International Journal for Housing Science and Its Applications
Publish August 4, 2025. Volume 46, Issue 3 Pages 2296-2305

https://doi.org/10.70517/ijhsa463192

Design and Application of a Motion Capture Analysis System for Intangible Cultural Heritage in Sports: A Case Study of Traditional Martial Arts

Haoxiang Xie[1,*]
[1] College of History and Culture,Chengdu Sports University, Chengdu, Sichuan, 641418, China
Corresponding authors: (e-mail: frankenstein510@163.com).

Abstract As an important component of China's intangible cultural heritage, the standardization of movements in traditional martial arts is the key to the inheritance of the art. In this study, a motion capture analysis system for traditional martial arts is designed with Kinect motion capture model as the core, combined with Unity3D technology and Dynamic Time Warping (DTW) algorithm. The system can capture human motion data stored in a standard database by Kinect device and provides videos related to traditional martial arts learning. Combined with the DTW algorithm to match the joint angle and velocity characteristics during movement, the automatic scoring of traditional wushu movements was realized. The average three-dimensional motion range of the spine in traditional wushu is between 137±13°, and the average range on two-dimensional recognition is between 138±17°, with better recognition effect on different levels. The correlation between the 3D coordinate-time curves of human joint points obtained by the system analysis of this paper and the manual analysis is strong, and the correlation coefficients are all greater than 0.90. The traditional wushu movement curves measured by this system are basically consistent with the actual movement curves. In addition, on the scoring of traditional wushu movements, the error between the scoring of this system and the scoring of experts was between ±0.2 points. The results of this study promote the digital inheritance of the intangible cultural heritage of sports and have practical value for the dissemination of traditional martial arts.

Index Terms Kinect, Unity3D, DTW algorithm, motion capture system, traditional martial arts

I. Introduction

Intangible cultural heritage is a concept and idea that emerged only when human society entered the period of civilization transition, and it has an inescapable historical mission, i.e. the protection and inheritance of civilization. In the face of the impact of globalization, the Chinese government has strengthened the protection and inheritance of outstanding national cultures through the legal means of setting up the “Intangible Cultural Heritage List”, in which a large number of outstanding traditional sports are included [1], [2]. Through heritageization, these outstanding traditional sports have gained wider attention and support from the government, society and the public.

At the same time, in the change of the great transformation of Chinese society today, the sports NRL is adapting to the social changes, which are breathless, multidimensional and progressive, promoting the ancient traditional culture to better modernization and globalization [3], [4]. At the same time, due to the wide area of China, the very uneven development of different districts and cities, and the coexistence of agricultural civilization, industrial civilization and post-industrial civilization, there are also many peculiarities and differences of recreational non-heritage projects in different regions and levels [5], [6]. As for the inheritance of sports non-heritage, the serious aging of the inheritor group and the lack of digital records have led to the distortion of most of the key movements in the inheritance, such as the palms and punches of wushu [7]-[9]. In this context, the innovative development of sports non-heritage protection and inheritance in sports academia has been promoted. Through wearable devices, multi-sensor fusion technology, motion capture technology, computer vision technology, artificial intelligence, etc., capturing movement data (including movement, physiology, psychology, interaction), constructing a three-dimensional model, systematic analysis of multi-dimensional data, and combining with artificial intelligence-type enhancement and other technologies to strengthen the movement, in order to pass on and protect the sports NRLs with more clarity and authenticity, which will also become the future sports NRL digital protection This will also become a key direction for the future digital protection and inheritance of sports NRLs [10]-[14].

In this paper, we design a sports intangible culture motion capture analysis system and take traditional martial arts as a case study for application. The system utilizes the depth image generation of Kinect motion capture device and the human bone recognition function to obtain the three-dimensional environment position and action information of human movement. Through the Kaiyuan Open NI framework, Kinect is combined with Unity3D engine to realize the 3D motion reconstruction of human joints. Then combined with the dynamic time regularization algorithm, the design of the system scoring module is completed, and the athlete's movement score is determined through the joint angle, speed and other characteristics. The system also carries a video teaching module in order to provide standardized traditional martial arts movement demonstrations.

II. Key technologies

II. A.Kinect Motion Capture Model

II. A. 1) Depth image generation

In the depth image, the gray value of each pixel point contains the depth value corresponding to that point, i.e., the vertical distance from the corresponding point in the object under test to the plane of the Kinect [15] depth camera. As a result, the acquisition of depth information can accurately realize the segmentation of the target object and the scene, so as to separate the front and back view of the test environment and obtain 3D data information.

