Applied Mathematics and Nonlinear Sciences

<https://www.sciendo.com>

Research on motion capture of dance training pose based on statistical analysis of mathematical similarity matching

Qingwen Chen¹, Abdullah Albarakati², Lanlan Gui^{1†}

¹ Department of Art, Shenzhen University, Shenzhen 518060, China

² Department of Information Systems, 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

In order to verify the effectiveness and feasibility of the combination of motion capture technology and teaching, based on dance teaching, this paper proposes a dance posture analysis method based on feature vector matching and applies it to dance teaching.. The main research work includes the following: (1) according to the characteristics of human motion poses-free editing, extracting human skeleton models, establishing a human motion model database, analysing the application of motion capture systems in dance training, and proposing a method of feature plane similarity matching to calculate model components and motion parameters. After verification, the method has high accuracy and robustness for the analysis of human posture, so that dancers can accurately compare the differences with standard dance movements, and provide theoretical support for scientific dance training. (2) Aiming at the complexity of learning dance, a dance teaching method based on motion capture technology is proposed. Using motion capture technology, a whole complex dance movement is decomposed into many small segments to make a teaching animation, which guides students to learn based on small dance movement. Imitation makes the abstract theory vivid, intuitive and easy to understand, which is conducive for the innovation of education and teaching methods.

Keywords: posture analysis, motion capture, feature plane, similarity matching, dance teaching

AMS 2010 codes: 18F25

1 Introduction

In the field of computer research today, people are increasingly hoping to use computer recognition to determine the external environment, to provide people with more intelligent information and services. Computer vision uses a single or multiple cameras to obtain external environmental information, then inputs the acquired

[†]Corresponding author.

Email address: llguian681@163.com

video or image information into a computer, converts it into digital signals, and processes it by the computer to analyse and understand possible data information. Thereby the process of judging the corresponding motion can be divided into three levels of low, medium and high, according to the computer research object. Also included are the edge gradient, texture grey value, contour, location information and deep understanding of the scene. Computer vision classification is shown in Figure 1 [1, 2].

Fig. 1 Schematic diagram of computer vision classification.

Computer vision can enhance the accuracy, effectiveness and robustness of information capture. It has universal application value and research significance in the following areas.

Motion analysis is one of the main research contents of computer vision. Applying motion analysis to dance teaching analyses and discriminates the dance movements of trainees, which helps trainers to find out where the movements are irregular in time, using dance technology. The action parameters are converted into a three-dimensional (3D) motion model, a teaching sample library is established and the analysis results are used as important indicators to measure the education level, which helps to improve the quality of education [3].

The research in this thesis combines motion analysis in computer vision with college dance teaching, researches on the capture and analysis of human motion characteristics, and detects and discriminates samples in motion in video, to achieve the analysis of human motion posture and realise the science behind the purpose of teaching and training. Using the high-precision infrared camera's accurate capture and the computer system's precise processing, more information can be obtained, such as the displacement, angle and speed of the human limb and joint movements. The intelligent method provides realistic and effective learning for dancing theoretical basis.

2 Human pose analysis

The human motion posture analysis first emerged in the West. With the advent of the new century, domestic research experts in China have also begun to study and analyse the 3D posture. Generally, the data obtained by means of motion capture devices or video are discretely changed. The attitude data parameters are extracted from the discrete data to determine the law of attitude changes. The use of mathematical similarity analysis is one of the main approaches. At this stage, there are several methods of posture analysis [4]:

