www.cadanda.com Computer-Aided Design AND APPLICATIONS

# Three-dimensional Digitizing of Modern Dance Based on Kinect Motion Capture System

Rujing Yao[1]

[1]Academy of Art and Design, Anhui University of Technology, China

Corresponding author: Rujing Yao, lomoyao@126.com

**Abstract.** The study of modern dance Three-dimensional motion capture data has a comprehensive acquisition method, and this method guides the Three-dimensional data acquisition of dance. The limb movement and facial expression capture system achieve the optimal balance point, and are widely used in animation, film production, virtual reality and other fields. Ethnic characteristics and supplementing with bone binding, realizing the goal of the animation of the dance motion-driven data model, and the efficiency of animation production is obtained. Research shows that motion capture technology can be used to protect modern dance art culture in China. The promotion of modern dance three-dimensional digital method based on Kinect's motion capture technology will produce positive social and economic benefits. In order to verify the feasibility and correctness of the Kinect-based motion capture system. The test results show that the work done can meet the expected goals.

**Keywords:** motion capture; Kinect; 3D human and facial expressions; real-time motion capture.
**DOI:** https://doi.org/10.14733/cadaps.2020.S2.145-157

## 1 INTRODUCTION

Modern dance is an important part of intangible cultural heritage. It is imperative that science be effectively protected and passed down through different means. Traditional methods of photography, recording, and video recording, although convenient to collect and easy to produce, do not allow for detailed recording of the dancer's movements, but they cannot talk about the scientific analysis of the dancers and apply them to choreographic creation. Therefore, aiming at the insufficiency of the two-dimensional recording and preservation of text, photographs, and video technology, proceeding from the characteristics of dance, motion capture technology is applied to protect the modern dance of our country in all aspects of three-dimensional digitization and truly preserve the artistic essence of each dance. It provides a precise digital platform for the protection research, creation, and film restoration of modern dance in the future. It opens up a new field for the protection of intangible cultural heritage and has far-reaching significance for the protection and scientific development of ethnic culture.

Although there have been researches on the protection of individual dances using motion capture technology, there is no unified normative technical process, and the three-dimensional digitization of modern dance has important guiding value for promoting three-dimensional digital protection of dance, relevant results proposed by Kim et al. [1, 2]. Therefore, this paper studies the process of dance three-dimensional digital protection based on motion capture technology and guides the collection of hand-swinging dances. At the same time, it also gives concrete application examples of the application of motion capture data animation.

## 2 HUMAN INFORMATION COLLECTION BASED ON KINECT

### 2.1 Introduction to Kinect

Microsoft launched Kinect Sensory Device in 2010, which consists of three independent motherboards. These three independent motherboards include the electronic components used by Kinect to process data, such as gravity accelerometer, data frame controller, 3D sensor processor, sound source pre-amplifier, USB data integration interface and data stream processor. Figure 1 shows the detailed hardware structure of three independent motherboards.

