbravská cesta 9, 841 04 Bratislava, Slovakia
[2]École Polytechnique Fédérale de Lausanne, Station 16,
CH – 1015, Lausanne, Switzerland
Email: julia.polovkova@savba.sk

Abstract. *YBCO polycrystalline samples doped with CeO_2 up to the level of 0.25 wt% have been prepared by the solid-state reaction method at the sintering temperature of 1007 °C in flowing oxygen. The critical temperature and transition width values slightly decrease with the increasing level of doping. Results from XRD patterns indicate that if cerium enters into the YBCO phase, its solubility limit is lower than 3.76 at % (x= 1.0 wt % of CeO_2). Starting from the CeO_2 level, the Y211 and $BaCeO_3$ phases could be identified. The addition of cerium leads to an increase of the magnetization and magnetization hysteresis and therefore to an increase of the critical current density for lower doping levels.*

Keywords: High-temperature Superconductors, $YBa_2Cu_3O_{7\pm\delta}$, Cerium Oxide, Resistance, XRD, Magnetization

1. Introduction

Although cerium belongs to rare earth elements, similarly as praseodymium, it does not form the $REBa_2Cu_3O_{7-\delta}$ superconducting phase, like other elements from this group [1]. On the other hand, cerium oxide together with the 211 phase is often used at the preparation of large mono-domains of $YBa_2Cu_3O_{7-\delta}$ in order to improve their superconducting properties [2]. While the Y211 phase inclusions have been found to act as the pinning centres and thus to increase critical current densities generally, the role of cerium oxide is to refine these inclusions in textured material. Cerium oxide is usually added at 0.5 - 1.0 wt% levels. As it was found, such amount of CeO_2 could slightly influence the critical transition temperature T_c and transition width ΔT_c of textured samples. Previously, it has been reported too, that additions of Ce and other combinations with Ce, as Pt or Sn, have different effects on the critical current densities j_c [3]. In this respect, the pinning by the $BaCeO_3$ inclusions has been proposed to have a main influence on intergrain j_c. The studied objects in most of the works represented melt-textured samples prepared from Y123 and Y211 phases and cerium oxide. In our study, we have investigated YBCO samples with the addition of cerium oxide, to reveal an effect of cerium oxide on superconducting and structural properties of the YBCO system itself.

2. Experimental

Polycrystalline samples of $YBa_2Cu_3O_{7-\delta}$ doped with 0, 0.25, 0.5, 1.0, 1.5, 2.0 and 2.5 wt% of CeO_2 were synthesized by the standard solid-state reaction method using Y_2O_3, $BaCO_3$, CuO and CeO_2 powders of the 99.99% purity. Mixtures of Y_2O_3, $BaCO_3$ and CuO in appropriate weight amounts were well mixed in an agate mortar in acetone, carefully dried and calcined at 930 °C for 40 hours in air. The obtained precursors were again homogenized in acetone, dried, mixed with CeO_2 powder up to 2.5 wt% and pressed isostatically into pellets. The samples were sintered in the horizontal tube furnace in flowing oxygen (10 ml/min) at 1007

°C for 24 h, followed by cooling to 520 °C and annealing at such temperature in oxygen for 24 h.
The phase composition of the samples was identified by powder X-ray diffraction using CuKα radiation. T_c(R = 0) and ΔT_c of the samples were determined using a standard resistance four-point method and the 10-90% criterion for transition width. Volume magnetization of the samples was measured at ~77 K after zero field cooling by a compensation method using the second-order SQUID gradiometer [4]. The applied magnetic field H_a was parallel to the axis of the sample.

2. Results and discussion

Temperature T vs. resistance R curves of CeO_2 doped $YBa_2Cu_3O_{7-\delta}$ samples are shown in Fig. 1. Values of T_c are in the range of 89.0-90.8 K and values of the transition width ~ 1.2-2.0 K for all the samples except for the undoped YBCO sample, which has the worst values. We believe that the bad superconducting properties of undoped sample were caused by a somewhat lower processing temperature with regard to its utmost value. Comparing the other doped samples, we can observe an increase of normal state resistance up to 1 wt% of doping with CeO_2 followed by a decrease in resistance for higher levels of doping. This could have been caused e.g. by changes in the oxygen content at the intergrain boundaries.

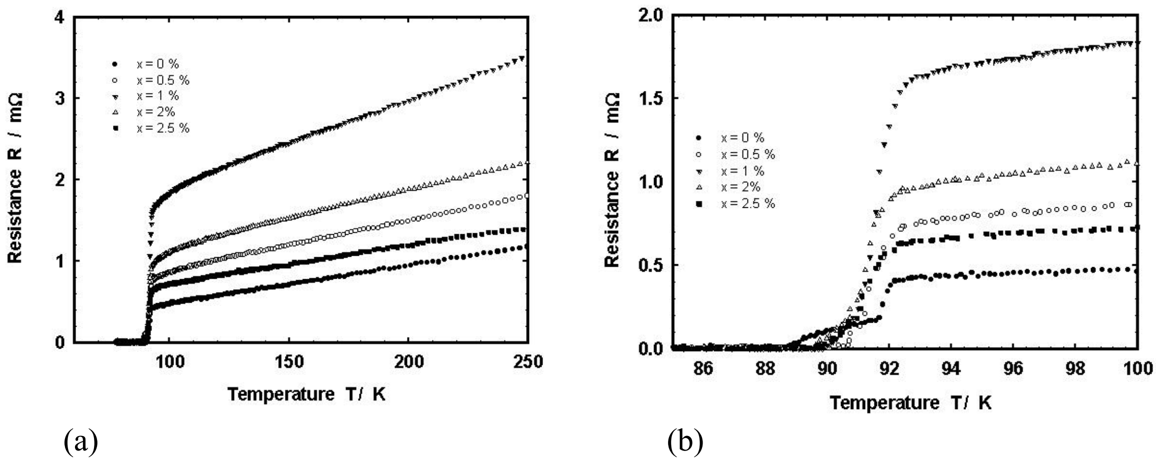
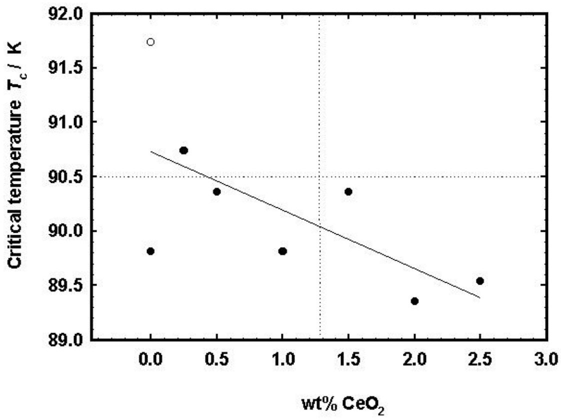


Fig. 1. R vs. T dependences of $YBa_2Cu_3O_{7-\delta}$ samples doped with x wt % of CeO_2 (a) and detail of R-T dependences in the temperature range of 85-100 K (b).

The T_c and ΔT_c dependences on increasing Ce-content in the YBCO compound are shown in Fig. 2 and 3, respectively. T_c slightly decreases with the increasing content of cerium oxide. This can be explained via the cerium atoms entering into the Y123 superconducting matrix and having mixed Ce^{3+} and Ce^{4+} valence with different ratios in the bulk and grain boundaries.
The powder XRD patterns of the undoped and doped with 2.5 wt% of CeO_2 YBCO polycrystalline samples are illustrated in Figs. 4 and 5, respectively. The diffraction patterns of undoped sample correspond to the characteristic orthorhombic Y123 phase. XRD data of the sample doped with 2.5 wt% of CeO_2, besides orthorhombic Y123 phase, show some peaks that can be assigned to green Y211 and $BaCeO_3$ phases. Tiny amounts of these phases were found already in the YBCO sample doped with 1 wt% of CeO_2. Results from XRD patterns indicate that if cerium enters into the YBCO phase, its solubility limit is lower than 3.76 at % (x = 1 wt % of CeO_2). This is also in accordance with results of Petrov et al. [5].



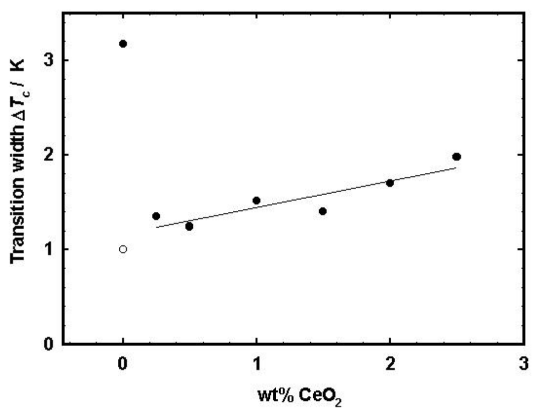
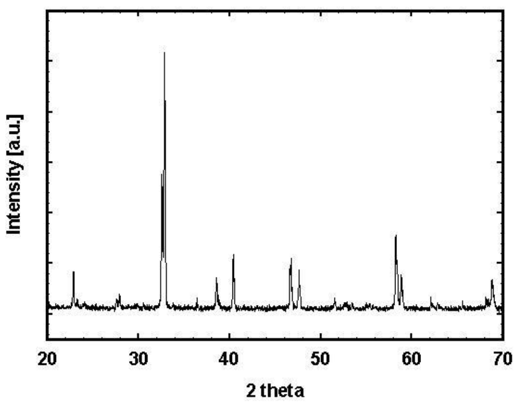
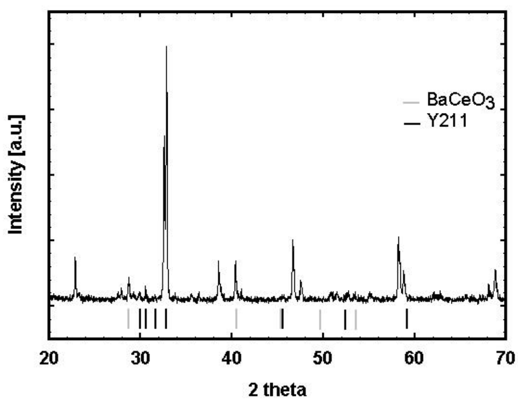


Fig. 2. and 3. T_c and ΔT_c vs. nominal content in wt% of added CeO_2 into YBCO samples, respectively. (○ – standard YBCO sample without Ce).

Changes in R vs. T dependences induced by an addition of cerium can also be explained by an entry of cerium into the YBCO matrix and formation of Y211 and $BaCeO_3$ phases.



(4)



(5)

Fig. 4. and 5. Powder X-ray diffraction patterns of the undoped YBCO superconductor (4) and YBCO sample doped with 2.5 wt% of CeO_2 (5).

AC volume magnetization M vs. the applied magnetic field H_a of undoped sample and samples doped with x = 0.00, 0.25, 0.5 and 1.5 wt% of CeO_2 are shown in Fig. 6. All the samples show the Z-shape magnetization curves typical for the polycrystalline samples. In the range of a higher magnetizing field amplitude, the hysteresis of magnetization loops grows with an increase of the Ce level up to x =0.50, which is followed by a decrease of magnetization and magnetization hysteresis with the continuation of adding CeO_2. The results are in accordance with the scheme of Ce entering into YBCO at lower levels and formation of Y211 and $BaCeO_3$ phases at higher contents**.** Substituted Ce atoms in the YBCO matrix can act as effective point pinning centres of magnetic flux resulting in an increase of magnetization hysteresis and so in an increase of critical current density. The same is true for Y211 and $BaCeO_3$ phases, however, the formation of these phases decreases the content of the YBCO phase that results in a decrease of magnetization values, see Fig. 5. plot with x = 0.25 wt % CeO_2.

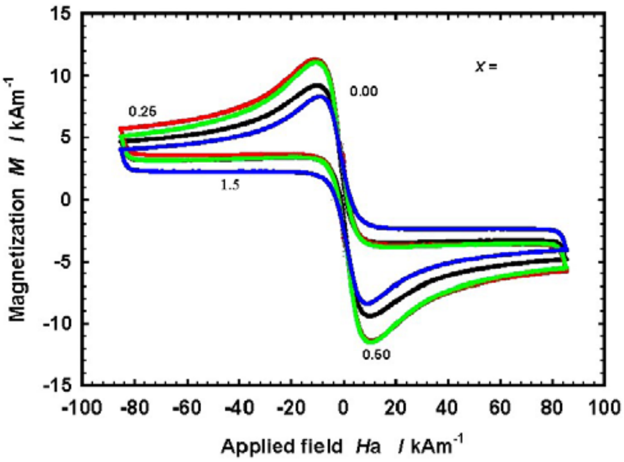


Fig. 6. AC volume magnetization M vs. the applied magnetic field H_a of undoped sample and samples doped with x = 0.00, 0.25, 0.5 and 1.5 wt% of CeO_2.

3. Conclusions

Effects of cerium oxide addition on the structural, transition, and magnetic properties of $YBa_2Cu_3O_{7-\delta}$ polycrystalline samples have been investigated by means of X-ray diffraction, resistance, magnetization and optical measurements. The addition of cerium causes a slight decrease in T_c from 90.8 to 89.0 K and an increase in ΔT_c from 1.2 to 2.0 K. Based on phase transition and XRD measurements, it can be inferred that cerium enters into YBCO compounds. Starting from the 1.0 wt% CeO_2, the Y211 and $BaCeO_3$ phases could be identified. The addition of cerium leads to an increase of the magnetization and magnetization hysteresis and therefore to an increase of critical current density, including significant intragrain or intergrain cluster contributions, with the maximum at x = 0.025 - 0.05 wt % CeO_2. Another increase causes a deterioration of the characteristics.

Acknowledgements

The authors thank to the Slovak Grant Agency for Science, which has provided the financial support for the 2/7083/27 project.

References

[1] Yang KN, Lee BW, Maple MB, Laderman SS. Compounds in mixed phase $CeBa_2Cu_3O_y$ and $TlBa_2Cu_3O_y$ systems. *Appl. Phys. A*, 46: 229-232, 1988.

[2] Monot I, Verbist K, Hervieu M, Laffez P, Delamare MP, Wang J, Desgardin G, Van Tendeloo G. Microstructure and flux pinning properties of melt textured grown doped $YBa_2Cu_3O_{7-\delta}$. *Physica C,* 274: 253-266, 1997.

[3] Harnois C, Hervieu M, Monot-Laffez I, Desgardin G. Relations between microstructure and superconducting properties of Ce+Sn doped YBCO. *J.Europ.Ceram. Soc.* 21: 1909-1912, 2001.

[4] Zrubec V, Cigáň A, Maňka J. Simplified fast method for magnetic characteristics measurement of the HT_c superconducting materials. *Physica C*, 223: 90, 1994.

[5] Petrov MI, Balaev DA, Gokhfeld YS, Dubrovskii AA, Shaikhutdinov KA. Effect of heterovalent substitution of rare-earth elements on the magnetic and transport properties of $YBa_2Cu_3O_7$. *Phys. Sol. State,* 11: 2047-2051, 2007.

Diffuse Reflectance and Transmittance Measurements Using a STAR GEM® Optical Accessory

[1,2]E.Kawate

[1]AIST, 1-1-1, Central 2, Umezono, Tsukuba, Ibaraki 305-8568, Japan
[2]TRAS, Inc., 725-28 Shimohirooka, Tsukuba, Ibaraki 305-0042, Japan
E-mail address: e.kawate@aist.go.jp

Abstract. *Four methods, a goniometer, integrating sphere, conical mirror and biconical mirror methods, for the measurement of the diffuse reflectance and transmittance are described and compared. These have been in use for a number of years. A STAR GEM, which has been recently developed in AIST, belongs to the category of the biconical mirror. The disadvantage of the biconical mirror is that it does not collect scatter at all angles, so anisotropic materials are not measured accurately. This disadvantage may be solved by such an improvement of the STAR GEM as the second ellipsoidal mirror can rotate around a GEM axis.*

Keywords: STAR GEM, biconical mirror, absolute reflectance, diffuse reflectance, BRDF, THR, absolute transmittance, diffuse transmittance, BTDF, THT

1. Introduction

Whether materials are almost purely diffusive, as in a lightly packed powder, predominately specular, as in an aluminum mirror, or something in between, such as a glossy paint sample, diffuse reflectance and transmittance spectroscopy tells us much about the physical and chemical characteristics that are not available by other analytical means. Four accessories to perform diffuse reflectance and transmittance measurements are based on the use of a goniometer, an integrating sphere, a conical mirror and a biconical mirror [1].

A STAR GEM accessory has been developed, at first, for the measurements of absolute reflectance and transmittance of specular materials in the wavelength region from 0.24 μm to 25 μm in AIST [2]. The STAR GEM is an acronym of Scatter, Transmission, and Absolute Reflection measurements using a Geminated Ellipsoid Mirror. This development succeeded beyond all expectations. Next step is to develop a STAR GEM as a scatterometer. It is ultimate that the scatterometer can measure absolute reflectance and absolute transmittance of a diffusive material as well as anisotropy of scatter from the diffuse material. Although the STAR GEM belongs to the category of the biconical mirror, it is expected that the STAR GEM may measure both the bidirectional transmittance/reflectance distribution function (BTDF/BRDF) and the total hemispherical transmittance/reflectance (THT/THR) by one accessory, because it carries two strong points of the goniometer and the integrating sphere. The first measurement about the BTDF using the STAR GEM was already published [3]. In this paper the possibility will be discussed to measure the THT/THR using the STAR GEM. Because there is not so much difference between diffuse reflectance and diffuse transmittance measurement methods, mainly the diffuse transmittance measurement method is described below.

Textiles, paper, and rough materials will show changes in the apparent reflectance just by rotation of the sample, which will not be observed using an integrating sphere. It is often said that the BRDF measurement is more a measurement of appearance than the THR. On the other hand, it is important to measure the THR of many materials in the R&D for the utilization of the solar energy in the wavelength region from 240 nm to 50 μm.

2. Comparison of four traditional accessories

Measurements by a goniometer involve tilting and rotating the sample or detector about the sample center in order to measure beam transmitted as a function of illumination and collection angles at selected point on the hemisphere above the sample. This function is called the bidirectional transmittance distribution function (BTDF). This function makes an anisotropy of the diffuse transmission clear. The time required to perform a goniometric measurements has decreased significantly due to advances in computer-controlled instrumentation. Nevertheless, the disadvantage is that goniometric measurements are still time consuming due to the large number of data points required, when people want to know not only the anisotropy (BTDF) but also the total hemispherical transmittance (THT). The THT can be obtained through the spatial integration of the BTDF all over the hemisphere.

Methods based on the integrating sphere are by far the most common in use today. A beam is focused onto the sample, which lies on a line tangent to the inside radius of the sphere. The transmitted radiation is collected by the sphere. The detector lies at another point on the inside tangent of the sphere and collects the diffusely transmitted radiation. As a result the THT can be measured, but the anisotropy can't be measured. Other disadvantages are that a sphere is by very nature an attenuating device and that the corrections of a sphere zero error and a substitution error are necessary and difficult [1].

In conical mirror accessories, the light source is placed at one focus of a highly specular hemiellipsoid and the sample is placed at the other focus. A reverse geometry also exists when the sample is illuminated by a directional beam with the detector lying at the other focus of the hemiellipsoid. The advantage is a great increase in efficiency of collection. The disadvantages are that the interreflections between the sample and detector or source and the position misalignments of the sample and detector or source become significant error sources. Another disadvantage is the cost required to produce such hemiellipsoid.

In biconical mirror accessories, an ellipsoidal mirror focuses an incident beam at a small point on a sample and the scattered transmitted light is caught by another ellipsoidal mirror that directs the beam to a detector. The advantage is that it allows one to get a spectrum on a far smaller sample that is required with an integrating sphere. The disadvantage is that it does not collect scatter at all angles, so anisotropic materials are not measured accurately. For this reason, biconical mirror accessories are generally thought of as qualitative.

3. Experimental details

An idea of a GEM is based on the fact that light emitted from a focus of the ellipsoidal mirror converges at the other focus [4]. The GEM consists of two ellipsoids of revolution (E1 and E2). A focus of the E1 coincides with that of the E2 and this common focus (F0) aligns with two remaining foci (F1 and F2). A sample is placed at F0 and two rotating plane mirrors (RM1 and RM2) are placed at F1 and F2, respectively. The RM1 and RM2 can rotate independently by each stepping motor.

We have already developed three types of the STAR GEM [3]. In a STAR GEM Type 3 in Fig. 1, a Ceiling was divided into two parts and each part was fastened on the E1 and E2 separately. The E1 was fixed on the Base in the previous manner. On the other hand, some part of the wall of a Base under the E2 was cut out and the E2 wasn't fixed on the Base, so the E2 was able to rotate around the GEM axis. A sample was placed vertically and incident light advanced along the GEM axis, as a result, the incident plane was a horizontal plane including the GEM axis. The STAR GEM Type 3 was an improved accessory to measure scattered radiation from a diffuse material not only in the incident plane but also out of the incident plane.

Fig. 1 STAR GEM Type 3 for scattered light measurements

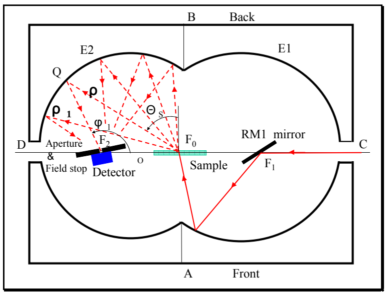


Fig. 2 Structure of GEM for BTDF measurements and beams

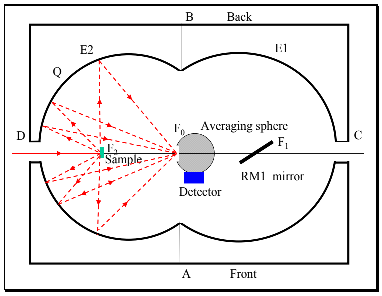


Fig. 3 Structure of GEM for THR measurements and beams

The STAR GEM Type 3 was used with a He-Ne laser and a silicon photodiode detector in Fig. 2. The laser was placed in the right hand outside the STAR GEM. The detector replaced the RM2 in Fig. 2. This configuration was to measure the BTDF of a transmissive diffuser. In order to measure the BTDF with high spatial (angular) resolution, a field stop and an aperture were attached in front of the detector. A laser beam was chopped to reduce both optical and electric noise. This was accomplished through the use of lock-in detection. A monitoring detector was also used to allow the computer to ratio out laser power fluctuations. The incident laser beam was focused onto the RM1 and also the sample surface and the detector following the nature of the ellipsoidal mirror. The source and the sample were fixed. The RM1 was rotated to a desired incident angle. The detector at F2 was also rotated to a desired collecting angle to collect scattered radiation from the sample.

Another configuration was to measure the THR of diffuse radiation using the STAR GEM Type 3. In Fig. 3 a sample, such as a reflective diffuser, replaced the detector in Fig. 2 and was settled to be perpendicular to the GEM axis at F2. A detector with an averaging sphere also replaced the sample in Fig. 2 and was settled at F0. The averaging sphere was used to eliminate the spatial anisotropy of the response of the detector. The laser beam entered into the STAR GEM along the GEM axis from the left side in Fig. 3. The ellipsoidal mirror of the E2 rotated around the GEM axis from χs= -65 degrees to χs= 65 degrees. It was impossible to rotate until ±90 degrees because of the existence of a rod to support the sample. In this configuration the detector received the total scatter field on the E2 due to a large aperture of the averaging sphere except for a specular component of reflection from the reflective diffuser because of an entrance hole for the laser beam on the E2, which is shown as D in Fig. 3. The sample diameter was limited within a solid angle, where the hole on the E2 was looked from F0. It was less than 10 mm. This configuration was like a conical mirror set up.

A sample was a ground glass diffuser (DFSQ1-30C02-240) made by SIGUMA KOKI CO. and was a transmissive diffuser. A front surface was a frosted glass and a back surface was polished. Its diameter was 30 mm and its thickness was 2 mm. Another sample was a reflective diffuser. Its diameter was 10 mm and its thickness was 1 mm.

4. Results and Discussions

Attentions must be paid to the calculation of the BTDF measured by the STAR GEM [3]. In the goniometer, the detector rotates around the sample and the detector field of view (FOV) is always constant, because the distance between the sample and the detector is constant. On the other hand, in the STAR GEM the detector rotates around the F2. The sum of two distances from a point of Q, which is on the E2 in Fig.2, to F0 and to F2 is constant because of the ellipsoid. These two distances change simultaneously, when a collection angle, θs, changes in Fig. 2. As a result the field of view (FOV), where the detector looks at the sample in the ellipsoid mirror of the E2, also changes depending on the θs.

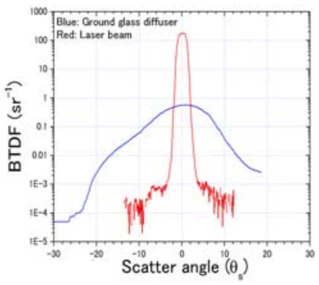


Fig. 4 BTDF and laser beam profile

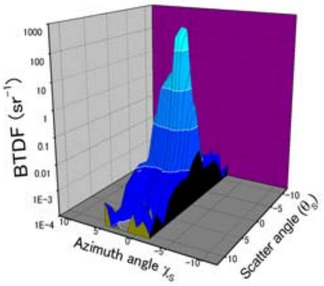


Fig. 5 Three dimensional profile of laser beam

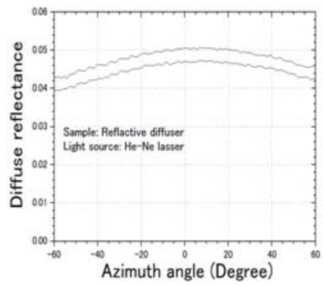


Fig. 6 Diffusive reflectance

In Fig. 2 a ground glass diffuser was illuminated using a 0.633 μm-He-Ne laser. In order to measure the angle dependence of scattered radiation from the transmissive diffuser, the detector was rotated around F2, so its FOV was swept through the scatter field on the E2. The BTDF is shown in Fig. 4. The spatial profile of the laser beam is also shown in Fig. 4. It is clear that the diffuser scatters the laser beam over a wide range of angles.

The STAR GEM Type 3 allowed the E2 also to rotate around the GEM axis. This rotation corresponded to the change of the azimuth angle, χs, from the incident plane. While the detector was simultaneously rotating around F2 and around the GEM axis, the intensity of the laser beam was measured. A three dimensional profile of laser beam is shown in Fig. 5. A full width at half maximum is 1.8º and its maximum value is 200 sr-1. The profile of the laser beam is not isotropic but is broad in the positive azimuth angle region. This result isn’t found from the measurements only in the incident plane in Fig. 4. The STAR GEM Type 3 may measure the anisotropy of the diffuse transmittance as well as the spatial profile of a laser beam.

The diffuse reflectance of the reflective diffuser, which was measured using the setup of Fig.3, is shown in Fig.6. The χs= 0 degree means that the E2 is horizontal. There are two curves, because the diffuse reflectance was measured two times by rotation of the sample around the normal to the surface by 90 degrees. Although both signals are decreasing with increasing the azimuth angle, this is not due to anisotropy of the sample but due to the shadow of the rod. In this setup the signal intensity is sensitive of the sample position. The intensity of both signals doesn’t correspond each other because of the shadow of the sample itself. In this configuration, it is difficult to measure a transmissive diffuser and there is another problem that the diffuse reflectance in Fig. 6 always includes the signal from the neighborhood of the D point in Fig. 3 independent of the azimuth angle.

When the configuration for the BTDF measurements in Fig. 2 compares with the configuration for the THR measurements in Fig.3, the configuration of Fig. 2 is superior because there is no shadow of the supporting rod and the sample itself and because not only diffuse reflectance/transmittance but also specular reflectance/transmittance can be measured. The disadvantage is a small aperture in Fig. 2. In the next step we will improve the field of view in Fig. 2 and find the method to obtain the THR/THT using the STAR GEM Type 3.

References

[1] Workman J. and Springsteen A.W., APPLIED SPECTROSCOPY, ACCADEMIC PRESS, San Diego, 1997.

[2] Kawate E. and Hain M. Regular reflectance and transmittance measurements of transmissive materials using a STAR GEM® optical accessory. To be submitted to the 7th International Conference of the Measurement 2009.

[3] Kawate E. Measurement method of optical scatter using a STAR GEM as a scatterometer. In proceedings of SPIE, vol. 7065, 2008, 706515-1-706515-9.

[4] Kawate E. An optical accessory for absolute reflection and transmission measurements in the wavelength region from 0.24 μm to 25 μm. In proceedings of SPIE, vol. 7065, 2008, 70650M-1-70650M-11

Dynamic Calibration of the Transient Plane Source - sensor

S. Malinarič

Department of Physics, Constantine the Philosopher University, Nitra, Slovakia
Email: smalinaric@ukf.sk

Abstract. *The Extended Dynamic Plane Source (EDPS) method is one of transient methods for measuring thermal conductivity and diffusivity in solids. This technique uses a transient plane source (TPS) sensor, which serves as the heat source and thermometer. Its calibration consists in measuring the temperature dependence of the TPS sensor resistance and computing the temperature coefficient of resistance (TCR) using the least squares (LS) estimation. The goal of this work is to calibrate the TPS sensor directly in the apparatus for EDPS measurement.*

Keywords: *transient method, thermophysical parameters, transient plane source - sensor, temperature coefficient of resistance*

1. Introduction

Transient methods [1] are used for measurement of thermophysical parameters, e.g. thermal conductivity λ and diffusivity a. Measurements are based on generation of the dynamic temperature field inside the specimen. The theoretical model of the method is characterized by a temperature function, which is a solution of the heat equation with boundary and initial conditions corresponding to the experimental arrangement. The principle of the evaluation is based on fitting of the temperature function to the experimental points (temperature response), determined by the TPS sensor resistance measurement. Hence, the calibration of the sensor is necessary for obtaining reliable values of thermal conductivity measurement.

2. Extended dynamic plane source (EDPS) method

The EDPS [2-4] method is characterized by one-dimensional heat flow into a finite solid body with low thermal conductivity. Fig.1 shows the TPS sensor in the form of a meander, made from a 20 μm thick nickel foil and covered on both sides with 25 μm kapton layer. The sensor is placed between two identical specimens with the same cross section. Heat sink, made of a very good heat conduction material (aluminium), provides isothermal boundary conditions for the experiment. Heat is produced by the passage of an electrical current in the form of a step-wise function through the TPS sensor.

The apparatus enables to increase the temperature of the experimental set-up. The electronics of the apparatus consists of a platinum thermometer (Pt100), 2 heating elements, a multichannel PC plug-in card (PCL-816) and a power DA converter. A proportional integral (PI) controller realized by PC software is used for temperature control. It is based on periodic measurement of the temperature and computing the manipulated variable (heating power). PI control allows good temperature stability and homogeneity in the heat sink and specimens (better than 5 mK).

The resistance of the platinum thermometer and TPS sensor are calculated by formulas

$$R_T = R_1 \cdot \frac{U_T}{U_1} \qquad R_S = R_3 \cdot \frac{U_S}{U_3} \tag{1}$$

where $R_1 = 136\ \Omega$ and $R_2 = 1.00\,\Omega$ are the constant resistors and voltages are measured as shown in Fig.2. To suppress quantization and electrical noise, voltages U_1, U_T and U_3, U_S are sampled and averaged 4000 and 1500 times per channel over the period of 1s and 60ms, respectively. Currents in the circuits were set to $I_1 = 1.2\,\mathrm{mA}$ and $I_3 = 300\,\mathrm{mA}$. The temperature of the Pt100 was determined by using the following formula [5]

$$R_T = R_0 \cdot \left(1 + \alpha_T \cdot T + \beta_T \cdot T^2\right), \qquad (2)$$

where $R_0 = 100.00\,\Omega$, $\alpha_T = 3.9092 \cdot 10^{-3} \cdot K^{-1}$ and $\beta_T = -5.917 \cdot 10^{-7} \cdot K^{-2}$.

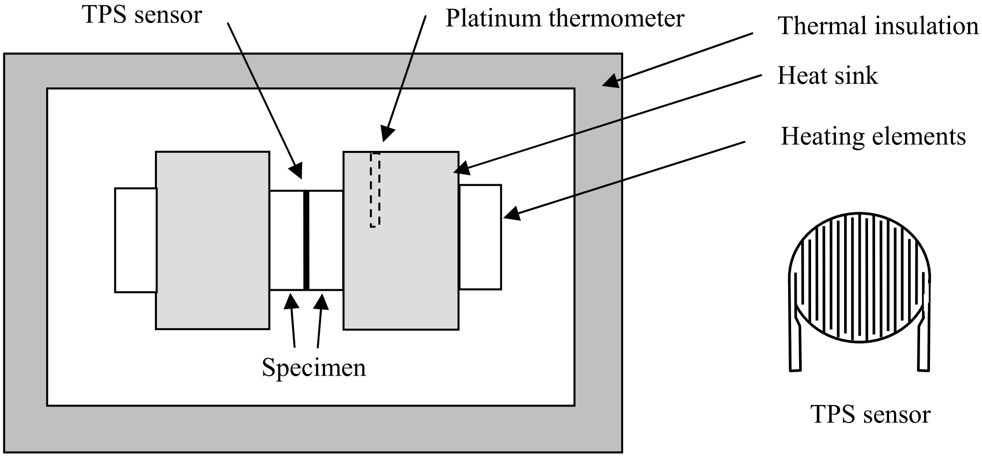


Fig. 1 The arrangement of the experiment

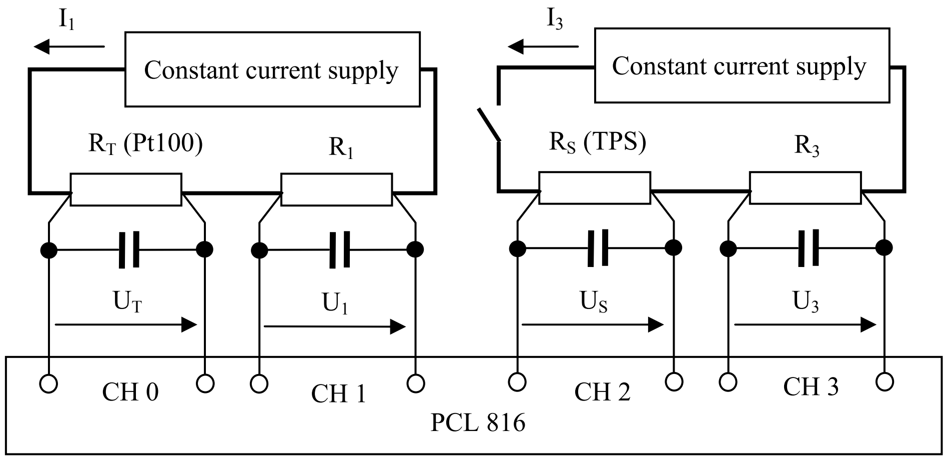


Fig. 2 Experimental circuit design

3. Temperature coefficient of resistance (TCR) measurement

In order to use the described apparatus for TCR measurement, the specimens were replaced by aluminium ones and coated with silicon oil to improve thermal contact between the TPS sensor and the heat sink. As the current I_3 caused the heating of the sensor itself, the extrapolation to the zero time was applied, as illustrated in Fig. 3. The measured data were fitted to the following polynomial [6]

$$R_S = a_0 + a_1 \cdot T + \cdots + a_k \cdot T^k. \qquad (3)$$

The LS estimate of the parameter vector is given by [7]

$$\vec{a}_{\mathrm{LS}} = \left(\mathbf{X}^T \cdot \mathbf{X}\right)^{-1} \cdot \mathbf{X}^T \cdot \vec{R} \tag{4}$$

where $\vec{R}$ is the observation vector of the TPS sensor resistance measured at 6 points T_i in the interval from 20 to 45°C and **X** is a sensitivity matrix defined by

$$\{\mathbf{X}\}_{ij} = T_i^{\,j} \tag{5}$$

Once we have the parameter estimates, the TCR of the TPS sensor can be computed using the relation

$$\alpha(T) = \frac{1}{R_S(T)} \frac{dR_S(T)}{dT} \tag{6}$$

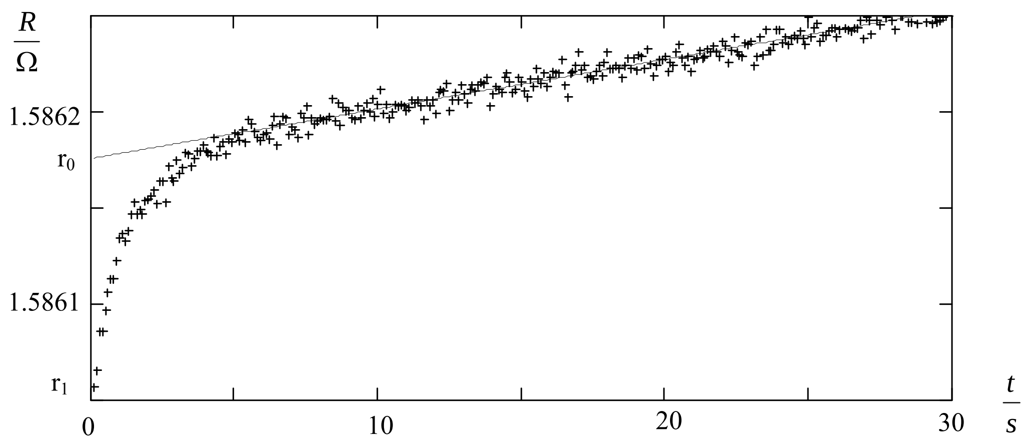


Fig. 3 Temperature dependence of the TPS sensor resistance after switching the current I_3 on.

Type A standard uncertainty (LS component) of the TCR estimate can be computed by [8,9]

$$u_A(\alpha) = s\sqrt{\vec{c}^{\,T} \cdot \left(\mathbf{X}^T \cdot \mathbf{X}\right)^{-1} \cdot \vec{c}} \tag{7}$$

where s is a standard deviation of residuals, $\vec{c}$ is a vector of sensitivity coefficients defined by

$$c_j = \frac{\partial \alpha}{\partial a_j}, \tag{8}$$

and α is given by Eq. 6 and 3.

4. Results and discussion

The evaluation was performed with both extrapolated values r_0 and first sample (0.1 s after switching) values r_1. Table 1 shows the results of fitting for three values of polynomial order k. A simplified uncertainty assessment is presented in Table 2. The total standard uncertainty of TCR estimation at temperatures from 20 to 45°C does not exceed $0.035 \cdot 10^{-3} \cdot \mathrm{K}^{-1}$ (0.7%), which is sufficient for thermal conductivity measurement of materials.

Acknowledgment

The author wishes to thank the Slovak Science Grant Agency for financial support under Contract 1/0302/09.

Table 1. Results of LS estimation, k is the polynomial degree, α is the TCR of TPS sensor and s is the standard deviation of residuals.

	extrapolated values r_0			first sample values r_1		
k	2	3	4	2	3	4
$\alpha_{20°}\cdot 10^3\cdot K$	4.828	4.842	4.840	4.828	4.840	4.829
$\alpha_{40°}\cdot 10^3\cdot K$	4.618	4.618	4.618	4.617	4.618	4.616
$s\cdot(\mu\Omega)^{-1}$	20.7	1.5	1.2	20.2	10.6	8.2

Table 2. Uncertainty budget for TCR of TPS sensor measurement (k = 2, extrapolated values)

Source of uncertainty	Standard uncertainty	Value	Standard uncertainty $u(\alpha_{20°})\cdot 10^3\cdot K$
LS component (Eq.7)	$u_A(\alpha)$		0.003
R_1 measurement + temperature stability	$u_B(R_1)$	26 mΩ	0.001
R_3 measurement + temperature stability	$u_B(R_3)$	4.4 mΩ	0.023
Pt100 calibration	$u_B(T)$	100mK	0.027
Combined uncertainty			0.035

References

[1] Kubičár Ľ, Boháč V. Review of several dynamic methods of measuring thermophysical parameters. In proceedings of the 24th Int. Conf on Thermal Conductivity, 1997, 135-149.

[2] Karawacki E, Suleiman B M, ul-Hag I, Nhi B. An extension to the dynamic plane source technique for measuring thermal conductivity, thermal diffusivity, and specific heat of dielectric solids. *Review of Scientific Instruments,* 63 (10): 4390-4397, 1992.

[3] Malinarič S. Parameter estimation in dynamic plane source method. *Measurement Science and Technology,* 15: 807-813, 2004.

[4] Malinarič S, Dieška P. Improvements in the Dynamic Plane Source Method. Internation Journal of Thermophysics, DOI 10,1007/s10765-008-0525-7

[5] Bentley J P. Temperature sensors characteristics and measurement system design. *Journal of Physics E: Scientific Instrument,* 17: 430-439,1984.

[6] Suleiman B M, Gustafsson S E, Börjesson L. A practical cryogenic resistive sensors for thermal conductivity measurements. *Sensors and Actuators A*, 57: 15-19, 1996.

[7] Beck J V and Arnold K J. Parameter Estimation in Engineering and Science. Wiley, New York, 1977.

[8] Guide to the Expression of the Uncertainty in Measurement. ISO, Geneva, 1993.

[9] Chudý V, Palenčár R, Kureková E, Hamaj M: Meranie technických veličín. Vydavateľstvo STU, Bratislava, 1999.

Development of a New Model for the Pulse Transient Technique

[1]V. Boháč, [2]P. Dieška , [1]Ľ. Kubičár

[1]Institute of Physics SAS, Dúbravská cesta 9, 845 11 Bratislava, Slovakia

[2]Department of Physics, Faculty of Electrical Engineering and Information Technology, STU, Ilkovičova 3, Bratislava, Slovakia

Email: bohac@savba.sk

Abstract. *The problems connected with deficiency in a large amount of testing material cause some problems in data evaluation as an ideal model usually assumes infinitively large specimen. The finite geometry of the specimen cause additional effects that harm the efficiency of the measurement evaluation. In this paper a new model was introduced that includes the effect of heat capacity of the heat source and completes the series of models introduced previously for the effects of heat losses from the sample surface.*

Keywords: pulse transient method, thermal conductivity, thermal diffusivity, specific heat

1. Introduction

Ideal model with infinite specimen geometry is used to keep a low number of unknown parameters but sometimes do not satisfy the real experiment. A detailed study has to be performed to find experimental circumstances when disturbing effects influence ideal model [1, 2]. Then, a modified model for the real experiment has to be used to take into account additional disturbing effects characterized by corresponding, and usually unknown parameters [3, 4]. In the past we introduced a difference in models based on ideal case when assuming non-infinite specimen geometry and the real pulse duration. New models assume real sample radius, heat capacity of the heat source and heat loss effect from the free sample surface. There were several papers published that covered the problems of the specimen geometry optimization on accuracy of the measurement. Ideal model assumes the infinitively large specimen as well as infinitively thin heat source having neglected heat capacity. Some effects were observed through measurements for thin thicknesses of the investigated material. These additional effects are influencing the measurement and thus the estimated parameters were covered by additional error [1, 2, 3].

For the lower and higher thicknesses the measured values of the thermophysical parameters e.g. specific heat, thermal diffusivity and thermal conductivity, were apparent, e.g. overestimated or underestimated. Analysis of this experiment show that heat losses from the sample surface lowered the measured temperature response and this effect rises with increasing material thickness. In case of lower thicknesses it was found influence of heat capacity of the heat source. Models that include the heat transfer coefficient from the sample surface to the surrounding were published previously [3, 4].

This paper concentrates on the effect caused by heat capacity of the heat source observed on the measurements at lower material thicknesses. In this case the evaluated parameters were below the recommended ones. This effect was caused by inadequate ratio between the heat capacity of the heat source and the measured material [5]. This was the main reason to introduce next parameter - a heat source capacity in a physical model.

2. Experimental method

The principle of the method is to record the temperature transient response to the heat pulse generated by a plane heat source and to calculate the thermophysical parameters from the characteristic features of the measured curve of the temperature response (Fig. 1. right). Transient temperature response measured at the distance h from the heat source is calculated according temperature function *T(h,t)* providing that the ideal model (Eq. 1.) is valid [1, 2]. The ideal model assumes that a planar temperature flow is not deformed as it penetrates into the depth of the specimen bulk (white-dotted area in the Fig. 1). The problem is that the temperature isotherms are not planar over the cross section of the specimen but they are deformed at the edges by the heat losses from the sample surface for large distances.

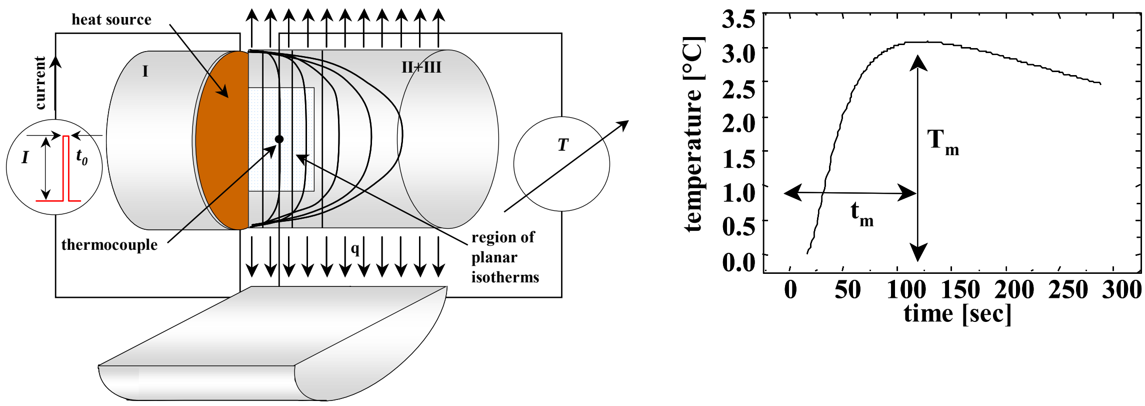


Fig. 1. The principle of the pulse transient method. Specimen set is drawn with heat flow paths when drawn isotherms are influence by heat loss effect (left). An example of the temperature response is on the right.

Model

In previous experiments a correction to the ideal model considering the real pulse width was applied to ideal model. The model is characterized by [1]

$$T(h,t) = \frac{2 \cdot Q}{c\rho\sqrt{\kappa}} \left[\sqrt{t} \cdot i\Phi^{*}\left(\frac{h}{2\sqrt{\kappa t}} \right) - \sqrt{t - t_0} \cdot i\Phi^{*}\left(\frac{h}{2\sqrt{\kappa (t - t_0)}} \right) \right] \tag{1}$$

where $i\Phi^{*} = \frac{e^{-x^2}}{\sqrt{\pi}} - x \cdot erfc(x)$. Here Q means heat flow density at source, c is specific heat, κ is thermal diffusivity, ρ is density and t is time. Equation 1 should be used for data evaluation by fitting procedure.

One point evaluation procedure

At the standard experiment due to fast calculations we use simple relations for the evaluation of thermal diffusivity, specific heat and thermal conductivity. These relations were derived for the maximum of temperature response on Fig.1. (one-point evaluation procedure). The thermal diffusivity is calculated according to the equation

$$\kappa = h^2 \cdot f_\kappa / 2t_m \tag{2}$$

and the specific heat

$$c = Q \cdot f_c / \sqrt{2\pi e}\rho h T_m \tag{3}$$

where f_κ and f_c are correction factors [5] and ρ is the density of material. Maximum temperature of transient response is T_m at time t_m (Fig. 1.) Thermal conductivity is given by

$$\lambda = a\rho\, c = h \cdot Q \cdot f_\kappa \cdot f_c \big/ 2\sqrt{2\pi e}\ t_m T_m \tag{4}$$

Real model assuming heat capacity of the heat source

A previous explanation of a heat capacity effect was solved in a new model by defining the initial and boundary conditions for the basic heat transport equation. Specimen set is arranged symmetrically in a form of planar boards inserted in between the heat exchangers having infinite large heat capacity stabilized at certain temperature. The thermal contact with the specimen is ideal (thermal contact resistance is zero). The heat source having non-zero heat capacity as well as perfect thermal contact with specimen is placed in between specimen boards. The solution of the heat equation is a temperature function in the form

$$T(t,x) = T_0 \left\{ \left(1 - \frac{x}{L}\right) + 2a \sum_{\upsilon} e^{-\frac{kt}{L^2}\upsilon^2} \times \frac{\upsilon \sin(\upsilon \frac{x}{L}) - a\cos(\upsilon \frac{x}{L})}{\upsilon^2 [\upsilon^2 + a(a+1)]} \right\}. \tag{5}$$

where $T_0 = \dfrac{qL}{\lambda}$, $a = \dfrac{\lambda L}{C\kappa}$; T is temperature increase, t is time, x Cartesian coordinate, L thickness of sample, q heat flow density at source, λ thermal conductivity, κ thermal diffusivity, C. heat capacity per unit area of source, v is a root of the equation $a\cos\upsilon - \upsilon\sin\upsilon = 0$ and k is the Stefan-Boltzmann constant. The relation (5) characterizes the step-wise measuring regime. After the duration of the heat pulse t_0, the temperature is expressed by the relation

$$T^*(t,x) = T(t,x) - T(t - t_o, x) \tag{6}$$

where $T(t,x)$ and $T(t - t_0, x)$ are given by the relation (5). The relation (6) characterizes the pulse transient regime.

3. Experiment and results

The specimen dimensions of calcium silicate reinforced by cellulose fibers were 150x150 mm in cross section and 40 mm in thickness (all 3 parts of the specimen set). The middle part was cut on two halves. The thermocouple was placed at different position from the heat source, to measure the thermophysical parameters for a range of thicknesses. The experimental details were described in [2, 4]. The theoretical temperature responses were calculated using thermophysical parameters evaluated by three different procedures – one point evaluation (Eq. 2 and 3.), fit of the data using Eq. 6. for the thicknesses up to 17 mm, and model described in [4] that account heat losses from the sample surface for the thicknesses bigger than 17 mm.

In the past the use of the ideal model (Eq. 1) and the one point evaluation procedure led to a problem with the heat capacity that was avoided by optimization of specimen geometry [2]. In practice this limit resulted in limits of the specimen thickness between 17 – 25mm for a class of porous building construction materials. Fig. 2 shows an illustration of this effect measured on a calcium silicate board reinforced by cellulose fibers (square points).

4. Conclusions

The new model assuming heat capacity of the heat source completes the series of models introduced previously for the effects of heat losses from the sample surface. In this case the values of thermophysical parameters were shifted towards the recommended one at lower thicknesses of the material and extended the range of data validity for a broader range of thicknesses. The statistical difference of data within ±5% is marked in Fig 2. for all parameters. The temperature responses theoretically calculated from the parameters obtained by fitting procedure are in good coincidence with the experimental one (Fig. 1.). The one point evaluation procedure that uses the ideal model is available only in the case when optimized specimen geometry is used [2].

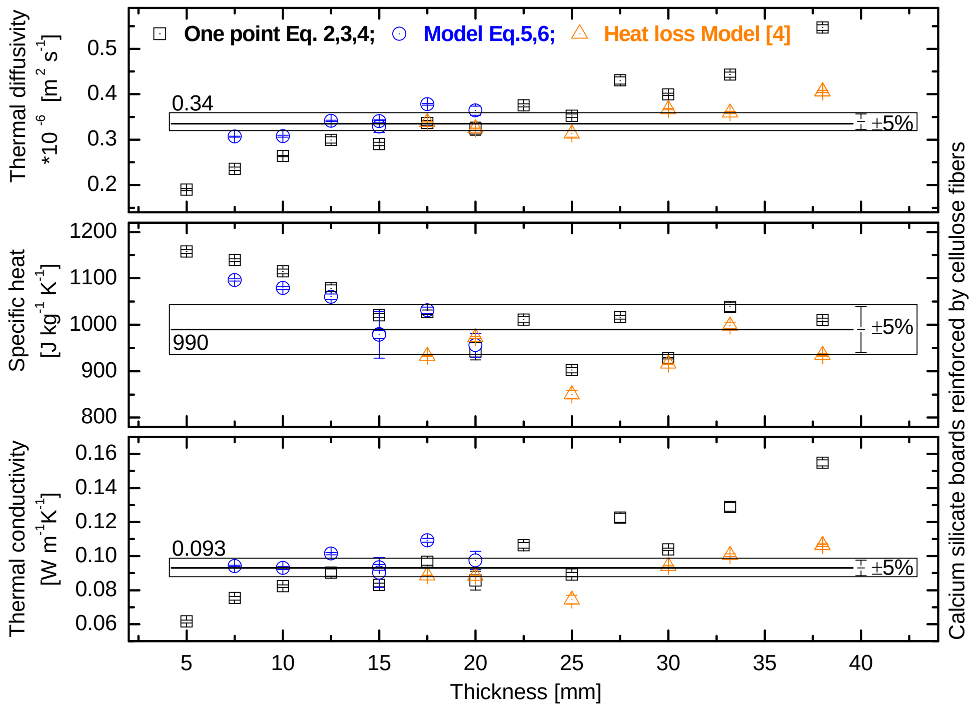


Fig. 2. Thermophysical parameters of calcium silicate boards reinforced by cellulose fibres. Values of one point procedure are compared with real models assuming heat capacity of the source (circles) as well as with model assuming real sample geometry and accounting heat losses from the sample surface [4].

Acknowledgements

The work was supported by Slovak grant agency VEGA under a project No. 2/0161/08 - Study of thermal-insulation properties of low thermal conductive materials.

References

[1] Boháč V.,. Kubičár Ľ,. Vretenár V, *HIGH TEMP-HIGH PRESS*. 2003/2004, Vol. 35/36, pp. 67-74

[2] Boháč V. et all. Methodology of parameter estimation of pulse transient method. Proc. *TEMPMEKO* 2004, 22 - 25 June 2004 Dubrovník, Croatia

[3] Diešková M, Dieška P, Boháč V, Kubičár Ľ. Determination of temperature field and an analysis of influence of certain factors on a temperature fields. In collection of manuscripts of the 17th ECTP, 2005, Bratislava, Slovakia. Vozár L., Medveď I. and Kubičár Ľ. editors. (CD-ROM).

[4] Boháč V, Dieška P, Kubičár Ľ. *MEASUREMENT SCIENCE REVIEW*, Volume 7, Section 3, No. 2, 2007

[5] Kubičár, Ľ. (1990): Pulse method of measuring basic thermophysical parameters. *Comprehensive Analytical Chemistry*, ed. G. Svehla, Elsevier, Amsterdam, Tokyo, Oxford, New York

Monitoring System of the Temperature-Moisture Regime Placed at Spis Castle

[1]Ľ. Kubičár, [1,2]V. Štofanik, [1]V. Boháč, [3]M. Brček, [3]J. Vlčko, [1]C. C. Anibarro

[1] Institute of Physics SAS, Dúbravská cesta 9, 845 11 Bratislava, Slovakia
[2] Department of Radioelectronics, FEI STU, Ilkovičova 3, 812 19 Bratislava, Slovakia
[3] Department of Engineering Geology, Comenius University in Bratislava, Faculty of Natural Sciences, Mlynská dolina G, 842 15 Bratislava, Slovakia
Email: ludovit.kubicar@savba.sk

Abstract. *Present paper deals with the principle of the hot ball method, with the construction of the moisture sensor and its application in areas of nondestructive testing of thermal-moisture regime within the rock mass. The sensor consists of a small cylinder having the diameter and length of 20 mm in which a small ball in diameter up to 2 mm is placed that delivers a heat in step-wise regime and simultaneously measures temperature. The sensor is prepared of the same material as the monitored rock. Spis Castle is placed on the travertine mound and is suffering deep-seated deformations. The paper introduces novel methodology for monitoring temperature and moisture in local position. The four sensors have been positioned in depths of 10, 40, 80 and 150 cm from the rock surface. Monitoring has been performed in the period from December 4th, 2008 up to January 17th, 2009.*

Keywords: hot ball method, thermal conductivity, moisture, rock mass

1. Introduction

Degradation of the stones is strongly influenced by moisture that in combination with temperature, salt and various biofactors, have high correlation with geographical location, microclimate and with the local hydrological conditions. Historic structures are predominantly built on rock massifs deteriorated by post genetic processes (predominantly weathering) and many decorative and ornamental parts of historic structures are constructed of easily cut porous materials. Both in the subgrade and the upper structure, the moisture plays a dominant role in their degradation and deterioration [1]. Therefore information regarding moisture-thermal regime of applied materials represent a basic assumption for a selection of a suitable procedure for preservation of historic monuments.

Spis Castle, a monument that was included in the UNESCO World heritage list (Eastern Slovakia, Hornádska kotlina Basin) is built on a travertine mound (Quaternary age) overlying the Tertiary soft rocks. Lateral spreading caused by the subsidence of strong upper travertines into soft claystone strata fractured and separated the castle rock into several cliffs. The differential movement of individual cliffs is the phenomenon influencing the instability of the monument.

In order to estimate the mode of failure and rate of displacement of rock cliffs several monitoring techniques are adopted:

- Monitoring of displacements carried out by the mechanical-optical crack gauge type TM-71 and the demec gauge (demountable mechanical crack gauge) type SOMET.
- Monitoring of displacements carried out by full automatic crack gauge type GEOKON 4.2.
- Monitoring of temperature-moisture regime within the rock mass body which is a subject of the presented paper.

Thermal conductivity sensor based on the hot ball method has been utilized for monitoring of moisture. This sensor in connection with the RTM 1.01 monitoring system records information on local temperature and thermal conductivity of the surrounding material. As thermal conductivity of a porous structure is a function of the pores distribution and their water content, this parameter can be used for monitoring of the moisture as well. This phenomenon allows the construction of a moisture sensor based on hot ball placed in a porous body. The moisture sensor can be calibrated in dry and water saturated conditions. For construction of moisture sensors, to cover various requirements of practice, a broad range of porous materials with different porosities can be utilized.

Monitoring methodology based on a simple instrument that is composed of a data logger in connection with a microprocessor is discussed. The monitoring system has been working under environmental conditions from December 4th, 2008 to January 17th, 2009.

2. Theory of the thermal conductivity sensor

Fig. 1. Model of the hot ball and its realization (left). Measuring cycle measured at sandstone for hot ball heat output q = 3.5 mW (right).

The principle of the thermal conductivity sensor is based on a model of the hot ball method that assumes a constant heat flux q from the empty sphere of radius r_b into the infinitive medium that starts to be delivered for times $t > 0$ (see Fig. 1 left) [2]. Then the temperature distribution within the medium can be characterized as follows

$$T(r,t) = \frac{q}{4\pi\lambda r^2}\left\{\sqrt{\frac{at}{\pi}}\left(1 - e^{\frac{-r^2}{at}}\right) + r \cdot erfc\left(\frac{r}{\sqrt{at}}\right)\right\} \tag{1}$$

where $erfc(x)$ is the error function defined by $erfc(x) = \frac{2}{\pi}\int_0^x \exp\left(-\zeta^2\right)d\zeta$ and λ and a are thermal conductivity and thermal diffusivity of the surrounding medium, respectively [3].

Function (1) gives a working relation of the measuring method based on the hot ball for long time approximation

$$\lambda = \frac{q}{4\pi r_b T_m(t \to \infty)} \tag{2}$$

where T_m is stabilized value of the temperature response that is reached in the long time limit at the surface of the empty sphere with radius r_b. During stabilization time, the heat produced by the ball penetrates in a material volume from which the information is gained. Penetration depth is a sphere of diameter around 20 mm for stabilized temperature considering sensor diameter of 2 mm. Typical measurement signal is shown in Fig. 1 along with the characteristic points used to calculate thermal conductivity.

The measuring procedure consists of the specimen temperature measurement (base line), switching on the heating and simultaneously scanning the ball temperature. When the ball temperature is stabilized, the heating is interrupted and a period of temperature equilibration follows. When the temperature in the specimen is equilibrated the next measurement can be realized. The highest repetition rate (frequency) of the measurements depends on the thermal conductivity and it takes from 3 up to 40 minutes.

The hot ball is composed of two parts that, on one of it generates constant heat and the second one measure the temperature response (patent pending). The RTM 1.01 instrument in connection with the thermal conductivity sensor was constructed in a manner that can work in the regime of a single measurement or in the monitoring regime. Simplified block diagram of the RTM 1.01 instrument is shown in Fig. 2. With assistance of any PC with standard USB interface, one can set the configuration of the instrument, as well as download measured data stored in the instrument's memory. Currently, additional electronic units that will support wireless data transfers are under construction.

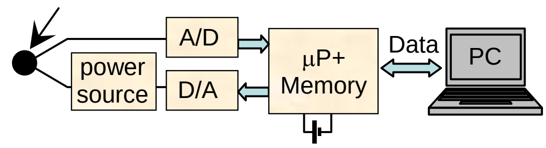


Fig. 2. Simplified block diagram of the RTM 1.01 instrument.

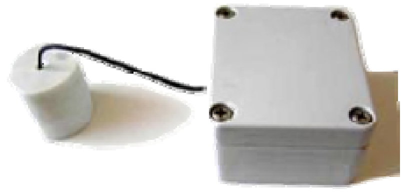


Fig. 3. The RTM 1.01 instrument in connection with the hot ball arranged as a moisture sensor.

3. The moisture sensor construction and in situ application

The true moisture sensor is based on the hot ball inserted into the cylinder made of porous material (Fig. 3). Such moisture sensor is then calibrated for dry and water saturated conditions. Later the sensor is fixed in a place, where one needs to have information on moisture. A choice of the cylinder material porosity allows measurement of a broad range of moisture. Such moisture sensors can be used for highly inhomogeneous wall structure as well as any natural stone or natural rock body.

Porous materials situated in natural conditions are exposed to sun radiation, precipitation, evaporation, frost and thaw phenomena. The mentioned processes form the water phases found in pores (gas, liquid, solid). Then the resulting thermal conductivity of a porous material is a function of the content of the pores.

We have applied the moisture sensors for monitoring the thermal – moisture regime in the rock mass forming a subgrade of Spis Castle (Fig. 4). Four holes were drilled (one for each moisture sensor) in a distance of 2 m and up to depths of 10, 40, 80, and 150 cm from the rock surface in a horizontal direction. The cylinder with 20 mm diameter for the moisture sensor was cut from the drilled core. The sensor configuration together with the drilled hole is shown in Fig. 5. The four sensors were previously calibrated in laboratory for dry and water saturated conditions. A paste made by mixing of water with travertine powder obtained during drilling was used to fix the sensors into the holes. The rest of each hole was filled by small pieces of travertine core combined with travertine paste up to the surface. To prevent diffusion of the water into the borehole, the opening heads were sealed by silicone paste. The four RTM 1.01 instruments have been configured appropriately considering measuring repetition rate, temperature and moisture changes. The in situ monitoring was performed in the period from December 4th, 2008, up to January 17th, 2009 (Fig. 6).

Fig. 4. Spis Castle.

Fig. 5. The moisture sensor and hole configuration.

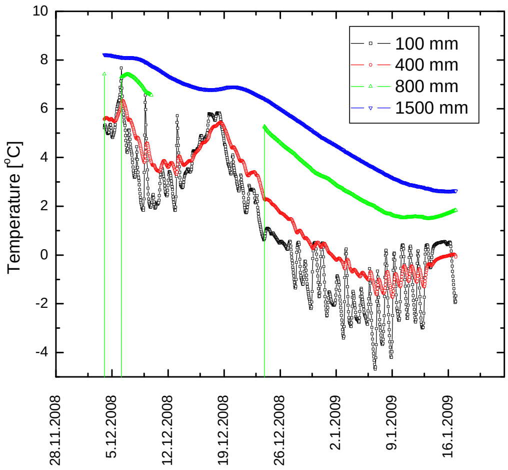


Fig. 6. Temperature variations of monitored rock mass for different depths, during the period of December 4th, 2008 – January 17th, 2009. Parts of data from the depth 40 cm are omitted due to the instrument failure.

4. Conclusions

Data on temperature obtained from all four positions during monitoring period are shown in Fig. 6. Temperature variations due to the day/night period are recognizable for depths of 10 cm and 40 cm, when sky was clear. Temperature at all four positions has fallen down within one month period. The moisture has risen in the period of monitoring for all four positions, however it is not possible to establish a measure of this rise due to need of additional calibration measurements.

Generally the accuracy of measurements is within 10% (compared with pulse transient method and data from literature) and the precision (reproducibility) is better than 1%.

The hot ball sensor belongs to the family of the multi-parametric ones. The sensor in combination with porous structure of the multifunction material can offer a range of information. In our case, the sensor can be used for determination of the moisture content. Processes, like thawing, freezing, drying, etc. can be monitored. It should be stressed that the listed processes are always starting from the material surface. One can monitor the fronts of the propagated effects by fixing the sensors to different material depths.

Acknowledgements

This work was supported by the EU project ACOUTHERM GRD3-2001-60001, by the project APVV 0497-07 and by the VEGA projects 2/0146/08 and 1/0499/08. Authors would like to express thanks to Mr. Markovič for his technical assistance.

References

[1] Vlčko J. Extremely slow slope movements influencing the stability of Spis Castle, UNESCO site. Landslides. - Vol. 1, No. 1 (2004), 67-71.

[2] Kubičár Ľ, Vretenár V, Štofanik V, Boháč V. 2008, Thermophysical sensors: Theory and application of the hot ball, Int. J. of Thermophysics, sent for publication.

[3] Carslaw HS, Jaeger JC. 1956, Conduction of heat in solids, Oxford, Clarendon Press, 348.

Precision Measurement of Cylinder Surface Profile on an Ultra-Precision Machine Tool

[1]J.C. Lee, [1]Y.J. Noh, [1]Y. Arai, [1]W. Gao, [2]C.H. Park

[1]Nano-Metrology and Control Lab, Department of Nanomechanics,
Tohoku University, Aramaki Aza Aoba 6-6-01, Sendai, 980-8579 Japan,
[2]Korea Institute of Machinery and Materials,
Sinseongno Yuseong-Gu, Deajeon, South Korea
Email: leejc@nano.mech.tohoku.ac.jp

Abstract. *This paper describes the measurement of the surface straightness profile of a cylinder workpiece on an ultra-precision machine tool which has a T-base design with a spindle, an X-slide and a Z-slide. The movement range of the X-slide is 220 mm and that of the Z-slide is 150 mm, which have roller bearings in common. Two capacitive sensors are employed to scan a cylinder workpiece mounted on the spindle along the Z-axis. The straightness error motion of the Z-slide is measured to be approximately 100 nm by the reversal method. The straightness profile of the cylinder workpiece is evaluated to be approximately 400 nm by separation of the motion error, simultaneously.*

Keywords: Measurement, Surface Profile, Error Motion, Straightness, Machine Tool, Slide, Reversal Method

1. Introduction

Ultra-precision machine tools are widely used in fabrication of various workpieces such as optical components with three dimensional microstructures. However, the motion error of the ultra-precision machine tool causes profile errors when fabricating the workpiece. It is thus necessary to measure the motion error of the ultra-precision machine tool for evaluation of the machine tool. Usually, a high precision gauge is needed for the measurement of the error motion. On the other hand, it is desired to measure the surface profile of the workpiece on the machine tool from the view point of machining efficiency [1-3]. For this purpose, it is desired to measure the motion error of the machine tool and the surface profile of the workpiece simultaneously [4-6]. In this research, a cylinder workpiece machined on an ultra-precision machine tool is used as the measurement target for simultaneous measurement of the error motion and the surface profile. To separate the two parameters from the sensor output, the reversal technique is employed. By using this method, both the error motion and the surface profile can be measured with a high precision. This paper presents the principle of the reversal method of using high-precision capacitive sensors. Experimental results on an ultra-precision diamond turning machine are also described.

2. Principle of measurement

The reversal method can be applicable for measurement of surface profile and error motion when the machine motion has a high repeatability [7]. The ultra-precision machine tool can satisfy this requirement because the motion of the machine tool slide is repeatable on the order of 10 nm. In this method, it is necessary to rotate the workpiece by 180 degrees and carry out two scans before and after the rotation (reversal). Fig. 1 shows the principle of the measurement by the reversal method. In the first scanning, two displacement sensors are employed to scan the workpiece before the reversal. The sensor outputs can be written as

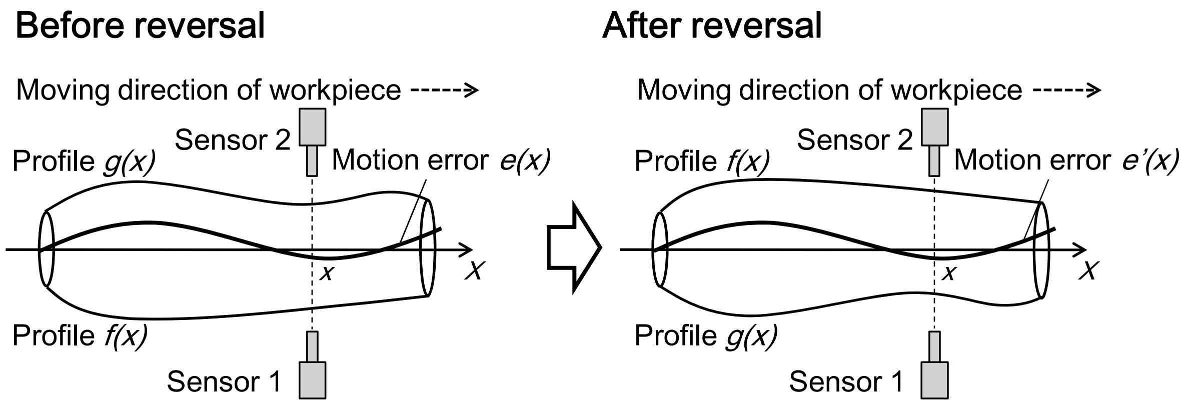


Fig. 1 Principle of the reversal method

$$m_{11}(x) = f(x) + e(x) \tag{1}$$

$$m_{21}(x) = g(x) - e(x) \tag{2}$$

where

- m_{11} sensor output of m_1 before reversal
- m_{21} sensor output of m_2 before reversal
- e(x) motion error of ultra-precision machine tool before reversal
- f(x) profile of the workpiece on the side of sensor 1
- g(x) profile of the workpiece on the side of sensor 2

When the workpiece is moved along the X-axis once again after the reversal, the sensor outputs are changed as follows

$$m_{12}(x) = g(x) + e'(x) \tag{3}$$

$$m_{22}(x) = f(x) - e'(x) \tag{4}$$

where

- m_{12} sensor output of m_1 after reversal
- m_{22} sensor output of m_2 after reversal
- e'(x) motion error of ultra-precision machine tool after reversion

From the above equations, the motion error e(x) and the surface profiles f(x), g(x) can be calculated as follows

$$e(x) = \left(m_{11}(x) - m_{22}(x)\right)/2 + \Delta e(x) \tag{5}$$

$$f(x) = \left(m_{11}(x) + m_{22}(x)\right)/2 + \Delta e(x) \tag{6}$$

$$g(x) = \left(m_{21}(x) + m_{12}(x)\right)/2 + \Delta e(x) \tag{7}$$

where

$$\Delta e(x) = \left(e(x) - e'(x)\right)/2 \tag{8}$$

Eq. 8 can be ignored because the motion of the ultra-precision machine tool has a high repeatability. Therefore, the motion error and profiles can be calculated in Eqs. (5)-(7) without considering Δe(x).

