

Research on Skeleton Data Compensation of Gymnastics based on Dynamic and Static Two-dimensional Regression using Kinect

Gang Zhao¹, Hui Zan^{1,2*}, Junhong Chen³

¹ Faculty of Artificial Intelligence in Education, Central China Normal University, Wuhan, 430079, China.

² Key Laboratory of Intelligent Education Technology and Application of Zhejiang Province, Zhejiang Normal University, Jinhua, 321004, China.

³ College of Mathematics and Computer Science, Zhejiang Normal University, Jinhua, 321004, China.

*Corresponding author: zanhui@zjnu.cn, zhaogang@ccnu.edu.cn

Abstract: The intelligent training and assessment of gymnastics movements require studying motion trajectory and reconstructing the character animation. Microsoft Kinect has been widely used due to its advantages of low price and high frame rate. However, its optical characteristics are inevitably affected by illumination and occlusion. It is necessary to reduce data noise via specific algorithms. Most of the existing research focuses on local motion but lacks consideration of the whole human skeleton. Based on the analysis of the spatial characteristics of gymnastics and the movement principle of the human body, this paper proposes a dynamic and static two-dimensional regression compensation algorithm. Firstly, the constraint characteristics of human skeleton motion were analyzed, and the maximum constraint table and Mesh Collider were established. Then, the dynamic acceleration of skeleton motion and the spatial characteristics of static limb motion were calculated based on the data of adjacent effective skeleton frames before and after the collision. Finally, using the least squares polynomial fitting to compensate and correct the lost skeleton coordinate data, it realizes the smoothness and rationality of human skeleton animation. The results of two experiments showed that the solution of the skeleton point solved the problem caused by data loss due to the Kinect optical occlusion. The data compensation time of an effective block skeleton point can reach 180 ms, with an average error of about 0.1 mm, which shows a better data compensation effect of motion data acquisition and animation reconstruction.

Keywords: Azure Kinect, motion capture, motion tracking, motion compensation.

1. INTRODUCTION

The intelligent training and evaluation system of gymnastics movements is conducive to reducing the duplication of a coach's work. Meanwhile, it is also helpful in providing guidance and accurate evaluation of athletes' learning. The core objective that this system strives to achieve is the capture and reconstruction of human movements. Human body capture devices emerge one after another from mechanical, and wearable to optical ones, but these industrial devices are expensive, either restraining actors' performance or limiting the space where actors can do gymnastics movements.

Microsoft Kinect (MK), the optical unmarked motion capture sensor device was released by Microsoft, combined with a proper SDK, provides color image data, depth image data, and skeleton data at the same time, which can convert the limb information during human movement into data that can be recognized by a computer, and then use the skeleton data to drive the three-dimensional model [1]. Unlike other motion capture devices, it can be applied to a variety of complex environments without too many restrictions,

effectively solving the problems of the long production cycle, low efficiency and low reduction in traditional animation production technology. It has made great achievements in many industries, such as film and television animation [2], physical training [2], [3], medical treatment [4]-[8], entertainment [9], [10], and education [11], [12] due to its advantages of convenient use and low price [13] since the release of Microsoft Kinect Xbox (MKv1).

In the motion reproduction experiment of gymnastic characters, it was found that if the captured data is not processed, the action will be distorted to some extent. However, even with the most popular equipment as unity for Kinect [15]-[17] and iPi Motion Capture [18]-[19] SDK that provides a depth data compensation algorithm, the human motion capture is not as smooth and reasonable as expected due to the influence illumination (ambient illumination direction, ground reflection degree) and body occlusion, as well as the angle and distance between the equipment and the captured object. Therefore, in actual application scenarios, the problem of data loss and error must be dealt with. Based on the above analysis, this paper proposes a dynamic and

static two-dimensional regression compensation algorithm (DSRA) solution according to the characteristics of gymnastics and human skeletons. The core idea includes the following three points:

1. Static frame space angle feature: The spatial position of occluded or lost skeleton points is estimated by the distance and angle characteristics of adjacent frames before and after the lost frame.
2. Dynamic frame motion characteristics: The spatial position of occluded or lost skeleton points is estimated by the angular velocity and acceleration of several adjacent skeleton points before and after the lost frame.
3. Least-squares polynomial regression operation: Through effectively captured coordinate data and missing point coordinate data estimated by spatial angle and motion characteristics, polynomial regression operation is carried out to obtain motion change curves to estimate skeleton point spatial coordinates in real-time.

2. RELATED WORK

A. MK development and performance analysis

Microsoft launched a new generation of Azure Kinect (MKv3) in 2019, which is the successor of MKv1 (2010) and MKv2 (2014). Various applications and research are gradually emerging based on MKv3. It helps us understand the dynamic capture sensor by comparing the performance parameters and differences among various versions. In addition, it is also helpful for researchers who still hold onto the previous version to improve the functions and parameters. The parameter characteristics are shown in Table 1., which are quoted from Microsoft's official website [19].

In addition to the detailed parameters officially released, the researchers also compared the performance stability [20] and availability [21], [22] of the three versions of MK.

The experimental results show that MKv2 is more stable than MKv1, and the average displacement error is less than 10 mm [23]. The range of motion (ROM) error is the smallest when MK rises 45 degrees in front of and tilts towards the player [24], Tölgyessy Dekan et al. proved that the Azure MKv3 surpasses its discontinued predecessors [25], [26] via well-designed experiments.

In terms of the practicality of motion capture, MK has an excellent performance in hand [27], upper limb exercise [28], [29], gait [30], [31], and even jump [6], [20]. Meanwhile, Guess confirmed that the MKv2 skeleton tracker could be used to calculate the relative skeleton kinematics of hip and knee joint angle [32], Bhateja et al. confirmed that the data provided by the MKv2 sensor has good stability in the air, glass, liquid, and other mediums [13].

B. Data compensation correction algorithm

Although each version of MK is optimized with respect to the previous version, it still cannot solve the problem of skeleton data capture loss caused by occlusion in the actual human motion capture process. Based on MKv2, researchers try to deal with the lost data through various algorithms to realize the smoothing of animation, which is the key to this research problem. At present, there are three main methods for data compensation and correction.

