Applied Mathematics and Nonlinear Sciences

<https://www.sciendo.com>

Research on aerobics training posture motion capture based on mathematical similarity matching statistical analysis

Qiuju Chen^{1†}, Rayan Atteah Alsemmeari²

¹ Hainan College of Economics and Business, Haikou Hainan, 571127, China

² Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

Submission Info

Communicated by Juan Luis García Guirao

Received June 17th 2021

Accepted September 24th 2021

Available online November 22th 2021

Abstract

Aiming at the freely editable characteristics of human motion posture, a human skeleton model is extracted, and a human motion posture model library is established. The application of motion capture system in dance training is analysed, and a method based on similarity matching between feature planes is proposed to calculate each model. Parameters of motion data between parts are obtained. After verification, the method has high accuracy and robustness for the analysis of human poses so that dancers can accurately compare the differences with standard dance movements and provide theoretical support for dancers to perform scientific dance training.

Keywords: motion capture, free editing, skeleton model, feature plane, similarity matching, pose analysis, scientific training

AMS 2010 codes: 34A34

1 Introduction

With the rapid development of human motion simulation and virtual reality, motion capture technology has been widely used in the fields of film and television production, medical rehabilitation, animation games and other fields, achieving the perfect combination of three-dimensional virtual world and reality technology. In recent years, researchers have conducted extensive research on the application of motion capture. Combining motion capture technology with Chinese puppet shows, a digital technical solution for puppet show is proposed. The motion capture system was used to study the hip and torso movement of the golfer during the swing, and the relationship between the movement of the torso, hip and swing was calculated, which provided theoretical support for the scientific training of golf. A virtual basketball training system was developed using motion

[†]Corresponding author.

Email address: hceb_cqj2010@163.com

capture technology, and users can train free throw courses by themselves [1–3]. An animation production method and process based on motion capture technology are proposed, and the techniques of animation composition and elimination of slippage in the animation production process are described. According to the normal range of motion of human joints, the motion postures of the main joint points of the human body are studied, and a method of human motion posture simulation is proposed. This method can use human motion data to drive a virtual human body model. Motion capture technology can help trainers to correct the movement of trainees by analysing and processing the motion data in real time. The traditional comparison method generally uses the Euclidean distance method to calculate the distance between two moving nodes to determine the similarity of the data of the two channels. However, this method requires equal length action time series, and the coordinate displacement caused by the trainees' height, fat and thinness will make the calculation result shift greatly [4].

In this paper, an optical motion capture system is used to track and detect the dancer's dance movements in segments, and a real-time analysis method of moving human poses based on similarity matching between feature planes is proposed. It not only improves the efficiency of posture analysis but also improves the teaching quality in real-time analysis, which has played an important role in the digital teaching of dance movements and scientific development.

2 3D data acquisition for motion capture

The motion capture system is a computer that processes data directly, sets up trackers in key parts of moving objects, uses a motion capture system to record the movement of objects and then processes them to obtain three-dimensional spatial data [5]. Motion capture can be divided into four types: mechanical motion capture, acoustic motion capture, electromagnetic motion capture and optical motion capture. The optical motion capture system gives the performer enough space to perform freely, without being restricted by space and mechanical equipment, and can capture high-speed motions or objects. This paper uses the optical motion capture system to obtain 3D motion data. It can be used to build models and skeleton databases, which fully implements high-precision real-time 3D motion data capture. The process is shown in Figure 1. Then, the 3D space coordinates of key points are used to obtain the motion of the model characteristics and finally realise the analysis of human posture.

Fig. 1 3D motion data capture process.

2.1 Near-infrared high-sensitivity digital camera captures human motion

The experimental research mainly uses an optical motion capture system, which uses a dedicated near-infrared high-sensitivity digital camera distributed in space to capture the movement of the performer in real time and transmits the captured motion to a computer to generate a virtual object. In the laboratory scenario, as shown in Figure 2, this article uses eight high-precision cameras connected to the DIMS controller with cables.