2.1 Attitude analysis based on simulated annealing particle swarm algorithm

The literature mainly studies the posture of the human arm. The video is used to obtain the parameters of the arm's posture. The particle filter algorithm is used as the weight of the sampling points to approximate the condition density. The approximate weights are randomly updated by Monte Carlo. Next, is to predict the random motion of the particles and measure the range of particle motion to extract the posture of the human arm. The literature introduced annealing particle swarm optimisation method into the analysis of human motion posture. First, the captured 3D parameters are reduced by PCA to obtain a low-dimensional compact attitude

space. Second, the fitness function is constructed using human silhouette features in low-dimensional space to measure the similarity between the model and image features. Finally, the annealing particle swarm optimisation algorithm is used to improve the analysis efficiency [5] as

$$E(y_t, x_t) = S \times \left((1 - \beta) \frac{B_t}{B_t + Y_t} + \beta \frac{R_t}{R_t + Y_t} \right) \quad (1)$$

where R_t is the silhouette area of the human body in the calculated image, B_t is the projected area of the model and Y_t is the overlap area of the two.

2.2 Model-based motion posture analysis

This method first needs to establish a corresponding representation model of the human body, and then map the model into a 2D image, and express the similarity between the model and the image data by some evaluation function. First, in the model selection and establishment stage, a good model representation can approximate the human body in an image. For example, the human skeleton is represented by a stick figure. In the method of least square curve fitting, the law of human motion is approximated by the obtained polynomial coefficients [6]. First, calculate the function $f(t) = a_0t + a_1$ of the discrete coordinate data point x_i, y_i, z_i as a function of time. To determine the approximate trend of the function, expect $\Phi = \text{span}\{\varphi_0t, \varphi_1t, \dots, \varphi_nt\}$ to find the most suitable function $f(t)$ in the coefficient space, that is,

$$s(t) = a_0\varphi_0t + a_1\varphi_1t + \dots + a_n\varphi_nt \quad (2)$$

The closest function $s(t)$ is the best fit to function B $f(t)$. as shown in Figure 2. Geometrically speaking, this method is to find the sequence of given points and find the curve with the smallest sum of squared distances.

Fig. 2 Least squares fitting curve.

2.3 Feature-based motion posture analysis

This type of method does not require modelling in advance, and feature selection is the basic work of attitude analysis. The common features are points, lines, faces, blocks, corners and other more complex features. For example, the centroid of the human body is used as a pose feature to represent the orientation of the frame where it is located, and the centroid position of the person in the next frame is estimated by changing the centroid position [7].

3 Basic theory of motion capture technology

3.1 Components of a motion capture system

With the continuous progress and development of computer hardware facilities, motion capture devices are becoming more and more commercialised. Different companies have relatively different configurations of motion capture devices, but in general, relatively classic capture devices often include the following components:

1. **Sensor:** In the motion capture system, the sensor is a kind of tracking equipment that is firmly set in the key part of the moving target. The motion capture device can capture the displacement, speed and angle of a moving object through the physical signals sent by the sensor. There are usually different requirements for the number of sensors for different inspection sites of different inspection objects. For example, to capture the movement of the human body, the demand for sensors is usually relatively small. For tracking research such as gesture capture, more sensors are needed [8].
2. **Signal capture device:** This kind of device will be different due to different types of motion capture systems. Its main function is to capture the position signal supplied by the sensor. As for the mechanical capture system, it uses a circuit board. To collect electrical signals, but now a relatively extensive optical capture system is used, which uses the higher resolution of infrared cameras to capture all signals.
3. **Data transmission equipment:** This equipment transmits all the motion parameters acquired in the motion capture device to the computer system, analyses and processes the relevant parameters in a timely manner through the computer system, and maintains the accuracy, efficiency and accuracy of the transmission process.
4. **Data processing equipment:** The motion data captured by motion capture equipment needs to be corrected and processed. After that, it is matched with the 3D model. The processed data makes the 3D model more realistic and natural.

3.2 Optical motion capture system

This article studies the optical capture system, which is the first practical optical capture product in China—Dongfangxin DIMS9100. The system uses multiple near-infrared high-sensitivity cameras, and uses the ‘S/D three-dimensional space calibration algorithm’ developed independently to improve the real-time 3D data. The main hardware equipment of the system includes a high-speed 3D data capture workstation, a DIMS controller, a dedicated near-infrared high-sensitivity camera, a high-sensitivity marking point, a high-precision 3D static and a dynamic calibration rod, as shown in Figure 3. The main operation flow of the system is shown in Figure 4 [9].