Time of Flight acquires the distance between the object under test and the focal plane of the infrared emitter d calculated by the formula:

$$d=\frac{1}{2}c\cdot t=\frac{1}{2}c\cdot\frac{\Delta\varphi}{2\pi f} \tag{1}$$

where c is the propagation speed of light in the atmosphere, i.e., $c\approx 3\times 10^{8}m/s$, f is the frequency of modulated light emitted by the infrared emitter, and $\Delta\varphi$ is the phase difference between the emitted and reflected light signals. The propagation time of an optical pulse can be expressed by calculating the phase difference between the two signals. For the unknown quantity $\Delta\varphi$, four equally spaced specific phase samples of the reflected optical signal are selected in one modulation period according to the cross-correlation function between the emitted and reflected signals. $t_i=ft$ ($t_0=0$, $t_1=\frac{1}{4}T$, $t_2=\frac{1}{2}T$, and $t_3=\frac{3}{4}T$), and the optical signal intensity is calculated separately to obtain the $\Delta\varphi$ value.

The light wave signal emitted at the modulation frequency f that has been set is:

$$S(t)=A\cdot\cos\left(ft\right) \tag{2}$$

The reflected light wave signal after propagation is:

$$r(t)=kA\cdot\cos\left(ft+\Delta\varphi\right)+B \tag{3}$$

The expression for the calculation of the cross-correlation function between the light wave signal emitted by the infrared emitter and the reflected light signal:

$$C_i=C\left(t_i\right)=\frac{1}{2}kA\cdot\cos\left(t_i+\Delta\varphi\right)+B \tag{4}$$

In the above expression, A is the amplitude of the emitted cosine wave signal, f is the modulation frequency of the emitted optical signal, k is the attenuation coefficient of the optical wave generated in the process of propagation and reflection of the optical wave, and B is the bias of the signal caused by the addition of background light. The phase difference is calculated based on the intensity of the optical signal at the four selected sampling points as:

$$\Delta\varphi=\arctan\frac{C_3-C_1}{C_2-C_0} \tag{5}$$

From the above calculation process, the depth value of the measured environment space is finally obtained, and the 3D position information of the measured environment can be obtained in real time.

II. A. 2) Human Skeleton Recognition Based on Depth Information

Kinect binarizes the depth information, adjusts the threshold to exclude background images other than human objects, and is the first to recognize the contours of “large” objects. Kinect 2.0 can capture six human objects at the same time, and through edge detection, noise processing and other algorithms to detect the input depth image, and finally the segmented human object outline output.

The construction of the intermediate representation layer of human objects, this stage is the process of recognizing and classifying different body parts of the separated human objects. Through the collection of data volume counted in terabytes of human object data information input into the system training model, the depth of the human body parts feature information is extracted and used to train a classifier containing 32 different parts of the human body features, distinguishing between the human body parts features formula is:

$$f_\theta(l,x) = d_l\left[x + \frac{\mu}{d_l(x)}\right] - d_l\left[x + \frac{v}{d_l(x)}\right] \tag{6}$$

where x is the pixel value of the depth map pixel point, $d_l(x)$ is the depth value of the x pixel value in the depth image l, $\theta = (\mu, v)$ is the lens offset vector, $1/d_l(x)$ is the regularization of the offset for handling body size scaling, and l is used as a marker image. The formula reflects the difference in the three-dimensional shape of the measurement pixel point and its surrounding region, e.g., if the offset pixel is a pixel point in the background image, $d_l(x)$ tends to positive infinity, and thus, depending on the influencing feature, body parts can be effectively distinguished.

II. B.Kinect combined with Unity3D

Unity3D [16] is a comprehensive multi-platform game development tool developed by Unity Technologies that allows developers to easily create interactive content such as 3D video games, architectural visualization, real-time 3D animation, and other types of interactive content, and is a fully-integrated professional game engine.

Calling Kinect in Unity3D is the difficulty of this system, this paper adopts Open NI of Open NI to combine the two.Open NI is an open source framework, there is not much implementation in itself, it mainly provides a set of specifications for human input and some hardware underlying APIs for realizing interfaces between different devices.The official website of Open Ni provides examples of how to develop in Unity3D. Unity3D development examples, on Open NI, need to implement Open NI-compliant modules, these modules use the Open Ni interface, self-implementation, to provide calls for their own top-level applications.

II. C.DTW-based motion capture evaluation algorithm

The Dynamic Time Warping (DTW) algorithm [17], [18] is an algorithm for evaluating the similarity of two time series. By finding the most matching path between two sequences, it efficiently handles local displacements, telescopes, or compressions of time sequences even if they exhibit deformations on the time axis. The steps to realize the motion capture evaluation are as follows:

(1) Create a distance matrix: suppose there are two action sequences of different lengths, the standard sample sequence $S = (s_1, s_2, \cdots, s_m)$ and the test sample sequence $T = (t_1, t_2, \cdots, t_n)$. where m and n represent the lengths of the two sequences, respectively. In order to compare these two sequences, a matrix D of size $m \times n$ is constructed, which is responsible for mapping the relationship between the two sequences. Where each element $D[i,j]$ in the matrix represents the distance between s_i between the standard sample sequences and t_j between the test sample sequences.