The first kind of data correction method is the Kalman filter or its extension method, which is used to extract the lower extremity joint angles [33] and get more accurate human skeleton data combined with the classical Set-Membership filter method [30]. Palmieri [34] also adopts a Kalman filter to make the velocity values of human arms more accurate. Similar to previous research, the Kalman filter is used to obtain the errors of the position and attitude outputs [35]. Some scholars [36] combine it with the multi-frame average method to solve the problem of image edge accuracy. Abbasi [37] also selects gradient descent and unscented Kalman filter methods to improve the cumulative error and occlusion problem successfully.

The second type of data correction method is a mathematical function. Filtering smoothing, for example, is adopted to capture the human joint position data [22], [46]. Ryselis, Petkus et al. [38] propose a data fusion algorithm using algebraic operations in vector space to analyze the dynamic characteristics of human motion during physical exercise. Lyu et al. [39] adopt the combination of a weight-based joint bilateral filter (WJBF) and depth compensation filter (DCF) for depth image restoration. Shotton et al. [42] present an error compensation method based on mathematical statistics, which matches the depth reference value with the depth measurement value to fit the quadratic curve.

The third method is to compensate for the data by combining the motion characteristics and the fitting curve. Humphrey et al. [22] propose a correction technique based on a two-point linear equation, which will analyze the captured data by regression to compensate for the inaccuracy in data collection. Polynomial curve fitting [40] is used to correct the angle error. When the corrected distance is within 8.5 m, the error is less than 20-50 mm. Li et al. [41] use curve fitting technology to compensate for the error. The experiment results show that the algorithm can quickly and effectively reduce the depth error of the Time-of-Flight (TOF) camera and that it is suitable for a real-time and high-precision large field of view 3D reconstruction.

Table 1. Comparison of MK function parameters of three versions.

	MK v1	MK v2	MK v3
Color map (Resolution/fps)	640*480/30 fps	1920*1080/30 fps	3840*2140/30 fps
Depth map (Resolution/fps)	320* 240/30 fps	512*424/30 fps	640 * 576/30 fps
Players tracked	6	6	6
Captured joints/players	20	25	32
Range of Detection	0.8~4.0 m	0.5~4.5 m	0.5~3.86 m
Horizontal Angle	57°	70 °	75 °
Vertical Angle	43°	60 °	65°

Through the above analysis, it is not difficult to find that the achievement of accurate data capture still requires building a new algorithm based on the physical motion characteristics of the captured object and mathematical model. These studies provide an important idea for this paper. The difference is that this paper will start from the two dimensions of static characteristics and dynamic attributes to regress and analyze the data.

3. METHOD

A. Research problem focus

From depth images to skeletal images, MK has to distinguish the human body from the background environment firstly and scan these regions' depth images pixel by pixel to determine to which parts of the human body they belong [42], it marks the 32 parts of the human body and obtains their spatial coordinates by utilizing deep learning. Then a skeleton topology is formed to drive the 3D model. The specific process is shown in Fig.1.

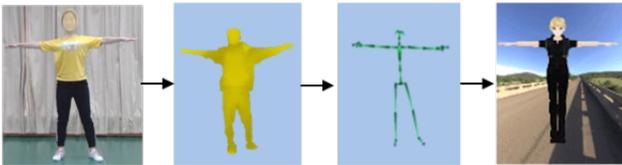


Fig.1. Schematic diagram of MK data-driven 3D model.

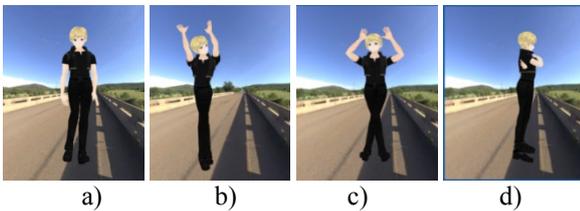


Fig.2. 3D model motion distortion caused by missing skeleton point data capture. a) Motion distortion, b) skeleton interpenetration, c) skeleton jitter, d) skeleton point loss.

In practice, if only the skeleton positioning provided by MK's SDK is used without additional adjustment and constraint, the skeleton movement will be misplaced due to ambient light and skeleton occlusion. In addition, when turning around or being blocked, it is difficult to avoid the loss of captured data. The 3D model will inevitably be

distorted. If the captured data is not processed, as shown in Fig.2., it will have a great negative impact on the actual engineering application effect of MK. Therefore, we propose a motion capture data compensation solution for skeleton occlusion.

B. Characteristic analysis of human skeleton nodes

Exploring the characteristics of skeleton data is very important for human motion imitation. However, how to extract discriminant features effectively is still challenging work. The movement of the human body is not caused by the change of the skeletons themselves but by the change of the position and direction of the bones attached to the joints. The length and shape of the bones remain unchanged during the movement, and they are connected by the joints. The topology of the human joint frame structure is shown in Fig.3.

Among the joint properties, the most important ones are the degrees of freedom and the range of possible motion of the lattice on the connected joints. The relationship between the length of two adjacent human joint points, the radius of gyration, and the centroid position are shown in Table 2., which gives the parameters of 6 joint points of the human body (height is expressed in H (m)).

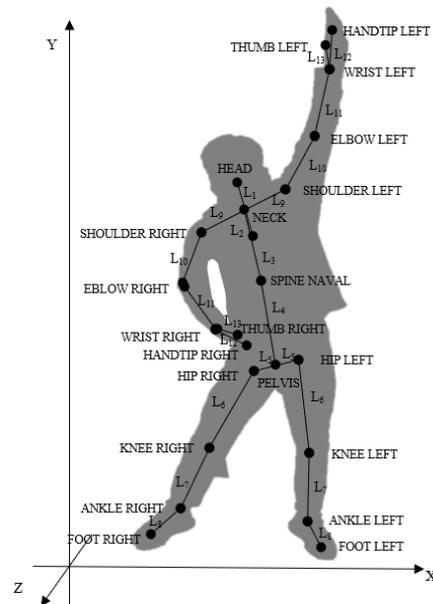


Fig.3. Topology of human joint frame structure.