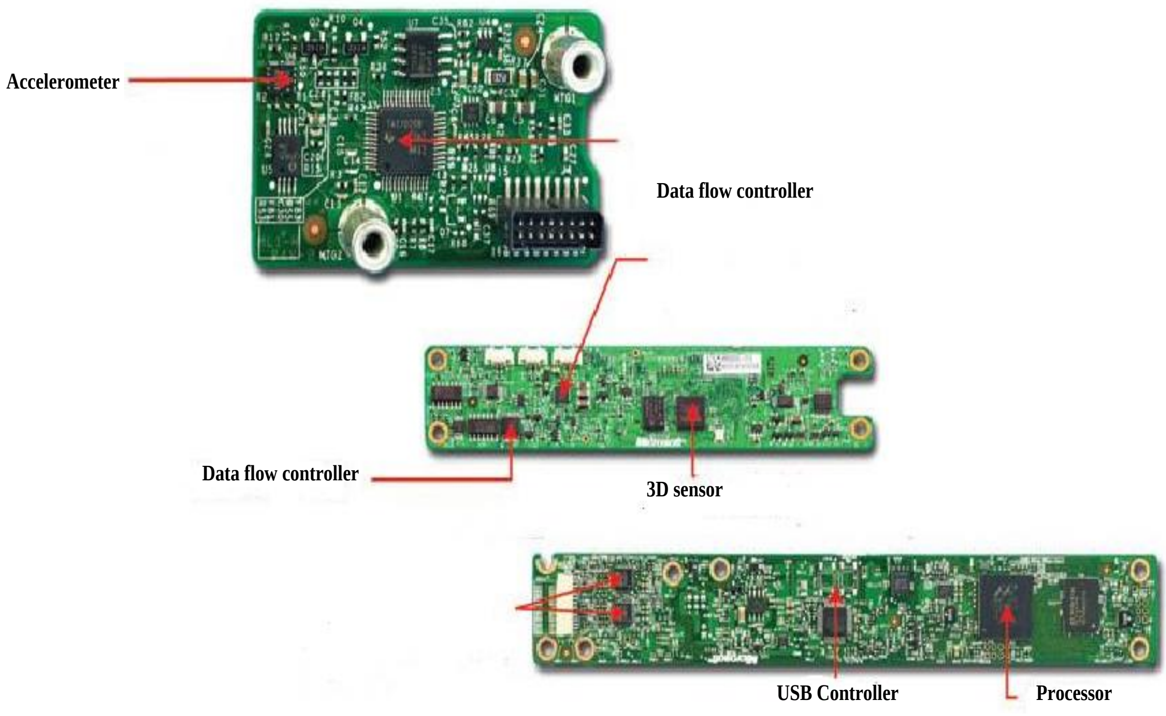


**Figure 1:** Kinect internal motherboard composition.

In the hardware of Kinect, the microphone array uses quaternion linear technology, which can capture multi-channel stereo sound and judge the location of the sound source. The infrared projector uses a common laser source to emit near infrared light with a wavelength of 830 nm after passing through ground glass and infrared filters. Infrared camera is constructed for CMOS, which receives speckle image, recognizes the depth field of the target, and transmits the infrared coding of the scene to PS1080 chip. The internal logic of PS1080 chip can synchronize color data stream, depth image stream and audio stream. The process of collecting depth information is accomplished on this hardware. The personal computer only needs to process USB communication program, which greatly reduces the requirement of hardware resources and improves the speed of software processing. The model of color camera is VNA38209015CMOS, which is used to collect color video images. Angle control motor can be programmed to adjust the camera pitch angle to obtain the best angle of view, to ensure that the test is fully captured, and the base motor can vertically adjust the positive and negative 28 degrees. Kinect uses PD720114 USB cable of Japan Electric Company as data integration interface. Its main controllers include low voltage stepping and single/double mainstream motor driver, on-chip system AP102 for controlling camera, imaging processor and audio controller. USB cable is mainly used to transmit color video data, depth data, audio data, and also can provide external power for PS1080 chip.

Kinect's infrared camera is a structured light measuring device based on light coding. Its basic principle is that infrared projection equipment projects laser scattering spots with three-dimensional depth on the surface of the object. The infrared camera captures the spatial information of the object by receiving the shape of structured light reflected by the object. The main principle is that the change of the object's surface makes the reflected light deform, so the position of the object can be calculated by changing the degree of the spot.

## 2.2 Data Stream Processing

Kinect can provide three kinds of data sources: original audio data, color video stream and depth data stream, which correspond to speech recognition, identity recognition and skeleton tracking in turn, as shown in Figure 2.

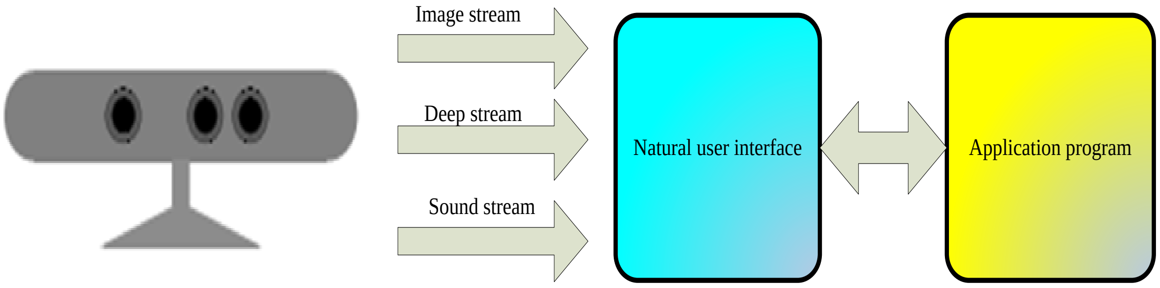


**Figure 2:** Data stream processing.

Kinect passes the original data to the application data development kit and develops the application on this basis. All the applications in this article are developed using SDK version 1.8. Image data information is a series of continuous static image sets. Kinect starts relevant modes in the initial stage according to the image type, resolution and buffer size set by the program. Applications can access color data streams. Users can segment data streams and deep data streams. High-quality color data requires larger buffer space and thus lower transmission rate. By setting up data buffer on Kinect device, the application program can obtain color image information, bone information and depth image information by calling related programs. When the data collected by Kinect is ready, it is immediately transmitted to the buffer. The application uses two ways to obtain the data stream, including polling mode and event mode. In polling mode, the image frame data is first opened, and then the data source is checked periodically for new data. If the data frame is returned to the successful application, the corresponding operation is performed. In event mode, the program registers the FrameReady event of the data frame in the Kinect initialization phase, and gets the data flow information whenever the event triggers.

Each depth image pixel captured by Kinect consists of 16 bits, with 13 bits high representing the distance between the captured object and the device, and 3 bits low representing the captured user index number. The depth data of the subject can be extracted from the original data by accessing the partitioned data. The form of the partition and storage of the depth data is shown in Figure 3. 0 means that no person has been found, and the other numbers 1 to 7 indicate that person 1 to 7. In principle, seven subjects can be tracked, but Microsoft's development kit only allows access to six subjects' information.