Fig. 3 Optical capture device hardware facilities.

Fig. 4 Flow chart of the optical operating system.

4 Analysis of dance pose based on feature vector matching

4.1 Acquisition of 3D data for motion capture

In this paper, an optical motion capture system is used to achieve the acquisition of motion data, to establish a database of human motion pose models and skeleton models. The basic flow is as shown in Figure 5.

Fig. 5 3D data acquisition flowchart.

4.2 Human motion pose analysis based on feature vector matching

Motion posture analysis is the process of tracking, capturing, acquiring and analysing the features of the human body to obtain the relevant parameters. Through the effective combination of motion analysis and teaching, the teaching system can be more personalised and special. It can also decompose the performance of the performers in detail, and gradually demonstrate each dance movement. The parameters obtained are helpful for quantitative analysis of the movement posture, providing good help for more scientific and intelligent dance teaching [10].

In order to better analyse the movement status of dance performers, a method of analysing human motion poses by using the principle of feature plane similarity matching is proposed. This method simplifies the traditional calculation of Euclidean distance based on multiple identification points to base on calculation of the feature plane feature vector and its included angle. In this paper, the identification points of 21 key parts are

simplified into seven feature planes to calculate the difference and correlation of motion. After verification, this method can quickly and effectively analyse human movement posture and apply it to dance teaching to improve the efficiency of dance teaching. The specific process is as shown in Figure 6, where the main steps of the analysis process include as follows:

Step 1. Real-time acquisition of skeleton data: The optical motion capture method is used to acquire the dance action sequence in real-time, and the coordinates of each identification point of the human model in the space coordinate system are stored.

Step 2. Posture analysis: Determine seven feature planes based on the feature points, extract the angle between the feature vector and the pose feature vector, and calculate the feature correlation coefficient of the human pose based on the movement characteristics of the key parts of the dance action.

Step 3. Analysis of feature pose difference: Through the correlation coefficients of feature vectors and their included angles, analyse the differences and accuracy of student dance movements and standard movements.

Fig. 6 Method flow chart.

4.3 Traditional 3D model similarity matching

The similarity matching of human poses is a measure of the pose difference or similarity between different human bodies. The most used method is the traditional Euclidean distance measurement.

The traditional 3D model similarity matching is based on Euclidean distance, and its calculation methods are as follows:

$$D = \text{sqrt} \left((x_1 - x_2)^2 + (y_1 - y_2)^2 \right) \quad (3)$$

$$D = \text{sqrt} \left((x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 \right) \quad (4)$$

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 marking points can be calculated by the formula based on the Euclidean distance. If the difference is less than the threshold, which is set by the coach, the two marking points are considered to be similar. If the difference is greater than this threshold, the two identification points are considered to be dissimilar. Comparison of the track of the same marked point between the standard action and the action are measured, as shown in Figure 7 [11].

The direct comparison method based on the traditional Euclidean distance is to compare the moving trajectories of two moving targets. During the comparison of each action sequence, the corresponding distance difference is obtained, and the degree of data matching is calculated according to a preset threshold. However,

Fig. 7 Comparison of motion trajectories in a single direction between a single identification point.

this method not only has a large amount of calculation, but also depends on the inherent characteristics of the test object. When the object's body proportion, such as height, weight and weight changes, the distance between the marking points also changes accordingly and must be repeated in the settings and calculations. Therefore, the strict requirements for moving objects limit the use of traditional methods, which not only greatly reduces the computational efficiency, but also lacks universality [12].