The camera resolution is 5.03 million pixels, the maximum acquisition frequency is 120 fps and the common acquisition frequency is 60 fps. The gray level is 10 bits parallel, equipped with a standard 3 to 18 mm variable focal length lens, using a marker ball with a diameter of 25 mm for human limb capture, and the human motion range is $3.0\text{ m} \times 3.0\text{ m} \times 3.5\text{ m}$ [6].

Fig. 2 Laboratory scene and main equipment.

2.2 3D motion data capture

Acquire data using an optical motion capture system. First, performers are required to wear monochromatic clothing with 21 markers on key parts, stand in a preset motion space, start high-precision 3D motion capture software, set a specified time and shoot a specified dance as required. For action, use the camera to capture and track the movement of 21 marker points, as shown in Figure 3.

Fig. 3 An actor in a monochrome costume with a marker.

Then save the captured 3D motion data, output the animation file in data editing software and save it as a TRC file. Finally, use Motion Builder to match the human motion data of the 21 marker points captured and edited with the actor model, activate the 21 marker point data to complete the matching with the actor and save the model in FBX format to complete the human motion posture. The database is established, as shown in Figure 4.

Fig. 4 Data and actor binding.

3 Human motion pose analysis based on feature vector matching

Human motion analysis is an important content in the field of computer vision research, and it is also the forefront of computer researchers' attention. The simulation of the actions of real performers needs to be supported by data. Human motion analysis is to capture, track and analyse the movement of the human body in some way to obtain relevant motion parameters. Applying motion analysis to teaching can not only establish a unique and personalised teaching system but also perform detailed decomposition of the performers' movements and demonstrate each dance step by step, which is convenient for quantitative analysis and provides scientific assistance for sports teaching help [7].

Fig. 5 Method flowchart.

This paper proposes a human motion pose analysis method based on similarity matching between feature planes. The specific process is shown in Figure 5. The main steps of the analysis process include the following:

1. Get skeleton data in real time. The dance motion sequence is acquired in real time by means of optical motion capture, and the coordinates of each landmark point of the human body model in the world coordinate system are stored.
2. Posture analysis. Determine 7 feature planes based on feature points, extract 13 feature vectors and calculate human poses based on the motion characteristics of key parts of the dance action feature correlation coefficient.
3. Analysis of feature attitude difference. Use the calculated feature correlation coefficients to analyse the differences and accuracy of students' dance moves.

3.1 Get skeleton data in real time

The analysis of human motion posture based on a motion capture system is to estimate the human motion posture from one or more perspectives. In order to analyse the posture data of dance movements, this paper uses feature points to calibrate the joint points of the human body and treats human bones as a multi-rigid body connection model. The connection between the feature points represents a rigid body; so, the adjacent joint points will not be deformed. The main action postures are mainly the movements of the head, torso, hips and limbs, as shown in Figure 6 [8].

Fig. 6 Dance gesture.

The movement of the human body is a complicated process. Without considering the conditions of the muscles and nervous system, the human movement can be abstracted into a simple chain system movement connected by some rigid bodies, as shown in Figure 7. Colour joint is used to represent the absolute position coordinates of the feature point in the world coordinate system. The upper limb is composed of two rigid bodies connected to the upper and lower arms of the elbow joint; the lower limb is composed of two rigid bodies connected to the thigh and calf connected by the hip joint. It is also connected by the knee joint; the head, torso and hip also use a line of joint points to represent a rigid body [9].

Fig. 7 Human skeleton and model.

Through the rigid body structure of the human body described above, the spatial coordinates of the landmark points are used as the data basis for calculating the vector, which is used as the basis for the subsequent similarity matching analysis of the feature plane.

3.2 3D human motion similarity matching algorithm

In pattern recognition and computer vision, the problem of similarity measurement is used to measure the difference or similarity between different objects and the movement at the marker point.