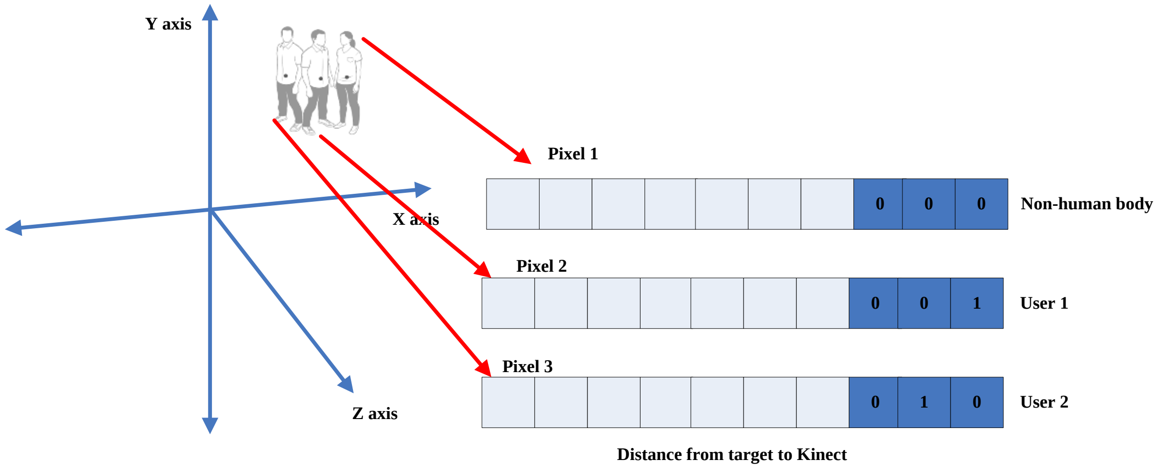


**Figure 3:** Deep data storage form.

Because human motion has the characteristics of time-varying, multi-dimensional and high degree of freedom, the corresponding motion data collected are highly complex. If all the change information of human body is captured, the data will be very huge. In order to effectively use human skeletal information, the program stores only the three-dimensional coordinates of the root node in the skeleton, while the remaining joint location information is calculated according to the translation transformation and rotation transformation matrix relative to the parent node. The method of matrix transformation can reduce the storage space of data acquisition, simplify the way of data organization, and improve the efficiency of operation.

For hardware performance and human-computer interaction, Kinect can track six human skeleton information at the same time. It consists of two human joint points and four human position information, as shown in Figure 4. Through the above user segmentation technology, the depth information of human body can be separated from the original data, and then the mapping between different coordinate systems can be realized. The color image of human body can be obtained from the original color image.

## 3 METHODOLOGY

### 3.1 Three-dimensional Data Acquisition of Ethnic Dance

The Kinect sensor is an RGB-D sensor, a sensor that can simultaneously obtain ambient values (RGB) and depth values. Its fast acquisition speed, high precision and low price make it quickly applied to many fields. The field of robotics has also begun extensive research on Kinect sensors. Mirosława Used the Kinect sensor to perform 3D reconstruction of the indoor environment, obtaining the 3D point cloud model of the environment is one [3]. Depth maps taken by Kinect have high resolution and low cost. Kinect Fusion's function is to capture and build real-life 3D models in real time, in other words, to turn Kinect into a 3D scanner, as shown in Figure 5.