4.4 Similarity calculation based on feature plane matching

This article uses a motion capture system to extract a human skeleton model. First, a feature plane can be determined based on three feature points, and 21 identification points on the human motion posture skeleton are converted into 7 feature planes as basic calculation planes. Each feature plane represents various parts of the human body, as shown in Figure 8. Seven feature plane normal vectors ($V_1 \sim V_7$) are extracted on the feature plane to determine the difference in the overall direction of the motion pose. Second, the local normality of the motion pose plane is determined by the angle ($\theta_1 \sim \theta_7$) of the feature plane edge vector. In addition, the feature plane edge vector and the trunk included angle ($\theta_8 \sim \theta_{12}$) in the vertical direction (V_{stand}) considers the local relationship between the limb and the torso. In this way, a two-tuple $\langle V, \theta \rangle$ is used as the input of the model similarity calculation, and the correlation parameter of $\langle V, \theta \rangle$ is used as the output. This method effectively solves the calculation errors caused by the inherent characteristics of the object to be measured, and reduces the calculation complexity, thereby improving the efficiency and stability of the human posture analysis [13].

Fig. 8 Plan view of human bone features.

Table 1 Discriminant vectors on the feature plane.

Feature plane	Feature vector
Eigenvector left arm (P1)	$V_1 = V_{LLarm} \times V_{LFarm}$
Right arm (P2)	$V_2 = V_{RLarm} \times V_{RFarm}$
Left leg (P3)	$V_3 = V_{LThigh} \times V_{LCrus}$
Right leg (P4)	$V_4 = V_{RThigh} \times V_{RCrus}$
Head (P5)	$V_5 = V_{LHead} \times V_{RHead}$
Chest (P6)	$V_6 = V_{LChest} \times V_{RChest}$
Hip (P7)	$V_7 = V_{LHip} \times V_{RHip}$

Table 2 Determining the included angle on the characteristic plane.

Feature plane	Geometric relationship	θ_{Max}	θ_{Min}
Left elbow (P1)	$\theta_1 = \langle V_{LFarm}, V_{LLarm} \rangle$	180°	40°
Right elbow (P2)	$\theta_2 = \langle V_{RFarm}, V_{RLarm} \rangle$	180°	40°
Left knee (P3)	$\theta_3 = \langle V_{LThigh}, V_{LCrus} \rangle$	180°	35°
Right knee (P4)	$\theta_4 = \langle V_{RThigh}, V_{RCrus} \rangle$	180°	35°
Head (P5)	$\theta_5 = \langle V_{RHead}, V_{Lhead} \rangle$	-	-
Chest (P6)	$\theta_6 = \langle V_{RChest}, V_{LChest} \rangle$	-	-
Hip (P7)	$\theta_7 = \langle V_{RHip}, V_{LHip} \rangle$	-	-

According to the normative requirements of dance movements and the relative range of movement of the human skeleton, the feature vectors of the main movement parts of the human body and their included angles are specified as shown in Tables 1 and 2:

1. The movement direction of the limbs and limbs can be determined by the inner product of the limb feature plane P_m ($m = 1, 2, 3, 4$), the normal vector V_m ($m = 1, 2, 3, 4$) and the vertical vector V_{stand} of the trunk to determine the direction of movement, and the angle θ_m ($m = 1, 2, 3, 4$) to accurately determine the regularity of limb movement.
2. Compare the normal vector V_5 of the head characteristic plane P_5 with the vertical standing direction V_{stand} to determine the included angle θ_5 of the head of the motion model. When the human body looks straight ahead, V_5 and V_{stand} are parallel [14].
3. The torso includes the chest and hips. It is mainly a turning movement and a bending movement. Among them, the turning movement compares the spine direction vector V_6 with the vertical standing direction V_{stand} and transforms the angle θ_6 . For the bending movement, when the human body stands upright, the hip plane P_5 remains horizontal, and its plane normal vector V_7 is parallel to the vertical direction V_{stand} .

The angle between the feature plane edge vector and the vertical direction is used to determine the relationship between the limbs and the trunk joints, and the connection angle is determined as shown in Table 3:

Table 3 The angle between the joints.