3. Experiments

Fig.2 shows the experimental system for measurement of the Z-slide straightness of an ultra-precision diamond turning machine and the surface straightness of the cylinder workpiece. Two capacitive sensors were employed in the measurement. Each of the sensors has a measurement range of ±50 μm, a resolution of 2 nm and a foot print of 1 mm. The sensor outputs were acquired by synchronizing them with the signal from a linear encoder which is used to the accurate position of the Z-slide. On the other hand, the Z-slide of the machine tool has a 150 mm stroke, a positioning resolution of 10 nm, and a positioning accuracy of ±50 nm. The air spindle has 50 nm rotational accuracies in the radial direction and the axial direction. Fig. 4 shows the results of the stability test, in which the machine tool was kept stationary. The test term was 10 min, and the sampling interval was 10 ms. As can be seen in the figure, the instability was measured to be 50 nm, which includes low-frequency components caused by thermal drift of the mechanical structure of the slide and long-term instability of the air pressure. The high-frequency components of the data came from the vibration of the machine tool. The cylinder workpiece was moved by the Z-slide for straightness measurement. The two capacitive sensors, which were kept stationary on the machine tool, scanned the both sides of the cylinder workpiece. After rotating the workpiece by 180 degrees, the workpiece was moved along the Z-axis once again. The length of measurement was 126 mm, the moving speed was 800 mm/min and the sampling interval was 180 μm. The reversal of the workpiece was carried out by the air-spindle. The straightness error of the Z-slide was measured to be approximately 100 nm. Fig. 6 shows the evaluated profiles of the cylinder workpiece by separation of the motion error. As shown in Fig. 6, the workpiece has a straightness error of approximately 400 nm.

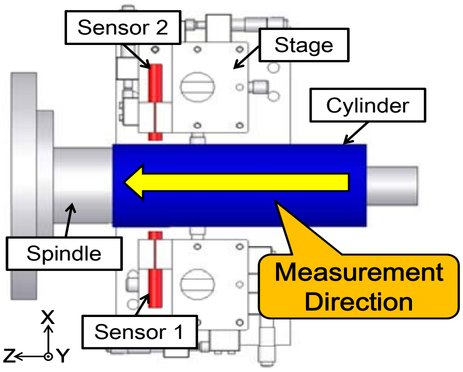


Fig. 2 Experimental system

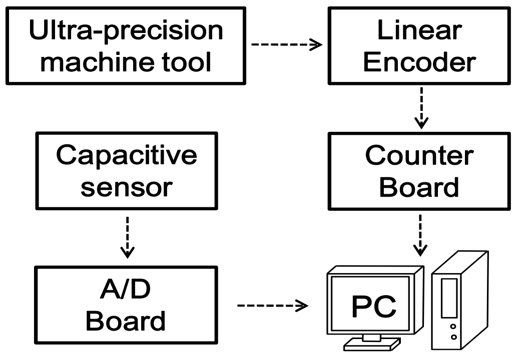


Fig. 3 Diagram of experimental setup

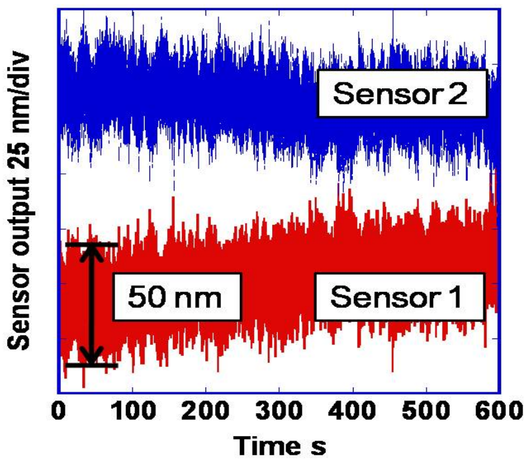


Fig. 4 Stability of sensor

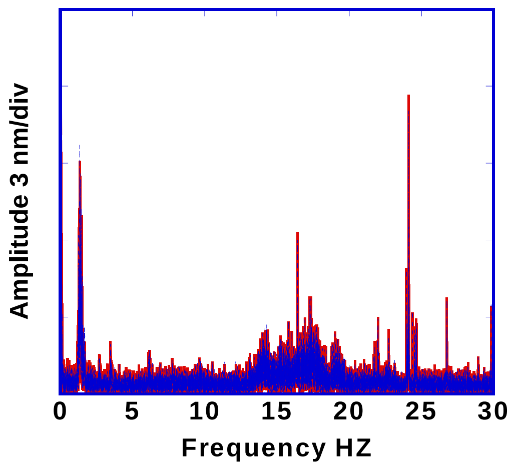


Fig. 5 Spectrum of the stability of sensor

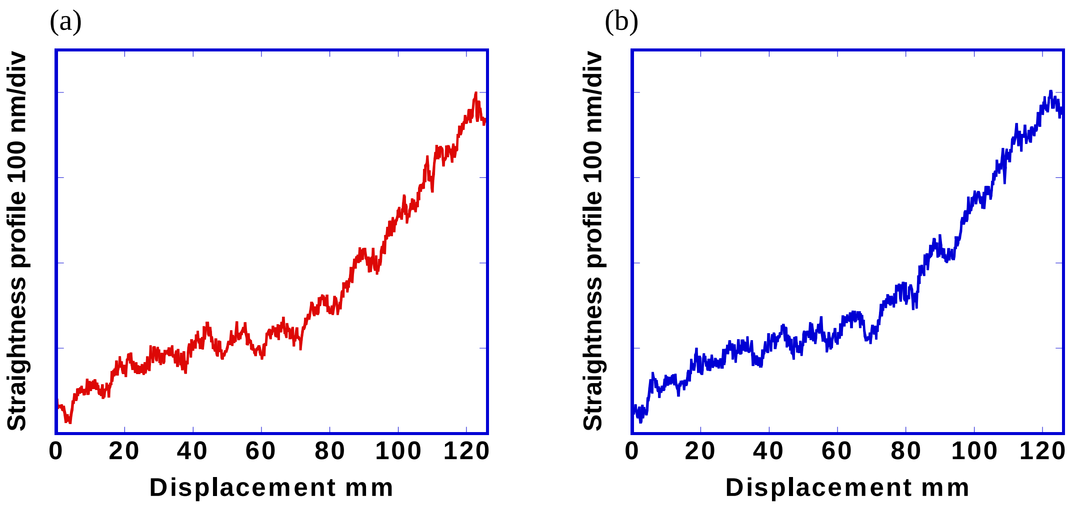


Fig. 6 Evaluated profile of cylinder: (a) profile of f(x) and (b) profile of g(x)

Conclusions

A measurement system that can measure the surface profile of a cylinder workpiece and the motion error of an ultra-precision machine tool at the same time has been established. It was confirmed that the straightness error of the Z-slide of the machine tool was approximately 100nm and the straightness profile of the workpiece was measured to be approximately 400nm.

References

[1]. Tanaka K, et al., Development of an ultra-precision machine tool, *Journal of Japan Society for Abrasive Technology*, 51(9):553-558, 2007.

[2]. Gao W, Precision nanometrology and its appications to precision nanosystems, *International Journal of Precision Engineering and Manufacturing*, 6(4):14-20, 2005.

[3]. Kiyono S, Gao W, Profile measurement of machined surface with a new dirrerential method, *Precision Engineering*, 16(3):212-218, 1994

[4]. Jywe W, Chen Cj, A new 2D error separation technique for performance tests of CNC machine tools, *Precision Engineering*, 31(4):369-375, 2007

[5]. Kono D, Matsubara A, Yamaji I, Fujita T, High-precision machining by measurement and compensation of motion error, *International Journal of Machine Tools and Manufacture*, 48(10):1103-1110, 2008

[6]. Tanaka H, Tozawa K, Sato H, O-hori M, Sekiguchi H, Application of a new straightness measurement method to large machine tool, *Annals of the CIRP*, 30(1):455-459, 1981

[7]. Whitehouse OJ, Some theoretical aspects of error separation techniques in surface metrology, *Journal of Physics E: Science Instruments*, *9*: 531-536, 1976.

Device for Ring Gauges Calibration

[1]J. Bartl, [1]R. Ševčík, [2]R. Fíra,

[1]Institute of Measurement Science, Slovak Academy of Sciences,
Dúbravská cesta 9, 841 04 Bratislava, Slovakia,
[2]Slovak Institute of Metrology, Karloveská 63, 842 55 Bratislava, Slovakia
E-mail: umerbart@savba.sk

Abstract. *A diameter measurement of the precise cylinder bores is the specific problem of the length metrology. In industrial practice this operation is being realised by means of three-pointed bore gauges - inside micrometers. The ring gauges - setting rings are used for testing of the inside micrometers. Ring gauges are manufactured from the wear resistant steel or zirconium ceramic with tolerance diameter of 1 μm and form tolerances according to DIN 2250. Since the setting rings of highest accuracy are the reference standards of the 2nd order, their precise calibration is of special interest. This article describes the optical design of the device for the ring gauges calibration.*

Keywords: Traceability, Setting rings, Interferometric resolution

1. Introduction

In the Slovak Republic, the length unit metre is realised by the radiation of the National Standard of Length (NSL No. 002/1997 SMU-1), HeNe/I_2 laser with stabilised optical frequency 473 612 353 607.9 kHz, corresponding to the vacuum wavelength of λ_o = 632 991 186.81 fm. Within the international comparison (BIPM.L-K11, 2005 – femtosecond comb measurements of HeNe/I_2 lasers), its frequency/vacuum wavelength was determined with the expanded uncertainty of 4,8 kHz or 0.063 fm respectively. This value was confirmed by the measurements at CMI one year later, using the fs comb again.

According to the traceability scheme [1], the length unit transfer from the NSL to the length of gauge blocks is realized by the Standard Interferometer Comparator (SIC), using the dynamic interference method, i.e. the counting of standard wavelength fractions. The standard gauge blocks then serve for the direct comparison for the tested gauge blocks of lower orders or for the calibration of the ring gauges up to 0.2 m on the interference device (IRG). It is well known that each step of the unit transfer results in the increase of the final measurement uncertainty. Therefore when IRG is the standard device of the 1st order, even in the ideal case the setting rings of the highest accuracy can be no better than of the 2nd order. In accordance with the MRA (Mutual Recognition Arrangement) among National Metrological Institutes (NMIs) and BIPM Sèvres, the proof of traceability of the National standards with declared uncertainties is strictly required. The device described in this paper is directly traceable to the NSL via frequency stabilized HeNe laser 633 nm, without intermediate stage of calibrated gauge blocks.

2. Laser interferometer

Special meachanical and optical instruments were developed for the calibration of the ring gauges many years ago [2]. One of them was the *Universal Komparator 200 mit Perflectometer Leitz Wetzlar*. Our design of laser interferometer is based on the principle of the mentioned Leitz Komparator-Perflektometer. The diameter of ring gauges can be observed from the shift of line scale by the measuring microscope of the Leitz comparator [3].

The measuring table carries the ring gauge. The line scale is fixed on the movable table in the direction of the vertical guide axis. The perflectometric optical part of instrument is serving for the localisation of the functional surfaces of the calibrated measure - ring gauges [4]. The perflectometric principle is based on the projection of cross wire plate to the gauge functional surface. After the reflection from the gauge surface, the reflected picture is projected to the reference cross wire plate. The coincidence of both cross wires indicates the localised position of the gauge functional surface. In such a way, we can localize the initial and final points of the vertical displacement of the table in line scale and thus determine the diameter of gauge rings.

According to our proposal, the device is directly traceable to the NSL. It consists of two optical parts. One part is represented by the laser interferometric measuring system LOS Limtek, Blansko, CZ. LOS serves for the measurement of the table displacement with resolution of 1.25 nm (□/512). The frequency of the HeNe laser from system LOS has been calibrated at Swiss NMI (National Metrological Institute) METAS with expanded uncertainty of U = 2.10-8 (k = 2). It means that the length of table displacement (i.e. the reading of a ring gauge diameter) is traced to the length standard of Swiss NMI. It would seem that the laser interferometry with resolution of about a few nanometres is able to solve a majority of problems in length metrology. However, a general problem in such application of the laser interferometry in an engineering industry lies in the insufficient localisation of the mentioned initial and final points. This problem is solved by the second part of our device, being a laser system localizing the initial and final position of the ring gauge edges during the vertical shift of the table. A simplified optical scheme of this part is on the Fig. 1A.

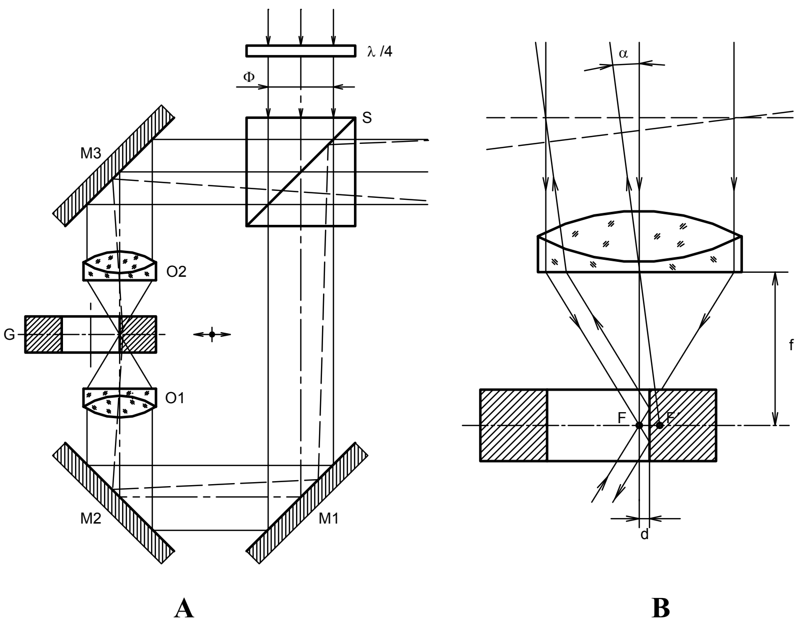


Fig.1

The expanded linear polarised laser beam passes through the $\lambda/4$ plate, transforming the linear polarisation to the circular one. By means of the non-polarising beam splitter S the beam is divided into two directions. One beam passes to the mirrors M1, M2, through objectives O1, O2, to the mirror M3 and again through the splitter S to the output. The second one is reflected by the splitter S to the mirror M3, passes through objectives O2 and O1 to the mirrors M2, M1 and then is reflected by the splitter S to the output. The path travelled by both beams is of the same optical length and therefore at the output no phase shift occurs between them, provided the gauge ring is not located in their path. In the case when the wave fronts of beams are strictly plane, both halves (upper and down) of the output viewing field have the same illumination. If the edge of gauge coincides with the optical axis from the right side, then left side of the beam passes down from the mirror M3 and right side of beam having been reflected from the mirror M2 is shaded by the gauge. The non-shaded parts of beams do not interfere, because the beams after their reflection from bright face of gauge are passing independently from upper and bottom halves of the output. If the asymmetrical beam splitter is used (e.g. ratio R/T = 30/60), the upper half of viewing field will be starker as the bottom half. It makes possible to distinguish that position when the gauge coincides with optical axis from the left or right side.

If the inside edge of the ring gauge is located closely to the optical axis of objectives at the distance d (Fig. 1B), then FF' = 2d and the inclination angles of output beams with respect to optical axis are $+\alpha$ and $-\alpha$ respectively. The angle between output wave planes W and W' is 2α (Fig. 2) and the Fizeau interference fringes (fringes of equal thickness) can be observed in the output visual field. Interference minimum corresponds to the odd number of half waves at the path difference Δ:

$$\Delta = (2k+1)\frac{\lambda}{2}, \text{ where } k = 1, 2, 3, \ldots \tag{1}$$

If the k +1 dark fringes in the visual field of a diameter Φ are observed, then for the angle 2α between both wave fronts it follows:

$$tan\, 2\alpha = \frac{(2k+1)}{\Phi} \cdot \frac{\lambda}{2} \tag{2}$$

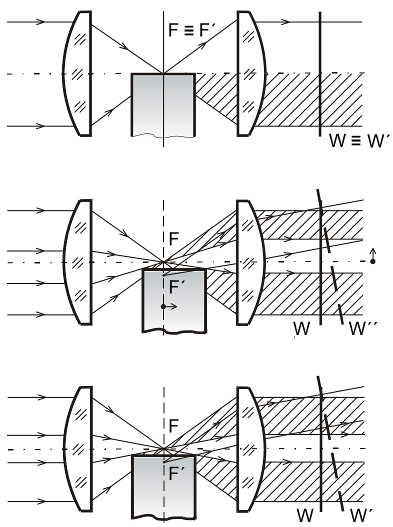


Fig. 2 Principle of a creation the interference fringes

The distance between two neighbouring fringes (interference minima) x is:

$$x = \frac{\lambda}{tan\, 2\alpha} \tag{3}$$

From the Fig.1B follows that $tan2\alpha = 2d/f$, where f is the focal length of objectives.
For small angles, $tan2\alpha = 2\ tan\alpha$. Therefore at the distance d of the gauge from the optical axis, the distance of neighbouring minima is:

$$x = \frac{\lambda . f}{4d} \tag{4}$$

If the number k + 1 fringes is observed at the output, then the distance d is given by the expression:

$$d = \frac{f.tan\,\alpha}{2} = \frac{(2k+1)}{8\Phi} \tag{5}$$

The resolution limit d_{min} is determined by the condition of k = 0, i.e. when the first interference minimum appears at the output.

$$d_{min} = \frac{\lambda . f}{8\Phi} \tag{6}$$

The output image can be scanned by the B/W CCD Camera. For example, the camera ORCA II-ER Hamamatsu has the following parameters:

Active area of CCD array is 8.67 mm (H) x 6.6 mm (V), (1344x1024 active pixels), high sensitivity, objective Canon T01-J624-000 f = 16 mm 1:1.6 (choice of lens system depends on the size of Φ), Peltier cooling, external control RS 232C, output signal RS 422A, pixel clock rate 10 MHz/ pixel, square pixel structure 10x10 μm.

At the wavelength of 633 nm we obtain the resolution limit d_{min} for objectives O1, O2 :

- Apochromat Meopta 10x0.30, f = 15.65 mm, free working distance 10.58 mm, Φ = 10 mm, d_{min} = 0.126 μm. The resolution of the table shift measurement by the system LOS is 1.25 nm.

For the comparison, the parameters of Perflectometer Leitz were:

- The uncertainty of the gauge edge localisation is 0.2 μm. The resolution of the table shift reading on the line scale by the measuring microscope is 0.3 μm.

3. Conclusion

In our arrangement the microscope objectives 10x0.30 have been used. The better resolution can be achieved by using of microscope objectives having both larger magnification and aperture (e.g. 20x0.45, 30x0.65, or 45x0.65). However, in practice it is not possible since it is limited by the height of calibrated ring gauges, well exceeding the working distance of the listed objectives (below 1 mm vs. the height of the ring gauges being up to 8 mm). Other way can be involving of the photographic camera objectives. The most frequently used objectives have the focal length of ~50-60 mm and diameter of visual field $\Phi \sim$ 42-50 mm. It corresponds to the aperture of 1:1.2. For example:

- Photographic objective Nikkor S, Nikon: f = 55 mm, free working distance 47.2 mm, input diaphragm Φ = 46.2 mm, d_{min} = 0.094 μm.
- Camera objective Tevidon, Zeiss: f = 25 mm, free working distance 12.3 mm, input diaphragm Φ = 18 mm, d_{min} = 0.109 μm.

Acknowledgements

The authors are grateful to the Slovak Grant Agency VEGA for the financial support of the research project No. 02/7081/27.

References

[1] Chudý, V.- Palenčar, R.- Kureková, E.- Halaj, M.: *Measurement of the technical quantities*. Bratislava: Slovak University of Technology 1999, pp. 612-619, ISBN 80-227-1275-2.(In Slovak).

[2] Brezina, I.: New trends in instrumental technique for the opening measurement. Fine mechanics and optics, **24**, 1979, No.4, pp.111-114 (In Slovak).

[3] Leinweber, P.: *Taschenbuch der Längenmeßtechnik*. Berlin, Gottingen, Heidelberg: Springer Verlag, 1954, pp.346-350.

[4] Ernst Leitz, GmbH, Wetzlar: Komparator 200 mit Perflektometer, 1982. 9p.

Rapid Measurement of Involute Profiles for Scroll Compressors

J.H. Yang, Y. Arai, W. Gao

Nano-metrology and Control Laboratory, Department of Nanomechanics, Tohoku University. Aramaki Aza Aoba 6-6-01, Aoba-ku, Sendai, 980-8579, JAPAN
E-mail: Jianhong@nano.mech.tohoku.ac.jp

Abstract. *Scroll compressors are widely used in air conditioners, vacuum pumps and so on. rapid measurement of Flank profile of a scroll compressor is important to improve the compression efficiency and decrease noises. A contact probe made of ruby was used for measurement of flank profile. The probe is moved by a linear slide along the X axis at a constant speed. The scroll workpiece was fixed on a precision rotary stage. The relationship between the stage rotational speed and the X axis moving speed complies with the Archimedean curve. The measurement data of the rapid measurement system were analyzed and measurement errors were removed by compensation of the offset between the coordinates of the rotary stage center and those of workpiece center. The measurement results were compared with those measured by a commercial coordinate measuring machine (CMM). The measurement time for the involute profile of the scroll is shortened to 153 s by the developed rapid measurement system from the 10 minutes measurement time by the CMM while the measurement accuracy is kept the same.*

Keywords: Scroll profile, Measurement, Contact probe, Rapid measurement, Scroll compressor

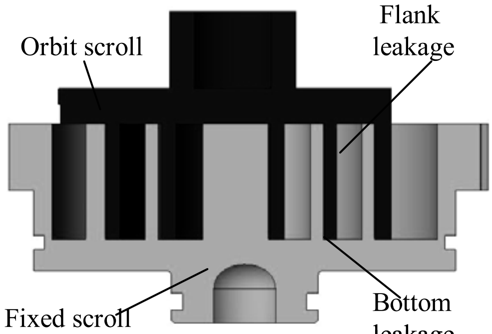
1. Introduction

Scroll compressors compress air by orbiting motion of scrolls. The air with a high pressure is taken out from a discharge opening by the orbiting scroll. The scroll compressor has a lot of advantages, including small variations of torque, low vibrations and noises. The high efficiency can also be made because there is no direct fluid path between suction and discharge opening [1, 2].

In order to further improve the efficiency of the scroll compressor, it is very important to reduce the leakages. Fig. 1(a) shows mainly two kinds of leakages. Fig1. (b) shows flank profile measurement including the inside involute profile, the outside involute profile and the non-involute profile. One of the leakages is the flank leakage caused by a gap between the flanks of the two scroll blades. The other is the tip leakage caused by a gap between the end plate and the scroll blade of the scrolls. These leakages can be decreased by increasing the manufacturing accuracy. Rapid measurements of the height and the flank profile are important to decrease manufacturing errors. Conventionally, scroll flank profiles are measured by a coordinate measuring machine (CMM), which is very time-consuming and expensive. The measurement time of the CMM for scroll profiles cannot meet the on-line machining measurement requirement [3]. The paper develops a rapid and accuracy profile measurement system for inside and outside involute scroll profiles. Measurement errors are analyzed by simulations. Measurement results of the rapid measurement system are compared with those of the CMM. It is verified that developed rapid measurement system can satisfy the required measurement accuracy (±3 μm) and measurement time (300s/per workpiece).

2. Measurement System and Measurement Method

Fig. 2 shows the platform of the developed rapid measurement system for the fixed scroll. The measurement system consists of X-Z-θ stages and a contact type scanning probe. Because there is very limited room for the on-line machining measurement of the scroll, the rapid measurement system was developed based on a roundness measurement system. The size of the measurement system is very small. The fixed scroll was fixed on a rotary stage by two taper pins. The rotational angle resolution of stage is 0.0025 degree. The X axis could be moved by a precision control board. The moving speed of the X axis and the rotational speed of the rotary stage were controlled by a PID controller [4]. The positioning error was small enough to be ignored. The position of each stage was measured by each encoder and taken into a personal computer via multi axis control board.



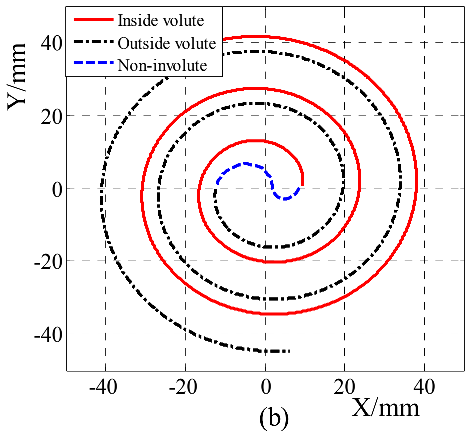


Fig.1.Leakages of scroll compressor and measurement profile

Taking into consideration the influence of cutting oil and chips, a contact probe made of ruby was employed. The probe fixed on the end of the X axis is used for scanning the flank involute profile mounted on the rotary stage [5]. A ruby ball with a 5 mm diameter was attached on the end of the probe axis. The scanning probe ball had three scales in the X, Y and Z directions. The outputs of X and Y direction were used for measuring the involute profiles and the output of Z direction was used for determining friction of the Z axis. The voltage outputs of the probe were taken into a personal computer via an A/D converter. The resolution and the measurement range of the probe were 0.1 µm and ±1 mm respectively. The measurement force of the probe was about 0.12 N. According to the required measurement time, the rotational speed of stage was set to 20 degrees/s. The increment of scroll radius can be described by the following equation.

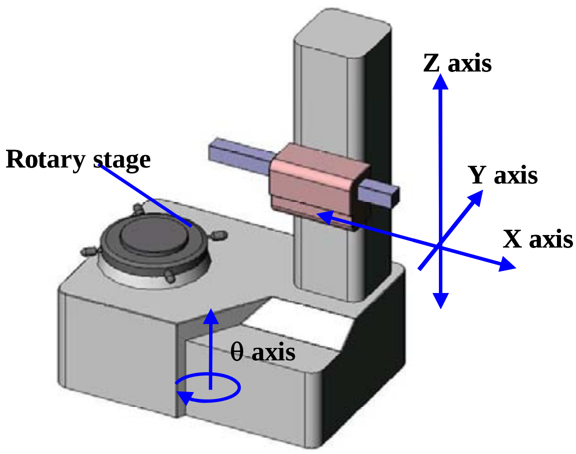


Fig.2. Measurement system of scroll profile

$$r_{\theta 1} - r_{\theta 2} = a \times (\theta 2 - \theta 1) \times \pi / 180 \qquad (1)$$

where

r_θ the scroll radius of the polar angle of measurement point

θ the polar angle of every measurement point

a the base circle radius of scroll

From the equation (1), the X axis moving speed can be set to 0.7923 mm/s. The measurement polar radius of the scroll profile can be got by the outputs of probe encoder and the outputs of the X axis encoder. Measurement errors of the profile can be described by the following equation (2).

$$r_{error} = r_{th} - r_{mea} \tag{2}$$

where

r_{error} measurement error of involute profile

r_{th} theoretical polar radius of scroll profile

r_{mea} measurement polar radius of scroll profile

The same scroll sample was also measured for comparison by a commercial CMM, which was installed in a temperature controlled metrology room. The probing accuracy of CMM was 0.6 μm.

3. Measurement Result and Error Analysis

The involute profile errors of the fixed scroll were measured by the developed rapid measurement system. Measurement results are showed in the Fig. 3. the measurement error range of the outside profile was about ±20 μm. the measurement error range of inside profile was about ±40 μm. It can be seen that there were large measurement errors in the rapid measurement system. The measurement accuracy couldn't meet the required measurement accuracy (±3 μm). There was an offset between the coordinates of workpiece centre and those of rotary stage centre. The offset had significant effect on the measurement results. The Y-directional offset of contact point △y was caused by the friction between the contact probe and the flank profile.. △y had significant influence on the measurement results too.

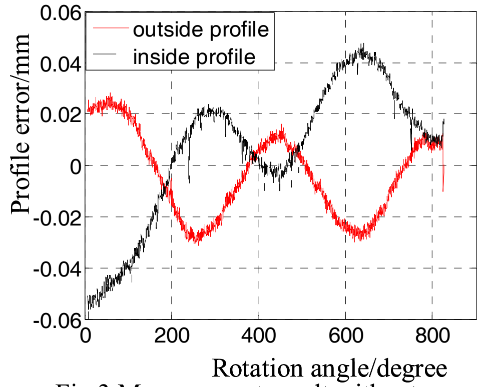


Fig.3.Measurement result without error separation

Simulations were carried out for analysis of the influence of the offset. Fig. 4 (a) shows the influence of the Y-directional offset of the contact point of the probe. Profile errors were calculated when △y=0.5 mm and 1 mm. It can be seen that the △y has large influence on the profile results. Fig. 4 (b) shows the influence of the offset between coordinates of workpiece centre and those of rotary stage centre. Here, xn and yn are defined by the offset between coordinates of workpiece centre and those of rotary stage centre. As can be seen in the figure, xn and yn generate a very large periodic measurement error even if the offset of coordinates is very small. To achieve the required measurement accuracy, it is confirmed that the Y-directional offset of contact point and the offset of coordinates between workpiece centre and rotary stage centre must be separated from the profile measurement result.

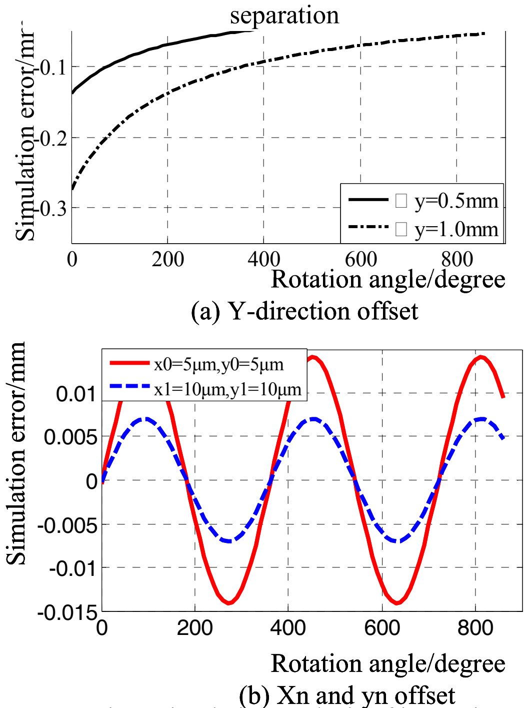


Fig.4.Simulation analysis of impacting measurement result

4. Measurement Result Compared with CMM

The Y-directional offset of the contact point can be measured by the output of the probe encoder. The actual scroll polar angle of the contact point can be calculated by compensation of the Y-directional offset. The offset between coordinates of workpiece centre and rotary stage centre can be got by an optimization algorithm of the measurement error shown in Fig. 3.

The outside and inside profile errors of the fixed scroll were also measured by a commercial CMM. About 1560 measurement points were got from the inside and the outside profiles respectively. The measurement time of the inside and outside profiles was about 20 minutes. About 4090 measurement points were got from the inside and outside profiles respectively by the rapid measurement system. The measurement time of the inside and outside profiles was about 150 seconds. The measurement results of the outside and inside profiles are showed by Fig. 5. As can be seen in the figure, the measurement results of the developed rapid measurement system were the same as those of CMM after removing the Y-directional offset and the offset of coordinates. But the measurement time of the rapid measurement system was much shorter than that of the CMM.. The rapid measurement system can satisfy the on-line machining measurement demands of the scroll.

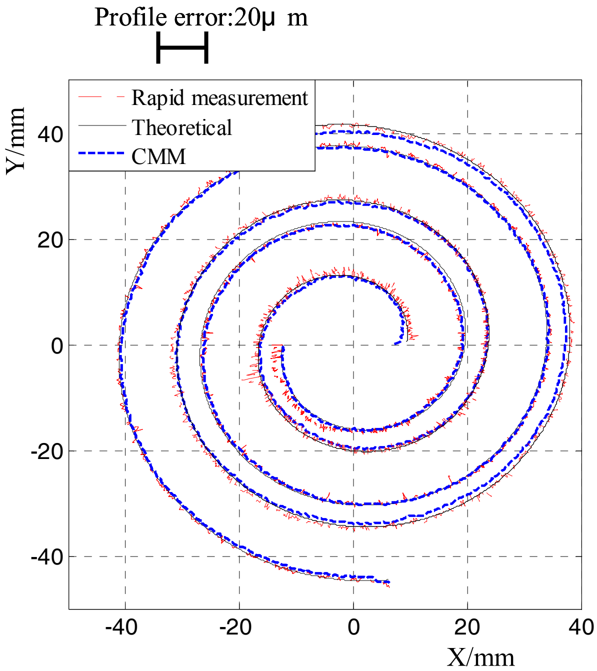


Fig.5.Measurement results of inside and outside profiles

5. Conclusions

A rapid measurement system for scroll compressors has been developed based on a precision three-coordinate probe. The measurement errors of two kinds of offsets were analyzed by simulations. The measurement results of the rapid measurement system were the same as those of the CMM. But the measurement time of rapid measurement system was much shorter than that of the CMM. The developed rapid measurement system can meet the on-line machining measurement requirement of scroll compressors. The rapid measurement of non-involute profile is the future work.

References

[1] E.Morishita. (1993). A basic technology of scroll compressor. Japan Industrial Publishing CO., LTD, Turbo machinery 21-12, 720-728.

[2] C. Yu. (2002). Mathematical modeling of scroll compressors part1. compression process modeling. International Journal of Refrigeration 25, 731-750.

[3] A. Inada, W. Gao. (2007). A high-speed profile measurement system for scroll compressors, Proc. of the 4th International conference on Leading Edge Manufacturing in 21st Century, 219-222.

[4] Wei Gao, Yoshikazu Arai, Atsushi Shibuya, Satoshi Kiyono, Chun Hong Park. (2006). Measurement of multi-degree-of-freedom error motions of a precision linear air-bearing stage. Precision Engineering, 30-1: 96-103.

[5] Gao Wei, Dejima Shuichi, Yanai Hiroaki, Katakura Kei, Kiyono Satoshi and Tomita Yoshiyuki. (2004). A Surface Motor-Driven Planar Motion Stage Integrated with an XYZ Surface Encoder for Precision Positioning. Precision Engineering 28(3): 329-337.

Compact Vibration Measuring System for In-Vehicle Applications

[1]G. J. Stein, [1]R. Chmúrny, [2]V. Rosík

[1]Institute of Materials and Machine Mechanics, Slovak Academy of Sciences,
Račianska 75, SK - 831 02 Bratislava, Slovak Republic
Email: stein@savba.sk

[2]Institute of Measurement, Slovak Academy of Sciences,
Dúbravská cesta 9, SK - 841 04 Bratislava, Slovak Republic
Email: umerrosi@savba.sk

Abstract. *Low frequency vibration occurs especially in ground transportation. It is of permanent interest in studies of driver's and passenger comfort issues and seating dynamics research. Piezoelectric accelerometers commonly used for vibration measurement are not suitable, hence other sensors capable of measuring accelerations down to sub-hertz region have to be used. Based on some previous experience with MEMS acceleration sensors a compact measuring system employing two three-axial MEMS accelerometers interfaced via a data acquisition unit to a light-weight notebook was designed and constructed. The digitised data were processed by scripts by Matlab® with the aim to analyse both the vibration influence on seated person and the dynamic properties of the seat. Some preliminary results from illustrative test runs with a passenger automobile are presented.*

Keywords*: vibration measurement, MEMS accelerometer, vibration measuring system, seating dynamics, human comfort*

1. Introduction

Low frequency vibrations occur, among others, in ground transportation; either as a cause in adjacent environment or within the vehicle itself. The frequency content may vary from well below 1 Hz up to, say, 20 Hz. Exposure to low frequency vibrations and shocks may affect comfort and in case of prolonged exposition, extending for many years, may impair health. Hence it is worthwhile to undertake technical efforts to minimize generation and/or propagation of low frequency vibrations. The comfort issues are of importance in design, development and marketing of vehicles, earth moving machinery and agricultural tractors.

To be able to assess the effectiveness of vibration attenuation by the seat during various field conditions a specific measuring apparatus is required. Various means are used, i.e. in-vehicle analogue data loggers, telemetry systems, instrumentation carried in an escort vehicle, etc.

In most cases the vibration transmission from the vehicle floor (seat base) to the seated driver via the driver's seat and the driver's buttocks is of interest. In general translatory and rotational vibration components are present; however, mostly the translatory components in mutually perpendicular coordinate axis are measured and evaluated. To facilitate their measurement the apparatus, described below, was developed in collaboration with the Institute of Measurement of the Slovak Academy of Sciences.

2. Fundamentals of vibration measurement in transport means

First of all it should be noted that in ground transportation the acting forces are of primary interest, followed by various vehicle parts relative movement. As a result the preferred measurand is the absolute acceleration followed by relative displacement measurement. Despite a well-defined physical relation between both characteristics experience shows (e.g. [1, 4, 6]) that the errors introduced by real measuring systems are of such magnitude that these exclude real application of this relation. In practical situations it is advisable to complement one measuring system by another one and use data fusion [1, 2].

From experience it is known that the *vibratory acceleration* a_v observed in ground transportation is smaller than the standard gravity acceleration $g_N \approx 9.81$ m.s^{-2}. It is also well known that the measured accelerations are of random nature with periodic components. The vehicles are in general subjected to traverse on arbitrary curvilinear trajectory with non-constant traversing velocity. Moreover the vehicle chassis plane may be inclined in respect to the horizontal plane. Both effects give rise to quasi-static acceleration, which is superimposed onto the mechanical vibrations due to road/track undulations and/or engine influence [3, 6].

The quasi-static *translatory acceleration* a_t frequency content extends down to sub Hertz frequencies. So measurement of the *total acceleration* $a_T = a_t + a_v$ poses some practical problems. The use of standard piezoelectric accelerometers with their well-known inherent low frequency limits [1, 2] is not feasible. Other accelerometers types have to be considered, capable measuring also quasi-static acceleration [1, 2, 5, 6]. The current approach is to employ sensors based on the so-called servo-accelerometer principle made using contemporary semiconductor manufacturing technologies which are called Micro-Electro-Mechanical Systems (MEMS) [1, 2, 5]. Further on the use of one type of MEMS type servo-accelerometer for this purpose will be illustrated. This is a continuation and extension of previous work on application of servo-accelerometers for measuring of mechanical and vibration quantities [4, 6].

3. Measurement system description

The compact vibration measuring system consists of following:

A/ Two identical three-axial MEMS accelerometers;

B/ Analogue acquisition unit with USB output/power supply;

C/ Standard notebook with acquisition program.

A/ Each of the used three-component MEMS accelerometer type CXL04LP3, made by Crossbow, San Jose, California USA is enclosed in a plastic box 19×47.6×25.4 mm of mass 46 g and delivered with a 244 cm long cable with a connector on the other end. The measuring range is ± 40 m.s^{-2} ($4 \times g_N$); output voltage is approx. 2.5 V for zero acceleration and sensitivity is ≈ 50 mV/m.s^{-2}. The accelerometers are factory calibrated. Accelerometer is fixed either to a 2 mm thick steel plate 120×100 mm (the base/floor sensor, denoted as "sensor A" (index "b")) or on a steel disk of 70 mm dia., located in a rubber disc "sensor B" (index "s"). The rubber disc is made according to the requirements of ISO 10326/EN 30326 standard, pertinent to laboratory tests of professional drivers' seats. This disc is located between the seated driver and the seat cushion. As described in the said standard, it is used to measure the vibratory input to the seated person. For comparable results the standard requirements was adhered to.

B/ The analogue acquisition unit with USB 2.0 communication interface consists of 8 input channels, each equipped with 24 bit sigma-delta convertors with sampling frequency of up to 50 kHz. The 8 channels are timed simultaneously, thus facilitating reliable phase/time delay estimation. The control is facilitated by a microcontroller. The unit draws some 250 mA from the 5 Volt power supply of the USB port. The acquisition software enables to sample either 32 thousand data points into a file or continuously much larger amount onto disk. For chosen sampling frequency of 200 Hz the first method facilitates measurement duration of 150 s, which suffices for foreseen application. Analogue anti-aliasing filtration is not implemented.

C/ Compact integration of the above units was facilitated. Because the sensors rely on constant voltage supply of required stability a stabilisation circuit of type ADP 3607 was used. No temperature compensation as in [6] was used. To increase system reliability no connectors were used. All connections were thoroughly fixed to each other to avoid any failures during field measurements. The electronics is thoroughly fixed in a rugged box of size 260 × 120 × 40 mm. The sensors' cables and the USB cable are entering the box via rubber bushings, so facilitating partial dust-proofness and compactness: see Fig. 1.

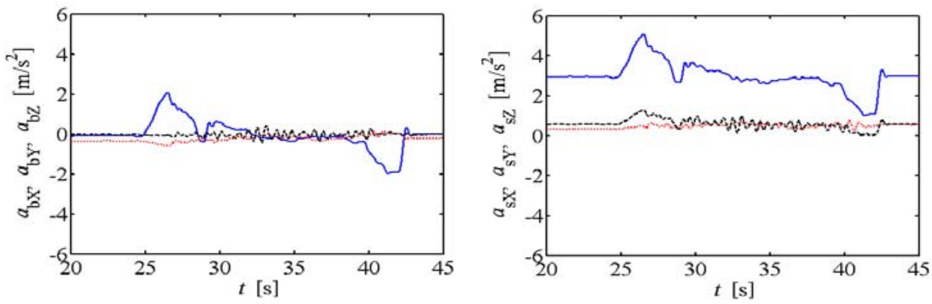
(a) (b)

Fig. 1. Photo of the sensors (a) and the box with the notebook (b).

To make the experimental conditions more transparent a straight traverse, including some original vehicle inclination, is assumed. This condition calls for a specific field measurement conduct: the vehicle stands for a while in standstill then accelerates to required velocity, maintains a straight course at this more-or-less constant velocity and then decelerates to standstill. This measurement organisation facilitates proper data analysis.

4. Measured signal processing in the laboratory

The data analysis program is written in the Matlab® environment. As indicated above the program first reads in raw data, containing a section on the beginning when the vehicle is in standstill. The signal is subjected to low-pass filtration by a FIR LP filter of 100th order with cut-off frequency 0.5 Hz. In this way the translatory acceleration a_t is obtained and displayed (Fig. 2 a, b). The program calculates the sensor inclination angles (in the illustrative example more pronounced for the sensor B on the seat; in this case some 14°, whereas the other ones are about 2° and hence negligible) and subtracts these, essentially DC components, from the time sequence to correct for sensor inclination. From Fig. 2 the operator can also assess when the vehicle was travelling with constant speed (a_t in the x-direction is approx. zero).



(a) (b)

Fig. 2. Pre-processing of acceleration signals for extraction of the translatory acceleration component a_t on seat base (a) and on seat surface (b): (——) fore-and-aft direction (x - direction); (• • • •) transversal direction y - direction); (— — —) upward direction (z - direction).

Then the time interval, wherein the signal is assumed to be stationary, is selected, in this illustrative example between 32nd s and 38th s. Signals from the selected time interval are further subjected to band-pass filtration by a FIR filter of 100th order to extract the vibratory acceleration a_v in the frequency band 0.5 Hz to 80 Hz in each of the three perpendicular sensors axes (Fig. 3). So pre-processed accelerations data are further subjected to calculation of power spectral densities (PSD), transfer function estimates (TFE) and respective coherence functions γ, root-mean-square values, etc., by Matlab® functions, which is a standard data analysis approach used in analysis of mechanical vibrations. Relative uncertainty of the measured acceleration is of the order of ± 5 % [7], as common in seating dynamics research.

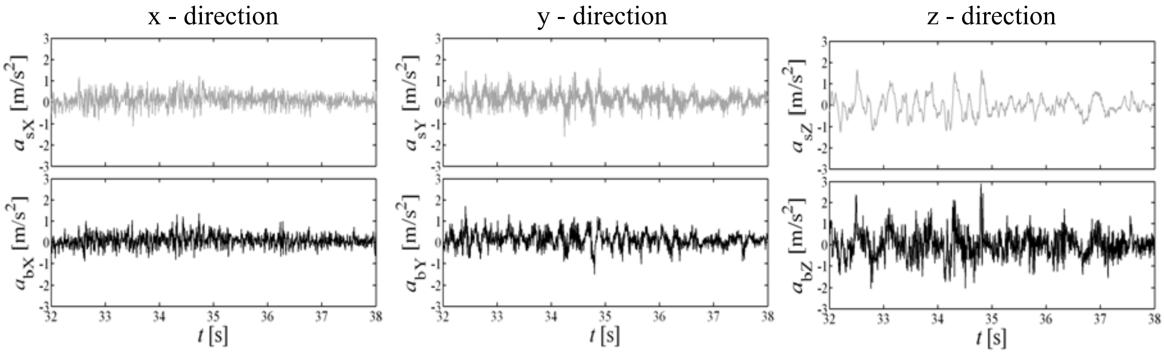


Fig. 3. Vibratory acceleration signals in mutually perpendicular axes: seat base (▬▬), seat (▬▬).

5. Results

The resultant PDFs and TFEs and coherence functions γ (to be shown in the presentation) are used by competent specialists for assessing seat dynamic properties and attenuation of vibration transmitted to the driver. This is the main purpose of the compact vibration measuring system for in-vehicle application. The illustrative results were gathered in test runs with voluntary test persons in a ŠKODA Fabia 1.2 HTP passenger car in academy premises.

Acknowledgement

The paper was prepared within project No 2/6161/27 of the VEGA Grant Agency for Science.

References

[1] Doebelin E. O.: Measurement Systems Application and Design. 5th Ed., McGraw Hill, New York, USA, 2003.

[2] Ďado S., Kreidl M.: Senzory a měříci obvody. ČVUT, Prague, Czech Republic, 1996.

[3] Anon.: Acceleration measurement, In Wiley Encyclopedia of Electrical and Electronic Engineering, Vol. 1 (2001). Webster JG, Editor, John Wiley & Sons, Inc., New York, USA, 2001.

[4] Stein G. J., Šperka M.: Position measurement by micro-electronic accelerometer. In Proceedings of IMEKO 2000, Vienna, Austria, ASMA, 2000, 385-388.

[5] Stein G. J.: Some recent developments in acceleration sensors. *Measurement Science Review*, 2001, 1, 183-186.

[6] Stein G. J., Chmúrny R., Rosík V.: Measurement and analysis of low frequency vibration. *Measurement Science Review*, 2007, 4, 47-50.

[7] Stein G. J.: On A-type (statistical) uncertainty estimation in Whole-Body Vibration measurements. In Symposium "Managing Uncertainty in Noise Measurement and Prediction", Le Mans, France, 2005, 1-6.

Device for Precise Measurement of Magnetic Microwire BH Loop at Low Frequency

[1]M. Butta, [1,2]J. Urdahl

[1]Czech Technical University in Prague, Faculty of Electrical Engineering – Dept. of Measurement, Prague, Czech Republic,
[2] Norwegian University of Science and Technology, Faculty of Information Technology, Mathematics and Electrical Engineering – Dept. of Electronics and Telecommunications, Trondheim, Norway
Email: buttam1@fel.cvut.cz

Abstract. *In this paper we present a precise and cheap system we developed to measure BH loops of magnetic microwires. We especially focused on low frequency BH loops because we are mostly interested in quasi-static magnetic properties. Due to low frequency, we had to face problems related to low signal to noise ratio. The extremely small cross-section of the microwires turned out in very small flux to be measured. We achieved our goal by realizing a BH tracer based on induction principle, which generates the current necessary to create H field, and it synchronously samples the induced voltage. Digitized voltages are then sent to PC, which performs numerical analysis and returns the BH loop. The system has proved to be cheap and it provided excellent measurement parameters.*

Keywords: BH loops, microwires, low frequency, magnetic properties

1. Introduction

Magnetic microwires have had a rapid development in the last years, due to several applications they can be employed for, and especially because they can be used to build several kind of magnetic field sensors [1]. We can define as magnetic microwires those wires composed by magnetic material, possibly combined with other non-magnetic materials, whose cross-section has diameter in order of tens of μm.

Researchers dealing with this kind of microwires are interested in precise characterization of their magnetic properties. Vibrating Sample Magnetometers (VSMs) can be used for this purpose, but they are extremely expensive, and they require up to several hours to measure a single BH loop. Induction principle can be used instead. The microwire is placed in a solenoid which generates a time-varying H field: a pick-up coil is wound around the microwire to measure the induced voltage V_{ind}, given by the change of magnetic flux in time. Integration of V_{ind} is then performed to obtain magnetic flux, then B field [Fig. 1]. The voltage due to the magnetic flux of the air is subtracted from V_{ind} connecting in anti-series an exact replica of the pick-up coil, without any magnetic wire inside.

This processing is usually performed by analog instruments [2]. The induced voltage is amplified and integrated by an analog integrator. An oscilloscope is used in XY mode to plot the BH loop. The main problem of this method is the noise and drift of the integrator. When the cross-sectional areal of the wire is so small, the signal is very small too. Moreover, any effort to reduce the voltage due to change of air flux, will never lead to perfect compensation. Finally, we will have a very small voltage mixed into some noise signal also due to air flux. When using this method, extremely high number of averaging is required, because of the low quality of the signal (usually thousands of periods).

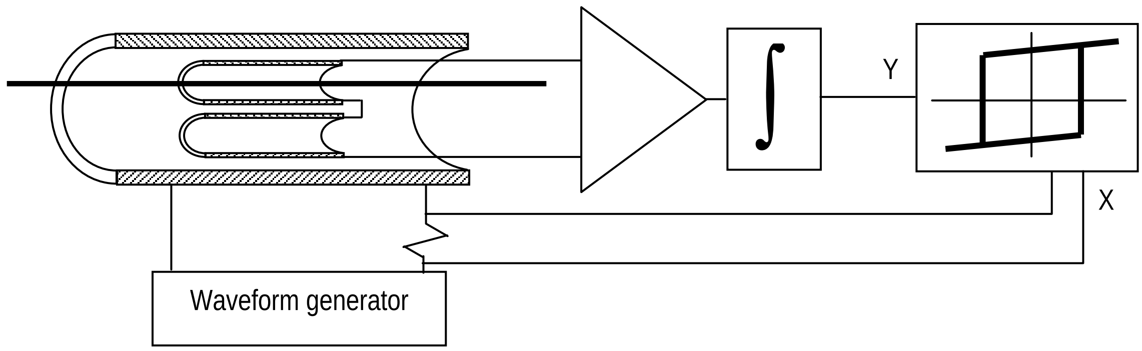


Fig. 1. Simplified scheme of set-up for measurement of BH loop in magnetic microwire.

Moreover, this system requires several instruments: a waveform generator, an amplifier, an integrator and a scope. It is pointless to use all these instruments for this measurement set-up, especially when one needs to constantly have to his disposition a measurement system to measure BH loops, rather than building the set-up once in a while.

Our goal was to develop a card, which performed all these functions, improving in the same time the quality of the measurement.

2. Structure of the developed BH tracer

Fig. 2 shows the simplified scheme of the BH tracer circuit. The core of the system is a PIC18F2550 microcontroller: it provides the connectivity to an host PC by USB bus and it manages all the components of the circuit. The whole circuit is powered by USB bus.

Generation of H field

In order to generate H field we must inject a time-varying current into the main solenoid. We use a programmable waveform generator based on DDS technology (AD9833) to generate both sine and triangular waveform. The frequency of the waveform is an exact fraction of the input Master Clock, provided by the PIC microcontroller. The microcontroller generates the Master Clock using a PWM module, based dedicated registers incremented by the microcontroller clock. In this way we have exact synchronization between the microcontroller and the generated waveform. Frequency range of the resulting H field is between 20 and 100 Hz.

Microcontroller communicates to the waveform generator on SPI bus, instructions such as the choice between sine and triangular waveform. Then, we amplify the waveform and we use a current source circuit (based on feedback operational amplifier and transistors) to feed the main solenoid.

Acquisition of induced voltage

The voltage induced into the pick-up coil is amplified by an AD620 instrumentation amplifier. The gain is set by a resistor, selectable by means of a multiplexer (controlled by the microcontroller as well). Then the amplified voltage is digitized by AD7685 ADC (16-bit, 250 kHz) and the result is sent to microcontroller by SPI bus. The acquired data is stored in internal flash memory of microcontroller and finally sent to PC by USB.

According to the composition of the magnetic material of the microwire, the induced voltage will include sharp peaks, due to sudden change of magnetic flux around coercivity field. Such peaks determine high frequencies into voltage's spectrum and require sampling frequency much higher than H's frequency. Despite the ADC has 250 kHz maximum sampling frequency, we found out that 80-90 kHz were more than enough for our application.

We had to use a 16-bit ADC because the signal could be so small that high resolution is necessary.

Acquisition of H field

In order to measure the H field we acquired the voltage on a shunt resistor. In this case the measured voltage is either sine or triangular waveform at 20÷100 Hz, therefore we used the internal 10-bit ADC of the PIC microcontroller, whose resolution was sufficient. Due to short spectrum we used much lower sampling frequency (2.2 kHz).

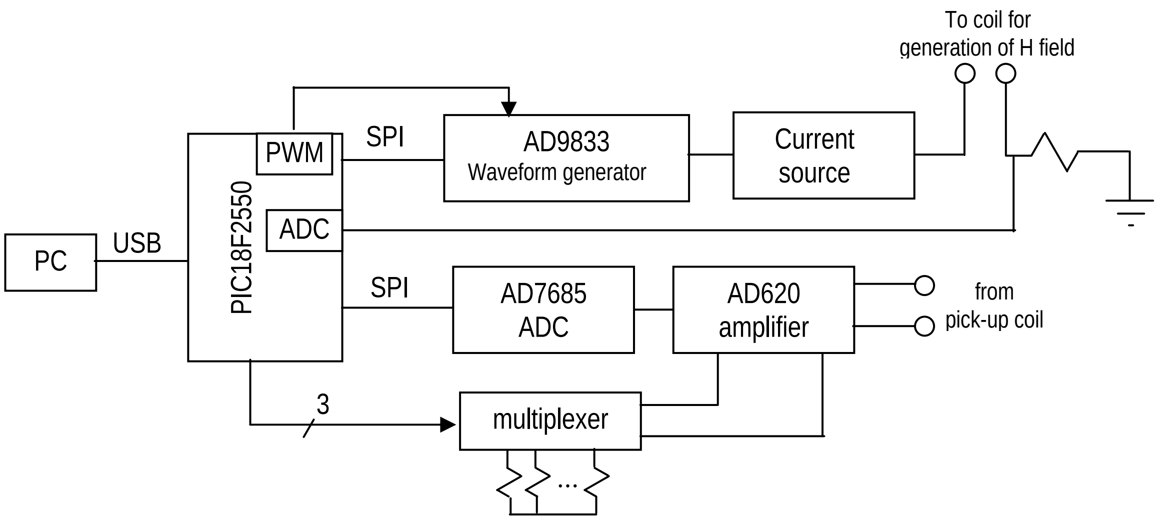


Fig. 2. Simplified scheme of the developed board. The core is a PIC18F2550 which manages all component of the circuit and provides connectivity to PC by USB.

3. Results

We have developed the firmware for the PIC microcontroller and a software in Labview which allows the user to set all the parameters of the circuit (such as frequency and shape of H field, gain of amplifier, sampling rate…); moreover this software acquires the data stored into the PIC and performs numerical integration of induced voltage to derive B field. The software includes the possibility to perform averaging, as well as numerical compensation of air flux.

Due to limited amount of flash memory available on the PIC we had to sample the waveforms in consecutive sections and transfer the data to PC every time. Thanks to already mentioned synchronization between H field and microcontroller we could sample the waveforms every time starting from the exact time when we ended sampling the previous time: no missing points at all. Assuming the signal to be stable, this technique was completely satisfactory.

Quality of the measurement

Thanks to the developed board, we were able to measure the BH loop of magnetic microwire with very high quality, compared with similar analog system. This is clear when we notice that averaging over 10 periods returns excellent BH loops for 40 μm diameter microwire. Previously mentioned analog system require thousands of periods for averaging. Moreover, we do not have problems related to drift of integrator: any average value is subtracted from digitized data before numerical integration.

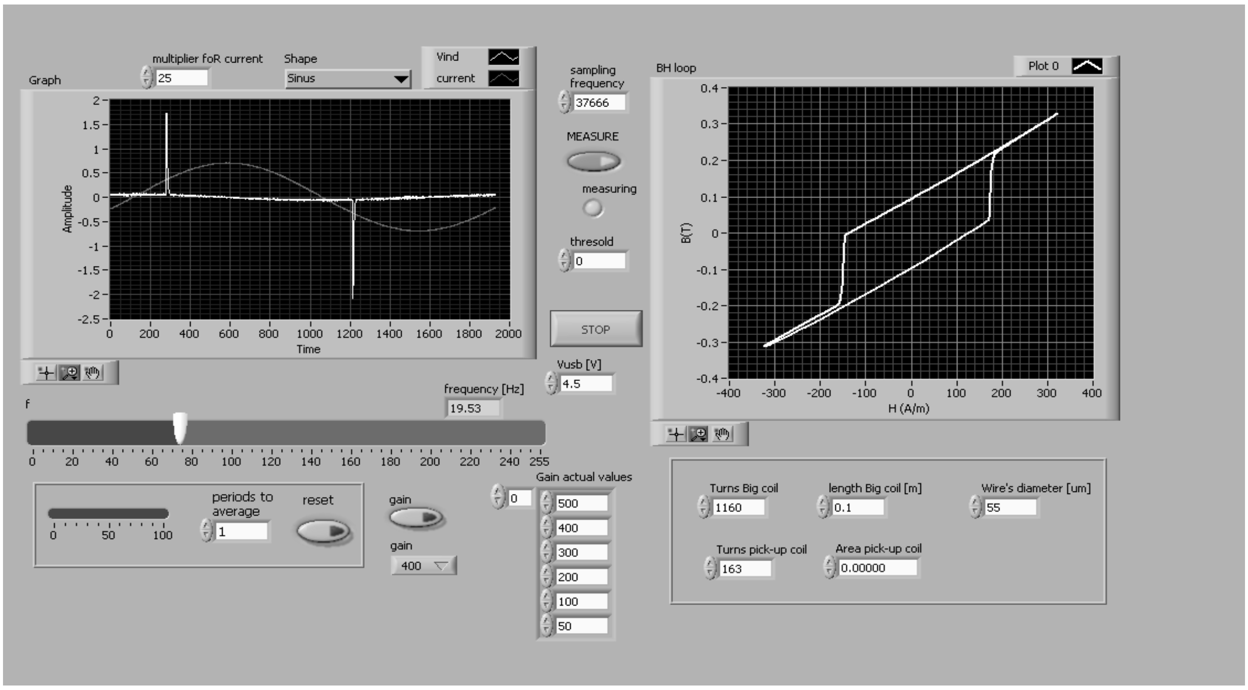


Fig. 3. Screenshot of the software used to acquire the data from the board and compute the BH loop.

4. Discussion

We have presented a board which can simplify the characterization of magnetic microwires. The user can just connect the board to the PC by USB and does not have to care about anything else. The implementation of digitalization of signals and numerical integration proved to improve the quality of the measured BH loop to a level never reached by analogue techniques. Thus, due to low number of averaging the measurement is much faster and reliable. Finally we should mention that all electronic components used to build this card can be bought for less than 50 EUR, that is orders of magnitude less than instruments traditionally used to measure a microwire BH loop by analogue techniques.

Acknowledgements

This project has been supported by the Czech Grant Agency grant No. GA102/08/0743 and from special scholarship for foreign students of the Rectorate of Czech Technical University in Prague.

References

[1] P. Ripka, M. Butta, M. Malatek, S. Atalay, F. E. Atalay, Characterization of magnetic wires for fluxgate cores, Sens. Actuators A: Phys. (2007), doi:10.1016/j.sna.2007.10.008

[2] L. Kraus, O. Chayka.: AC hysteresis loop tracer for soft magnetic thin films, J*ournal of* Electrical Engineering, vol 57. no 8/s, 2006, 73-76

3D Magnetic Field Measurement, Visualisation and Modelling

T. Mikolanda, M. Košek, A.Richter

Technical University of Liberec, Liberec, Czech Republic
Email: tomas.mikolanda@tul.cz

Abstract. *Automated apparatus for the measurement of 3D magnetic field was constructed using Hall probes and 2D positioning system. The apparatus was tested by the field of ring permanent magnets. The measured field is visualised by graphs for 1D lines, surfaces for 2D areas and vectors or flux lines for 3D space. The magnetic field was modelled by coupled currents and then by the application of Biot-Savart Law. The agreement with experiment is good either for the simplest assumption of uniform magnetisation. The model can be refined.*

Keywords: Hall probe, Permanent magnets, Automated measurement, Scientific visualization, Permanent magnet modelling

1. Introduction

The magnetic force in the form of magnetic coupling is used in many technical and scientific applications, especially for contact-less force, momentum, contact-less shaft bearing, levitation, as examples. The survey of magnetic coupling methods and applications can be found in literature [1]. The effective realisation of such apparatus is complicated, money and time consuming. Therefore, ways of optimum theoretical design are welcomed. The first step is the modelling of magnetic field. In the literature either the method of non-existing magnetic charges is used in analogy with electrostatic field [2] or a lot of finite element methods (FEM) are applied. However, other methods exist.

On the other hand magnets are complicated components and their material parameters (magnetization) cannot be known in details. Magnets need to be modelled and the models are subjects of numeric computation. The calculated magnetic field and other properties are approximate ones. The experiment is necessary in order to find the difference between reality and model. It also makes possible the model improvement. Visualization of complete calculated or measured magnetic field is necessary for perfect understanding of the apparatus. It is the reason, why we concentrate to three areas: modelling, visualization and measurement.

2. Subject and Methods

According the introduction three areas, magnetic field measurement, visualization and modelling, will be mentioned here. We limit here to permanent magnets, but the approach can be used for electromagnets or any other combinations as well.

As for magnetic field measurement, the 3D Hall probe was found to be the best practical device for magnetic field scanning. In order to make 3D measurement, two perpendicular positioning system of Hall probe is used for its location in a horizontal plane. The vertical probe shift is made manually. The step motors are used for the realization of horizontal motion. The apparatus is controlled by the electronic system that realizes the instructions from computer program that is written in the low level C language for step motor motion and in high level MATLAB for automated measurement. Many parameters can be set in order to make the measurement as quick as possible.

The simplest way of magnetic field visualization is the graph. Graphs are effective quantitative visualization means, but they deal only with a small 1D part of the magnetic field. Therefore, the surface graphs are used for illuminating prezentation of the 2D field in the plane. For 3D field the flux density vectors or lines were found as the best vizualization mean. All the objects are produced by MATLAB after the finishing the measurement or calculation.

As for the permanent magnet model, the magnet is given very simply by its geometrical parameters and magnetization $\boldsymbol{M}$ that fully describes its magnetic properties. There are two basic ways how to calculate all the magnet effects. They can be derived by superposition either from elementary magnets of magnetic momentum $\boldsymbol{M}\mathrm{d}V$ in the magnet volume or from the coupled elementary volume $\boldsymbol{i_m}\mathrm{d}V$ and surface $\boldsymbol{j_m}\mathrm{d}S$ currents. Their densities are derived from volume and surface distribution of magnetization by formulas [3]

$$\vec{i}_m(\vec{r}_0) = \mathrm{rot}\,\vec{M}(\vec{r}_0) \qquad \vec{j}_m(\vec{r}_0) = \mathrm{Rot}\,\vec{M}(\vec{r}_0) = \vec{n}\times\left(\vec{M}_2(\vec{r}_0) - \vec{M}_1(\vec{r}_0)\right) = -\vec{n}\times M(\vec{r}_0)\,. \tag{1}$$

Thanks to its clarity, the method of elementary coupled currents was preferred. The magnetic flux density $\boldsymbol{B}$ due to coupled currents is then given by Biot-Savart Law

$$\vec{B}(\vec{r}) = 10^{-7}\int_{(S)} \frac{\vec{j}_m\times(\vec{r}-\vec{r}_0)}{\left|\vec{r}-\vec{r}_0\right|^3}\,\mathrm{d}S + 10^{-7}\int_{(V)} \frac{\vec{i}_m\times(\vec{r}-\vec{r}_0)}{\left|\vec{r}-\vec{r}_0\right|^3}\,\mathrm{d}V\,, \tag{2}$$

where $\boldsymbol{i_m}$ and $\boldsymbol{j_m}$ are coupled volume and surface current density, respectively, S and V are the surface and volume of the magnet, respectively, $\boldsymbol{r}$ is a position vector of the point, where the flux density $\boldsymbol{B}(\boldsymbol{r})$ is calculated, $\boldsymbol{r_0}$ is the position vector of surface and volume elements dS and dV, respectively. The formula requires numerical integration and can be used for the calculation not only of magnetic field, but also for magnetic forces.

3. Results

The permanent magnet has a shape of a thin ring with inner diameter of 25 mm, outer diameter of 70 mm and height of 4 mm. The magnetization was 1.2 T according to the data sheet of producer. It is the only known material parameter.

We realized fully automated apparatus shown in Fig. 1 together with its block scheme of control. Both the transport devices and 3D Hall probe are the basic parts. The driving step motors, control system and computer make possible the automated measurement. The scanned area is relatively large, about 0.5 times 0.5 meters. Theoretical accuracy is about 0.1 mm; in practice it reduces to a value bellow 1 mm. Due to the small, but non negligible, dimensions of Hall probe, the average flux density is measured rather than its true point value.

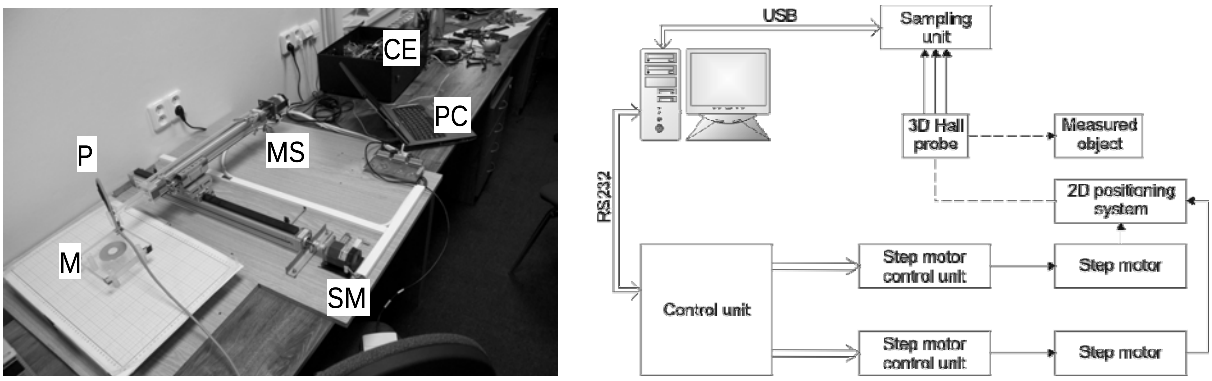


Fig. 1. Apparatus and its block schema: M – magnet, P – probe, MS – mechanical shift, SM – step motor, PC – controlling computer, CE – control electronics.

All the components of the magnetic field were measured by the step of 1 mm in both the *X* and *Y* directions. Typical results in the form of surface graphs are shown in Fig. 2. The *x* component of magnetic flux density, B_x, is in left hand side of Fig. 2, while the most important component B_z is in the right hand side of Fig. 2

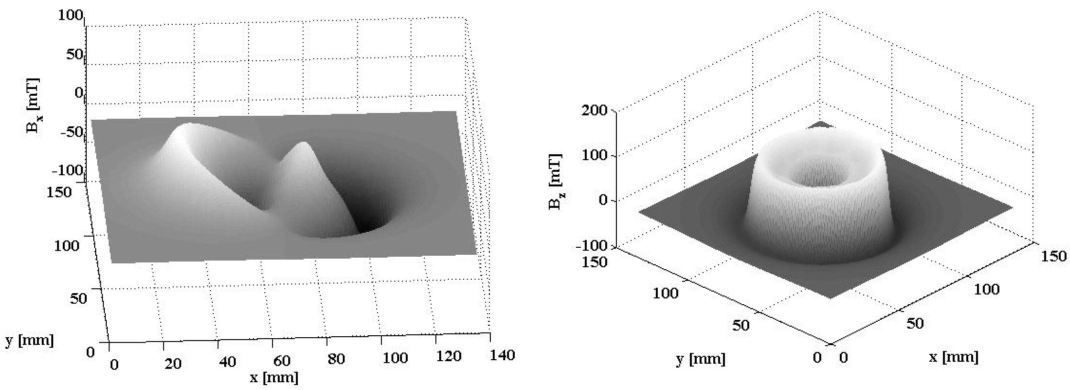


Fig. 2. Surface graphs of magnetic flux density 2 mm from magnet surface. Component B_x is on the left hand side and the B_z one on right hand side.

The same field in the vector form is at Fig. 3. Left hand side shows vector representation of magnetic flux and right hand side displays several important flux lines of magnetic field.

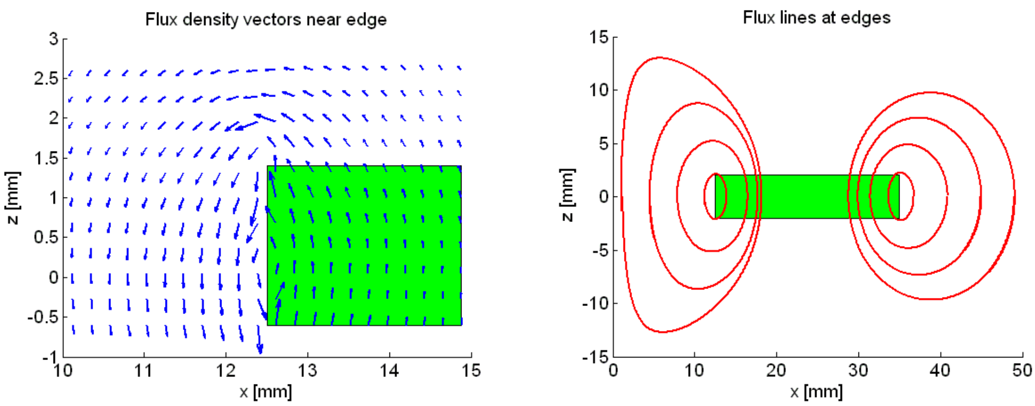


Fig. 3. 2D cut of flux density near permanent magnet edge (left) and flux lines near the face (right)

Comparison of experiment and model for central cut in the form of classical graph is in Fig. 4.

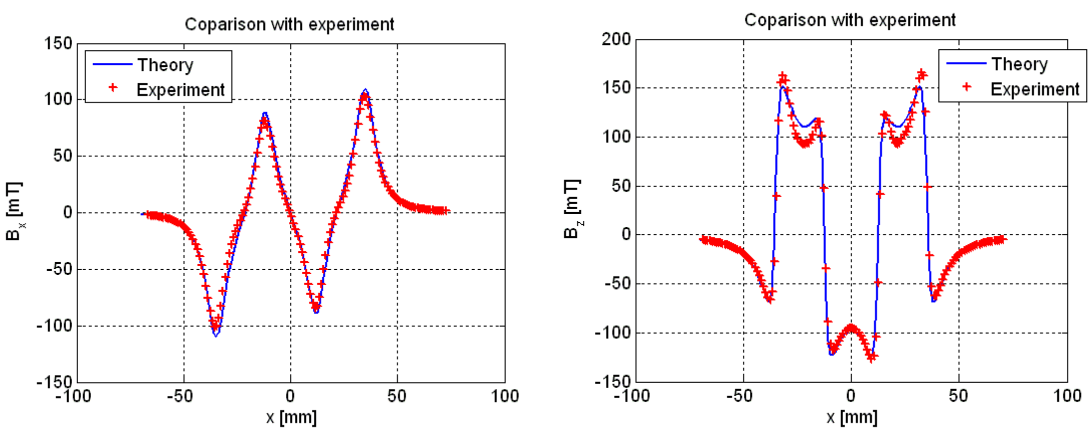


Fig. 4. Comparison with experiment of magnetic flux at permanent magnet surface. X-component of magnetic flux on the left side, z -component on the right side.