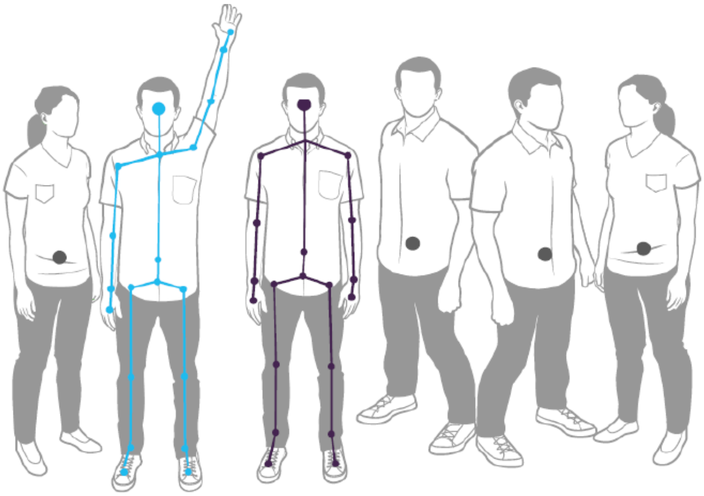

**Figure 4:** Tracking skeleton maps.

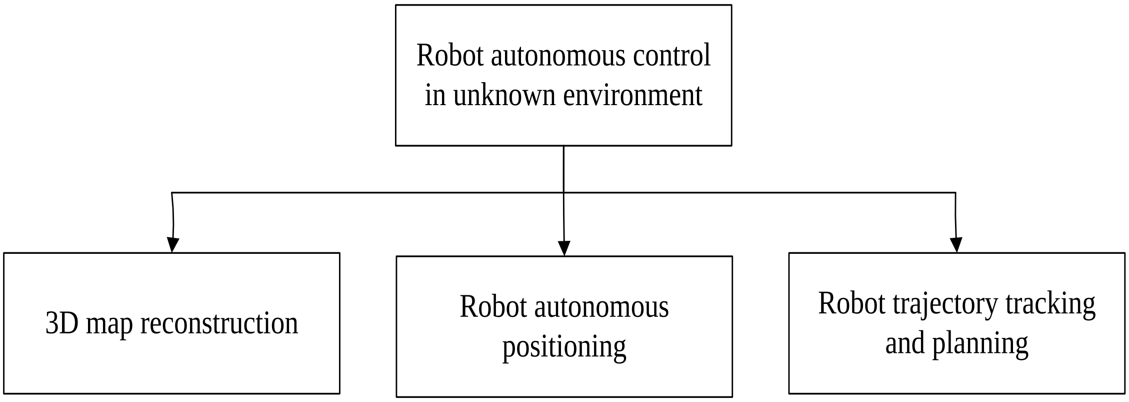


**Figure 5:** Kinect's real-time capture.

Compared to these human motion capture systems that rely on specific hardware devices, the system combined with Kinect's human motion capture system does not require any special hardware. It only uses 3D cameras to capture human motion. This reduces system hardware costs and debugging. Maintenance costs. The most important is the ease of use, the smoothness of the effects, and the ability to observe motion capture in real time. Therefore, revising the human motion capture system will have broad application prospects in the advanced interactive interfaces of video and video games, movie industry, and personal computers.

Zhu et al. introduced Current motion-capturing devices require a large number of infrared devices and body sensing points for body positioning [4-6]. Therefore, the need for smooth and accurate motion presentation often requires a lot of space and expensive funds. In order to use the Microsoft XBox360 Kinect as a motion capture device, most of the motion recognition can be achieved in terms of accuracy, and the technical threshold for developing motion capture can be more widespread. The author combines the depth value and skeleton recognition with each other to achieve similar effects to the traditional motion capture system. The world coordinate system is defined as the camera coordinate system of the first frame, k(u) is the vertex coordinate of the u point in the current frame, and the u point is the corresponding point in the predicted frame, normal to point. Obtain the best relative pose by formula (1).

$$v_i(u) = D_i(u)K^{-1}[u,1] \tag{1}$$

A linear method is used to convert the optimization problem into a least-squares optimization. An optimal solution is calculated by computing a linear equation group such as equation (2).

$$n_i(u) = (v_i(x+1, y) - v_i(x, y)) * (v_i(x, y+1) - v_i(x, y)) \tag{2}$$

Monnet et al. Introduced that the point-to-plane error mechanism is used to measure the accuracy of the current relative pose. In the 3-dimensional case, the error between x and y is the distance d from the tangent of x to y points, and then the two point clouds pass through the camera center of k [7]. Projecting onto the image plane, the point in the two point clouds with the same projection point on the image plane is the corresponding point algorithm. The corresponding point is also filtered by the Euclidean distance between the corresponding points and the normal direction angle, as follows Formula 3 shows.

$$v_i^g(u) = T_i v_i(u) \text{ and } n_i^g(u) = R_i n_i(u) \tag{3}$$

Since these three-dimensional coordinate data come from the actors' real actions, the animation generated using this data looks very realistic and natural. Using the above method, the author decomposed the dance movement according to the dance beat, and placed 29 “Marker” points on the major joints of the dancers according to the scale of the motion capture equipment, and successfully collected the hand dances in the national non-material cultural heritage list [8]. Figure 6 is a three-dimensional recording of "Marker" points on dancers in real time using motion capture software data, a dancer with "Marker" points recording dance moves in real time through a motion capture device.