(2) Search for the best path: through this matrix, the DTW algorithm searches for a path from the point $(1,1)$ to the point (m,n) that satisfies the constraints and has the smallest total distance, which represents the best alignment between the two sequences.

(3) Calculate the path cost: the cumulative distance of the points on the best path is the DTW distance between two sequences, which reflects the similarity of the two sequences, and the smaller the distance, the more similar the two sequences are.

(4) Normalization: In order to make the DTW distance not affected by the path length, the distance is usually normalized.

The DTW algorithm defines the regularized path D from the starting point $(1,1)$ to the point (m,n) by applying local path constraints that minimize the cumulative mismatches while adhering to specific constraints. The p th element of D is defined as $D_p = (i,j)_p$ with $D = d_1, d_2, \cdots, d_p, \cdots d_P$. where $\max(m,n) \le P < m+n-1$.

The regularized path used by the DTW algorithm in calculating the similarity between two time series is subject to a series of constraints to ensure the rationality of the path and the effectiveness of the algorithm. These constraints mainly include:

(1) Boundary condition: The regular path must start at the first element $D[1,1]$ of the matrix and end at the last element $D[m,n]$, ensuring that the path covers all elements of both sequences.

(2) Continuity constraint: each step of the regularized path can only move in the forward (horizontal, vertical) or forward-right (diagonal) direction. Ensure the continuity of the path and avoid any backward operation, i.e. from any point $D[i,j]$, the next step can only be $D[i+1,j]$, $D[i,j+1]$ or $D[i+1,j+1]$.

(3) Monotonicity constraint: as time advances, the regularized path must move forward monotonically, allowing neither lateral nor vertical jumps back. If $D_k=(i,j)_k$ is a point on the path, then a and b must increase monotonically for all k. This ensures that the sequence maintains its original temporal order.

Subject to satisfying the boundary conditions, continuity and monotonicity, DTW finds the path λ of which minimizes the cost of regularization $\lambda=\min\left\{\sqrt{\sum_{p=1}^{P}\frac{d_p}{P}}\right\}$. Starting at the beginning of the matrix, each subsequent point is an accumulation of the computed distances of the points on the path. When the end point (m,n) is reached, the cumulative distance $D(s_i,t_j)$ is the final distance $D(S,T)$.

III. Experimental data sets

(1) The COCO2017 human pose dataset is part of the MS COCO2017 dataset, which is annotated for training pose estimation models, and includes a total of 58,945 image data, of which 56,599 images in the training set and 2,346 images in the validation set, which is a large-scale dataset that can be used to detect a target and localize its keypoints at the same time.COCO keypoints for the human body are defined as 17 joints, which are nose, left_eye, right_eye, left_ear, right_ ear, left_shoulder, right_shoulder, left_elbow, right_ elbow, left_ wrist, right_ wrist, left_hip, right_ hip, left_knee, right_knee, left_ankle, right_ ankle.

(2) The traditional wushu movement dataset is a collection of movement data from 20 traditional wushu masters in the College of Physical Education of our university. The dataset contains a total of 4,410 images of various traditional wushu movements. The dataset was divided according to the 7:3 ratio. The images were manually labeled by using Labelme annotation tool, and JSON format files were generated and converted into corresponding TXT format files for subsequent model training and validation. The traditional posture dataset for the human body keypoints is defined as 14 joint points, which are head, neck, left_shoulder, right_shoulder, left_elbow, right_ elbow, left_wrist, right_ wrist, left_hip, right_ hip, left_knee, right_knee, left_ankle, right_ ankle.

The experiments in this paper use COCO2017 human posture dataset and traditional martial arts movement dataset, this experiment is conducted under Windows 10 operating system, CUDA version 12.2, Pytorch version 2.0.1, Python version 3.8, the hardware environment is CPU: IntelCore i7-13700F, the GPU uses NVIDIA GeForce RTX 4090 graphics card.

IV. Traditional martial arts motion capture analysis system design

IV. A. Logical framework of the system

In this section, based on the techniques mentioned above, the design of the functionality, internal structure, and development content of the traditional martial arts motion capture analysis system was carried out, which in turn led to the realization of a training system that meets the user's needs. The framework of the traditional martial arts motion capture analysis system is shown in Figure 1.

(1) Video teaching module

The design of the video teaching module mainly provides users with video instruction of standard movements of traditional martial arts, which is convenient for users to learn by watching videos and to choose the traditional martial arts contents they need to learn.

(2) Movement Acquisition Module

The action acquisition module includes data acquisition, data filtering and noise reduction processing, and skeleton data preservation. The collected movement data of traditional Wushu masters are saved after data filtering and noise reduction processing to establish a standard movement database. The collected skeleton data of testers are saved after filtering and noise reduction processing, waiting for the test scoring stage.