4. Discussion and Conclusions

Only preliminary results of automated magnetic field measurement were presented. Main problem is the correct position of sample that is necessary for simple comparison of experiment and theory. The use of absolute coordinates with exact and rigid sample positioning should be realized. The speed of measurements is relatively low. Measurement in a plane by 140 x 140 points takes about 300 minutes. Therefore, automation of the shift in the 3rd dimension is not necessary.

Several methods of magnetic field visualization were presented. Although the standard graph contains the smallest part of information with respect to the 3D vector field, the information is ready to direct use, see Fig. 4, for instance. Qualitative information on plane regions is in surface graphs, typical representatives are in Fig. 2. The full description by the use of 3D vectors is not usually transparent. It can be valuable in special cases. However, the use of vectors and flux lines in selected 2D regions is very descriptive, see in Fig. 3.

The main purpose of the work is the comparison of model and experiment. The outlined model needs only geometrical dimension of magnets and one material parameter, magnetization. In the first step we suppose the uniform magnetisation. As it follows from Fig. 4, the agreement with experiment is good with an exception of the field near sample edges. The magnetization is not uniform in rectangular sample cross section and the highest deviations can be expected at sample edges. The agreement can be improved, if we refine the model by the non-uniform magnetization. The only straightforward way how to find the correct distribution of magnetization, is the use the FEM.

We have used the model of coupled currents and integral formulae for the calculations. Usually the FEM is preferred. Our approach has several advantages. The programming is relatively simple, the user has full check at all steps of computation, and there are no problems with boundary conditions especially in infinity. All the quantities can be calculated at any given point and the accuracy can increase to any reasonable value. The only disadvantage is relatively long computation time, but the cluster can be used, if necessary.

The measured and modelled results were used for the numerical calculations of repulsive force between ring magnets used in practice. The results agree well with experiment in several orders of the force [4]. Also the momentum, which is difficult to measure, was found.

Acknowledgements

This work was supported by the fund of Ministry of Industry No. FT-TA3/017 and the fund of Czech Science Foundation No. GAČR 102/08/H081.

References

[1] Overshott KJ. The comparison of the pull-out torque of permanent magnet couplings. *IEEE Transactions on Magnetics,* 25 (5): 3913 – 3915, 1989.

[2] Furlani EP. Formulas for force and torque of axial couplings. *IEEE Transactions on Magnetics,* 29 (5): 2295 – 2301, 1993.

[3] Haňka L. Theory of electromagnetic field, SNTL, Praha, 1975. (in Czech).

[4] Košek M, Mikolanda T, Richter A, Škop P. Intelligent mechatronical system using repulsive force produced by permannet magnets. In Proceedings of the 13th International Conference on Material Engineering, Mechanics and Design, 2008, 1 – 4. ISBN 978-80-969728-2-1.

Magnetic Field Measuring Devices for Low Field MARY Spectrometer

[1]E.V. Kalneus, [1,2]D.V. Stass, [2,3]A.F. Pavlov, [2]A.V. Ivannikov

[1]Institute of Chemical Kinetics and Combustion SB RAS, Novosibirsk, Russia
[2]Novosibirsk State University, Novosibirsk, Russia
[3]Geophysical Survey SB RAS, Novosibirsk, Russia
Email: kalneus@kinetics.nsc.ru

Abstract. *The paper reports the development of magnetic field measurement system for MARY spectrometer specifically intended for studies of chemical reactions in low magnetic fields. The MARY spectrometer has a scanning magnetic system (up to 50 mT) and a system for compensation of the residual magnetic fields (~ 0.1 mT). We discuss Hall sensor, fluxgate meter and magnetoresistive sensor as the transducers for measuring magnetic field in these ranges. The details of construction and testing of the actually built measurement devices based on these sensors are provided.*

Keywords: MARY Spectroscopy, Low Magnetic Field, Magnetoresistive Sensor, Hall Sensor, Fluxgate Meter.

1. Introduction

The influence of weak magnetic fields on chemical reactions passing via short-lived paramagnetic intermediates is intriguing both from chemical and biological points of view. The method of Magnetically Affected Reaction Yield (MARY), or level-crossing, spectroscopy as developed in the authors' lab takes advantage of narrow resonance-like lines on the dependence of the intensity of recombination fluorescence from X-irradiated sample on external static magnetic field. The lines arise in zero and in weak (1-10 mT) magnetic fields due to coherent evolution of the spin systems of radical ion pairs that are formed upon ionization of molecules in the sample [1]. Of particular interest now is the region of very weak magnetic fields comparable to Earth's field (0-0.5 mT) [2]. A dedicated MARY spectrometer optimized for the region of weak magnetic fields is currently being developed, and its magnetic system has been described in [3]. In this work we describe the systems of measurement and control of the field created by MARY spectrometer in 0-50 mT range.

2. Subject and Methods

The scanned DC magnetic field in the MARY spectrometer is swept along one axis from -50 mT to +50 mT with accurate passage through the zero of the field. As the developed spectrometer has no ferromagnetic elements, the stabilization of the created field can be conveniently performed by stabilization of the currents through its coils. Although this approach avoids the feedback loop locked on the magnitude of the field, a careful control of the created field is still an essential feature of the spectrometer. The main problem for the field measuring device here is the wide dynamic range of field changes, as it is necessary to measure with relative accuracy of at least 10^{-2} both rather weak (of the order of 0.1 mT) and rather high (of the order of 50 mT) magnetic fields almost simultaneously. The sensors should also have good linearity and be compact (not larger than 1 cm^3). The field is swept along the z axis of the spectrometer, and the remaining two components (x and y) of the residual magnetic field are compensated by a dedicated subsystem. The setting of the compensation system also requires measuring magnetic field, but the requirements here are somewhat different. In this case it is sufficient to provide a stable and sensitive (sensitivity at least 1 µT) indicator of the zero of the field.

Unfortunately a single magnetic field sensor meeting all these requirements is still not available. We thus decided to combine several magnetic field transducers based on different physical principles and divide the task of magnetic field measurement into three ranges: linear measurement of the scanned magnetic field in the range 0.5 – 50 mT, linear measurement of the field in the range 0 – 0.5 mT, and a stable zero field indicator.

3. Results and Discussion

Hall sensors are convenient transducers for wide range of measuring magnetic field strength. They are easy to use, compact, have good linearity, but suffer from poor temperature stability and zero offset. Furthermore, sensors with rather large working surface are required to measure the sub-gauss fields, which are not commercially available. Control sensor temperature helps improve its parameters, and special schematic solutions are used to compensate the offset [4]. Several chips are now commercially available that integrate the sensor with the compensating solutions to substantially improve specifications of the transducer. One of the producers of chopper-stabilized linear Hall sensors is Allegro MicroSystems Inc. To measure magnetic field in the range 0.5 – 50 mT we selected their Linear Hall sensor A1321 having sensitivity 5 mV/G, temperature stability and zero offset at room temperature about 10^{-2} [5]. However, the intrinsic noise of up to 40 mV (Peak-to-Peak) of the sensor makes it unsuitable for measurement of the lower fields.

To measure magnetic field in the range 0-0.5 mT with accuracy better than 10^{-2} we used magnetoresistive sensors HMC1052 from Honeywell [6]. The sensor has working range ±0.6 mT and sensitivity 12 nT, with linearity in the field range ±0.3 mT of at least 10^{-3} and temperature coefficient 10 ppm/ºC. The transducer itself is a bridge of four thin-film Ni-Fe magnetoresistive elements, complemented by proprietary integrated electronics for remagnetization of the material of the magnetosensitive resistors. This approach allowed the producer exploit the linear portion of the dependence of sensor elements resistance on external magnetic field and thus substantially improve the parameters of the device. To turn the sensor HMC1052 into a computer-controlled measurement system we developed a controller with RS 232 interface for a PC and the required software. Fig. 1 shows a picture of the working prototype.

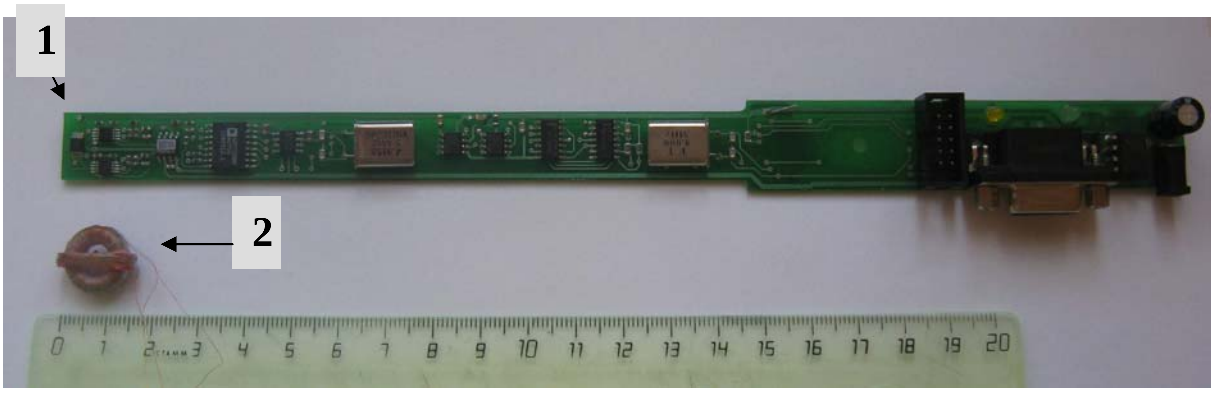


Fig. 1. Photo of magnetic field meter based on magnetoresistive sensor (1) and fluxgate sensor (2).

Both Hall sensor and magnetoresistive sensor require the compensation of the zero offset and sensitivity calibration. The offset was measured by placing the sensor in a mu-metal screen.

To calibrate the sensitivity we built a solenoid with a large aspect ratio to calibrate longitudinal sensitivity, and a system of coils with a clearance in-between to calibrate the transversal sensitivity. The latter was required to calibrate the second half of the magnetoresistive sensor (the crystal of HMC1052 contains magnetoresistive bridges measure two mutually orthogonal field components). Both calibration systems were calculated to provide the relative field homogeneity better than 10^{-3} over the size of the sensor (within ±5 mm from the center) allowing for manufacturing tolerances and thus allowed to perform the calibration with the required accuracy. To take into account the presence of the residual external fields the calibration was performed with reversal of the current in the coils.

The performed measurements demonstrated that the actual magnetoresistive sensor used in our meter had quite low deviations from the nominal sensitivity and offset values as provided in its technical specifications, with the relative values not exceeding $3\text{x}10^{-3}$.

A fluxgate was chosen as the zero indicator for regulation of current in the compensation coils. The fluxgate that we built is a coil of thin annealed mu-metal foil (3 x 200 mm, thickness 30 µm) on a 13 mm quartz former for mechanical stability, with the excitation coil having 100 turns of 0.1 mm copper wire. The detection coil was wound over the ring and has 50 turns of 0.1 mm copper wire. The picture of the transducer is also shown in Fig. 1. The fluxgate was tested in the Geophysical Survey SB RAS, Novosibirsk, on a working fluxgate meter and showed the sensitivity of about 10 nT. Currently we are building the stand-alone fluxgate meter based on this transducer to be used as the zero field detector.

4. Conclusions

While developing the systems of magnetic field measurement and control for the low field MARY spectrometer we analyzed the available field sensors and chose those most suitable for the particular tasks in the spectrometer. The meter for the scanned field of the spectrometer was built using an A1321 Hall sensor (Allegro MicroSystems Inc) and a HMC1052 magnetoresistive sensor (Honeywell). A fluxgate transducer was built and tested for compensating the transversal components of the residual magnetic fields in the working region of the spectrometer.

References

[1] Stass D.V., Tadjikov B.M., Molin Yu.N. Manifestation of quantum coherence upon recombination of radical ion pairs in weak magnetic fields. Systems with equivalent nuclei, *Chem.Phys.Letters* 235, 511, 1995.

[2] Maeda K., Henbest K.B., Cintolesi F., Kuprov I., Rodgers C.T., Liddell P.A., Gust D., Timmel C.R., Hore P.J. Chemical compass model of avian magnetoreception, *Nature* 453, 387, 2008.

[3] Kalneus E.V., Stass D.V., Grishin Yu. A. Yoke-free magnetic system for low field studies in magnetically affected reaction yield spectroscopy, *Rev. Sci. Instrum.*, 76, 084102, 2005.

[4] Bilotti A., Monreal G., Vig R. Monolithic magnetic Hall sensor using dynamic quadrature offsetcancellation *IEEE Journal of Solid-State Circuits*, 32, 6, 829–836, 1997.

[5] http://www.allegromicro.com/en/Products/Part_Numbers/1321/1321.pdf.

[6] http://www.ssec.honeywell.com/magnetic/products.html.

[7] Geyger W.A. New Type of Flux-Gate Magnetometer, *J.Appl.Phys.*, 33, 1280, 1962.

Effect of Cable Termination on EMI Measurement Results

M. Bittera, V. Smieško

Department of Measurement, Slovak University of Technology, Bratislava, Slovakia
Email: mikulas.bittera@stuba.sk

Abstract. *The paper deals with effect of terminations of interference cable of equipment under test to results of EMI measurements. It is evident that interference cables represent the potential sources of undesired radiation, which can be suppressed or emphasized by choice of termination load value. The behaviour of terminated two-wire cable is analysed in term of its input impedance. The radiation is also surveyed using numerical techniques based on models that are verified by analytical calculation and measurement.*

Keywords: EMI measurement, emissions measurement, cable termination, two-wire cable

1. Introduction

Especially due to many disturbing factors, electromagnetic interference (EMI) measurement is a very complex measurement. During the measurement we shall identify the maximal intensity of unwanted electromagnetic field radiated from the equipment under test (EUT), measure it and compare with limit values. Because of reproducibility, the measurement is performed strictly according to international standard [1]. Also EUT shall be in typical mode of operation and in a test arrangement that is representative of typical installation practice. It is because the influences of equipment, which are necessary for performing EMI measurement, and of test site arrangement, can be included into measurement correction or its uncertainty, but the effect of the EUT cannot [2].

The EUT cables and their arrangement are the most problematic part of EUT by emission measurement. Interference cables have to be connected to each interference port of EUT, while the type and length of cables have to be specified by equipment manufacturer. The cables should be no longer than 0.4 m. If it is necessary for normal operation of EUT the ends of interference cables that are not connected to other auxiliary equipment can be terminated by real impedance. In the paper, we concentrate on the role of termination of two-wire cable as prospective radiator of disturbance and its effect on EMI measurement reproducibility. We suppose that the cable is attached to electrically small tabletop EUT and the differential mode disturbance is applied on.

2. Problem analysis

In term of radiation, the behavior of the mentioned cable can be expressed using transmission line method. For a transmission line, it can be shown that the input impedance Z_{in} of a cable of length l and loaded by an impedance Z_L is

$$Z_{in} = Z_0 \frac{Z_L + Z_0 \tanh \gamma l}{Z_0 + Z_L \tanh \gamma l} \qquad (1)$$

where γ is the propagation constant and Z_0 is the characteristic impedance of the transmission line. Having the distance D between the wires of the cable and the radius d of the conductors, the characteristic impedance of the two-wire cable in medium with permittivity ε is given

$$Z_0 = \frac{120}{\sqrt{\varepsilon}} \cosh^{-1}\left(\frac{D}{d}\right) \qquad (2)$$

while propagation constant γ is also function of resistance R, conductivity G, inductance L and capacitance C of the cable, which are dependent on material parameters of the cable and its surrounding

$$\gamma = \sqrt{(R + j\omega L)(G + j\omega C)} \tag{3}$$

The using of transmission lines theory was verified also by the real measurement with network analyzer on chosen model of the two-wire line (see Figure 1). Using substitution (2) into (1) and the Ohm's law consequently, it is possible to calculate the current I through the transmission line.

The two-wire cable as antenna can be represented by rectangular loop antenna. In this case the magnitude of E-field component of electromagnetic field E in arbitrary point of the surrounding space specified in spherical coordinate system (r, ϕ, θ) is then given [3]:

$$E = \frac{8\eta I}{r} \frac{\sin\left(\frac{ka}{2}\sin\theta\cos\phi\right)\sin\left(\frac{kb}{2}\sin\theta\sin\phi\right)}{\sin\theta\sin 2\phi} \tag{4}$$

where I is feed current flowing through the loop antenna, η free space wave impedance, k phase constant; a and b represent the dimensions of the rectangle. In our case, it is evident that $a \gg b$, so the transmission wire can be compared with folded dipole. Hence, if the current is uniform along the antenna, the radiation is very weak for small b, since the radiation from the two long arms of the antenna nearly cancels. Of more interest would be the case when the current is not uniform (if $a \approx \lambda$), consequently the currents in the two long arms flow in the same direction. While the shorter side of the loop $b \ll \lambda$, the E-field radiated from the transmission line we get by simplification of equation (4):

$$E = \frac{\eta kbI}{r} \frac{\sin\left(\frac{ka}{2}\sin\theta\cos\phi\right)}{\sin\phi} \tag{5}$$

The equations (4) and (5) has two main disadvantages. They are suitable only for computing the E-field in far-field zone, which is not fulfilled in case of lower frequencies of our interest and they do not include other surrounding material as free space is. Therefore it is necessary to find other method. The [4] shows the advantage of numerical simulations that constitute powerful tool to analyze such structures as wire cables. To analyze the two-wire cable effect by numerical simulator, it is enough to build a proper model of such a cable.

3. Results

As the model of the interference cable simple 40 cm long two-wire cable was chosen, which consists of two parallel wires, with 50 Ω termination on one end and with point voltage source on the other one. The load impedance is chosen not equal to characteristic impedance to survey its radiation behavior. The diameter of these wires is 0.4 mm and the distance between them 1.8 mm. The long two wire cables are coated by plastic insulation. The input impedance Z_{in} of this model was obtained using all the mentioned methods. In Fig. 1 one can see the conformity of the results. Even though the first impedance maximum is expected at first resonance frequency about 187 MHz, the permittivity of the isolating material, which is higher than permittivity of free space, causes the shift of resonance frequency to the lower values (about 110 MHz). Also values of supply current I and electric field E in frequency range 30 ÷ 500 MHz obtained by theoretical computation using (1) and (4) and numerical simulation of chosen model, but without isolation, are shown in Fig. 2. We assume that signal voltage has the constant level of 1 V in whole frequency range. In case of current comparison their frequency dependences have the same tendencies; only frequency of maximum calculated current is slightly moved to higher values of frequencies. On the other hand some differences of E-field values in distance 3 m are evident especially at lower frequencies. It is because equation (5) is suitable only for computation of E-field in far-field. The validation of this model using real measurement can be found in [3].

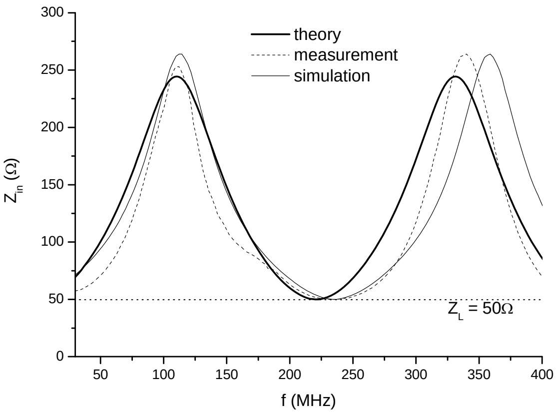


Fig. 1. Frequency dependence of input impedance of 40 cm two-wire cable for 50Ω load.

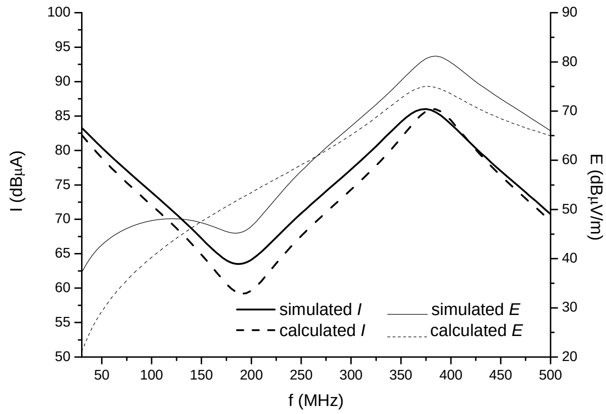


Fig. 2. Frequency dependence of supply current of 40 cm two-wire cable and radiated field in 3 m distance.

It is expected that cable terminated by load impedance Z_L, which is equal to characteristic impedance Z_0, has the lowest unwanted radiation. In most cases unfortunately we do not know the impedance Z_0. Then the determining parameter is input impedance Z_{in} of the analyzed cable. It can be obtained simply by measurement using network analyzer (see Fig. 3). The input impedance Z_{in} determines the behavior of such cables; we expect usually the maximum of radiation if the Z_{in} is in its minimum and vice-versa.

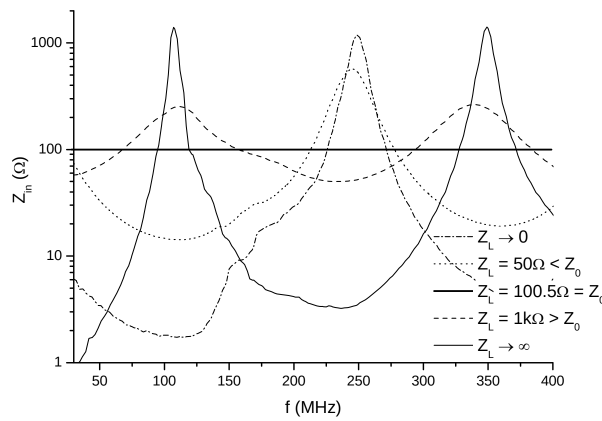


Fig. 3. Measured values of input impedance Z_{in} of 40 cm two-wire cable for different terminations Z_L.

To confirm our expectation we used numerical simulation with chosen model - simply hanging 40 cm long two-wire cable with given parameters. Two different simulations were performed – to obtain E field in 3 m distance from the cable in free space and in presence of ground plane (1 m over the plane as it is given for standard EMI measurement). The results of simulations are shown in Figures 4 and 5. As we can see from these dependences, the E field from cable terminated by Z_0 is not constant but without evident or sharp extremes. Using terminators with lower values of Z_L the cable behaves as folded dipole and for higher values of Z_L as classic half wave dipole, but with half value of maximal radiation frequency. The level of radiated E field depends on difference between values of terminating impedance Z_L and characteristic impedance Z_0 of the cable. It means that the worst situation of EMI measurement is when we used the interference cable with open or short circuit. Note in Fig. 4 that for lower values of Z_L there are no minimums in the E field frequency dependence, they appear just in case of ground plane (Fig. 5). Otherwise, the ground plane presence does not influence the frequency dependences of E field very evidently. Results shown in Fig. 5 can help us to find the frequency with maximal radiation.

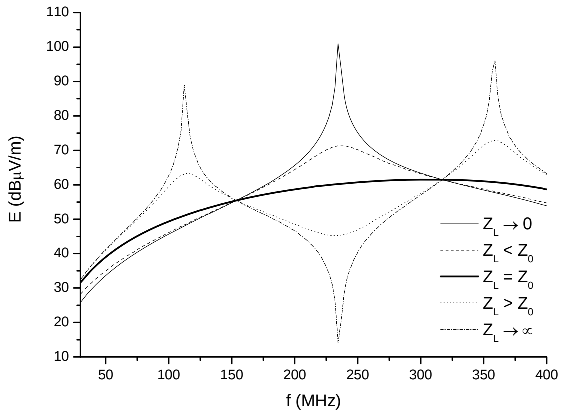


Fig. 4. Frequency dependence of radiated E-field from 40 cm two-wire cable as a function of terminator value Z_L in free space

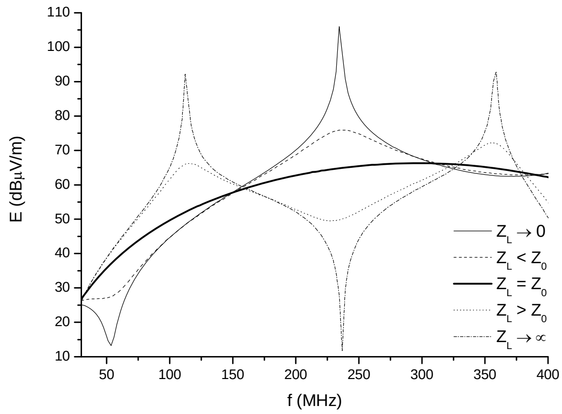


Fig. 5. Frequency dependence of radiated E-field from 40 cm two-wire cable as a function of terminator value Z_L with ground plane

4. Discussion

The models of two-wire interference cable were presented, verified and analyzed to survey the properties and behavior in term of potential radiation due to differential mode disturbance. In general, the character of cable radiation is given by properties of the cable. The wrong termination can cause additional maximums of radiation and the fail of the EMI measurement test consequently. To get the frequencies of potential radiation maximums quickly one can use mentioned method – measurement of input impedance of the cable. Hence to minimize the radiation from the cables of EUT, it is recommended to terminate them by impedance with value close to characteristic impedance (it is necessary to get as straight frequency dependence of impedance as possible).

Acknowledgements

This research was financially supported by the project VEGA VG 1/0551/09 and by the Ministry of Education of the Slovak Republic under grant 2003SP200280202.

References

[1] CISPR 22 Information technique equipment – Radio disturbance characteristics – Limit and methods of measurement. 5th edition, IEC, 2005.

[2] Bittera, M., Smieško, V., Kováč, K. Problem of bundled two-wire cable of tested equipment in emission measurement, *Radioengineering*, 15(4): 22-26, 2006

[3] Bittera, M., Smieško, V., Kováč, K. Interference cable effect to EMI measurement results. In proceedins of the Meditarreanen Microwave Symposium MMS, 2007, 159-162.

[4] Hallon, J., Kováč, K., Szolik, I. Geometrical configuration of cabling as factor influencing the reproducibility of EMC imunity tests. *Radioengineering*, 15(4), 27-33, 2006

Non-reciprocal passive protection of HF transmitters from reflection failure by the help of semiconductor isolators

L. Laurinavičius

Department of Electrical Engineering Vilnius Gediminas Technical University
Sauletekio av.11, LT-2040 Vilnius, Lithuania
Email: laurinavicius@el.vtu.lt; lagnel4@gmail.com

Abstract. *The fundamental problem of protection the transmitters against microwave energy reflections from aircraft antenna is discussed. The paper presents some applications of active and passive isolators and circulators as a tool for protection reflections. It has been stated that electrical properties of ferrite isolators an circulators grow worse at lower operational frequencies (below 50 MHz) and temperatures (cryogenic). Semiconductor magnetoplasma exhibits gyrotropic effects similar to those of magnetized ferrites and can solve the problem non-reciprocity in the HF range from 3 to 30 MHz, which is important for HF communication. The semiconductors devices realization, optimal parameters, operational temperature range, based on minimum forward loss α and maximum reverse attenuation β, were chosen. Major electromagnetic parameters of semiconductor isolators were determined.*

1. Introduction

Ferrite circulators and isolators are the key elements in modern microwave engineering. Their fundamental property of non-reciprocity is capable of simplifying the construction and improving the stability, efficiency and accuracy of radar, communication and testing systems.

An isolator is a passive non-reciprocal 2-port device which permits signal energy to pass through it in forward direction whilst absorbing energy in the reverse direction.

A circulator is a passive non-reciprocal device with 3 or more ports. Energy introduced into first port is transferred to second port, the others ports being isolated.

The frequency range within the non-reciprocal ferrite devices meets the guarantied specifications is no less 50 MHz.

Non-reciprocal devices placed between the transmitter and antenna. RF signal can reach the antenna but returning signal goes to the load (circulator) or dissipating at the core (isolator).

Insertion loss α (expressed in dB) is the attenuation that results from including the device. It is also defined as the ratio of signal power at the output of the inserted device to the signal power at the input of the inserted device.

Isolation β is the ratio, expressed in dB of the input power to output power for signal in the reverse direction.

Maximum power and temperature range are the same important parameters for the non-reciprocal devices.

Electrical properties of ferrite isolators and circulators grow worse at lower operational frequencies (below 50 MHz) and temperatures (cryogenic).

Semiconductor magnetoplasma exhibits gyrotropic effects similar to those of magnetized ferrites and can solve the problem non-reciprocity in the HF range from 3 to 30 MHz, which is important for HF communication [1].

A different approach to the basic nonreciprocal mechanism is used in solid-state plasma devices. This approach makes use of solid-state plasma waves, so-called helicon waves which are nonreciprocal by their nature [2].

An active isolator the same can solve the protection problems [3]. The active devices have noise, input and output, linearity problems, but they are useful solution over wide bandwidths, where ferrite and semiconductor isolators are unavailable.

2. Physical background of investigation

Physical background of the operation of semiconductor isolator is based on the effect of the dimensional resonance of helicon waves in plate placed in the static magnetic field.

The propagation of magnetoplasma wave $k_{\pm}$ may be detected from the characteristic equation [4].

$$c^2 k_{\pm}^2 = \frac{\omega_p^2 \omega}{\pm \omega_H + i\tau^{-1}}, \; \omega_p^2 = \frac{4\pi Ne^2}{m}, \quad \omega_H = \frac{eH}{mc}. \tag{1}$$

where k is wave vector, ω is frequency, ω_H is cyclotron frequency, ω_p is plasmas frequency, H is magnetic field, τ is relaxation time and m is effective mass carries. The items with the argument k_+ are caused by the helicon waves. In the case $\omega_H \tau >> 1$ we have

$$k_+ = \sqrt{\frac{\omega_p^2 \cdot \omega}{\omega_H \cdot c^2}} \left(1 - \frac{i}{\omega_H \tau}\right). \tag{2}$$

And effective magnetic permeability

$$\mu_{eff} = \frac{\Phi_x}{2aH} = \frac{tg_- a}{2k_- a} + \frac{tg_+ a}{2k_+ a}. \tag{3}$$

The μ_{eff} has maximum in the case of resonance

$$(\mathrm{Re}\, k_+) a = \frac{n\pi}{2}, \; n = 1, 3, 5, ... \tag{4}$$

3. Modeling of the semiconductor isolators

We have developed a model of the helicon isolator. We considered the electromagnetic parameters of the helicon resonator material, dimensional effects, the impact of the mine carriers, and some other issues.

The main element of the model represented is the nonreciprocal transformer, characterized by the matrix of impedances.

The impedance matrix for the nonreciprocal transformer is calculated according to the formula [5].

$$|Z| = \begin{vmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{vmatrix} = \begin{vmatrix} i\omega L_1 & i\omega M_{21} \\ i\omega M_{12} & i\omega L_2 \end{vmatrix}, \tag{5}$$

where $L_1 = \mu_L L_{01}$ and $L_2 = \mu_L L_{02}$ are inductances of coils whose physical dimensions are approximately the same as helicon resonator size a $M_{21} = -M_{12} = M_L\mu_T$, $M^2_L = L_{01} L_{02}$.

μ_L and μ_T are „magnetic permeabilities“of semiconductor core in parallel and perpendicular directions, when resonance (n = 1) appear

$$\mu_L = 1 - 4/\pi^2 (1 + 2iQx_m)/ [1 + iQ(x_m - x^{-1}_m)], \tag{6}$$

$$\mu_T = 4u/\pi^2 (1 / [1 + iQ(x_m - x^{-1}_m)], \tag{7}$$

where $u = R_H|B|/\rho = \mathrm{tg}\Theta$, Θ are the Hall angle and R_H is the Hall constant.

Quality of the helicon resonator depends on the Hall constant and magnetic induction B

$$Q = ½ (1 + u^2)^{1/2}. \tag{8}$$

The angular frequency of the mine dimensional resonance (resonance Fabry-Perot) ω_r (0)

$$\omega_r (0) = (2\rho Q/\mu_0)(\pi^2/d^2). \tag{9}$$

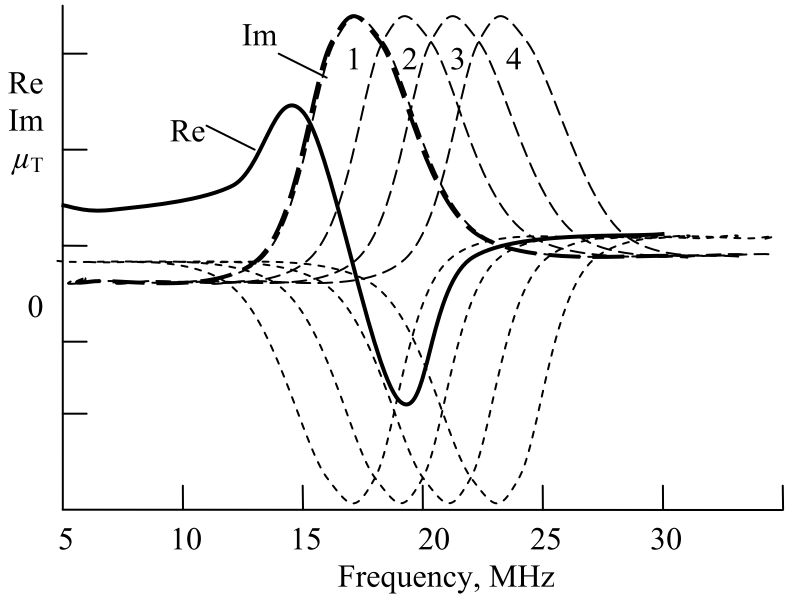


Fig. 1. Calculated frequency responses of magnetic permeability μ_T of a nonreciprocal transformer.

If $u = \mu B > 1$, $\omega_r (0) << \omega_H$, where μ is carrier mobility, the frequency of the mine dimension resonance can be written in simple form [5],

$$f_r = AB/ (Nd^2), \tag{10}$$

where A = 7,82 · 10^{24} is the constant, N is density (concentration) of the carrier of the semiconductor material.

The quadripole of this type will resemble a gyrator. The half wave helicon resonator with the inductance coils represents passive device - the nonreciprocal transformer with nonreciprocal phase shift equal to 180^o in the direct ($+ \pi/2$) and reverse ($- \pi/2$) directions.

The rate of coupling between the inductance coils and the resonator depends on Q factor of the dimensional resonator, number of turns in the inductance coils, on the filling of coils by core, and the cross section between the coils [5].

By transmitting the helicon waves through a semiconductor plate the effect of a resonator shape and size on resonance characteristics was investigated. It was found that $d/a < 0.2$, where a is the least value of the plate with, resonance frequency may be determined by the equation (10), used for semi – infinite resonator.

In Fig. 1 we can see calculated by (7) the frequency responses of the real and imaginary parts of the magnetic permeability μ_T of a semiconductor core - helicon resonator. The curves 1, 2, 3, 4 cover different electrons densities N of semiconductor: $2{,}3 \cdot 10^{-22}$ m^{-3} (4); $2{,}0 \cdot 10^{-22}$ m^{-3} (3); $1{,}7 \cdot 10^{-22}$ m^{-3} (2); $1{,}4 \cdot 10^{-22}$ m^{-3} (1).

On Fig. 2 typical dependences of the Hall constant R_H of n-InSb, doped with tellurium, on temperature are shown. We can see, that only for doped n-InSb (density donors $N_D = 1{,}0 \cdot 10^{23}$ m^{-3}) in temperatures T ~ 300K the Hall constant R_H and density N are stable. For doped n-InSb (density donors $N_D = 1{,}3 \cdot 10^{22}$ m^{-3}) stable range of R_H do not exceed 250 K.

If mine carriers (electrons) density $N > 2 \cdot 10^{-23}$ m^{-3}, (T = 300K), the temperature stability is better, but mobility of electrons (Fig. 2) are less 4 $m^2V^{-1}s^{-1.}$ High magnetic fields are necessary.

We can supplement nonreciprocal phase transformer with new circuit elements depends on the schematic realization of the isolator as nonreciprocal LC-filters different types. The isolators (Fig. 3) can be made-up with nonreciprocal low pas filter (L1-C2-L2), nonreciprocal high pas filter (L1-C4-L2) or with band pass or band rejection filters [7, 8].

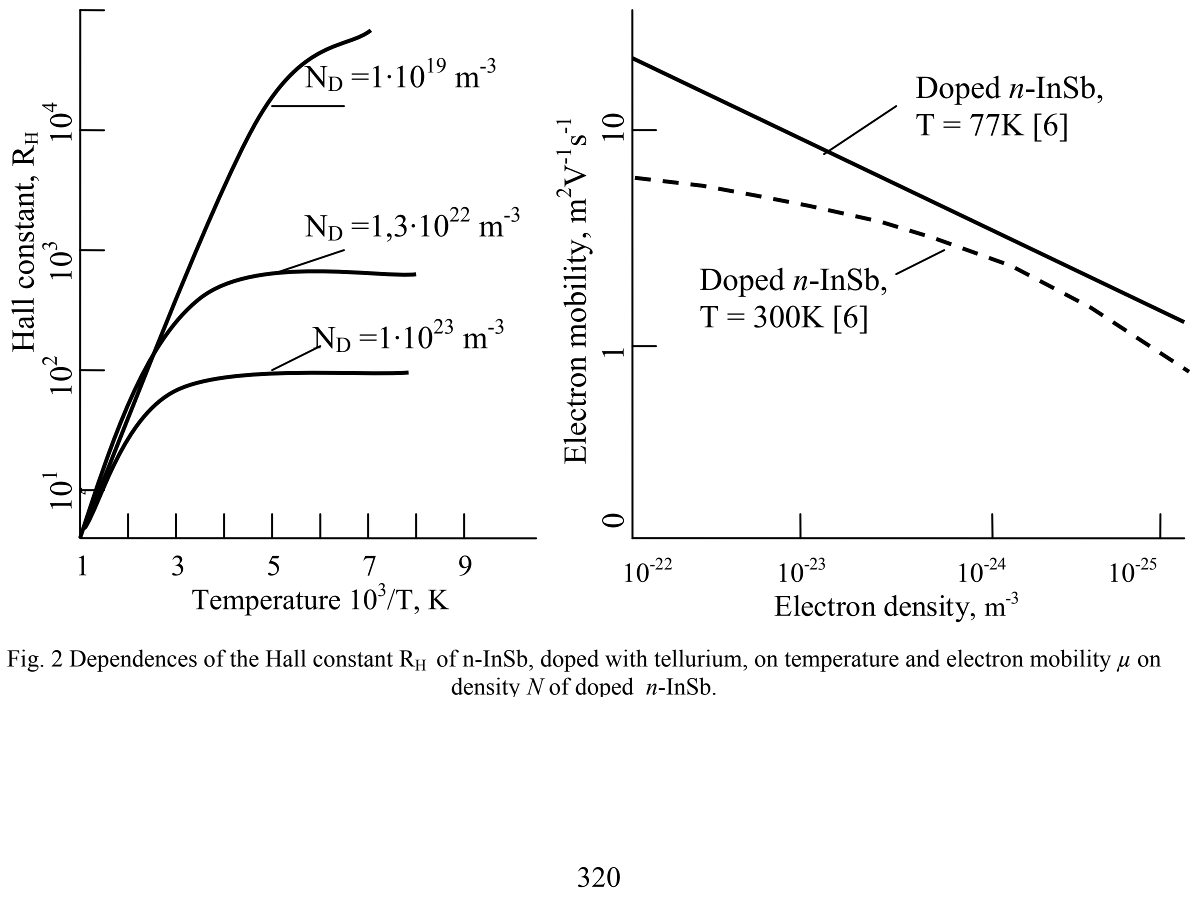


Fig. 2 Dependences of the Hall constant R_H of n-InSb, doped with tellurium, on temperature and electron mobility μ on density N of doped n-InSb.

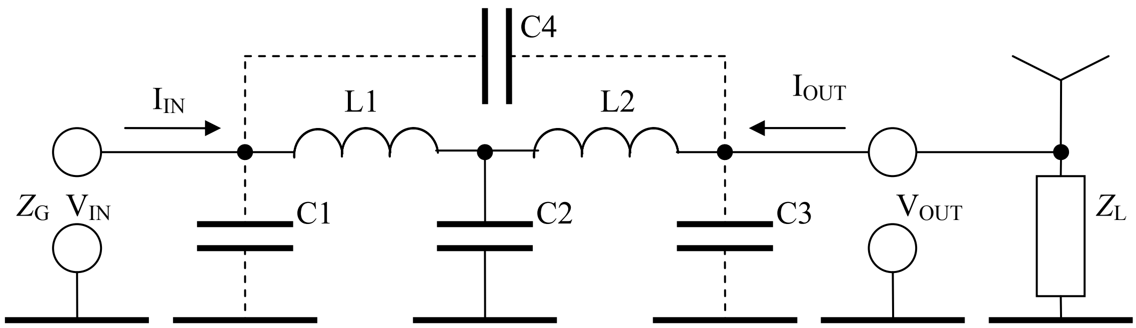


Fig. 3. Semiconductor isolator

The scattering matrix for the isolator:

$$|S| = \begin{vmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{vmatrix} = \begin{vmatrix} 0 & Ue^{i\xi} \\ 0 & 0 \end{vmatrix}, \tag{11}$$

where U is invariant of the isolator system and for ideal isolator $U = 1$, because insertion loss $\alpha = 0$.

For the nonreciprocal low pas filter (L1-C2-L2) isolator $U^2 = 1/\,(1 + 1/u^2)$. The electrical parameters of the circuit components (for example the capacitance C2) can be found by minimum loss α.

We can calculate direct loss α and reverse attenuation β in decibel by the formula.

$$\alpha, \beta = 10\lg(\mathrm{Re}P_L/\mathrm{Re}P_{G)} = 10\lg\{\mathrm{I}Z_{IN}\mathrm{I}\ \mathrm{Re}Z_L/\ [\mathrm{I}Z_T\mathrm{I}(Z_{IN} - Z_G)\}\ , \tag{12}$$

where $P_{G,}$ Z_G are power and impedance of the generator, P_L is output power and, Z_L is load impedances and $Z_{IN} = V_{IN}/I_{IN}$; $Z_T = V_{IN}/I_{OUT}$.

4. Problem of temperature stability. Choice and optimization of the material

According to (10) the density of carrier should be temperature stable. The most appropriate semiconductor is the *n*-type indium antimony (*n*–InSb) doped with tellurium. If there are higher the electron density the temperature stability better. However, an increase in concentration of donor impurity results in diminishing the electrons mobility and in the resonator quality factor. All this will cause the isolator forward loss to increase. The temperature characteristics of helicon resonators were considered for a low-band filter isolators because loss of these isolators is lower.

Based on the chosen model, we have calculated the main characteristics of the isolator: direct loss α and reverse attenuation β in the wide range of temperatures and frequencies. It has been shown that for the development of mobile non-reciprocal devices a semiconductor material of high carrier mobility $\mu > 4\ m^2V^{-1}s^{-1}$ and high mine carrier density $N > 10^{-22}\ m^{-3}$, (T = 300K), is needed. The electrical parameters of semiconductors should be stable in operational temperature range. Te – doped *n*-InSb, semiconductor alloy $Cd_XHg_{1-X}Te$ (x~0,1) and anisotropic alloy $Bi_{1-X}Sb_X$ (x~0,1) offers valuable properties.

In our calculations, we considered the frequency responses and temperature characteristics of the helicon resonator material, the helicon resonator geometry, physical parameters and dimensions of the inductance coils, characteristics of insulation between the coils, parameters of the transformer filling, electrical parameters of the circuit components, etc.

5. Conclusions

Nonreciprocal passive devices as isolators and circulator based on ferrite medium are used in microwave engineering and wireless communication. One of the problems is protection the transmitters against microwave energy reflections from aircraft antenna in decametric range of waves. The low-frequency limit of applicability of ferrite nonreciprocal devices is 50 MHz. At lower frequency range, application of nonreciprocal ferrite devices is questionable. The difficulties arise from absence of ferrite materials with the necessary properties. In the metric and decametric range of waves, the helicon devices may serve as a substitute for ferrite nonreciprocal devices. The main point is that the helicon devices use the gyroelectric effects in semicondustors instead of the gyromagnetic one in ferrites. A mathematical model describing the electrical properties of non-reciprocal quadripole has been presented. It has been shown that for the development of mobile non-reciprocal devices a semiconductor material of high carrier mobility $\mu > 4\ m^2V^{-1}s^{-1}$ and high mine carrier density $N > 10^{-22}\ m^{-3}$, ($T = 300K$), is needed. The electrical parameters of semiconductors should be stable in operational temperature range. The characteristics of the helicon isolators containing nonreciprocal transformer with helicon dimensional resonator are calculated. The experimental results are in compliance with calculations what confirms adequacy of used model.

References

[1] Final Report of NATO Research and Technology Organization (RTO) Information System Technology (IST). Panel Research Task Group IST-050/RTG-022 (or Research Task Group IST-050). June 2007.

[2] J. Gremillet. Propagation des ondes metriques et decametriques dans les semi-conducteurs et presence d'une induction magnetique continue. Effect Helicon. Ann. de Radioelect., 1964. No 76, 122.

[3] I.Polivka. Wideband UHF/Microwave Active Isolators. High Frequency Electronics. 2006, December, 70-71.

[4] Z. Jankauskas, L. Laurinavičius. Magnetic and Electric Excitation of Magnetoplasmic Waves. Electronics and Electrical Engineering. 2002. No 2(37), 32-34.

[5] L. Laurinavičius. Non-Reciprocal and Control Devices based on the Semiconductor Magnetoplasma – Vilnius: Technika, 1995. 200 p. (in Russian).

[6] Litwin-Staszewska, E., W. Szymanska, and P. Piotrzkowski, Phys. Status Solidi (b) 106, 2. 1981, 551- 559.

[7] V. S. Vountesmery, Y.V. Vountesmery. Temperature Characteristics of Broadband Helicon Isolators for Meter and Decameter Waves. IEEE Transactions on Microwave Theory and Techniques. 2007.Vol. 55. No. 10, 2097 – 2102.

[8] L. Laurinavičius. "High frequency transformer" U.S.S.R. Patent 930398 BI 19, May. 25, 1982.

Measurement of Cole-Cole Plot for Quality Evaluation of Red Wine

K. Watanabe, Y. Taka, O. Fujiwara

Graduate School of Engineering, Nagoya Institute of Technology,
Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan

Email: taka.yoshinori@nitech.ac.jp

Abstract. *For the aim of evaluating the quality of red wine, we measured the complex relative permittivity in the frequency range from 100 MHz to 40 GHz with a network analyzer, and showed that the Cole-Cole plot of red wine consists of a semicircle at frequencies above around 1 GHz and straight line at frequencies below 1 GHz, which come from the dispersion properties for the water solution of alcohol and ingredients peculiar to red wine, respectively. Based on the measured Cole-Cole plots for seven kinds of red wines made from the same brand of Merlot in different production years, Debye dispersion parameters were estimated to reveal that the alcohol concentration and ingredient property can simultaneously be estimated from the parameters for the semicircle and straight line, respectively.*

Keywords: Red wine, Cole-Cole plot, dispersion parameters, alcohol concentration, ingredients.

1. Introduction

Quality that guarantees the taste of alcoholic drinks is being evaluated by measurement of alcohol concentration, chemical analyses of ingredients and expert's sensuality test, which require a great deal of time and labour.

For the aim of evaluating the quality of red wine, we previously measured the complex relative permittivity in the frequency range from 10 MHz to 6 GHz with a network analyzer, and estimated the Cole-Cole plot parameters of red wine, which exhibited that the alcohol concentration of red wine can be estimated from the parameters for semicircle dispersion in the Cole-Cole plot, and other ingredients can also be evaluated by the parameters for straight line dispersion [1, 2].

The developed method allows to simply estimating the concentration of alcohol in drinks containing various flavour substances (ingredients) like in the wine. No fractional distillation is necessary for test. We assumed that the shape of semicircle and straight line on Cole-Cole plot should characterize the dispersion properties of the water solution of alcohol and ingredients peculiar to red wine, respectively

In the present study, to validate this assumption, we measure the complex relative permittivity of red wine, its distillation and residue after distillation in the frequency range from 100 MHz to 40 GHz, and compare their Cole-Cole plots to show how the alcohol component and other ingredients should affect the dispersion properties of the Cole-Cole plot. Based on the measured Cole-Cole plots, Debye dispersion parameters are also estimated to show their dependence on elapsed production years of seven kinds of red wines made from the same brand of Merlot.

2. Method

Figure 1 shows a setup for measuring complex permittivity with a dielectric probe. Liquid samples used for estimating alcohol concentration were pure water and dilute ethanol solution, whose complex relative permittivity was measured in the frequency range from 100MHz to 40GHz with the dielectric probe connected to a network analyzer.

The complex relative permittivity of red wine including conductive impurities [3] can be expressed as

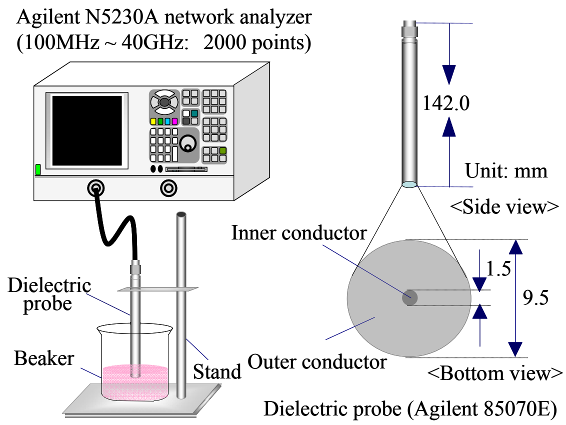


Fig. 1. Setup and configuration of dielectric probe for measurement of complex permittivity.

$$
\left.
\begin{aligned}
&\varepsilon_r^* = \varepsilon_r' - j\varepsilon_r'' = \varepsilon_{r\infty} + \frac{\varepsilon_{r0} - \varepsilon_{r\infty}}{1 + (j\omega\tau_0)^\beta} + \frac{1}{(j\omega\tau)^\alpha} \\
&\varepsilon_r' = \varepsilon_{r\infty} + \left\{ 1 - \frac{\sinh(\beta \ln \omega\tau_0)}{\cosh(\beta \ln \omega\tau_0) + \cos\frac{\beta\pi}{2}} \right\} \times \frac{\varepsilon_{r0} - \varepsilon_{r\infty}}{2} + \{\cosh(\alpha \ln \omega\tau) - \sinh(\alpha \ln \omega\tau)\} \times \cos\frac{\alpha\pi}{2} \\
&\varepsilon_r'' = \frac{\sin\frac{\beta\pi}{2}}{\cosh(\beta \ln \omega\tau_0) + \cos\frac{\beta\pi}{2}} \times \frac{\varepsilon_{r0} - \varepsilon_{r\infty}}{2} + \{\cosh(\alpha \ln \omega\tau) - \sinh(\alpha \ln \omega\tau)\} \times \sin\frac{\alpha\pi}{2}
\end{aligned}
\right\} \quad (1)
$$

where ε_{r0} is the DC relative permittivity, $\varepsilon_{r\infty}$ is the relative permittivity at infinite frequency, α and β represent the degree of relaxation distribution, τ and τ_0 are the relaxation time constants. For pure water or dilute ethanol solution without DC conductivities, the first and second terms on the right hand side of Eq. (1) are used.

Eq. (1) shows that Cole-Cole plot or ε_r' - ε_r'' curve consists of a semicircle and straight line, which can be represented by the first and second terms and the third term, respectively, on the right hand side of Eq. (1).

According to Ref. [2], the alcohol concentration of red wine is assumed to be evaluated from the parameters of ε_{r0}, $\varepsilon_{r\infty}$, τ_0 and β in Eq. (1), which was estimated in the following way. Measurement of the Cole-Cole plots was made for pure water and dilute ethanol solution, which were fitted to Eq. (1) without the third term to reveal the dependence on alcohol concentration of the above-mentioned parameters for calibration data. The parameters in Eq. (1) were also obtained from the Cole-Cole plot measured for red wine, whose alcohol concentration was estimated from the calibration data for pure water and dilute ethanol solution, and was validated by using a distillation method [2].

In order to verify the validity of the above-mentioned assumption, we compared the Cole-Cole plots measured for pure water, red wine, its distillation and residue after distillation in order to examine how alcohol component can affect these Cole-Cole plots. For red wines to be measured, we used seven kinds of Japanese red wines made from the same brand of Merlot in different production years, which were named here A, B, C, D, E, F and G.

3. Results and discussion

Figure 2 shows the Cole-Cole plots measured and calculated for pure water, wine A, its distillation and residue. Table 1 summarizes Cole-Cole parameters and their estimated values of wine A, distillation and residue. The estimated alcohol concentration of wine A was 12.2 %, which agrees well with the measured alcohol concentration (12.1%) from a distillation method. We found from Fig. 2 that there is good agreement between measurement and calculation by Eq. (1). The results also show that the pure water and distillate have only semicircles in the Cole-Cole plot, while wine A and its residue have both of semicircles and straight lines. It should be noted that the semicircles of wine A and the residue approximately agree with those for the distillation and pure water, respectively. This means that the semicircle and straight line are essentially derived from the dispersion properties of alcohol component and specific ingredients of red wine.

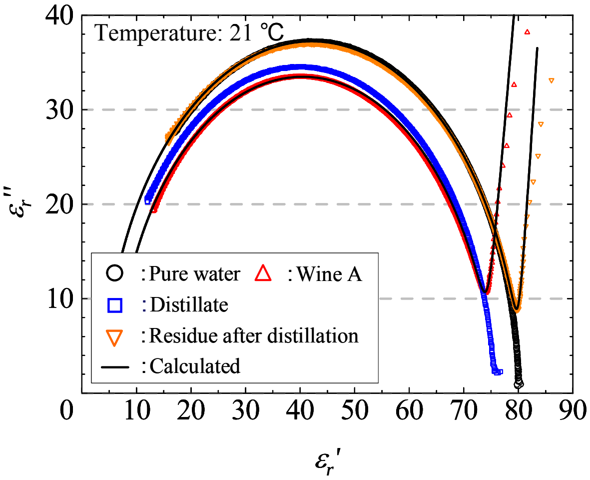


Fig. 2. Measured and calculated Cole-Cole plots.

Figure 3(a) shows the Cole-Cole plots measured for six kinds of wine B to G with the same brand of Merlot in different production years. Also shown in Fig. 3(b) is an enlargement of the Cole-Cole plots around the minimum points. These Cole-Cole parameters are summarized in Table 2 together with the estimated values of alcohol concentration. We found that all the semicircle parts almost coincide, which shows the same alcohol concentration for wine B to G despite different production years. There are some differences between the minimum points and gradients of straight lines for the production years, which suggests a possibility that ingredients peculiar to red wine could be evaluated from the corresponding Cole-Cole parameters: τ, α and f_m.

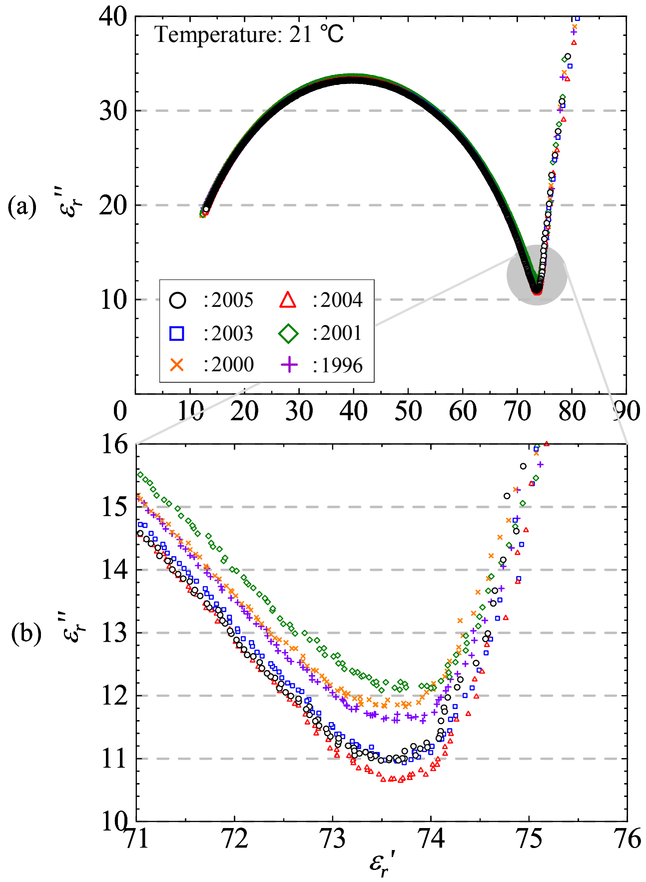


Fig. 3. (a) Measured results of Cole-Cole plots for red wines of different production years and (b) enlargement of the Cole-Cole plots around the minimum points .

Table 1. Cole-Cole parameters and their estimated values of wine A before/after distillation.

Merlot wine	ε_{r0}	$\varepsilon_{r\infty}$	β	τ_0 [ps]	τ [ps]	f_m [MHz]	α
Wine A[2006]	73.6(12.2)	6.2	0.986	12.76	25.70	798	0.910
Distillate	75.5<12.1>	4.9	0.983	12.61			
Residue	79.5	4.0	0.986	9.64	33.50	898	0.930

[ ]: Production year
( ) :Estimated alcohol concentration [%]
< > : Measured alcohol concentration with a distiller [%]
f_m: frequency at which Cole-Cole plot reaches the minimum point

Table 2. Estimated values of alcohol concentration and parameters for Cole-Cole plot of red wines.

Merlot wine	ε_{r0}	$\varepsilon_{r\infty}$	β	τ_0 [ps]	τ [ps]	f_m [MHz]	α
Wine B[2005]	73.4(12.4)	5.6	0.976	12.65	25.30	938	0.916
Wine C[2004]	73.5(12.5)	5.7	0.978	12.94	26.90	858	0.912
Wine D[2003]	73.5(12.5)	5.6	0.976	12.96	25.30	839	0.910
Wine E[2001]	73.5(12.5)	5.5	0.978	13.21	21.00	998	0.911
Wine F[2000]	73.4(12.6)	5.5	0.978	13.24	22.50	918	0.912
Wine G[1996]	73.4(12.6)	5.6	0.979	12.83	21.50	798	0.908

[ ]: Production year
( ) :Estimated alcohol concentration [%]

4. Conclusions

In comparison of the Cole-Cole plots of red wine, its distillation and residue, we confirmed that the semicircle and straight line in the Cole-Cole plot are essentially based on the dispersion properties of the water solution of alcohol and the ingredients peculiar to the red wine, respectively. The dependence on elapsed production years of the Cole-Cole parameters implies a possibility that the maturity of red wine may be evaluated from the dispersion properties of wine characterized by the straight line in Cole-Cole plot.

The future subject is to evaluate the quality of red wine from the Cole-Cole plot in the low frequency region based on TDR (Time Domain Reflectometer) measurement.

References

[1] K.Watanabe,Y.Taka, and O.Fujiwara: “Measurement of Cole-Cole Plot and Parameter Estimation for Quality Evaluation of Wine”,IEICE Technical Report,EMCJ2007-78,MW2007-125(2007-10) (in Japanese).

[2] Kota Watanabe, Yoshinori Taka, Osamu Fujiwara : Estimation of Alcohol Concentration of Red Wine Based on Cole-Cole Plot, IEEJ Trans. FM, Vol. 129, No. 5 (2009-05 to be published in Japanese).

[3] K.Baba, S.Tsunekawa, T.Fukuda, T.Takagi:“Time Domain Spectroscopy for Electrical Properties of a Slightly Reduced Rutile Crystal”Jpn. Appl. Phys. Vol.32, pp.3649-3652 (1993-8).

Innovative Digital Eyes for Objective Color Inspections and Measurements

[1]D. Hofmann, [2]G. Linss, [2]M. Rosenberger

[1]Steinbeis Transfer Center QA & QM, Jena, Germany
[2]Technical University Ilmenau, Department QA, Ilmenau, Germany
Email: dietrich-hofmann@t-online.de

Abstract. *An actual problem in research and industry, food processing and health care, traffic and environmental protection, security and administration is the implementation of objective quality inspections and measurements. A new class of smart vision sensors for dimensional and color measurements will be introduced. The advantages of an innovative programming method in natural languages with see & click instructions and solution guarantee will be presented.*

Keywords: Smart Vision Sensors, Color Measurements, Quality Inspection.

1. Introduction

An actual problem in research and industry, food processing and health care, traffic and environmental protection, security and administration is the implementation of objective quality inspection and measurements with digital image processing systems. Innovative color sensitive smart vision sensor systems are new chances especially for small and medium sized enterprises. The new vision sensor systems are characterized by small dimensions, true color sensitivities, integrated cameras and controllable lightings [1]. They are convenient, reliable and affordable. They must not be programmed in higher programming languages. The measurement functions are predefined. Programming (operation) is accomplished windows-like with intuitive icons and/or soft keys in natural language on touch screens. Nevertheless one problem must be mentioned:

Conventional image processing systems for quality inspections and measurements are operated by trained specialists. They have a profound qualification and technical documentations as usual [2]. The fast distribution of innovative vision sensor systems in the hands of lower qualified users strongly demand a new kind of technical documentations and programming software.

As a first step we developed the so called see & click operation instructions (S&COI). They are solution dependent driven by icons and soft keys. The user has only to follow red, yellow and green frames in the innovative operation and programming instructions. Red, yellow and green frames are used in analogy to traffic lights: Red (see), Yellow (wait), Green (click).

A second step to support the fast distribution of smart vision sensor systems was an open access to these see & click instructions. We applied the new possibilities of browser-based software as a service [3] and cloud computing [4].

A third step was the implementation of the see & click programming procedures in a specialized platform as a service for digital color image processing and spectral imaging.

We took the platform www.spectronet.de [5] to provide the intuitive programming instructions online which can be accessed from a web browser, while the software and data are stored on SITEFORUM servers [6].

More than 30 practical examples are given in www.spectronet.de > Akademie (academy) > BV Anwendungen (image processing applications) > Bedienungsanleitungen (operation instructions) > OMRON Xpectia from optical character recognition over data matrix code and pattern recognition through geometrical and spectral measurements. The practical operation of innovative smart vision sensors with intuitive see & click programming for color measurements will be demonstrated.

2. Subject and Methods

Smart vision sensor systems are a new kind of measuring instruments on the market. The OMRON Xpectia is a smart vision sensor system for visual quality control that includes everything from camera with integrated light sources to image-processing units (Table 1, left). With Omron's newly developed proprietary measurement algorithms, the parameters can be set through only a few steps involving the operation of touch-panel color monitors, colored functional icons and/or full-text functional soft keys in natural language and see & click programming instructions. (Table 1, right).

Table 1. Xpectia smart vision sensor system (left) and see & click programming system (right)

Xpectia vision sensor system
http://www.visquanet.de/servlets/SITEFORUM?i=1122540720771&&e=UTF-8&l=1&t=/contentManager/selectCatalog&intro=1&active=no&ParentID=1222358585640

05	A B C	M01. Zeichenerkennung / Character Inspection m01_xpectia_print.pdf m01_xpectia_web.pdf
06		M02. Winkelmessung / Circle Angle m02_xpectia_print.pdf m02_xpectia_web.pdf
07		M03. Musterklassifikation / Classification m03_xpectia_print.pdf m03_xpectia_web.pdf
08		M04. Farbflächenmessung / Gravity and Area m04_xpectia_print.pdf m04_xpectia_web.pdf
09		M05. Farbtonmessung / Color Data m05_xpectia_print.pdf m05_xpectia_web.pdf
10		M06. Farbobjekterkennung / Labeling m06_xpectia_print.pdf m06_xpectia_web.pdf

Xpectia see & click programming system (extract)
http://www.spectronet.de > Akademie > BV Anwendungen > Bedienungsanleitungen > Xpectias

The new smart vision sensor systems and the new intuitive programming methods are milestones in the digitalization of measurements and therefore a breakthrough in measurement science as well as in measurement education and training.

Visual quality inspections are fast, contactless and non-destructive. Most visual quality inspections still provided by human eyes. Visual quality inspections by man have five significant disadvantages:

1. The human inspector decides subjectively
2. The human inspector must be specialized and trained
3. The human inspector is expensive
4. The human inspector becomes tired
5. What human inspect in subjective visual perception is not simply a translation of the image on the retina. Thus people interested in perception have long struggled to explain what visual processing does to create what we actually see. This problem is unsolved till today.

Therefore an increasing interest is observable to use objective vision sensors and image processing software for visual quality measurements [7].

Special interest is focused on visual quality assurance **with color image processing** and spectral imaging. The simple reason is the huge information content of color scales in comparison with gray scales (Tab.2).

Table 2. Information content of different scales

$2^8 = 256$	$(2^8)^3 = 16.777$

Source: http://www.visquanet.de/servlets/SITEFORUM?t=/contentManager/selectCatalog&e=UTF-8&i=1122540720771&l=1&ParentID=1223122702488&active=no

Another reason is the availability of affordable color cameras under the influence of an exploding color camera market for consumer goods. The production of industrial color cameras grew by 13% in 2008 and has topped first in history the monochrome cameras [8]. But design and programming of industrial cameras are different to consumer goods. Therefore the see & click programming instructions has been elaborated for industrial color and other measurements.

3. Results

To understand the philosophy of the new programming method selected instruction steps for color inspections with the smart vision sensor system Xpectia are shown (Table 3).

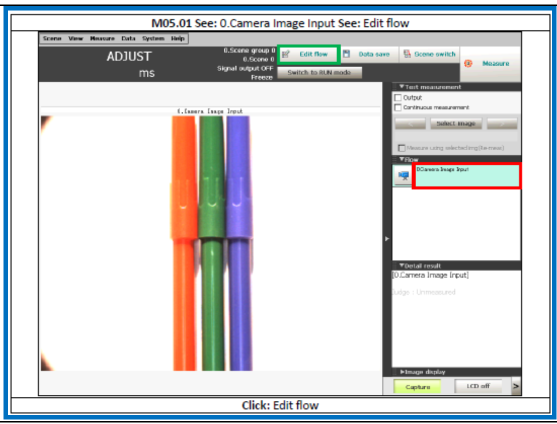
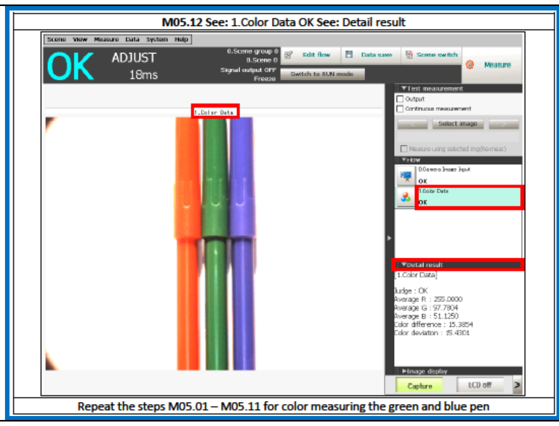
The task is to measure the color of the red pencil.

The results of the measurements are displayed on the graphical user interface in Table 1. The measurement results concerning the red pencil are:

- Average R(ed): 255.0000
- Average G(reen): 97.7804
- Average B(lue): 51.1250.

For measurements of the green and the blue pencils the same procedures have to be run again.

Table 3. Selected instruction steps for see & click instructions in color quality measurements

	
Step 1 of programming the color measurement with Xpectia See: Red frame; See: Camera Image Input & wait; See: Green frame; See:Edit flow & click Edit flow	**Step 12** of programming the color measurement with Xpectia See: Red frame: See Color Data OK & wait, See: Red frame: See Detail result & the END

4. Conclusions

In the paper has been shown that innovative instrumentation and programming facilities are enriching and changing the fundamental methods in measurement science and education.

Smart vision sensor systems and see & click programming instructions making the visual quality control objective, convenient, reliable and affordable. Together with cloud computing and a specialized platform as a service we are at the beginning of a new era for the application of external eyes in research and industry, food processing and healthcare, traffic and environmental protection, security and administration.

Acknowledgements

Thanks are due to Ralf Lattuch from OMRON EUROPE B.V. for his reliable collaboration. The NEMO project has been financially supported by Bundesministerium fuer Wirtschaft und Technologie (BMWi) under FKZ 011902N.

References

[1] J. Oelsner, „More accurate, faster and more reliable (Genauer, schneller, zuverlaessiger)“, INSPECT 9, 2008, H. 4 pp. 85-87

[2] D. Juhl, „Technical Documentation (Technische Dokumentation)”. 2. Aufl. Springer, Berlin, Heidelberg, New York 2005

[3] Wikipedia, the free encyclopedia, http://en.wikipedia.org/wiki/Software_as_a_service

[4] Wikipedia, the free encyclopedia, http://en.wikipedia.org/wiki/Cloud_computing

[5] D. Hofmann, “spectronet & visquanet visual quality assurance (Visuelle Qualitaetssicherung)”, platform with web2 technologies, www.spectronet.de

[6] D. Schlenzig, SITEFORUM Smart up your business http://www.siteforum.com

[7] M. Rosenberger, Digital image processing of the next generation for intelligent measurments and quality assurance (Digitale Bildverarbeitung der nächsten Generation für die intelligente Messtechnik und Qualitätssicherung) www.tu-ilmenau.de/qualimess

[8] S. Zimmermann, The shroud is doffed (Den Schleier gelüftet). INSPECT 10, 2009, H. 1-2, p 5.

The Use of Colour Image Properties in the Classification of Transparent Polymeric Foils

R. Ravas, L. Syrová, J. Grman

Department of Measurement, Faculty of Electrical Engineering and Information Technology, Ilkovičova 3, 812 19 Bratislava, Slovakia
Email: rudolf.ravas@stuba.sk

Abstract. *The contribution treats the problem of experimental analysis of statistical properties of colour image of visualized optically transparent polymeric foil by means of schlieren method. Intensity images of red, green, blue component and grey-scale representation of colour image were used. Subject of the analysis was recognizability of the type of foil by the features calculated from characteristics of histogram of intensities level and intensities co-occurrence matrix of foil image. Minimal numbers of incorrect classified foil samples into feature convex subspaces were compared. Subspaces were separated by the linear hyper-spaces. Experiment showed that number of errors of foil sample classification depends on the used colour component, eventually on the combination of colour components and combination of characteristics describing the image. The use of colour images increases reliability of the classification of the foil type.*

Keywords: polymeric foil, classification, colour images

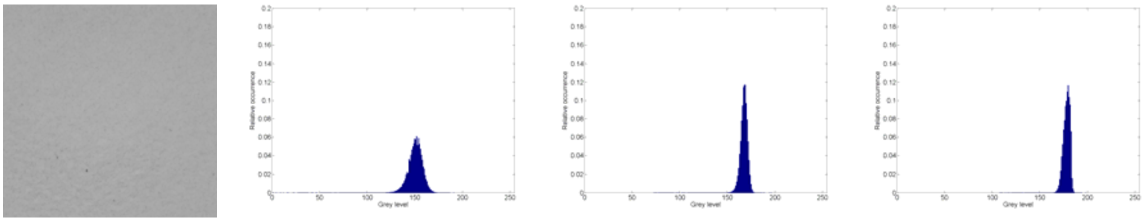
1. Introduction

Optical visualization methods are sensitive to the changes in the absorption of photons during their transmission through mechanically loaded regions of transparent materials such as optically transparent polymers and, especially, polymeric foils. The methods use the fact that the material density is a function of the refractive index of tested foil. The intensity of photon absorption changes especially in the area of deformation, which results in a change of the refraction index. The change of the refraction index can be visualized [1]. Therefore the main advantage of all visualized methods is, that they provide useful information for further processing after photographic or digital recording. Considering the available instrumental facilities and required sensitivity of measurement, we have chosen the schlieren method. The schlieren system [2] enables to measure the amount of light deflection generated by a transparent optical phase object. When there is a disturbance in the optical path the light is deviated from its nominal course in the absence of refractive index variations. Optical disturbance in the test object will produce variations of recorded light intensity that are a measure of the deflection experienced by the light in the test object. Obtained results can be processed using methods of image processing. Results of papers [3], [4] refer to possibilities of schlieren visualization method in the field of the diagnostics and classification of optically transparent polymeric foils by means of statistical methods.