	Geometric relationship	θ_{Max}	θ_{Min}
Left shoulder	$\theta_8 = \langle V_{LFarm}, V_{Stand} \rangle$	180°	0°
Right shoulder	$\theta_9 = \langle V_{RFarm}, V_{Stand} \rangle$	180°	0°
Left hip	$\theta_{10} = \langle V_{LThigh}, V_{Stand} \rangle$	45°	0°
Right hip	$\theta_{11} = \langle V_{RThigh}, V_{Stand} \rangle$	45°	0°
head	$\theta_{12} = \langle V_5, V_{Stand} \rangle$	45°	0°

In this paper, the cosine similarity method is used as the similarity function. This method can not only measure the degree of difference between vectors, but also measure the similarity and difference between angles. According to the angle cosine of the product of the two vectors in space, the difference between the feature vectors can be measured. The calculation method is as follows:

$$\theta_{\langle i,j \rangle} = \arccos(\text{similarity}(V_i, V_j)) \quad (5)$$

The effective movement range $\theta_{\langle i,j \rangle}$ between the feature vectors and the value range $[\theta_{\min}, \theta_{\max}]$ are used to judge the standard and formativeness of the dance movement amplitude. If $\theta_{\langle i,j \rangle} < \theta_{\min}$ or $\theta_{\langle i,j \rangle} > \theta_{\max}$, then the corresponding limb movement is wrong and the next operation cannot be performed. If $\theta_{\min} \leq \theta_{\langle i,j \rangle} \leq \theta_{\max}$, then the corresponding limb movement is within the discriminable range, that is, the next operation is performed.

Because there are individual differences between human bodies, such as different problems such as height, weight and arm length, the proportion of human bodies is constant. Therefore, it is necessary to measure whether the limb's motion amplitude meets the standard through the similarity of the angles. The calculation method is as follows:

$$\text{Corr}(\theta_{\langle i,j \rangle}) = 1 - \left(\frac{\arccos(\text{similarity}(V_i, V_j))}{\pi} \right) \quad (6)$$

Taking the movement of the left arm of the dance trainer as an example, and taking the characteristic plane P1 of the left arm as the basic calculation plane, three discriminative parameters $\{\text{Sim}(V_1, V_{\text{stand}}), \text{Corr}(\theta_1), \text{Corr}(\theta_8)\}$ can be obtained, and the overall posture of the left arm can be determined based on the above three parameters. Experiments show that the calculation error is effectively reduced. The results are shown in Table 4 [15].

Table 4 Correlation parameters of left arm gesture (0–1 s).

Test subject	Sim (V_1, V_{Stand})	Corr (θ_1)	Corr (θ_8)
Standard object	0.6943	0.9729	1
Test object	0.7025	0.8299	0.9815

5 Analysis of experimental results

This experiment was completed on a PC with a Core (TM) i5-3470 3.2 GHz CPU and 4GB of memory, and MATLAB was used as the development environment. The created motion database contains 18 sets of dance action clips, and each set of dance actions is about 1200 frames. The experimental subjects are randomly selected college students. The experimental subjects have a basic dance.

The subjects are required to imitate the standard actions of the dance teacher, make corresponding dance moves under the optical motion capture system and extract the kinematic characteristics of the joint points of the left arm of the subjects. Compared with the standard actions, this article mainly uses single-segment dance moves (0–10 s (Inner) The final movement of the left arm is compared experimentally. Taking a real-time captured local motion sequence as an example, as shown in Figure 9, the differences between the main movement changes of the test object and the standard movements are compared [16].

Through the comparison analysis of the degree of difference, it is concluded that the degree of elbow flexion of the test object in the 4–6 s time interval and the left arm swing amplitude in the 6–8 s time interval are significantly different from the standard movements, as seen in a comparison chart of the differences in the left arm's movement posture (Figure 10). From the figure, the difference between the action to be measured and the standard action is clearly seen [17].