(3) Auxiliary Scoring Module

The auxiliary scoring module mainly performs angle feature extraction on the standard skeleton data collected by the action acquisition module respectively. For the test skeleton data, the angular features are extracted and the DTW algorithm matches the corresponding frames. Then the speed features are extracted from the two sets of skeleton data, combined with the previously extracted angular features, to provide an auxiliary judgment basis for the test module, comparing the test skeleton data with the standard skeleton data and judging the scores, and giving suggestions for traditional Wushu training, so as to achieve the purpose of auxiliary training.

(4) Three-dimensional reconstruction module

Three-dimensional animation display module is mainly to obtain the tester's traditional wushu movement data produced into the movement file assigned to the built character model, drive the virtual character, reproduce the animation of traditional wushu students. It makes the traditional martial arts training get rid of the traditional training by virtue of experience, and increases the fun of training.

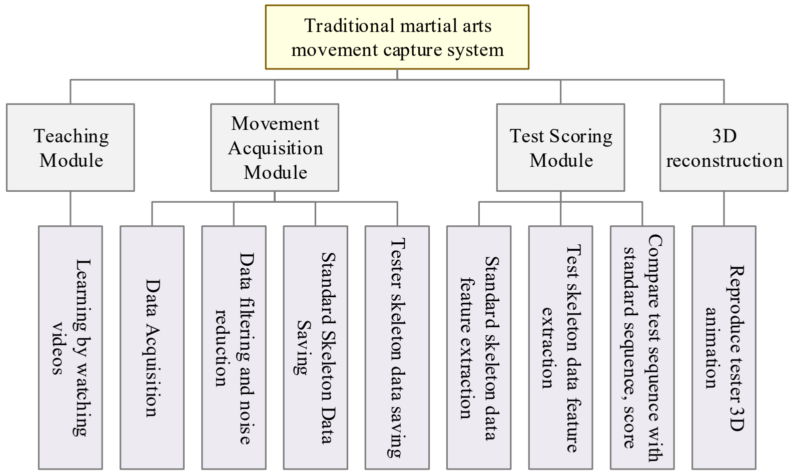

Figure 1: Traditional martial arts movement capture system framework

IV. B. Standardized databases

This system collects motion data through Kinect somatosensory devices that the human body does not need to wear other devices, and records the collected athlete data in the form of virtual data. The key factors for the establishment of the standard database are as follows:

(1) The collection process must select professional traditional martial arts practitioners, and professional traditional martial arts practitioners are selected to act as samples for the data collection of this system, which can make the traditional martial arts movement data relatively standard.

(2) In the process of action data settlement, it is necessary to match the task model and the user debugging, action comparison debugging needs to do three different actions, the use of the three action poses to calibrate the relative position of the body parts and the skeleton of the virtual character in Unity.

(3) In the process of collecting data with Kinect, in order to use the video as a reference when correcting the data at a later stage, it is necessary to collect the whole set of movements of the athlete with Kinect.

IV. C. Video teaching module

According to the characteristics of learning traditional wushu, it is necessary to follow the traditional wushu instructor to improve the level of training in the process of repeated learning, so one of the main functions of this system is to provide standardized traditional wushu movement demonstrations. The teaching module of this system is categorized into traditional video teaching, which is widely applicable, and users only need a computer or a mobile client to learn through the form of video playback.

The teaching of this system also introduces the traditional wushu culture through the form of video playback to include the cultural background of traditional wushu, famous masters, theoretical knowledge, the origin of the practice methods and so on. Users can choose the learning content according to the demand, and through the system interface design button to the control part of the command, the program through the user input instructions to control the video source of the playback, pause, fast forward, fast forward, fast forward and other screen control, to achieve the purpose of learning.

IV. D. System scoring module

The core of this system is the design of the scoring function, which combines DTW matching of corresponding frames with joint angle and velocity characteristics. The angle difference between the tester's action sequence and the standard action sequence is calculated as a basis for scoring. Considering the nature of the movements of traditional martial arts, this design also extracts the velocity features of the movement sequence, and takes the difference in velocity as another basis for assessment. In addition, this system will also give the duration of the traditional martial arts movements, and finally the score will be calculated by comprehensive consideration. The design flowchart of the scoring algorithm is shown in Figure 2.