Table 2. Centroid of the radius of gyration of limb joints and their relationship.

Joint	Length H(m)	Gyratory radius	Centroid position
Upper arm	$L_{torso} = 0.545H$	$0.830 L_{torso}$	$0.6600 L_{torso}$
Crus	$L_{crus} = 0.237H$	$0.528 L_{crus}$	$0.4049 L_{crus}$
Thigh	$L_{thigh} = 0.211H$	$0.540 L_{thigh}$	$0.4550 L_{thigh}$
Palm	$L_{palm} = 0.108H$	$0.578 L_{palm}$	$0.6306 L_{palm}$
Forearm	$L_{forearm} = 0.151H$	$0.526 L_{forearm}$	$0.4306 L_{forearm}$
Upper arm	$L_{upper arm} = 0.202H$	$0.542 L_{upper arm}$	$0.5506 L_{upper arm}$

According to statistics [43], [44], human skeleton nodes have a certain quantitative relationship. As expressed in formula (1), it can be used as the judgment basis for ratio constraints between different skeletons and effective skeleton data.

$$\begin{cases} L_{\text{torso}} = L_2 + L_3 + L_4 \approx 4L_1 \\ L_{\text{Upper arm}} \approx 2.2L_1 \\ L_{\text{forearm}} \approx 1.9L_1 \end{cases} \quad (1)$$

$L_{\text{Torso}}, L_{\text{Upper arm}}, L_{\text{forearm}}$ represent the length of the spine, upper arm and forearm, respectively. L_1 represents the sum of the distance between the center of the head and the neck, the same as formula (2). Before forming the human action description feature vector, it is necessary to determine whether the data is invalid. The basis of judgment is shown in formula (2).

$$\begin{cases} \frac{L_1 + L_3}{L_1} \in (4(1 - a), 4(1 + a)) \\ \frac{L_{10}}{L_1} \in (2.2(1 - a), 2.2(1 + a)) \\ \frac{L_{11}}{L_1} \in (1.9(1 - a), 1.9(1 + a)) \end{cases} \quad (2)$$

where a is the allowable error range, and its value will change according to the body shape difference of athletes; through calculation, the estimated value in the experiment is 0.1. The last two equations are the ratio of the left and right two upper arm lengths and the left and right two forearm lengths to the head length, respectively. When the conditions are satisfied, the data are valid, and the human action description feature vector is finally constructed.

C. Dynamic and static two-dimensional compensation algorithm

Before the formal human motion capture, a skeleton rotation angle constraint and a collision box are added to the whole-body skeleton. The skeleton point data will be automatically adjusted and repaired when the captured skeleton orientation exceeds the constraint limit, or a collision occurs. The core idea is to use the skeleton point data before and after the collision and make corrections by combining the movement characteristics. The basic process is shown in Fig.4.

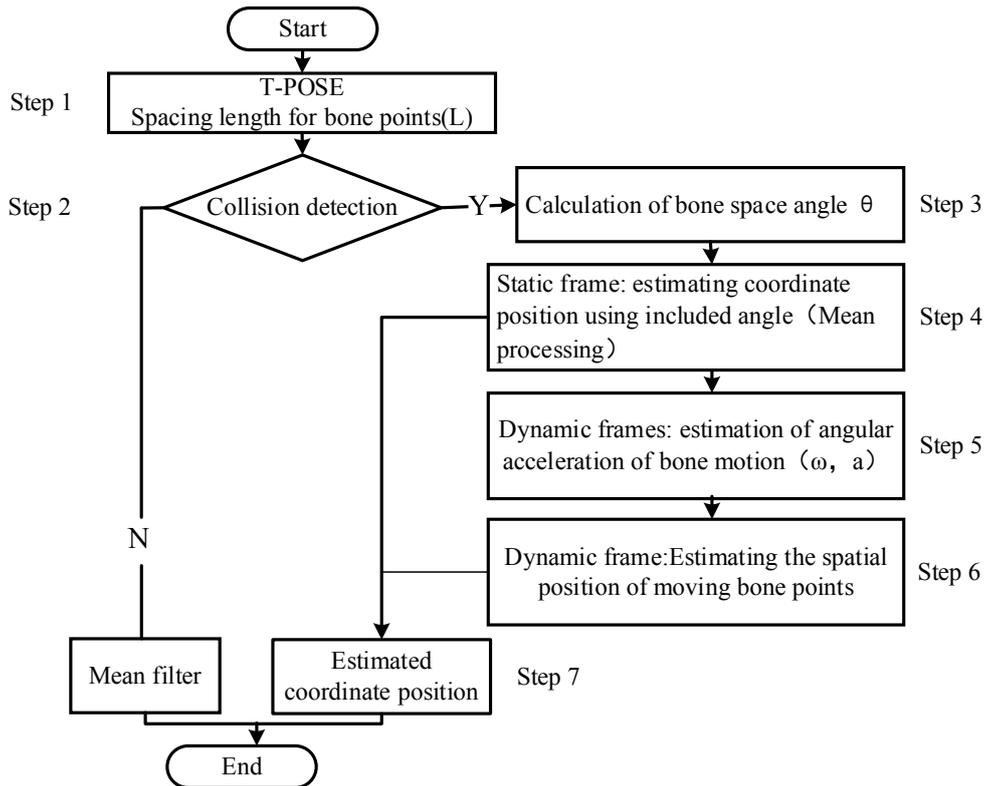


Fig.4. Flow chart of DSRA.

Step 1. According to the routine, t-pose is used to enter the data acquisition state, and Euclidean distance is used to calculate the distance between two skeleton points. The quantitative relationship between skeletons is shown in Fig.5. Euclidean distance to express the distance between them

directly: for any two points, $P_i(x_i, y_i, z_i)$ and $P_j(x_j, y_j, z_j)$, the distance can be calculated using formula (3).

$$d_{x,y,z} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \quad (3)$$

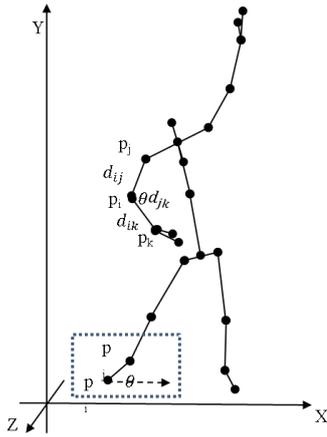


Fig.5. Schematic diagram of mathematical relationship of human skeleton nodes.