2. Image Processing of Colour Visualized Polymeric Foils

At the experiments whose results already have been published, the image characteristics were calculated by different methods only from the greyscale images. One of the problems of experimental analysis is to find whether numerical characteristics calculated for the individual colour components would have similar properties (values) to the characteristics of greyscale image. It can be of course expected that properties wouldn't be the same because of two reasons. At first the source of homogenous light won't have the same intensities of RGB components after decomposition and for the second schlieren method is based on the visualization of anomalies due to light beam deflection caused by the change of refractive index in tested material. The angle of deflection depends on the value of refractive index and wavelength of light beam. The purpose of experimental analysis is to find, whether the numerical characteristics calculated for the different spectral component would be redundant to the characteristics calculated for the greyscale image or whether they would be useful for the classification.

In Fig.1 are presented sample of corrected image of visualized polymeric foils and its histograms of colour components.



a. b. c. d.

Fig. 1. a/ Sample of corrected image of visualized polymeric foil KXE30, b/ Histogram of intensity of blue c/ of green d/of red component

One of the groups of classifiers are signature classifiers. They are based on the object description by the vector $\boldsymbol{x}$ of numerical characteristics x_s belonging to the classified object. Purpose of the classification is on the basis of the vector $\boldsymbol{x}$ describing the object to classify the object into one of the finite set of classes ω_r. Under the class we understand the subset of objects with the common property. In our case the common property is the type of foil of classified sample represented by the image. Objects are assigned to the class ω_r on the basis of decision rules. For the analysis of classifiability of foils we use simple decision rule (1)

$$d(\mathbf{x},\omega_r) = \arg\min_r \sum_{s=1}^{S} \frac{(x_s - E_{rs})^2}{D_{rs}} \qquad (1)$$

where

E_{rs} mean value of feature x_s of images of foil from the class ω_r

D_{rs} variance of feature x_s of images of foil from the class ω_r

The rule (1) separates the S- dimensional signature space into disjunctive convex subspaces bounded with linear hyper-planes. In the case when the signature subspace is separated in such way that all patterns of the objects belong to the subspace for their class we reach 100% successfulness of classification. In opposite case we can use either the statistical approach of classification or the space can be separated by the nonlinear hyper-planes.

3. Results

The effect of colour of light on the properties of classifier was experimentally examined. We used 6 samples of foils. One sample (OOL30) belongs to foils with marked structure of foil image. For 5 of other samples the indistinctive image structure is characterized by weak changes of intensities that are typical for very thin foils. Recognition of this group is more difficult what is caused by the imperfection of intensity correction of schlieren apparatus and by the presence of small deformation on the foil surface. Proportion in deformations intensities of foil images is very small for foil with marked structure. There were 16 samples of foils KXE20, OOL30, 15 samples KXT21, ON25, 14 samples ONE21 and 6 samples KXE30 in the set of analyzed samples. We used a colour digital camera with resolution 5 mega-pixels. The signatures were calculated from the part of images with dimensions 1400x1400 pixels. The signature vector consists of 13 components for each colour component (r, g, b) and greyscale, together S=52. Signatures were calculated from histograms of intensities level and from co-occurrence matrix [5]. In table 1 is a survey of characteristics,

Table 1. Survey of characteristics calculated from histograms and co-occurrence matrix and their indices

s	Feature	s	Feature	s	Feature
1	$m=\sum_i ip(i)$	6	$\sum n^2 \left\{\sum_i \sum_j q(i,j)\right\}_{\lvert i-j\rvert=n}$	11	$\sum_i \sum_{j\neq i} q(i,j)/(i-j)^2$
2	$\sigma=\sqrt{\sum_i (i-m)^2 p(i)}$	7	$\left(\sum_i \sum_j (ij)q(i,j)-m_C^2\right)\Big/ s_C^2$	12	$\sum_i \sum_j q(i,j)/(1+\lvert i-j\rvert)$
3	$\left(\sum_i (i-m)^4 p(i)\right)\Big/\sigma^4$	8	$-\sum_i \sum_j q(i,j)\log_2 q(i,j)$	13	$\sum_i \sum_j (i+j-2m_C)^2 q(i,j)$
4	$m_c=\sum_i i\sum_j q(i,j)$	9	$\sum_i \sum_j q^2(i,j)$		
5	$s_C=\sqrt{\sum_i .(i-m_C)^2 \sum_j q(i,j)}$	10	$\max\{q(i,j)\}$		

where

i, j intensity level of image i, j $\in <1, I_{max}>$

p(i) probability of occurrence of intensity level i in the image

q(i,j) element of intensity co-occurrence matrix defined by (2)

$$q(i,j)=\frac{Q(i,j)}{\sum_i \sum_j Q(i,j)} \qquad Q(i,j)=\#\{[(x,y),(u,v)] \mid y=v, \lvert x-u\rvert=1, I(x,y)=i, I(u,v)=j\} \qquad (2)$$

where

\# denotes number of elements of set

I(x,y) intensity function of image

Signature space was normalized for all coordinates to the interval <-1, 1>. At first subspaces they were sensed for each component separately. It is not possible to realize systematic search of the whole space in regard to the number of the possible combinations 2^{52}. We should find such a combination of features that would reach minimal non-successfulness of classification at the minimal number of features. One of the possible results for each type of image is presented in table 2. Table contains results of several experiments. In the first part of the table there are numbers of incorrect classified foils for the classes marked by the index 1-KXE20, 2- KXE30, 3- KXT21, 4- ON25, 5- ONE12, 6- OOL30.

Those marked features, whose combination with the minimal number of features gave minimal number of non-correct classified foils samples are listed in the second part of the table. The total number of non-correct classified foils for the features calculated from red, green, blue, grey images and their combination are presented in the third part. The value behind the slash means the number of non-successfully classified samples in the case that all 13 features were used.

Table 2. Number of incorrect classified foil samples into the individual classes for minimal number of features and for minimal total number of incorrect classified samples

	Index of class						Index of selected feature													Sum of
	1	2	3	4	5	6	1	2	3	4	5	6	7	8	9	10	11	12	13	errors
Red	1	4	3	8	0	0	1	1	1		1	1								16/18
Green	0	2	1	2	1	0			1				1			1	1			6/10
Blue	1	1	2	1	1	0	1	1	1		1		1							6/10
Gray	0	2	1	4	1	0	1		1			1		1						8/12
Blue& Green	0	0	1	1	0	0			1				1	1		1 1	1 1			2

4. Conclusions

Better results were reached for all types of images when compare the case of selected features and the case when all features have been used. The decomposition of the signature space to the convex subspaces separated by the linear hyper-planes was more successful for intensity images from blue and green component of acquired colour image. When we compared results for grey-scale image (that was formally used by means of all features when the non-successfulness of the correct classification was 14,6 %) the non-successfulness dropped to 2,4% at the selected combination of features of green and blue component. Experiments showed that the statistical properties of visualized images of optically transparent polymeric foils were dependent on the choice of the colour component of the image.

Acknowledgements

We gratefully acknowledge that the results presented here were solved as partial goals of the research task VEGA No. 1/0480/09.

References

[1] Smith A. J., Lim T. T. Flow Visualization. Imperial College Press, London, 2000.

[2] Bolf J., Bajcsy J., Bolf P.: Methods of the Properties Testing of Optically Transparent Materials. *Journal of Electrical Engineering* , 44 (8): 244-247, 1993.

[3] Syrová L., Ravas R., Jambrich, M., Grman J. Use of Statistical Approach for Image Analysis of Visualized Transparent Polymeric Foils. *Vlákna a textil,* 12(4): 177-184, 2005.

[4] Syrová, L., Ravas, R., Grman, J. The Use of Schlieren Visualization Method in the Diagnostic of Optically Transparent Polymeric Foils. *Journal of Electrical Engineering*, 58, (5): 257-264, 2007

[5] Haralick, R.M., Shapiro, L.G. Computer and Robot Vision. Addison-Wesley Publishing Company, New York, 1992.

Dynamic Evaluation of Pedestrian Flows by Traffic Line Distributions

S.Takayama, T.Iwata, H.Nishikado, K.Kariya

Ritsumeikan Univ.BKC, Shiga, JAPAN
Email: s-tkym@se.ritsumei.ac.jp

Abstract. *This paper describes the system to take traffic lines of pedestrian flows in traffic field and the investigation of dynamic characteristics of the flows by using the time transition of the density, direction and velocity distribution of them. The dynamic characteristics are so useful to evaluate the congestion and latent dangerousness of traffic field and also the arrangement of structures and public facilities.*

Keywords: Traffic Lines, Pedestrians Flow, Congestion Evaluation

1. Introduction

To realize smooth and safe walking of people outdoors in city or town, it is necessary to know latent characteristic (congestion and dangerousness) of the field where they are walking. The latent characteristic of field is estimated by analyzing pedestrian flows. Pedestrian flows are affected by the formation of road and square, field conditions, the arrangement of structures (buildings, houses, railway stations and so on), the position of entrances and exits of them and the public facilities (stairs, monuments, bus stops, crossing points, fountain and so on) . It is recognized that these are static factors. Additionally, pedestrian flows are also affected by the congestion density, walking directions and velocities of themselves. The direction, velocity and density of pedestrian flows change dynamically by time transition. It is recognized that these are dynamic factors. The dynamical state of pedestrian flows expresses the changing of congestion and the latent dangerousness of field where they are. To know the latent state of field is very useful to predict frequent accidents for pedestrians and to redesign the arrangement of structures/facilities and the formation of road/field.

This paper describes

1) the system to find the traffic lines of pedestrian flows and to construct the density, direction and velocity distribution of them in open square and street,
2) the static state (congestion) evaluated by density, direction and velocity distributions of traffic lines in a constant measurement time,
3) the dynamic state (congestion) evaluated by monitoring the time transition of density, direction and velocity distributions of traffic lines.

2. Measurement of Traffic Lines of Pedestrian Flow

To recognize pedestrians outdoor, several kinds of research works are reported[1-5]. To recognize the pedestrian flow in traffic field, this paper shows the measurement system to detect the moving track of pedestrian flow as a traffic line. Traffic lines were generated by image processing of the pedestrian flows recorded with digital video camera. Fig.1 shows the sample of frame image of video taken the situation of pedestrians walking in morning at Student University Festival in autumn 2007 in morning. In Fig.1, the right-top side area (circle_1) is the entrance of restaurant building. In this area, it was found that standing students and walking students in/out the building are confusing. And the five areas shown with circle_2 are street food stalls. In these areas, the students standing in the food stall and walking slowly to see foods and goods for sale were confusing. Around these areas, the density of students is higher and the walking velocity of them is lower. Video camera was fixed on wall on the roof of building (21m height). The recording was continued in several hours from morning. The record was stored into several video tapes. By special image processing, pedestrian movements in the video image were expressed as the connections of traffic line segments. In this experiment, the image processing was done by off-line. The image processing was done by 7Hz. The data of traffic line segments were stored to internal memory (RAM) and external database in storage (HD) each 1/7sec. At the same time, each 5min., the picture of traffic lines superimposed on static image (Fig.4) was stored as a JPEG file. By using these data, the dynamical situation of pedestrian flows and the effect of facilities/ field characteristics were evaluated in consideration on the density distribution, direction distribution and velocity distribution.

Fig.1 Static Video Image


Fig.2 Differential Image Frame


Fig.4 Traffic Lines of Pedestrians (5 min.)

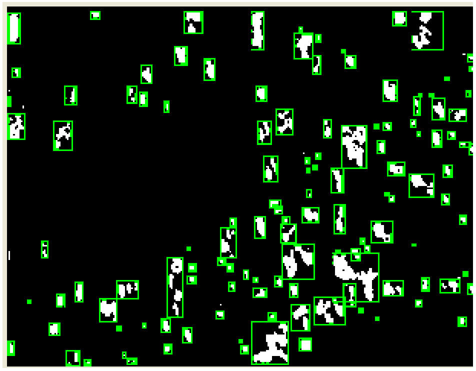


Fig.3 Recognition of Pedestrian Movement with Green Boxes

3. Density Distribution of Pedestrian Flows

The static congestion of pedestrian flows is evaluated with the density distribution of traffic lines. Fig.5 shows the image of absolute density distribution of traffic lines. The density distribution was constructed by counting the number of traffic line segments in each divided image block (10 dots square). Thick color areas show the state of high density of traffic lines. The density is expressed by 8bit. It is confirmed that entrance area of restaurant building (circle_1) and the street food stalls (circle_2) was crowded. The reason of this state is by students standing and walking slowly around the entrance of restaurant building and the street food stalls. In addition to that, upper and right parts at edge (dot circle_3 and circle_4) of lawn area centered in the image are also crowded. This means that pedestrians were walking slowly at edge of lawn, especially at the corner. Their behaviors became carefully near the off limit area like lawn. Around the bottom, there is high density area formed thick straight line (dot circle_5). Here is the Main Street of campus. Many pedestrians walked through freely not affected by area and facility characteristics. As mentioned above, the situations of density distributions of the characteristic four areas were expressed. But these density distributions were not constant. These situations are invisible in a static image and changes according to time transition.

4. Direction Distribution of Pedestrian Flows

Fig.6 shows the image of direction distribution of traffic line segments measured in 5 min. at 13:00. The directions of traffic line segments are expressed with the color distribution referred by Fig.7. In Fig.6, there are many traffic line segments drawn to various directions. Then, to clarify characteristics of direction distribution, the colored traffic line segments were classified into four regions as follows;

1) right side region : -45 to +45 [degree],
2) upper side region : +45 to +135 [degree],
3) left side region : +135 to +225 [degree],
4) down side region : +225 to -45 [degree].

According to this classification, Fig.6 was divided into four kinds of distinguished direction distributions (Fig.8). Around left and right side of lawn (dot circle_4 and 6), most of pedestrians were walking to the direction of left-down (colony of green line segments). Also, around upper side of lawn (circle_3), most of pedestrians were walking to the direction of left (colony of light blue line segments). In street at the bottom of image (circle_5), two kinds of pedestrian flows to right and left side were confused. In other regions, various kinds of directions were overlapped. That means that the pedestrians who walked to various directions were crossing.

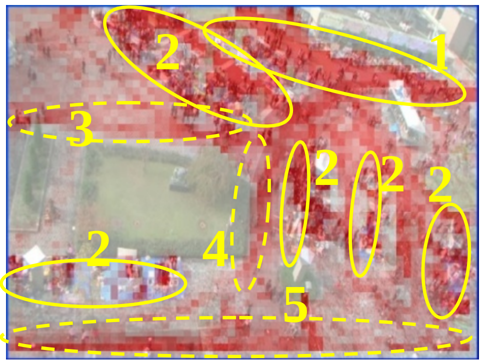


Fig.5 Density Distribution of
Traffic Lines

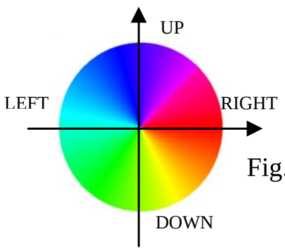


Fig.7 Matching Circle between
Direction and Color

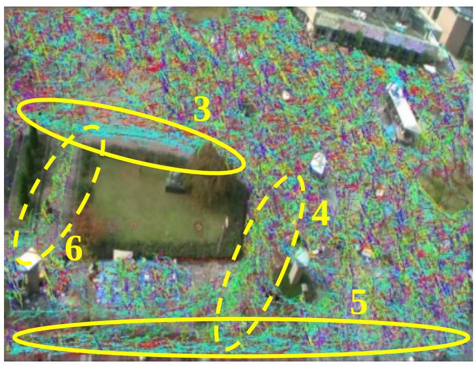


Fig.6 Direction Distribution of
Colored Traffic Lines

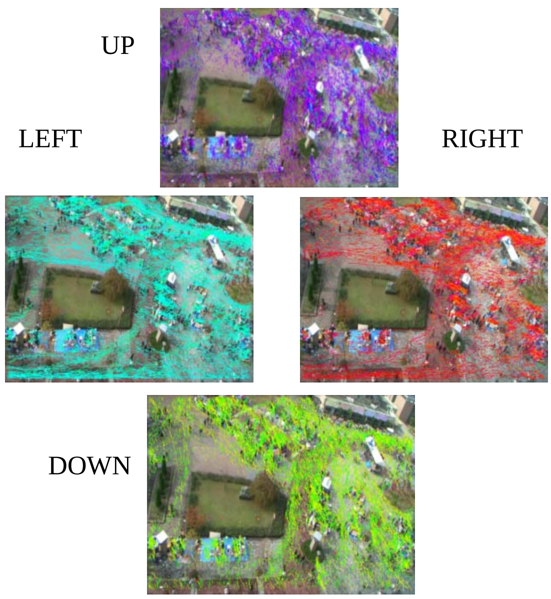


Fig.8 Direction Distributions classified into Four
Kinds of Regions (UP, DOWN, LEFT, RIGHT)

5. Velocity Distribution of Pedestrian Flows

Fig.9 shows the image of velocity distribution of pedestrian flows in 5 min. at 09:45. The velocity distribution of walking pedestrians was constructed by dividing length of traffic line segments by flame rate (1/7 sec.) in each image block (10 dots square). Thick color areas show the state that pedestrians were walking by high speed. The velocity was expressed by 8bit. In morning, as pedestrians were not so much, the density of velocity distribution was low in average. The high and low strengths of velocity were express well.

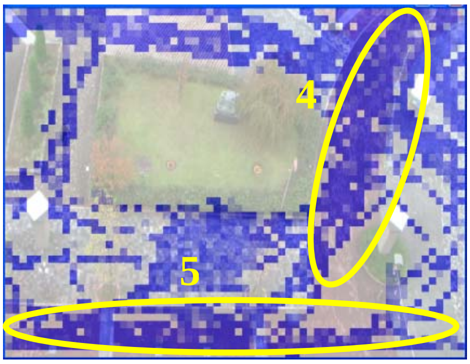


Fig.9 Velocity Distribution of Pedestrian Flows

6. Conclusions

This paper describes the measurement system to detect traffic lines of pedestrian flows and evaluates the latent characteristics (dynamic states of density, direction and velocity) of pedestrian field. These characteristics are not able to recognize with video and static images. They relate to the congestion and latent dangerousness in pedestrian field/area. They are not defined by only arrangement of buildings and facilities. They must be recognized as the dynamism of density, direction and velocity distribution of pedestrian flows.

By experiments and evaluations to analyze pedestrian flow in university campus, some invisible congestions and dangerous areas were confirmed. The situation of these areas were changing corresponding with the time transition. In future, authors will reconstruct the measurement system as an on-line system, and investigate the definition of the congestion and dangerousness of field/area estimated by the dynamism of density, direction and velocity distribution of pedestrian flows.

References

[1] O.Masoud and N.P.Papanikolopoulos, “Robust Pedestrian Tracking using A Model-Based Approach”, Proc. of the IEEE ITSC 1997, pp.338-343, 1997.

[2] G.Ma, S.B.Park, A.Loffe, S.M.Schneiders and A.Kummert, “A Real Time Object Detection Approach Applied to Reliable Pedestrian Detection”, Proc. of the 2007 IEEE Intelligent Vehicles Symposium, pp. 755-760, 2007.

[3] H.Fujiyoshi, T.Komura, I.E.Yairi, K.Kayama and H.Yoshimizu, “Object Detection Based on Temporal Differencing for Pedestrian ITS and Its Evaluation”, Trans. of IPSJ, Vol.45, No.SIG 13, pp. 11-20, 2004.

[4] T.Kinugawa and M.Ishimura, “Simple Noise-resistant Method of Motion Detection by Interframe Difference – Elimination of Randomly Fluctuating Backgrounds by Checking Space-Time Connectivity of Extracted Pixels -”, IEICE Technical Report, CS2007-54, IE2007-137, pp. 1470152, 2007.

[5] D.M.Gavrila, “Pedestrian Detection from a Moving Vehicle”, Proc. European Conf. Computer Vision, Vol.2, pp.37-49, 2000.

Active Infrared Thermography in Non-destructive Testing

M. Hain, J. Bartl, V. Jacko

Institute of Measurement Science Slovak Academy of Sciences,
Bratislava, Slovak Republic
Email: hain@savba.sk

Abstract. *The contribution deals with state of the art non-destructive testing method – active infrared thermography. The method can be used for revealing of defects and inhomogenities inside the bodies, e.g. bubbles and cracks in materials, which can have crucial importance from the point of view of their physical-mechanical properties. Goal of the research is theoretical evaluation of the performance of the pulse active thermography as non-destructive testing method and its experimental validation.*

Keywords: temperature, emissivity, heat propagation, Finite Element Method, subsurface defect

1. Introduction

During the past years both passive and active infrared thermography have become powerful and effective tools in a wide range of applications, including science, medicine, industry and non-destructive testing. More common from these two methods is the passive IR thermography, where the thermal radiation emitted from object's surface (Planck's law) is scanned by infrared camera, and gives information about the surface temperature of the object in thermal equilibrium. On the contrary, by the active infrared thermography the object under test is thermally excited - usually irradiated by a source of infrared radiation. In this case the object is in non-equilibrium state and the transient temperature field measured by thermographic camera can provide information about the thermo-physical properties, defects and inhomogenities inside the object. Contemporary the main application fields of passive IR thermography can be found in civil engineering (inspection of thermal insulation of buildings), predictive maintenance (e.g. inspection of pumps, motors, bearings and electric insulators), medicine (diagnosis of diseases), agriculture and in a variety of research experiments. The active thermography is mainly used in non-destructive testing to reveal defect (e.g. bubbles, cracks) inside various materials and objects without harming them.

2. Pulse active infrared thermography

There are several techniques used in the active thermography, the main of them are pulse thermography PT and modulated (lock-in) infrared thermography MT. Combination of these two approaches is pulse phase infrared thermography [1]. In the pulse thermography the object under test is during a limited time thermally excited by an infrared source of radiation (see Fig.1). Some part of the IR radiation incident on the object's surface is absorbed (depends on the emissivity of the surface) and transformed into a thermal energy, which propagates by thermal diffusion from surface to the inward of the object. The thermal energy propagation within the

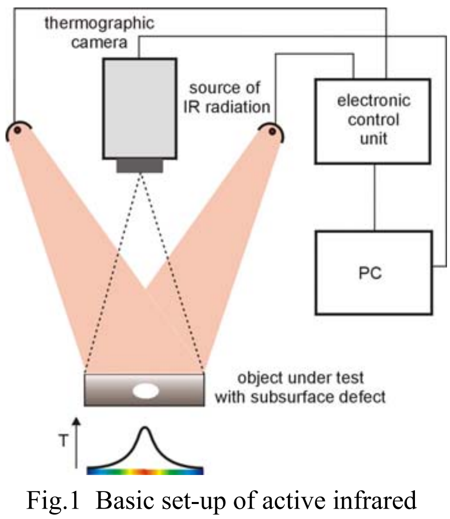


Fig.1 Basic set-up of active infrared thermography

material can be described by Fourier's partial differential equation

$$\rho.C_p \frac{\partial T}{\partial t} + \nabla.(-k.\nabla T) = Q \tag{1}$$

where T is temperature (K),
t is time (s),
Q is supplied thermal energy (J),
ρ is mass density of the material ($kg.m^{-3}$),
C_p is specific heat capacity of the material at constant pressure ($J.kg^{-1}.K^{-1}$),
k is thermal conductivity ($W.m^{-1}.K^{-1}$).

After the thermal excitation and in a simplified case (isotropic material), the Fourier's equation can be written as:

$$\frac{\partial T}{\partial t} = \alpha.\nabla^2 T \tag{2}$$

where

$$\alpha = \frac{k}{\rho.C_p} \tag{3}$$

is thermal diffusivity of the material ($m^2.s^{-1}$).

The principle of active infrared pulse thermography can be described as a measurement of time evolution of surface temperature differences, which arise as a result of different (e.g. reduced) thermal diffusion inward the object in places, where subsurface defects are present. In other words, subsurface defects change thermal diffusion rate, mathematically described by Fourier's equation (2), and therefore subsurface defects appear as surface areas of different temperature with respect to normal areas.

It is apparent that the thermal diffusion is not stationary and therefore exist some lower limit of observation time τ, when the effect (surface temperature differences) arise after thermal activation of the surface. This time is approximately equal to the time interval of the thermal pulse propagation from the surface to subsurface defect and can be estimated from the Fourier's equation [1]

$$\tau \approx \frac{z^2}{\alpha} \tag{4}$$

where z is depth of the subsurface defect.

For selected materials the estimated propagation times τ for depth of the subsurface defect z=5mm, together with other thermo physical constants, are introduced in the Tab. 1.

material	thermal conductivity k ($W.m^{-1}.K^{-1}$)	mass density ρ (kg/m^3)	specific heat capacity C_p ($J.kg^{-1}.K^{-1}$)	thermal diffusivity α ($m^2.s^{-1}$)	propagation time for z=5mm τ(s)	emissivity 8-14 μm ε
aluminium	250	2700	870	1.06E-04	0.23	0.09
gypsum	0.47	1150	1090	3.75E-07	67	0.85-0.94
concrete	1.4	2000	1050	6.60E-07	38	0.85-0.94
epoxy	0.35	1300	970	2.78E-07	90	0.8

Tab.1 Estimated times of the heat pulse propagation to the subsurface defect with a depth 5mm.

3. Thermal modelling with the use of Finite Element Method

For the optimisation of the active IR thermography parameters (the amount of thermal excitation energy, duration of the pulse, observation time), and for the correct interpretation of the measured results it is necessary theoretically describe relevant thermo-physical processes running in the object, including absorption of incident thermal energy, propagation of the heat in the inhomogeneous environment with defects and thermal losses caused by conduction, convection and radiation to the surrounding. In very simple cases it is possible to find analytical solution of the problem, but in the real 3D complex objects it is necessary to use numerical solution of the Fourier's differential equation. One of the suitable methods is the finite element method – FEM [4]. The basic concept of this method is to divide the area of solution into smaller parts called finite elements, connected at nodal points (mesh generation). In each node the physical equilibrium equation is defined, and on the object's boundary the boundary conditions are defined. In this way the physical problem described by partial differential equation PDE is break down into linear system of equations. For thermal FEM simulation of active thermography method the Comsol Multiphysics program was used [4]. The geometry of the 3D object for simulation purposes was chosen as it is shown on Fig.2a,b.

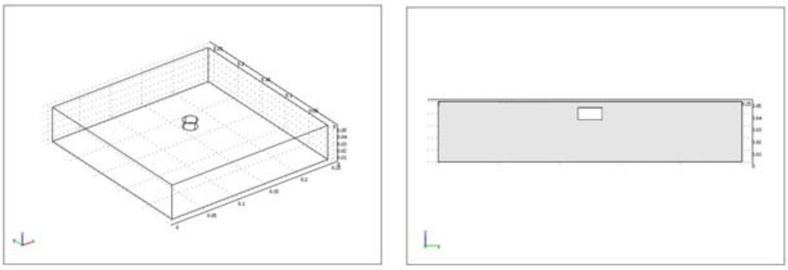


Fig. 2a,b Geometry of the object with subsurface defect

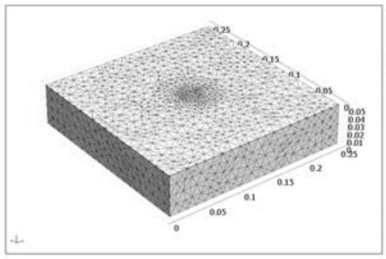


Fig. 2c Finite elements – 3D mesh

The generated mesh (see Fig 2.c). was created from 74843 tetrahedral finite elements

For the heat propagation model the PDE equation (1) was used with the following assumptions:

1. heat energy is exchanged between object and surrounding only through upper side of the object

2. upper side of the object is irradiated by the IR radiation flux 500 W.m^{-2}

3. the energy equilibrium through the upper side of the body is described by the border equation

$$-\mathbf{n}.(-k.\nabla T) = \varepsilon.\Phi + \varepsilon\sigma\left(T_s^{\,4} - T^4\right) + h(T_a - T) \qquad (5)$$

where T is temperature of the object,

T_s is temperature of the surrounding (background),

T_a is temperature of the air layer next to object,

ε is emissivity of the object's surface,

h is heat transfer coefficient (W.m^{-2}.K^{-1}).

As a result of model solving we obtain the transient temperature distribution on the object's surface with the geometry shown on the Fig.2, irradiated by the defined IR radiation. Model was solved for various materials (construction materials, metals, plastics, etc.), presented results bellow were obtained for construction material gypsum (see Fig. 3 and Fig.4).

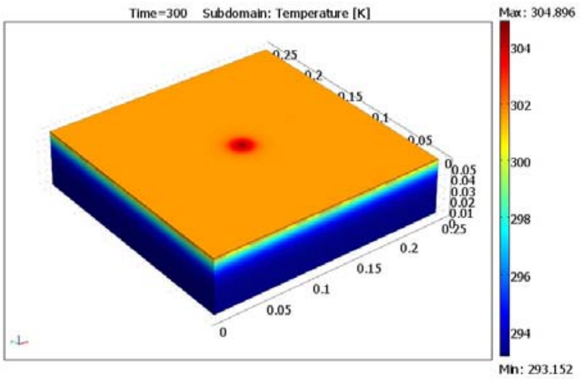


Fig. 3 Temperature field of the object with subsurface defect irradiated by IR radiation – pseudo-colour presentation of FEM model results

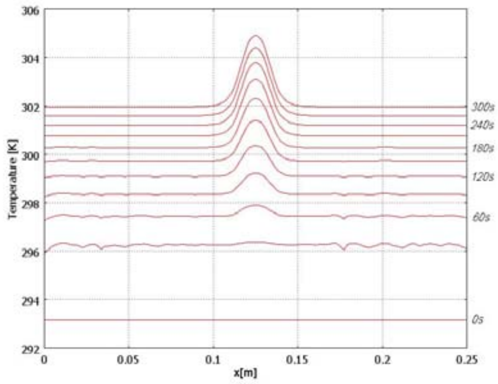


Fig. 4 Time evolution of the temperature along the surface line of the object with subsurface defect irradiated by IR radiation – results from FEM model

To find the optimal time for observation, we will evaluate thermal contrast defined as:

$$C_T(t) = \frac{T_d(t) - T_s(t)}{T_s(t) - T_s(0)} \tag{6}$$

where $T_d(t)$ is surface temperature of defect area,
$T_s(0)$ is surface temperature of sound area before heating pulse is applied,
$T_s(t)$ is surface temperature of sound area at the time t.

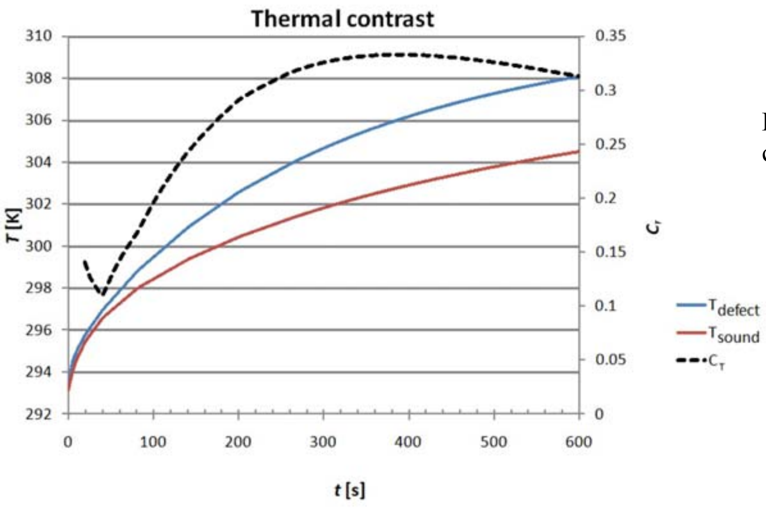


Fig.5 Time dependence of the thermal contrast between defect and sound area

Analysis of the thermal contrast shows in this case local extreme at the time 360s after start of the heating pulse.

4. Experimental results

To prove the theory and FEM modeling results, the experimental active thermographic measurements were done. Set-up of the measuring system was similar as on the Fig.1, thermographic camera NEC San-ei Thermo Tracer TH7102WX was used. The camera is equipped with micro-bolometric array 320x240; camera's instantaneous field of view is 1.6 mrad. Distance of the camera from the object under test was 0.625 m; it corresponds to the spatial resolution 1 mm on the object's surface. For the heat pulse generation two IR panel heaters were exploited. As our team is involved in the non-destructive testing of cultural heritage (frescos, wall paintings); in the first stage the emphasis in the experiments was put on the testing of building materials like plaster, gypsum, concrete, bricks, etc. The testing objects with subsurface defects (filled with air) with the shape as shown on the Fig.2, were made from various constructive materials. The subsurface defect has a cylindrical shape, 20 mm in diameter and is located 5 mm under the surface. Thermogram on the Fig.6 shows the temperature field of the object's surface after IR heating pulse was applied; in the middle of the thermogram a defect area is clearly identifiable.

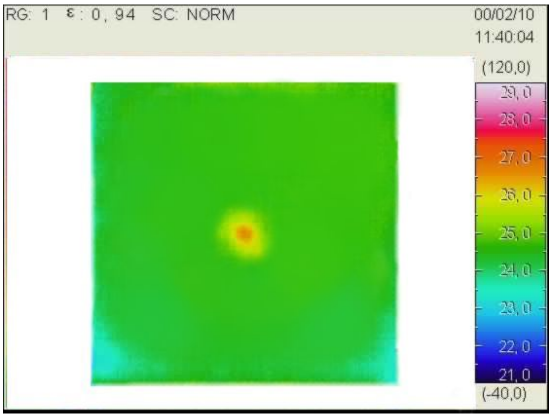


Fig.6 Thermogram of the object under test with subsurface defect in the middle

Temperature difference between defect and sound area is around 2.5 K, as it was predicted by the theory and FEM simulations.

5. Discussion and conclusions

The finite element method was successfully used for the modelling of thermal energy propagation in 3D objects with subsurface defects. Simulation results allow optimization of the active infrared thermographic method, optimal setting of heat pulse parameters and the thermogram observation time. In the active thermography experiments special care should be taken if the emissivity of the surface is non-uniform, or the surface of the object is non-uniformly irradiated by the IR source. This can introduce temperature gradients on the object's surface similar to changes induced by subsurface defects, and therefore these non-uniformities should be eliminated. Results of FEM simulations were experimentally verified in laboratory experiments and the active pulse thermography method was also successfully used at the investigation of frescos and mural paintings (co-operation with restorer J. Dorica).

Acknowledgements

The authors are grateful Scientific Grant Agency of Ministry of Education of the Slovak Republic and Slovak Academy of Sciences VEGA for financial support of the project No. 2/7082/27 "Infrared methods of measurement and non-destructive testing – active infrared thermography, reflectography and thermometry".

References

[1] Maldague X, Marinetti S. Pulse Phase Infrared Thermography. *J. Appl. Phys*, 79(5): 2694-2698, 1996.

[2] Meola C, Carlomagno G. Recent advances in the use of infrared thermography. *Meas. Sci. Technol,* 15: R27–R58, 2004.

[3] Vincent J D. Fundamentals of Infrared Detector Operating and Testing. John Wiley & Sons, New York, 1990.

[4] Comsol Mutiphysics: http://www.comsol.com.

Calorimetric Measurements of AC Losses in BSCCO and YBCO Tapes

P. Mozola, M. Polak

Institute of Electrical Engineering, Slovak Academy of Sciences, Bratislava, Slovakia
Email: pavol.mozola@savba.sk

Abstract. *Using of calorimetric method we measured AC losses of commercially available BSCCO and YBCO tapes exposed to external magnetic field with amplitudes up to ~ 100 mT. The evaporation rate of the liquid nitrogen, caused by AC losses, was measured by electronic flow-meter. We studied the losses in single tapes as well as in tapes arranged in stacks of several tapes. The losses measured with single tapes are compared with the data calculated using the theoretical model of Brandt. The calibration procedure, measurement accuracy and detection limits of this method will be also discussed.*

Keywords: Calorimetry, superconductor, BSCCO, YBCO, AC losses.

1. Introduction

Measurements of AC losses of high temperature superconductors are of high importance for the design of electro-magnetic devices operating in AC regimes. Electrical methods of AC loss measurements are often used [1], however, the results are affected by various errors and the experimental set-up is quite expensive. Moreover, the measurements are difficult in the case of an non-regular shape of the sample.

Calorimetric measurements are less sensitive and less precise, however, the results are quite reliable. Schmidt [2] presented a set-up for calorimetric ac loss measurements at 4.2 K. A slight temperature increase of the sample in an ac-field, thermally insulated from the liquid helium bath, is used as a measure of losses. Okamoto et al. [3] developed a measurement set-up based the measurement of the nitrogen boil-off method with the sensitivity limit of ~ 0.5 W, using a wet gas counter. The sensitivity of their set-up is determined by the background flow rate, which is about 100 mW on a 5 hour average.

2. Subject and Methods

For measurements of AC losses in superconducting tapes and pancake coils we developed a simple measurements set-up using a sensitive electronic flow-meter to measure the evaporation rate of LN_2. The measured sample is placed in a calorimeter box and the evaporation rate of the cooling medium (nitrogen) is proportional to the sample losses. The evaporation rate is measured by an electronic flow-meter Setaram [4]. The calibration of the system using a resistor is quite easy and reliable.

AC magnetic field is generated by a copper magnet. To illustrate the properties of the set-up we measured AC losses of BSCCO tapes (American superconductors) in AC magnetic field with frequency 256.4 Hz at 77 K.

The sketch of the calorimeter is shown in Fig. 1. It is a cylindrical vessel made of a glass-fibre epoxy material. The evaporated nitrogen flows through a glass - fibre epoxy tube to the flow-meter. The measured samples are fixed to the bottom plate of the calorimeter using polystyrene spacers, as shown in Fig.1.

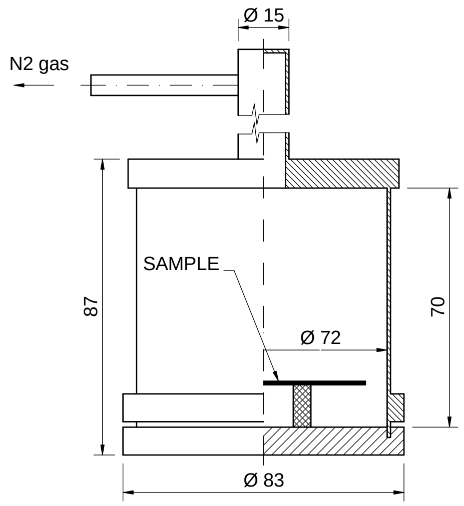


Fig. 1. The sketch of the calorimeter.

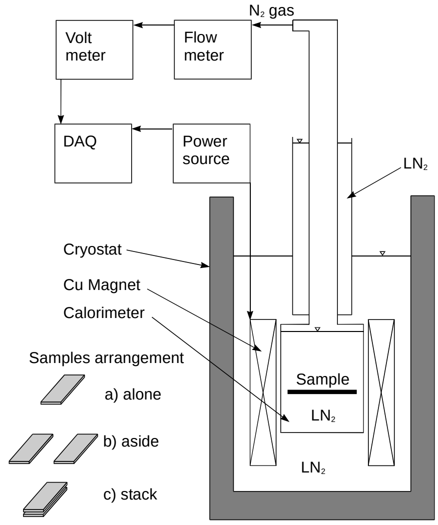


Fig. 2. Schematic view of experimental set-up and samples arrangement in the calorimeter.

The measurement set-up is shown in Fig.2. The calorimeter is placed in the working space of the copper coil with the working diameter of 85 mm.

The coil is supplied by an AC power supply (FUG) controlled by a signal generator. Due to large inductive voltages across the coil (coil inductance is 22.15 mH, its reactance at 256.4 Hz is 1610 Ohm) the coil must be supplied via a battery of condensers. If $\omega L = 1/\omega C$, the total voltage across LC circuit is small, as the voltages across L and C are in antiphase.

Calibration

To calibrate the calorimeter we used a resistor R = 100 Ohm inserted in the liquid nitrogen. A current I flowing through the resistor generates heat $W = RI^2$, which evaporates the liquid nitrogen. The output voltage of the flowmeter, V_f, is proportional to the evaporation rate of LN_2. In Fig.3 we show the dependence of V_f on the power $W=RI^2$ dissipated in the liquid nitrogen.

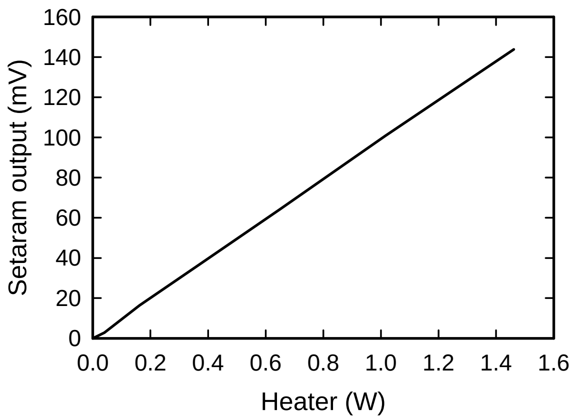


Fig. 3. The calibration curve of the calorimeter. The calibration constant is K_c= 98.84 mV/W.

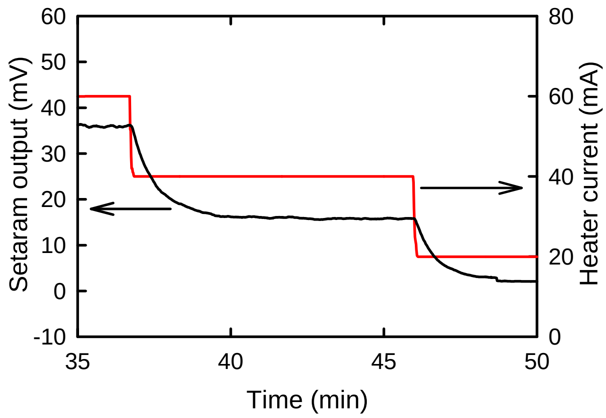


Fig. 4. The dependence of V_f on time recorded during the change of the heating current I_h from 60 mA to 40 mA and from 40 mA to 20 mA.

To measure the saturated evaporation during the calibration, we have to wait of about 5 minutes, as it follows from the measured dependence of V_f on time (see Fig.4).

The stability of the backround nitrogen flow measured during 1 hour is illustrated in Fig.5. The signal change was less than ~ 0.2 mV, which corresponds to losses of about 2 mW. We believe that measurement set up is able to detect losses about 10 to 20 mW approximately.

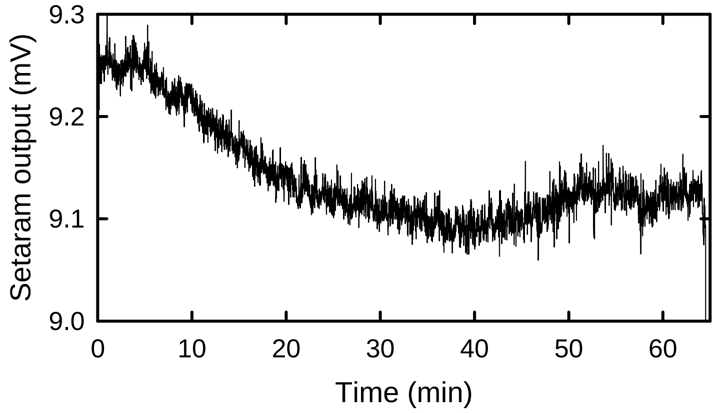


Fig. 5. The measured time variation of the voltage V_f.

3. Results and discussion

In Fig. 6 we show the losses measured for 1 and 2 pieces of 50 mm long BSCCO tapes vs. the amplitude of the external magnetic field with frequency of 256.4 Hz at 77 K. The filamentary tape with the cross-section of 0.25 mm x 4 mm was manufactured by American Superconductors, its critical current at 77 K is I_c ~ 110 A in zero field. The mean value of I_c in magnetic field interval from 0 to 30 mT is ~ 70 A. Due to silver matrix the filamentary tape behaves like a monocore tape. The samples were placed parallel in parallel arrangement at the distance of about 15 mm, so that the mutual influence of the tapes is negligible. For comparison, we show also the results for copper clad YBCO tape 4 mm wide with the total thickness of 0.1 mm (manufacturer Super Power).

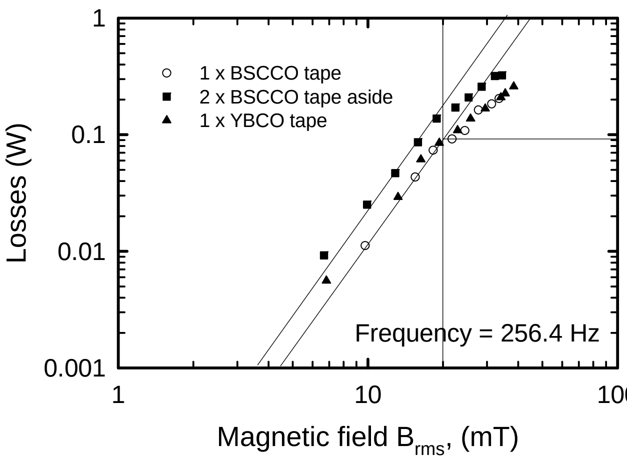


Fig. 6. The AC losses of one and two pieces of BSCCO tape 50 mm long. Also losses of 4 mm wide YBCO tape are shown for comparison.

At low values of the external magnetic field B the losses are proportional to ~ B^3 as predicted by the theory of Carr [5]. The measured AC losses are quite well proportional to the number of tapes supposing they are in the sufficient distance to avoid the mutual influence of the tapes.

We calculated hysteresis losses, P_h, using Carr formula $P_h=2\cdot f\cdot a\cdot I_c\cdot B_m$ [W/m], where f – frequency, a – half tape width, I_c – mean value of critical current of the tape in the measured field interval from 0 to B_m, B_m – magnetic field amplitude. For parameters f = 256.4 Hz, a = 2 mm, I_c = 70 A, B_m = 28.3 mT (B_{rms} = 20 mT), losses are 2.1 W/m, we measured losses 0.09 W for 50 mm long tape, which corresponds to 1.8 W/m. This difference is reasonable.

4. Conclusions

In the calorimeter described above we were able to detect losses above 20 mW approximately. The calibration constant of the calorimeter is ~ 100 mV/W and the stability of the signal for zero losses is of about +/- 1 mV, which corresponds to the loss of ~ 10 mW. Further improvement will be focused on the reduction of the zero loss signal and its stability.

Acknowledgements

We acknowledge the support of, European Office of Aerospace Research and Development (EOARD), Award No. FA8655-07-1-3005.

References

[1] Polak M, Kvitkovič J, Mozola P, Ušák, E, Barnes P.N, and Levin G.A. Frequency dependence of hysteresis loss in YBCO tapes. Supercond. Sci. Technol. 20: S293-S298, (2007).

[2] Schmidt C. Calorimetric ac-loss measurement of high T_c-tapes at 77 K, a new measuring technique. Cryogenics: 40, 137-143, (2000).

[3] Okamoto H, Hayashi H, Sumiyoshi F, Iwakuma M, Izumi T, Yamada Y, Shiohara Y. Nitrogen boil-off method of measuring AC losses in YBCO coils. Physica C 463-465: 795-797, (2007).

[4] Flowmeter Setaram, http://www.setaram.com/

[5] W J Carr Jr. Loss in a striated coated conductor. Supercond. Sci. Technol. 20: 168-175, (2007)

Continuous vs. Discrete Height Scan Method in Normalized Site Attenuation Measurement for EMC Test Site Evaluation

[1]M. Kollár, [2]K. Kováč

[1]Austrian Research Centers GmbH – ARC, RF-Engineering,
A-2444 Seibersdorf, Austria,
[2]Department of Measurement, Faculty of Electrical Engineering and Information Technology, Ilkovičova 3, 81219 Bratislava, Slovak Republic
Email matej.kollar@arcs.ac.at

Abstract. *Normalized Site Attenuation is a standardized procedure for anechoic chamber validation in a frequency range from 30MHz to 1GHz. As an antenna positioning is a part of the measurement procedure, discrete or continuous scanning method can be used. The continuous height scan method has a significant speed advantage over the discrete approach, but the measurement accuracy comes to a question if antenna movement speed is not set properly. This paper deals with the theoretical analysis of the measurement error caused by antenna movement speed as well as the real measurement comparison of both methods. The final conclusions can be applied on various radiofrequency measurements.*

Keywords: RF measurement, site validation, frequency & height scan method

1. Introduction

The radiofrequency department of the Austrian Research Centres GmbH – ARC [1] deals with various site validation and calibration measurement tasks. As accredited test laboratory it performs on demand measurements all around the world. One of such measurements is the Normalized Site Attenuation (NSA), which is used for semi anechoic chambers validations. This type of measurement is designed for frequency range from 30MHz to 1GHz and is further defined by various standards [2]. The goal is to find out how much reflections occur from obstacles existing in measuring space and imperfections of the tested chamber [3]. In the ideal case if the signal is transmitted by transmit antenna the receive antenna should receive only interfered combination of the direct wave and wave reflected from the ground plane. To come near to the ideal behaviour, walls of the chamber are covered with absorber material (usually ferrite or graphite pyramids) to eliminated reflections as much as possible. Of course this intention is never fully achieved, and thus some reflections always occur. The purpose of the validation procedure is to find out if the chamber meets the standard specification of 4dB deviation from the theory.

2. Subject and Methods

The measurement procedure itself involves frequency scans at various receiving antenna heights from 1 to 4 meters to measure the maximum intensity of the wave interference (see Fig. 1). For antenna positioning a specialized mast device is used. In principle there are two ways how the receive antenna movement can be realized, using the discrete or the continuous method.

The discrete height scan is a classical method. It works the way the receiving antenna is positioned in specific height, then a frequency sweep is performed and afterwards the antenna is moved to the next height determined by a height step. This procedure is repeated until 4 meters height is reached. Then for each frequency the maximum received signal (= minimum attenuation) is calculated among all sweeps.

The continuous height scan performs frequency sweep at the bottom 1m height then the mast starts its movement until reaches the top 4m height. During the mast movement frequency sweeps are continuously executed and when mast stops one more sweep is done at the top height. Finally the results are computed in the same way as mentioned above.

As complete chamber validation process requires performing described procedure for up to 10 positions and 2 polarisations of the transmitting antenna, the time duration of the measurement is critical. Comparing both described height scan options, the continuous method is approximately 3 times faster than its discrete alternative. It is mainly because there is no needed to wait until the scan is finished and no need to start/stop mast movement many times. When a real NSA site validation is performed using the continuous height scan several days of work can be saved in this manner. The only think one should have in mind is that the sweep speed must be fast enough in relation to the mast movement to ensure finding the maximum signal. Accordingly the continuous height scan procedure accuracy is examined and optimal antenna movement speed is determined in the following.

A reason why the size of the height step has an impact to the NSA measurement results has its origin in the physical phenomena of the interference of direct and reflected signals. Due to this the signal attenuation depends on constructive or destructive interference occurrence as the receive antenna position (height) changes. This way attenuation minimums and maximums alternate over the height range. Additionally the higher the signal frequency is the more such alternations occur as shown in Fig. 1. The height step size is more critical for higher frequencies than for lower ones, because if the attenuation extremes alter more frequently it is probable to miss it if the height step is too big.

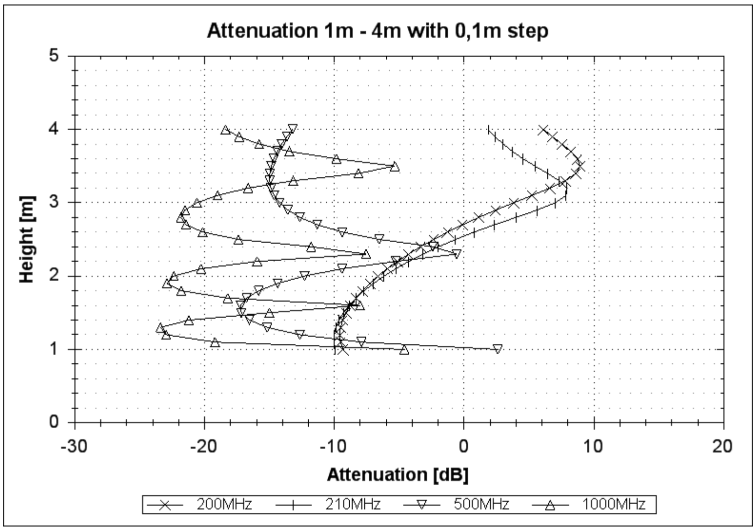


Fig. 1. Attenuation value for various signal frequencies when receive antenna changes its height from 1 to 4 meters with 10cm height step.

In the continuous height scan simulation the duration of a sweep is assumed to be equal to a certain mast movement. This means the movement speed can be defined as distance per sweep. We get 30 frequency sweeps if mast moves from 1m to 4m with a distance per sweep of 0.1m. In addition sweeps at top and bottom height positions with antenna not moving are taken.

The reference simulation was in fact the discrete height scan with 0.01m step, what resulted in 301 sweeps. In Fig. 2 results are displayed for both reference and continuous height scan using various height steps. Obviously by increasing height step size the error grows.

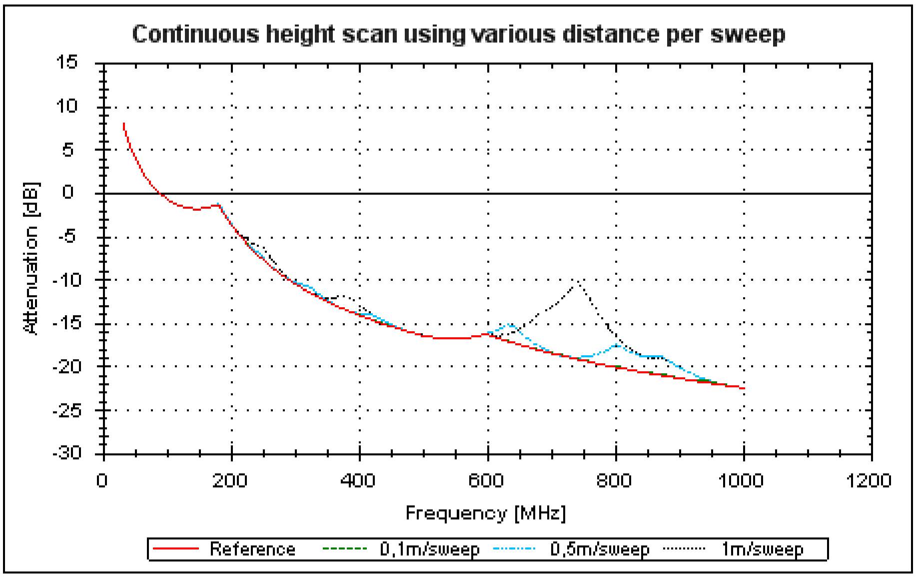


Fig. 2. Results of the continuous height scan using various mast speed (antenna setup: vertical polarization, 1m transmit antenna height, 3m transmit/receive antenna distance).

By computing maximal error from reference for various antenna positions (as defined in [2]) and movement speeds, it can be shown (see Fig. 3) that with distance per sweep of 0.2m the error keeps below 1.4dB. Using distance per sweep of 0.1m the error resides below 0.4dB.

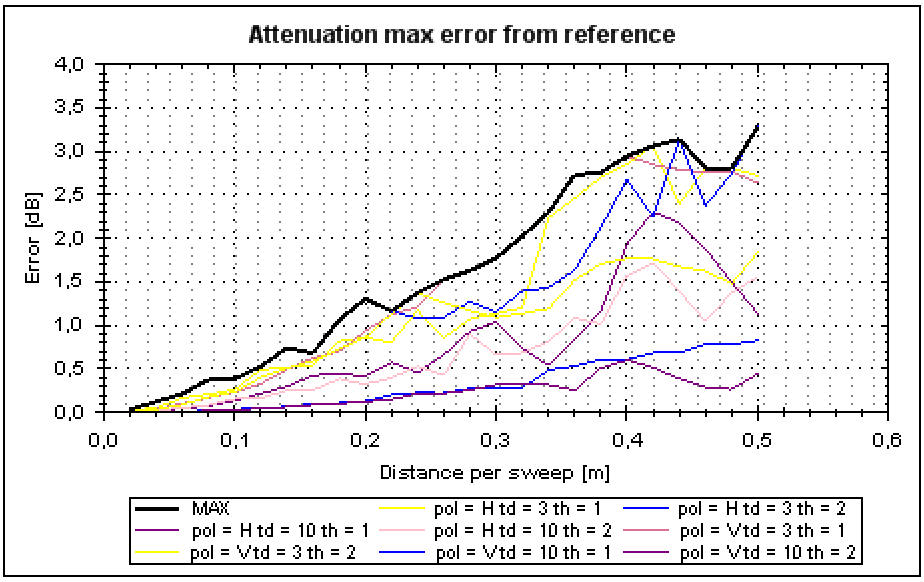


Fig. 3. Maximum deviation of the continuous height scans from the reference for various antenna positions using various movement speeds in the frequency range 30MHz - 1GHz. In the legend pol = polarization (horizontal/vertical), td = transmit antenna distance [m], th = transmit antenna height [m]

3. Results

The signal attenuation depends on the frequency range in which the measurement is performed. As seen in Fig. 2 in a range of 30MHz to 200MHz the receive antenna movement speed has negligible influence on the measurement results. On the other hand for frequencies above 200MHz the influence is much more significant. Performed simulations implied that continuous height scan method is close to the discrete method precision if appropriate mast speed is used. To support our theoretical findings we present real on site measurements using both discrete and continuous height scan methods (see Fig. 4).

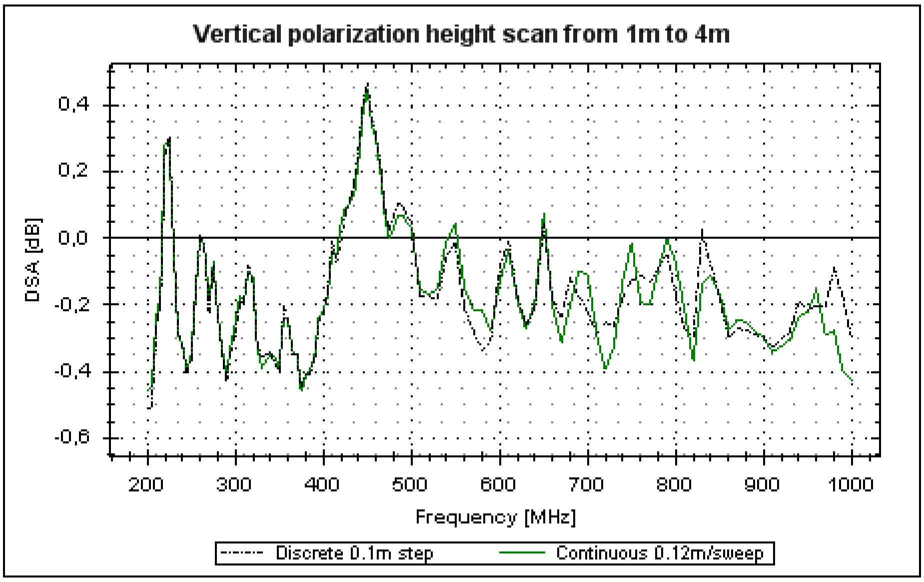


Fig. 4. Continuous vs. discrete height scan in the frequency range 200MHz - 1GHz.

4. Discussion

Performed simulations as well as real measurements have proven that the discrete height scan method can be substituted by significantly faster continuous height scan method if appropriate mast speed is used. The determined error/speed functionality is important for the measurement technicians to identify NSA measurement results validity. The continuous height scan method was implemented in automated measurement software Calstan [3] and is used in various measurement procedures.

References

[1] Austrian Research Centers GmbH – ARC, RF-Engineering, A-2444 Seibersdorf Austria, “www.seibersdorf-rf.com”

[2] CISPR 16-1-4 Ed.2, IEC, www.iec.ch ; ANSI C63.4, IEEE, www.ieee.org

[3] Bittera M., Hallon J., Smieško V. Measurement and Simulation of Field Homogenity Inside Semi-Anechoic Chamber.In: Measurement Science Review. Vol. 3, sect. 3 (2003). - pp. 143-146

[4] Kollar M. Calstan – A software for Automation of Radio Frequency Measurements and Calibrations. Measurement 2007, Proceedings of the 6th International Conference, Smolenice, Slovakia

Roundness: Determining the Reference Circle for MCCI and MICI System

M. Petrík, J. Kováč, P. Kaťuch, L. Bednarčíková, R. Hudák, J. Živčák,

Technical university of Kosice, Department of Biomedical Engineering, Automation and Measurement, Letna 9, 040 01 Kosice, Slovakia, Email: martin.petrik@tuke.sk

Abstract. *Roundness measurements are used to evaluate and control the quality of cylindrical objects. There are several methods how to find the reference circle. The roundness error is then defined as the maximum peak-to-valley distance from the reference circle. In this article we present methods with their basic principles. The aim is focused on finding the minimum circumscribed circle and maximal inscribed circle.*

Keywords: circumscribed, inscribed, circle, Voronoi, roundness.

1. Introduction

Roundness measurements are used to evaluate and control the quality of cylindrical objects. Methods used for determining the roundness are the least-squares circle (LSC), maximum inscribed circle (MIC), minimal circumscribed circle (MCC) and minimum zone circle (MZC) method. These methods are defined as follows [2][4].

- Least square circle (LSC): It is a circle which separates the roundness profile of an object by separating the sum of total areas of the inside and outside it in equal amounts. The roundness error then can be estimated as the difference between the maximum and minimum distance from this reference circle

- Minimum circumcised circle (MCC): It is defined as the smallest circle which encloses whole of the roundness profile. Here the error is the largest deviation from this circle

- Maximum inscribed circle (MIC): It is defined as the largest circle that can be inscribed inside the roundness profile. The roundness error here again is the maximum deviation of the profile from this inscribed circle.

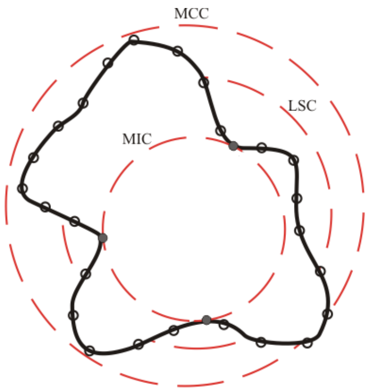


Fig. 1. Overview of roundness methods

- Minimum Zone circle (MZC): Here two circles are used as reference for measuring the roundness error. One circle is drawn outside the roundness profile just as to enclose the whole of it and the other circle is drawn inside the roundness profile so that it just inscribes the profile. The roundness error here is the difference between the radius of the two circles.

2. Subject and Methods

For evaluating the roundness error from the actual measurement, a reference circle must be established from the measurement data to minimize the maximum deviation between the reference circle and the actual one. The roundness error is then defined as the maximum peak-to-valley distance from the reference circle.

In this chapter we describe the basic principle of finding the reference circle for deviation calculation.

Taylor's maximum material principle

This principle is invented by Wiliam Taylor in 1905 as GO gage. This principle is still used to inspect the maximal material condition (MMC) and least material condition (LMC). Principle states: "The maximum material limits of as many related elements or dimensions as practicable should be incorporated in the go gage, whereas the minimum material limit of each related element or dimension may be gaged only by individual minimum material limit gages or gaging methods." On cylindrical parts, size of gage limits roundness. This is accomplished through checking procedures of two opposing points [2].

Voronoi – Voronoi diagram

Voronoi diagrams are fundamental data structures that have been extensively studied in Computational Geometry. A Voronoi diagram can be defined as the minimization diagram of a finite set of continuous functions. Usually, each of those functions is interpreted as the distance function to an object. The associated Voronoi diagram subdivides the embedding space into regions, each region consisting of the points that are closer to a given object than to the others. We may define many variants of Voronoi diagrams depending on the class of objects, the distance functions and the embedding space. Voronoi diagrams are named after Russian mathematician Georgy Fedoseevich Voronoi who defined and studied the general n-dimensional case in 1908. The application of the diagram can be used in many fields, for example accuracy of parameters, biology, chemistry, network systems, modeling, planning [3].

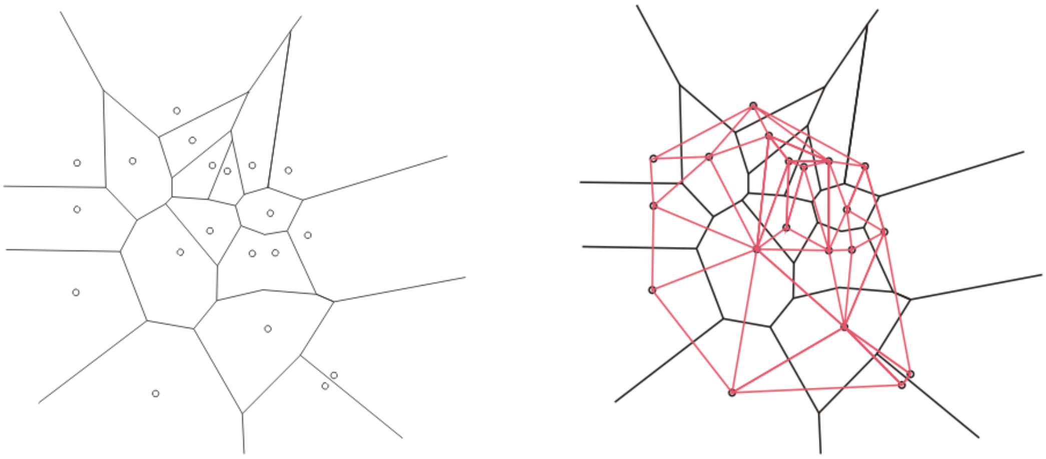


Fig. 2. Voronoi diagram (left), relationship between Voronoi and the Delaunay triangulation (right).

The Voronoi diagram we can use for determining the minimal circumscribed circle or maximal inscribed circle. The circumscribed circle we can construct from the boundary points of diagram, which lies farthest from center of diagram. The inscribed circle has the center on Voronoi vertex or on Voronoi edge, which intersect convex hull. The radius is a distance of circle center and convex vertex. For a calculating the convex and Voronoi diagram we use the Delaunay triangulation. Delaunay triangulations maximize the minimum angle of all the angles of the triangles in the triangulation. The Delaunay condition states that a triangle net is a Delaunay triangulation if all the circumcircles of all the triangles in the net are empty. The triangulation was invented by Boris Delaunay in 1934. The relationship between Voronoi and the Delaunay is on the fig 2. The convex vertex represents the measured value.

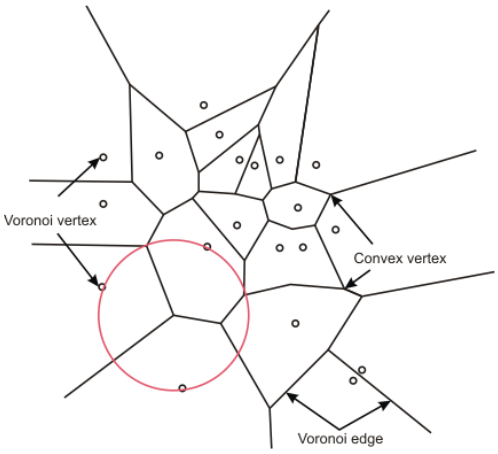


Fig. 3. Determining the inscribed (left) and circumscribed (right) circle using Voronoi.

Permutation method

Simplest way to find the MIC or MCC is to make the permutation from measured points. The each three point set are calculated. The condition for these points is that the centre of circumscribed triangle by circle must be within these tree points and all point must lie in (MIC) or out (MCC) from the circle (Fig.4). The second pass calculates with the pairs of points, where the circle is in the middle and the radius is the half distance of these points. This is not the best for many points and not fastest method to find the circles.

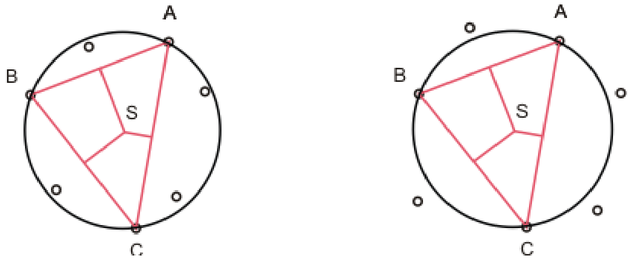


Fig. 4. Demonstration of condition for circumscribing(left) and inscribing(right) circle

The diameter and the centre position of this circumscribed circle around the triangle are following [1][5]:

$$diameter = \frac{|AB|.|BC|.|CA|}{2.S_{abc}} \tag{1}$$

where

A, B, C vertex of triangle

S_{abc} area of triangle

$$Sx = \frac{1}{2\begin{vmatrix} B_x - A_x & B_y - A_y \\ C_x - A_x & C_y - A_y \end{vmatrix}} \begin{vmatrix} |A|^2 & A_y & 1 \\ |B|^2 & B_y & 1 \\ |C|^2 & C_y & 1 \end{vmatrix} \qquad Sy = \frac{1}{2\begin{vmatrix} B_x - A_x & B_y - A_y \\ C_x - A_x & C_y - A_y \end{vmatrix}} \begin{vmatrix} A_x & |A|^2 & 1 \\ B_x & |B|^2 & 1 \\ C_x & |C|^2 & 1 \end{vmatrix} \tag{2}$$

where

S_x, S_y centre of circle

Linear time method

In computational complexity theory, an algorithm is said to take linear time, or O(n) time. We can state that, there exist O(n3) such circles, and each takes O(n) time to check, for a total running time of O(n4). Elzinga and Hearn gave an O(n2) algorithm in 1972, and the first O(n logn) algorithms were discovered by Shamos and Hoey (1975), Preparata (1977), and Shamos (1978). Nimrod Megiddo showed in 1983 that his ingenious prune-and-search techniques for linear programming that adapted to find the minimal enclosed circle. This method reduced the number of circles (points) from calculation in every step by the special formula. The latter algorithm finds the median (and in general, the k-th smallest element) [6][7].

