

Design and implementation of Peking Opera action scoring system based on human skeleton information

Hang Zhao^{1,a}, Youdong Ding^{1,2}, Bing Yu¹, Chenfeng Jiang^{1,2}, Wanying Zhang^{1,2}

¹Shanghai Film Academy, Shanghai University, Shanghai 200072 P.R. China

²Shanghai Engineering Research Center of Motion Picture Special Effects, Shanghai University, Shanghai 200072 P.R. China

Abstract. At present, most of the preservation records of Peking Opera remain in the ways of video and text, and the digitalization degree is far lower than the development level of science and technology. The immaterial cultural heritage cannot be fully displayed and Peking Opera's value is weakened. Therefore, adopting advanced motion capture technology is of great significance to the protection and inheritance of Peking Opera. We use optical motion capture equipment to record the movement information of Peking Opera actors, then keep the human skeleton information in a specific file format. After that, the hierarchical human action skeleton model was analysed, and the final score was obtained by comparing the change sequence of information of reference action and training action skeleton with the improved DTW algorithm. We have realized the graphical interface of the system, and the trainer can easily select the action segments to train or select a specific body part for specific action training. This paper introduces the overall design framework of our Peking Opera action scoring system, including the collection of action information, the implementation of scoring algorithm and the design of software interface.

1 Introduction

On November 16, 2010, Peking Opera was included in the "representative list of intangible cultural heritage of human beings". However, in recent years, the frequency of Peking Opera's appearance in people's field of vision has become lower and lower, and the audience of Peking Opera has become less and less, especially for young viewers. For the younger generation, few people pay attention to the effective inheritance and healthy development of Peking Opera.

Some excellent Peking opera performance techniques are on the verge of loss and demise. It is need to provide scientific and effective protection and inheritance through different means [2]. The traditional method for the preservation of Peking Opera is mainly to record various repertoires through video recording. This method can only be used to observe the movements of Peking Opera from the perspective of shooting at that time, and it is impossible to carry out three-dimensional all-round display. The shortcoming is that the recovery of character movement is difficult to achieve, the network communication and display effect is poor, the reusability is low, and the degree of digitization is low [1].

In addition, the use of shooting instruments for shooting records, the editability of the data is poor, which is not conducive to new generation and development. With the continuous development of digital technology, the data of Peking Opera action is collected by motion capture technology to form three-dimensional data. Later, through editing and processing,

the Peking Opera action data can be re-developed and utilized. The rapid development of motion capture technology not only effectively broadens the protection of Peking opera culture, but also provides a new and optional means for the development and reuse of Peking Opera action [2].

At present, most of the skeletal model representation methods are represented by the three-dimensional coordinates of the joint points of the human body or the three-dimensional rotation angle of the skeleton. And most of them use machine learning methods to divide and match actions. This paper uses the three-dimensional rotation of the human skeleton to describe the movement changes of the human body posture.

2 System design framework

With the rapid development of motion capture technology, the motion capture device interaction system has become an important technical means in the field of education and training.

Our Peking Opera Action Scoring System uses our optical motion capture equipment as the Peking Opera action skeleton information collector, including ten high-speed cameras, professional sportswear with 50 feature points and corresponding software facilities.

In general, our Peking Opera action scoring system includes the following parts: the collection of human motion information, the establishment of a standard action database, the implementation of the action scoring

* Corresponding author: ^a Hang Zhao: 18621621921@163.com

algorithm, and the production of the system software interface.

The system design framework flow chart is shown in Figure 1.

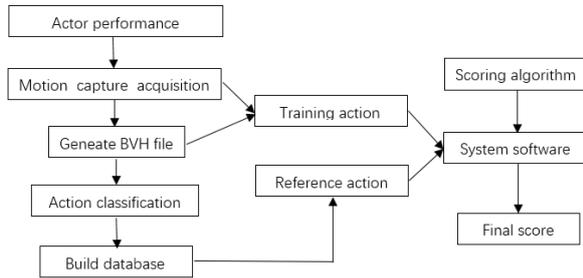


Figure 1. System design framework flow chart.