Step 2. The Mesh Collider model can be established according to the maximum constraints of human joint movements-the maximum angle between the fore and rear arm of the left hand is 120 degrees-and the critical movement characteristics of occlusion generated by MKv3. When the body movement is close to the model, as shown in Fig.6., the data compensation algorithm will be activated.

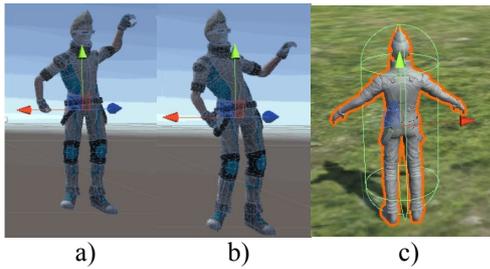


Fig.6. The Mesh Collider model a) maximum angle of 120° between the fore and rear arm of the left hand in gymnastic movements; b) the critical movement generated by the right arm; c) Mesh Collider of the human body.

There are some limitations when using the Mesh Collider. When Mesh is marked as Convex, it can collide with other Mesh Colliders. In unity 3D, the specific parameters are set to Kinematic RigidBody Collider and Kinematic RigidBody Trigger Collider.

Step 3. According to the coordinates of skeletal points in N frames before the collision, the spatial angle is calculated by the cosine theorem, and the last effective angle before and after the collision is read, as in formulas (4) and (5).

$$\theta = \arccos \frac{d_{ij}^2 + d_{ik}^2 - d_{jk}^2}{2d_{ij}d_{ik}} \quad (4)$$

$$\theta = \arccos \frac{x_j - x_i}{d_{ij}} \quad (5)$$

Step 4. Static frame spatial coordinate estimation. The static method is used when the lost node has neighboring nodes to capture, by being able to capture the last frame before and after the collision, the angle of the proximity node pinch, the capture lost skeleton node coordinate position is

calculated, and the two angle values before and after are obtained, and the average of the spatial position coordinates is achieved via formula (6).

$$\cos \theta = a * b / (|a| * |b|) \quad (6)$$

If a skeleton node is turned around, a capture loss situation occurs when the parent node is successfully captured, but the child node loses the capture. In this case, the static child node frame can be calculated directly based on the capture data of the parent node.

Step 5. Dynamic frame space motion characteristics parameter estimation. Through the collision data before the proximity of the effective capture of several frames, skeletal motion inertia characteristics parameters are calculated, and the speed of motion and angular acceleration is estimated by displacement. Formulas (7) and (8) are as follows.

$$\omega = \lim_{\Delta t \rightarrow 0} \frac{\Delta \theta}{\Delta t} = \frac{d\theta}{dt} \quad (7)$$

$$a = \lim_{\Delta t \rightarrow 0} \frac{\Delta \omega}{\Delta t} = \frac{d\omega}{dt} \quad (8)$$

Step 6. Dynamic frame space motion coordinate estimation. If the angular velocity and acceleration are obtained, the inverse of the Euclidean distance formula (9) can be performed.

$$S_{t+\Delta t} = S_t + \frac{1}{2} a \Delta t^2 \quad (9)$$

t represents a certain time and Δt represents a time interval.

Step 7. Least-squares fitting operation: according to the spatial location coordinates of the skeletal points estimated in Step 4 and Step 6, the least squares method polynomial fitting budget is carried out, and the calculation formulas (10), (11), (12) are listed as follows.

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_k x^k \quad (10)$$

$$\text{Loss} = \sum_{i=1}^n [y_i - (a_0 + a_1 x_i + a_2 x_i^2 + \dots + a_k x_i^k)]^2 \quad (11)$$

$$\frac{\partial \text{Loss}}{\partial a_k} = -2 \sum_{i=1}^n [y_i - (a_0 + a_1 x_i + a_2 x_i^2 + \dots + a_k x_i^k)] x_i^k = 0 \quad (12)$$

The polynomial parameters are calculated by deriving the parameters to obtain the complete fitting curve formula, which can be used to estimate skeletal coordinates during other motions, and a combination of dynamic methods (pre- and post-collision detection motion acceleration) and static methods (pre- and post-collision spatial pinch angles) is achieved in the fitting process.

The unmasked spatial coordinate positions can be obtained directly, and a simple mean smoothing process can be done to make the overall action smoother. Formula (13) is listed as follows.

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i \quad (13)$$

4. EXPERIMENT & DISCUSSION

We present a critical analysis, interpretation, and evaluation of the obtained results. Two groups of experiments of dimension reduction data compensation test and gymnastics movement capture, and correction experiment were carried out to verify the feasibility of the method in the MKv3 environment.

A. Experimental equipment

1. Hardware device: The experiment was conducted on a mobile computer. The operating system is Ubuntu 18.04 with NVIDIA Quadro p 4000. The memory card is configured as DDR3-1600MHZ-32GB.

Azure Kinect DK: including a 1-megapixel TOF depth camera and 12-megapixels RGB HD camera.

2. Software: Unity 3D 5.5, Azure Kinect examples for unity package, data analysis software MATLAB 2016a.

B. Experiment 1: Dimensionality reduction data compensation test

A 10-year-old elementary school student participated in the experiment to perform an arm swing test by intercepting the spatial position data of the skeletal action points every 30 ms, including the arm unfold-occlusion-unfold process, as shown in Fig.7.



Fig.7. Skeleton point occlusion test of arm swing (unfold-occlusion-unfold).

The skeletal data rotation vector is a three-dimensional vector, which is downsampled to one-dimensional values for the sake of illustration, and the data are mainly the spatial position data of the Hand Tip nodes in the Y-axis plane of the world coordinate axis, with a total of 24 data points and a time of 25 seconds.

