

Research on the Positioning Technology of Sports 3D Teaching Action

Based on Machine Vision

LIU-Hao

Zhongnan University of Economics and Law, Wuhan 430064, China

liuhao@zuel.edu.cn

Hari Mohan Pandey*

Department of Computer Science, Edge Hill University, UK.

Pandeyh@edgehill.ac.uk

*Corresponding author: Pandeyh@edgehill.ac.uk

Abstract: This paper presents a method of action location in three-dimensional motion teaching. The machine vision technology is used to solve the problems of low positioning accuracy and long positioning time in the traditional motion three-dimensional teaching method. The work of this method is as follows: (a) using machine vision method to determine the world coordinate system of the image; (b) using MRF algorithm to extract the features of 3D teaching action image; (c) determining the spatial correlation of 3D teaching action data. In the three-dimensional teaching action image, the smooth filtering technology is used to suppress and eliminate the noise. Then the convolution neural network (CNN) is used to reconstruct the three-dimensional teaching action image. The entropy of three-dimensional teaching behavior of physical education is determined by CNN. Through a large number of computer simulations, the effectiveness of the proposed system is confirmed. The results show that the system achieves 95% accuracy when the positioning time is 1.9s.

Keywords: Machine vision; 3D teaching action; Positioning technology; Smoothing filtering technology.

1. Introduction

With the development of video technology and biometric recognition technology, human posture recognition technology has been widely used in traffic, sports, medical

and other fields. Among them, the application of human motion recognition is more extensive. In recent years, with the rapid development of biometric recognition technology, the recognition and location of human motion has become a hot research field, and it is a technology with potential development space. In the field of sports, the orientation of three-dimensional teaching movement can help athletes and coaches improve their training level. Coaches give effective guidance according to the performance of athletes in sports. Since the end of the 20th century, the distance knowledge and skills teaching based on the network support environment has become a reality, and various network teaching courses and related research have also arisen. The network can break through the limitation of time and space, provide personalized teaching and diversified interactive mode, and make up for the defects of traditional teaching activities. This paper analyzes and summarizes the sports skills from the theory of sports skill teaching. Therefore, people have carried on a lot of research on the orientation of the three-dimensional teaching behavior of physical education and achieved some achievements [1-3].

In reference [4], a method of action location for CNNs based on global spatiotemporal features is proposed. This method is designed to solve the problem of poor local feature extraction of human motion. This method first constructs a dual channel recognition model of global time domain and spatial features. Through the design of spatial feature channel, multi frame fusion method is introduced to fuse the motion action. Then the global time domain channel is used to analyze the motion action, and the fused data is effectively trained to realize the deficiency of motion action. This method can quickly complete the prepositioning of action recognition through the set recognition model, but the scope of this method is narrow, and the effect is good only in a certain range, which has some limitations. In reference [5], a human motion localization method based on multi-scale trajectory of optical flow key points is proposed. Aiming at the problem of complex feature extraction of smooth light in traditional recognition methods, this method is designed. In this method, the scale trajectory of key points in optical flow region is combined with human motion trajectory, and the constraint conditions of human motion recognition are designed to

determine the local characteristics of the action, and then the obtained action related data is input into the designed classifier to complete the motion recognition. This method can effectively improve the positioning accuracy of the action, but there are many interference factors in the operation process of this method, which takes a long time to de-noising, and the positioning time is long, so it still needs further improvement. In reference [6], an action feature tracking algorithm based on Markov random field is proposed. The algorithm extracts the feature sub blocks of various parts of the human body in motion, selects the tracking sub block with higher significance as the basis of action location, obtains the spatial constraint relationship between different sub blocks through Markov random field, and realizes the action feature location based on the constraint relationship. The algorithm can accurately divide different feature sub blocks of action, but the resolution of the action image obtained by the algorithm is low, which leads to low positioning accuracy and has some defects.

Based on the aforementioned problems, this paper proposes a novel motion location technology by using machine vision to determine the world coordinate system of the image, using MRF algorithm to both extract the characteristics of three-dimensional teaching action image and determine the spatial correlation of three-dimensional teaching action data. With the help of smooth filtering technology, the noise in the three-dimensional teaching action image is effectively suppressed and eliminated. Therefore, the three-dimensional teaching action image is reconstructed. The entropy value of the three-dimensional image is determined by using convolution neural network. The visualization of teaching behavior is realized, and the three-dimensional positioning of physical education teaching behavior is realized.

The key contributions of this paper are presented as follows:

- Firstly, the world coordinate system of the three-dimensional teaching action image is determined by machine vision, which is used to locate the motion. In this paper, the three-dimensional sports teaching action image is first obtained through the camera, and then located the action by using the artificial intelligence algorithm.
- Secondly, MRF algorithm is used to extract the features of 3D teaching action image of physical education, and to determine the spatial correlation of 3D teaching

action data of physical education. With the help of smooth filtering technology, the noise in 3D teaching action image of physical education is effectively suppressed and eliminated, and the preprocessing of 3D teaching action image of physical education is completed.