3 Experimental data representation and processing

We invited the professional Peking Opera teacher of Shanghai Film Academy to collect the Peking Opera action data. We selected some classic movements in Peking Opera, including Yunshou, Yunbu, Laobu, Huazizibu and so on.

3.1. Marked optical motion capture system

Motion capture is a high-tech that accurately measures and records the motion trajectory or attitude of an object in the real three-dimensional space in real time and reconstructs the motion state of the moving object at each moment in the virtual three-dimensional space.

Optical motion capture is a technique for capturing motion by monitoring and tracking target feature points from different angles by multiple high-speed cameras based on the principle of computer vision.

Marked optical motion capture systems generally consist of optical markers, motion capture cameras, signal transmission devices, and data processing workstations. Optical motion capture systems, commonly referred to as optical motion capture systems, are commonly referred to as point-and-click motion capture systems.

Marker points are attached to key parts of moving objects (such as joints of human body). Multiple motion capture cameras detect Marker points from different angles in real time. The data is transmitted to the data processing workstation in real time. The space coordinates of Marker points are accurately calculated according to the triangulation principle. Then, from the principle of biokinetics, the 6-degree-of-freedom movement of the skeleton is calculated.

The motion capture system we use is a marked point optical motion capture system that is arranged by 10 high-speed digital cameras around the performance venue. The field of view overlap of these cameras is the performance range of the performer.

Peking Opera action actors are required to wear professional tight-fitting clothing, and put 50 reflective marking points on each key part of the body. The marking points should be placed on the basic skeleton of the capturing movement, which can be recognized by

computer [3]. The Actor and virtual model movement is shown in the Figure 2.

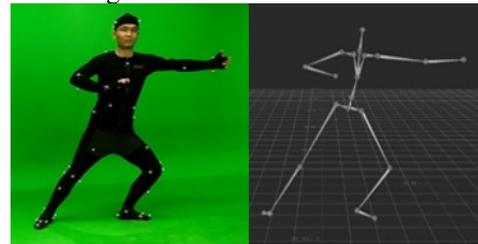


Figure 2. Actor and virtual model.

3.2 Motion capture data format BVH

The data captured by the motion capture describes the human skeleton structure and the motion parameters at various time points.

The most commonly used motion capture data format is the BVH[4] format, which uses a hierarchical motion description method to record human motion data.

Standard BVH file data is stored in frames and consists of two parts: skeleton information and motion data.

The skeleton information section defines the skeleton structure used for the motion, and the motion data section defines the motion data required for the motion [5].

3.3 Human body skeletal model

We use the Motion Builder software to bind the human body model to the skeleton model of the moving actor. The purpose of this is to record the human body's motion trajectory, record the three-dimensional state of the human body motion, and observe the human body movement from different angles through software. The action data facilitates the reuse and research of others. And we use the software to export the BVH format file, parse the content of the file, and build the human body skeleton model.

Based on the self-property of the human skeleton, the skeleton can be regarded as a rigid body, and the skeleton joint, that is, the joint can be regarded as a connecting mechanism. Therefore, when kinematically modeling the articulated skeleton model, the human body can be considered as a motion mechanism composed of multiple links and multiple joints. Human body skeletal model is shown in Figure 3.

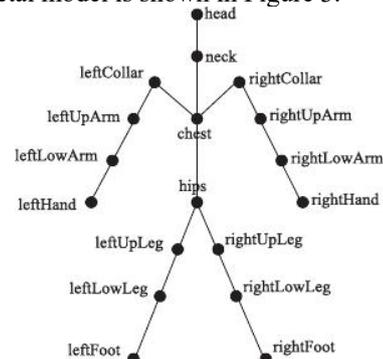


Figure 3. Human body skeletal model.

The human joint skeleton model is described by means of kinematics in the field of robotics [6]. The ankle joint point is the root node of the entire hierarchy and contains six degrees of freedom, namely three rotational degrees of freedom and three degrees of freedom in three-dimensional space [7].