The captured data is compensated and corrected according to the dynamic and static two-dimensional regression algorithm to obtain a set of visual diagrams of data analysis, as shown in Fig.8. The data processing is explained below.

In Fig.8.a), at seconds 8-16, the hand skeletons are obscured, and a disconnect occurs, and the goal of data processing is to connect the missing data smoothly.

Fig.8.b) directly ignores those missing data due to occlusion and performs a polynomial fitting operation using the least-squares method. Since the predicted output is estimated based on the limited valid data before and after, the predicted data may not always be accurate and can be detached from the human motion characteristics in terms of motion trends, leading to large errors.

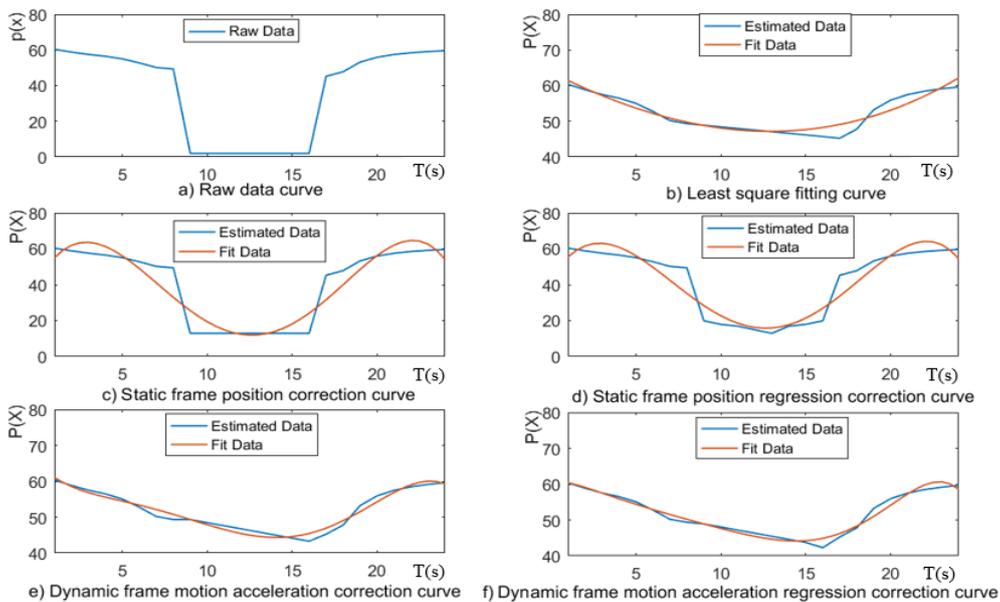


Fig.8. Visual diagram of curve comparison between skeleton point estimation data and fitting data.

Therefore, when estimating the lost skeletal data, we cannot disregard movement characteristics, or regard the estimation of the lost skeletal data as pure mathematical calculations. Especially when the motion suddenly starts or stops, pure mathematical predictions may result in a further spread or even amplification of the noise already in the data into future data, which may increase the noisiness of the data.

Fig.8.c) and Fig.8.d) show the static skeletal frame estimation and static skeletal frame regression estimation, respectively. Due to the increase in the amount of data, the trend of the fitted curve has been fundamentally transformed, and the data curve and the fitted curve behave more consistently.

Fig.8.e) and Fig.8.f) introduce motion acceleration, which is estimated by using the motion characteristics of dynamic skeletal frames, and then weighted superposition is performed based on static frames, and the estimated value and the fitted curve are the same after the regression superposition.

Fig.8.f) shows that the actual data and the fitted data perform better, and the curve is relatively smooth, which meets the data estimation requirements.

C. *Experiment 2: Gymnastics movement capture and correction experiment*

According to the characteristics of gymnastic movements, the effect simulation test was arranged for gymnastic movements with occlusion movements. The experimenter completed the turning movement and side movement, including hand masking and leg masking, and completed the data redirection experiment in a unity 3D environment according to the test data, in which smoothing processing and dynamic and static 2D regression processing for the original data mean were done, respectively. The effect of data-driven animation is shown in Fig.9.

In the effect pictures of animation redirection of the above three groups of real gymnastics movements, it can be seen that there are skeleton positions that do not conform to the characteristics of human movement. The overall effect is improved after correction, and the optimization comparison of action details is shown in Fig.10.

The comparative results showed that the motion reconstruction is approaching close to the real state after data supplement and calculation.

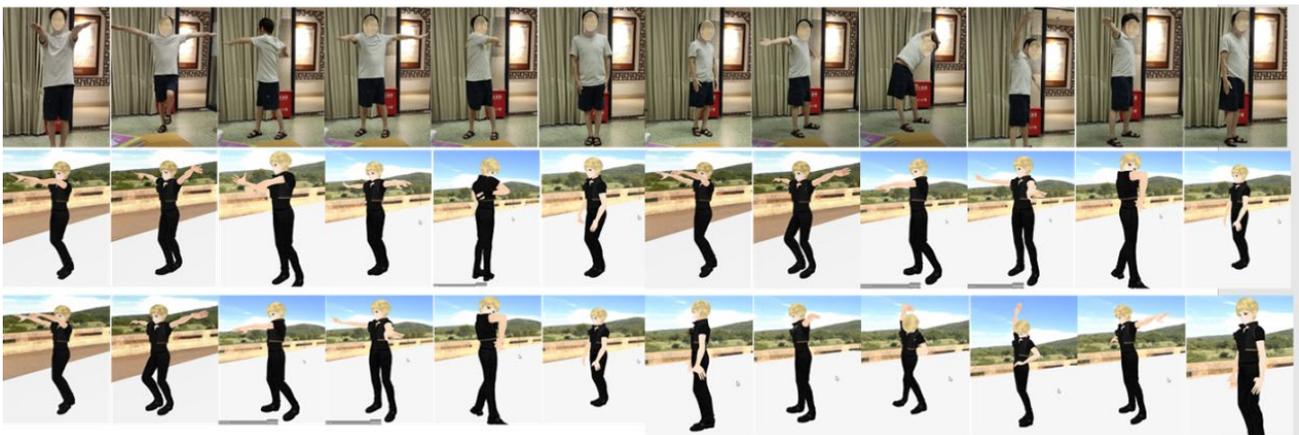


Fig.9. Comparison of skeleton data-driven animation effects (including real action (1st group of pictures), animation processed by unity for Kinect SDK (2nd group of pictures), and animation processed by DSR (3rd group of pictures)).