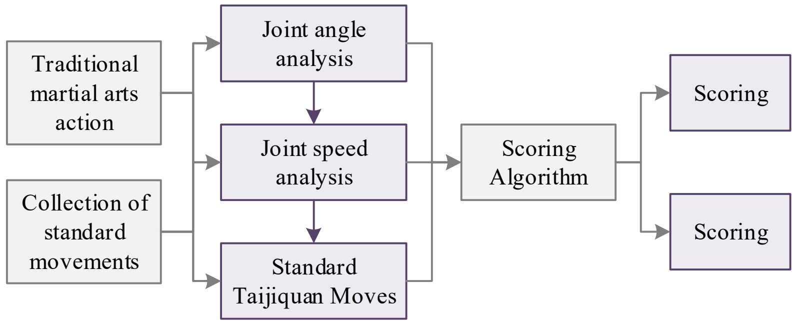

Figure 2: Grading module design process

IV. E. 3D Reconstruction Module

The design of the 3D module tries to solve the problem of how to assign the rotation data of each joint point acquired by Kinect to the corresponding joint point of the Unity3D character model. After acquiring the data of a certain joint point, the data needs to be processed with noise reduction due to Kinect noise. Then for the possible problem that the character pose does not match the physiological structure, the rotation of each joint point is forced to be constrained, which can make the generated model pose more reasonable. Then through the Kinect coordinate system and Unity3D's mirror image compensation and rotation compensation, to realize the reproduction of the three-dimensional movements captured by Kinect. Kinect and Unity module establish communication through the control part of the program, the Unity module mainly refers to the use of graphic design software to build the virtual scene and the motion unit, and the module is encapsulated in the Unity3D engine, data processing and interaction control in the Unity3D engine, which in turn can reproduce the user's 3D animation in real time.

IV. F. Motion Capture Case Study Results and Analysis

IV. F. 1) Measurement of average range of motion in traditional martial arts

Ten experimental subjects were recruited for this study, and each subject had no injuries or surgeries within three months prior to testing and was in good health. Before the test, each subject was informed of the experimental procedure and the potential risks during the experiment, and signed the experimental informed consent form after the subjects agreed to the test. And then changed into motion capture suits to prepare for filming in the venue. Maker reflex markers were attached to the subjects' body joints, and the elbow, wrist, knee and ankle joint centers were assigned to the midpoints of the lateral and medial markers, while the shoulder joint centers were assigned to the midpoints of the anterior and posterior shoulder markers. Three-dimensional motion capture and two-dimensional video filming were performed using signal light as the synchronized filming signal, and synchronized frame calibration was performed at a later stage. According to the video observation method, the locations where typical movements of traditional martial arts mainly appeared in the frame were recorded. Subjects were asked to complete the action filming within a specific orientation within the experiment, with a filming distance of 3 meters in front of the equipment. In order of the most common joints assessed in the study, the spine (trunk), shoulders, wrists, elbows, hips, knees, and ankles, with the upper extremities and trunk having more directional faceted movements than the other joints.

The 3D motion analysis provides an export of skeletal data that includes joint positions at 52 points. In order to analyze the stance angles of typical movements, the 3D motion recognition of 10 athletes was analyzed individually. The average range of motion for 3D recognition of traditional martial arts is shown in Figure 3. Among them, the average range of motion of the spine was 137±13°, and the average range of motion of the upper limbs: 144±15° for the left elbow, 126±14° for the right elbow, 60±20° for the left shoulder, and 134±14° for the right shoulder. The mean range of motion of the lower limbs: left hip 107±13°, right hip 113±17°, left knee 119±18°, right knee 84±16°. The average range of motion for two-dimensional identification of traditional wushu is shown in Figure 4. The average range of motion of the spine is 138±17°, and the average range of motion of the upper limbs: left elbow 127±16°, right elbow 124±14°, left shoulder 60±20°, right shoulder 139±13°. The average range of motion of the lower extremities: left hip 100±20°, right hip 126±10°, left knee 123±16°, right knee 93±13°. The results of the study established that the system in this paper has good recognition of the angles of the joints of traditional Wushu movements in both three-dimensional and two-dimensional motion analysis.