3.4 Quaternion

We use quaternions to represent the posture rotation of the human joints.

Hamilton[4] proposed a method of using quaternions to represent rotation. Quaternion rotation means that the interior consists of four numbers (x, y, z, w) that do not represent angles or axes and represent rotation in three-dimensional space.

Let Q be the four-dimensional vector space on the real number field R, and its orthogonal basis (1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1) use i, j, k, e respectively. means that the elements in Q can be expressed as four-bit vector q:

$$q = w + ix + jy + kz \quad (1)$$

Then the 3D rotation matrix is represented as R:

$$R = \begin{pmatrix} 1 - 2y^2 - 2z^2 & 2xy - 2wz & 2xz + 2wy \\ 2xy + 2wz & 1 - 2x^2 - 2z^2 & 2yz - 2wx \\ 2xz - 2wy & 2yz + 2wx & 1 - 2x^2 - 2y^2 \end{pmatrix}$$

(2)

It is found that the quaternion can completely complete the rotation change in the 3D space. Compared with the rotation matrix and the Euler angle, it has higher operating efficiency than the rotation matrix, and also completely avoids the problem of the universal locking of the Euler angle [4].

The x, y, and z in the BVH file do not correspond to the individual components of R. In other words, the three values of the quaternion x, y, z, and R have no relationship at all, but there is an expression that can be converted.

$q = ((x, y, z), w) = (v, w)$, where v is a vector and w is a real number. Given an Euler rotation (x, y, z), its corresponding quaternion is

$$x = \sin\left(\frac{Y}{2}\right) \sin\left(\frac{Z}{2}\right) \cos\left(\frac{X}{2}\right) + \cos\left(\frac{Y}{2}\right) \cos\left(\frac{Z}{2}\right) \sin\left(\frac{X}{2}\right) \quad (3)$$

$$y = \sin\left(\frac{Y}{2}\right) \cos\left(\frac{Z}{2}\right) \cos\left(\frac{X}{2}\right) + \cos\left(\frac{Y}{2}\right) \sin\left(\frac{Z}{2}\right) \sin\left(\frac{X}{2}\right) \quad (4)$$

$$z = \cos\left(\frac{Y}{2}\right) \sin\left(\frac{Z}{2}\right) \cos\left(\frac{X}{2}\right) - \sin\left(\frac{Y}{2}\right) \cos\left(\frac{Z}{2}\right) \sin\left(\frac{X}{2}\right) \quad (5)$$

$$w = \cos\left(\frac{Y}{2}\right) \cos\left(\frac{Z}{2}\right) \cos\left(\frac{X}{2}\right) - \sin\left(\frac{Y}{2}\right) \sin\left(\frac{Z}{2}\right) \sin\left(\frac{X}{2}\right) \quad (6)$$

$$q = ((x, y, z), w) \quad (7)$$

Through the above method, the quaternion of the human body rotating motion can be obtained, and further processing is required thereafter.

4 Action scoring algorithm

The DTW algorithm was proposed by Japanese scholar Itakura [11] and is widely used in non-equal time series matching, such as speech recognition and motion recognition.

The behavioral recognition using the DTW algorithm is essentially template matching. It is assumed that the training sample (standard template) TR and the test sample (to be matched sample) TE are:

$$TR = (r_1, r_2, \dots, r_m), \quad TE = (e_1, e_2, \dots, e_m)$$

To achieve the matching of two unequal-length sample sequences, first calculate the similarity distance $d(r_n, e_m)$ between any two frames in the two samples, and obtain the similar distance matrix $dN * M$, and obtain the final similarity distance of the two samples. $Dist[TR, TE]$. The DTW algorithm is simple, but it has a large amount of computation and low computational efficiency. Therefore, there are many improved algorithms for its computational complexity.