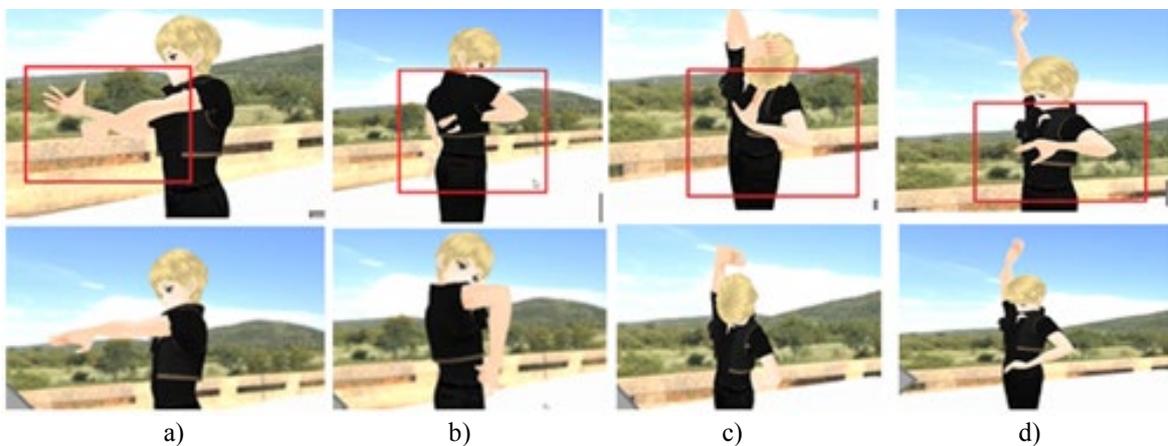


Fig.10. Details comparison of skeleton data redirection: a) skeleton interpenetration, b) side body skeleton loss, c) skeleton interpenetration from the front and from behind, d) side body skeleton loss.

Table 3. Comparison table of SDRA performance test.

	Solution	Device	Max interval	Average error	Root mean square error	Average gradient
Humphrey J et al. [23]	MF	MKv2	—	< 10 mm	—	—
Li Liangfu L et al. [36]	KF	MKv2	—	—	38.1025	0.4713
Abbasi J et al. [37]	GD&UKF	MKv2	—	0.03 mm	—	—
Xiuqi x et al. [40]	PCF	MKv3	—	0.20-0.5 mm	—	—
Zhanli L et al. [41]	CF	MKv2	90 ms	0.63 mm	—	0.7004
Ours	DSRA	MKv3	180 ms	0.45 mm	37.0200	—

It was found in the experiment that there is a lot of noise in the coordinate data of the joints obtained by Kinect, and the causes underlying these noisy data may be related to the position of the sensor array, the noise reduction mode, and the accuracy of data rounding that has been introduced in this study. Researchers analyze and optimize specific practical problems from different perspectives. For example, some researchers emphasize the compensation time while others focus more on data errors. This paper compares the compensation time and average error (the average value of the difference between the estimation points and the fitting curve), as shown in Table 3.

It is found that the dynamic and static two-dimensional regression data algorithm is better than other algorithms in many indicators. This method not only simply modifies and compensates the data but also makes full use of the data that can be accurately measured for iteration and estimation.

5. CONCLUSION

The paper proposes a DSRA method that uses MKv3 to extract human skeletal information (3D coordinates of each joint point), then connects two joint points to define skeletal vectors and skeletal point angles according to the principle of human structure. In the static aspect, the adjacent frames before and after the end of the collision are used for estimation and regression operation using vector relations; in the dynamic aspect, the angular velocity and angular acceleration are estimated by the adjacent frames before and after the collision, and the coordinates of the skeletal points are estimated by time, which are found by least-squares polynomial fitting. The method makes up for the problem of single processing of raw data in previous studies, and the experiment proves that the method effectively solves the problem of data loss during turning, provides a good method for motion capture, completes the data correction of motion capture, and is quite helpful in the field of practical use. It is worth mentioning that the algorithm proposed in the article may not necessarily be adapted to other specific environments as well, taking into account the characteristics of human motion, which also provides a wide space for our future research.

ACKNOWLEDGMENT

This study was funded by the Research on Automatic Segmentation and Recognition of Teaching Scene with the Characteristics of Teaching Behavior of National Natural Science Foundation of China (Grant number-61977034), as well as supported by open fund of Key Laboratory of

Intelligent Education Technology and Application of Zhejiang Province (Grant number-jykc20057). Zhejiang Education Science Planning Project (Grant number-2021SCG309). Special Program for Guiding Local Science and Technology Development by the Central Government, China. (Grant number-2019ZYD012).

REFERENCES

- [1] Hu, X., Yang, Z.W., Liu, X.P. (2016). A real-time algorithm of virtual animation character driving based on Kinect. *Journal of Hefei University of Technology*, 39 (6), 756-760. (in Chinese)
<http://dx.chinadoi.cn/10.3969/j.issn.1003-5060.2016.06.008>
- [2] Han, L., Zhang, M.C. (2017). Research on the teaching design and experiment of sports micro course based on the fusion of motion capture technology. *Journal of Liaoning Normal University (Natural Science Edition)*, 40 (2), 199-206. (in Chinese)
https://oversea.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&filename=LNSZ201702010&dbname=CJFD_LAST2017
- [3] Eltoukhy, M., Kuenze, C., Oh, J., Wooten, S., Signorile, J. (2017). Kinect-based assessment of lower limb kinematics and dynamic postural control during the star excursion balance test. *Gait & Posture*, 58, 421-427.
<https://doi.org/10.1016/j.gaitpost.2017.09.010>
- [4] Yang, W.L., Li, X.R., Xia, B. (2018). System of swing-arm therapy auxiliary training based on unity 3D and Kinect. *Modern Computer*, (20), 79-84. (in Chinese)
<http://www.cnki.com.cn/Article/CJFDTotal-XDJS201820018.htm>
- [5] Xing, M.M., Wei, G.H., Liu, J., Zhang, J.Z., Yang, F., Hui, C. (2020). A review on multi-modal human motion representation recognition and its application in orthopedic rehabilitation training. *Journal of Biomedical Engineering*, 37 (1), 174-178, 184. (in Chinese)
<http://dx.doi.org/10.7507/1001-5515.201906053>
- [6] Gray, A.D., Willis, B.W., Skubic, M. (2017). Development and validation of a portable and inexpensive tool to measure the drop vertical Jump using the Microsoft Kinect V2. *Sports Health: A Multidisciplinary Approach*, 9 (6), 537-544.
<https://doi.org/10.1177%2F1941738117726323>
- [7] Gambi, E., Agostinelli, A., Belli, A., Burattini, L., Cippitelli, E., Fioretti, S., Pierleoni, P., Ricciuti, M., Sbröllini, A., Spinsante, S. (2017). Heart rate detection