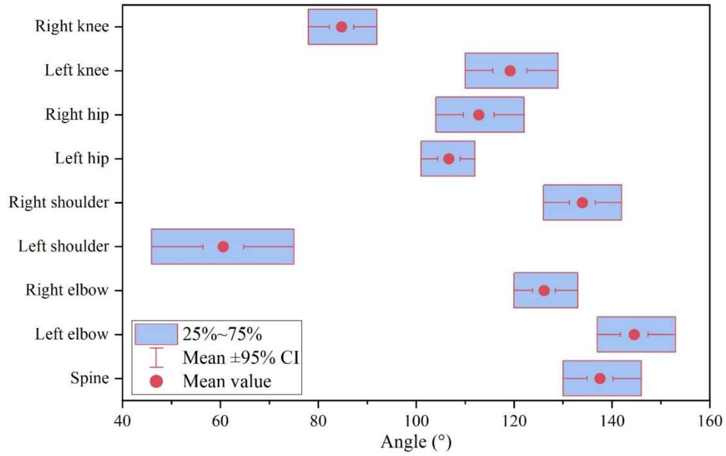

Figure 3: The three-dimensional average of traditional martial arts

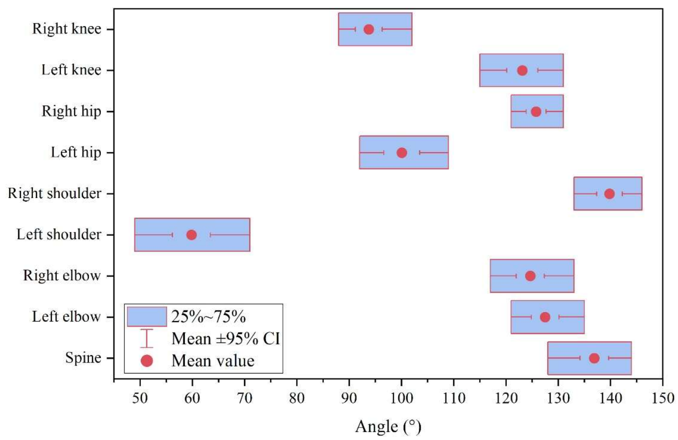

Figure 4: The two-dimensional average of traditional martial arts

IV. F. 2) Validation of the Effectiveness of Traditional Wushu Motion Capture

The system in this paper automatically analyzes the video of a traditional martial arts simulation match, obtains the 2D coordinates of 18 joint points of the human body in 3 cameras, uses a direct linear transformation method to synthesize the analyzed data into 3D coordinates under the geodetic coordinate system, and uses Butterworth low-pass filtering to filter the analyzed 3D coordinates - time curves, with a truncation frequency of 10 Hz. In addition, the raw images obtained are manually analyzed frame-by-frame and point-by-point by professional In addition, a professional researcher manually analyzes the original images frame by frame and point by point to obtain the 3D coordinates - time curves of the 18 joints of the human body. The multiple correlation coefficients (MCCs) and the mean values of the differences between the automatically analyzed and manually analyzed 3D coordinates - time curves of human joints were calculated. Table 1 shows the correlation coefficients and differences of the 3D coordinates - time curves of human joints.

The correlation coefficients between the automatically resolved and manually resolved 3D coordinates - time curves of human joints are all greater than 0.90, and the correlation coefficients between more than 90% of the automatically resolved curves and the corresponding manually resolved curves are greater than 0.95. The differences between the automatically resolved and manually resolved curves for all joints except for the Z coordinate of the right shoulder, the X coordinate of the right heel, the XY coordinate of the left shoulder, and the Z coordinate of the left ankle, are shown in Table 1, which is the average value of the differences between the automatically resolved and manually resolved curves for all joints. The difference between the resolved curve and the manual resolved curve is less than 0.015m.

Table 1: The correlation coefficient and difference of the node coordinate

Body part	X-coordinate		Y-coordinate		Z-coordinate	
	Coefficient	Difference value(m)	Coefficient	Difference value(m)	Coefficient	Difference value(m)
Right shoulder	0.999	0.0026	0.999	0.0127	0.999	0.0175
Right elbow	0.999	0.0016	0.999	0.0078	0.934	0.0067
Right wrist	0.999	0.0053	0.999	0.0022	0.999	0.0039
Right hand	0.999	0.0072	0.999	0.0016	0.999	0.004
Right hip	0.913	0.0068	0.999	0.0051	0.999	0.0078
Right knee	0.999	0.0029	0.999	0.0048	0.999	0.0108
Right ankle	0.999	0.0061	0.999	0.0193	0.999	0.0011
Right foot heel	0.921	0.0185	0.999	0.0027	0.999	0.004
Right foot tip	0.999	0.0126	0.999	0.0055	0.999	0.0104
Left shoulder	0.999	0.0191	0.999	0.0196	0.999	0.0018
Left elbow	0.999	0.0121	0.999	0.0107	0.999	0.005
Left wrist	0.999	0.0051	0.999	0.0112	0.957	0.0145
Left hand	0.999	0.0122	0.962	0.018	0.999	0.009
Left hip	0.999	0.0104	0.999	0.009	0.999	0.0057
Left knee	0.999	0.0023	0.999	0.0145	0.999	0.0144
Left ankle	0.999	0.0108	0.984	0.0137	0.999	0.0166
Left heel	0.969	0.0115	0.999	0.0077	0.999	0.0017
Left foot tip	0.999	0.0039	0.999	0.0158	0.999	0.0069

IV. F. 3) Motion Capture System Motion Curve Tracking Accuracy

In order to verify the overall effectiveness of the system for traditional wushu movement curve tracking test, the test was carried out in the Mat-lab platform. Tracking the training trajectory of the athlete is the basic function of the movement analysis system, and the results of the traditional martial arts movement curve tracking test are shown in Figure 5.

Analyzing the data in the figure, it can be seen that the application of the system based on this paper, the obtained movement curve is consistent with the actual movement curve. After measurement, the absolute errors of measurement on elbow flexion and extension, shoulder flexion and extension, shoulder internal and external rotation, and shoulder abduction and internal retraction are 0~1.3°, 0~1.7°, 0~2.3°, and 0~2.4°, respectively. This is due to the fact that the Kinect-based system adopts Holt's double-exponential smoothing algorithm to smooth the coordinates of the skeleton joint points, which improves the stability of the skeleton joint point coordinates data which in turn improves the tracking accuracy.