The low computational efficiency of the DTW algorithm matching process is mainly reflected in the calculation of similar distance and the matching of the optimal path. We use Zhang Youmei's improved DTW algorithm for "one-time planning, secondary refinement" for these two aspects [10]. The algorithm improves the drawbacks of the traditional DTW algorithm, and improves the recognition efficiency while improving the matching and recognition accuracy.

5 System software implementation

We use MySQL to establish a relational database, classify the collected human skeleton information and build a database based on the human body hierarchical relationship model. The established database is mainly used as a training tool for learning reference and our software to make calls.

Our system software is produced by QT5.10. The software interface is simple and intuitive, easy to operate, and the data effect is clear. The improved DTW algorithm is used to perform an alignment of the standard action and the training action skeleton motion sequence, and is scored in conjunction with the time matching of the action. Software interface is shown in the Figure4.

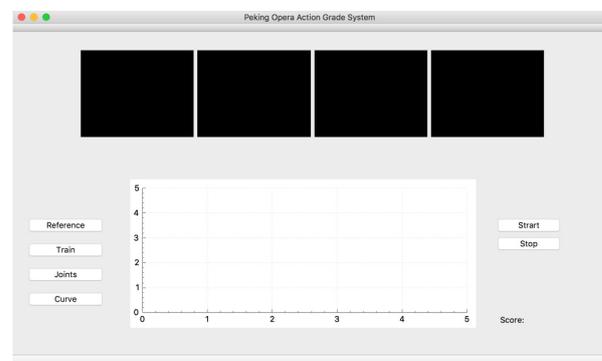


Figure 4. Software interface.
 One of the score results is shown in the Figure 5.

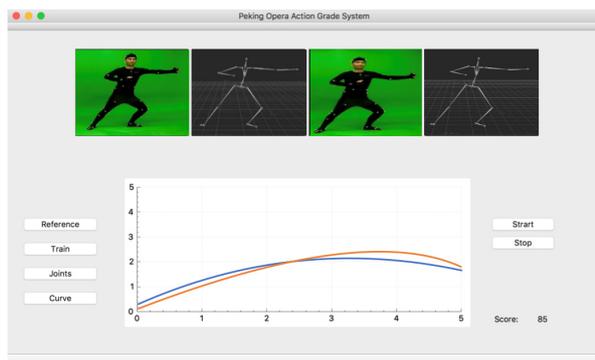


Figure 5. Score result.

The selected action is to raise the hand to the left, the trainer's arm is not straight enough, and the action time is inconsistent with the reference action time, resulting in a score of only 85 points.

6 Conclusion

The functions that the software can implement include: selecting reference training actions, selecting training parts, displaying the sequence of joint rotation changes, displaying the final score, and so on.

The whole system can basically achieve real-time scoring, which provides a good reference for Peking Opera performance exercise trainers.

References

1. Q. Cheng, Research on Digital Protection of Lingnan Dance Based on Motion Capture Technology, 8, 35-37(2012)
2. W. Wang. Digital Development of Minority Dance Based on Motion Capture Technology, 11(2017)
3. Y. Shen, Protection of National Dance Based on Motion Capture Technology, 20(2013)
4. Q. Zhang, Study on the rapid assembly simulation of aero-engine equipment disassembly, 3, 35(2018)
5. Y. Nie, Motion Simplification Based on BVH File, 2, 33(2009)
6. R. Murry, Z. Li and S. Sastry. A mathematical introduction to robotic manipulation (1993)
7. S. Huang, Human motion division and matching based on three-dimensional skeleton model, 6(2013)
8. H. Li, Virtual character control method based on bone information, 1,28(2016)
9. S. La, A Kinect-based interface to animate virtual characters,7,4(2013)
10. Y. Zhang. Motion recognition based on 3D human skeleton, 4,45(2017)
11. Itakura F. Minimum prediction residual principle applied to speech recognition, 23, 1, 67 -72(1975)
12. C. S, sDTW: computing DTW distances using locally relevant constraints based on salient feature alignments, 5(2012)
13. X. Ruan, Dynamic gesture recognition based on improved DTW algorithm, 2134 -2138(2015)