- using Microsoft Kinect: Validation and comparison to wearable devices. *Sensors*, 17 (8), 1776. <https://doi.org/10.3390/s17081776>
- [8] Eltoukhy, M., Oh, J., Kuenze, C., Signorile, J. (2017). Improved Kinect-based spatiotemporal and kinematic treadmill gait assessment. *Gait & Posture*, 51, 77-83. <https://doi.org/10.1016/j.gaitpost.2016.10.001>
- [9] Chang, X.J., Ma, Z.G., Li, M., Hauptmann, A.G. (2017). Feature interaction augmented sparse learning for fast Kinect motion detection. *IEEE Transactions on Image Processing*, 26 (8), 3911-3920. <https://doi.org/10.1109/TIP.2017.2708506>
- [10] Yang, Y.H., Xu, W., Zhang, H., Zhang, J.P., Xu, M.L. (2014). The application of KINECT motion sensing technology in game-oriented study. *International Journal of Emerging Technologies in Learning*, 9 (2). <https://doi.org/10.3991/ijet.v9i2.3282>
- [11] Wu, X., Zhang, Y., Shen, Y., Yan, X. (2013). National dance 3D digitizing protection method based on motion capture technology. *Computer and Modernization*, (1), 112-114, 118. (in Chinese) <http://dx.chinadoi.cn/10.3969/j.issn.1006-2475.2013.01.032>
- [12] Kataoka, H., Satoh, Y., Aoki, Y., Oikawa, S., Matsui, Y. (2018). Temporal and fine-grained pedestrian action recognition on driving recorder database. *Sensors*, 18 (2), 627. <https://doi.org/10.3390/s18020627>
- [13] Bhateja, A., Shrivastav, A., Chaudhary, H., Lall, B., Kalra, P.K. (2021). Depth analysis of Kinect v2 sensor in different mediums. *Multimedia Tools and Applications*, 1-26. <https://doi.org/10.1007/s11042-021-11392-z>
- [14] Antico, M., Balletti, N., Laudato, G., Lazich, A., Notarantonio, M., Oliveto, R., Ricciardi, S., Scalabrino, S., Simeone, J. (2021). Postural control assessment via Microsoft Azure Kinect DK: An evaluation study. *Computer Methods and Programs in Biomedicine*, 209, 106324. <https://doi.org/10.1016/j.cmpb.2021.106324>
- [15] Kean, S., Hall, J.C., Perry, P. (2011). Microsoft's Kinect SDK. In *Meet the Kinect*. Berkeley, CA: Apress, 151-173. https://doi.org/10.1007/978-1-4302-3889-8_8
- [16] Rahman, M. (2017). *Beginning Microsoft Kinect for Windows SDK 2.0: Motion and Depth Sensing for Natural User Interfaces*. Berkeley, CA: Apress. <https://doi.org/10.1007/978-1-4842-2316-1>
- [17] Gabbasov, B., Danilov, I., Afanasyev, I., Magid, E. (2015). Toward a human-like biped robot gait: Biomechanical analysis of human locomotion recorded by Kinect-based Motion Capture system. In *2015 10th International Symposium on Mechatronics and its Applications (ISMA)*. IEEE. <https://doi.org/10.1109/ISMA.2015.7373477>
- [18] iPi Soft LLC. (2012). *iPi Motion Capture™ Version 2.0*.
- [19] Microsoft. (2022). *Azure Kinect DK*. <https://azure.microsoft.com/en-us/services/kinect-dk/>
- [20] Dehbandi, B., Barachant, A., Smeragliuolo, A.H., Long, J.D., Bumanlag, S.J., He, V., Lampe, A., Putrino, D. (2017). Using data from the Microsoft Kinect 2 to determine postural stability in healthy subjects: A feasibility trial. *PLoS One*, 12 (2), 0170890. <https://doi.org/10.1371/journal.pone.0170890>
- [21] Zulkarnain, R.F., Kim, G.Y., Adikrishna, A., Hong, H.P., Kim, Y.J., Jeon, I.H. (2017). Digital data acquisition of shoulder range of motion and arm motion smoothness using Kinect v2. *Journal of Shoulder and Elbow Surgery*, 26 (5), 895-901. <https://doi.org/10.1016/j.jse.2016.10.026>
- [22] Amini, A., Banitsas, K. (2019). An improved technique for increasing the accuracy of joint-to-ground distance tracking in Kinect V2 for foot-off and foot contact detection. *Journal of Medical Engineering & Technology*, 43 (1), 8-18. <https://doi.org/10.1080/03091902.2019.1595762>
- [23] Mortazavi, F., Nadian-Ghomsheh, A. (2018). Stability of Kinect for range of motion analysis in static stretching exercises. *PLoS One*, 13 (7), 0200992. <https://doi.org/10.1371/journal.pone.0200992>
- [24] Seo, N.J., Fathi, M.F., Hur, P., Crocher, V. (2016). Modifying Kinect placement to improve upper limb joint angle measurement accuracy. *Journal of Hand Therapy*, 29 (4), 465-473. <https://doi.org/10.1016/j.jht.2016.06.010>
- [25] Tölgyessy, M., Dekan, M., Chovanec, E., Hubinský, P. (2021). Evaluation of the Azure Kinect and its comparison to Kinect V1 and Kinect V2. *Sensors*, 21 (2), 413. <https://doi.org/10.3390/s21020413>
- [26] Tölgyessy, M., Dekan, M., Chovanec, E. (2021). Skeleton tracking accuracy and precision evaluation of Kinect V1, Kinect V2, and the Azure Kinect. *Applied Sciences*, 11 (12), 5756. <https://doi.org/10.3390/app11125756>
- [27] Sharma, P., Anand, R.S. (2020). Depth data and fusion of feature descriptors for static gesture recognition. *IET Image Processing*, 14 (5), 909-920. <https://doi.org/10.1049/iet-ipr.2019.0230>
- [28] Simonsen, D., Popovic, M.B., Spaich, E.G., Andersen, O.K. (2017). Design and test of a Microsoft Kinect-based system for delivering adaptive visual feedback to stroke patients during training of upper limb movement. *Medical & Biological Engineering & Computing*, 55 (11), 1927-1935. <https://doi.org/10.1007/s11517-017-1640-z>
- [29] Hsu, S.C., Huang, J.Y., Kao, W.C., Huang, C.L. (2015). Human body motion parameters capturing using Kinect. *Machine Vision and Applications*, 26 (7), 919-932. <https://doi.org/10.1007/s00138-015-0710-1>
- [30] Hazra, S., Pratap, A.A., Tripathy, D., Nandy, A. (2021). Novel data fusion strategy for human gait analysis using multiple Kinect sensors. *Biomedical Signal Processing and Control*, 67, 102512. <https://doi.org/10.1016/j.bspc.2021.102512>
- [31] Shani, G., Shapiro, A., Oded, G., Dima, K., Melzer, I. (2017). Validity of the Microsoft Kinect system in assessment of compensatory stepping behavior during standing and treadmill walking. *European Review of Aging and Physical Activity*, 14, 4. <https://doi.org/10.1186/s11556-017-0172-8>
- [32] Guess, T.M., Razu, S., Jahandar, A., Skubic, M., Huo, Z.Y. (2017). Comparison of 3D joint angles measured with the Kinect 2.0 skeletal tracker versus a marker-