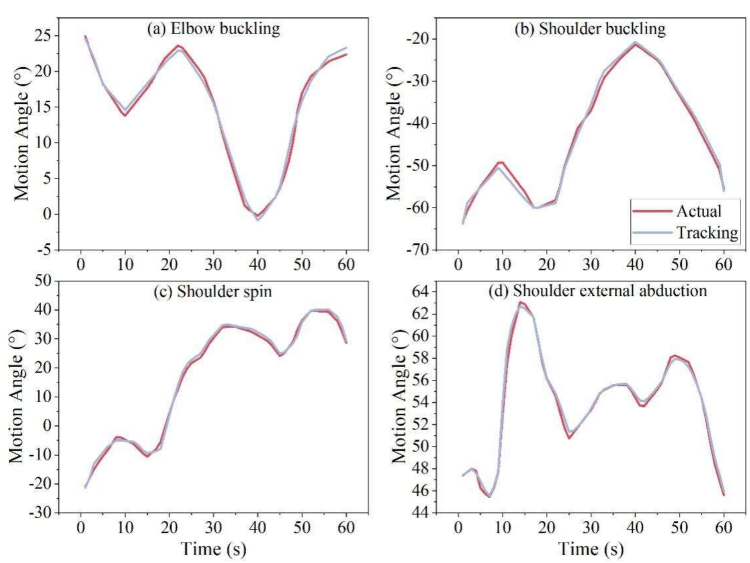

Figure 5: Traditional martial arts action curve tracking test results

IV. F. 4) Evaluation of System Application Effectiveness and Action Scoring Module

Five athletes with the same level of historical physical training were selected to apply the system of this paper for traditional wushu training, and the training level was scored by the system of this paper. At the same time, 3 traditional wushu experts were invited to conduct artificial scoring to verify the validity of the scoring module of the system.

The training videos of the above five athletes were automatically parsed, and Fig. 6 shows the movement characteristics of the five athletes during their transmission wushu training. The results showed that the athletes were able to complete the corresponding movement transitions according to the system prompts during the traditional wushu training, which led to the corresponding changes in the angles of elbow flexion and extension, shoulder flexion and extension, shoulder internal and external rotation, and shoulder abduction and adduction. However, due to the differences between individual athletes, there are slight differences in the angles of movement transitions, but the overall trend is the same.

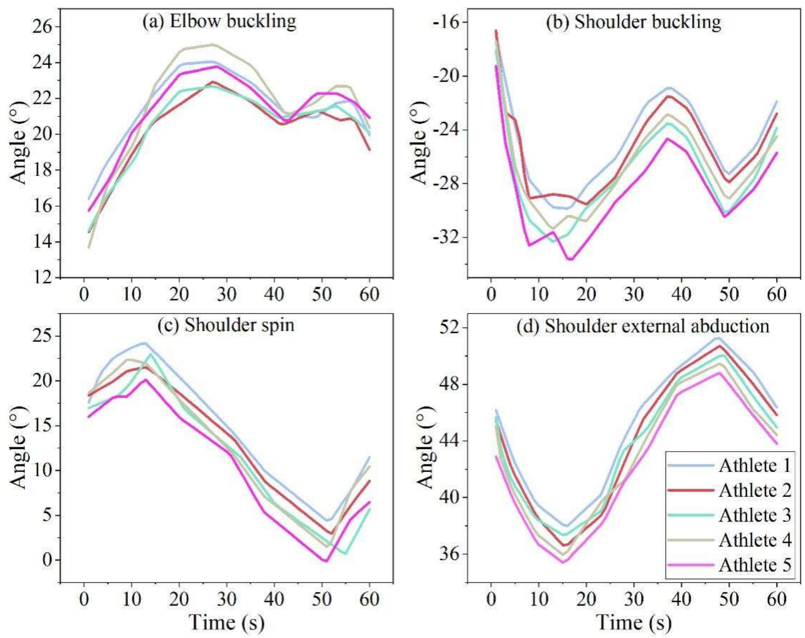

Figure 6: Five athletes' driving characteristics of martial arts

The results of the system and expert scoring in this paper are shown in Table 2. According to the data in the table, it can be seen that five athletes achieved very good results in applying the Kinect-based construction system for traditional martial arts training. The system score is basically consistent with the scores of 3 experts, and the scoring error is between ±0.2 points. It shows that the system scoring module equipped with dynamic time regularization algorithm can objectively analyze the key movements of athletes according to their joint angles, speed and other characteristics during training. The system in this paper can be targeted according to the athlete's skeletal joint coordinates data for relevant training, improve the quality of training and training effect, and verify the effectiveness of the system.

Table 2: The results of this paper and the experts score

Project	Our system	Expert 1	Expert 2	Expert 3
1	8.3	8.4	8.2	8.3
2	9.5	9.5	9.6	9.6
3	8.7	8.7	8.6	8.7
4	8.9	9.1	9.0	8.9
5	9.2	9.2	9.3	9.1

V. Conclusion

The article designs a sports intangible cultural heritage motion capture analysis system based on Kinect motion capture device and Unity3D engine, and applies it to traditional martial arts research. The system is divided into standard database, video teaching, system scoring and 3D reconstruction modules. The movement information collected by Kinect motion capture device will be stored in the standard database, and the angle and speed characteristics of the movement sequence will be extracted by combining with the dynamic time regularization algorithm to realize the scoring of traditional wushu movements. At the same time, the effectiveness of this paper's system in the application of traditional martial arts is verified through simulation experiments.