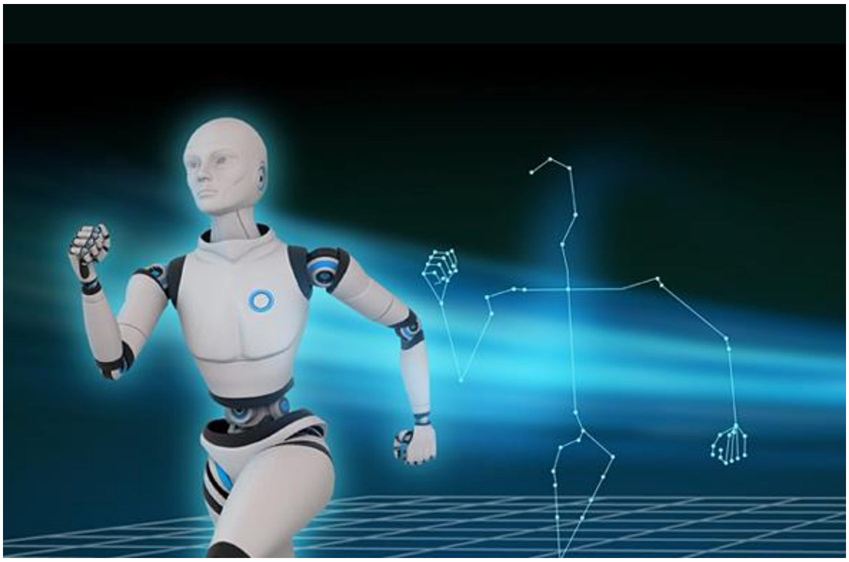

**Figure 6:** Three-dimensional data acquisition.

### 3.2 Motion Capture Data Post Processing

Chen et al. introduced that the construction and implementation of the integrated service platform for the smart campus of higher vocational colleges based on cloud computing and the internet of things technology may cause confusion and obstruction due to the "Marker" points of different parts of the human body [9]. The three-dimensional dynamic capture data collected directly is the original data. Data repair can be completed through two channels: The first is to use the post-processing function of the acquisition device's own software to remove noise points and smooth movement trajectories. The second is the use of animation software frame-by-frame or on the key frame to fine-tune data to achieve data repair, that is, according to the animation curve will be one by one to correct the error in the action, and to ensure the fluency and accuracy of the action, as shown in Figure 7.

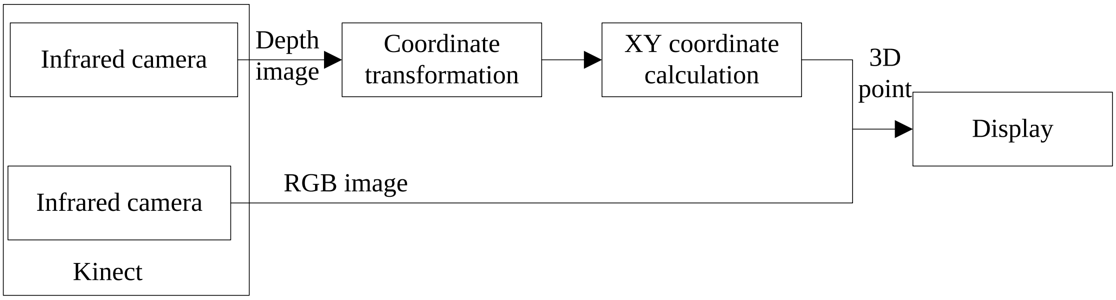


**Figure 7:** Three-dimensional data acquisition.

The theory of Light Coding is to use continuous light (near-infrared) to encode the measurement space. The sensor reads the coded light and passes it to the chip for operation and decoding to generate an image with depth. In the final analysis, Light Coding technology is still structured light technology. However, unlike the traditional structured light method, his source is not a cyclical two-dimensional image coding but a "body coding" with three-dimensional depth. This light source is called a laser speckle and is a random diffraction spot formed when a laser strikes a rough object or penetrates a frosted glass. These speckles have a high degree of randomness, and will change the pattern with distance. The speckles at any two places in the space will be different patterns. It is equivalent to marking the entire space, so any object enters the space, you can accurately record the position of the object.

## 4 RESULTS AND DISCUSSION

### 4.1 Create a Role Model

The modern dance three-dimensional data obtained through the use of motion capture technology can be used to construct a few modern dance databases, choreographed choreographies, and interactive games. At the same time, it can also greatly increase the level of animation production and the efficiency of animation production and reduce costs. Make the animation process more intuitive and more vivid. A flow chart of 3D character animation generation based on human motion capture data has been proposed by Chen [10]. This figure shows the whole process from the acquisition of modern dance 3D motion data to character animation generation. The 3D motion capture data of modern dance and the character model with bone binding are used for motion data binding to generate the character animation.

The system design scheme uses a grating located in front of the camera through the Kinect somatosensory camera to evenly disperse the coded infrared rays in the measurement space. The infrared camera is used to record each speckle with spatial marks. The original data is calculated by the built-in chip. As image data with 3D depth, as long as the object has a movement change in the space, the amount of displacement change can be completely recorded.

The depth information is converted into 3D image data for conversion to the skeletal system. The system can detect up to 6 people at the same time in the measurement space. For each person's skeletal features, 20 sets of data can be recorded, including body torso, hand and foot extremities and fingers and other tracking data. These data will be used to achieve full body sensation. Microsoft uses machine learning function Machine Learning to understand the user's actions, establish an image database to form intelligent identification capabilities, and improve the accuracy of human motion as much as possible，as shown in Figure 8.