3. Conclusions

Each of these methods described above has some advantages and disadvantages. The best choice depends on quantity of points to be checked for reference circle. The question is “Which method is more precise, less time consuming, and reliable?”

Acknowledgements

This research has been supported by the research project 1/0829/08 VEGA - Correlation of Input Parameters Changes and Thermogram Results in Infrared Thermographic Diagnostic.

References

[1] Godfrey C, Onwubolu, B,. V. Babu: New Optimization Techniques in Engineering, Springer, 2004, 615-620

[2] Meadows J. D.: Geometric Dimensioning and Tolerancing: Applications and Techniques for Use in Design, Manufacturing, and Inspection, CRC Press, 1995, 46-50

[3] O'Rourke J.: Computational geometry in C, Cambridge University Press, 1998

[4] http://en.wikipedia.org/wiki/Roundness, 27.2.2009

[5] http://en.wikipedia.org/wiki/Circumcircle, 27.2.2009

[6] http://www.cs.mcgill.ca/~cs507/projects/1998/jacob/solutions.html, 27.2.2009

[7] http://www.personal.kent.edu/~rmuhamma/Compgeometry/MyCG/CG-Applets/Center/centercli.htm, 27.2.2009

Measurement of Electromagnetic Shielding Efficiency of Composite Materials

J. Dřínovský

Department of Radioelectronics,
Faculty of Electrical Engineering and Communication, Brno University of Technology
Purkyňova 118, 612 00 Brno, Czech Republic
Email: drino@feec.vutbr.cz

Abstract. *This paper deals with the theoretical and practical aspects of the shielding efficiency measurements of construction materials. In this contribution is described the alternative test method of these measurements by using the measurement circular flange. There are also discussed the measured results and parameters of coaxial test flange. The measurement circular flange is described by measured scattering parameters in the frequency range from 9 kHz up to 1 GHz. The accuracy of the used shielding efficiency measurement method was checked by brass calibration ring. The whole measurement of shielding efficiency was controlled by the program from the personal computer. This program was created in the VEE Pro environment produced by © Agilent Technology.*

Keywords: Shielding Efficiency, Composite Materials, Scattering Parameters, Measuring Flange

1. Introduction

The measurement of the shielded, absorbing and EMC chambers or boxes are usually done by the setup which contains the transmitting and receiving antennas, test signal generator and test signal receiver. As the test signal receivers are usually used the EMC receivers or spectral analysers. The measurement itself runs by the following way. The receiver with the receiving antenna and also with essential cable is situated inside the chamber or tested box. The transmitter (signal generator) and transmitting antenna are placed at the outer side of the tested object. The location of the antennas is changed around the chamber or box. The worst case, when the shielding efficiency is lowest, is reliably identified by this positioning of antennas [1] and [2].

The problem could appear when it is necessary to measure the shielding effectiveness of the material from which will be the chamber or box constructed. Especially in the development stage is not possible to construct the whole chambers or boxes with the huge sizes for accurate measurements. This approach is very expensive and also time consuming.

The similar problem appears when it is necessary to know the shielding efficiency of the construction materials like bricks, plasterboard, concrete etc. This material could be also called as composite material, especially during its development stage. The construction of the chambers or boxes from these types of materials for the measurement setups mentioned above has the main problems with the door construction. The door of these chambers or boxes has usually the main influence on the whole shielding effectiveness, in the other words the doors always represent the weakest part of these chambers. But the construction of the doors from the concrete is really complicated, in same cases nearly impossible.

2. Subject and Methods

The alternative test method for the testing of shielding efficiency of shielding materials is discussed in [3]. The presented coaxial test apparatus is suitable for thin materials like plastic or metallic board, fabric material and so on. But this setup is not suitable for the construction materials (concrete, bricks etc.), because it is very complicated to produced the thin concrete plain with the maximal height around 1 mm. The modified test setup according to Fig. 1 was produced, after analyses of commonly available measurement solutions and setups. The Fig. 1 shows the technical drawing of the measurement coaxial flange. This flange was mainly designed for frequency range from 9 kHz up to 1 GHz. The shape and dimensions of the flange were calculated for the 50 Ω input and output impedances [4].

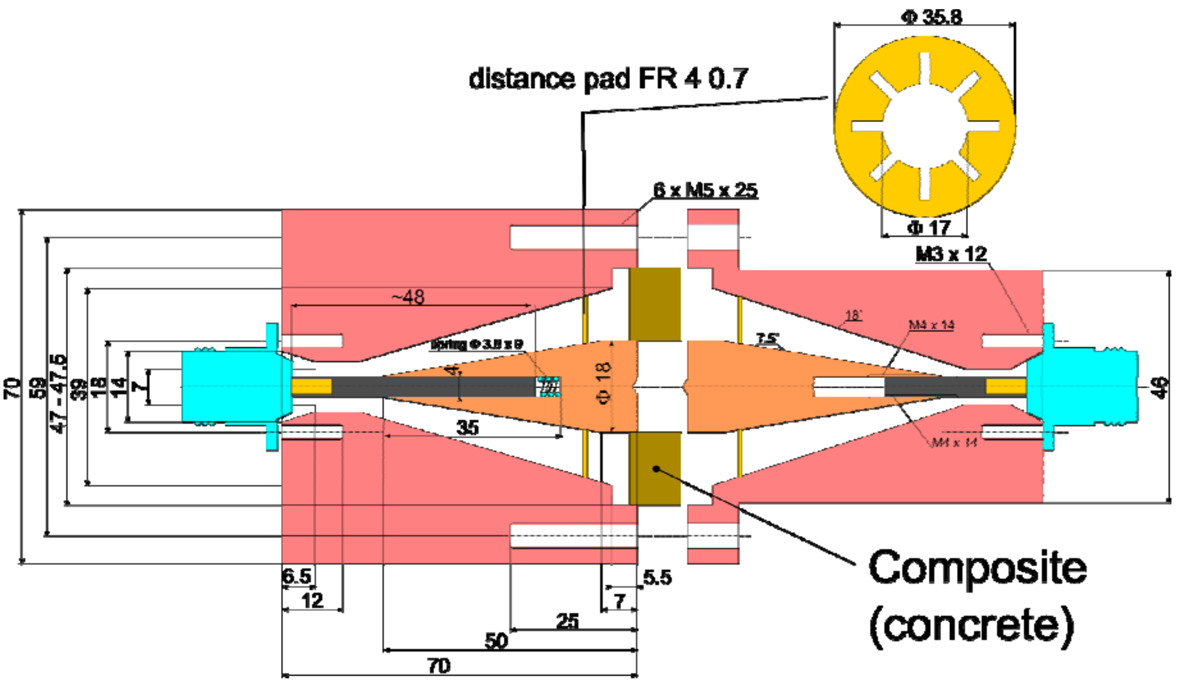


Fig. 1. Basic dimension drawing of the circular flange (dimensions are given in mm).

The design of the flange was done according to the basic mathematical relations [4]

$$Z_{\mathrm{m}} = \frac{60}{\sqrt{\varepsilon_{\mathrm{r}}}} \ln \frac{a_2}{a_1} \tag{1}$$

where

- Z_M is the characteristic impedance of the measurement system (50 Ω);
- ε_r is the relative permittivity (in this case is equal to 1, air);
- a_1, a_2 are the radius of the coaxial line (flange).

The transition from the N-type connector to the opposite end of the flange has the linear shape for both parts central and external. This shape was chosen for the better fabrication. The liner shape could be optimised for the better impedance matching especially at frequencies over 1 GHz. The central flange conductor is fabricated from the brass. The rest of the flange is made from the aluminium alloy. The flange is tightened by the torque wrench after the inserting the test composite by the same torque every time. This setup increases the accuracy of each measurement and also increasing the repeatability during the several measurement.

The measured scattering parameters of the flange itself are given in the Fig. 2. The S_{11} and S_{22} are in the whole range of interest under the -15 dB which refers about the good matching of the both test ports with the measuring system. The insertion losses in both directions (S_{21} and S_{12}) are in the whole measuring frequency range, under 1 dB. This data refers about the accurate design of the whole flange. The flange itself will have the insignificant influence on the total dynamic range of the whole measurement setup. The dynamic range will be mainly affected by the used measuring devices (generator and spectral analyser).

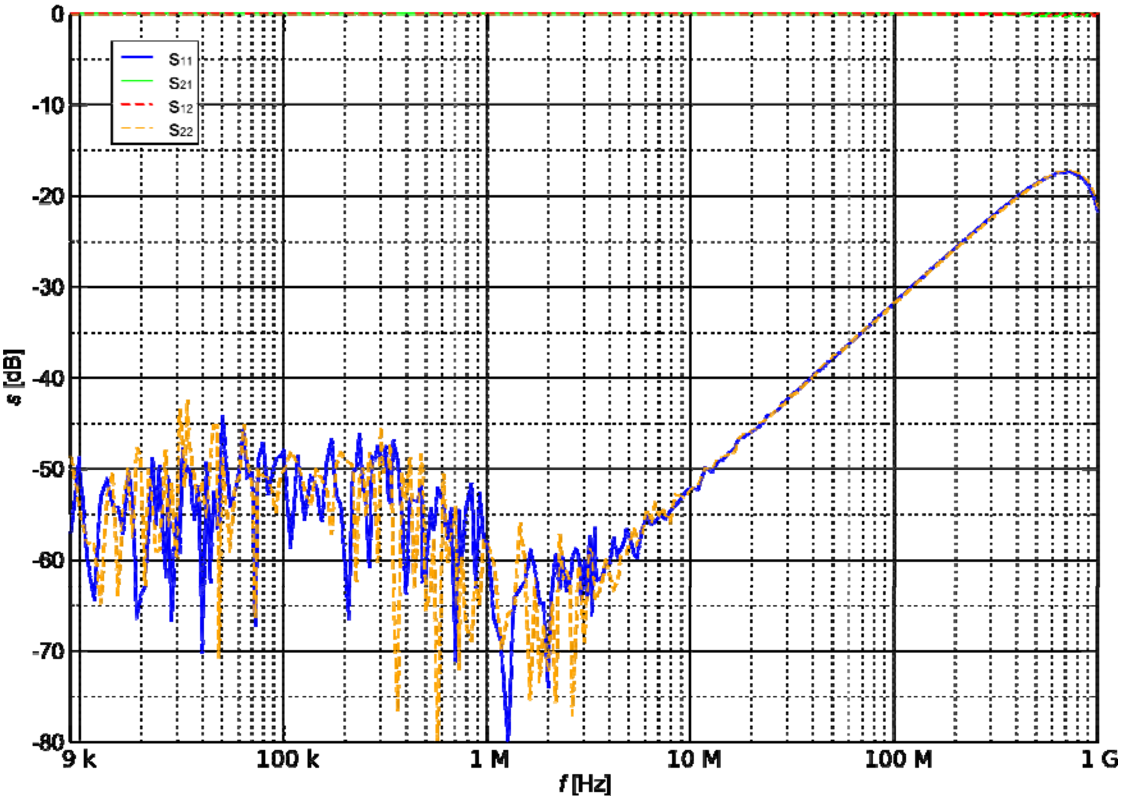


Fig. 2. Measured scattering parameters of the realised coaxial flange.

3. Results and Discusion

The measured scattering parameters refer about the accurate design of the coaxial flange. Several problems could appear during the prefabrication of the test sample concrete ring. This ring has to be produced with the high accuracy of its dimensions. The example of measured shielding efficiency of the composite concrete material is depicted in the Fig. 3. There is also shown the data which was measured with the brass disc. The shielding efficiency of the brass disc is the 115 dB at the kHz range and around 70 dB at the GHz range. The shielding efficiency of the composite concrete material is only several dB in the range from 100 MHz up to 1 GHz. The so low shielding efficiency of the concrete material is mainly caused by the small thickness of the material (only 8 mm).

4. Conclusions

The shielding efficiency of material is composed from several parts. The reflection loss, absorption loss and multiple path reflection losses are the main three parts of whole electromagnetic shielding. For the accurate classification of the shielding efficiency of composite concrete material will be necessary to measure each part of the whole electromagnetic shielding effectiveness. This measurement could be done by the vector network analyser. The dependency of the thickness of the material and shielding efficiency could be determined in the harmony with measured data. The future work will be focused on this problem and also on the composite concrete materials compound. There will be also tested the different thickness of the concrete samples.

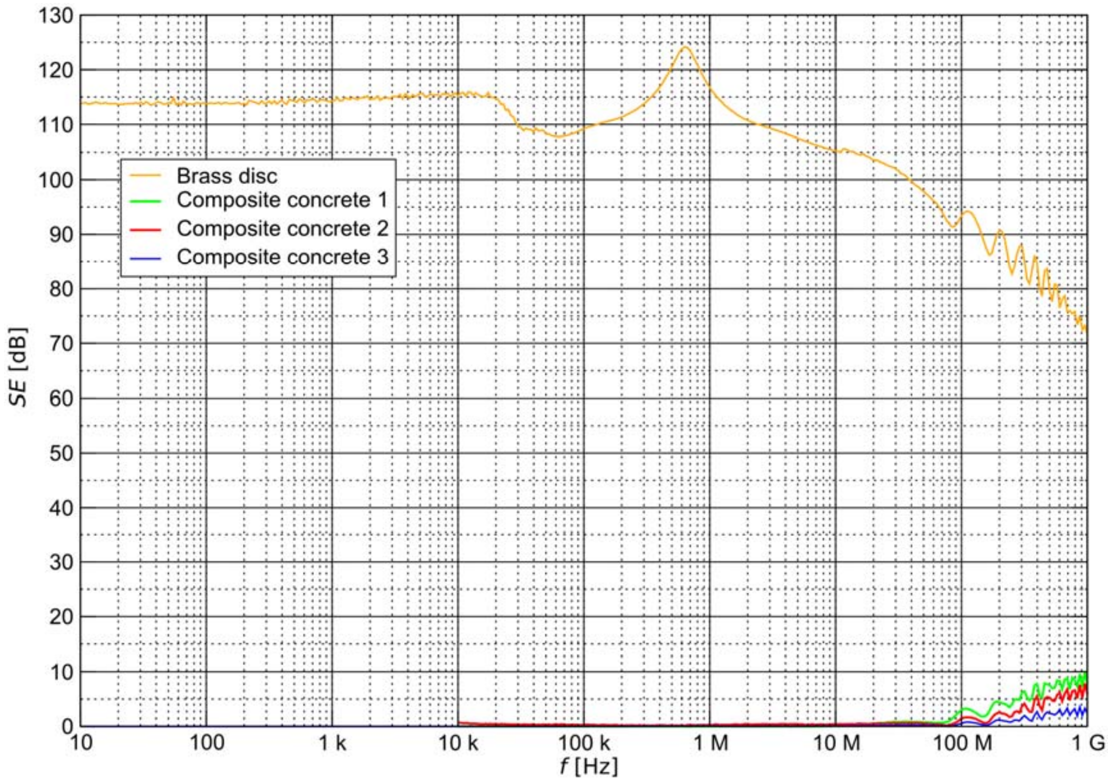


Fig. 3. Shielding efficiency of the brass calibration test disc and the composite concrete test sample.

Acknowledgements

This work has been prepared as the part of the solution of the grant no. 102/07/0688 ”Advanced microwave structures on non-conventional substrates” of the Czech Science Foundation, grant no. 102/09/P215 ”Advanced EMI filters Insertion Loss Performance Analyses in System with Uncertain Impedance Termination” of the Czech Science Foundation and with support of the research plan MSM 0021630513 ”Advanced Electronic Communication Systems and Technologies (ELCOM)” and with the support of the Ministry of Industry and Commerce under the contract FT-TA3/027.

References

[1] IEEE Standard: *Method for Measuring the Effectiveness of Electromagnetic Shielding Enclosures*. EMC Society, New York 2006, p. 39.

[2] Svačina, J., *Electromagnetic Compatibility, Principles and Methods*. (printed in Czech) Brno University of Technology 2001. “Join to European Union”.Volume 2, 156 p. IBSN 80-214-1873-7.

[3] Haga, A., Kamo, Y., Kawamata, K., Minegishi, S., *A New Measurement Method of Shielding Effectiveness*. In International Symposium on Electromagnetic Compatibility 14 – 18 September Roma Italy. 1998. p. 254 – 257.

[4] Černohoský, D., Nováček, Z., Raida, Z., *Electromagnetic waves and lines*. (published in Czech). Brno: Vutium 1999. 136 p. ISBN 80-214-1261-5.

Analysis of Transformer Humidity by Capacitance Measurement

[1]M. Kučera, [1]M. Gutten, [1]S. Kučera, [2]J. Kúdelčík,

[1]Department of Measurement and Application of Electrical Engineering
[2]Department of Physic
Faculty of Electrical Engineering, University of Žilina, Univerzitná 1, 01026 Žilina
Email: kuceram@fel.uniza.sk

Abstract. *To prevent failure states of transformers, we perform different types of measurements. They shall illustrate a momentary state of the measured equipment and if necessary to draw attention in advance to changes of parameters, which have specific relationship to no-failure operation of the equipment. Water presence in oil transformer causes deterioration of its insulation and finally thermal defect of solid insulation.*

Keywords: humidity, transformer, dissipation factor, capacity

1. Introduction

State of new insulation in operation mostly deteriorates due to surface contamination of insulators and insulation, their moistening and ageing. If no measures are taken in time so as to avoid this degradation, the situation usually results in damage of insulator and consequently in stop of electrical device. State of insulation of important electrical devices, such as transmission transformer which bring huge economic cost due to each stop in operation, needs to be checked regularly.

Water presence in oil transformer causes deterioration of its insulation and finally thermal defect of solid insulation. Dielectric warming can be so high that the temperature increase is out of control and transformer becomes dangerous for its surrounding.

2. Theoretical Analysis

The measurements of dissipation factor and capacities of transformer windings are used for additional determination of insulation quality as whole or only of some parts of transformer. The value of dissipation factor indicates presence of polar and ion compounds in oil and it also determinate the aging of oil. The degree of oil humidity can be measured by temperature dependence of $\mathrm{tg}\delta$ [1].

The frequency dependence of capacity is next method for determination the degree of oil humidity (to 10 kHz – Fig. 1). In wet isolation, the absorption current is negligible to leakage current, which is independent on frequency. The stage of insulation can be determined as the ratio of capacities at two different frequencies (for example 50 and 100 Hz).

The next method on the determination of oil humidity was the measurement the value of capacity at various temperatures. The capacity is also the function of the absorption processes, which are characterized by their time constants, and distribution of absorption charges. This method is base on the determination of the ratio $\frac{C_{75} - C_{20}}{C_{20}}$, where C_{20} and C_{75} (or C_{80}) are capacities at 20 and 75 °C (or 80 °C). For some disadvantages of experiment this method was substitute by previous method base on the determination of the ratio of capacities at two different frequencies.

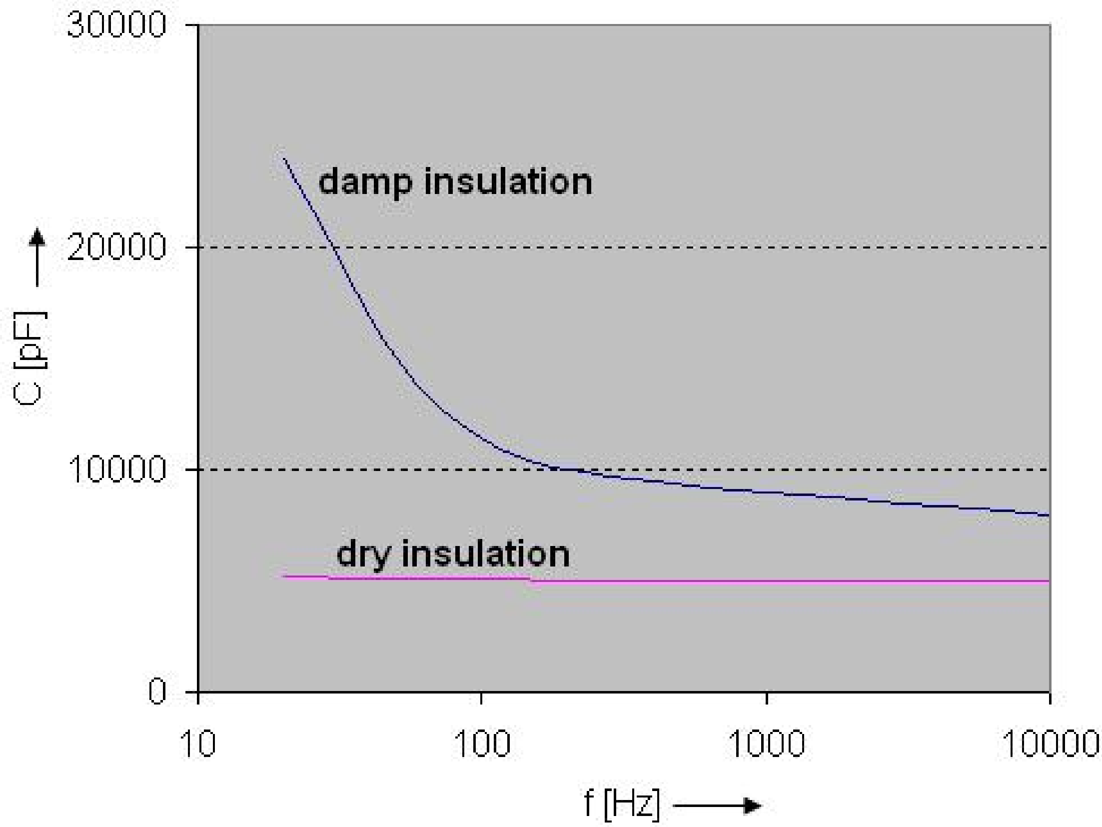


Fig. 1. Frequency dependences of capacity at dry and damp insulation

3. Description of experimental measurement

Experimental measurement on test transformer 60 VA, 220/52 V immersed in tank with transformer oil ITO 100 is provided as an example of safety and reliability inspection of transformer based on insulation humidity.

State of insulation was measured step by step - with no water content in oil and after water was added in amount of 0,05%; 0,15% and 0,25% of transformer tank volume filled with oil. The aim of experiment was to verify relation between increase in oil humidity and capacity-frequency characteristics (see Fig. 3).

In the second step, we measured frequency dependence of capacities at different temperatures in oil: 25, 35, 45, 55, 65 and 75 °C. The experimental measured values showed the connection between the increase of temperature and the change of frequency development of capacity (see measured values Fig. 4).

We were measured capacities as function of oil temperature, applied frequency (to 10 kHz) and oil humidity (water content in oil). The results were compared with the value of insulation resistance. The measurements were automatic using RLC meter and computer.

Computer program set the values of voltage and its frequencies for RLC meter and then it read electrical parameters of selected transformer, which were determined by RLC meter. All measured values were plotted using Excel.

3.1 Description of experimental instrument

Automatic RLC meter Fluke was used to measure dependence of capacity on frequency (Fig.2). Measuring principle of RLC meter is based on measuring selected voltage and current values of measured impedance between transformer's Z_X, see Fig. 2.

The component measurement is based on the current and voltage technique [4]. The component voltage and the component current are measured and converted into binary values. Each measurement cycle lasts approximately 0,5 seconds. For AC measurements one cycle consists of seven single measurements, the results of which are stored and arithmetically evaluated. Finally, the result is transferred to the display.

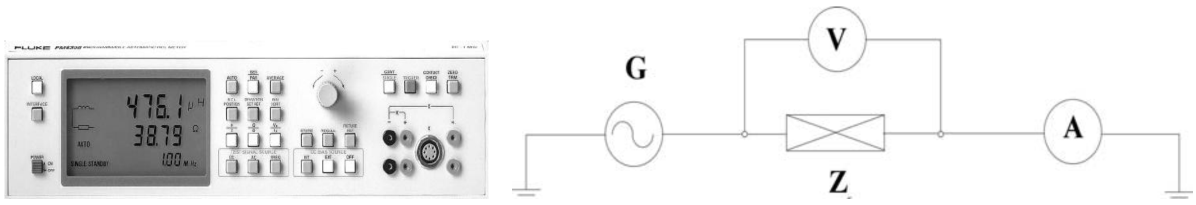


Fig. 2. RLC meter and setup of experiment

The microprocessor uses the measured values to calculate the equivalent series resistance R_s, the equivalent series reactance X_s, and the quality factor $Q = X_s/R_s$ of the component. In AUTO mode, the microprocessor determines the dominant and secondary parameter, calculates its value; and displays it together with the equivalent circuit symbol.

3.2 Results of measurements

Based on measurements, we proved correctness, reliability and high sensitivity of the method for determining humidity in transformer. After water was added in amount of 0,05% of transformer tank volume filled with oil, we already experienced significant change in curve upwards (increase in humidity – Fig. 3).

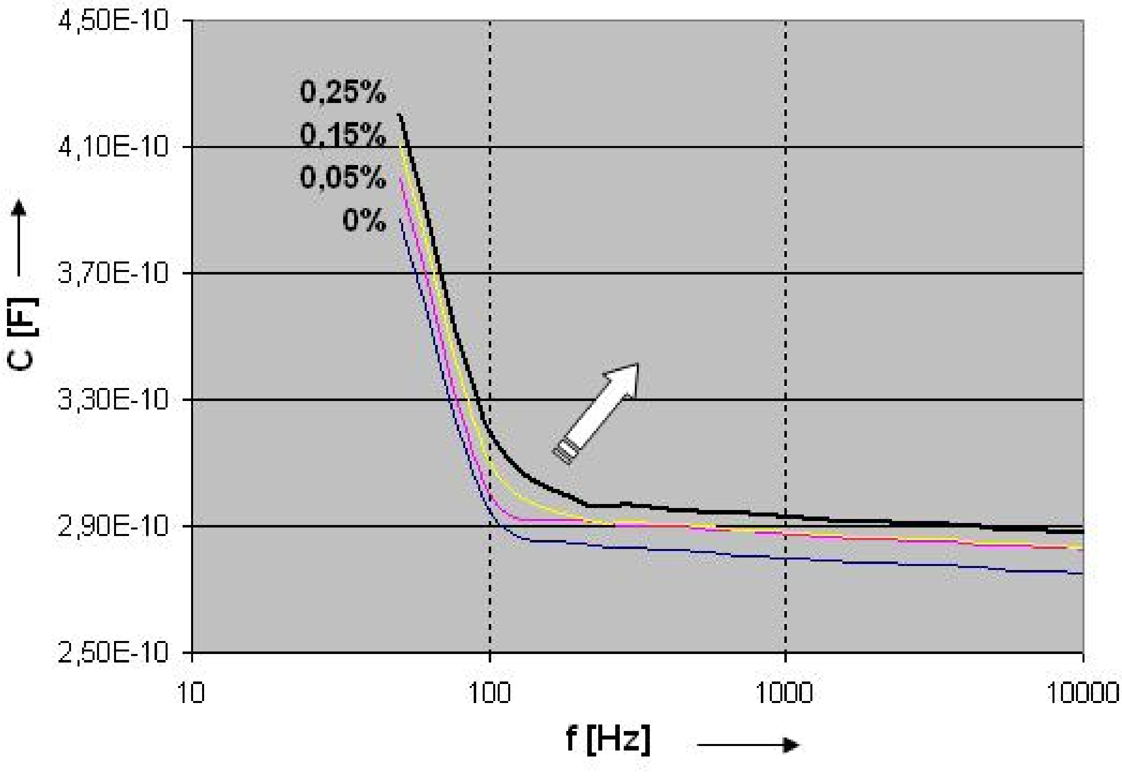


Fig. 3. Frequency developments of capacity as function of water contents in oil: 0; 0,05; 0,15 and 0,25 % (curve increase with water content); ratio $C_{50\,Hz}/C_{100\,Hz} > 1{,}2$ - damp insulation for all the examples

In the second step of measurements (see Fig. 4) it is different frequency dependence of capacities at temperatures in oil: 25, 35, 45, 55, 65 and 75 °C. The experimental measured values we verified the connection between the increase of temperature and the change of frequency development of capacity, i.e. the capacity is also the function of the absorption processes, which are characterized by their time constants, and distribution of absorption charges.

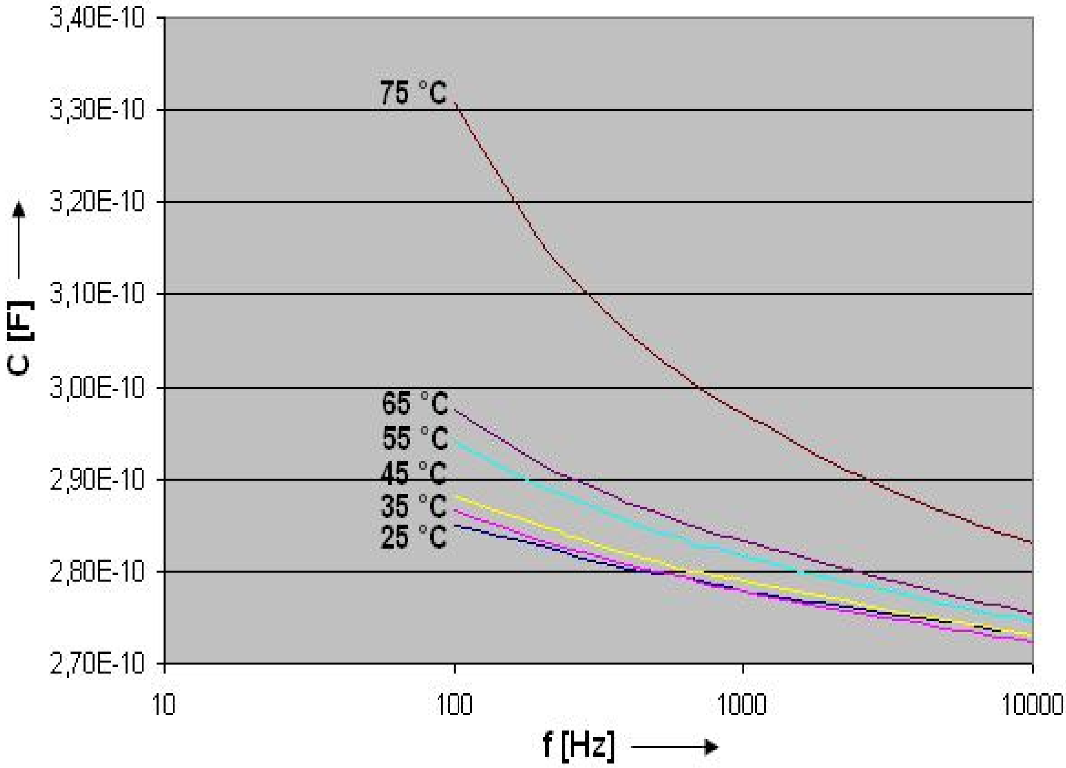


Fig. 4. Frequency developments of capacity as function of oil temperature

4. Conclusions

On the base of measurements, we proved correctness and high sensitivity of both methods for determination of oil moisture in the transformer, i.e. the frequency monitoring of capacity to 10 kHz.

These methods can be utilized to determine insulation state during short term layoff of transformer and thus increase its reliability and safety.

Acknowledgements

This work was supported by the Grant Agency VEGA from the Ministry of Education of Slovak Republic under contract 1/0515/09 and 1/0548/09.

References

[1] Artbauer J., Šedovič J., Adamec V.: Insulants and Insulations, ALFA, Bratislava, 1969
[2] Gutten M.: Diagnostic and Monitoring of Power Transformers, In *5-th International Conference Elektro 2004*, Žilina, 2004, Slovak Republic
[3] Gutten M., Kúdelčík J.: Influence of transformer humidity for its safety and reliability, In *ELDICOM 09*, Žilina, 2009
[4] http://assets.fluke.com/images/products/fpm/bench_instruments/ds_rcl_6304_6.pdf

Slow Rotations in Earthquake Motions Detected in Nuclear Power Plants

[1]Ľ. Ondriš, [2] E.Juhásová, [1]D. Krušinský, [1]V.Rusina, [1]J. Buzási, [1]M. Trutz

[1] Institute of Measurement Science, Slovak Academy of Sciences,
Dúbravská cesta 9, 841 04 Bratislava, Slovakia,
[2] Institute of Construction and Architecture, Slovak Academy of Sciences,
Dúbravská cesta 9, 845 03 Bratislava, Slovakia
E-mail: umerivan@savba.sk

Abstract: *The paper informs about seismic behaviour of the reinforced concrete reactor shaft body of WWER Type Slovak NPPs. The information concerns observed displacement motions due to the earthquakes obtained by instrumentation that serves for the control of reactor vessel position. The nuclear reactor shaft tilt measuring sensitive systems, that utilise the hydro-levelling and pendametric methods, have registered motions excited by different far large earthquakes. The paper presents some of the records illustrating the effects of earthquake Sumatra, 26th Dec., 2004.*

Keywords: Earthquake, Seismic Response, Hydrostatic Sensor, Pendametric Sensor, Nuclear Power Plant (NPP)

1. Introduction

The different monitoring systems are running in every NPP. They support the control and safety of respective NPP during its operation. One from such systems is devoted to monitoring of the concrete reactor shaft tilt as required by technical directives. The tilt vector of reactor main jointing plane is usually measured once each year during reactor shutdown by method, called 'precise levelling'. It involves measuring relative height differences at the points and the immediate calculation of the reactor tilt vector from this data. Besides this classical method , two independent methods were incorporated to the authorised methods. These are hydro-levelling and pendametric methods. These methods are advantageous as they allow practically continuous measurement of tilt vector even during reactor operation. Such systems are nowadays operating in both Slovak NPPs Jaslovské Bohunice and Mochovce [2]. The measurement of the pressure vessel considers the quasi-static process, therefore the used time step of recording was chosen as 5 minutes.

2. Measuring methods and systems used for tilt control

Tilt measuring systems running in the Slovak EBO, EMO power plants use methods of hydro-levelling and pendametry with unified optoelectronic measuring of hydrostatic liquid level and damped pendulum wire position, respectively (Fig.1). In contrast with classical levelling method, which allows measurement of this quantity only during the reactor shut-down state (which is not the operational state), the used measuring principles and applied methods allow a practically continuous measuring of reactor tilt vector. The amplitude of tilt vector in this paper is defined as maximum height difference of two points in the reactor flange plane on the flange diameter basis. The angle of tilt vector is given by angle orientation of line with maximum tilt lying in flange plane to the reactor coordinate system.

Hydro-levelling is a well known levelling method, and in connection with the application of optoelectronics and information technology has some advantages to purely optical methods. A minimum of three hydro-levelling sensors are necessary in order to measure reactor shaft tilt vector. Three hydro-levelling sensors that were developed and made in the Institute of Measurement Science, Slovak Academy of Sciences are used for reactor tilt vector measurement [1, 2]. The optoelectronic sensor in system of connected vessels consists of a flat window cell and an optoelectronic system for hydrostatic liquid level height measurement. All sensors are interconnected with the central industrial computer and power supply. This optoelectronic part is based on a CCD line image sensor. Liquid temperature in sensors is measured by means of thermometers built in the sensors. This temperature is used for the correction of measured data that are influenced by different temperature of sensors. The liquid levels in individual vessels lie in one horizontal plane when certain physical conditions are fulfilled. This plane can then serve as the reference plane for the elevations differences measured between individual points. The range of height difference measurement is ±10 mm, the precision is 0.01 mm.

The pendametric method uses the properties of a damped pendulum. The pendulum hinging point is connected to the measured object. The position of the pendulum wire is measured bi-axially in a reference horizontal plane. The wire's position in the reference plane and its length determine the reactor shaft tilt vector. The optoelectronic method is used for this position measurement similarly as in the hydro-levelling method. Although one biaxial pendametric sensor is sufficient for tilt vector measurement, the measuring system incorporates two pendametric sensors as a purpose-built redundancy. Interconnection of sensors with computer and power supply is common with hydro-levelling sensors one. The measuring range of the pendameter wire position is ±2 mm in both coordinate axes; the resolution is 0.001 mm.

The topology of the three hydro-levelling sensors located on the reactor shaft perimeter allows measurement of its tilt vector in the power plant coordinate system. All sensors are fixed in the reactor shaft body.

Central personal computer allows an user-friendly computation, presentation, and archiving/storage of measured data of the reactor shaft tilt. The measurement regime, including the measuring (sampling) interval, may be programmed according to the user requirements.

3. Records obtained during different seismic events from reactor tilt measurements

In the next text are presented the records obtained from tilt measurement due to Sumatra earthquake. These records show the time dependences of the reactor tilt vector reference amplitudes given in relative units (mm per flange diameter).

The signals recorded by the tilt measuring systems are depicted in Figure 2. The Figure 2 shows two perturbations: the first one started on 23rd December 2004 at 15.40 UTC caused by the earthquake between Australia and Antarctica and the second one started on 26th December 2004 at 01.20 UTC caused by the earthquake off the Indonesian coast with

magnitude 9.3. These records were obtained in NPP EBO-V2. Amplitudes detected with pendametric sensors P1, P2 reached higher values, smaller amplitude values were recorded by hydro-levelling sensors. The reason of smaller amplitudes is due to the delay that was caused by time needed for reaching the equilibrium positions of used liquid.

The time of this earthquake effect corresponds to regular seismological records obtained by national seismic stations network.

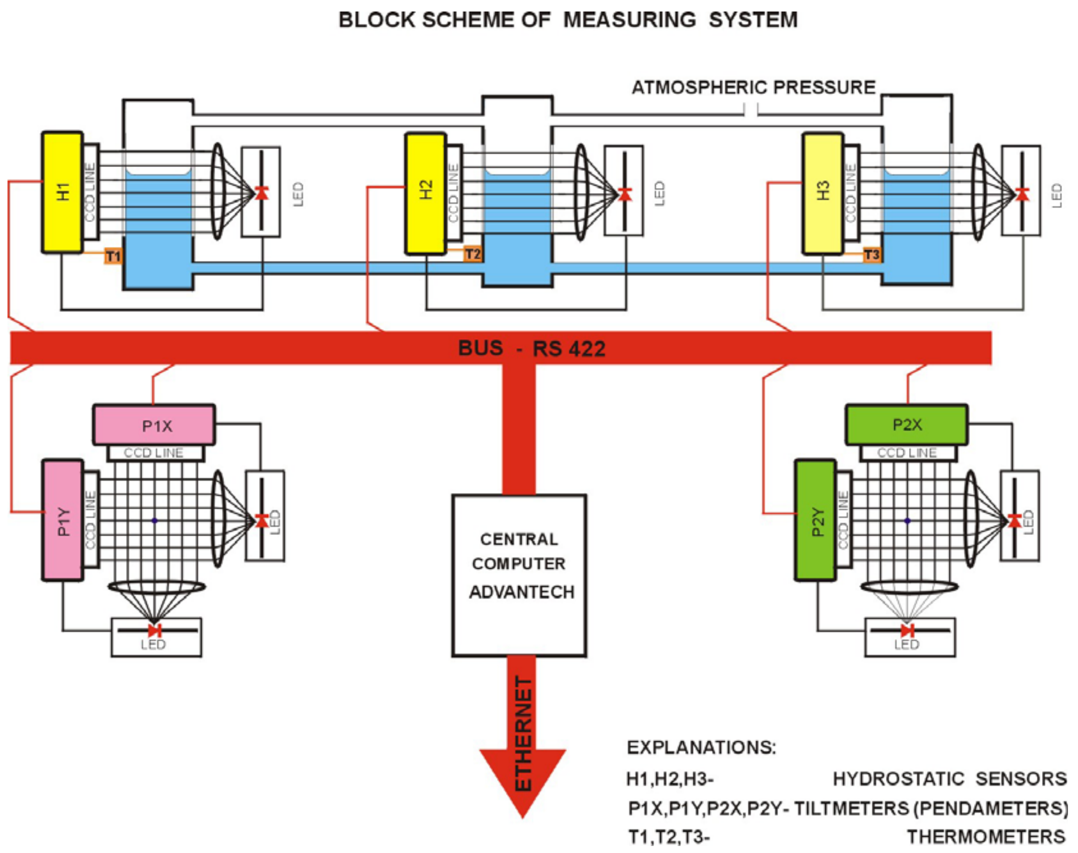


Fig.1 . Measurement system block scheme

4. Data analysis and discussion of obtained results

From Figure 2 can be seen the effects of seismic events on tilt motion of reactor shaft. Either these values are low and pretty far away from those that could call for any action, this information can be generally utilised. The obtained pendametric data can be converted into horizontal plane considered as a part of rotation seismic motions. However, time histories of the registered tilt vectors and their shapes are influenced by the sampling time step that was about 5 minutes. Actually, the measuring system was designed and serves for quasi-static tilt measurement. Therefore, the obtained data suggest that there exist the intermediate values of these slow motions which can be smaller or higher than the registered values. Philosophy of tilt measurements is based on incremental values related to the geodetic measurements executed during the last shut-down of the respective NPP unit. Figure 3 give original uncorrected position of pendametric wire in horizontal plane which corresponds to starting points of pendametric records P1 and P2. Either registrations of P1 and P2 were not executed at the same moment, their common traces indicate the existence of variable intermediate values in quasi-static tilt motions. It has been proved that the safety margin is sufficiently high and in all cases the quasi-static tilt motions have returned back to the original position. In spite of that, the idea of intermediate values recording should be re-estimated and re-assessed.

In terms of safe operation of the power plant, it would be feasible and more valuable to have directly measured information about reactor position changes in terms of displacements, than those obtained by doubled integration with filtering of recorded response accelerations, e.g. from the NPP seismic response monitoring system.

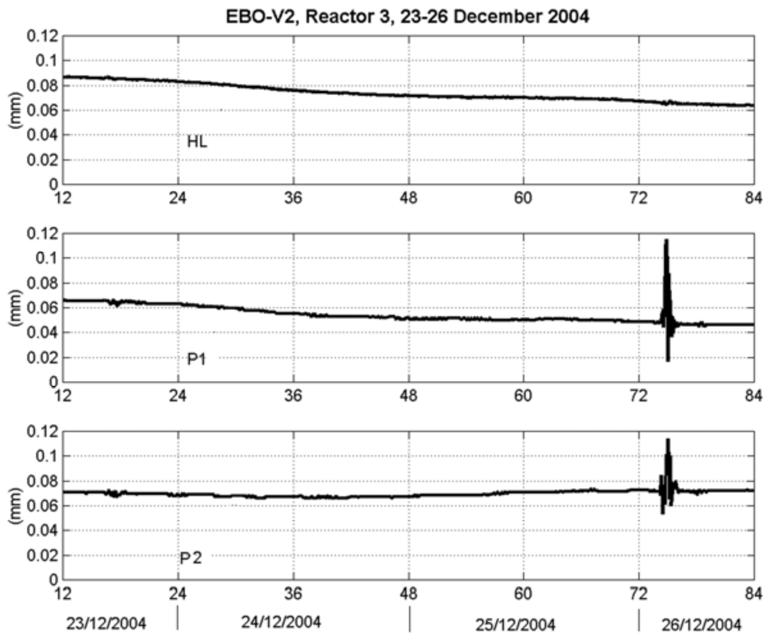


Fig. 2 NPP Jaslovské Bohunice EBO-V2. Reference shaft tilt amplitude records. Records of hydro-levelling (HL) and two pendametric subsystems (P1 and P2) from 23rd – 26th December 2004. The larger perturbation corresponds to the Sumatra earthquake on 26th December 2004, the smaller perturbation corresponds to earthquake between Australia and Antarctica on 23rd December 2004.

5. Conclusions

The structures built on stiff subsoil (hard rock) and on the soft one present different response to the actions of far or near source earthquakes. Any information available from monitoring systems can contribute to the knowledge about motions of NPP structures or, in general, about seismic effects at a site. Uncertainties connected with the very low frequency components in acceleration records can be limited when displacement/rotation supporting records are available.

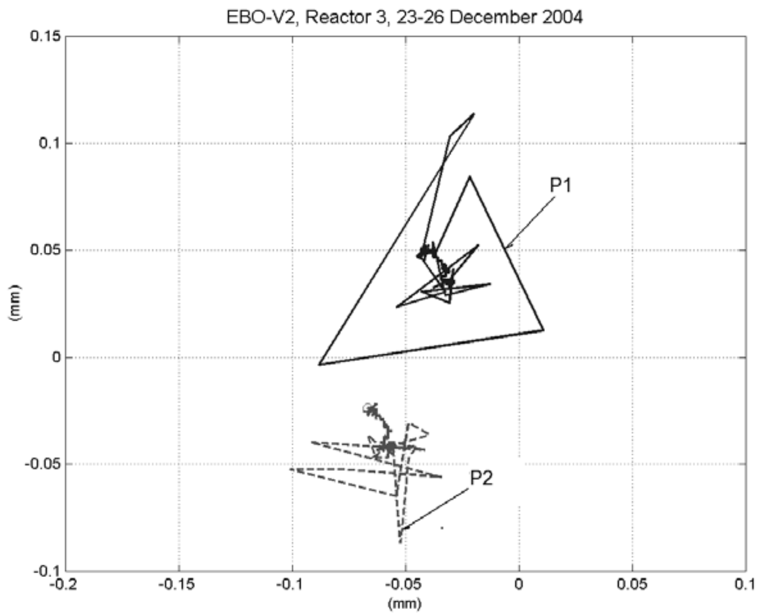


Fig.3 . NPP Jaslovské Bohunice EBO V-2. Reactor 3. Records from Figure 1 transferred into horizontal plane.

Records of such static-dynamic displacements based on measured data of reactor shaft tilt can be utilised in two directions: firstly giving the proof that the temporary tilt of shaft during any earthquake has not exceeded the limit value; secondly providing the information whether seismic displacement or seismic rotation at a site has not exhibited residual quasi-static components. To this end, the prepared special set of instrumentation could cover the frequency range that content is usually unknown and therefore is disregarded.

The movements detected on reactor shaft suggest the option of assessment and evaluation of seismic motions in view of direct displacement records rather than those obtained by acceleration and/or velocity sensors. This approach gives additional data that could be utilised like the proof that the temporary tilt of reactor shaft during the earthquake has not exceeded the limit value. In addition to the basic purpose of these measuring systems, it has been observed their capability to record time histories in very low frequency domain caused by earthquake events. Then, the challenging task is in the analysis of such data and their potential use both for safety control of NPP and for general objectives of earthquake engineering and seismology.

For the sake of more credible and precise measurement of the reactor body position during the earthquake, the option exists to extend and/or improve the operating reliable measuring system. The solution can follow the prevention of the existing measuring system, with consequent completion with the additional units. In such complementary recording units, firstly the measuring frequency should increase app. up to 10 Hz and then faster pendametric sensors would support higher sampling frequency or smaller sampling time step. The full synchronisation of local power plant time with UTC is to be completed.

In view of safe operation of the power plant, is important the fact, that the reactors' positions after the tilt low frequency motions due to far large earthquakes were the same as before the seismic event.

Acknowledgements

The research was partially founded by European Commission DG XII Science, Research & Development in the framework of E-CORE Studies and partially by the Slovak Academy of Sciences and Ministry of Education of SR in the framework of projects VEGA. The financial support is gratefully acknowledged.

References

[1] Ondriš Ľ, Trnovec M, Keppert M, Rusina V, Buzási J. An optoelectronic hydrolevelling system, *Meas. Sci. Technol.*, 5, 1287-123, 1994.

[2] Ondriš Ľ, Lukáč Š, Rusina V, Buzási J, Krušinský D. The optoelectronic systems for vertical shifts and tilt measurement of objects and technological equipment. Proceedings of FIG Working Week 2004 "The Olympic Spirit in Surveying", Athens 22-27 May, Technical Chamber of Greece, Athens. Paper TS3.6/1-13, 2004.

[3] Juhásová E. Seismic Effects on Structures. Elsevier, Amsterdam, 340 p. 1991.

[4] Juhásová E, Ondriš Ľ, Rusina V. Displacements caused by earthquake motions detected in nuclear power plants. In Proceeding of SMIRT 19 – Structural Mechanics in Reactor Technology, Toronto 2007, paper F01/4 -8p.

SFRA Method - Frequency Analysis of Transformers

[1]M. Gutten, [1]M. Brandt, [2]R. Polanský, [2]P. Prosr

[1] Dept of Measurements and Applied Electrical Engineering,
University of Žilina, Slovak republic
[2] Faculty of Electrical Engineering, ZČU Plzeň, Czech republic
Email: gutten@fel.uniza.sk

Abstract. *This article deals with a description of methods of an experimental analysis concerning the actual condition of windings and magnetic circuit of the transformer, which is required in power distribution companies all over the world.*

Keywords: transformer, frequency analysis, measurement methods

1. Introduction

Considering a significance of power transformers in the electric system, their price and possible damages arising in accidents, it is necessary to pay attention to higher prevention of these devices. Windings of power transformers should be designed to avoid various mechanical or thermal deteriorations caused by short-circuit currents occurring in operation.

Besides the permanent deformation effects of short-circuit current, there is also gradual aging process of the electrical device, which can worsen its mechanical properties. Heat shocks can cause decrease of mechanical strength of transformer and consequent unexpected damage of transformer during the operation.

To prevent a damage state of transformers, we perform different types of the measurements that should illustrate an actual condition of the measured equipment. It is therefore important to choose a suitable diagnostics for the right prediction of such conditions.

2. Theory of SFRA method and its importance in transformer diagnostics

SFRA method belongs to current most effective analyses and allows to detect the influences of short-circuit currents, overcurrents and other effects damaging either winding or magnetic circuit of the transformer. This all can be performed without a necessity of decomposition of device and subsequent winding damage determination, which is very time consuming.

The method of the high-frequency analysis (Sweep Frequency Response Analysis – SFRA [2]) is also one of the methods of undisassembling diagnostics of transformers. No intervention to the construction of tested device is demanded, the whole measurement is performed on detached device (not under the voltage). This method is applicable mainly for determination and measuring immediately after the manufacturing of device, i.e. for measuring of reference values. These parameters are consequently compared to the other measurements performed on the transformer, which is decommissioned, after the damages or revisions of transformer etc.

There is possible to detect by SFRA:

- a deformation of winding and its movements or a partial breakdown of winding,
- a short-circuited turn or opened winding,
- a loose and a damaged switching system,
- a core connection problem and a core movement or its wrong grounding.

Results measured on the new transformer can be used as the reference parameters for further comparison with values measured later after certain operation time of the transformer. They can be also compared with the test results performed after the transformer breakdown (or after the n-short-circuits) or repair or it can be used as a diagnostic test, when vibration sensors indicate some potential problem in transformer.

According to [3] SFRA method determines the transformer responses in a time or frequency area. The time response measurement provides curve determination of the time response to the specific voltage impulse applied to winding input connection. The frequency response measurement consists in determination of amplitude eventually phase response to the harmonic voltage of variable frequency applied to winding input. While the time response is the record of time behavior of voltage, frequency response is the amplitude response dependence on frequency.

A relation between the response and the winding condition is definite, otherwise it is complicated. It is impossible to expect the assessment of concrete damage of winding from differences in response behaviors. The measurement results lead us only to a statement of the fact that some change of winding condition really occurred. Such test results are very helpful to decide, whether it is unavoidable to open and revise the transformer or not.

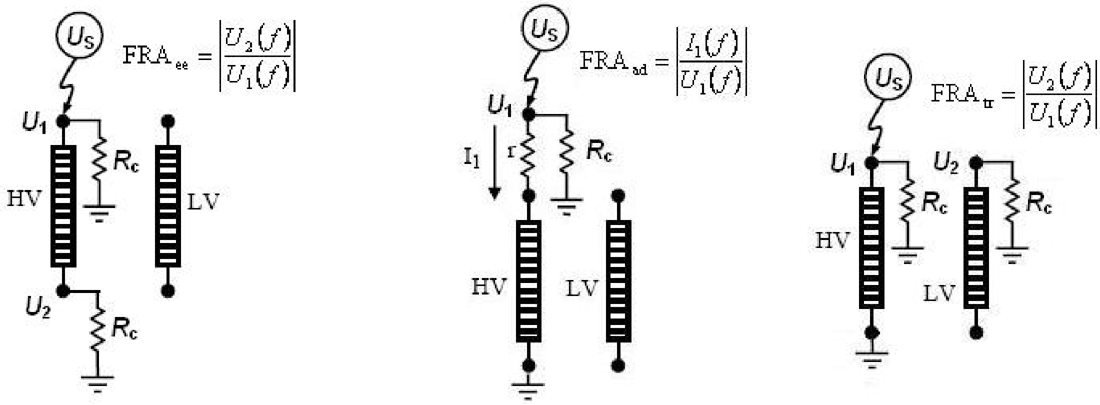
3. Measuring principles

The frequency response characteristics of windings can be obtained using either the impulse frequency response analysis (IFRA) method in the time domain or the sweep frequency response analysis (SFRA) method in the frequency domain [4]. In principle the two methods give the same results if the same connection method is used. However a frequency domain measurement using a method which records the ratio of the input and output voltages over the frequency range by using a sequence of narrow band spot measurements has been found to be particularly suitable for obtaining measurements in an electrically noisy environment. Making a series of narrow band measurements increases the signal to noise ratio and the dynamic range available. Measuring only at the exciting frequency also prevents any non-linearity of the test object (not usually a problem at the small signal levels employed) from affecting measurements at different frequencies. The measurement using this technique is conveniently made using a network analyser or similar instrument. This produces a frequency-varying sinusoidal voltage signal, applied to one terminal of the test winding with the input voltage being measured by a separate cable at that terminal and the response to this input measured at another terminal.

The three FRA measurements (SFRA or IFRA) connection methods most commonly employed in the industry are chosen for comparison [4]. These FRA connection methods are identified as:

a) End-to-end voltage ratio measurement (FRA_{ee})
b) Input admittance measurement (FRA_{ad})
c) Transfer voltage ratio measurement (FRA_{tr})

FRA measurements can be expressed using terminal voltages and currents as shown in Fig. 1. All measuring cables are terminated with matching impedance of the cable, R_c (typically 50 Ω) and the winding terminal voltages are measured across R_c as shown in Fig. 1. For input admittance measurement, a resistive impedance r (typically $r = R_c$) is used to measure the input current.



a) End-to-end voltage ratio b) Input admittance c) Transfer voltage ratio

Fig. 1. Test connection for FRA (SFRA or IFRA) measurements.

4. Behaviors of transformer winding responses by SFRA method

The behaviour of transformer winding response reflects e.g. electromagnetic couplings between the winding and transformer tank, also between the primary and secondary (eventually tertiary) winding, between the windings of particular phases or be-tween turns themselves of particular windings.

The power transformer measurement requires a setting up of the frequency range from 10 Hz to 2 MHz (Fig. 2), whereas there is necessary to follow the right measuring technique to prevent various inaccuracies and faults. Input parameter of measuring system is voltage with value 10 V and its output parameter is current response (0÷90 dB) to change impedance for respective default frequency.

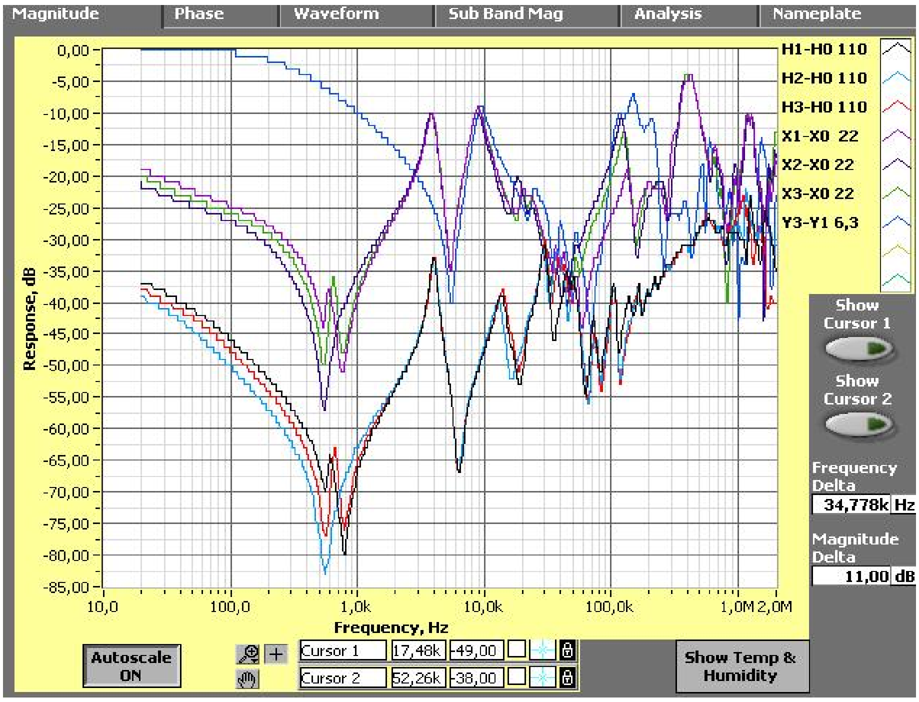


Fig. 2. Magnitude Chart of three-winding transformer by M5100 measuring system

During the open circuit tests a mechanical condition of tested winding and ferromagnetic core is detected. The following curves typical for this measurement provide us important information about changes in the core, which are visible in low frequencies, while higher frequencies refer to problems such as winding movements or turn-to-turn fault.

The application of analysis of phase attenuation depending on frequency is suitable for more complete evaluation of winding condition. This analysis enables to assess the processes of winding movements during the particular short-circuits influences.

Problems with core grounding or shorted laminates in the core will typically change the shape of the lowest section of the curve (to 10 kHz). Mid frequencies (from 10 kHz to 200 kHz) represent axial or radial movements in the windings and high frequencies indicate problems such as e. g. winding knocking or problems with contacts.

During the short circuit tests only the winding condition in primary or secondary part of transformer is detected. This measurement notifies reliably of deformation of inner winding and its movement as a result effects of short-circuit currents.

5. Conclusions

Problem of the frequency analysis of transformers by SFRA method is very comprehensive and its application becomes interesting for many transformer manufacturers and operators. From the long-term point of view the SFRA method is supposed to be very useful and it provides enough information on tested transformers. These transformers have their reference data obtained by the manufacturers, suitable for the comparison with further data of particular transformer.

SFRA testing method represents one of the most effective alternative diagnostic methods compared to visual check. This method allows to detect the effects of the short-circuit currents, whereas we are able to evaluate the mechanical strength action on the transformer winding during previous operation It is also possible to identify the specific winding phase, which has been mostly influenced by the short-circuit currents, without a necessity of transformer dividing, which would be very time consuming.

There is also possible to identify the size of parameters of other electric machines as well as the position of their resonance frequency from particular curves (analysis of damage winding).

Acknowledgements

This work was supported by the Grant Agency VEGA from the Ministry of Education of Slovak Republic under contract 1/0548/09.

References

[1] Gutten M., Brandt M., Polanský R., Prosr P.: High-frequency analysis of three-winding autotransformers 400/121/34 kV, ADVANCES in EEE, No.1-2, 7/2008, Žilina, Slovakia

[2] http://www.doble.com/products/sweep_frequency_response_analysis.html, 15. 1. 2007.

[3] Kvasnička V., Procházka R., Velek J.: Verification of method frequency characteristics in control room of distribution system Czech Republic, In Diagnostika 05, Plzeň 2005, CR

[4] Jayasinghe J.A.S.B., Wang Z.D., Jarman P.N., Darwin A.W.: The Winding Movement in Power Transformers: A Comparison of FRA Measurement Connection Methods. IEEE Transactions on Dielectrics and Electrical Insulation Vol. 13, No. 6; 2006, Canada

[5] Kúdelčík J.: Generation of breakdown in the liqiud dielectrics, 16.Conference of Slovak Physics, 2007, Žilina, Slovakia

A Novel Proceeding for Optical Coordinate Measuring Machines to Locate Deviations Behind an Undercut

M. Rückwardt, A. Göpfert , S. Lerm, M. Rosenberger , M. Schellhorn, G. Linß

Technical University of Ilmenau, Ilmenau, Germany,
Email: matthias.rueckwardt@tu-ilmenau.de

Abstract. *The precise measurement of geometrical dimensions represents the basis of quality assurance in the industrial production. Tactile measurement machines are often replaced by optical non-contact measuring procedures to speed up the measurement process and to minimize the effort needed to ensure the product quality. Within a very short time, optical coordinate measuring machines acquiring numerous measuring points. However the feasibility of a measurement depends on the optical path of light. For this reason it is extremely difficult or partly not possibly to measure undercuts. In these cases it is necessary to deflect the beam of light.*

This paper presents a novel proceeding for optical coordinate measuring machines to locate deviations behind an undercut developed by the Department of Quality Assurance at the Faculty of Mechanical Engineering of the Technical University of Ilmenau. The new approaches are described by the example of measurement of an undercut groove on the inside of a cylinder. Therefore several possibilities for beam deflexion, illuminations scenes, foci criteria, minimum of detecting field and the needed optical magnification were discussed.

Keywords: optical measurement, detection of hidden quality features, deflexion of light and camera beam

1. The defficulties of tactile measurement behind an undercut

Dimensional measurements behind undercuts are always very difficult. As described in [1] tactile coordinate measuring machines are not fast enough for total inspection in the productions process. Beyond that there are further problems with the tactile sampling process. For the measurement behind undercuts it is necessary to use a star or angulate stylus to reach the area of interest. This can be quite impossible when the inner diameter of the cylinder is not enough space for the geometrical form of the stylus. Even if sufficient place is present, a correct measurement is not guaranteed. Special problems result for example in the case of the measurement of grooves in particular through the stylus tip radius [Fig.1].

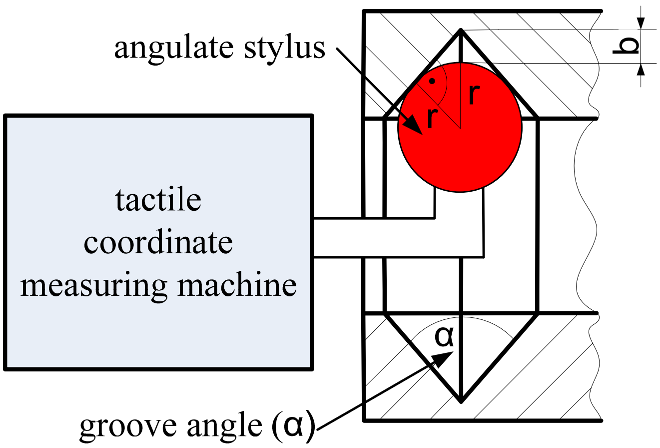


Fig.1 A principal design to measure behind an undercut by using tactile coordinate measuring machines.

For Example a fine stylus with a radius of 500 µm and a typical groove angle α of 90° the radial error of measurement is almost b = 210 µm, [Eq.1],[Eq.2]. Although it is a systematic error and could be corrected, the groove angle can vary and is - in worst cases - unknown.

$$\sin\frac{\alpha}{2} = \frac{r}{b+r} \tag{1}$$

$$b = r \cdot \left(\frac{1}{\sin\frac{\alpha}{2}} - 1 \right) \tag{2}$$

α groove angle
r radius of the stylus
b error of measurement to the ground of the groove

For this reason it is expedient to use optical coordinate measuring machine for this problem. They are able to measure in very small inner diameter of a cylinder with an appropriate beam deflexion element like a prism [Fig.2]. Due to the general character of this paper boundary conditions were neglected and the constructive beam deflexion solutions for special matters are not mentioned here.

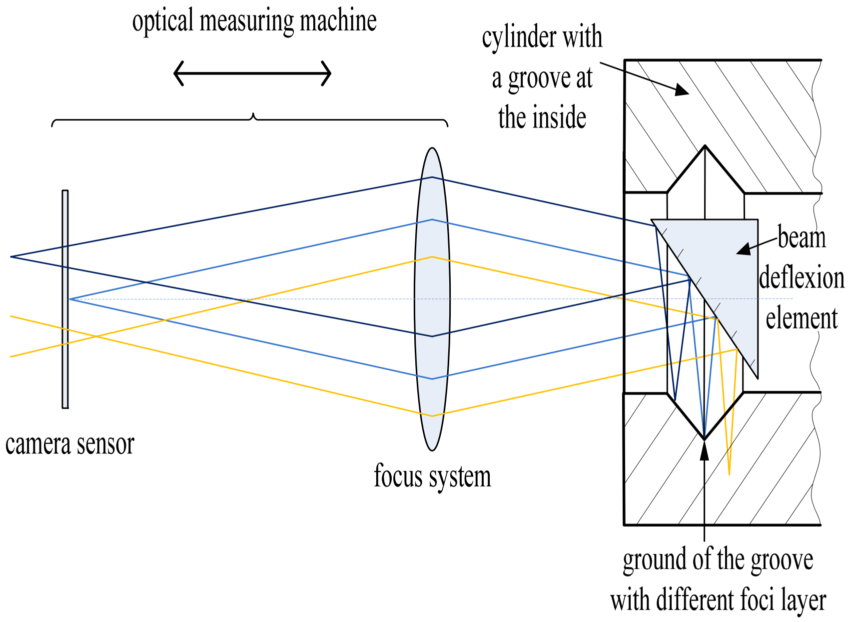


Fig.2 A principal design to measure behind an undercut by using optical coordinate measuring machines.

2. Finding the parameters to measure behind an undercut with an optical coordinate measuring machine

In the field of this research the main focus lay in the determination of the optimal measuring parameters. This covered optical magnification, illuminations scene, focus criteria and minimum of detecting field to find plenty enough and stable measuring points on the ground of a groove which is found on the inside of different cylindrical elements. Different extreme cases for the optical measurement were examined. The selected units under test covered different material classes for example metal with high reflexion character, plastic in different colours and surface roughness and even transparent plastics. The transparent objects represent a special challenge for optical measurement. That is because of the fact that detected measuring points can, as desired, part of the ground of the groove but in addition they could occur through material inclusions, reflexions on a particle or dust or at the backside of the groove. With problematic measuring parameters it is possible that no detectible reflexion occur. On account of this the results of transparent object will be named.

A typical optical measuring machine has got an uncertainty of measurement of 4 up to 5 µm and a confocal incident light, as in [2]. The presented results were accomplished on the optical coordinate machine UNI-VIS 250 from Mahr OKM GmbH Jena. With a planar mirror, which is placed at an angle of 45° to the optical axis, it is possible to detect the ground of the groove. By using different magnifications and foci criteria variable areas of interest (AOI) are placed on the ground of the groove to find detectable points. The measurements were ten times repeated.

Table 1. Researched parameters for finding the fewest standard deviation for the z-coordinate (in ascending order in µm).

parameter			**tranpsrent cylinder**	
size of AOI (µm)	**magnifi-cation**	**foci-criteria**	**standard deviation**	**range (max.-min.)**
40x40	5x	scattering	0,56	1,60
80x80	5x	scattering	1,79	4,80
80x80	5x	noise	1,95	5,30
40x40	3x	noise	2,23	6,70
100x100	3x	scattering	2,28	7,60
80x80	5x	contrast	2,30	6,60
80x80	1x	scattering	2,53	8,80
80x80	3x	scattering	2,63	9,00
100x100	5x	contrast	2,70	9,20
40x40	5x	contrast	2,71	9,60
40x40	3x	scattering	2,82	8,00
40x40	3x	contrast	3,25	10,50
80x80	1x	noise	3,48	11,00
80x80	3x	contrast	3,59	13,10
40x40	5x	noise	3,66	11,90
100x100	1x	scattering	3,81	13,30
100x100	3x	contrast	3,87	11,80
100x100	1x	noise	3,92	13,70
100x100	1x	contrast	3,93	12,50
40x40	1x	scattering	4,59	12,60
80x80	1x	contrast	4,69	14,30
40x40	1x	noise	4,74	15,50
100x100	5x	noise	4,77	14,30
40x40	1x	contrast	4,88	16,80
80x80	3x	noise	5,07	16,00
100x100	5x	scattering	6,31	18,20
100x100	3x	noise	8,37	29,50

The encountered coordinates, especially the z-coordinate recovered during the focusing, is needed to get the diameter of the groove by using a coordinate transformation into a polar coordinate system. Thus, the standard deviations and the ranges of the z-coordinate are listed in Table 1 for all measurements to find the best parameter. As seen there, it is possible to get very stable results with a standard deviation under 1 µm.

But therefore it is necessary to heed some parameters. A higher magnification depends on a smaller AOI, because resolution is rising and the field of view shrinks. Thus more details are visible whereby its more difficulty for the focus algorithms to detect accurate an point. As consequence several small AOIs should be used as a large AOI to detect those quality features.

It is very interesting that the contrast focus criterion does not provide the best results although it is so often used in machine vision. Theoretical the focus point depends only on the object-wide because the frequency content of the image does not change, see also [2]. But on average the scattering criterion is about 1 µm better than the contrast criterion. So it is the best in the test with also the lowest range for the maximum to the minimum.

3. Conclusions

Optical coordinate measuring machines are faster and so more appropriate than tactile in the productions process for total inspections. With a beam deflexion element precision measurements behind an undercut are possible. Therefore it is shown here that a standard deviation under 3 µm can be easily reached. The best parameters for this aim are a high magnification with a small sized AOI from about 40 x 40 µm in combination with the scattering focus criterion.

Acknowledgments

First of all my thanks go to Prof. Dr.-Ing. Gerhard Linß, for giving me the opportunity to work on this project. Furthermore I want to thank my colleges and my family.

References

[1] H.-J. Gevatter, U. Grünhaupt, Compendium for measurement and automation technology, Handbuch der Mess- und Automatisierungstechnik in der Produktion, Springer, Berlin, 2006.

[2] S. C. N. Töpfer, Automated edge probing for high resolution optical geometric measurement, Automatisierte Antastung für die hochauflösende Geometriemessung mit CCD-Bildsensoren, Ilmenau University of Technology, Dissertation, 2008

Two Solutions of the Quantum Imaging X–ray CT System based on GaAs Radiation Detectors

[1]J. Přibil, [2]B. Zaťko, [1]I. Frollo, [2]F. Dubecký

[1]Institute of Measurement Science, SAS, Dúbravská cesta 9, Bratislava, Slovakia
[2]Institute of Electrical Engineering, SAS, Dúbravská cesta 9, Bratislava, Slovakia
Email: umerprib@savba.sk

Abstract. *The work presents two generations of developed portable quantum X–CT mini–systems which utilizes monolithic semi–insulating GaAs radiation detectors. This contribution describes present status of the assembling of new portable X–ray CT mini system. Performed measurement of the performance of the SI GaAs detectors and the integral spectra of ASIC DX64 readout chips are also mentioned.*

Keywords: X–ray imaging, Single photon counting detection, Image reconstruction

1. Introduction

The X–ray computer tomography (X–CT) is a non–destructive testing method able to evaluate the inner structure of investigated objects [1], [2]. The cross–sectional imaging is achieved by the mathematical reconstruction of projections of a tested object. Such a projection is an intensity image of transmitted X–ray photons through the evaluated object. Small CT imaging instruments for the evaluation of small animals have been developed preferably for operation in the positron emission mode (PET). Almost all commercial X–CT systems use photodiodes as detectors that are covered by a scintillator [3].

This paper makes a comparison of two developed portable quantum X–CT mini–systems (XCTMS) which utilizes monolithic semi–insulating (SI) GaAs detectors. The solution of these two X–CT devices was also different in the approach of sample and X–ray source/detector arrangement. Both devices are using the same principle of image reconstruction from projections [1], but modifications of the basic reconstruction algorithm to eliminate some imperfections following from physical properties of used X–ray detector and utilized scanning technique must be applied [4].