The system in this paper has good recognition effect on the traditional wushu movements of each joint angle in three-dimensional and two-dimensional levels, in which the average motion range of the spine in three-dimensional and two-dimensional recognition is 137±13° and 138±17°, respectively. The correlation coefficients between the three-dimensional coordinate-time curves of the human joint points and the artificial resolution of more than 90% of the systems are greater than 0.95. In the motion curves obtained by the system in this paper, the absolute errors of measurement on the elbow flexion-extension, shoulder flexion-extension, shoulder internal-external rotation, and shoulder abduction-internal retraction are between 0 and 2.4°. And the system can accurately score the athletes' traditional Wushu movements, which is similar to the results of expert scoring.

References

[1] Xu, Y., Tao, Y., & Smith, B. (2022). China’s emerging legislative and policy framework for safeguarding intangible cultural heritage. International Journal of Cultural Policy, 28(5), 566-580.

[2] Xiao-rong, C. H. E. N. (2022). Sports Intangible Cultural Heritage: Concept, Characteristic, Function, Classification. China Sport Science, 42(4), 14-21.

[3] Kunpeng, L. (2023). Spatial and temporal changes and influence mechanisms of sports intangible cultural heritage in the Yangtze river delta region. J Sociol Ethnol, 5(10), 39-50.

[4] Jike, G. (2023). The connotation and characteristics of Tibetan traditional sports intangible cultural heritage in the context of rural revitalization. Academic Journal of Humanities & Social Sciences, 6(17), 108-112.

[5] Gao, J. (2023). Construction of digital protection path for intangible cultural heritage of ethnic minority traditional sports: take Tibetan traditional sports as an example. J. Sociol. Ethnol, 5(9), 27-32.

[6] Li, K. (2025). Sports Intangible Cultural Heritage Promotes the Sense of Community of the Chinese Nation. Open Journal of Social Sciences, 13(4), 282-295.

[7] Zhang, H. (2023). Research on the History, Dilemma and Optimization Strategies for the Preservation of Intangible Cultural Heritage of Sports in China. Journal of Humanities, Arts and Social Science, 2023, 7 (4).

[8] Liu, L., Zhang, H., & Zhu, L. (2023). Evaluation on the coordination of protection and development of Yunnan ethnic traditional sports intangible heritage projects. Journal of Sociology and Ethnology, 5(7), 139-149.

[9] Liu, X. (2018). International communication of intangible cultural heritage in central plains: a case study of Chinese Wushu. International Journal of Social Sciences and Humanities, 2(3), 196-204.

[10] Chen, H. (2024). Advancements in Human Motion Capture Technology for Sports: A Comprehensive Review. Sensors & Materials, 36.

[11] Rana, M., & Mittal, V. (2020). Wearable sensors for real-time kinematics analysis in sports: A review. IEEE Sensors Journal, 21(2), 1187-1207.

[12] Ouyang, J., Zhang, J., Zhao, Y., & Yang, C. (2024). Multi-Sensor Data Fusion Analysis for Tai Chi Action Recognition. International Journal of Advanced Computer Science & Applications, 15(11).

[13] Host, K., & Ivašić-Kos, M. (2022). An overview of Human Action Recognition in sports based on Computer Vision. Heliyon, 8(6).

[14] Li, Z., Wang, L., & Wu, X. (2025). Artificial intelligence based virtual gaming experience for sports training and simulation of human motion trajectory capture. Entertainment Computing, 52, 100828.

[15] Wenchuan Jia,Hanyang Wang,Qi Chen,Tianxu Bao & Yi Sun. (2025). Analysis of Kinect-Based Human Motion Capture Accuracy Using Skeletal Cosine Similarity Metrics. Sensors,25(4),1047-1047.

[16] Lunlong Zhong,Ping Han,Qingyan Shi,Junhui Li & Wenpan Li. (2025). Development of Virtual Simulation Platform for Automatic Flight Control Based on Unity3D. Curriculum and Teaching Methodology,8(2),

[17] Honglian Li,Suwan Jiang,Mengli Wang & Jiaxiang Lei. (2025). Research on the generation method of outdoor design parameters for summer air-conditioning based on machine learning and dynamic time warping. Building and Environment,278,112968-112968.

[18] Floor van der Does,Wessel van Eeden,Laura F. Bringmann,Femke Lamers,Brenda W. J. H. Penninx,Harriëtte Riese... & Erik Giltay. (2025). Dynamic time warp versus vector autoregression models for network analyses of psychological processes. Scientific Reports,15(1),11720-11720.