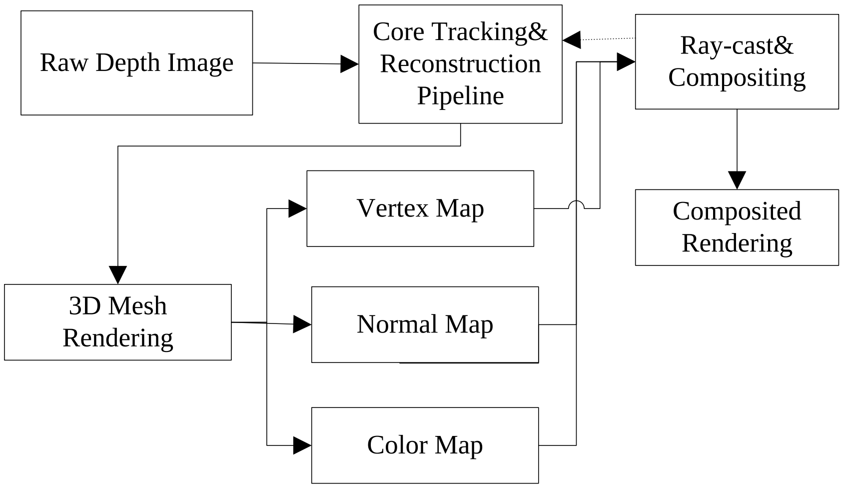


**Figure 8:** The position of the dark side from the light passing through the pixel.

Motion capture data ultimately requires the use of specific models as vectors to show the spatial performance of the action. Accurate and appropriate models have a significant impact on the binding of motion capture data, and thus the establishment of the role model itself is also a very important process. In this paper, with the background of ethnic dance, a realistic personage model with ethnic characteristics has been established. The software used here is Autodesk 3DS Max, and the role is the image of Tujia dancers. The first step in human is to start with a simple geometric shape and gradually add lines and nodes so that the model gradually approaches the expected image. Since the model will be danced. If you can't deal with the relationship between clothing and human body and bone properly, it will affect the texture afterwards, making the texture appear cracking or warping and other errors. The clothing here is relatively tight and fits as close to the body as possible to avoid the problem of dress management in dance movements. At the same time, in order to highlight the ethnic characteristics of the Tujia nationality, the data was captured in a gesture that fits the dance of the Tujia nationality. A round hat was added to the model. The model at this time is only a completely static model. Important bone binding is required. The biped skeletal system in Autodesk 3DS Max is selected here.

Biped is meant to be bipedal and is used extensively for human and other bipedal walking characters, and the biped skeletal system can be easily combined with Actor in Autodesk Motion-Builder.

### 4.2 The Binding of Character Models and Dance Movement Data

Kinect emits class 1 laser light, which is invisible to the human eye, using infrared rays. The laser light is evenly projected in the measurement space through the diffuser in front of the lens, and each speckle in the space is recorded by an infrared CMOS sensor. In combination with the original speckle pattern, it is then transmitted through the wafer as an image with a 3D depth. The sensor senses the environment through black and white spectra. Pure black represents infinity, pure white represents infinity, and black and white corresponds to the physical distance from the object to the sensor.

When Kinect collects data, the data may be mixed with noise due to factors such as the environment, light, and the Kinect device itself; the joints will be lost due to blockage. Therefore, before the model is driven, the data must be optimized, including noise reduction smoothing and lost joint repair processing. For the problems such as joint loss and limb extremity jitter, this topic proposes a lost joint repair algorithm and a skeleton motion smoothing algorithm. Lost joint repair algorithms and skeletal motion smoothing algorithms all use skeletal joint prediction values. Finally, it explains how to use the skeleton motion smoothing algorithm to solve the end joint jitter problem.

Skeletal joint prediction is based on several frames of joint data captured in the previous period as a training sample to estimate the joint position information of the next frame. This project uses an algorithm that Microsoft officially recommends. The algorithm uses the n-order derivative of the Taylor formula (without the remainder) to derive the skeleton joint prediction formula, as shown in Equation 4 below:

$$x_{n+l}=\sum_{i=0}^{N}\frac{f^{(i)}(n)}{i!} \tag{4}$$

n is the number of predictive samples, $x_{n+l}$ is to use nth frame of bone data and its first few frames of data to predict the n+1th frame of bone data. $f^{(i)}(n)$ is the i-order derivative of Taylor's formula. Taylor's first derivative, second derivative, and third derivative can be approximated by the formula in Equation 5, respectively:

$$\begin{cases} f^{(1)}(n)=\mathrm{x}_n-x_{n-1} \\ f^{(2)}(n)=f^{(1)}(n)-f^{(1)}(n-1)=\mathrm{x}_n-2x_{n-1} \\ f^{(3)}(n)=f^{(2)}(n)-f^{(2)}(n-1)=\mathrm{x}_n-3x_{n-1}+3x_{n-2}-x_{n-3} \end{cases} \tag{5}$$

Nx is the input nth frame of bone data. For example, when N is 3, a predictive expression can be obtained as shown in Equation 6 below:

$$\mathrm{x}_{n+l}=\frac{8}{3}x_n-\frac{5}{2}x_{n\text{-}1}+x_{n\text{-}2}\text{-}\frac{1}{6}x_{n\text{-}3} \tag{6}$$

Here only the x coordinate is used as an example and the other two coordinates are the same. This allows the three-dimensional coordinate position of the next skeleton point to be predicted from the previous frames of data. In view of the real-time requirements and accuracy requirements, the system predicts the sample size is set to 7 (N=7, the prediction formula is too long, not given here). When the prediction sample size is too large, the system is time-consuming; the prediction sample size is too small and the prediction value is not accurate enough.