2. Subject and Methods

In the first generation of realized X–CT device was used the construction solution, where the sample is rotated and can be linear moved in the *x*-axis for changing the reconstruction geometry (zoom-in/zoom out), the X–ray source and detector are stationary [4] (see Fig.1). This device use the active X–rays source with energy of photons up to 80 keV, and the digital X–ray scanner consisting of 17 monolithic 24 strip line detectors named "SAMO" based on bulk SI GaAs [5].

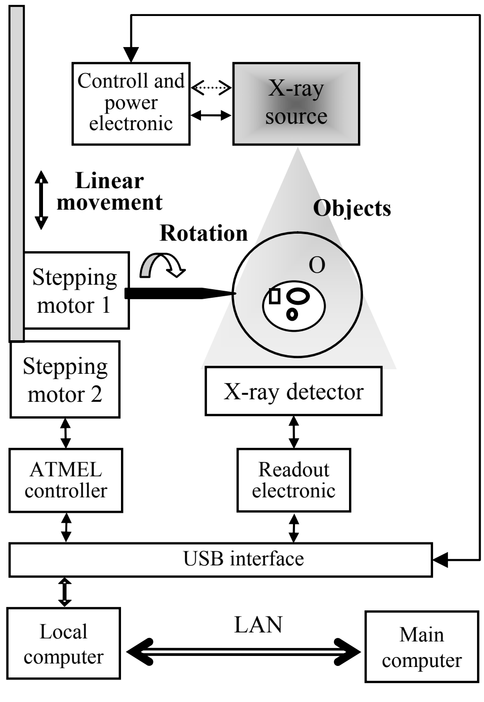


Fig. 1. Block diagram of the 1st X–CT

In the second generation of XCTMS the different construction approach was applied. An evaluated object is located between the X–ray source and the detection unit. One of the stepper motors ensures a full rotation of the X–ray source & detection unit coupled system around the object (see Fig. 2a). The other stepper allows for its linear positioning along the *z* direction. As a fan–beam configuration (diverging beam) is used, the micro–adjustable slit ensures that the width of the X–ray beam is little wider than detector double array width and irradiates only active area of the pixel detectors.

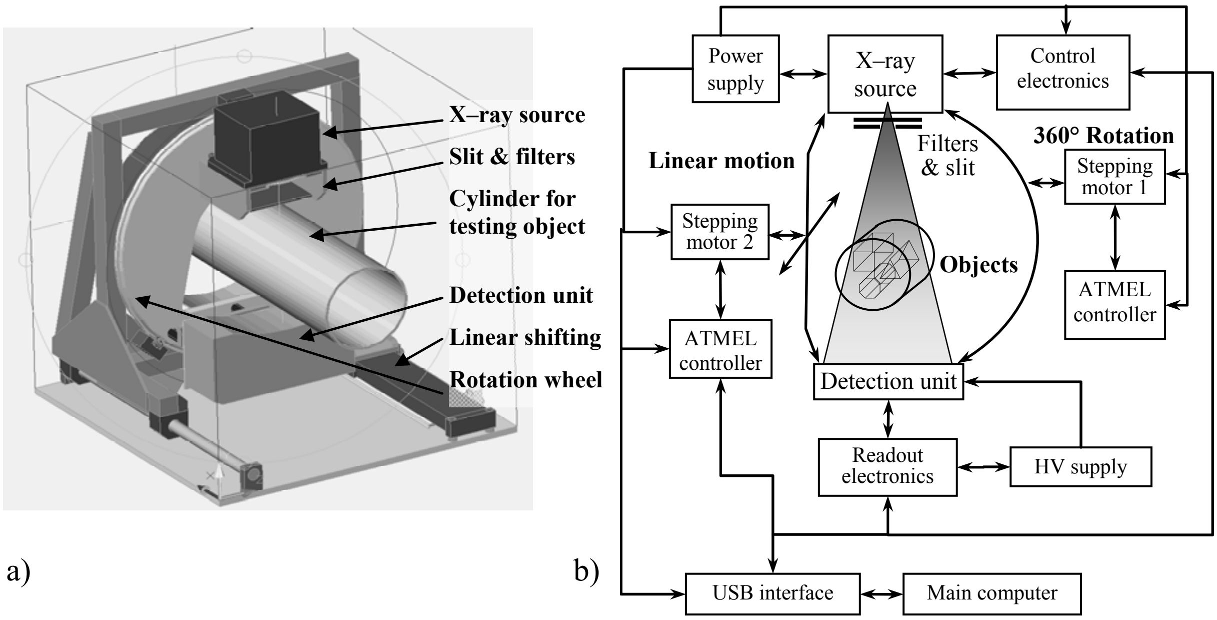


Fig. 2. The developed new portable XCTMS: 3-D visualization (a), control block diagram (b).

The mechanical construction of the XCTMS has following structure: the X–ray source, connected with the filter and the slit holder, is positioned in the top part of an aluminium wheel, which is driven by the stepper motor. The cylindrical holder is used to fix the tested objects. This XCTMS device is portable and consists of the following parts (see block diagram in Fig. 2b): the X–ray source, the detection unit, two stepper motors, which control the rotation and the linear movement of the X–ray source coupled with the detection unit around the object evaluated, and a local high–performance PC with a touch–screen monitor allowing for comfortable control of all system elements.

The X–ray source (Source–Ray Inc.: Model SB–80–500) with a tungsten anode operates in an accelerating voltage range of 35–80 kV at a maximum current of 500 μA with a maximum focal spot size smaller than 46 μm at a maximum power of 40 W [6]. X–rays generated by the source pass through two filters, selectable from 2×8 positions in two rotatable carousels, followed by a micro–adjustable slit, which restricts the width of the cone (into which X–ray photons are emitted) into a narrow, almost one–dimensional beam.

The detection unit consists of 16 input mini–modules, each with 2×64 monolithic SI GaAs pixel detectors with a pitch of 250 μm arranged along two lines on the chip and assembled into an arc. The detection double array incorporates 2×1024 pixels over a total length of 261.5 mm. A small gap, corresponding to 1.5 pixels, is maintained between each two neighbouring modules to prevent the outer pixels from damage. The total length of the double array on the chip is 16.25 mm. The operational bias of the SI GaAs detectors is ranged between 150 and 300 V. The pixel detectors are dc–coupled to the inputs of two ASICs–type DX64 readout chips [7], by wire bonding via pitch adapters, which adjust the pitch of the pixel detectors to the input pads of the ASICs. The readout chips have two discrimination levels and 20–bit counters for each channel; hence one energy window is available within one evaluation scan. The devices are glued and wire–bonded onto an input PCB (printed circuit board) with dimensions of about 16 mm × 120 mm. The PCBs are fixed onto Peltier coolers, which stabilize the working temperature of the detectors and the input electronic circuitry at about 15° C using a thick Cu holder.

We have used the filtered back projection (FBP) for image reconstruction from projections reconstruction because of sufficient speed to accuracy ratio [1]. Nevertheless the FBP needs many projections (several hundreds or more) to obtain a correct image because of many effects coming into X–ray sinogram such as fluorescence, X–ray scattering, beam hardening, phase effects, etc. Perspective approach uses an iterative scheme of tomographic reconstruction. Its advantage is that we can get a good image even from noisy data and low number of projections [1]. In the case of first generation of the X–CT device, the parallel beam or fan beam with equidistal detectors geometry can be used. For the new XCTMS with rotating X–ray source and detection unit system, the fan beam with equiangular detectors geometry will be used.

3. Experiments and Results

Before implementation of image reconstruction algorithm to the developed portable XCTMS, used approach was tested on the experimental equipment realized in two year ago [7]. For giving good X–ray images some imperfections must be eliminated. The imperfections follow from physical properties of used devices and applied scanning technique. Used type of SI GaAs detector generates the frequency noise, which also depends on the temperature. The high frequency component of this noise brings a “background” texture of reconstructed object. For that reason, the low–pas (LP) filtration on the slices of projection data matrix can be performed. On another hand, the filtration produces the blur edge effect. Therefore, the parameters of designed LP filter must be set as a compromise between two mentioned antagonistic requirements. The tested objects were placed in the maximum distance from the X–ray source and near upon the scanner. Consequently, emitted beams can be considered as parallel and the image reconstruction method based on parallel beam can be used (that be characterized by a lower computing complexity than the fan beam method). Finally, for comparison, the fan beam reconstruction method was also applied – see Fig. 3.

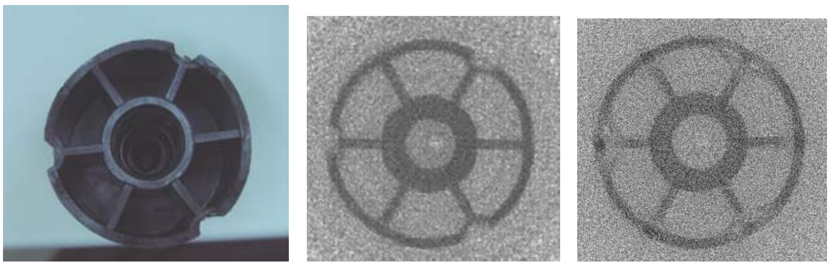


Fig. 3. Image reconstruction on the testing object: photography of plastic tube with six capsules, diameter 40 mm (left), finally reconstructed image by the parallel beam method – 100 projections per 3.6 deg (middle), reconstructed image by fan beam method (right).

The second generation of XCTMS is now almost finished. The measurement of the performance of the SI GaAs detectors and the integral spectra using ASIC DX64 readout chips were also performed. In our experiments we used eight detectors with four different diameters of the Schottky contact. The measured reverse I–V characteristics of all detectors at RT are depicted in Fig. 4a). Measured integral spectra (dependence between the total count of the channels and adjusted threshold) of SI GaAs detectors are demonstrated in Fig. 4b). Right edge of each spectrum represents impulses generated by 59.5 keV γ–photons.

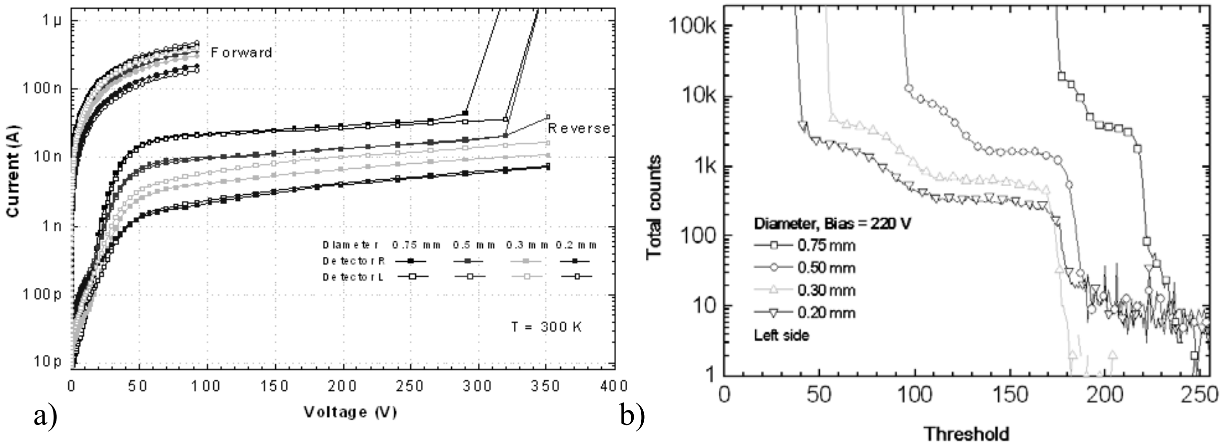


Fig. 4. Measurement of used SI GaAs detectors and ASIC DX64 readout chips: current–voltage characteristics of fabricated detectors (a), integral spectra of ^{241}Am: detector set 1 (b).

4. Conclusions

Two portable quantum X–CT mini–systems based on monolithic GaAs radiation imaging detectors have been realized. The modulation transfer function (MTF) of the first generation X–CT device was not tested up to now, because we have not an applicable phantom object. The detection homogeneity and the performance of the GaAs monolithic detectors connected to 24 channel readout chain were measured and tested; obtained results are introduced in [8].

The current assembling system is able to operate as an X–ray scanner or X–CT equipment with fixed objects up to a diameter of 180 mm and 250 mm in length. The detection part consists of a double array of 2×1024 SI GaAs detectors with a pitch of 250 μm. Two discrimination levels of the readout chip enable the energy selection of incident photons. A fine tunable slit mounted behind the X–ray source output guarantees a minimal dose for living objects. The expected spatial resolution of the device is 125 μm.

Acknowledgements

This work was partially supported by the Agency for Promotion of Research and Development under contracts no. APVV–0459–06 and APVV-99-P06305.

References

[1] Kak A. C.,Slaney M., Principles of Comp. Tomogr. Imaging. IEEE Press, 1999.

[2] Jakůbek J., *Nucl. Instr. and Meth.* A 576, 223, 2007.

[3] Wu H. et al., Book of Abstracts, IEEE NSS–MIC–RTSD. Dresden, 2008, p. 545.

[4] Přibil J. et al., Experiments with Application of Image Reconstruction Method Based on Perspective Imaging Techniques in X-ray CT Mini System. In proceedings of the *15th* International Conference IWSSIP'08, 2008, 33-36.

[5] Dubecký F. et al. Digital X-ray portable scanner based on monolithitic semi-insulating GaAs detectors. *Nucl. Instr. and Meth. in Phys. Res.* A 546: 118-124, 2005.

[6] The Model SB-80-500 SourceBlock Inst./Oper. Manual. Source-Ray, Inc. ©2002.

[7] Gryboś P. et al., Conference Record, IEEE Nucl. Sci. Sympos., 2006, N17-5, 693-696.

[8] Zaťko B. et al. Performance study of monolithic line radiation detector based on SI GaAs using X-ray source. *Nucl. Instr. and Meth. in Phys. Res.* A 551: 78–82, 2005.

Preparation and Properties of YBCO-PE Composites

[1]A. Cigáň, [1]J. Polovková, [2]I. Krupa, [1]A. Koňakovský, [1]J. Maňka

[1]Department of Magnetometry, Institute of Measurement Science
Slovak Academy of Sciences, Dúbravská cesta 9, 841 04 Bratislava, Slovakia

[2]Department of Special Polymers and Biopolymers, Polymer Institute
Slovak Academy of Sciences, Dúbravská cesta 9, 842 36 Bratislava, Slovakia
E-mail: julia.polovkova@savba.sk

Abstract. *Composite samples consisting of YBCO superconductor and low-density polyethylene have been prepared. The standard resistance four-point method and contactless induction method based on Meissner-Ochsenfeld effect were used for characterizing transition properties, e.g. critical temperature of sintered YBCO precursor and composite samples, respectively. In all composite samples, no decrease of the critical temperature of YBCO precursor filler was observed.*

Keywords: High temperature superconductor, YBCO, polyethylene (PE), composite, magnetization, critical temperature

1. Introduction

It has been more than twenty years since a new class of high temperature superconductors, HTS, has been discovered in ceramic materials containing Cu-O_2 layers. Today, many interesting future industrial applications are already being laboratory prepared and tested, such as superconducting conductors, motors, and generators with high power and torque at smaller dimensions, superconducting fault-current limiters, superconducting magnets, superconducting energy storage systems (rotation, magnetic), superconducting levitation transportations, bearings, etc. [1-3]. On the other side, HTS show poor mechanic properties as low fracture toughness and high brittleness as well as chemical instability under ambient conditions, generally. So, more different routes as to improve the properties are being found. One of them is using polymers. Composites, where a superconductor serves as a filler of polymer matrix, represent more frequently studied compositions. The $YBa_2Cu_3O_y$ (YBCO) superconductor is the most frequently used superconducting filler in the superconductor-polymer (S-P) composites [4-6]. This is a result of its relatively good phase stability and good control of processing conditions. In addition, many different polymer agents were studied in S-P composites, e.g., PVC, PVB, PVA, PA, PPS, epoxy resins, etc., [7-9]. However, with respect to the preservation of superconducting properties of the S-P composites, it is important to use polymer agents or components with relatively low processing temperature, e.g., melting, curing, and decomposition temperatures.

In this paper, we present the preparation and properties of a composite superconducting material containing $YBa_2Cu_3O_y$ and low-density polyethylene using the solid-state reaction method of syntheses of YBCO precursor and contactless induction method of measurement of superconducting properties of composites. The induction method is more suitable for the composites that consist of electrically insulated superconducting grains.

2. Subject and Methods

Superconducting precursors with the nominal composition $YBa_2Cu_3O_{7-\delta}$ have been prepared by the standard solid-state reaction method from Y_2O_3, CuO oxides and $BaCO_3$. The powder mixture in appropriate weight amounts has been homogenized in an agate mortar and calcined in air at 915 °C for 24 h. The obtained precursors have again been homogenized in the agate mortar, pressed into pellets (with the diameter of 12 mm) under the pressure of 200 MPa and sintered in flowing oxygen (10 ml/min) at 1007 °C for 24 h; then cooled to 520 °C and held at this temperature for 24 h, and thereafter cooled in a furnace to the room temperature. As polymer matrix, we used the low density polyethylene Bralen RA 2-19 product of Slovnaft Petrochemicals, Ltd., Bratislava, Slovakia. (MFI = 1.7 g/10 min., density = 0.916 g cm^{-3}, particle size < 50 μm, specific enthalpy of melting = 109.8 J g^{-1}, melting temperature = 107 °C). The superconducting and PE powder precursors have been homogenized in an agate mortar and the mix of powders was used to obtain (1-x) wt. % $YBa_2Cu_3O_{7-\delta}$ +x wt. % PE composite samples, by hot pressing at temperature 130 °C under the pressure of 3 MPa for 3 minute. The samples in the form of the pellets of the diameter of about 12 mm and of the thickness of 1 mm had mass from 0.13 to 0.34g depending on the content of superconductor. The content of YBCO in composite changed from 30 to 85 % of the sample weight.

The critical temperature $T_c(R=0)$ of the precursor YBCO samples was determined by the standard resistance four-point method and the transition width ΔT_c was characterized by the 10-90 % criterion. Another contactless electromagnetic induction method based on Meissner effect was used for the measurement of transition properties of the S-P composite samples. The method using a change of the mutual inductance of two coils separated by a sample is based on the Meissner-Ochsenfeld effect - exclusion of the magnetic field from the interior of superconductor resulting in changes of induced voltage (U_1) in the secondary coil [10]. AC volume and mass magnetization (M and M_m) characteristics were measured in detail by a compensation method using the second-order SQUID gradiometer [11]. All magnetization characteristics were measured at 77.3 K after the zero-field cooling in applied magnetization field H_a with the frequency of 0.1 Hz and amplitude ranging from 10^{-1} to 10^5 Am^{-1}. H_a was parallel to the axis of the sample. The superconducting precursor for the composite was prepared by powdering of the $YBa_2Cu_3O_{7-\delta}$ sintered samples.

3. Results

Typical dependence of the electrical resistance (R) vs. temperature (T) of a precursor sintered sample before powdering in the range of transition into the superconducting state is in Fig.1.

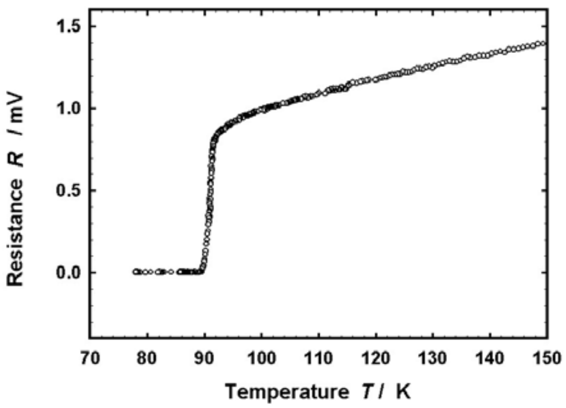


Fig. 1. R vs. T dependence of the $YBa_2Cu_3O_{7-\delta}$ precursor sintered sample before powdering.

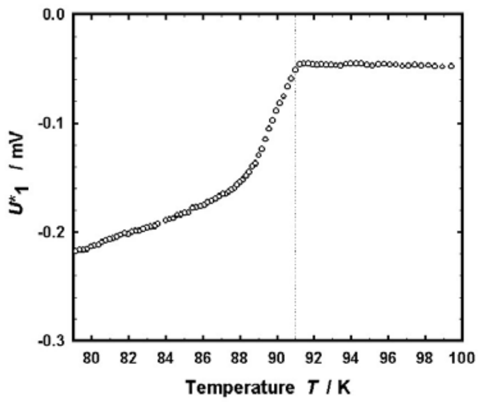


Fig.2. U^*_1 vs. T dependence of YBCO-PE composite sample with 60 wt. % of YBCO powder at frequency of 1008 Hz near the transition into the superconducting state.

From measured data, values of the critical temperature Tc(R=0) ~ 91 K and the transition width ΔTc ~ 1.3 K can be determined.

The dependence of the induced voltage U*1 vs. T of composite sample with 60 wt.% of YBCO powder is shown in Fig. 2. The symbol * denotes that an effect of background (a sample-free probe) was subtracted from measured data. The characteristic temperature Tcon determined from the onset of a diamagnetic behaviour (decrease in the induced voltage U*1) is usually used as critical temperature in case of induction method. So, Tcon is about 91 K for all composite samples.

The volume magnetization M versus applied field Ha dependences of YBa2Cu3O7+δ precursor sample before powdering and composite sample consisting of 85 wt. % YBa2Cu3O7+δ and 15 wt. % polyethylene powders for the low amplitude of Ha are in Fig.3.

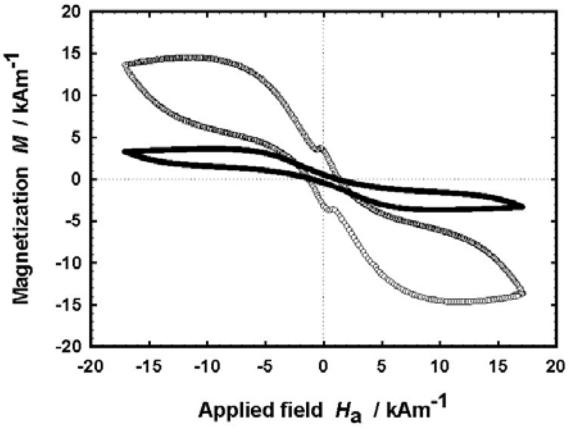


Fig. 3. M vs. H_a dependences of $YBa_2Cu_3O_{7+\delta}$ precursor sample before powdering (○) and composite sample consisting of 85 wt. % $YBa_2Cu_3O_{7+\delta}$ and 15 wt. % polyethylene powders (●).

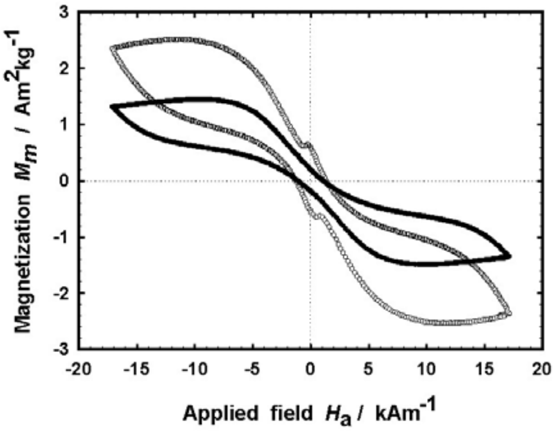
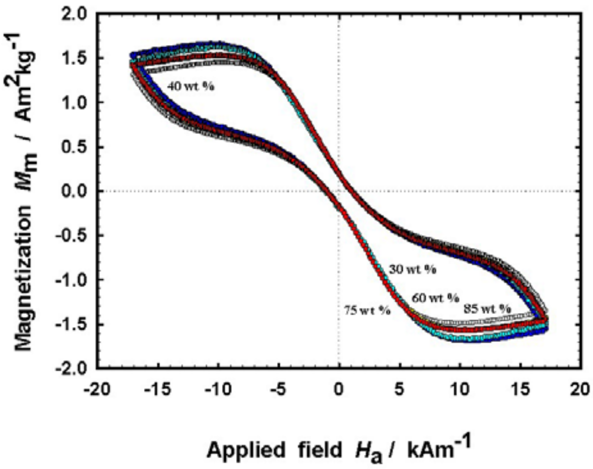


Fig. 4. M_m vs. H_a dependences of $YBa_2Cu_3O_{7+\delta}$ precursor sample before powdering (○) and composite sample consisting of 85 wt. % $YBa_2Cu_3O_{7+\delta}$ and 15 wt. % polyethylene powders (●).

The difference between sample magnetization curves can be ascribed to inter-grain junctions and to different YBCO content. The magnetization peak nearly at zero field of YBCO precursor sample before powdering results from intergrain junctions, whereas the composite samples show no peak. From this, it could be inferred that superconducting grains in the composite sample are not significantly electrically connected. The rate of grain junction magnetization contribution of the samples can be inferred from Fig. 4, where the mass magnetization Mm vs. Ha dependences of the samples are shown. Nevertheless, different sizes or morphology of grains in the two samples could also play some role.



Mass magnetization Mm vs. applied field Ha composite samples with 30, 40, 60, 75, and 85 wt. % of YBCO are illustrated in Fig. 5. It can be seen that the graphs are nearly the same.

Fig. 5. M_m vs H_a dependences of composite $YBa_2Cu_3O_{7+\delta}$ superconductor-polyethylene samples with 30, 50, 80, and 85 wt. % content of YBCO.

4. Conclusions

We developed a technique of the preparation of a composite material consisting of ceramic high temperature superconductor YBCO and the low-density polyethylene. All the prepared composite samples show superconducting properties with critical temperature about 91 K. Based on results of detailed magnetization and inductive measurements, it can be inferred that the presented technique of the preparation has no degradation effect on basic superconducting properties. Moreover, the superconducting intergrain junctions effects have not been observed for composite samples, even for the sample with the highest YBCO content. Magnetization measurements suggest possible applications of the composites materials, such as electro-magnetic shielding, or protection of superconductors against moisture or chemical agents. In addition, to conserve superconducting properties of $YBa_2Cu_3O_{7-\delta}$ ceramics, no antioxidant organic compound has to be added in the initial mixture of superconductor and PE.

Acknowledgements

This work was supported by the Slovak Grant Agency for Science, projects VEGA No. 2/7083/27 and No. 2/6114/26.

References

[1] Chen M, Donzel L, Lakner M, Paul W. High temperature superconductors for power applications. *J. European Ceram. Soc.* 24: 1815-1822, 2004.

[2] HAUGHT d, Daley J, Bakke P, Marchionimi B. Overview of the U.S. Department of Energy (DOE) High-Temperature Superconductivity Program for Large-Scale Applications. *Int. J. Appl. Ceram. Technol*., 4: 197-202, 2007.

[3] Lee K, Kim B, Ko J, Jeong S, Lee S. Advanced design and experiment of a small-sized flywheel energy storage system using a high-temperature superconductor bearing. *Supercond. Sci. Technol*., 20: 634-639, 2007.

[4] Bhadrakumari S, Predeep P. High-T_c superconductor/linear low density polyethylene (LLDPE) composite materials for diamagnetic applications. *Supercond. Sci. Technol.,* 19: 808-812, 2006.

[5] Bhadrakumari S, Predeep P. PTCR Characteristics in $YBa_2Cu_3O_{7-x}$-Linear Low Density Poyethylene (LLDPE) Composite Materials. *J. Supercond. Nov. Magn*., 21: 313-316, 2008.

[6] Acosta JL, Moure C, Duran P. High T_c superconductor polymer composites based on $YBa_2Cu_3O_{7-x}$. *J. Mater. Sci.,* 30: 2888-2894, 1995.

[7] Namadneh I, Abdullah I, Halin SA, Abd-Shukor R. Dynamic magnetic propreties of superconductor/polymer NdBCO/PVC composites.*J.Mater.Sci.Lett.,* 21:1615-1617,2002

[8] Davtyan SP, Tonoyan AO, Hayrapetyan SM, Manukyan LS. A note on the peculiarities of producing of high-temperature super-conducting polymer-caramic composites. *J. Mater. Proces. Technol*., 160:306-312, 2005.

[9] Bhadrakumari S, Predeep P, Sheela A. Magnetic levitation force measurement on HTSC ceramic/thermoplastic polymer composites. *Comp. Sci. Technolog.,* 68: 3230-3233, 2008.

[10] Koňakovský A, Cigáň A, Maňka J, Bartkoviak J. Automatic Apparatus for Simultaneous Measurement of Resistance and Inductance Temperature Transition Dependences of Superconductors. In Measurement Science Review: journal published by the Institute Measurement Science, Slovak Academy of Sciences. ISSN 1335-8871, 2008, vol. 8, sec. 3, No. 6, p. 158-160. (http://www.measurement.sk/)

[11] Zrubec V, Cigáň A, Maňka J. Simplified fast method for magnetic characteristics measurement of the HT_c superconducting materials. *Physica C*, 223: 90, 1994.

Development of Precise Measurement Methods of Electrical Quantities in Macedonia and Croatia

[1]R. Malarić, [2]M. Čundeva-Blajer, [2]Lj. Arsov, [1]H. Hegedušić, [1]P. Mostarac

[1] Faculty of Electrical Engineering and Computing, Zagreb, Croatia
[2] Ss. Cyril and Methodius University, Faculty of Electrical Engineering and Information Technologies, Skopje, R. Macedonia
Email: roman.malaric@fer.hr, mcundeva@feit.ukim.edu.mk

Abstract. *In this paper two high precision methods developed at the National laboratories of Macedonia and Croatia for the measurements of electrical quantities will be presented. The metrological infrastructure is primarily located in the Laboratories for Electrical Measurements at the Faculty of Electrical Engineering and Information Technologies-Skopje, R. Macedonia and the Faculty of Electrical Engineering and Computing-Zagreb, R. Croatia. Both laboratories tend to become complementary regional centres of metrology in the field of electrical quantities and have developed original methods for precise electrical measurements, which will be described in this contribution. An emphasis in this article will be given on precise measurements of high voltages, currents and resistances.*

Keywords: electrical measurements, CAD, instrument transformer, resistance measurements

1. Introduction

The development of high precision methods for measurements of electrical quantities suitable for national calibration laboratories is time consuming and costly. Therefore a mutual bilateral scientific project between the Faculty of Electrical Engineering and Information Technologies-Skopje, R. Macedonia (FEIT) and the Faculty of Electrical Engineering and Computing-Zagreb, R. Croatia (FER) has been launched, supported by the governments of the both countries, for metrological research and development of methods and procedures for electrical measurements. The Laboratories for Electrical Measurements at FEIT and FER are part of the metrological infrastructure for electrical quantities of the both countries, which has been described in the first part of this contribution. Two very different high precision methods will be presented in this paper: measurements of high currents and voltages at FEIT and measurements of resistances at FER.

2. Development of methods for precise electrical measurements at AMLEQ, Macedonia

The Authorised Metrological Laboratory for Electromagnetic Quantities (AMLEQ) at FEIT-Skopje is a part of the metrological infrastructure in R. Macedonia, as described in the first part of this contribution. At the AMLEQ a prototype of a combined current-voltage instrument transformer (CCVIT) for measurement of high voltages and currents has been developed. The CCVIT must comply with the rigorous metrological requirements of the IEC 60044-3 standard, [1]. The CCVIT is very complex electromagnetic system compring: current measurement core (CMC) and voltage measurement core (VMC) with four windings and two magnetic cores. The full geometrical and electromagnetic description of the CCVIT, as well as the thorough FEM-3D analysis of its electromagnetic field distribution has previously been given in [2]. The electromagnetic analysis has been done for the calculation of the main contribution factors to the CCVIT uncertainty budget, the leakage reactances of the four CCVIT windings, [2]. The main metrological parameters of the CCVIT are: p_u VMC voltage error, p_i CMC current error, δ_u VMC phase displacement error, δ_i CMC phase displacement error. The optimisation objective function f_{opt} minimises the CCVIT measurement errors (p_u VMC voltage error at rated regime, $p_{u0.25}$ VMC voltage error at 0,25 of rated load, p_i CMC current error at rated regime). The GA program maximises f_{opt}, therefore f_{opt} is defined as in (1).

$$f_{opt} = \frac{1}{1+|p_i|} + \frac{1}{1+|p_u|} + \frac{1}{1+|p_{u0.25}|} + \frac{1}{1+|p_u + p_{u0.25}|} \tag{1}$$

The phase displacement errors of the both measurement cores are restrictive functions in the mathematical model. The FEM-3D results have been input data into the genetic algorithm. In the mathematical model of the CCVIT all quantities which affect the objective function are made to be dependent on the 11 optimisation variables. The results derived by the GA optimisation process are the so called *optimal design project*.

A prototype of the CCVIT has been realised in the Instrument transformer production company, EMO A. D.-Ohrid, R. Macedonia. High voltage testing has been done over the prototype according to the IEC 60044-3 standard, [1]. The comparison of the metrological characteristics at different working regimes of the both measurement cores of the initial and optimal design project and the prototype are given in Tables 1 and 2.

Table 1. Comparison of the design and the experimental VMC parameters (CMC at rated regime)

$\frac{U_u}{U_{ru}}$	Initial (Analytical + FEM-3D)		Optimal (FEM-3D+GA)		Prototype (experiment)	
	p_u	δ_u	p_u	δ_u	p_u	δ_u
	[%]	[min]	[%]	[min]	[%]	[min]
0,2	-7,64	-181,68	-5,56	-188,61	-1,338	-23,72
0,4	-5,18	-148,82	-3,38	-150,91	-0,855	-16,97
0,6	-3,75	-111,58	-2,07	-112,91	-0,631	-13,52
0,8	-2,96	-89,43	-1,34	-90,46	-0,53	-12,1
1,0	-2,47	-75,48	-0,86	-75,44	-0,465	-11,2
1,2	-2,10	-64,00	-0,53	-64,80	-0,429	-10,97

Table 2. Comparison of the design and the experimental CMC parameters (VMC at rated regime)

$\frac{I_i}{I_{ri}}$	Initial (Analytical + FEM-3D)		Optimal (FEM-3D+GA)		Prototype (experiment)	
	p_i	δ_i	p_i	δ_i	p_i	δ_i
	[%]	[min]	[%]	[min]	[%]	[min]
0,2	-1,08	0,25	-0,86	22,61	-0,66	8,3
0,4	-0,81	6,38	-0,12	10,02	-0,46	5,3
0,6	-0,76	3,55	-0,07	3,43	-0,34	3,1
0,8	-0,72	1,23	-0,03	3,87	-0,27	2
1,0	-0,68	0,32	-0,01	3,47	-0,22	1,1
1,2	-0,65	1,27	0,00	2,92	-0,18	0,6

A thorough transient analysis of the CCVIT has also been done by using a universal transformer model in the Matlab/Simulink program, [3]. The results from the FEM-3D analysis and the construction parameters derived from the GA optimal design are an input data for the CCVIT transient analysis. The main metrological characteristics of the CCVIT are derived for the most common transient regime, the *in-rush regime,* of the initial and the optimal design project at rated load of the both cores and industrial frequency of 50 Hz. The results confirm the metrological improvement of the optimal CCVIT during the transient regimes in comparison to the initial design project, as shown in Figures 1-4. Some of the results for the peak values of the voltage and current error (for the most rigorous moment at β=0, of the input voltage period, i.e. at t=5 ms) of the initial and the optimal design project of the CCVIT transient analysis are comparatively given in Table 3. It can be concluded that the CCVIT metrological improvement has been achieved from transient regimes' aspect, by application of the above universal analysis and design methodology.

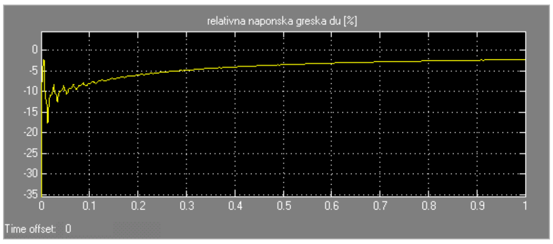


Fig. 1. Metrological transients of the initial CCVIT design project voltage error

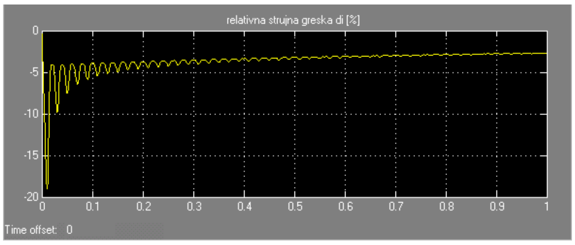


Fig. 2. Metrological transients of the initial CCVIT design project current error

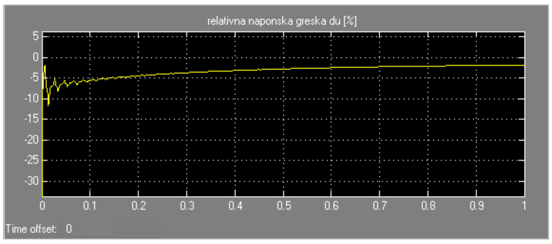


Fig. 3. Metrological transients of the optimal CCVIT design project voltage error

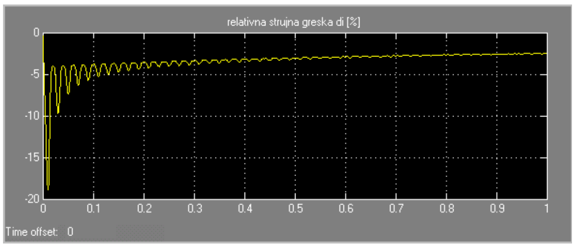


Fig. 4. Metrological transients of the optimal CCVIT design project current error

Table 3. Comparison of the CCVIT transient metrological parameters at rated regimes of the both cores

	Initial design project	Optimal design project
β [rad]	0	0
p_u [%]	-17,5	-11,5
p_i [%]	-19	-17,5

3. Development of high accuracy resistance standard measurement at PEL, Croatia

The resistance standards are compared with digital voltmeter method already described in several papers [4, 5]. Whole measurement process is automatized using personal computer. Recently the software has been rewritten from Pascal to Labview (Fig. 5.).

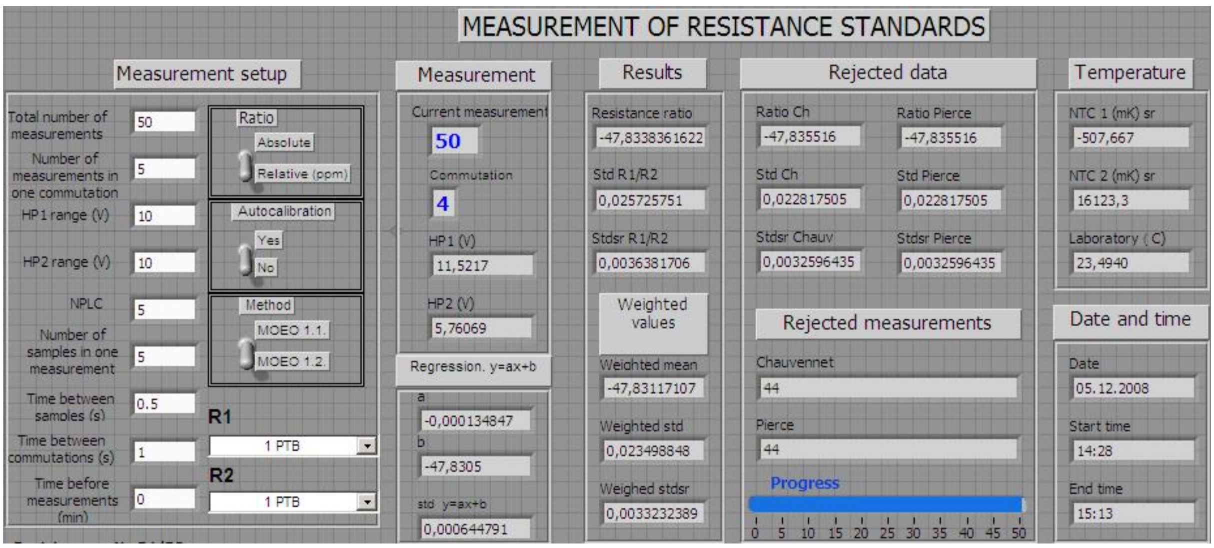


Fig. 5. Labview front panel for the measurement of resistance standards

This is offering new possibilities, especially in graphical presentation of the data. Also the data rejection methods developed by Pierce and Chauvennet are used to reject some data that are off limits. The front panel contains all the necessary data for the measurement report, including measurement setup, results, time and date, standards temperature, etc.... The results are presented graphically even during the measurement which greatly enhances the measurement process. The results obtained with this method are very good. The standard deviation of such comparison is usually less than 0,01 ppm (Fig. 6.). Each ratio consists of four measurements, as DVMs are interconnected to each resistance and also the current is reversed. This gives the possibility to calculate the standard deviation of each measurement (red line in Fig. 6). Any deviation from usual measurements can be thus easily monitored. In the end the weighted mean and standard deviation can be calculated which is usually gives better results than normal average ratio.After ratio is determined, then the standard resistance value of unknown standard is easily calculated using the known value of resistance standard used for comparison.

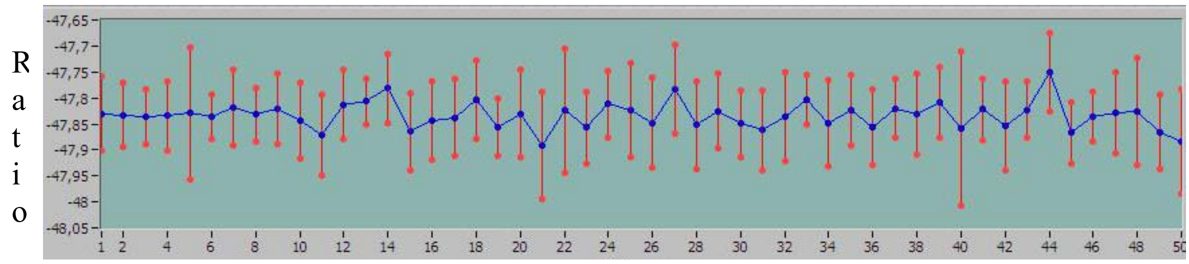


Fig. 6. Comparison of two 10 kΩ resistance standards.

4. Conclusion

The first part of the paper has presented the metrological infrastructure for electrical quantities in the R. of Macedonia and the R. of Croatia. The original and metrologically verified methods for precise electrical measurements at two laboratories have been presented in the second part of this contribution. The precise methods for measurements of high voltages, current and resistance are original. The methodology for estimation of the uncertainty budget is advanced, computer aided, universal and can be applied for verification of other procedures for precise electrical measurements. The optimisation techniques can be applied reduction of the contribution factors to the unceratinty budgets of the measurements' methods.

References

[1] IEC 60044-3: "Instrument transformers-Part 3: Combined transformers", Geneva, 1980

[2] Cundeva M, Arsov L, Experimental Verification of the CAD Analysis and Design Results of Combined Current-Voltage Instrument Transformer, In the proceedings of the 13th IMEKO TC4 Int. Symp. on Measurements for Research and Industry Applications , Athens, Greece, 2004, 487-492

[3] Ong C. M.: Dynamic Simulation of Electric Machinery Using MATLAB/SIMULINK, Prentice Hall PTR, Upper Saddle River, New Jersey, 1998

[4] Malarić R, Leniček I, Šala A, Scaling of resistance standards in the primary electromagnetic laboratory of Croatia, In the proceedings of the Measurement. 34 (2003), Volume 4, 347-353

[5] Bego V, Butorac J, Malarić R, Maintenance of Croatian resistance standards, In the proceedings of the XIV IMEKO World Congress, Volume IV A, 1997, 161-166

Electronic Thermometer

J. Foit, J. Novák

Department of Microelectronics, Faculty of Electrical Engineering,
Czech Technical University, Prague, Czech Republic
Email: foit@fel.cvut.cz

Abstract. *In evaluations of interactions between the electromagnetic fields and other objects, a significant simplification of approach is possible and reasonable: the degree of interaction can be followed by both global and local temperature variations. Therefore, the theoretical ideal thermometer has infinitely small dimensions and a negligible heat capacity. A small point contact realized by a thermocouple junction as the thermal sensor is the device of choice for the judgment and tracing of such interactions This paper deals with the principle, design, and obtained properties of an electronic thermometer fulfilling these general requirements.*

Keywords: thermometers, thermocouples, microvolt DC signal processing.

1. Introduction

Since common objects, and biological objects in particular, under normal living conditions hardly ever keep their own long-term internal temperature stable within a range narrower than about ± 0.2 K, the absolute accuracy of the planned instrument need not be better than about ± 0.1 K. On the other hand, in order to detect minor effects, its temperature variation resolving power should be at least an order of magnitude better, i.e. about ± 0.01 K. Another important factor, in terms of the possibility of accurate spatial temperature mapping, the temperature sensor size (linear dimensions, volume, and thermal capacitance) should be as small as possible. In other words, the sensor should resemble a "point-contact" sensor. A third important requirement says that the sensor should not, as far as technically possible, interfere with the biological object standard conditions, in other words it should be inert in terms of the internal chemistry of the measured object. Fourth, not least important factor, is the stability of the temperature sensing device, in terms of both short-term and long-term stability, permitting durable and reliable calibration of the system.

2. Basic design deliberations

Technically, such a system can be designed in several different ways. After a careful weighting of all possible aspects of the solution, we have decided to use classical sensors based on the bi-metallic thermocouple. A large number of material pairs usable in thermocouple applications is known, but in terms of long-term stability, the best are metal-to-metal thermocouples. Several time-proven combinations come in question, differing in output voltage yield, resistance to high temperatures, and chemical immunity. The most common are the combinations iron-constantan, copper-constantan and chromel-alumel, in decreasing voltage yield and increasing chemical and temperature immunity. For easy accessibility, we have selected the compromise choice of copper-constantan thermocouples. Such a thermocouple can be prepared relatively easily by simple welding and it can be passivated easily by sealing in a thin-wall pointed glass tube. At the same time, using moderately thin wires (diameter 0.3 mm), such a thermocouple can easily be prepared as small as about 1 cubic millimeter, permitting the required low mechanical and chemical interference with the measured object.

3. Electronic circuit solution approach

The characteristic property of most thermocouples is their rather low output voltage. Typically, a Cu-Const thermocouple generates a voltage of about 5 mV at 100 0C [1], [2]. From that follows that in order to obtain the required accuracy, the electronic system should be capable of accurate measurements of voltages better than ±5 μV and have a sensitivity (resolution) better than 0.5 μV. Though it is possible today to design and build semiconductor electronic signal processing devices with D.C. errors within these limits, it is not an easy task, especially in terms of the required operation in the common environmental temperature range, typically 0 to +40 0C. To avoid similar problems, we have decided to convert the thermocouple D.C. output to an A.C. signal which can be amplified easily and accurately while disregarding any D.C. voltage or current errors in the processing electronic devices. One of the best devices capable to perform such a conversion is a simple mechanical contact – of course under the condition that it is

a) rapid enough to operate accurately at a reasonable switching frequency,
b) mechanically adjustable to provide an optimization possibility of the switching process,
c) made of materials selected in such a way as to suppress its own thermoelectric voltages,
d) reasonably free of crosstalk between the processed signal and the contact drive signal.

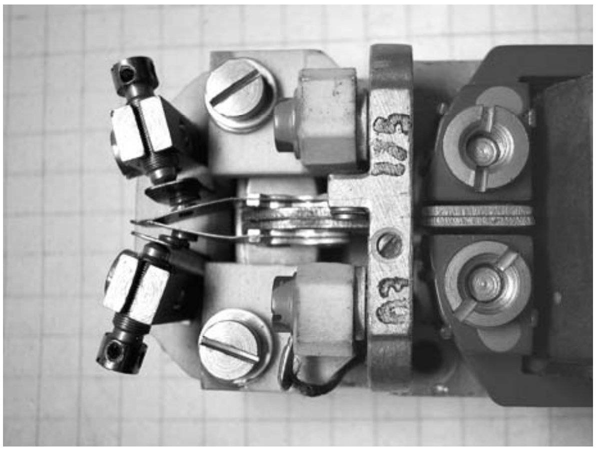


Fig. 1: chopper contacts

When looking for the optimum solution, we have selected the signal repeater relays ("chopper" in Fig. 2), used in classical teletype systems. These relays are quite rapid, they have micrometer screw-adjustable SPDT contacts, they are polarized (permitting the adjustment for a perfect symmetrical bistable operation), and their magnetic circuit is located relatively far away and symmetrically relative to the contact system, minimizing crosstalk. An example of such a teletype repeater relay contact arrangement is shown in Fig. 1. A slight modification of the connecting leads to the relay contacts, bringing them straight away axially instead of through the relay socket further minimizes the residual residual crosstalk. Moreover, the relay contacts are gold-plated, assuring reliable operation free of contact-bouncing over long time periods. The design sequence of construction materials (metals) of the contacts of this type of relays is symmetri-cal, assuring a fair suppression of parasitic thermoelectric effects. It results in no thermal protection or stabilization of the relay contacts necessary. In the D.C. to A.C. conversion process we can make good use of the relatively low electrical resistance of common thermocouples, typically below 10 Ω. Their low resistance permits stepping up the chopped thermocouple output voltage by a simple transformer.

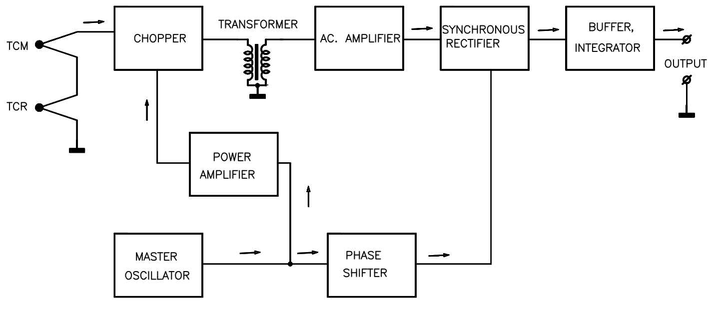


Fig. 2: System block diagram

The principle of operation of the whole system is shown by the block diagram in Fig. 2. The D.C. output from the thermocouple is converted to A.C. by the mechanical switch (chopper) , stepped up by a transformer, then amplified by an A.C. amplifier, and finally rectified by a synchronous rectifier to the analog D.C. output voltage. In order to suppress external interference, mostly caused by by the common A.C. power supply lines, the conversion (chopping) frequency is selected as an uneven multiple of the power line frequency, and the amplified signal is rectified by a synchronous rectifier. The temperature is measured by a differential thermocouple that generates its output as a difference voltage between the measurement (TCM) and reference (TCR) junctions. The reference junction is stabilized at the reference temperature 0 ^{0}C. So the operation is quite straightforward, the only block requiring some explanation is the phase shifter. The phase shifter forms short-duration (about 1 ms) switching pulses controlling the synchronous rctifier and permits to adjust the timing of the pulses to a position corresponding to the A.C. amplifier output signal section where the switching transients have already settled down. Depending on the particular design of the step-up transformer, the transients can be minimized very effectively by tuning the transformer secondary to the operating frequency by a parallel capacitance (not shown in the diagrams). The operating frequency was set to 77 Hz, a compromise value low enough for the mechanical switch speed, high enough for the design of the step-up transformer, and, as mentioned above, it is non-synchronous with the expected main source of external interference (power supply line frequency and its harmonics).

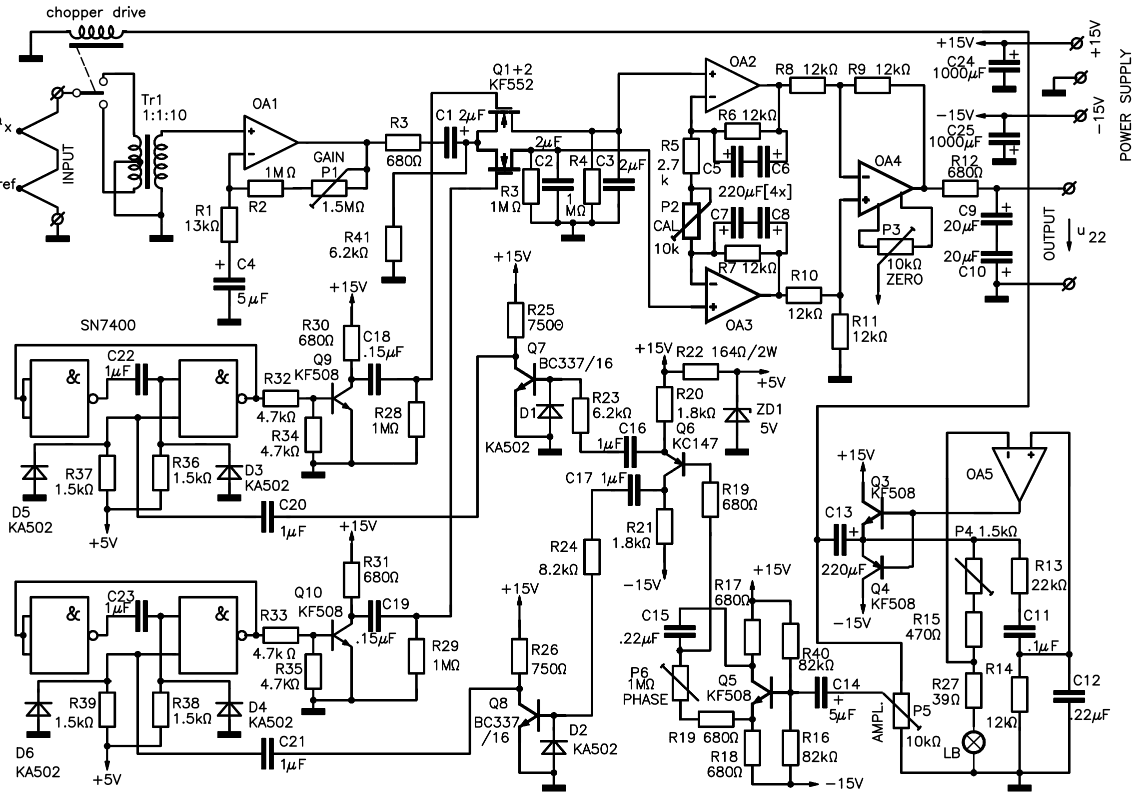


Fig. 3: Electrical connection of the thermometer

The complete electrical circuit connection of the thermometer is shown in Fig. 3. All operational amplifiers are of the LF356 type. Their type, however, is not at all critical; we have tested several versions, using even the cheap 741 or 725 types, and they all worked well. The same is true for the synchronous rectifier since it processes a relaively large low-frequency signal. Instead of a mechanical contact, almost any common small P-channel MOSFET type can be used; N-channel MOSFETs can also be used, only the switching pulses must be changed to positive.

4. Results

Fig. 4: Response to 1µV input step

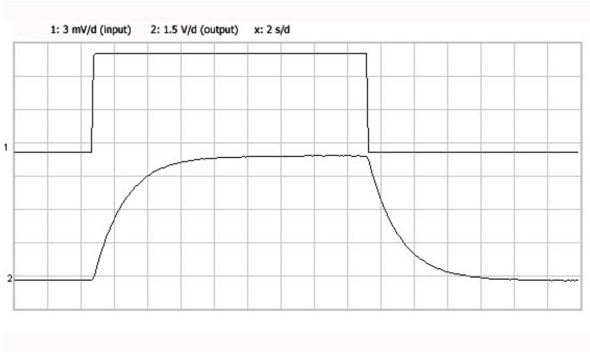


Fig. 5: Response to 10 mV input step

The following figures show the critical properties of the final design. Fig. 4 shows the response to an input voltage step of 1 µV, corresponding to a temperature variation of about 0.02 K at 300 K. Fig. 5 shows the response to an input voltage step of 5 mV, corresponding to +100 ^{0}C at the measurement thermocouple. Fig. 6 shows the long-term stability (zero input voltage at average room conditions, i.e. random temperature variations by approximately ±5 K during the whole 8-hour measurement interval). Of course no thermocouples are linear in a large temperature range. Fig. 7 shows the calibration curve of our *Cu-Const* themocouples. A numerical calibration table must be used for accurate absolute temperature measurement; an accurate A-D linearizing converter for the oupt voltage is currently being designed for the whole -200 ^{0}C to +200 ^{0}C temperature range.

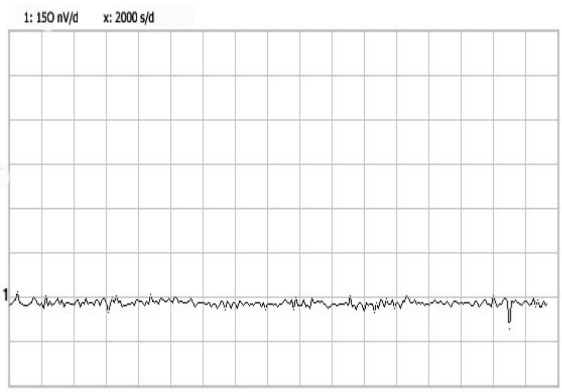


Fig. 6: Stability during 8 hours of operation

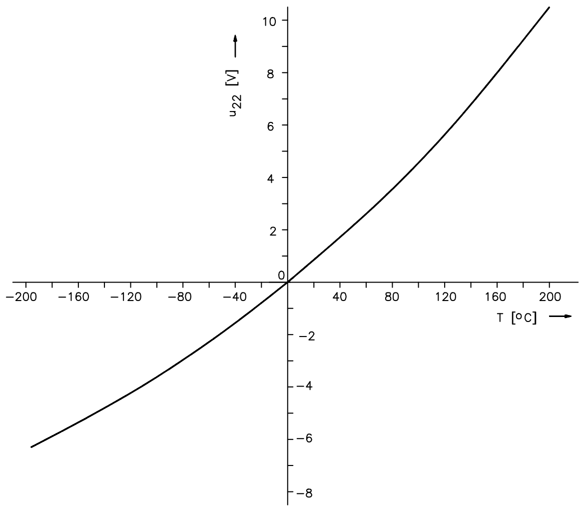


Fig. 7: Steady-state calibration curve

Acknowledgment

This research has been supported by the research program No. MSM6840770015"Research of Methods and Systems for Measurement of Physical Quantities and Measured Data Processing" of the CTU Prague, and partially by the Czech Science Foundation project No. 102/09/1601 "Micro- and nano-sensor structures and systems with embedded intelligence".

References

[1] *Reference handbook for radio engineers* Federal Radio Corp., New York, 1944.

[2] *Hoskins Co., USA*: Chromel-Alumel/Ni/Cr-Ni thermocouple calibration table.

Measurement of Electrical Parameters of Breakdown in Transformer Oil

J. Kúdelčik, M. Gutten

University of Žilina, Faculty of Electrical Engineering, Žilina, Slovakia
Email: kudelcik@fyzika.uniza.sk

Abstract. *The initial state of breakdown development is explained on the basis of bubble theory. Application of HV-DC voltage to electrodes immersed in oil results to creation of small channels, in which streamers can develop. In the next phase a plasma channel between the electrodes can be formed. The electrical resistance of plasma channel changes from a few ohms to a few hundred milliohms due to Joule heating caused by high arc current which flows through the plasma. The dynamics of the arc current depends on the parameters of outer circuit and is represented by RLC circuit.*

Keywords: plasma channel, arc resistance, arc current, RLC circuit

1. Introduction

The electrical breakdown in transformer oil and characteristic properties of this process are very important for many applications. Insulating liquids such as transformer oils are critical components for high voltage and pulsed power system. It was reported in numerous publications that dielectric breakdown is based on complex interactions of hydrodynamic and electronic phenomenons [1,2].

In present it is well known the initial stage of breakdown in transformer oil can be described using bubble mechanism. This theory assumes that a bubble of gas is formed by vaporization of liquid by local heating in the strong field region at a surface of the electrode. So the formed bubble will grow and a breakdown will take place inside the bubble. The breakdown processes are also dependent on mechanisms, which play role on interface of the liquid and the surface of electrodes. During breakdown a plasma channel with high initial resistance value is formed. This stage of breakdown - the creation of the plasma channel is similar for various types of liquid or gaseous at enough high applied voltage, although times and processes leading to this stage of breakdown are different.

The aim of the research reported in this article is to describe the time development of breakdown and its current in transformer oil using electrical detection techniques.

2. Experimental setup

Fig. 1 shows the schematic diagram of the experimental setup, which includes HVdc power supply TESLA BS 221 (max voltage 10 kV and current 3 mA), electrode system, electric and optical diagnostics. Sphere-to-sphere Cu electrodes with radius 1 cm were used as the electrode system. The distance of electrode was measured by metric gauge blocks with accuracy of 0.01 mm. New and unfiltered transformer oil - ITO 100 was filled into discharge chamber (0.2 dm^3) and electrodes were cleaned after series of 5 breakdowns. Time intervals between breakdowns were 15 minutes. The capacitor bank contained up to 4 HV capacitors connected in parallel, with nominally capacitance 0.05 μF. This allowed capacitances of between 0.05 and 0.2 μF to be used. The applied voltage and current were measured using a high voltage probe (E253/01, 10 MHz) and a Rogowski coil (Pearson Current monitor 110A, 10 kA, 20 MHz, 50 ns). Development of current and voltage were measured using 150 MHz external oscilloscope ETC M621.

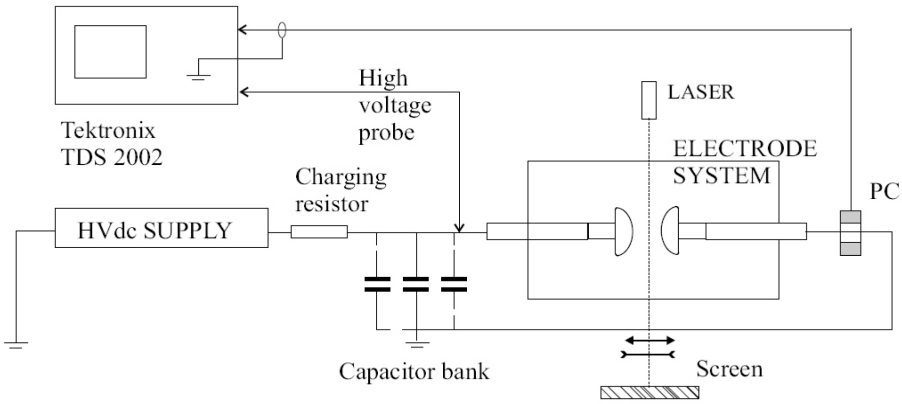


Fig. 1. Experimental setup

3. Results

Various transport phenomena were observed at electric field below 10^6 V/m. At around 25 % value of breakdown voltage (3×10^6 V/m) a small channel with diameter of some micrometers was detected between electrodes. Number of channels rose with increasing voltage. Their shapes were not stable and they were changing and moving. Shapes of these narrow channels illuminated by the laser are displayed in Fig. 2. Number and distribution of channels were dependent on electrode distance and applied voltage. Scattering of laser light on the interface of channels and the oil was caused by lower density of channels than that of the oil. The channels were concentrated along the electrode axis at voltage over the breakdown voltage.

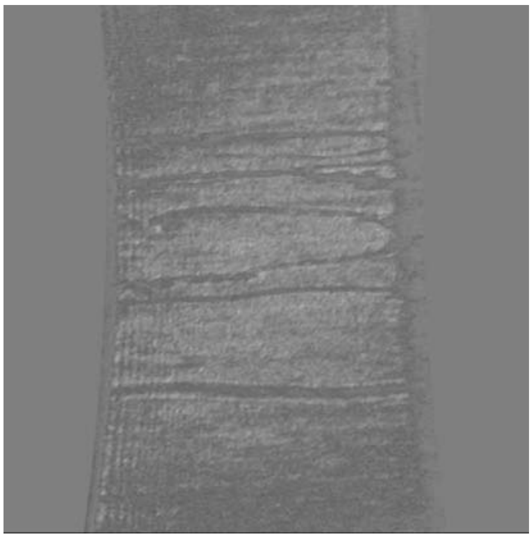


Fig. 2. The picture of discharge gap at the applied voltage 1.2 kV and gap distance 0.4 mm.

Development of the arc current at various capacitances is presented in the Fig. 3. For this case and type of electrode configuration there is almost homogeneous electric field with the electric intensity 74 kV/cm. The arc current is characterized by under-damped oscillation and its angular frequency depends on the value of capacitance, as it can be seen in Fig. 3. At equal capacitance only amplitude and duration of arc current changed with applied voltage. The measurements were also made at various electrode distances (0.1, 0.2 and 0.3 mm) and similar developments of arc currents as were observed in the Fig. 3. Simple measurements in transformer oil were also made by Marton [5], in water by Timoshkin et al. [1] and in air by Kijonka et al. [3].

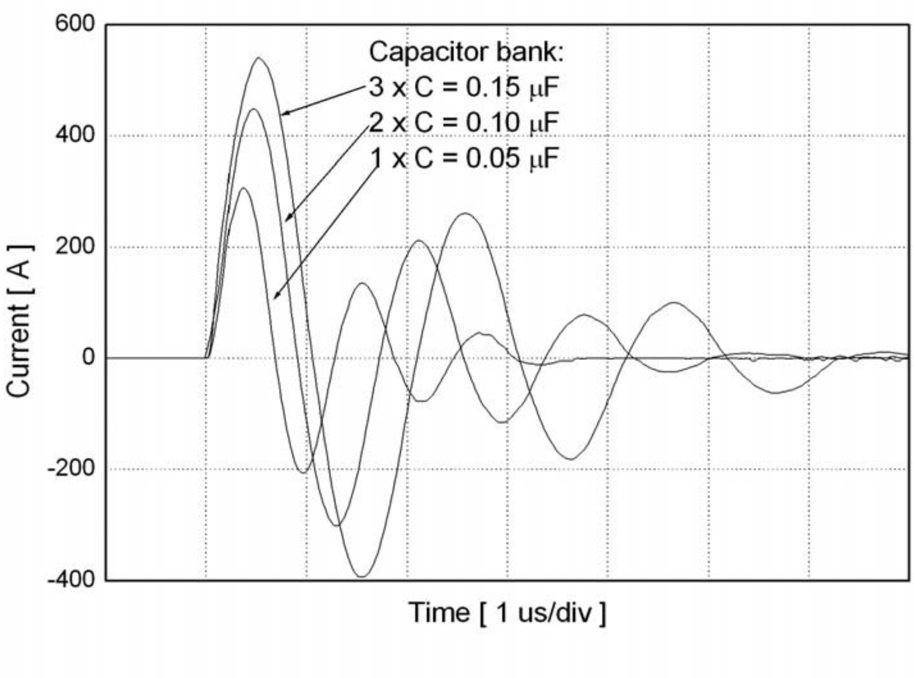


Fig. 3. Development of arc current and voltage across gaps at voltage 3000 V in ITO 100 and gap distance of 0,3 mm

During breakdown the experimental setup can be described by a electrical circuit, in which the arc current flows. Time dependence of the arc current can be fitted by function Sine Damp:

$$I(t) = I_0 e^{-\alpha t} \sin(\omega t), \tag{1}$$

where

I_0 the amplitude of current,
α the damping ratio,
ω the angular frequency of under-damped oscillation.

These parameters are determined by interpolation of measured pulses of arc current. Similar development of current can be observed in RLC circuit. On the basis of this similarity, corresponding values R, L and C* of the electrical circuit were calculated using the previous parameters and the breakdown voltage U_B as:

$$L = \frac{U_B}{I_0 \omega}, \quad R = 2\alpha L, \quad C^* = \frac{1}{L(\omega^2 + \alpha^2)} \tag{2, 3, 4}$$

where

R the total resistance of the electrical circuit.

Table 1. Electrical parameters of arc discharge

C [μF]	I_0 [A]	α [μs^{-1}]	ω [rad μs^{-1}]	R [mΩ]
0.05	303	0.719	4.976	1162
0.10	450	0.484	3.955	753
0.15	542	0.378	2.998	561
0.20	630	0.352	2.548	165

This circuit can be divided into two parts: external part with $R_{circuit}$ (HVdc power supply, capacitor bank, connecting cords and experimental equipment) and the arc resistance of the plasma channel R_{pl} (t): $R = R_{pl} + R_{circuit}$. Capacitance C^* calculated using previous formula was equal to the capacitance C of capacitor bank with accuracy better than 2%. Fitted parameters I_0, α, ω and calculated R are listed in the Table 1 for different capacitances C. The same identifications as in work [1] were used. Resistance R of the electrical circuit reduces with increasing input energy.

4. Discussion and Conlusion

From electrical point of view the whole breakdown can be represented by RLC circuit. Parameters of this circuit were calculated from experimental results using the theory of RLC circuit. During this stage, energy stored in capacitor bank was discharged to the system and under-damped arc current oscillations (Fig. 3, Eq. 1) were observed. The model of RLC circuit was also used in [1]. After transient phase, the value of arc resistance $R_{pl\text{-}const}$ was order of severel mΩ, so it had minimal effect to development of the arc current. Development of the arc current depends on capacitance of capacitor bank (see Fig. 1) and parameters of outer circuit [4]. Similar development of the arc current was observed in oil by Marton et al. [5]. The breakdown was accompanied with other various processes as: acoustical effect, light flash, shock and acoustic waves. Many bubbles of various magnitude were observed after breakdown.

The breakdown characteristics of oil ITO 100 were measured. At voltage over the breakdown voltage the arc channel was created. Its arc resistance changed from a few ohms to a few hundreds milliohms due to Joule heating caused by the arc current flowing through the arc channel. On the basis of experimental results it can be said that breakdown in transformer oil is represented by RLC circuit on a microsecond scale. The development of the arc current is modified by parameters of outer circuit and effect of nonlinear characteristics of breakdown was neglected.

References

[1] TimoshkinV, Fouracre R, Given M, MacGregor S. Hydrodynamic modeling of transient cavities in fluids generated by high voltage spark discgarges, *J. Phys. D.: Appl. Phys.*, 39: 4808-4817, 2006.

[2] Jones H, Kunhardt E. Development of pulsed dielectric breakdown in liquids, *Journal Physics D: Applied physics*, 28: 178-188, 1995.

[3] Kijonka J, Petr O, Šimandl L. Diagnostic of energy of arc, (in Czech), *32. Sešit KTE,* Vysoká Škola báňská - TU Ostrava, 2005.

[4] Kúdelčík J. Development of breakdown in transformer oil, *ADVANCE* 6: 35-39, 2007.

[5] Marton K. Prebrakdown stage of insulating fliuds in combination of electric and magnetic fileds, (in Slovak), Proc. DISEE, Inter. Conference, Bratislava, pp. 57-60, 2004.

Non-contact Measurement of Carrier Density and Mobility of Magnetized Materials

L. Laurinavičius

Department of Electrical Engineering, Vilnius Gediminas Technical University
Sauletekio av.11, LT-2040 Vilnius Lithuania
Email: laurinavicius@el.vtu.lt; lagnel4@gmail.com

Abstract. *The microwave density and mobility meter with constant magnetic field source and its measurement possibilities are described. It is shown that method magnetoabsorbsion of magnetoplasmic waves can be used for measurement thin films of magnetic materials. The accuracy of concentration and mobility measurements free charge carriers depends on magnetic induction relative errors. Also it was determined that solid state plasma specimen affects on measurement accuracy.*

Keywords: magnetic field, carrier density, mobility

1. Introduction

Main electrical properties of semiconductors are determined by values of concentration and mobility of free charge carriers in semiconductors. Microwave magnetoplasma methods are successfully used for investigations of semiconductors and semimetals due to their universality [1]. One of well known methods is Rayleigh interferometry when transmitted via investigated specimen microwave signal is compared with reference signal. The reference signal has a constant phase and amplitude. The amplitude and wave length of transmitted wave depend on electrical properties of a semiconductor. Therefore observing dimensional resonance of magnetoplasma waves in a semiconductor specimen, placed in pulsed magnets, which can generate a strong magnetic field it is possible to determine the concentration and mobility of free charge carriers of semiconductors. But such as it is measurement equipment is complicated and there are problems to measure thin specimen, because transmitted signal is weak. The accuracy of measurement can't be so high, because there are so many measured parameters: amplitudes of oscillations, magnetic induction and geometrical size of specimen.