In the trajectory, the movement data curves of the marked points are independent of each other. The two sets of motion data are compared by means of direct comparison. The traditional Euclidean distance measurement method is used to calculate the similarity between the marked points. It cannot be true and reflects the accuracy of the comparison [10].

3.2.1 Similarity matching of traditional 3D models

Based on the Euclidean distance comparison method, for two-dimensional and three-dimensional spaces, the Euclidean distance calculation formulas are as follows:

$$D = \text{sqrt} \left((x_1 - x_2)^2 + (y_1 - y_2)^2 \right) \quad (1)$$

$$D = \text{sqrt} \left((x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 \right) \quad (2)$$

where $X_1 = (X_{11}, X_{12}, X_{13}, \dots, X_{1n})$, $X_2 = (X_{21}, X_{22}, X_{23}, \dots, X_{2n})$ is n-dimensional data. Every two landmarks can be calculated by the formula based on Euclidean distance. If the difference is less than the threshold (set by the coach), the two landmarks are considered similar; if the difference is greater than this threshold, the two landmarks are different. Figures 8 and 9 show standard actions and actions to be tested. Figures 8 and 9 show the comparison of the trajectory of the same mark point between the standard action and the action to be measured.

The direct comparison method based on Euclidean distance is to compare the data of two actions, and the corresponding distance value will be obtained during the comparison of each action sequence. According to the preset threshold, the percentage of data matching degree can be obtained. However, this method is too computationally intensive. In addition, the displacement of such landmarks is a reference standard, which is of little significance in the comparative training process [11].

Fig. 8 Comparison of motion trajectories in a single direction between individual landmarks.

Fig. 9 Marker data based on Euclidean distance.

3.2.2 Similarity calculation based on feature plane matching

The similarity between the dance performers and standard movements in this paper can be attributed to the similarity of geometric shapes between objects. It is known that three feature points determine a feature plane, and the skeleton of the human pose is mainly composed of seven feature planes, as shown by the shadow in Figure 10. In this paper, the angle relationship between the vector of joint points is used to evaluate the difference between dancer's movement and standard movement. According to the principles of ergonomics, the human body uses the spine as the main axis, the spine as the z-axis of the spatial rectangular coordinate system and the x-axis and y-axis of the horizontal plane as the ground plane of the motion capture device. The comparison of the similarity of the edge vectors of a plane and the similarity of normal vectors between the planes is given.

Fig. 10 Plan view of human bone features.

The main gestures of the human body include the following:

1. Movement of limbs. Taking the left arm as an example, comparing the amplitude of the left arm's elbow swing can compare the similarity between the side vector V_{LLarm} and the side vector V_{LFarm} of the left arm plane $P1$; compare with the vector V_{LFarm} correlation; the rotation direction of the arm changes the inner product of the left arm plane normal vector $V1$ and the vertical V_{stand} , as shown in Table 1.

Table 1 Characteristic vectors of limbs.

Limb movement posture	Feature vector
Stand upright	V_{stand} = Positive Z axis direction in world coordinate system
Left upper arm	$V_{LFarm} = \text{CoorLElbow} - \text{CoorLShoulder}$
Left lower arm	$V_{LLarm} = \text{CoorLElbow} - \text{CoorLHand}$
Right upper arm	$V_{RFarm} = \text{CoorRElbow} - \text{CoorRShoulder}$
Right lower arm	$V_{RLarm} = \text{CoorRElbow} - \text{CoorRHand}$
Left thigh	$V_{LThigh} = \text{CoorLKnee} - \text{CoorLHip}$
Left calf	$V_{LCrus} = \text{CoorLKnee} - \text{CoorLAnkle}$
Right thigh	$V_{RThigh} = \text{CoorRKnee} - \text{CoorRHip}$
Right calf	$V_{RCrus} = \text{CoorRKnee} - \text{CoorRAnkle}$
Left arm feature plane	$V_1 = V_{LLarm} \times V_{LFarm}$
Right arm feature plane	$V_2 = V_{RLarm} \times V_{RFarm}$
Left leg feature plane	$V_3 = V_{LThigh} \times V_{LCrus}$
Right leg feature plane	$V_4 = V_{RThigh} \times V_{RCrus}$