When the action performer's limbs occlude each other, some joints cannot be positioned by the Kinect and the joints are lost. Using the missing joint's movement data to drive the model will cause the model to be distorted, so it is necessary to repair the missing joints. The application scenario of this algorithm is that both the parent node and the child node of the missing joint point exist. The left shoulder joint point A and the left wrist joint point C exist, and the middle left elbow joint point B is lost. During the movement of the performer, the lengths of the bone lengths AB and CB are fixed, the lengths are LAB and LCB, the ball aS with radius LAB is centered on point C, and The LCB is the radius of the ball cS intersecting the circle on bC, as shown in Figure 9.

Under the idea of infinite differential, it can be approximated that each joint of the human body performs a uniform linear motion in a very short time. Assume that the position information of the elbow joint before the k-th frame exists, and the k-th frame of the k-th frame can be used to estimate the elbow joint k-1 frame elbow motion speed.

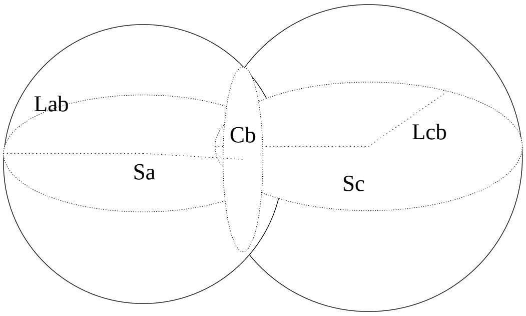


**Figure 9:** The original algorithm determines the range of B-points.

## 4.3 Experimental Results and Analysis

After completing the skeleton model with the bone binding, you can use the three-dimensional motion data of the folk dance to drive the skeleton-binding character model. The concrete step of the motion data binding is to first import the collected three-dimensional motion capture data of ethnic dances into Autodesk and match it with the human shape model Actor provided by the software to match the dimensions, proportions, and angles. The Actor is the same as the body size and initial movement of the actor performing the corresponding action. This will make the actions after the motion capture data and model binding more accurate and natural. Then in the software, the various tracking points in the capture data are mapped with the key nodes on the corresponding Actor's skeleton, so that the Actor's bones can be moved by the spatial position data of the tracking points in the motion capture data. After the exercise binding and debugging on the Actor is over, you can build it. The actual movement of the dancer and the resulting animation are compared as shown in Figure 10.

The character model includes two parts of the model skin and skeletal system. The model skin consists of several polygons. The vertices of each polygon are called grid vertices. The premise of the low-distortion motion of the skeletal data-driven character model is that the character model needs a skeletal structure that matches the Kinect skeleton. The standard is similar to the Kinect skeleton hierarchy, and is based on joint points, which use a tree structure to layer the skeleton. In addition to the root joint, each joint has a parent node, and every two joint points in the character model have specific bones connected to each other. The principle of movement of the model is that the skeleton takes the joint point as a fulcrum, performs a constrained rotational movement, and drives the next layer of joint motion.


**Figure 10:** Comparison of dancer's actual movement and generated animation.

Before driving the character model, first bind the bone joint points acquired by Kinect with the skeletal points of the character model. Kinect V2 can locate the 10 joint points of the human body. In this project, the number of human bone joints created is 10, and when binding, the extra joint points in Kinect are ignored. When binding, the joint details are shown in Table. 1.

| Kinect V2 positioned joints | Model joint variables | Speed ratio |
|---|---|---|
| Joint 0 (Spine Base) | HIP | 45 μm sieve residue 14.16% |
| Joint 1(Spine Mid) | NECK | 45 μm sieve residue 23.72% |
| Joint 2(Neck) | HEAD | Fineness modulus Medium sand |
| Joint 3 (Head) | SHOULDER_LEFT | Length 60 mm Diameter 75 μm |
| Joint 4 (Shoulder Left) | ELBOW_LEFT | Length 60 mm Diameter 75 μm |
| Joint 5 (Elbow Left) | WRIST_LEFT | Length 9 mm Diameter 18 μm |
| Joint 6 (Wrist Left) | HAND_LEFT | Length 45 μm Diameter 74 μm |
| Joint 7 (Hand Left) | SHOULDER_RIGHT | Length 45 μm Diameter 74 μm |
| Joint 8 (Shoulder Right) | ELBOW_RIGHT | Length 60 mm |

| | | |
|---|---|---|
| | | Diameter 75 μm |
| Joint 9 (Elbow Right) | WRIST_RIGHT | Length 60 mm<br>Diameter 75 μm |
| Joint 10 (Wrist Right) | HAND_RIGHT | 45 μm sieve residue<br>23.72% |