In present article the author offer method to measure concentration and mobility of free charge carriers of semiconductors based on magnetoplasma phenomenon to absorb the HF signal. HF resonator lumped elements is proposed.

2. Physical background of investigation

If semiconductor specimen is placed in external magnetic field, radio microwaves can propagate in semiconductor along the direction of magnetic induction B. The propagation of magnetoplasma wave is determined by dispersion equation [2]

$$c^2 \frac{k^2}{\omega^2} = \varepsilon_{\pm}^{'} + i\varepsilon_{\pm}^{''} = \varepsilon_L \left(1 - \frac{\omega_p}{\omega\left[(\omega \pm \omega_c) + i\nu\right]} \right), \qquad (1)$$

where k is wave vector, ω is frequency, ε_L is lattice constant of semiconductor, $\omega_c = eB/m^{\otimes}$ is cyclotron frequency, $\omega_p = \left(\frac{e^2 n}{m^{\otimes}\varepsilon_0 \varepsilon_L} \right)^{1/2}$ is plasmas frequency, $\nu = 1/\tau$ is frequency of carrier's collisions.

The conditions when magnetoplasma helicon waves can propagate in magnetic material are as followings:

$$\omega \pm \omega_c >> \nu; \qquad \omega_c >> \omega; \qquad \omega_c \tau \equiv \mu B >> 1. \tag{2}$$

The length λ of helicon waves is determined by equation [3]

$$\lambda / 2 = \pi / k' = \frac{c}{\omega} \frac{1}{\sqrt{\varepsilon'}} = \frac{c}{\omega} \left(\varepsilon_0 \omega B \Big/ eN \right)^{1/2}. \tag{3}$$

It means that it is possible to determine the density of free charge carriers by observation of dimensional resonances of magnetoplasma microwaves in magnetic materials.

When mine (half – length) resonance is observed, the value of density N is determined by simple equation

$$N = \frac{A \cdot B}{d^2 \cdot f_R}, \tag{4}$$

where B is magnetic induction, f_R is exciting frequency ($d = \lambda_H/2$), d is thickness of the specimen, λ_H is wavelength of helicon and A = const.

When exciting frequencies $f << f_R$, we have non-resonance conditions in the specimen. In this case values of concentration N and mobility μ it is possible to determine by analysing tensor of specimen conductivity $|\sigma|$. Tensor of conductivity $|\sigma|$ looks as following [4]:

$$|\sigma| = \begin{Vmatrix} \sigma_{xx} & \sigma_{xy} & 0 \\ \sigma_{yx} & \sigma_{yy} & 0 \\ 0 & 0 & \sigma_{zz} \end{Vmatrix} = \begin{Vmatrix} \dfrac{\sigma_{zz}}{1+\omega_c^2\tau^2} & \dfrac{\omega_c \tau \sigma_{zz}}{1+\omega_c^2\tau^2} & 0 \\ \dfrac{\omega_c \tau \sigma_{zz}}{1+\omega_c^2\tau^2} & \dfrac{\sigma_{zz}}{1+\omega_c^2\tau^2} & 0 \\ 0 & 0 & \sigma_{zz} \end{Vmatrix}. \tag{5}$$

3. Realization of non – contact HF meter and experiment

Experimentally curves of magnetoplasma wave responses can be observed with high frequency meter whose simplified block diagram is shown in Fig. 1

A semiconductor specimen 1 is put in pulsed magnetic field generated by two axial solenoids of electromagnet 2. High frequency generator 3 is connected with coil 4 of LC resonator. The magnetoplasma wave is excited in local area of a semiconductor plate put in a constant magnetic field. Propagating across semiconductor plate magnetoplasma wave is indicated by the same coil 4. A receiving signal is indicated by memorized oscilloscope 5. Magnetic field control system 6 is connected to the oscilloscope 5. Described magnetoplasma meter has also a pulsed current source, magnetic field sensor and a scanning mechanism which are not shown.

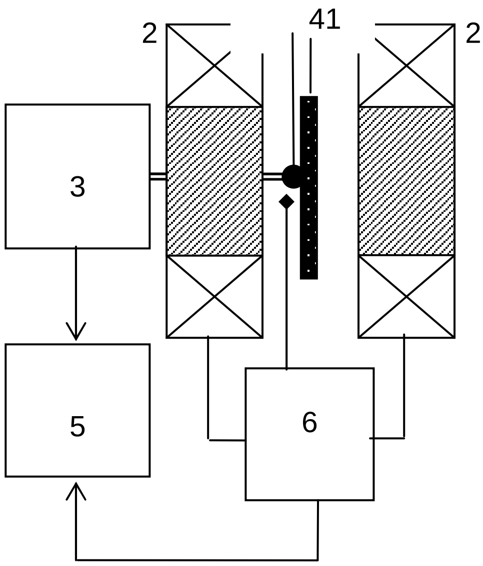


Fig. 1. Block diagram of magnetoplasma meter.

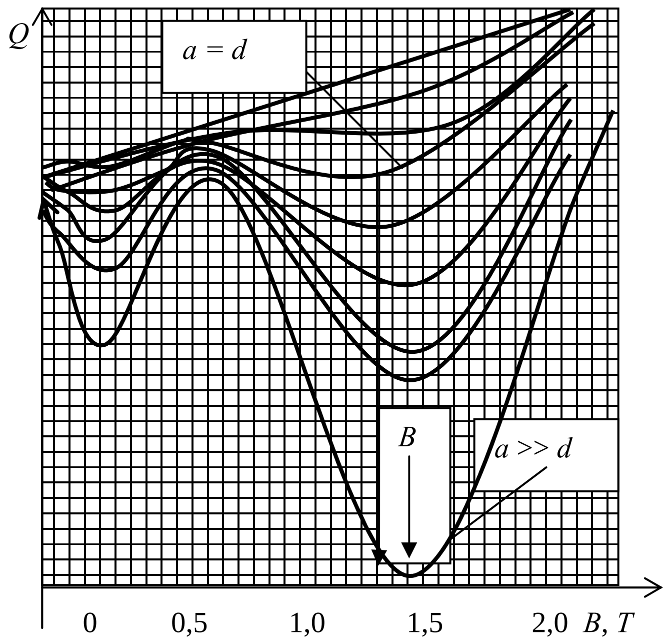


Fig. 2. Measurement of density. High frequency signal (quality of resonator Q) dependence from magnetic induction B.

Quality of induction coil (LC – resonator) placed on the surface of semiconductor in magnetic fields, when frequency of generator is $\omega_R \cong \omega$ is very low, because $Q = \frac{\sqrt{L/C}}{R}$ resistance losses R is high (mine dimensional resonance of helicon waves). The value of density N is determined (see Fig. 2) by equation (4).

When exciting frequencies $f << f_R$, we have non-resonance conditions in the specimen and dependence $Q = f(B)$ is measured.

Quality of LC – resonator Q is described by equation:

$$Q \equiv \sigma_{xx} = \frac{\sigma_{zz}}{1+\mu^2 B^2} = \frac{e \cdot N \cdot \mu}{1+\mu^2 B^2}, \tag{6}$$

where e, N, μ are mobility, density and charge of electrons respectively.
Mobility of electrons can be calculated very simply

$$\mu = \frac{1}{B_{1/2}}, \tag{7}$$

And quality $Q_{1/2} = (Q_B - Q_0)/2$.

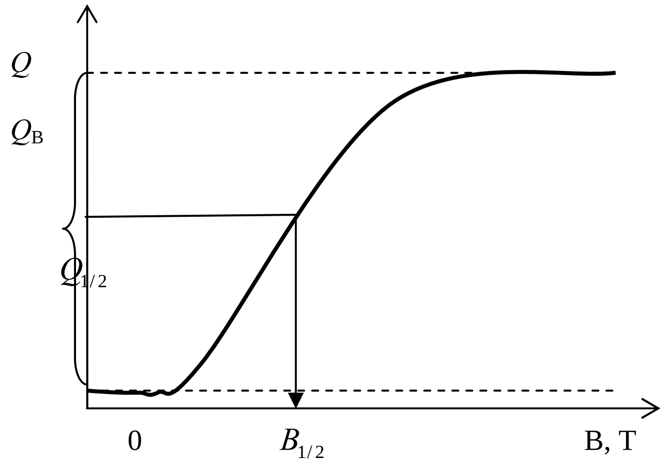


Fig. 3. Measurement of mobility. High frequency signal dependence from magnetic induction B.

4. Accuracy evaluation

The relative error of concentration can be determined by formula

$$\left(\frac{\Delta n}{n}\right)=\sqrt{\left(A_f\left(\frac{\Delta f}{f}\right)\right)^2+\left(A_d\left(\frac{\Delta d}{d}\right)\right)^2+\left(A_B\left(\frac{\Delta B}{B}\right)\right)^2}\ , \tag{8}$$

where $\frac{\Delta f}{f}$, $\frac{\Delta d}{d}$, $\frac{\Delta B}{B}$ are relative errors of frequency, sample thickness, magnetic induction respectively. A_f, A_d, A_B are influence coefficients proportional to partial derivatives of output signal. Determined values of these coefficients are A_f = -1, A_d = - 2, A_B = - 1.

A relative error of frequency determination of modern generators is less then ± 1 %. The thickness of the specimen can be measured with relative error less then ± 1 %. Magnetic field induction is controlled by memorized oscilloscopes. So magnetic field can be determined with relative error no more than $\pm$ 5 %.

Taking $\frac{\Delta f}{f}=\pm 1\%$, $\frac{\Delta d}{d}=\pm 1$ %, $\frac{\Delta B}{B}=\pm 5\%$ it was cleared that the relative error of concentration is determined in general and one does not exceed ± 5 %.

The relative error of mobility depends only of relative error of magnetic induction $\frac{\Delta B}{B}$. It was calculated, the relative error of mobility does not exceed ± 10 %.

Solid state plasma specimen size have influenced on measurement accuracy of density. If the geometrical size *a* of measurement aperture (coil) is less than *d*, the curve extreme is moving to the side of less induction. If the size of aperture is equal to *d* (*a* = *d*) relative error of magnetic induction are increasing.

5. Conclusion

Measurements of density and mobility of free charge carriers were executed in monocristalic specimens of n-Insb, n-GaAs, CdHgTe. We used magnetic field with induction B = 2 T. Good coincidence with technical specification of specimens was confirmed. Proposed method for measurement mobility of free carrier can be used in express measurement of thin films of magnetized materials.

References

[1] L. Laurinavichus, J.Novickij, V. Filipavichus. The accuracy of pulsed magnetoplasma interferometer measuring electrical properties of semiconductors. Proc. of 3-th International Conference on Measurements. Smolenice Castle. 2001. P. 57-60.

[2] Palik E. D., Furdina J. K. Infrared and microwave magnetoplasma effects in semiconductors. *Report on Progress in Physics*. V. 33. 1970. P. 1195 - 1322.

[3] Z. Jankauskas, L. Laurinavicius. Magnetic and Electric Excitation of Magnetoplasmic Waves. *Electronics and Electrical Engineering*. 2002. No 2(37). P. 32-34.

[4] L. Laurinavicius. Non – destructive HF Magnetoplasma Methods for Testing Semiconductor Materials. *Electronics and Electrical Engineering*. 1995. No 3. P. 39-46.

Microwave Characterization of Frequency and Temperature Dependences of Beef Bone Dielectric Properties Using Waveguide Measurement System

D. Faktorová

Department of Measurement and Applied Electrical Engineering,
Faculty of Electrical Engineering, University of Žilina,
Univerzitná 1, 010 26, Žilina, Slovak Republic
Email: faktor@fel.uniza.sk

Abstract. *The article focuses on the measurement and calculation of dielectric properties of biological tissue in the frequency band 4,5 GHz – 16 GHz. Special attention is paid to the frequency and thermal dependence dielectric constant and loss factor. Observed frequency changes of biological tissue dielectric properties obey causality, i.e. Kramers-Krönig relationships which relate changes of dielectric constant with conductivity changes. Our results for frequency dependence of complex permittivity prove the fundamental Cole-Cole model but also the necessity to investigate the biological tissues like heterogeneous material with not only one relaxation phenomena.*

Keywords: Biological Tissues, Dielectric Properties, Microwave Frequencies, Waveguide Method.

1. Introduction

Dielectric properties of biological tissues and cell suspensions have been of interest for over a century for many reasons. They determine the pathways of current flow through the body and, thus, are very important in the analysis of wide range biomedical applications. To develop the use and additional applications of microwave energy, it must be considered that the dielectric properties of the tissues which determine the absorption and propagation of electromagnetic energy through the tissues. So from the knowledge of dielectric constant, tissue properties can be characterized in the microwave frequency range. To analyze the response of a tissue to electric stimulation, we need data on the variation of specific conductivity and relative permittivity of the tissues and organs, with the frequency and temperature changing [1].

A microscopic description of the response is complicated by the variety of cell shapes and their distribution inside the tissue as well as different properties of the extra cellular media. A macroscopic approach is most often used to characterize field distributions in biological systems. Moreover, even on a macroscopic level, the electrical properties are complicated. They can depend on the type of tissue, on the tissue orientation relative to the applied field (directional anisotropy), the frequency of the applied field (the tissue is neither a perfect dielectric nor a perfect conductor), on the tissue temperature, or they can be time- and space-dependent (e.g., changes in tissue conductivity).

2. Dielectric properties of biological tissues

Large differences exist in dielectric properties of biological materials. These differences are determined, to a large extent, by the fluid content of material. For example, blood and brain conduct electric current relatively well. Lungs, skin, fat and bone are relatively poor conductors. Liver, spleen, and muscle are intermediate in their conductivities.

The dielectric properties of biological tissues are highly dispersive due to the cellular and molecular relaxation, generated by different parts of the tissues at different frequencies. In the microwave region the dominant relaxation is the dipolar relaxation of free water molecules. Therefore, the dielectric properties of the tissues in microwave region are highly correlated to the water content. At the frequencies in microwave region (~10^9 Hz) the rotations of the polar molecules in the water begin to lag behind the electric field oscillations.

At frequencies for which the loss angle δ differs significantly from 90° the water has the dual role. It functions both as a dielectric and as a conductor and dielectric properties of materials are quantified by their bulk permittivity ε which has a complex character

$$\varepsilon = \varepsilon' - \mathrm{j}\varepsilon'', \qquad (1)$$

where the ε' is relative permittivity (dielectric constant). It is determined by the magnitude of overall net polarization in material. ε' determines the amount of energy stored per unit volume in material for a given applied field. The imaginary part in (1) is loss factor given as $\varepsilon'' = \sigma / \omega\varepsilon_0$, where σ is the electrical conductivity, ε_0 the dielectric constant for free-space and ω angular frequency. The loss factor represents the energy loss in a material and is governed by the lag in polarization upon application of the field and the energy dissipation associated with charge polarization. In solids (in our study of biological tissue) and liquids consisting polar molecules (those retaining a permanent dipole moment, e.g. water), the resonance effect is replaced by relaxation. A phenomenological approach for the mathematical modelling of dispersion is the Debye theory. The theory suggests a first order of differential equation system, similar to charge of a linear *RC* circuit. The complex permittivity in the frequency domain reduces to the Debye equation

$$\varepsilon = \varepsilon_\infty + \frac{\varepsilon_\mathrm{s} - \varepsilon_\infty}{1 + \mathrm{j}\omega\tau}, \qquad (2)$$

where ε_s and ε_∞ are the static and optical dielectric constants and τ, i.e. the relaxation time, that is a time constant of this first order system. Biological materials like markedly heterogeneous material do not exhibit single time constant relaxation behaviour. In concentrated systems, as well as biological tissues the electrical interaction between the relaxing species will usually lead to a distribution of relaxation time, $p(\tau)$ and with the help of this distribution, the following relation is used

$$\varepsilon = \varepsilon_\infty + (\varepsilon_\mathrm{s} - \varepsilon_\infty)\int_0^\infty \frac{p(\tau)}{1 + \mathrm{j}\omega\tau}\mathrm{d}\tau - \mathrm{j}\frac{\sigma_\mathrm{i}}{\omega\varepsilon_0}, \qquad (3)$$

where σ_i is the static ionic conductivity of the medium placed in a constant field that influencing very low frequencies. To enable a more wide-band model of the properties the time constant can be divided in several regions to match different type of relaxation. Due to the complexity and composition of biological tissues [2] extended Cole-Cole model is commonly used as physically based compact representations of wideband frequency dependent dielectric properties. In this case the complex dielectric constant is

$$\varepsilon = \varepsilon_\infty + \sum_n \frac{\Delta\varepsilon_n}{1 + (\mathrm{j}\omega\tau_n)^{1-\alpha_n}} - \mathrm{j}\frac{\sigma_\mathrm{i}}{\omega\varepsilon_0}, \qquad (4)$$

where $\Delta\varepsilon_n = \varepsilon_\mathrm{s} - \varepsilon_n$, and ε_n is relative permittivity appertaining to one relaxation process, α_n represents the distribution parameter which is a measure of broadening of dispersion and ionic conductivity is for microwave frequencies in (3) and (4) ignored.

Along with frequency effects, the complex permittivity is also affected by temperature. This dependence is governed by the effect of temperature on the individual polarization mechanisms [3]. Most notably, because charge mobility is affected by temperature, the electrical conductivity of a material will increase with an increase of the temperature.

3. Experimental results

For our measurement we have chosen the waveguide Hippel's method which has proved successful for the measured material and it behaves as the most accurate methods for dielectric properties of biological materials measurement. The measurements were carried out in frequency of two microwave bands - 4.5 GHz to 16 GHz. The experimental set-up for both bands are drawn in Fig. 1.

As a microwave source there was alternatively used besides the reflex klystron as well as HP broadband generator. All measurements were carried out on beef bones and the samples were taken from one homogenous part of femoral bone (without marrow). The samples were formed in such way to fit close to the waveguide walls (they were impressed into the waveguide). Different temperatures were obtained by immersion the waveguide with the sample into warmed water and the waveguide was safeguarded against water penetration into the sample, while the short - circuit was maintained.

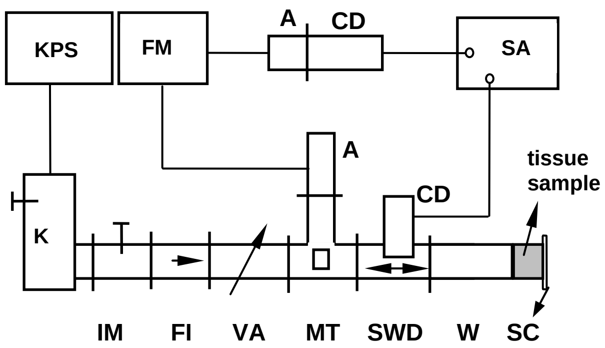


Fig. 1. Experimental set-up for complex dielectric constant measurement. K – klystron, KPS – klystron power supply, FM – frequency meter, SA – selective amplifier, IM – impedance match, FI – ferrite isolator, VA – variable attenuator, MT – magic T, SWD – slotted section, W – waveguide, SC – short circuit, A – adapter, CD – crystal detector.

The appropriate values were calculated and dependences ε' and tgδ on frequency and temperature are plotted in Fig. 2 up to Fig. 5.

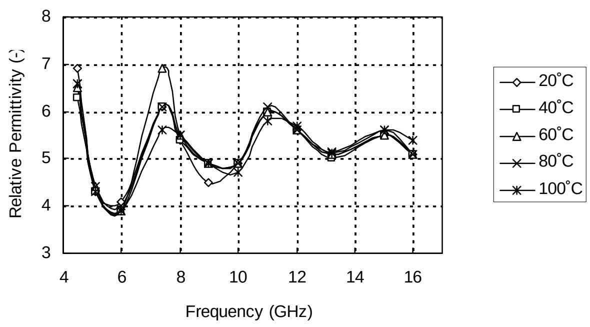


Fig. 2. Frequency dependence of relative permittivity for different temperatures.

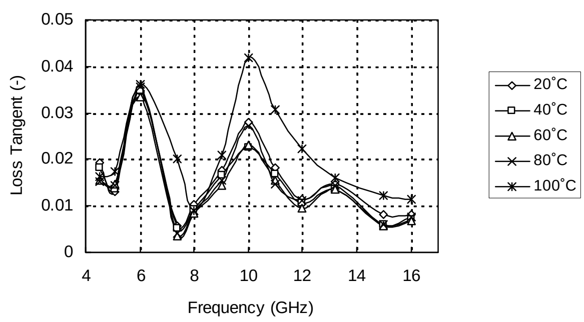


Fig. 3. Frequency dependence of loss tangent for different temperatures.

It can be seen from the Fig. 2 and Fig. 3 that the frequency dependences of ε' and tgδ show similar courses for all temperatures. Values of ε' and ε'' cannot vary independently with frequency, since their frequency variations are connected through the Kramers-Krönig relationship: a drop in ε' with increasing frequency is necessarily associated with a peak in

ε'', Fig. 2 and Fig. 3.While the values of ε' at different temperatures remain almost constant at all measured frequencies, the values of tgδ rise at higher temperatures except the frequency 6 GHz, Fig. 4 and Fig. 5.

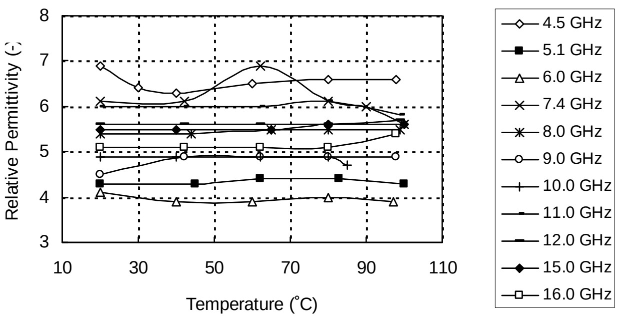


Fig. 4. Temperature dependence of relative permittivity for different frequencies

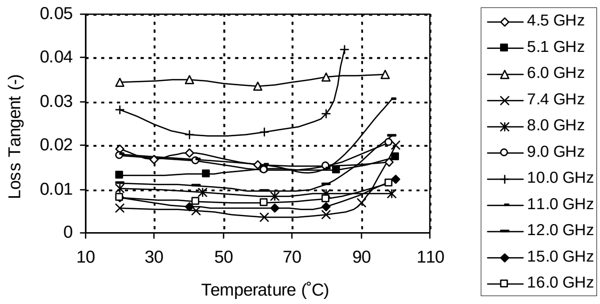


Fig. 5. Temperature dependence of loss tangent for different frequencies

The temperature dependence of relative permittivity ε' does not show expressive changes.

4. Conclusions

Although there are being published more information concerning dielectric properties different biological tissues accomplished data about bones occur only rarely. At the same time the temperature and frequency dependence of bone relative permittivity and loss tangent at the present time developing microwave therapy, will call for more detailed data about every part of human tissue. From this standpoint we have proceeded to the choice of the presented work and our results have brought some closer information for these applications. We submit data about temperature and frequency dependences more comprehensively occurring so far rather separately. Apart from that our results we point out that biological tissue is an inhomogeneous material and for the description of frequency dependence of dielectric constant a simple Cole-Cole model is not sufficient. It can be seen from our figures that several relaxation mechanisms go on simultaneously.

Acknowledgements

The author would like to thank MSc. Pavol Žirko the Director of High School for Agriculture and Fishing in Mošovce for technical support in our experiments.

The work has been done in the framework of Grant VEGA 1/0761/08 “Design of Microwave Methods for Materials Nondestructive Testing” of the Ministry of Education of the Slovak Republic.

References

[1] Foster K. R., Schwan H. P. Dielectric properties of tissues In: C. Polk, E. Postow, eds., Handbook of Biological effects of electromagnetic fields, John Wiley & Sons, New York, 1996.

[2] Gabriel S., Lau R. W., Gabriel C. The dielectric properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz. *Phys. Med. Biol.*, 41: 2251-2269, 1996.

[3] Pethig R. Dielectric and electronic properties of biological materials John Wiley & Sons, New York, 1979, 23-147.

Modeling of Radiation Heat Transfer in Indirect Heating Process

[1]J. Barglik, [2]I. Dolezel, [1]R. Przylucki, [3]B. Ulrych

[1]Silesian University of Technology, Katowice, Poland
[2]Czech Technical University, Prague, Czech Republic
[3]University of West Bohemia, Pilsen, Czech Republic
Email: roman.przylucki@polsl.pl

Abstract. *Correct evaluation of radiation in modeling of indirect or induction heating process is discussed. A newly offered approach introduces so called generalized convection and experimentally-numerical determination of its coefficient. The suggested algorithm is illustrated by typical example.*

Keywords: indirect heating, induction heating, thermal convection and radiation, boundary element method, generalized convection.

1. Introduction

The boundary conditions of temperature field of bodies heated indirectly or by induction in common gaseous media (air) are mostly given by convection and radiation. The intensity of the convective heat transfer between the heated body and ambient gaseous medium depends on the local physical circumstances and thickness of the laminar boundary layer of fluid surrounding the solid body. The heat flow obeys the Newton law

$$q_{\mathrm{c}} = -\lambda \frac{\partial T}{\partial n} = \alpha_{\mathrm{c}} \left(T - T_0 \right) \tag{1}$$

where q_{c} denotes the heat flow from the body to the ambient medium due to convection, λ the thermal conductivity of the body, n the outward normal, T the surface temperature of the body, T_0 the temperature of the distant gas medium and α_{c} the convective heat transfer coefficient. Coefficient α_{c} is usually found experimentally or from the theory of similarity.

For the solid body in an unbounded gas medium the heat transport by radiation may approximately be quantified by the Stefan-Boltzmann law in the form

$$q_{\mathrm{r}} = C_{\mathrm{r1}} C_{\mathrm{r2}} \sigma \left(T^4 - T_0^4 \right) = \bar{C}_{\mathrm{r12}} \sigma \left(T^4 - T_0{}^4 \right) \tag{2}$$

where q_{r} denotes the heat flow from the body to the ambient medium due to radiation, $\sigma = (5.67032 \pm 0.0071) \cdot 10^{-8}\,\mathrm{Wm^{-2}K^{-4}}$ is the Stefan-Boltzmann constant, T the absolute temperature of the surface of the radiating body and T_0 the absolute temperature of the distant gas medium. Symbol $C_{\mathrm{r1}} \in \left(0,1\right)$ stands for the emissivity of the surface of the radiating solid body and $C_{\mathrm{r2}} \in \left(0,1\right)$ the emissivity of the ambient medium. The product of both above quantities $\bar{C}_{\mathrm{r12}}$ may be interpreted as the "effective" emissivity of the interface between the body and ambient medium.

Now the principal question is whether it is necessary to consider both mechanisms in every case or neglect one of the mechanisms with respect to the other. In older references (i.e. [1]) we can often find the opinion that up to certain temperatures (about 200 °C) the radiation can be neglected in comparison with the natural convection. Nevertheless, the elementary experience (see, for example, a radiating body of central heating) provides another idea. This was demonstrated by an example described in [2]. In some cases, however, the problem of induction heating should be completed by analysis of multiple reflection phenomena describing heat transfer between the heated body and surrounding surfaces. Some examples are described in [3], [4].

2. Generalized convection

We can say that the total transfer of heat from the solid body into the unbounded gas medium is characterized by heat flow q_{tot} given by the sum of the convective (1) and radiation (2) flows (the third mechanism – conduction – being neglected due to very poor thermal conductivity of common gaseous media)

$$q_{\mathrm{tot}} = q_{\mathrm{c}} + q_{\mathrm{r}} = \alpha_{\mathrm{c}}\left(T - T_0\right) + \sigma \overline{C}_{\mathrm{r}12}\left(T^4 - T_0^4\right). \tag{3}$$

After a simple rearrangement we can formally write

$$q_{\mathrm{tot}} = \alpha_{\mathrm{tot}}\left(T - T_0\right),\ \ \alpha_{\mathrm{tot}} = \alpha_{\mathrm{c}} + \alpha_{\mathrm{r}},\ \ \alpha_{\mathrm{r}} = \sigma \overline{C}_{\mathrm{r}12}\left(T + T_0\right)\left(T^2 + T_0^2\right). \tag{4}$$

This relation can be found, for example, already in [5]. But only in association with a simple experimentally-numerical algorithm for determining of the generalized coefficient α_{tot} it becomes practically applicable.

3. Experimentally-numerical determination of coefficient α_{tot}

The thermal flux at an arbitrary point M on boundary Γ between the solid body and unbounded gas medium is given by relation

$$q_{\mathrm{tot}} = -\lambda \frac{\partial T}{\partial n} = \alpha_{\mathrm{tot}}\left(T - T_0\right). \tag{5}$$

Hence we can find the value of α_{tot} at point M provided that we know the values of T, T_0, λ and $\partial T / \partial n$. The experimental determination of the distribution of the surface temperature T, as well as temperature T_0 of the unbounded gas medium can sufficiently accurately be realized in any laboratory equipped with common apparatus. On the other hand, the direct measurement of $\partial T / \partial n$ on the surface of the same body is difficult by principle and always will be burdened by an error. More advantageous is to find this distribution by the Boundary Element Method (BEM) [6] that starts from the knowledge of the distribution of the surface temperature T. The particulars are described in [2].

4. Practical example of determination of α_{tot}

To simplify the description of the experiment, we will further work with temperatures in °C. Consider an unpolished steel shell in unmoving (steady) air, whose surface is heated to temperature $T_{\max}$ B 500 °C. The cylindrical shell **1** (see Fig. 1, left part) is assumed to be infinitely long, so that its temperature varies only in the radial direction, which can be expressed as $T = T(r)$. The aim of the experiment is to illustrate the experimentally-numerical algorithm for determination of coefficient $\alpha_{\mathrm{tot}}(T)$ described in the previous paragraph. Another partial aim is to compare the amount of heat transferred from the wall of the tube by convection and radiation.

The measuring device is depicted in Fig. 1, right part. It consists of the mentioned shell **2** on whose surface we investigate the values of α_{tot}, α_{c} and α_{r} as functions of the surface temperatures. These temperatures are measured by six thermocouples **1** placed on the surface of the measuring cylinder NiCr_NiAl and registered on the display of the measuring device **4**. The temperature of the ambient medium T_0 is measured by a mercury dilatation thermometer **3** that is not influenced by radiation from the measuring cylinder (combination glass + mercury is a very good reflecting surface for infrared radiation). The internal power heating the measuring cylinder **2** (and affecting also its surface temperatures) is controlled by temperature control unit **5** and evaluated by wattmeter **6**. The measuring cylinder used for the described experiment is depicted in details in Fig. 1 left. It consists of an electrical resistive heating element **2** of maximum power 500 W, inserted by means of the distance rings **3** (asbestos tape) into a steel (carbon steel, its thermal conductivity $\lambda(T)$ is shown in Fig. 2 left) cylindrical shell **1**. Its surface is equipped by six thermocouples NiCr_NiAl **8**. For suppressing the variations of temperature of the measuring cylindrical shell **1** in the axial direction both ends of the heating element **2** are equipped with thermal insulation **6** (fiberglass). The cylindrical shell **1** is also closed by asbestos insulating flanges **7**.

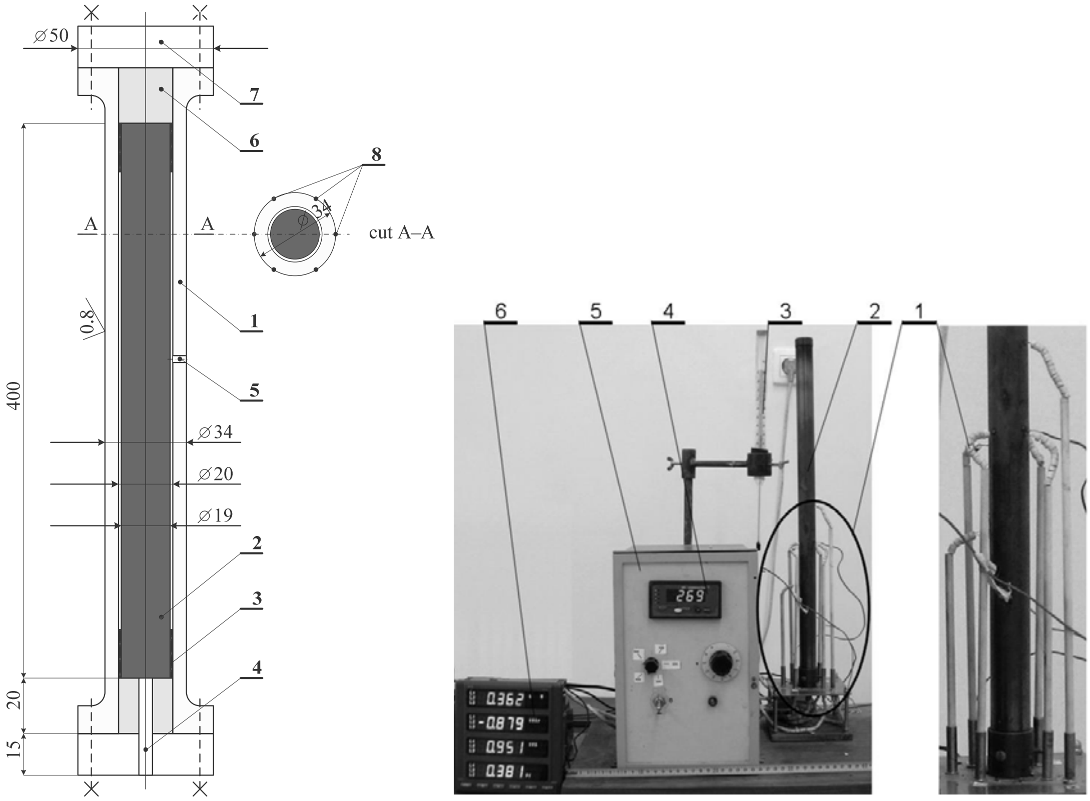
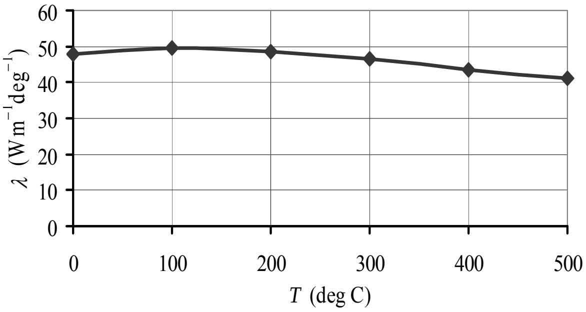


Fig.1 left: measured sample: **1** –cylindrical shell from carbon steel**, 2** – heating element 500 W, **3** – distance rings (asbestos tape), **4** – inlet for feeding of the heating element, **5** – inlet for measurement of the internal temperature, **6** – thermal insulation (fiberglass), **7** – asbestos insulation flanges, **8** – thermocouples

Fig.1 right: measuring stand: **1** – thermocouples, **2** – measuring cylinder, **3** – mercury dilatation thermometer, **4** – measuring device, **5** – temperature control unit, **6** – wattmeter

With respect to the capabilities of the registration device, all measurements could always be carried out only in steady state. Therefore, the measuring cylinder was designed in such a manner that would ensure (at relatively uniform thermal field) low thermal inertia. Fig. 2 left shows the temperature dependence of thermal conductivity of the cylindrical shell, Fig. 2 right depicts the results of the experiment. As the distribution of temperature in the wall of the shell depends on radius r only, it is not necessary to apply the boundary element method to find $\partial T / \partial n$ because this value may easily be determined from the analytical solution.



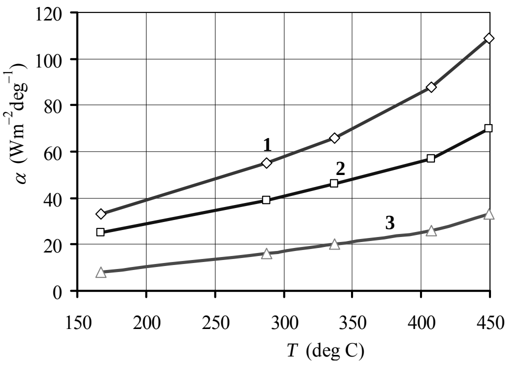


Fig. 2 left: the temperature dependence of thermal conductivity of the used steel shell

Fig. 2 right: measured temperature dependencies of particular coefficients of heat transfer
1 – α_{tot} (experiment), **2 –** α_{c} (see [7]), **3 –** α_{r} calculated from (4)

Fig. 2 right shows that the effect of radiation in the investigated range of the temperatures $T \in \langle 150, 450 \rangle$ °C is somewhat smaller than the effect of convection, but in no case negligible. All these mechanisms of heat transfer grow with the increase of the surface temperature T.

5. Conclusions

The above theoretical conclusions agree well with the results described in references. Nevertheless, in order to confirm their general validity, the authors prepare an experiment making possible to experimentally find the value of α_{tot}. But it would be also possible to prepare specific experiments providing (with a sufficient accuracy) the values α_{c} and $\overline{C}_{\mathrm{r12}}$.

Acknowledgements

The financial support of the Polish Ministry of Science and Higher Education (3 T08C 01630), Grant Agency of the Czech Republic (project No. 102/07/0496) and Ministry of Education of the Czech Republic (Research Plan MSM 6840770017) is gratefully acknowledged.

References

[1] Boley, B.A., Wiener, J.H.: Theory of Thermal Stresses. Wiley and Sons, Inc. NY, 1960.

[2] Barglik, J., Karban, P., Doležel, I., Ulrych, B., Smagór, A.: Simultaneous Effects of Convection and Thermal Radiation as a Generalized Boundary Condition of Temperature Fields of Bodies in Unbounded Gaseous Media. Electrotechnical Review 84 (2008), No. 11, pp. 134–139.

[3] Barglik, J., Czerwiński, M., Hering, M., Wesołowski, M.: Radiation in Modeling of Induction Heating Systems. IOS Press, 2008, pp. 202–211.

[4] Kachel A., Przyłucki R.: Simulation of Induction Heating Process with Radiative Heat Exchange. Journal of Achievements in Materials and Manufacturing Engineering, Vol. 22, 2007, pp. 54–56.

[5] Lykov, A.V.: Exchange of Mass and Heat. Energija, Moscow, 1972 (in Russian).

[6] Brebbia, C.A., Telles, J.C.F., Wrobel, L.C.: Boundary Element Techniques. Springer-Verlag, NY, 1984.

[7] Yamaguchi, T. et al: Study of the Cooling Rate of Steel in Heat Treatment. Mitsubischi Tech. Rev., Vol. 60, No. 8, 8, pp. 1–11, 1969.

Classification of Mixed Indications in the Defectoscopy by Eddy-Current

J. Grman, R. Ravas, L. Syrová

Department of Measurement, Faculty of Electrical Engineering and Information Technology, Ilkovičova 3, 812 19 Bratislava, Slovakia
Email: jan.grman@stuba.sk

Abstract. *The contribution treats the topic of classification of indications in the field of non-destructive defectoscopy by eddy-currents. One of the fields is classification of mixed indications into classes that are characterized by the signal shape, eventually by the signatures relating to the signal shape. The contribution concentrates on the choice of transformation of mixed indications into representation suitable for classification.*

Keywords: defectoscopy, classification, mixed indications

1. Introduction

Presented paper describes experiments with output signal from testing heat-exchanger tubing by a differential probe. It is non-destructive testing method based on eddy-current. The shape of output signal from the probe reflects properties of tested material. Potential locations of the defect in the signal are called indications. In our previous work we presented several algorithms for localisation and classification of indications. Different transformations of indications info vectors of signatures were presented too [1][2][3].

This paper is focused to classification of mixed indications. These indications are result of composition of two or more signals influenced by tube structure changes (defects or construction elements). The most common signal compositions are mixes of support plate and defects of different type. We identify six basic classes of mixed indications in our reference database of indications (Table 1.).

Table 1. Classes of mixed indications

Class ID	Class name	Dominant indication
17	Defect 100% near support plate edge	Defect
18	Defect 100% under support plate	Support
19	Defect 48% near support plate edge	Defect
20	Defect 48% under support plate	Support
21	Outer groove 20% near support plate edge	Defect
22	Outer groove 20% under support plate	Support

2. Data representations

At the experiments whose results have been published, the indication characteristics were calculated by different methods. Analysis in this paper will be specifically focused on mixed indications. Transformations were compared for ability of classification of mixed indications using different criteria. In the next section we try to briefly summarize calculation and basic properties of these representations. Representation peak-angle-peak is intuitive and was inspiration for design of representation std-cov-std. Fourier based signatures were inspired by [4] and wavelet coefficients by [5] and localization algorithm [1].

Table 2. Survey of indications transformations to signature vectors R

Transformation	Calculation
Peak-angle-peak 3components	$dist(\tilde{s},m,n)=\sqrt{((x(m)-x(n))^2+((y(m)-y(n))^2}$; $(\bar{m},\bar{n})=\underset{m,n}{\operatorname{argmax}}(dist(\tilde{s},m,n))$ $\Delta x(\bar{m},\bar{n})=\lvert x(\bar{m})-x(\bar{n})\rvert$; $\Delta y(\bar{m},\bar{n})=\lvert y(\bar{m})-y(\bar{n})\rvert$; $\phi(\tilde{\mathbf{s}}(t))=\arctan\left(\frac{\Delta y(\bar{m},\bar{n})}{\Delta x(\bar{m},\bar{n})}\right)$ $R=[\Delta x(\bar{m},\bar{n}),\Delta y(\bar{m},\bar{n}),\phi(\tilde{\mathbf{s}}(t))]$
Std-cov-std 3components	$E(x)=\frac{1}{L}\sum_{k=1}^{L}x(k)\,;E(y)=\frac{1}{L}\sum_{k=1}^{L}y(k)\,;D(x)=E[(x-E(x))^2]$; $D(y)=E[(y-E(y))^2]$ $\operatorname{cov}(x,y)=E[(x-E(x)).(y-E(y))]\,;R=[\sqrt{D(x)},\sqrt{D(y)},\operatorname{cov}(x,y)]$
Fourier transf. 30components	$\tilde{X}[k]=\sum_{n=0}^{N-1}\tilde{x}[n]e^{-i\frac{2\pi}{N}nk},k=0,1,\dots N-1;\tilde{x}[n]=\frac{1}{N}\sum_{k=0}^{N-1}\tilde{X}[k]e^{i\frac{2\pi}{N}nk},n=0,1,\dots N-1$ $R=[\tilde{\mathbf{x}}[0],\tilde{\mathbf{x}}[1],\dots,\tilde{\mathbf{x}}[15]]$
Wavelet transf. 3components	$C_{x(t)}(s,p)=\int_{-\infty}^{\infty}x(t).\psi(s,p,t)dt$; where $\psi(s,p,t)$ is wavelet function $M(s,p)=\sqrt{C_x^2(s,p)+C_y^2(s,p)}$; $(\bar{s},\bar{p})=\underset{s,p}{\arg\max}(M(s,p))$; $\alpha=\arctan(C_{y(t)}(\bar{s},\bar{p}),C_{x(t)}(\bar{s},\bar{p}))$; $R=[M(\bar{s},\bar{p}),\bar{s},\alpha]$

3. Experiments

Using reference database of mixed indications we were able to calculate reference vectors of signatures. Then we analyzed separability of signature vectors into subspaces corresponding to six selected classes from table 1. Figure 1 shows projections of calculated signature vectors for transformations peak-angle-peak and std-cov-std. It's easy to see that classes 21 and 22 can't be separated using peak-angle-peak, but using std-cov-std method we are able to classify all indications into proper classes.

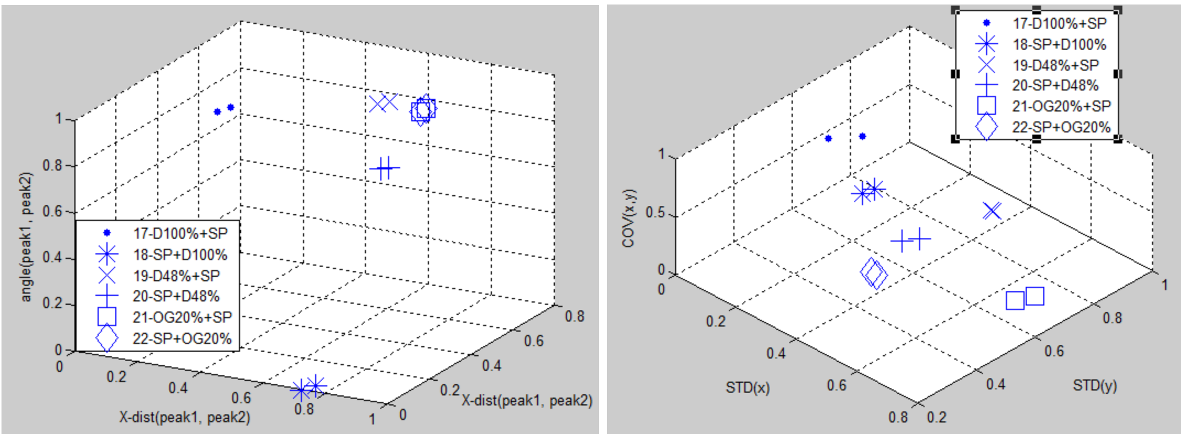


Fig. 1. Projection of signature vectors for peak-angle-peak (left) and std-cov-std (right) transformations calculated from 25kHz signal

Table 3 present results for all transformations. These results were calculated for indications from signal measured using 25 kHz frequency. It is important for interpretation of results. Low frequencies better describe changes on outer side of the tube. It means that influence of support plate outer groove is dominant.

Table 3. Description of classification errors for 25 kHz measuring signal

Transformation	Classification errors	Description
Peak-angle-peak	21/22	No separation for outer groove near or under support plate
Std-cov-std		
Fourier transf.	19/20	No separation for defect 48% near and under support plate
Wavelet transf.	19/21, 18/20, 21/22	No separation for defect 48% and outer groove near support plate edge No separation for defect 100% and 48% under support plate

Our next experiments was done using 700 kHz measuring signal. High frequencies better maps tube internal structure and changes. Signal influence by inner defects is higher. Figure 2 again presents projections of indications using peak-angle-peak and std-cov-std.

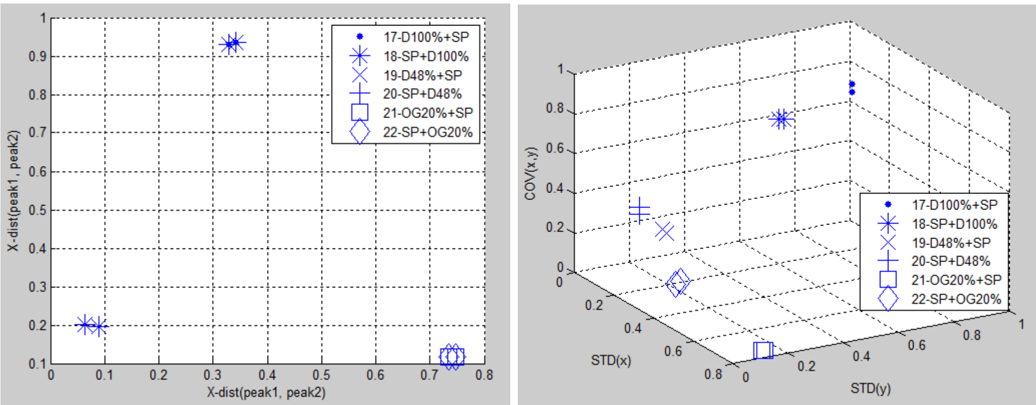


Fig. 2. Projection of signature vectors for peak-angle-peak (left) and std-cov-std (right) transformations calculated from 700kHz signal

Table 4 describes classification errors. Representation std-cov-std has again 100% success. It is interesting that representation based on Fourier transformation with dimension 30 succeed too.

Table 4. Description of classification errors for 700 kHz measuring signal

Transformation	Classification errors	Description
Peak-angle-peak	17/18,19/20,21/22	No difference if defect is near or under support plate
Std-cov-std	--	
Fourier transf.	--	
Wavelet transf.	17/18, 20/22	Defect 48% and outer groove near support plate edge Defect 100% and 48% under support plate

Presented (and omitted because of limited paper range) results indicate std-cov-std as universal transformation of mixed indications into signatures suitable for classification. But it is important to analyze other transformations. When using different frequencies, different transformations are able to produce good results.

Very good example is peak-angle-peak transformation used on 700 kHz signal. For the first look mentioned transformation produce a lot of classification errors. Bud detailed analysis shows that using this representation we are able to separate indications into very good and clear clusters (figure 2, left image). These clusters represent type of defect mixed with support plate without respect to their relative position.

4. Conclusions

Classification of mixed indications is very complicated but very important. Contact of support plate with tube is very often source of corrosion process. Experiments show that mixed indications can be successfully identified. All transformations use original signal from measurement probe. Using presented transformations we are able to omit signal mixes (linear compositions of signals in different frequencies used to suppress indications of support plate). Not each transformation is suitable for classification of mixed indications, but some of them can be used in specific signal frequencies to increase classification success, stability and reliability.

Acknowledgements

We gratefully acknowledge that the results presented here were solved as partial goals of the research task VEGA No. 1/0480/09.

References

[1] GRMAN, J. & SYROVÁ, L. : Alternative Localization of Defects Using Wavelets in the Defectoscopy by Eddy-Curents. Proc. of the 19th International DAAAM Symposium „Intelligent Manufacturing & Automation: Focus on Next Generation of Inteligent Systems and Solutions“, 22.-25. October, 2008, Trnava, Slovakia. - ISBN 978-3-901509-68-1. - s. 573-574.

[2] GRMAN, J., RAVAS, R. & SYROVÁ, L.: The Role of Wavelet Analysis in the Nondestructive Defectoscopy by Eddy-Current. In: Radioelektronika 2007 : Proceedings of 17th International Conference. Brno, Czech Republic, 24.-25.4.2007. - Brno : VUT, 2007. - ISBN 978-80-214-3390-8. - S. 471-473.

[3] GRMAN, J., RAVAS, R. & SYROVÁ, L.: Data Representation the Field of Multifrequency Eddy-Current Testing. In: Annals of DAAM for 2003 & Proceedings. Vienna : DAAAM International, 2003. - ISBN 3-901509-34-8. - s. 173-174. Konferencia: 14th International DAAAM Symposium "Intelligent Manufacturing & Automation : Focus on Reconstruction and Development". Sarajevo. Bosnia and Herzegovina, 22.-25.10.2003

[4] YAHN, Ch.T. & ROSKIES, R.Z.: Fourier Descriptors for Plane Closed Curves. IEEE Transactions on computers, March 1972.

[5] STRANG, G. & NGUYEN, T. : Wavelets and Filter Banks. Wellesley-Cambridge Press, 1996.

Utilization of Miniature Multilayer Ceramic Inductors in the Dual-Mode Crystal Oscillator

[1, 2]V. Stofanik, [1]M. Minarik, [2]V. Bohac

[1] Dept. of Radio and Electronics, Faculty of Electrical Engineering and Information Technology, Slovak University of Technology, Bratislava, Slovak Republic.
[2] Institute of Physics, Slovak Academy of Sciences, Bratislava, Slovak Republic.
Email: vladimir.stofanik@savba.sk

Abstract. *This paper deals with dual-mode crystal oscillator (DMXO) based on two bridge-type oscillators using a quartz resonator as frequency stabilization element and sensor element of its own temperature as well. We have investigated a possibility of miniature multilayer ceramic inductors employment in the DMXO. Employment of such miniature inductors instead of air-core wire-wound ones enables further reduction of the oscillator dimensions, as well as a reduction of temperature gradients among the DMXO elements.*

Keywords: Dual-Mode Crystal Oscillator, SC-cut resonator, multilayer ceramic inductors

1. Introduction

The conventional method for sensing resonator's temperature in Temperature Compensated Crystal Oscillators (TCXO), for example, utilizes a thermistor, placed in close proximity to the resonator. This method suffers from inaccuracies due to thermal lag stemming from differences in time constants and thermal gradients between the resonator and the thermistor, as well as thermistor aging. Simultaneous excitation of two modes of vibration in a piezoelectric resonator enables to realize self-temperature-sensing of the resonator. The self-temperature-sensing method eliminates temperature offset and lag effects, since no external temperature sensor is used. The history and different applications related to the dual-mode excitation have been reviewed in [1], [2].

Self-temperature-sensing of Stress Compensated (SC) quartz resonator (or SC-cut) utilizing simultaneous excitation of the fundamental c-mode (the slow thickness-shear mode) together with the 3rd overtone c-mode in the resonator has been introduced in [3]. This method has been employed in the Microcomputer Compensated Crystal Oscillator (MCXO). Since the MCXO was primary intended for military applications, it have to operate reliably in the wide temperature range between -55^{o}C and +85^{o}C [4]. Optimal quartz resonators (not standard SC-cut) with the lower turnover temperature of the 3rd overtone frequency close to +20^{o}C have been designed especially for the MCXO [5]. However later, author in [6] has presented that the differences between the aging of the two exited mode frequencies in the resonator cause an offset with a tilt in the MCXO output frequency over the operating temperature range; it limits the accuracy of the correction process implemented in the MCXO.

We have designed and investigated novel DMXO that employs a standard 10-MHz 3rd overtone SC-cut resonator with lower turnover temperature of the 3rd overtone between +80^{o}C and +85^{o}C [7], [8]. Such quartz resonators are utilized in the highest-stability Oven Controlled Crystal Oscillators (OCXO). Possible applications of the DMXO with excitation of the two overtones include stabilization of the SC-cut resonator's temperature as well as compensation for frequency shifts due to the variations of the temperature in the resonator surrounding.
In this paper we illustrate a possibility of employment of nowadays miniature multilayer ceramic inductors in the DMXO we have developed.

2. Implementation of the SC-cut Resonator Self-Temperature-Sensing

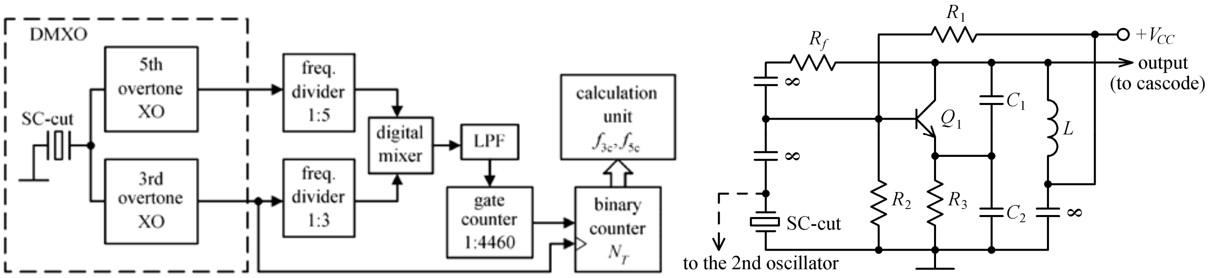


Fig. 1. Block diagram of the SC-cut self-temperature-sensing implementation (left) and simplified schematic diagram of the bridge-type crystal oscillator (right); two similar structures form the DMXO.

Block diagram of our SC-cut self-temperature-sensing implementation is shown in Fig. 1 (left). The 5th overtone oscillator's frequency divided by five is subtracted from the 3rd overtone oscillator's frequency divided by three, with assistance of the digital mixer and low pass filter (LPF). The difference frequency f_d at the output of the LPF can be expressed as follows

$$f_d(\vartheta) = \frac{f_3(\vartheta)}{3} - \frac{f_5(\vartheta)}{5} \tag{1}$$

where ϑ is the temperature of the SC-cut resonator.

The gate counter, shown in Fig. 1, produces approximately one-second time interval, during which the binary counter accumulates clock pulses with frequency f_3 (i.e. frequency of the 3rd overtone XO). After the clock pulses accumulation, the content of the binary counter can be expressed by following formula

$$N_T(\vartheta) = \mathrm{int}\left(\frac{f_3(\vartheta)}{f_d(\vartheta)} \cdot 4460\right). \tag{2}$$

The content of the binary counter (2) is used to form an independent variable $N=N_T-N_o$ that represents actual temperature of the SC-cut resonator. The integer N_o is the constant, which represents content of the binary counter at selected temperature of the SC-cut resonator, e.g. at the lower turnover temperature of the 3rd overtone frequency (Fig. 3).

The data representing the f_3 vs. N dependency (or f_5 vs. N dependency) are collected during the calibration run (with assistance of temperature chamber, precise counters and computer). The dependency usually can be approximated by simple polynomial. The calculation unit (shown in Fig. 1) then can determine an actual value of the frequency f_3 (or f_5) according to actual value of the independent variable N by computing of appropriate polynomial.

3. Employment of Multilayer Ceramic Inductors in the DMXO

We have investigated an impact of instabilities of the miniature (L×W×T=1.6×0.8×0.8mm) multilayer ceramic inductors (HK1608 series from Taiyo Yuden Co. Ltd.) due to temperature variations on the DMXO's frequencies. To be the influence more obvious, we replaced the SC-cut resonator in the particular bridge-type oscillator (Fig. 1) by the resistor with the resistance approximately equal to motional resistance of the appropriate mode, normally excited in the SC-cut resonator. Corresponding frequency vs. temperature characteristics of the particular oscillators are shown in Fig. 2 (left). The temperature was measured with assistance of miniature platinum resistance temperature detector (PT100) placed in the vicinity of the oscillator elements.

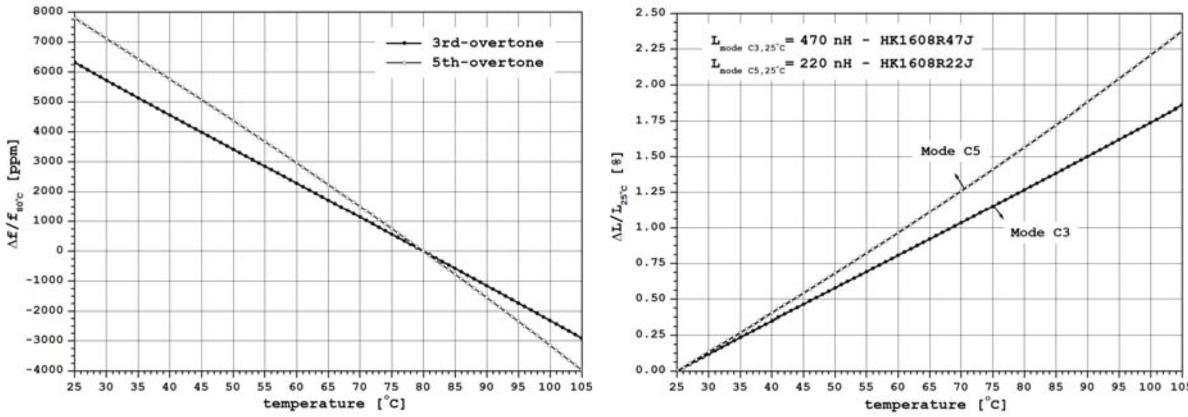


Fig. 2. Measured frequency vs. temperature of the particular oscillators forming the DMXO without the SC-cut (left) and estimated values of the inductance vs. temperature of the used multilayer ceramic inductors (right); the SC-cut resonator was replaced in the DMXO by the two appropriate resistors.

If we consider that the temperature dependency of the used inductor in the bridge oscillator is dominant, then we can estimate its inductance at the temperature ϑ using following formula

$$L(\vartheta) = L_{25°\mathrm{C}} \cdot \left(\frac{f(25°\mathrm{C})}{f(\vartheta)} \right)^2 = L_{25°\mathrm{C}} \cdot \left[1 + \alpha_L \cdot (\vartheta - 25) \right] \tag{3}$$

where α_L represents a temperature coefficient of the inductor inductance; $f(25°\mathrm{C})$ and $f(\vartheta)$ are the measured frequencies of the appropriate oscillator at room temperature (25°C) and at given temperature ϑ, respectively; and $L_{25°\mathrm{C}}$ represents value of the inductor L inductance at room temperature (i.e. its nominal value).

Estimated values of inductance of the particular multilayer ceramic inductors vs. temperature are shown in Fig. 2 (right).

4. Results and Conclusions

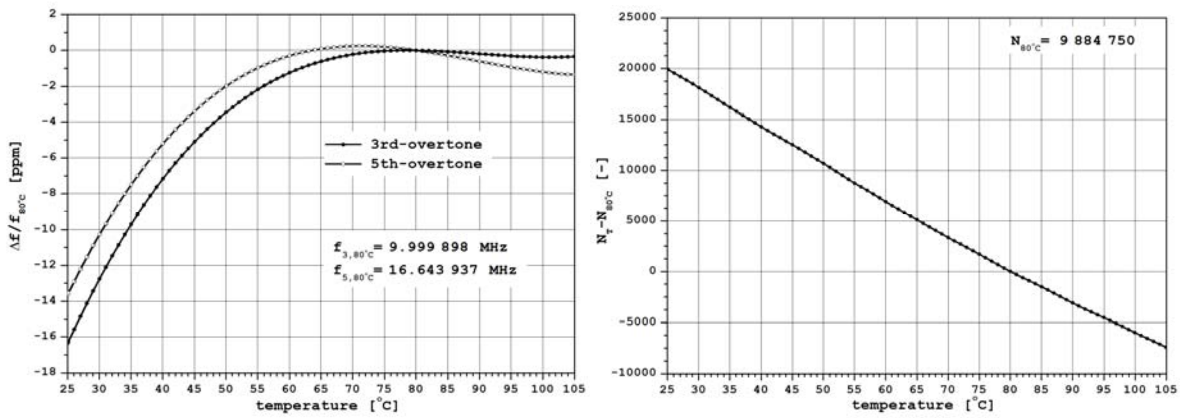


Fig. 3. Measured resonant frequency of the two investigated overtones (c-modes) of the SC-cut vs. temperature (left). Number of clock pulses accumulated in the binary counter during the time interval 4460 / f_d vs. temperature (right).

We have evaluated possibility of employment of nowadays miniature multilayer ceramic inductors in the DMXO with simultaneous excitation of two overtones in the SC-cut resonator. Inductance of investigated multilayer ceramic inductors is more temperature dependent in comparison with the wire-wound air-core inductors. However, our experimental results (Fig. 3 and Fig. 4) illustrate that the nowadays miniature ceramic inductors can be reliably employed in the DMXO circuit.

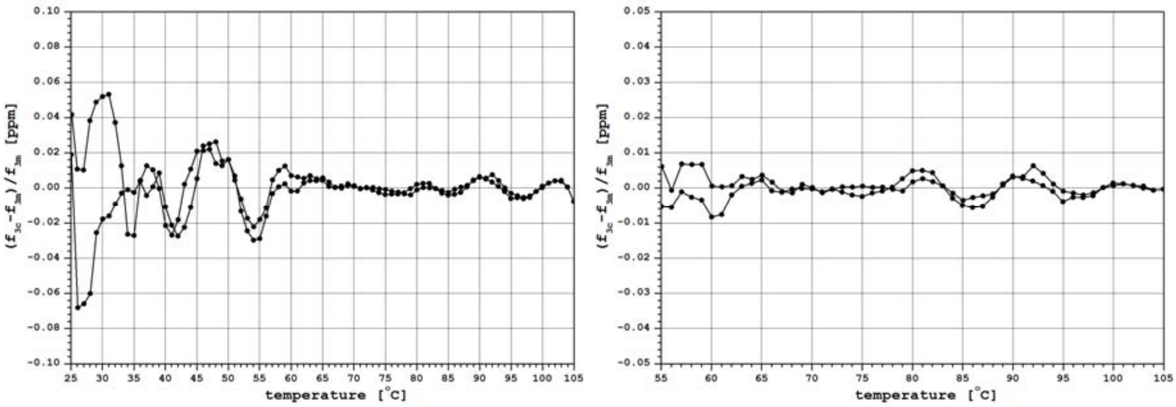


Fig. 4. Residuals vs. temperature in the case of 3rd overtone XO; data from calibration-run were fit to a single-segment 9th order polynomial over the temperature range between +25°C and +105°C (left), and over the reduced temperature range between +55°C and +105°C (right).

Consequently, the DMXO dimensions, as well as temperature gradients among the circuit elements may be reduced. Investigation of a long-term frequency stability of both excited overtones in the developed DMXO is our aim also; however, it requires a long-time evaluation of the realized prototypes.

Acknowledgements

This work was supported by the Ministry of Education of Slovak Republic under the grants No. 1/0688/09 and No. 3/7411/09.

References

[1] Vig JR. Dual-mode Oscillators for Clocks and Sensors. In proceedings of the 1999 IEEE Ultrasonics Symposium, 1999, 859-868.

[2] Vig JR. Temperature-Insensitive Dual-Mode Resonant Sensors - A Review. *IEEE Sensors Journal*, 1 (1), 62-68, 2001.

[3] Schodowski SS. Resonator Self-Temperature-Sensing using a Dual-Harmonic-Mode Osci-llator. In proceedings of the 43rd Annual Symposium on Frequency Control, 1989, 2-7.

[4] Benjaminson A, Stallings SC. A Microcomputer Compensated Crystal Oscillator using a Dual-Mode Resonator. In proceedings of the 43rd Annual Symposium on Frequency Control, 1989, 20-26.

[5] Filler RL, Vig JR. Resonators for the Microcomputer Compensated Crystal Oscillator. In proceedings of the 43rd Annual Symposium on Frequency Control, 1989, 8-15.

[6] Kim Y. Aging of Dual Mode Resonator for Microcomputer Compensated Crystal Oscillator. In proceedings of the 2005 Joint IEEE International Frequency Control Symposium and Precise Time and Time Interval Systems and Applications Meeting, 2005, 171-175.

[7] Stofanik V, Balaz I, Minarik M. Dual-Mode Crystal Oscillator with Simultaneous Excitation of Two Overtones in a Stress Compensated Quartz Resonator. In proceedings of the joint 2007 European Frequency and Time Forum and the 2007 IEEE International Frequency Control Symposium, 2007, 227-229.

[8] Stofanik V, Minarik M, Balaz I, Cocherova E. Investigation of Self-Temperature-Sensing of SC-Cut with Simultaneous Excitation of Two Higher Overtones in Dual-Mode Crystal Oscillator. In proceedings of the 2008 IEEE International Frequency Control Symposium, 2008, 170-173.

Preparation and Characterisation of Al_2O_3-Y_2O_3 Glass Microspheres

A. Prnová[1, 2], R. Karell[2], D. Galusek[2], I. Vávra[3]

[1]Department of Magnetometry, Institute of Measurement Science,
Slovak Academy of Sciences, Dúbravská cesta 9, 841 04 Bratislava, Slovakia
[2]Vitrum Laugaricio – Joint Glass Centre of the IIC SAS, TnU AD,
FChFT STU, and RONA, a.s., Študentská 2, 911 50 Trenčín, Slovakia
[3]Department of Superlattices, Institute of Electrical Engineering,
Slovak Academy of Sciences, Dúbravská cesta 9, 841 04 Bratislava, Slovakia

Abstract: *Al_2O_3-Y_2O_3 microspheres with high alumina content were prepared and characterised. Because Al_2O_3 is not a typical glass former, preparation of aluminate glasses in bulk is difficult due to high melting temperatures and high tendency to crystallisation, which requires high cooling rates. For this reason the studied glasses were prepared in the form of micropsheres with diameters of up to 40 μm, which ensured the quenching rates at the order of magnitude of 1.10^3 K^{-1}. Despite high cooling rates most compositions comprised, apart from fully amorphous microspheres, also microspheres containing crystallites of α-Al_2O_3, YAG (yttrium aluminium garnet, and YAP (yttrium aluminium perovskite). The degree of crystallisation was found to depend from the composition, and diameter of prepared microspheres.*

1. Introduction

Yttrium-aluminate, and generally rare earth aluminate, glasses with high alumina content are considered as potential candidates for various applications ranging from transparent ballistic protections, through infrared transparent windows, to hosts for rare earth elements in materials used for solid state lasers, due to their good mechanical properties, high refraction index and good corrosion resistance. However, the preparation of aluminate glasses in bulk is difficult because these systems have high melting temperatures and high tendency to crystallisation.

The work of Rosenflanz et al. [1] describes preparation of high alumina glasses and glass ceramics with aluminate glass matrix and dispersed nanosized crystals of rare earth aluminates, with hardness between 14.4 and 18.3 GPa and the fracture toughness between 2.1 and 4.2 $MPa.m^{1/2}$ by pressure-assisted viscous flow sintering of glass microbeads at temperatures above their glass transition region.

Our previously published work describes the synthesis of Al_2O_3 – Y_2O_3 glass microspheres with various Al_2O_3 content in O_2-CH_4 flame and with the use of in-house built experimental equipment, and the attempts to prepare fully dense transparent bulk glass by hot pressing in the temperature interval between Tg and the onset of crystallisation. As yet, all attempts were unsuccessful due to extensive crystallisation during heat treatment as the consequence of the presence of nuclei formed in the synthesised microspheres during cooling. The present work is focused at more detailed characterisation of the microspheres from the point of view of their morphology, size, phase composition, and microstructure by optical microscopy (OM), scanning electron microspcopy (SEM), transmission electron microscopy (TEM) and by X-ray powder diffraction (XRD).

2. Experimental

The compositions of the synthesised specimens (as weighed) are presented in the Table 1. High-purity oxide powders (Y_2O_3 - Treibacher Industrie AG, Austria, Al_2O_3 – Taimicron TM

DAR, Krahn Chemie GmbH, Germany) were used as starting materials. The alumina powder was mixed with yttria dissolved in nitric acid, the mixture homogenised by ball milling, yttrium nitrate converted to hydroxide by the addition of NH_4OH, dried, sieved, and pre-reacted at 1650 °C, and then fed into a modified gas burner axially into the centre of methane-oxygen flame. The microspheres quenched by spraying with water were collected in a sedimentation vessel, dried, and characterised. The morphology was examined by a Nikon Eclipse ME 600 optical microscope and the diameter size distribution determined with the use of the image analysis software Lucia v. 4.82. Detailed morphology was studied with the use of a Zeiss EVO 40 SEM. The internal structure of microspheres was examined by a JEOL 1200EX TEM at the acceleration voltage 120 kV. X-ray diffraction was carried out at the powder diffractometer STOE with CoKα radiation at the wavelength of 1.78897 Å in the 2Θ range 15 – 85°. The weight fractions of crystalline phases were determined from X-ray diffraction patterns from integral intensities of diffraction peaks by the method of standard addition.

3. Results and discussion

Preliminary examination of the prepared aluminates after flame synthesis by OM indicated they were molten thoroughly in the flame, resulting in nearly ideal spherical morphology. The spheres were transparent, and without any obvious traces of crystallisation (Fig. 1a, b). However, the SEM examination revealed striking differences among surface morphologies of microspheres of various compositions and sizes (Fig. 1c, d). In some compositions, especially in those close to the composition of the pseudobinary eutectic YAG – Al_2O_3 (A60, T_m = 1760 °C, Fig. 1c), all micospheres irrespective of their size were smooth, indicating their highly amorphous character.