2. Head movements. The normal vector $V5$ of the plane $P5$ is compared with the vertical standing direction V_{stand} . When the human body looks straight ahead, $V5$ is parallel to the V_{stand} direction.
3. Chest movement. The main reason is to compare the transformation angle between the spine direction vector $V6$ and the vertical standing direction V_{stand} when doing a turning movement.
4. Hip movement. When the human body stands vertically, the hip plane $P7$ remains horizontal, and its plane normal vector $V7$ is parallel to the vertical direction V_{stand} , as shown in Table 2 [12].

Table 2 Torso pose feature vectors.

Torso movement posture	Feature vector
Head direction	$V_5 = V_{LHead} \times V_{RHead}$
Chest direction	$V_6 = V_{chest}$
Hip direction	$V_7 = (Coornavel - CoorLHip) \times (Coornavel - CoorRHip)$

To sum up the related calculation process, this paper uses the cosine similarity as the similarity function. By measuring the cosine of the angle of the inner product of two vectors in space, it can be measured by obtaining the size of the difference between them. Compared with the Euclidean distance metric, the cosine similarity pays more attention to the difference in direction between the two vectors. The calculation method is as follows:

$$similarity(\theta_i) = \frac{\sum_{t=1}^n A_t \times B_t}{\sqrt{\sum_{t=1}^n (A_t)^2} \times \sqrt{\sum_{t=1}^n (B_t)^2}} \quad (3)$$

Among them, θ_i is the joint angle; A_t and B_t are the corresponding feature plane edge vectors. The calculated cosine value is $[0,1]$ (if the dancers are performing dance moves in the same direction). If the value is close to 1, it indicates that the movement of the dancer to be tested is consistent with the standard movement and dance norms; if the value is close to 0, it indicates that the movement of the dancer to be tested is too large.

Cosine similarity can not only measure the degree of difference between vectors but also measure the similarity and difference between angles. Because there are individual differences between human bodies, such as different problems such as height, weight, arm length, etc., the proportion of human bodies is constant. Therefore, the similarity of the angles can also be used to measure whether the motion amplitude of the limb meets the standard. The calculation method is as follows [13]:

$$Corr(A, B) = 1 - \left(\frac{\arccos(similarity(\theta_i))}{\pi} \right) \quad (4)$$

According to Eqs. (3) and (4), calculate the correlation parameters of the main motion postures at the arms and compare the key motion correlation parameters of the test object with the standard motion.

Table 3 Posture correlation parameters of the left arm standard movement.

Timing	Sim (V2, Vstand)	Corr (VLLarm, VLFarm)	Corr (VLFarm, Vstand)
0–1 s	0.6442	0.9365	0.9586
1–2 s	0.7004	0.9623	0.9956
2–3 s	0.7257	0.9726	0.9659
3–4 s	0.7126	0.9759	0.9546
4–5 s	0.6897	0.9588	0.9849
5–6 s	0.6416	0.9659	0.9546
6–7 s	0.5659	0.9588	0.9659
7–8 s	0.7258	0.9649	0.9785
8–9 s	0.6896	0.9876	0.9416
9–10 s	0.6585	0.9659	0.9755

Table 4 Parameters related to the posture of the left arm to be measured.

Timing	Sim (V2, Vstand)	corr (VLLarm, VLFarm)	corr (VLFarm, Vstand)
0–1 s	0.6357	0.9433	0.9576
1–2 s	0.6958	0.9613	0.9879
2–3 s	0.7253	0.9715	0.9646
3–4 s	0.7110	0.9698	0.9559
4–5 s	0.6906	0.8399	0.9848
5–6 s	0.6506	0.9648	0.9547
6–7 s	0.5998	0.9498	0.9646
7–8 s	0.7199	0.9650	0.9689
8–9 s	0.6826	0.9798	0.9398
9–10 s	0.6595	0.9698	0.9698