**Table 1:** Properties of raw materials.

Determine if the node is a root node. If it is a root node, assign its coordinate position to the root joint point of the model skeleton, assign the global coordinate system to the initial coordinate matrix of the model coordinate system, and assign the root node to the quaternary element. The corresponding rotation matrix is assigned to the transition matrix. The quaternion-to-rotation matrix conversion process uses the get Matrix function to complete the contrast between the dancer's actual dance movement and the resulting animation to test whether the motion capture technique performs a full and accurate record of modern dance movements. The figure shows the modern dance three-dimensional digitization method described in this article. It uses the optical motion capture device produced by Devil Motion Analysis Inc. in the United States to collect a frame of action of a dancer with 10 "Marker" points in real time.

In this way, three-dimensional data corresponding to the dance of the frame can be obtained. According to the animation generation method described in this article, a corresponding animation frame effect generated by a three-dimensional dynamic capture data driven character model is obtained, and the animation effect is consistent with the actual dance movement. The experimental results show that the motion capture technology truly records the actor's dance moves, and the three-dimensional data can be used for animation to improve the animation production efficiency.

## 5 CONCLUSION

The three-dimensional data of modern dance has played an active guiding role in the protection of dance culture and the integration of resources, and has played a positive guiding role in research and promotion and application of contemporary digital means in the protection of dance art. It is a product of the mutual promotion of culture and science and technology and the idea of common development. It helps to promote the construction of cultural digital construction. Taking the Kinect somatosensory camera as the core of the system, the limb movement and facial expression capture system achieves the best balance between development costs, operational effects, and development efficiency. Through the application of motion capture technology, it is possible to effectively avoid the incompleteness of the information in the conventional impact recording method. The problem can also provide more analysis data for modern dance choreography to obtain the original entity of the dance posture that can be accurately reproduced by traditional methods, reproduce the artistic style of the year, and have a short replication cycle and low cost. It is a very There are promising artistic protection technologies. At the same time, the realization of three-dimensional digitization of modern dance can effectively promote the integration and optimization of modern dance culture resources and play an important role in promoting the sustainable development of modern dance culture.

*Rujing Yao*, https://orcid.org/0000-0002-1687-039

**REFERENCES**


[1] Kim, Y.: Dance motion capture and composition using multiple RGB and depth sensors, International Journal of Distributed Sensor Networks, 13(2), 2017, 155-183. https://doi.org/10.4218/etrij.17.2816.0045.

[2] Chang, Y. J.; Chen, S. F.; Huang, J. D. A.: Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities, Research in developmental disabilities, 32(6), 2011, 2566-2570. https://doi.org/10.1145/2049536.2049627.

[3] Mirosława, M. D.; Panek, D.; Paweł M.; Wiesław C.; Alda, W.: An improved kinematic model of the spine for three-dimensional motion analysis in the vicon system, Studies in health technology and informatics, 176, 2012, 227-231. https://doi.org/10.3233/978-1-61499-067-3-227.

[4] Zhu, Y.; Li, Ti.; Taoruan, W.: Automatic polygon mesh repair and simplification for three-dimensional human modeling, Modern Physics Letters B, 31(19-21), 2017, 1740077. https://doi.org/10.1142/S0217984917400772.

[5] Chang, X.: Feature interaction augmented sparse learning for fast kinect motion detection, IEEE Transactions on Image Processing, 26(8), 2017, 3911-3920. https://doi.org/10.1109/TIP.2017.2708506.

[6] Yoshimura, M.; Murasato, H.; Kai, T.; Kuromiya, A.; Yokoyama, K.; Hachimura, K.: Analysis of Japanese dance movements using motion capture system, Systems and Computers in Japan, 37(1), 2006, 71-82. https://doi.org/10.1002/scj.20250.

[7] Monnet, T.; Samson, M.; Bernard, A.; David, L.; Lacouture, P.: Measurement of three-dimensional hand kinematics during swimming with a motion capture system: a feasibility study, Sports Engineering, 17(3), 2014, 171-181. https://doi.org/10.1007/s12283-014-0152-4.

[8] Ji-Su, Y.; Jae-Seop, O.: Greater trochanter location measurement using a three-dimensional motion capture system during prone hip extension, Journal of Physical Therapy Science, 29(2), 2017, 250-254. https://doi.org/10.1589/jpts.29.250.

[9] Chen, C.; Roozbeh J.; Nasser K.: A survey of depth and inertial sensor fusion for human action recognition, Multimedia Tools and Applications, 76(3), 2017, 4405-4425. https://doi.org/10.1007/s11042-015-3177-1.

[10] Chen, G.: Three-dimensional reconstruction of digitized human liver: based on Chinese Visible Human, Chinese Medical Journal, 123(2), 2010, 146-150. https://doi.org/10.3760/cma.j.issn.0366-6999.2010.02.004.