However, in compositions with higher alumina contents and higher melting temperature (e.g. A75, T_m = 1900 °C), in microspheres with larger diameters SEM revealed regular features suggesting at least partially crystalline nature of these particles.

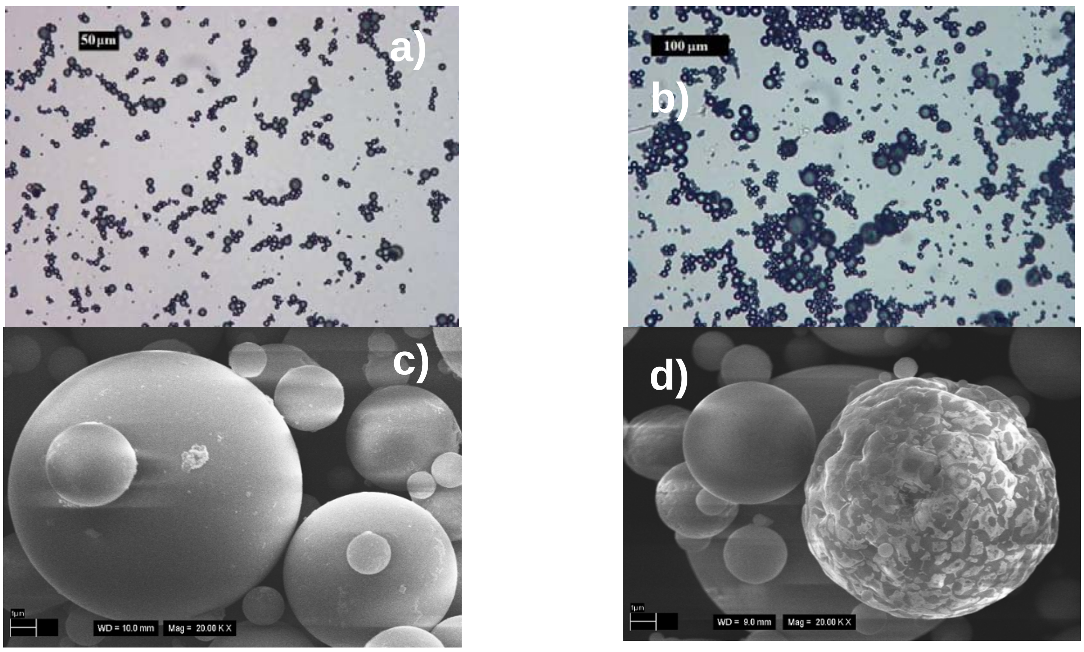


Fig. 1 Optical micrographs of the microspheres A60 (a) and A75 (b), and the detailed SEM micrographs of the same specimens, A60 (c) and A75 (d).

Another distinct feature is different distribution of diameters of prepared microspheres (Fig. 2) with various compositions. As all precursor powders underwent the same treatment, including sieving through a 40 μm mesh screen, the reason for such behaviour is not clear at the moment. It might be related to a different tendency to agglomeration in powders with different compositions, which then results in different granulometry of powder agglomerates fed to the flame. Generally speaking, the compositions closer to the eutectic showed more narrow distribution of microspheres' diameters, with sizes up to 15 μm, while those further apart contained particles with sizes up to 40 μm.

Sample	Al_2O_3 (wt.%)	Al_2O_3 (mol.%)	T_m/ºC	Amorphous phase / wt %	Crystalline phases/wt.%
A43	43	62,6	1925	92	YAG – 100 %
A57	57	74,6	1790	87	YAG - 88 % α-Al_2O_3 - 12 %
A60	60	76,8	1760	94	YAG - 100 %
A63	63	79,0	1775	96	YAG - 56 % YAP - 32 % α-Al_2O_3 - 12 %
A67	67	81,8	1825	89	YAG - 57 % YAP - 14 % α-Al_2O_3 - 16 % trans. Al_2O_3 - 13 %
A75	75	86,9	1900	85	YAG - 26 % YAP - 29 % α-Al_2O_3 - 19 % trans. Al_2O_3 - 26 %

Table 1. The compositions (as-weighed) of prepared microspheres and their basic characteristics: T_m-melting temperature as determined from the phase diagram, the content of amorphous phase, and phase composition of the crystalline fraction.

XRD examination revealed some relations between the composition, granulometry, phase composition and degree of crystallinity of prepared microspheres. Generally, the compositions closer to the eutectic point (A60, A63) were about 95 % amorphous. Several batches of the eutectic composition A60 were even 100 % amorphous. This is easily explained by the lowest tendency to crystallisation in eutectic mixtures, or in compositions close to them. The composition A43, corresponding to stoichiometric YAG, was 92 % amorphous, the rest being crystalline YAG. On the high alumina end of the composition spectra the specimens showed the degree of crystallisation higher than 10 %. The composition of the crystalline part corresponded roughly to the phase diagram, the main crystalline phases being YAG and α-Al_2O_3. The presence of the YAP phase is the result of incomplete YAG formation according to the reaction scheme:

$2Y_2O_3 + Al_2O_3 \rightarrow Y_4Al_2O_9$ (mellilite) T = 900 – 1100 °C (1)

$Y_4Al_2O_9 + Al_2O_3 \rightarrow 4YAlO_3$ (perovskite) T = 1100 – 1250 °C (2)

$3YAlO_3 + Al_2O_3 \rightarrow Y_3Al_5O_{12}$ (garnet) T = 1400 – 1600 °C. (3)

The presence of transition aluminas (δ, θ-Al_2O_3) is related to reaction of alumina with quenching water during flame synthesis of microspheres. The higher extent of crystallisation is related to higher tendency to crystallisation in compositions further from the eutectic, or in those identical to stoichiometric crystalline phases (YAG), and to insufficient cooling rate in larger spheres, which were documented in these compositions.

The TEM examination confirmed completely amorphous character of smaller microspheres, revealing heterogeneous nucleation and growth of nanosized crystals from the surface of larger ones into their interior. In some cases polycrystalline microspheres with the grain size of about 200 nm and the microstructure resembling polycrystalline α-Al_2O_3 or YAG were observed. Detailed TEM investigation of the present crystalline phases by electron diffraction is in progress.

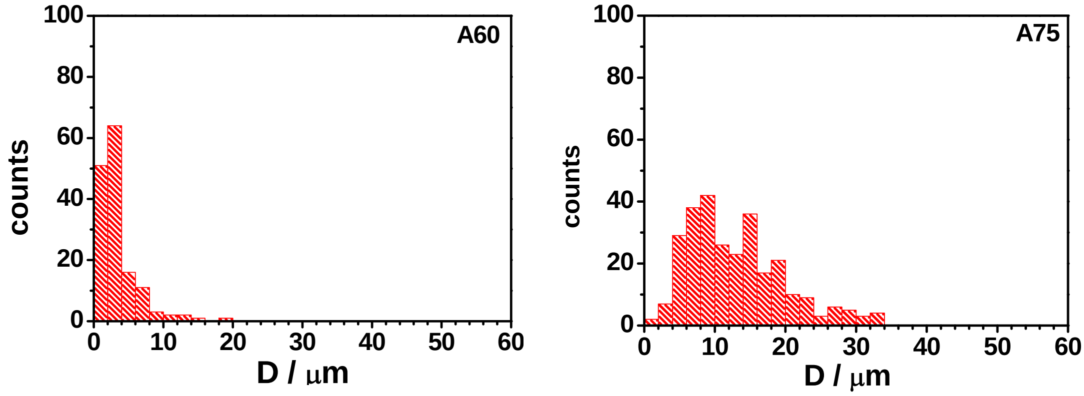


Fig. 2 Distribution of diameters of microspheres of the eutectic composition A60 and the composition A75 at high alumina end of prepared compositions.

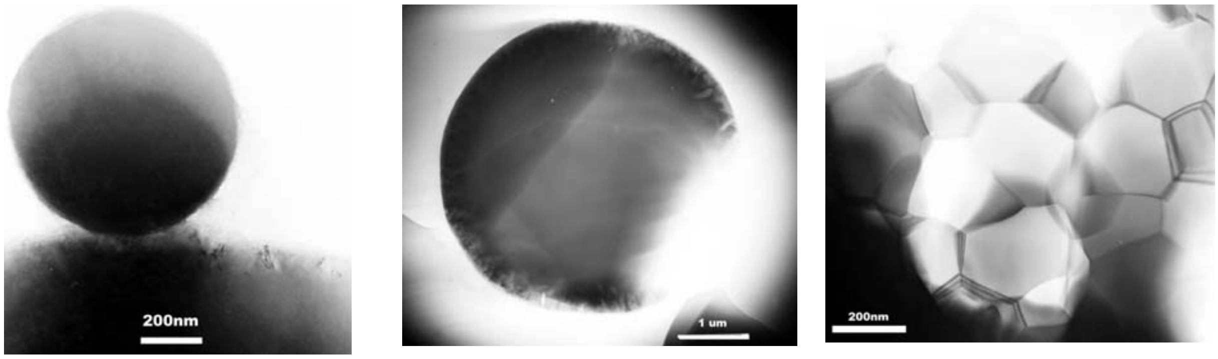


Fig. 3 TEM micrographs of prepared glass microspheres: a), A43, small, completely amorphous microsphere, b), A75, larger microsphere with surface crystallization, c) A57, polycrystalline microsphere with microstructure resembling polycrystalline α-Al_2O_3 or YAG.

4. Conclusions

Amorphous yttrium aluminates of various compositions and with high alumina contents were prepared by flame synthesis from powder precursors, and characterised by OM, SEM, TEM and XRD. The methods confirmed spherical morphology of prepared specimens, with diameters up to 40 µm. Spheres with smaller diameter were completely amorphous, while those with lager diameter were often partly, or completely crystalline, containing nanocrystals of YAG, YAP, or α-Al_2O_3 growing from the surface to the spheres' interior. In some case completely polycrystalline spheres were observed with microstructures resembling polycrystalline α-Al_2O_3 or YAG. In addition, transition aluminas were detected in microspheres with higher alumina contents.

Acknowledgement

The financial support of this work by the grants VEGA 1/0603/09 and VEGA 2/7083/27 is gratefully acknowledged.

References

[1] Rosenflanz A., et al., *Nature*, **430**, 761-764 (2004)

[2] A. Prnová, R. Karell and D. Galusek, Y_2O_3 – Al_2O_3 binary glass microspheres: Synthesis and Characterisation, *Advanced Materials Research,* **39-40** 189-192, (2008)

[3] A. Prnová, PhD Thesis, Alexander Dubček University of Trenčín, 2008

[4] Zhydachevskii Ya., et al., *J.Phys.Condens.Matter.*, **20**, (2008)

Investigation on Comparability of Surface- and Material Dependent Measurements in Multi-Sensor Coordinate Measuring Machines

M. Vogel, M. Rosenberger, G. Linß, S.C.N. Töpfer

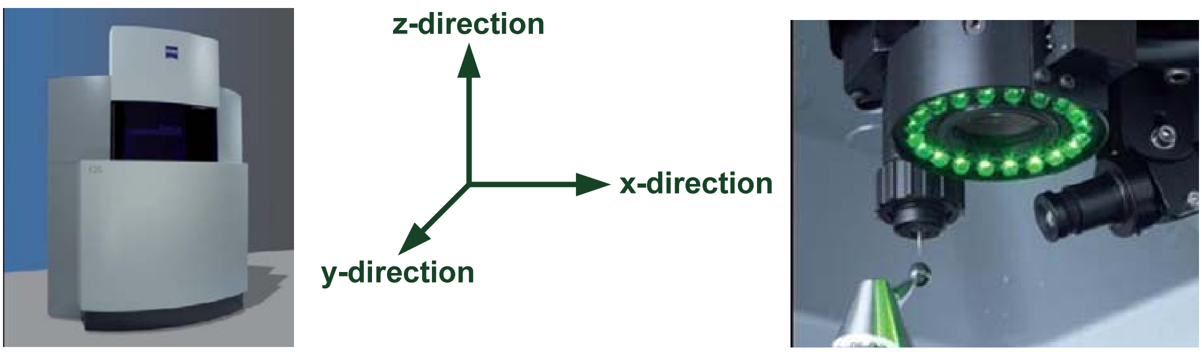
Technische Universität Ilmenau, Ilmenau, Germany,
Email: Michael.Vogel@tu-ilmenau.de

Abstract. *Multi-sensor coordinate measuring machines (MS-CMM) offer new possibilities for probing work pieces. The advantage of these sensor systems lies in the different working principles of its individual sensors. Tactile and non tactile sensors are often combined in such a sensor system. Most non tactile sensors deployed are using optical principles to acquire the measuring points. The quality of these points depends besides various factors significantly on the surface characteristics of the sample. Typically different measurement results are attained for different work piece material and surface characteristics. Consequently, there exists a characteristic displacement between a tactilely and optically measured surface point. This paper describes an experiment exploring this displacement between the point detection of a touch probe and an optical CCD matrix sensor.*
Furthermore the issue of the displacement stability in relation to the material properties is discussed.

Keywords: sensor displacement, tactile point detection, optical point detection, 3D coordinate measurement techniques, focus based height measurements with CCD matrix sensors;

1. Introduction

The coordinate measurement technique is a very important technology for quality assurance in wide ranges of industrial quality assurance applications. The measuring range spreads from few metres up to few nanometres [1]. The range of different samples demands tactile and non tactile sensors. Tactile sensors utilise the direct contact to the surface of the measuring object. During the measuring process only one point without moving the positioning stage can be sampled. Furthermore there are samples like thin membranes which can not be sampled with a tactile sensor, because the measuring object would be destroyed.



Deployed multi-sensor coordinate measuring machine (on the left), the orientation of the machine coordinate system (in the middle), probing sensors of the MS-CMM (on the right) [1]

In that case an optical sampling method is to be preferred. Choosing an optical matrix sensor for this task has the advantage getting more measuring points without moving the positioning stage. Additionally there is no contact to the sample surface. Thus, it is advantageous to combine these two measuring principles in relation to the intended measuring task. There are many applications in industry which require multi-sensor capabilities. One of the most discussed topics is the objectivity of optical measurements in comparison to tactile measurements. In our experiment the results of tactile point detection are compared to optical point detection. The focus in this paper lies on the deviation in the z-coordinate of the detected points in dependence of surface and material properties [Fig.1]. In the following discussions the coordinates are oriented as shown in Fig.1.

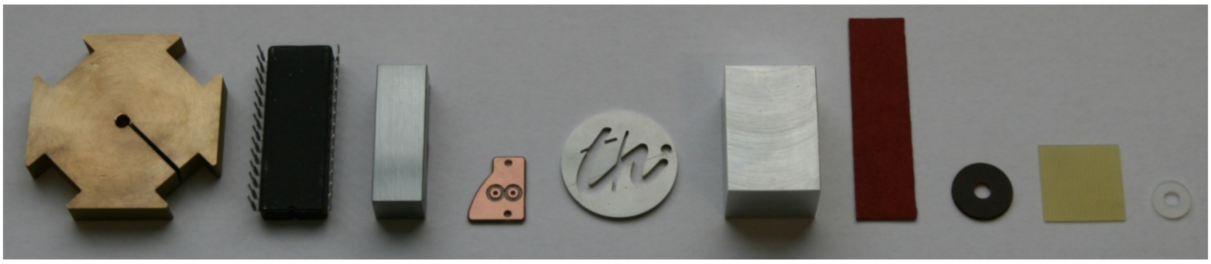
Demands for optical measurements with optical matrix sensors in comparison with tactile measurements

In cases of optical point detection and objectivity in optical measurements a wide range of influencing factors must be considered. In [2], [3], [4], [5] the most influencing factors for optical measurements were described. In summary it is necessary to follow the hints regarding illumination and focus criteria. For our experiments always the same focus criterion was used. Furthermore a comparable illumination was established by controlling the gray scale diagrams for each measuring object individually. Beside this, the same light source was used, id est bright field top light illumination.

2. Subject and Methods

Measuring objects and experimental setup

The experiment was done with a 3D coordinate measurement machine F25. It is a machine from the Carl Zeiss Company. This machine is equipped with a tactile sensor and a CCD camera with different magnifications. The measuring volume is 130x130x100 mm³ with a length measuring deviation of 0,25+l/666 µm (MPE in comparison to DIN EN ISO 10360-2 (MPE_p = 0,3 µm)) [1].



The measuring objects for the experiment (from left to right: first six metallic materials; last four non-metallic materials)

For the experiment the tactile sensor with a probing sphere diameter of 120 µm was used. The sensitivity of the tactile probe is configurable and had a value of 0,6 mN during the tests. For the optical measurements the optical sensor ViScan was used. During the experiment the magnification was 10.0 and the illumination wavelength 532 nm. Both sensors were probing the sample from above (in negative z-direction). For the analysis the measuring software Calypso was used. For the experiment ten different material samples were chosen, six types of metal and four types of non metal material [Fig. 2.]. The focus hereby lies to get information about typical surfaces in industrial measuring applications. The manufacturing process differs from laser beam cutting, electric discharge machine, and turning centre up to injection moulding machines. So a wide band of characteristic surface properties were tested. The surface properties are characterized in [6]. Especially the bidirectional reflectance distribution function (BRDF) describes the interaction of light source with the material surface. It shows the difficulty of optical conditions in case of measuring on a surface structure [7],[8],[9]. Recent works are using the Cook-Torrance BRDF model [10] in order to model specular reflection [11]. Additionally to the height measurements with the coordinate measurement machine, roughness-measurements were done.

Sensor calibration method

Both sensors were calibrated on an ultra precise sphere. Furthermore the optical sensor was calibrated under the test rules of the VDI 2617 part 6. At first the operator had to measure four points with the tactile sensor. After that the sphere is scanned automatically collecting many measuring points. The same procedure on the same sphere had to be done with the optical sensor. The middle point coordinates of the sphere were stored in the machine. After that a software algorithm inside the machine software can correct the displacement between the two sensors. During the calibration process also a material and surface depending offset occurs.

Measuring process chain

Every sample was touched ten times with the tactile sensor at three different points. The identical three points were focused with the video optical sensor. As focus criterion the area probing criterion was used. Considering the optical transfer characteristics of the tested materials the illumination had to be adjusted to keep a constant grey level. The grey level amounted to 160±10 (increments), (camera resolution 8 bit).

The measured values for the coordinate “z” (height) of both measurements were compared. The average displacement between both measuring sensors and the standard deviation are given in Tab.1 . The table shows the results of one detected point of these measurements.

Tab. 1. Measurement results at different materials

Nr.	**material**	**R_z**	**R_a**	**touch probe**		**optical probe**		**average value Δz (mm)**	**$б_z$ (mm)**
				average value z (mm)	**$б_z$ (mm)**	**average value z (mm)**	**$б_z$ (mm)**		
1	CuZn	10,7	2,2	-1,822269	0,000153	-1,974719	0,000457	0,152450	0,000535
2	Al203	3,8	0,7	-0,000337	0,000027	-0,148015	0,000272	0,147678	0,000284
3	hardened steel (DIN 861)	3,5	0,6	0,001737	0,000033	-0,142271	0,000216	0,144008	0,000205
4	Cu	1,8	0,3	-0,155648	0,000032	-0,299356	0,000485	0,143708	0,000490
5	CrNi	1,0	0,1	-0,137287	0,000013	-0,280664	0,000405	0,143377	0,000405
6	Al	0,7	0,2	0,000529	0,000007	-0,142545	0,000555	0,143074	0,000552
7	PET (red)	10,2	2,0	-0,584480	0,000026	-0,732391	0,000345	0,147911	0,000351
8	PET (brown)	2,0	0,3	-0,187274	0,000041	-0,334853	0,000491	0,147578	0,000498
9	FR4 synthetic resin	2,3	0,5	-0,636396	0,000010	-0,783443	0,000869	0,147047	0,000866
10	PET (transparent)	0,8	0,1	-0,378031	0,000272	-0,520717	0,001399	0,142687	0,001387

Afterwards the roughness measurements were done. Therefore the samples were also proofed ten times. The results of the roughness measurements are given in Tab.1 with R_a and R_z.

3. Results and Discussion

Table 1 contains the measurements with the touch probe. It shows the stability and the precision of the machine. Mostly the standard deviation lies under 153 nm for tactile and 1.44µm for optical probing. The offset between the optical and tactile measurement depends on the material, the material surface and the calibration process. A look at the standard deviation on a combined optical-tactile measurement shows for metal surfaces a maximum value of 552nm. This is corresponding to a high degree of stability. It shows the possibility to correct these offset values. For the non metal samples the standard deviations are up to factor six higher.

Overall the measurements showing a material and surface characteristics related deviation during the point detection in the z-coordinate. The roughness measurements reveal the dependencies on the surface roughness and the error behavior. A high roughness value leads to greater deviations in the height measurements. It is the predominant influence for the tactile measurements (compare columns 3 and 4 with column 6). As for optically probed points this is not true (column 8). Roughness has a significant influence but it is not the only one. One other important factor for optical measurements is the type of reflection on the surface. For optical probing it is a great difference whether diffuse or specular reflection is the predominant component of the observed reflection image. For metal surfaces specular reflection may be the most predominant reflection component depending on the observation angle, the illumination angle and on the local surface structure. Using the Cook Torrance model describing a surface with a microfacet model specular reflection can be well modeled. Additionally a higher surface roughness results in a more efficient light trapping and absorption. Thus, a lower brightness is observed. At measuring objects 7 to 10 diffuse reflection is predominant. Measuring object 10 is a special case for it is transparent. Considering the tactile results on measuring object 10 it poses the interesting question which influence leads to the high value of the standard deviation σ_z? Whether it is an outlier or not shall be investigated in future experiments.

4. Conclusions

The experiment shows to what extent an optical measurement depends on the optical surface characteristics of the sample.

Our future research will deal with these effects using tactile sensors with different probing ball diameter and different optical magnifications. Furthermore a cross-validation with results acquired with direct mounted sensors is interesting and shall be done. Additionally the separation of different surface roughness and different optical surface properties has to be done. This research will provide a new quality and stability of measurement results of multi-sensor measurements.

Furthermore the trend in mechanical engineering techniques goes to Computer Aided Designs (CAD). During the construction process, during the production process and also during the measurement process CAD data are available. Therefore a lot of additional information is given. It can be used for automated inspection planning. So the material and the processing parameters are known. With this knowledge and the results of our ongoing research a new quality in measurements can be achieved.

Acknowledgements

We thank the Faculty of Mechanical Engineering for the support during the experiments. We express our gratitude to the research programme Innoprofile which is funded by the BMBF (Federal Ministry of Research and Education).

References

[1] Carl Zeiss Industrielle Messtechnik GmbH: „ F25 Messen im Nanometerbereich“. f25_60-20-126-iii-d.pdf, http://www.zeiss.de/imt, Oberkochen 2007

[2] Nehse U., Linss G., Kühn O.: Method for automatic of focus and lighting and for objectivated scanning of edge site in optical precision measuring technique, PCT WO02/084215 A1, international publication date 10/2002.

[3] Töpfer S.C.N., Linß G., Nehse U.: Quality measures for depth-from-focus approaches. In: Proceedings of SPIE, vol. 6180 (2006), pp. 55-60.

[4] Töpfer S.C.N.: Automatisierte Antastung für die hochauflösende Geometriemessung mit CCD-Bildsensoren. Weißensee Verlag, 2008.

[5] Beets S., Roithmeier R., Weckenmann A.: Ausbildungskonzept Koordinatenmesstechnik – AUKOM. In: FQS-DGQ, vol. 81-01 (2002), p.7.

[6] EN ISO 8785:1999 Geometrical Product Specifications (GPS) ” Surface imperfections - Terms, definitions and parameters (ISO 8785:1998)”; German version

[7] Nehse, U.: Beleuchtungs- und Fokusregelungen für die objektivierte optische Präzisionsantastung in der Koordinatenmesstechnik. Technische Universität Ilmenau, Dissertation, 2001

[8] Shree K. Nayar, Eatsushi Ikeuchi, and Takeo Kanade. Surface Reflection: Physical and Geometrical Perspectives. Technical report, The Robotics Institute, Carnegie Mellon University, 1989

[9] Gall D.: Grundlagen der Lichttechnik - Kompendium. 1.Auflage. München: Pflaum Verlag, 2004

[10] Cook R., Torrance K.: A reflectance model for computer. ACM Transactions on Graphics, 1, 1 (1982), pp. 7-24

[11] Obein G., Leroux T., Knoblauch K., Vienot F.: Visually relevant gloss parameters. Proc. of 11th Int. Metrology Congress, Toulon, 2003, p. 6

Evaluating the Potential of Infrared Thermography in the Study of Peripheral Arterial Occlusive Disease

[1]M.G. Sowa, [1]J.R. Friesen, [2,1]M. Hain

[1]Institute for Biodiagnostics, National Research Council Canada, Winnipeg, Canada,
[2]Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovakia
Email: mike.sowa@nrc-cnrc.gc.ca

Abstract. *The diagnosis of peripheral vascular disease is not an easy one to make. It typically begins with an in office assessment that commonly includes reporting of presenting symptoms (with questioning) followed by inspection and palpation (physical examination). This in-office assessment is followed by vascular lab testing and may include Doppler ultrasound, ultrasound scanning, transcutaneous oxygen monitoring, x-ray, CT or magnetic resonance angiography. However, not all technologies can be used reliably in all patients. Because this disease can be difficult to stage, challenges designed to alter blood flow to the extremity (whether they be exercise-induced, occlusion-induced, or supine/dependency states) are occasionally used for additional insight. We investigate the potential of infrared thermography in the evaluation of peripheral arterial occlusive disease and compare a conventional challenge with no challenge at all.*

Keywords: Infrared thermography, peripheral vascular disease, nonlinear fitting

1. Introduction

Peripheral arterial occlusive disease (PAOD) is characterized by poor circulation of blood to the extremities and most generally the lower leg and foot. The disease carries a significant patient morbidity and it is generally the leading cause of amputation in adults [1]. Atherosclerosis exacerbated by the vascular consequences of diabetes and/or smoking is usually the underlying cause of the disease. Narrowing or blockage of the vessels impedes blood flow to the extremities. Starved for blood, fingers or toes are often the first tissue to suffer leading to open, chronic wounds that can be further infected; left unchecked the entire limb may require amputation. At the outset the symptoms of the disease usually consist of the patient developing pain in the calf muscles of the leg when they walk with the pain subsiding on resting. This is known as intermittent claudication and patients are usually treated conservatively at this stage. As the disease advances, pain persists at rest, non-healing wounds may appear and gangrene may develop. At this stage, known as critical ischemia, more aggressive treatment intended to restore blood flow to the extremities is necessary: bypass surgery or less invasive procedures such as angioplasty or atheroectomy can be carried out to restore blood flow.

Treatment is based on the severity of the disease therefore accurate diagnosis and the ability to monitor its progression is essential. Angiographic methods are difficult and costly to carry out routinely and are typically reserved to determine if and what type of revascularization procedure should be used and if a revascularization procedure was successful. More commonly ankle pressure, ankle brachial index (ABI; ankle pressure/arm (brachial) pressure), toe pressures, and transcutaneous oxygen measurements are used to help assess the disease. Each measurement technique has drawbacks or limitations and often more than one of these various measurement methods is used to provide an understanding of the severity of the disease. For example, toe pressures which measures the magnitude of the pulse of blood at the toe using photoplethysmography are an important measure in diabetic patients who generally show falsely elevated or incompressible ankle pressure readings due to extensive vascular calcification [5]. The reproducibility of some of the measurements has also been criticized. For instance, the variability of ankle-brachial pressure index measurements made by trained staff has been reported as 15% [6]. Simple and reliable methods to measure peripheral perfusion that could be used routinely could enhance our ability to diagnose PAOD and result in improved clinical management of the disease. In this paper we examine infrared thermographic measurement of perfusion to the extremities to help study PAOD.

2. Subject and Methods

Infrared Thermography

For the contactless temperature measurement a ThermaCAM SC 3000 (FLIR Systems AB) thermographic camera equipped with GaAs quantum well infrared photon detector Stirling cooled to 70 K was used. The detector has a spectral range between 8-9 μm, providing 0.03 ^{o}C thermal sensitivity at 30 ^{o}C and an accuracy of 1% of the measured value or ±1 ^{o}C (for measurement range up to 150 ^{o}C).

Clinical Study & Experimental Protocol

The PAOD study took place at St. Boniface General Hospital in Winnipeg, Manitoba and was approved by the Winnipeg Research Ethics Board of the Institute for Biodiagnostics, the Research Review Committee of St. Boniface General Hospital and the Research Ethics Board of the University of Manitoba, Faculty of Medicine. A total of 60 subjects were enrolled in the study; age range was (47-88), mean age was 71. Based on clinical examination of presenting symptoms and patient history, pressure and transcutaneous oxygen measurements, subjects were divided into three groups, 20 with claudication or rest pain (without tissue loss), 20 with critical ischemia (possible tissue loss, ulcers, and/or gangrene) and 20 controls with normal peripheral blood circulation. All subjects provided written informed consent.

Perfusion in the feet of patients was assessed with infrared thermography in combination with ankle pressure, ankle brachial index, toe pressures, and transcutaneous oxygen measurements. A pressure – cuff challenge was used to invoke changes in perfusion to the foot and examine perfusion recovery times. The challenge consisted of three phases: 1. Rest; 2. Ischemia where a pneumatic cuff on upper thigh was inflated to pressure of 30 mmHg above thigh systolic pressure for 3 minutes to occlude arterial flow; 3. Reperfusion where cuff pressure was released. Infrared thermographic measurements were made throughout this protocol. Figure 1 outlines the expected results from such a challenge.

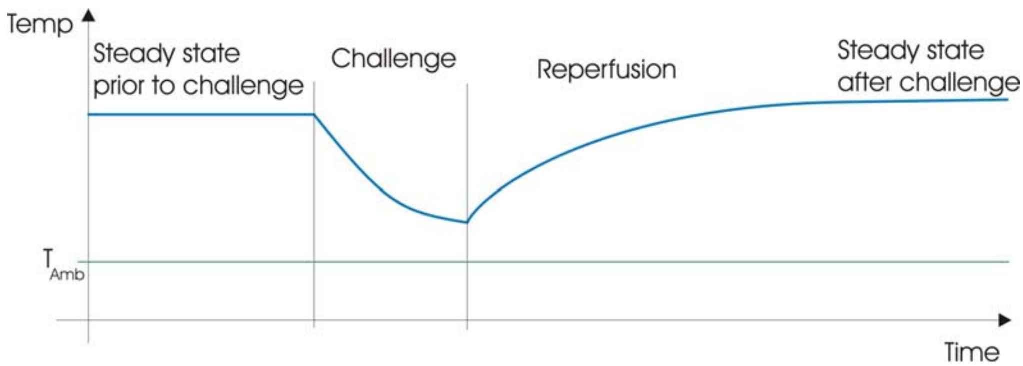


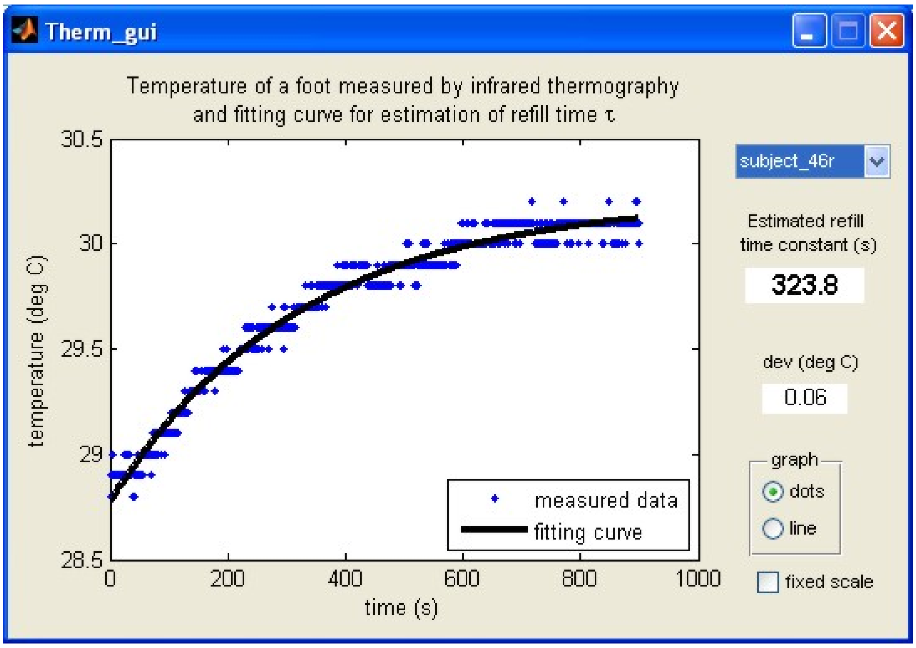
Fig. 1. Schematic of the expected change in the temperature of the foot over the experimental protocol.

Data Analysis

The time dependence of the change in temperature, *T(t)*, of the foot as a result of the challenge was fit a function

$$T(t) = T_0 + \Delta T.\left(1 - e^{-\frac{t}{\tau}}\right)$$

where T is temperature of the foot at the selected point, T_o is steady state temperature of the foot prior to challenge, t is time, ΔT is temperature difference prior and after challenge, and τ is the time constant. The unconstrained multivariable nonlinear optimization method



supported by the MATLAB function *fminsearch* was used to estimate the parameters of this model.

Fig. 2. Example of the thermographically measured time dependence of temperature of a lower extremity during reperfusion phase and nonlinear fitting curve with estimated refill time t_{TH}

Correlation analysis was performed to understand the association between the thermographic measurements as well as with the assessment of PAOD disease. An ordinal score, DSI = Disease Seriousness Index, was introduced where 1= control group, 2= claudicant group and 3= critical group. Correlation analysis was performed on the following variables, DSI, t_{TH} the thermographic refill time, T_{TOE} the temperature of the big toe prior to challenge, T_{ANK} the temperature of the ankle prior to challenge and $T_{ANK} - T_{TOE}$ the difference between the temperature of the ankle and temperature of the big toe prior to challenge. Correlation analysis was performed in the MATLAB environment.

3. Results

Correlation analysis demonstrates a strong association (r > 0.8) between the static thermographic measurements made at the ankle and toe prior to the challenge with the static thermographic measurement at the toe showing a modest negative correlation (r = -0.33) with the disease seriousness index. Static thermographic measurement of the ankle were less correlated with the disease index (r = -0.19) but both ankle and toe measurements displayed the expected negative correlation with disease severity. People with more advanced disease showed lower temperatures in general. When examining the difference between the thermographic measurements at the toe and ankle, the difference measure showed a modest positive correlation with the disease index (r = 0.34). Surprisingly, the estimate of the refill time from the nonlinear fit to the temperature curve from the challenge, t_{TH}, showed no significant correlation with the disease seriousness index and only a modest positive correlation with the static temperature measurement made at the ankle (r = 0.39). Table 1 summarizes the results from the correlation analysis.

Table 1. Correlation analysis examining the association between DSI – disease seriousness index, t_{TH} – thermographic refill time, value of this variable is obtained as a result of mathematical evaluation (estimation of refill time from nonlinear fitting curve) of thermographic assessment of the PAOD disease, T_{TOE} - temperature of the big toe prior to challenge, T_{ANK} - temperature of the ankle prior to challenge, $T_{ANK} – T_{TOE}$ - difference between the temperature of the ankle and temperature of the big toe prior to challenge.

	DSI	**t_{TH}**	**T_{TOE}**	**T_{ANK}**	**T_{ANK}-T_{TOE}**
DSI	1	-0.06	-0.33	-0.19	0.34
t_{TH}	-0.06	1	0.16	0.39	0.12
T_{TOE}	-0.33	0.16	1	0.80	-0.83
T_{ANK}	-0.19	0.39	0.80	1	-0.33
T_{ANK}-T_{TOE}	0.34	0.12	-0.83	-0.33	1

4. Discussion

The results from this limited clinical study indicate that infrared thermographic measurements can be made at the time of the physical examination of patients suspected of have lower limb PAOD. Static thermographic measurement of ankle and toe temperatures under controlled conditions provide additional information on the state of lower limb perfusion in the patient. These thermographic measurements and their difference showed a modest correlation with the severity of the disease. Thermographic measurements during a pressure cuff challenge and the vascular refill time based on the temperature profiles did not show a correlation with disease seriousness. These preliminary results cast doubt on the value of a pressure cuff challenge to be used in conjunction with infrared thermography in the assessment of PAOD.

Acknowledgements

The authors would like to acknowledge support from the National Research Council Canada, Genomics and Health Initiative.

References

[1] Krupski WC, Skinner HB, and Effeney DJ. Amputation, in Current Surgical Diagnosis and Treatment. Way LW, Editor. Appleton and Lange, San Mateo, CA, 1988.

[2] European Working Group on Critical Leg ischemia. Second European Consensus Document on Critical Leg ischemia. *European Journal of Vascular and Endovascular Surgery*, 6 (Supp A), 1992.

[3] Jacobs MJHM, Ubbink DTh, Kitslaar PJEHM, et al. Assessment of the microcirculation provides additional information in critical limb ischemia. *European Journal of Vascular and Endovascular Surgery*, 6: 135-141, 1992.

[4] Rutherford RB, Baker JD, Ernst C, et al. Recommended standards for reports dealing with lower extremity ischemia: Revised version. Journal of Vascular Surgery, 26: 517-538, 1997.

[5] Raines JK, Darling RG, Buth J et al. Vascular laboratory criteria for the management of peripheral vascular disease of the lower extremities. *Surgery,* 79: 21–29, 1976.

[6] Ray SA, Srodon PD, Taylor RS, Dormandy JA. Reliability of ankle:brachial pressure index measurement by junior doctors. *British Journal of Surgery*, 81:188-190, 1994.

Regular Reflectance and Transmittance Measurements of Transmissive Materials Using a STAR GEM® Optical Accessory

[1,3]E.Kawate, [1,2]M.Hain

[1]AIST, 1-1-1, Central 2, Umezono, Tsukuba, Ibaraki 305-8568, Japan
[2] Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovakia
[3] TRAS, Inc., 725-28 Shimohirooka, Tsukuba, Ibaraki 305-0042, Japan
E-mail address: e.kawate@aist.go.jp

Abstract. *So far, five methods for the measurement of absolute reflectance have been developed. All these methods are analyzed and compared with particular emphasis on application to transmissive materials, such as windows and filters. Four of the methods are commonly used: V-W, V-N, goniometer, and integrating sphere method; the fifth, new developed method, is a STAR GEM method which is expected to achieve the greatest accuracy in the measurement of regular reflectance and transmittance. This is because beams, spatially shifted due to multi-reflections inside the material, can be directed to a detector by automatic adjusting of an output mirror angle. The absorptance calculated from reflectance and transmittance measured by the STAR GEM is compared with the absolute absorptance measured by a laser calorimeter. Measurement results from these two methods agree well each other.*

Keywords: STAR GEM, Symmetry X, absolute reflectance, laser calorimeter

1. Introduction

Light incident on a material can interact with it in five main ways: it can be absorbed, transmitted either regularly or diffusely, and reflected either regularly or diffusely. In this paper we will consider that a sample under test is specular, and in the other paper we will discuss the measurements of diffuse transmittance and reflectance [1]. Most commonly the transmittance (T) and reflectance (R) are measured by a spectrophotometer. The absorptance (A) can be then calculated from equation $A=1-T-R$. On the other hand, the absolute absorptance can be directly measured by a laser calorimeter.

Although the techniques of the absolute reflectance measurement are discussed in many reference books [2], they do not pay so much attention to the regular transmittance measurements. All spectrophotometers and contemporary accessories on the market have been developed according to the idea that it is possible to separately measure transmittance and reflectance. However, when the absorptance is calculated from $1-T-R$ and optical constants of a sample are evaluated from solutions of simultaneous equations for refractive index and extinction coefficient using measured values of T and R, the accuracy of the results becomes worse. A few researchers discussed the measurement methods of T and R with the equal accuracy. [3, 4]

Up to now, four methods for the measurement of absolute reflectance have been developed: V-W, V-N, a goniometer and an integrating sphere methods. For a transmissive sample the additional beams, coming from the back surface reflections, are displaced with respect to the beam reflected from the front surface. Displacement depends on the incident angle, on the refractive index of the sample, and the sample thickness. So far, in the V-W, V-N and the goniometer accessories no solution employing a fine adjustment of the mirror angle in front of a detector was introduced. Usually only an averaging sphere is used to reduce sensitivity to alignment, but the spatial nonuniformity of the response of the integrating sphere makes the solution of the displacement more complicated.

2. Description of a STAR GEM

The fifth method is a STAR GEM. The STAR GEM is an acronym of Scatter, Transmission, and Absolute Reflection measurements using a Geminated Ellipsoid Mirror. The STAR GEM has been developed on the basis of two ideas in AIST. The first idea is the Symmetry X to measure the absolute reflectance [4]. For the transmittance measurement, an optical path of a transmission overlaps that of a background and the ratio gives us the absolute transmittance. The Symmetry X was inferred from the transmittance measurement. We make two reflection measurements at the same angle of incidence from a front side (RFF) and a back side of the sample (RBB) and we also make two background measurements at the same angle from a front side (BFB) and a back side (BBF), so these four paths overlap everywhere. The shape of the four paths is like a "X" over the other "X". A geometric mean of two reflectance values gives us the absolute reflectance.

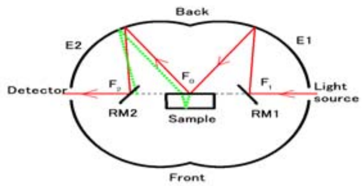


Fig.1 Structure of GEM and beams

The second idea is a Geminated Ellipsoid Mirror (GEM). The GEM consists of two equivalent ellipsoids of revolution (E1 and E2). A focus of the E1 coincides with that of the E2 and this common focus (F_0) aligns with two remaining foci (F_1 and F_2) in Fig. 1. A sample is placed at F_0 and two rotating plane mirrors (RM1 and RM2) are placed at F_1 and F_2, respectively. The RM1 and RM2 can rotate by each stepping motor independently. The STAR GEM Type 1 is shown in Fig.2 and its size is 198mm×166mm×120mm. A sample holder has two equivalent holes and a sample is attached to the one hole. Six modes in Fig. 3 are carried out for measurements of both absolute reflectance and transmittance according to the Symmetry X. As a result, reflectance and transmittance can be measured with the equal accuracy. An incident angle to a sample can be changed continuously from 0 to 90 degrees. The STAR GEM can be used not only with an outside light source and a detector, but also with an installation into a sample compartment of a commercial spectrophotometer without any change of the spectrophotometer itself.

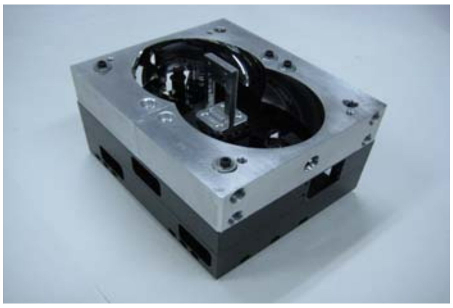


Fig. 2 STAR GEM Type 1

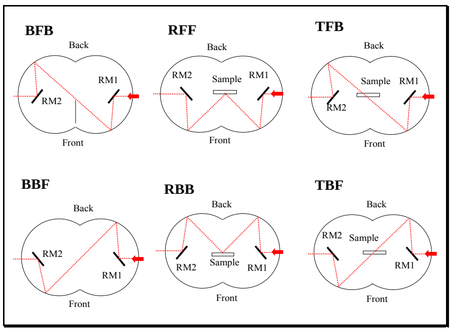


Fig. 3 Six modes for absolute reflectance and transmittance measurements

3. Experimental details

The STAR GEM optical device was inserted inside a sample compartment of the Varian FTIR spectrophotometer (FTS 7000e). The FTS 7000e was configured with a ceramic lamp, a Ge/KBr beam splitter and a Deuterated Tri-Glycine Sulfate (DTGS) detector for the spectral measurements in the region from 1.6 μm to 25 μm. The IR beam was focused onto the RM1 mirror. STAR GEM with an additional lens was also inserted inside a sample compartment of a Shimadzu grating spectrophotometer (SolidSpec 3700). The SolidSpec 3700 was configured with two lamps: tungsten, and deuterium D2 and with three detectors: a photomultiplier (PMT), an InGaAs detector, and a cooled PbS detector. An averaging sphere was used in front of the detectors for the spectral measurements in the region from 0.24 μm to 2.6 μm. The lens, which focused the original beam from the SolidSpec 3700 onto the RM1, was settled near the entrance port of the STAR GEM.

In order to directly measure absolute absorptance, we prepared and used a laser calorimeter [5]. An addenda was made of aluminum foil. A platinum thin film resistor as a thermometer and an electric resistor as a heater were folded in the addenda and a sample was attached to the addenda with a small amount of grease to increase thermal contact. The addenda was hung using thin nylon thread inside a vacuum chamber with an entrance and an exit windows for a laser beam. The laser was a diode-pumped frequency-doubled, Nd:YAG laser that emitted an output beam at a 532 nm wavelength. The maximal output power was 100 mW. At first, the sample was irradiated by laser beam for more than ten minutes and the temperature increment induced by light absorption was measured. It was very important to take care of the addenda, which must not be directly irradiated by laser beam and also must not be irradiated by scattered laser light from the chamber's windows and the sample. Secondly, the temperature increment induced by Joule heating was measured. The absolute absorptance was obtained from the two temperature increments and the measurement of laser power at the sample.

Spectral and calorimetric measurements for the same fused quartz disk were carried out. Its size was 10 mm diameter and 1 mm thickness.

4. Results and Discussions

For a transmissive sample, the first beam is reflected from an incident surface and focuses at F_2 (red solid line in Fig. 1). The additional beams coming from the back surface reflections and focus near F_2 (green dotted line in Fig. 1). The angle of the RM2 mirror, which directs all beams reflected from the sample under test to the detector, depends on the incident angle, the refractive index of the sample and the sample thickness. The calculation result of this angle difference is shown in Fig. 4 (the thickness is 1 mm and the index is 1.457). This result means that the Symmetry X is not sufficient for the transmissive sample. The suitable measurement procedure should always search a rotation angle (ϕ_0) of RM2 for the maximum signal intensity at each mode in Fig. 3. Angle of the RM2 mirror is rotated by 0.1 degree between ϕ–10 and ϕ+10 degrees, where ϕ is the expected angle from the Symmetry X, and then the spectral measurements at ϕ_0 are done.

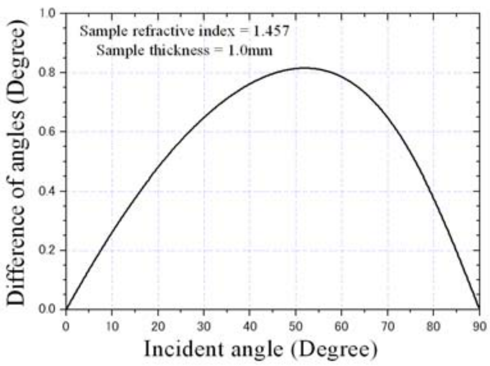


Fig. 4 Difference of RM2 rotation angles for reflection beams from a front and a back surfaces

The automatic measurement procedure using the STAR GEM is as follows: (1) the sample holder is rotated by 180 degrees, (2) the background measurement of a BFB mode is carried out for an empty hole, (3) the sample holder is raised up, (4) the sample measurements of a TFB and an RFF modes are carried out for a sample, (5) the sample holder is rotated by 180 degrees, (6) the sample measurements of an RBB and a TBF modes are carried out for the sample, (7) the sample holder is held down, (8) the background measurement of a BBF mode is carried out for the empty hole. The maximum signal is found by the adjustment of the RM2 rotation angle for each measurement.

Figure 5(a), (b) and (c) are the measured spectra of reflectance and transmittance and the calculated spectrum of absorptance. The abscissa is a logarithmic scale. Green and red curves are spectra measured by the grating and the FTIR spectrophotometers, respectively, and black solid circles are spectral values calculated using the optical constants from the handbook [6]. In these measurements the incident angle was 10 degrees. The difference between the optimized RM2 angle and the RM2 angle expected from the Symmetry X was 0.1 degree. This difference is about half of the value estimated from Fig. 4 and it is reasonable. When the spectrophotometers were changed, each spectrum measured using the STAR GEM agrees well with the corresponding spectrum in the overlapping wavelength region and the difference is less than 0.1 %. Each measured spectrum also agrees well with the calculated spectrum. In the visible region from 0.3 µm to 1.38 µm the absorptance spectrum is less than ±0.05 % and the noise level of the measured spectrum is estimated to be ±0.2 %, so that the sample didn't absorb light. The transmittance and reflectance measured using the STAR GEM are independent each other and have the same accuracy of the measurement.

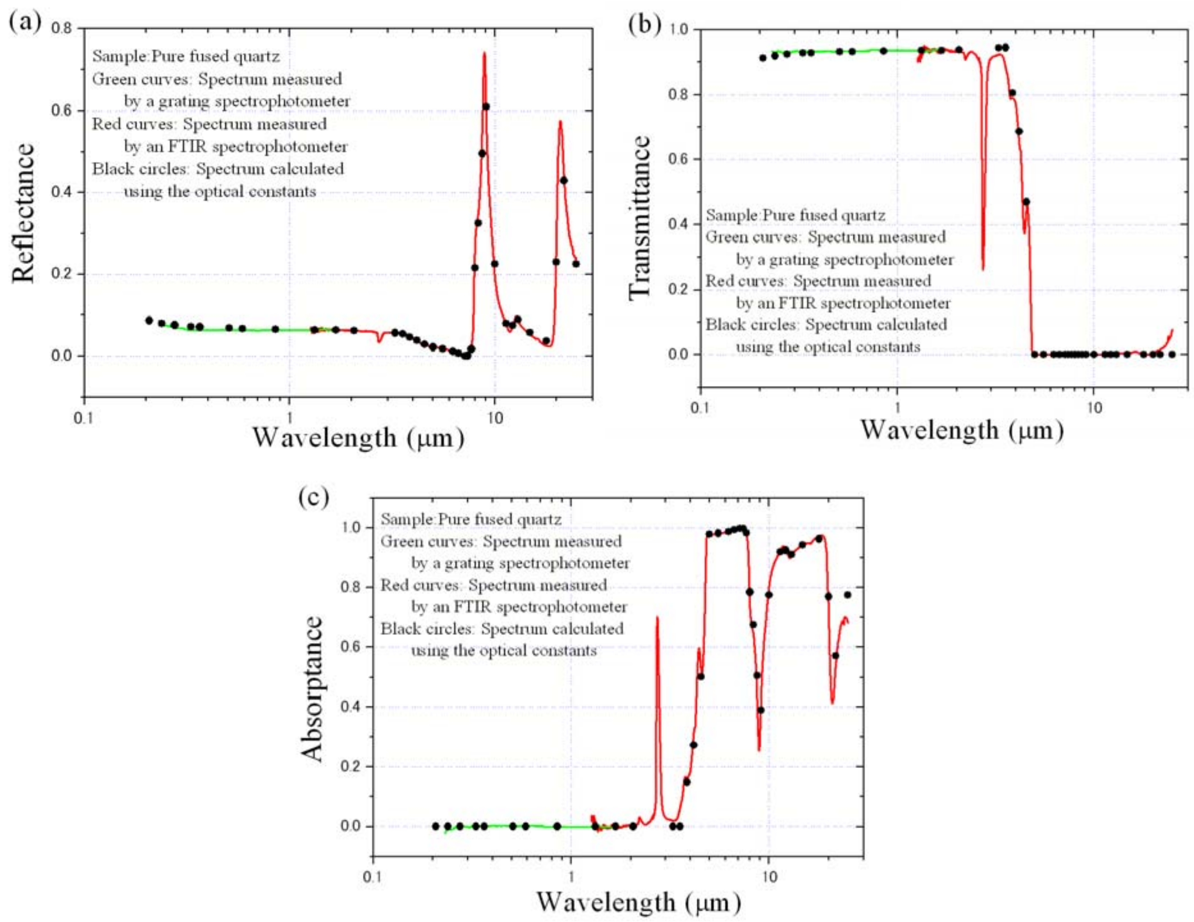


Fig. 5(a), (b) and (c) Reflectance, transmittance and absorptance spectra of a pure fused quartz

The temperature increments of the same fused quartz measured by the laser calorimeter are shown in Fig. 6. Two black curves in Fig. 6(a) are temperature variations without laser irradiation measured two times and two red curves are temperature increments measured with laser irradiation of 29.5 mW power. Figure 6(b) is temperature increments induced by Joule heating from 0.04 mW to 0.25 mW. The noise level of the laser calorimeter is estimated to be

±0.03 %. The absolute absorptance of 1mm thick fused quartz at 532 nm measured by the laser calorimeter was 0.00±0.000.

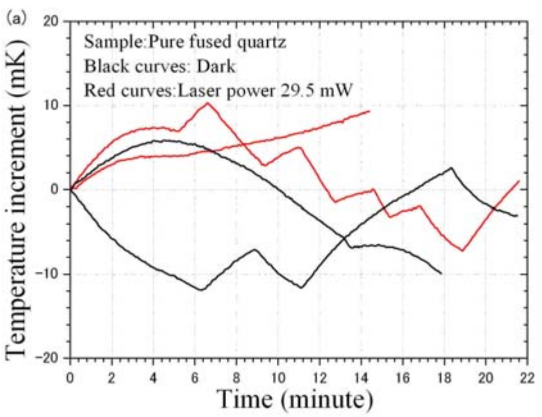


Fig. 6(a) Black curves are temperature variations at dark and red curves are temperature increment by laser irradiation.

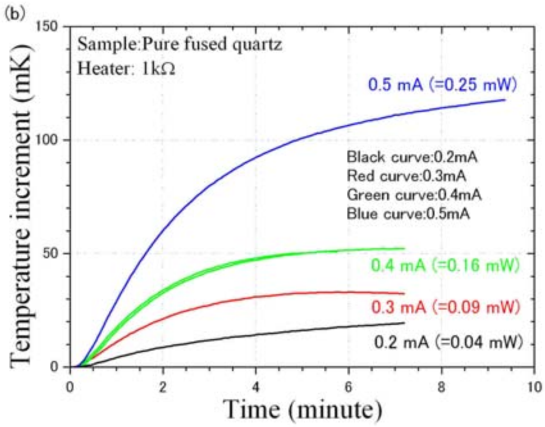


Fig. 6(b) Temperature increment by Joule heating depending on the current.

5. Conclusions

The Symmetry X is an innovative idea which was employed in the STAR GEM optical device for the absolute reflectance and transmittance measurement. In order to evaluate the accuracy of the STAR GEM, the absorptance calculated from reflectance and transmittance measured by the STAR GEM was compared with the absolute absorptance measured by the laser calorimeter at the wavelength 532 nm. Both absorptance measurements were carried out for the same sample made from fused quartz and results show good agreement. As it was analyzed above, measurements by means of STAR GEM optical device for transmissive materials need to find the maximum signal intensity. This procedure was fully automated in the STAR GEM optical device and therefore it can be used as a new innovative accessory of a FTIR or grating spectrophotometer for fully automated and precise optical measurement of absolute reflectance and transmittance of materials.

References

[1] Kawate E. Diffuse reflectance and transmittance measurements using a STAR GEM® optical accessory. To be submitted to the 7th International Conference of the Measurement 2009.

[2] Workman J. and Springsteen A.W., APPLIED SPECTROSCOPY, ACCADEMIC PRESS, San Diego, 1997.

[3] Hansen L. Integrating-sphere system and method for absolute measurement of transmittance, reflectance, and absorptance of specular samples. *Applied Optics*, 40, 3196-3204, 2001.

[4] Kawate E. Symmetry X system and method for absolute measurements of reflectance and transmittance of specular samples. *Applied Optics*, 42, 5064-5072, 2003

[5] Kawate E, Hanssen L, Kaplan S, Datla R, Development of an adiabatic laser calorimeter. In Proceedings of SPIE, vol.4103, 2000, 133-144.

[6] Palik E.O. Handbook of Optical Constants of Solids. Academic, San Diego, 1985

Magnetic Field of Saddle-shaped Coil

[1]P. Andris, [2]J. Weis, [1]I. Frollo

[1]Institute of Measurement Science, Slovak Academy of Sciences,
Dubravska cesta 9, Sk-841 04 Bratislava, Slovak Republic
[2]Department of Radiology and Department of Medical Informatics and Bioengineering, Uppsala University Hospital, SE-751 85 Uppsala, Sweden
Email: peter.andris@savba.sk , jan.weis@radiol.uu.se

Abstract. *Saddle-shaped coil is one of several basic RF coils used in magnetic resonance instruments. Magnetic field of saddle-shaped coil was calculated using the Biot-Savart law and numerically computed. Magnetic field of a sample coil was measured using magnetic resonance method. To verify the calculation correctness, the calculated and measured values were compared showing excellent results.*

Keywords: magnetic field, magnetic resonance, saddle-shaped coil, NMR

1. Introduction

RF coil is an important part of Nuclear Magnetic Resonance (NMR) tomograph. It must excite protons in a sample and/or convert RF magnetic field from the sample into electrical signal. Several parameters of the coil must be considered during a design. One of them and very significant is magnetic field of the coil and its homogeneity. Homogeneity or sensitivity of the coil can be increased by optimisation. Knowledge of the magnetic field is inevitable in the both cases. Magnetic field can be calculated in more manners. Calculation using the Biot-Savart law is rather frequent because it can be applied on a coil with an arbitrary shape. The saddle-shaped coil was very frequently used RF coil in the beginning of the NMR. Nowadays it was replaced by some new kinds of coils in the machines with high magnetic field but it is still important for imaging at low magnetic field. The purpose of the article is to present equations for computing the magnetic field of saddle-shaped coil based on the Biot-Savart law. Correctness of the equations was verified by NMR method described in [1].

2. Subject and Methods

Magnetic field of a one-turn saddle-shaped coil was calculated using the Biot-Savart law (see Fig. 1). The calculation was made in a vector form, that's why the magnetic field for a more-turn coil can be derived very simply adding the magnetic field from the next turns. Calculations in vector form were performed very successfully using the program package Mathematica (Wolfram Research Inc., Champaign, IL). The computed magnetic field was compared with the magnetic field measured on a sample of the saddle-shaped coil using the NMR method [1]. Measurement was performed on 1.5 T NMR system (Gyroscan NT, Philips, Best, the Netherlands).

3. Results

A one-turn saddle-shaped coil is depicted in Fig. 1. The application of the Biot-Savart law [2] yielded the following equations:

$\mathbf{B} = \dfrac{\mu_o I}{4\pi} \oint \dfrac{d\mathbf{s} \times (\mathbf{P} - \mathbf{s})}{|\mathbf{P} - \mathbf{s}|^3}$ is magnetic field calculated using the Biot-Savart law, where

I is current through the coil,
P is vector pointing to the observer and
s is vector pointing to centerline element of the coil conductor d**s**.

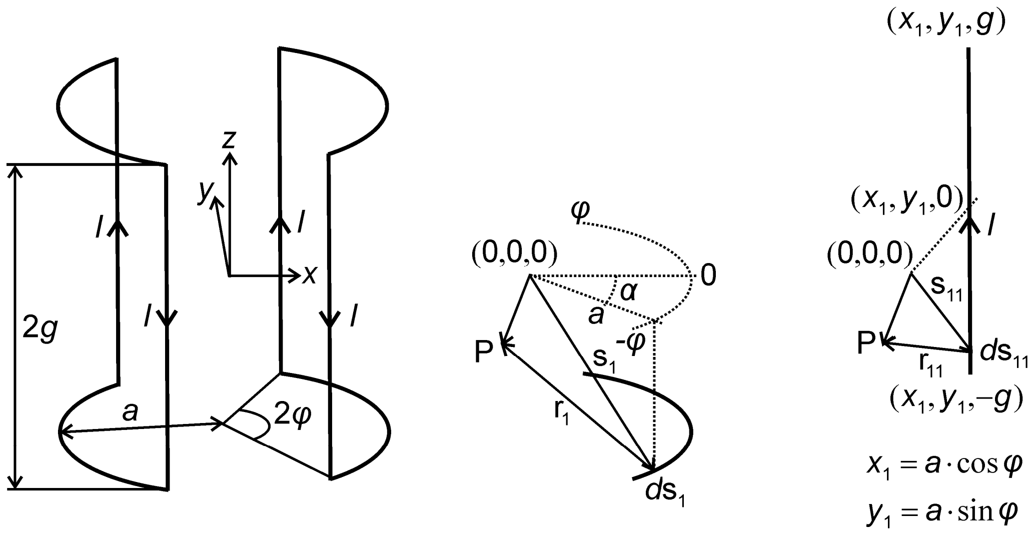


Fig. 1 One-turn saddle-shaped coil with details for the first arc and vertical.

Magnetic field of a saddle-shaped coil arcs:

$\mathbf{P} = \mathbf{i}p_x + \mathbf{j}p_y + \mathbf{k}p_z$ is point of the observer in which the magnetic field is calculated.

$\mathbf{i}, \mathbf{j}, \mathbf{k}$ are unity vectors in the directions of the x, y, z axis,

$\mathbf{s}_1 = \mathbf{i}a\cos\alpha + \mathbf{j}a\sin\alpha - \mathbf{k}g$,

$\mathbf{s}_2 = \mathbf{i}a\cos\beta + \mathbf{j}a\sin\beta - \mathbf{k}g$,

$\mathbf{s}_3 = \mathbf{i}a\cos\gamma + \mathbf{j}a\sin\gamma + \mathbf{k}g$ and

$\mathbf{s}_4 = \mathbf{i}a\cos\xi + \mathbf{j}a\sin\xi + \mathbf{k}g$ are points of the arcs.

$d\mathbf{s}_1 = (-\mathbf{i}a\sin\alpha + \mathbf{j}a\cos\alpha)d\alpha$,

$d\mathbf{s}_2 = (-\mathbf{i}a\sin\beta + \mathbf{j}a\cos\beta)d\beta$,

$d\mathbf{s}_3 = (-\mathbf{i}a\sin\gamma + \mathbf{j}a\cos\gamma)d\gamma$ and

$d\mathbf{s}_4 = (-\mathbf{i}a\sin\xi + \mathbf{j}a\cos\xi)d\xi$ are vector elements of the arcs.

$\mathbf{r}_1 = \mathbf{P} - \mathbf{s}_1$,

$\mathbf{r}_2 = \mathbf{P} - \mathbf{s}_2$,

$\mathbf{r}_3 = \mathbf{P} - \mathbf{s}_3$ and

$\mathbf{r}_4 = \mathbf{P} - \mathbf{s}_4$ are vectors between the observer and the arcs.

$\mathbf{B}_{arc} = \left(I \cdot 10^{-7}\right) \cdot \left(\int_{-\varphi}^{\varphi} \frac{d\mathbf{s}_1 \times \mathbf{r}_1}{|\mathbf{r}_1|^3} + \int_{\pi+\varphi}^{\pi-\varphi} \frac{d\mathbf{s}_2 \times \mathbf{r}_2}{|\mathbf{r}_2|^3} + \int_{\varphi}^{-\varphi} \frac{d\mathbf{s}_3 \times \mathbf{r}_3}{|\mathbf{r}_3|^3} + \int_{\pi-\varphi}^{\pi+\varphi} \frac{d\mathbf{s}_4 \times \mathbf{r}_4}{|\mathbf{r}_4|^3} \right)$ is the magnetic field from the arcs of the one-turn saddle-shaped coil.

Magnetic field of a saddle-shaped coil verticals:

$\mathbf{s}_{11} = \mathbf{i}a \cos \varphi + \mathbf{j}a \sin \varphi + \mathbf{k}z$,

$\mathbf{s}_{22} = \mathbf{i}a \cos (-\varphi) + \mathbf{j}a \sin (-\varphi) + \mathbf{k}z$,

$\mathbf{s}_{33} = \mathbf{i}a \cos (\pi - \varphi) + \mathbf{j}a \sin (\pi - \varphi) + \mathbf{k}z$ and

$\mathbf{s}_{44} = \mathbf{i}a \cos (\pi + \varphi) + \mathbf{j}a \sin (\pi + \varphi) + \mathbf{k}z$ are points of the verticals.

$d\mathbf{s}_{11} = d\mathbf{s}_{22} = d\mathbf{s}_{33} = d\mathbf{s}_{44} = \mathbf{k}dz$ are the vector elements of the verticals.

$\mathbf{r}_{11} = \mathbf{P} - \mathbf{s}_{11}$,

$\mathbf{r}_{22} = \mathbf{P} - \mathbf{s}_{22}$,

$\mathbf{r}_{33} = \mathbf{P} - \mathbf{s}_{33}$ and

$\mathbf{r}_{44} = \mathbf{P} - \mathbf{s}_{44}$ are vectors between the observer and the verticals.

$\mathbf{B}_{vert} = \left(I \cdot 10^{-7}\right) \cdot \left(\int_{-g}^{g} \frac{d\mathbf{s}_{11} \times \mathbf{r}_{11}}{|\mathbf{r}_{11}|^3} + \int_{g}^{-g} \frac{d\mathbf{s}_{22} \times \mathbf{r}_{22}}{|\mathbf{r}_{22}|^3} + \int_{-g}^{g} \frac{d\mathbf{s}_{33} \times \mathbf{r}_{33}}{|\mathbf{r}_{33}|^3} + \int_{g}^{-g} \frac{d\mathbf{s}_{44} \times \mathbf{r}_{44}}{|\mathbf{r}_{44}|^3} \right)$ is the magnetic field from the verticals of the saddle-shaped coil.

$\mathbf{B} = \mathbf{B}_{arc} + \mathbf{B}_{vert}$ is the resulting magnetic field of the saddle-shaped coil in the point **P**.

The NMR method used for verification is described in [1]. The magnetic field was measured on the four-turn saddle-shaped coil with the following parameters:

$\varphi_1 = 1.46\,\text{rad}$, $\varphi_2 = 1.25\,\text{rad}$, $\varphi_3 = 1.0\,\text{rad}$, $\varphi_4 = 0.66\,\text{rad}$, $a = 0.038\,\text{m}$, $g_1 = 0.05\,\text{m}$, $g_2 = 0.045\,\text{m}$, $g_3 = 0.04\,\text{m}$, $g_4 = 0.035\,\text{m}$, I=50 mA.

The magnetic field was calculated and measured in the three basic planes, Fig. 2 depicts the both calculated and measured values on centrelines in the planes *xy* and *yz*. It is evident very good agreement between the calculated and the measured values.

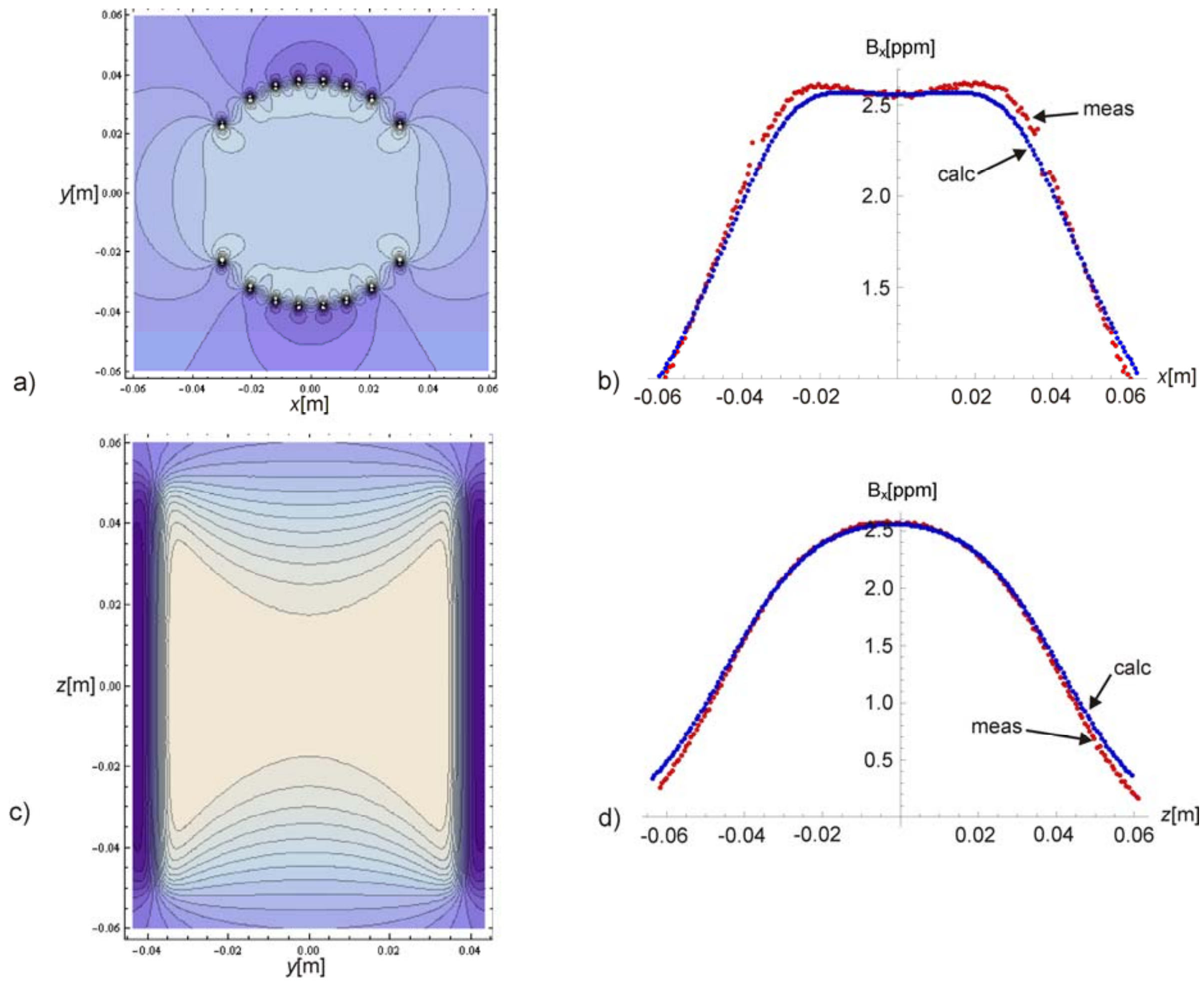


Fig. 2 Comparison between the calculated and the measured magnetic field of the four-turn saddle-shaped coil. The calculated magnetic field in the central plane *xy* a); the measured and calculated values at the centreline in the *xy* plane b); the calculated magnetic field in the central plane *yz* c); the measured and calculated values at the centreline in the *yz* plane d).

4. Discussion and Conclusions

Experiments confirmed correctness of the calculated values. A problem can be the numerical integration; it can spend much time and complicate a subsequent optimisation in such way. The presented equations can be simple used for the magnetic field of more-turn saddle-shaped coils calculation – the magnetic field of the next turns is simply added to the magnetic field of the first turn.

Acknowledgements

Support of the Grant agency of the Slovak Academy of Sciences, grant no 2/0142/08 is gratefully acknowledged.

References

[1] Weis J., Andris P., Frollo I., Ahlström H.: A simple method for mapping the B_1 distribution of linear RF coil. *MAGMA*, 2005, 18: 283-287.

[2] Ilkovic D.: Physics. ALFA, Bratislava, 1972. (in Slovak)

AUTHORS INDEX