3.3 Calculation of differences between feature poses

This experiment uses dance music under the same rhythm to ensure that the two groups complete the same specified dance movements in the same time period, regardless of time regularity. First, two groups of people complete an action at the same time and calculate the correlation coefficient $corr$ of each part of its action. Second, calculate the correlation coefficient error, using the relative error method, and the calculation method is as follows:

$$\Delta corr_t = \frac{|corr_i - corr_j|}{corr_i} \% \quad (5)$$

Among them, t is the time point of the dance action, $corr_i$ is the standard relative coefficient of the dance action and $corr_j$ is the relative coefficient of the dance action to be measured. The basis for judging the action standard is the relative error in calculating the correlation coefficient. The condition for error convergence is given as follows:

$$\Delta corr_t \leq C \quad (6)$$

Among them, C is the selected error threshold, and the corresponding error threshold can be modified according to the teacher's requirements for the dance action. Calculate the relative error according to formula (5) to indicate the deviation of the dance movements of each object from the standard movements. The calculation results are shown in Table 5.

Table 5 Relative error of main movement attitude of test object/%.

Timing	Sim V2, Vstand)	Corr VLLarm, VLFarm)	Corr VLFarm, Vstand)
0–1 s	1.32	0.72	0.11
1–2 s	0.65	0.10	0.77
2–3 s	0.05	0.11	0.13
3–4 s	0.22	0.63	0.14
4–5 s	0.13	12.54	0.01
5–6 s	1.41	0.11	0.01
6–7 s	5.65	0.95	0.13
7–8 s	0.82	0.01	0.99
8–9 s	1.02	0.79	0.19
9–10 s	0.15	0.40	0.58

4 Analysis of experimental results

The experimental platform is Core i5-3470. Then, 2 GHz CPU, 4 GB memory PC and MATLAB were used as the development environment. The created motion database contains 18 sets of dance action fragments, each of which is about 600 frames. The experimental subjects are randomly selected college students. The experimental subjects have a basic dance. Firstly, the subject is required to imitate the standard movements of the dance teacher, make corresponding dance movements under the optical motion capture system, extract the motion characteristics of the joint points of the left arm of the subject and take a local motion sequence captured in real time as an example. The main movement changes of the subject were analysed differently from the standard movements. This article mainly compares the final movement of the left arm with a single dance movement (within 0 to 10 s) for experimental comparison [14].

In the experiment, the error threshold is set to $C \leq 2$, and the feature pose differences are compared to select the parts that do not meet the threshold. The degree of bending of the elbow in the time interval of 4 to 5 s and the swing amplitude of the left arm of the time interval of 6 to 7 s were significantly different from the standard movement. In this paper, the feature plane is used as the basic calculation plane, and three discriminative parameters are calculated. Figure 11 shows the comparison chart of the left arm movement posture difference.

Fig. 11 Timing chart of left arm motion correlation parameters.

The traditional three-dimensional model similarity comparison method based on Euclidean distance is used to directly compare the two sets of data. For dance moves with large amplitudes, the posture analysis results are often not accurate enough. Persons with different heights, who are fat and who are lean with different proportions of the human body will have deviations in the displacement of the model itself, as shown in Figure 12. Without specifying the movement position of the person to be tested, the spatial displacement deviation is often too large, as shown in Figure 13.

Fig. 12 x-axis distance difference of a single feature point.

Fig. 13 Euclidean distance difference of a single feature point.

Through the verification of the experimental results, the motion pose analysis method based on the similarity matching of feature planes can clearly and efficiently detect the differences and norms between moving objects, has high robustness and provides scientific theoretical support for dance scientific training [15].