- based motion capture system. *Journal of Applied Biomechanics*, 33 (2), 176-181.
<https://doi.org/10.1123/jab.2016-0107>
- [33] Chakraborty, S., Nandy, A., Yamaguchi, T., Bonnet, V., Venture, G. (2020). Accuracy of image data stream of a markerless motion capture system in determining the local dynamic stability and joint kinematics of human gait. *Journal of Biomechanics*, 104, 109718.
<https://doi.org/10.1016/j.jbiomech.2020.109718>
- [34] Palmieri, P., Melchiorre, M., Scimmi, L.S., Pastorelli, S., Mauro, S. (2020). Human arm motion tracking by Kinect sensor using Kalman filter for collaborative robotics. In *Advances in Italian Mechanism Science: Proceedings of the 3rd International Conference of IFToMM ITALY*. Springer, 326-334.
https://doi.org/10.1007/978-3-030-55807-9_37
- [35] Li, H., Wen, X., Guo, H., Yu, M. (2018). Research into Kinect/inertial measurement units based on indoor robots. *Sensors*, 18 (3), 839.
<https://doi.org/10.3390/s18030839>
- [36] Li, L.F., Zou, B., Zhou, G.L., Wang, C., He, J.F. (2018). Repair and error compensation method for depth image based on optimization estimation. *Journal of Applied Optics*, 39 (1), 45-50. (in Chinese)
<http://www.yygx.net/en/article/doi/10.5768/JAO201839.0101008>
- [37] Abbasi, J., Salarieh, H., Alasty, A. (2021). A motion capture algorithm based on inertia-Kinect sensors for lower body elements and step length estimation. *Biomedical Signal Processing and Control*, 64, 102290.
<https://doi.org/10.1016/j.bspc.2020.102290>
- [38] Ryselis, K., Petkus, T., Blažauskas, T., Maskeliunas, R., Damaševičius, R. (2020). Multiple Kinect based system to monitor and analyze key performance indicators of physical training. *Human-centric Computing and Information Sciences*, 10, 51.
<https://doi.org/10.1186/s13673-020-00256-4>
- [39] Lyu, C., Shen, Y., Li, J. (2016). Depth map inpainting method based on Kinect sensor. *Journal of Jilin University (Engineering and Technology Edition)*, 46 (5), 1697-1703.
<https://doi.org/10.13229/j.cnki.jdxbgxb201605046>
- [40] Xie, X.Q., He, Y.Q., Feng, Y.W. (2020). Research on the Azure Kinect DK deep sensor error analysis and correction method. *China Plant Engineering*, (16), 24-25. (in Chinese)
<https://doc.taixueshu.com/journal/20201735zgsbgc.html>
- [41] Li, Z.L., Zhou, K., Mu, Q., Li, H.A. (2019). TOF camera real-time high precision depth error compensation method. *Infrared and Laser Engineering*, 48 (12), 263-272. (in Chinese)
<https://doc.taixueshu.com/journal/20201735zgsbgc.html>
- [42] Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., Blake, A. (2011). Real-time human pose recognition in parts from single depth images. In *CVPR 2011*. IEEE, 1297-1304.
<https://doi.org/10.1109/CVPR.2011.5995316>
- [43] Li, J.F., Xu, Y.H., Chen, Y. (2006). A real-time 3D human body tracking and modeling system. In *2006 International Conference on Image Processing*. IEEE, 2809-2812. <https://doi.org/10.1109/ICIP.2006.312992>
- [44] Gu, J.H., Li, S., Liu, H.P. (2018). Human action recognition algorithm based on angle of skeletal vector. *Transducer and Microsystem Technologies*, 37 (2), 120-123. (in Chinese)
[https://doi.org/10.13873/J.1000-9787\(2018\)02-0120-04](https://doi.org/10.13873/J.1000-9787(2018)02-0120-04)

Received December 12, 2021

Accepted May 30, 2022