Through comparison and verification of the experimental results, the method of feature plane similarity matching can be used to clearly and efficiently detect the differences and standards of moving objects in the

analysis of motion poses. It has high robustness and lays the foundation for scientific training in dance basis [18].

Fig. 9 Movement differences of the skeleton model.

Fig. 10 Timing diagram of left arm motion correlation parameters.

6 Conclusion

The advantages of this research are mainly connected with dance teaching through the optical motion capture system, which improves the intuitiveness of learning effects, real-time data collection and analysis, and provides timely feedback for teaching. At the technical level, the teaching form and student acceptance of innovative and other aspects provide scientific theoretical support, get rid of other interference factors of the traditional teaching mode, provide a reliable basis for the improvement of the teaching mode, and help the system to improve personalised teaching system. The next major research work is to complete the real-time analysis of human motion posture with the assistance of an optical motion capture system.

Acknowledgements.

The authors acknowledge the teaching research and reform project of Guangdong Province, 2018 (Guangdong higher education [2018]180) and the general project of Guangdong Education Science Planning Project '13th five year plan' in 2020 Guangdong Hong Kong Macao Big Bay area international education demonstration area construction research (2020wqyb063).

References

- [1] 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.
- [2] 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.
- [3] 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.
- [4] Barnes, A., Mental, R., Macnamara, B., & Vick, J. MATLAB visualization program for analysis of c3d motion capture data of speech and language., 2016. 140(4):pp. 3403-3403.
- [5] Hedwig, B. Motion capture on the web., 202(3): pp. 159-162.
- [6] Lindsay, A. Kaydara's motionbuilder: this product bridges the gap between keyframing and motion capture. review. (brief article., 2016. 49(2):pp. 708-708.
- [7] 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.
- [8] 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.
- [9] Ronald Yager, Fred Petry, & Paul Elmore. Multiple attribute similarity hypermatching. *Soft Computing*, 2017. 22(5):pp. 2463-2469.
- [10] Shaofei Wu, A Traffic Motion Object Extraction Algorithm, *International Journal of Bifurcation and Chaos*, 2015. 25(14):pp. 154-169.
- [11] Shaofei Wu, Mingqing Wang, Yuntao Zou. Research on internet information mining based on agent algorithm, *Future Generation Computer Systems*, 2018. 86:pp. 598-602.
- [12] Shaofei Wu, Nonlinear information data mining based on time series for fractional differential operators, *Chaos*. 2019. 29:pp. 113-114.
- [13] Abdul Rauf Nizami, Afshan Perveen, Waqas Nazeer, Mahnoor Baqir. WALK POLYNOMIAL: A New Graph Invariant, *Applied Mathematics and Nonlinear Sciences* 2018. 3(1): pp. 321-33.
- [14] Sk. Sarif Hassan, Moole Parameswar Reddy, Ranjeet Kumar Rout. Dynamics of the Modified n-Degree Lorenz System, *Applied Mathematics and Nonlinear Sciences*, 2019. 4(2), pp. 315-330.
- [15] Dushko Josheski, Elena Karamazova, Mico Apostolov. Shapley-Folkman-Lyapunov theorem and Asymmetric First price auctions, *Applied Mathematics and Nonlinear Sciences*, 2019. 4(2), pp. 331-350.
- [16] M. Salai Mathi Selvi, L. Rajendran. 2019. Application of modified wavelet and homotopy perturbation methods to nonlinear oscillation problems, *Applied Mathematics and Nonlinear Sciences*, 2019. 4(2):pp. 351-364.
- [17] Weiguo Zhao, Tiancong Shi, Liying Wang, Fault Diagnosis and Prognosis of Bearing Based on Hidden Markov Model with Multi-Features. *Applied Mathematics and Nonlinear Sciences*, 2020. 5(1):pp. 71-84.
- [18] Yaya Sagna, Multidimensional BSDE with Poisson jumps of Osgood type. *Applied Mathematics and Nonlinear Sciences*, 2019. 4(2):pp. 387-394.

This page is intentionally left blank