5 Conclusion

This paper proposes a human pose analysis method based on similarity matching between feature planes. First, the optical motion capture system is used to collect the movement sequences of dance performers in real time, to obtain human skeleton data and to establish a human motion model; then, to use the method of feature plane similarity matching to calculate the correlation degree of motion data, finally, use this method to achieve the posture of a moving human body. Real-time analysis and effective application in dance teaching, comparing the formativeness and consistency of teaching movements, is conducive to the realisation of computer-aided dance teaching mode and improve teaching quality [16]. The experimental data verify that the method has good accuracy for real-time analysis of human posture and is of great significance for the realisation of digital teaching of dance. It also lays a good foundation for the next step of retrieval of dance movement data.

References

- [1] Shaofei Wu, A Traffic Motion Object Extraction Algorithm, *International Journal of Bifurcation and Chaos*, 2015,25(14):pp. 154-169.
- [2] Shaofei Wu, Mingqing Wang, Yuntao Zou. Research on internet information mining based on agent algorithm, *Future Generation Computer Systems*, 2018. 86:pp. 598-602.
- [3] Shaofei Wu, Nonlinear information data mining based on time series for fractional differential operators, *Chaos*, 2019. 29:pp. 113-114.
- [4] Minxiang Ye, Cheng Yang, Vladimir Stankovic, Lina Stankovic, & Andrew Kerr. A depth camera motion analysis framework for tele-rehabilitation: motion capture and person-centric kinematics analysis. *IEEE Journal of Selected Topics in Signal Processing*, 2016, 10(5):pp. 877-887.
- [5] Schurr, S. A., Marshall, A. N., Resch, J. E., & Saliba, S. A. Two-dimensional video analysis is comparable to 3d motion capture in lower extremity movement assessment. 2017, 12(2):pp. 163-172.
- [6] Youn-Kyoung Joung, Quanri Li, & Sang Do Noh. Xml-based neutral schema for automated ergonomic analysis with digital human simulation and inline motion capture. *International Journal of Computer Integrated Manufacturing*, 2016, 29(1):pp. 25-41.
- [7] Barnes, A., Mental, R., Macnamara, B., & Vick, J. Matlab visualization program for analysis of c3d motion capture data of speech and language. *Motion capture on the web.*, 2016, 202(3): pp. 159-162.
- [8] Lindsay, A. Kaydara's motionbuilder: this product bridges the gap between keyframing and motion capture. review. (brief article. 2016, 49(2): pp. 708-708.

- [9] Shaghayegh Zihajehzadeh, & Edward J. Park. A novel biomechanical model-aided imu/uwb fusion for magnetometer-free lower body motion capture. *IEEE Transactions on Systems Man & Cybernetics Systems*, 2016, 47(6): pp. 1-12.
- [10] S. Rojas-Lertxundi, J. R. Fernández-Lopéz, S. Huerta, & P. G. Bringas. Motion capture systems for jump analysis. *Logic Journal of IGPL*, 2017, 25(6): pp. 890-901.
- [11] Ronald Yager, Fred Petry, & Paul Elmore. Multiple attribute similarity hypermatching. *Soft Computing*, 2017, 22(5): pp. 2463-2469.
- [12] Hafidha Harraga, Mustapha Yebdri. Attractors for a nonautonomous reaction-diffusion equation with delay, *Applied Mathematics and Nonlinear Sciences*, 2018, 3(1):pp. 127-150.
- [13] M. K. Ammar, D. A. Oda. Design of Gravity Assist trajectory from Earth to Jupiter, *Applied Mathematics and Nonlinear Sciences*, 2018, 3(1):pp. 151-160.
- [14] P. K. Pandey, A new computational algorithm for the solution of second order initial value problems in ordinary differential equations, *Applied Mathematics and Nonlinear Sciences*, 2018, 3(1):pp. 167-174.
- [15] R. A. Mundewadi, Kumbinarasaiah S, Numerical Solution of Abel's Integral Equations using Hermite Wavelet. *Applied Mathematics and Nonlinear Sciences*, 2019. 4(2):pp. 395-406.
- [16] Shanmukha B. Venkatesha, Some Results on Generalized Sasakian Space Forms. *Applied Mathematics and Nonlinear Sciences*, 2020. 5(1):pp. 85-92.

This page is intentionally left blank