- Thirdly, a convolution neural network is used to determine the entropy value of sports three-dimensional teaching action image, and to realize the positioning of sports three-dimensional teaching action.
- Finally, computer simulations have been performed to evaluate the performance of the proposed system.

The rest of the paper is organized as follows: **Section 2** presented the acquisition and preprocessing of 3-D teaching action image of the physical education; **Section 3** discusses the physical education using 3-D teaching action positioning; Experimental analysis is described in **Section 4**; and finally, conclusion has been drawn in **Section 5**.

2. Acquisition and Preprocessing for 3-D Teaching Action Image of Physical Education

2.1 Acquisition method

In order to realize the three-dimensional teaching action positioning of physical education, this paper uses machine vision to locate it. Machine vision is a key technology in industrial detection, target recognition, positioning, tracking and other fields through robots [7]. This method has the advantages of non-contact, high precision, flexibility and reliability, and good environmental adaptability. Machine vision obtains the scene image through the camera, and then preprocesses the image to determine the location information of the research object, which can effectively solve the problem of human detection and positioning. In this paper, we first obtain the three-dimensional sports teaching action image through the camera, and then use the artificial intelligence algorithm to locate its action [8-9].

In this paper, firstly, through the camera calibration, determine its coordinate system, and then obtain the sports three-dimensional teaching action image. It mainly includes three-dimensional teaching action image coordinate system, camera

coordinate system and world coordinate system. The coordinate system of sports 3D teaching action image is represented by pixel unit, and the coordinate origin is located in the lower left corner of the whole action image.

Set (u, v) to represent the coordinates of a certain point of the action image, and both represent the unit pixel of the action image. Set the origin of the coordinate system as (u_0, v_0) pixel and set the coordinates of a point representing the action image using equation (1).

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{k} & 0 & u_0 \\ 0 & \frac{1}{l} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (1)$$

Among them, the pixel size of the action image is $k \times l$.

Through the transformation of its coordinates by equation (1) and it is obtained by equation (2).

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix} f & -f \cot \theta & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & \sin \theta & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (2)$$

In the equation (2), f represents the focal length of the camera, θ represents the skewness of the camera coordinate system. On this basis, the relationship between the action image and the camera coordinate system is determined using equation (3).

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix} \frac{f}{k} & \frac{-f \cot \theta}{k} & u_0 & 0 \\ 0 & \frac{f}{l \sin \theta} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \frac{1}{Z_c} M \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (3)$$

Where, M represents the intra-camera parameter matrix. $(k, l, u_0, v_0, f, \theta)$ represents the camera internal parameters.

Through the world coordinate system, the acquisition of three-dimensional teaching action image of physical education is completed. The world coordinate system

is a three-dimensional coordinate system with any point in space as the origin. The relationship between the camera coordinate system and the world coordinate system is represented through equation (4).

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} R_{3 \times 3} & t_{3 \times 1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \quad (4)$$

Where, $R_{3 \times 3}$ represents the rotation matrix, $t_{3 \times 1}$ represents a translation vector, (X_C, Y_C, Z_C) represents the camera coordinates, (X_W, Y_W, Z_W) represents the world coordinate system.

During the three-dimensional space, the two-dimensional rotation of the x, y, z coordinate axis obtained above is rotated, the coordinate angle of the three-dimensional teaching action image of physical education is determined, and the image acquisition is completed using equation (5).

$$\begin{bmatrix} x^1 \\ y^1 \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (5)$$

2.2 Image Feature Extraction

In order to realize the 3D teaching action positioning of physical education, it is necessary to obtain the 3D teaching action data [10] of physical education in the coordinate system of determining the action image above. According to the three-dimensional teaching action image of physical education, the characteristic data of three-dimensional teaching action image of physical education are obtained.

Physical education 3D teaching action image imaging is more complex because physical education 3D teaching action characteristics are not typical features of target images. Therefore, it is necessary to extract the features of the three-dimensional teaching action image of physical education. This paper uses the MRF algorithm [11] extract the features of the three-dimensional teaching action image of physical education and determine the spatial correlation of the three-dimensional teaching action data of physical education.

Suppose $A = \{a_1, a_2, a_3 \dots a_i\}$ represents the set of three-dimensional physical education action data, the threshold of any value is $\Delta(1, 2, \infty)$. The selected value is the random value in the three-dimensional teaching action data of physical education, which is regarded as the joint value of the three-dimensional teaching action data of physical education. At this time, the probability of the joint value in the MRF algorithm is G_a as represented by equation (6).

$$G_a = G\left(\frac{a_i}{a_n}\right) / \sum_{a=1}^i \left(\frac{a_i}{a_n}\right) \quad (6)$$

In the formula, $\frac{a_i}{a_n}$ represents the conditional probability of random values in the three-dimensional physical education action data.

According to the joint value of physical education 3D teaching action random field obtained above, the image threshold of physical education 3D teaching action is determined. Suppose the three-dimensional teaching action image of physical education contains three parts: target, shadow and background [12], set the threshold of these three Z , divide them into two pixels $p(e, r)$ as shown by equation (7).

$$p(e, r) = \begin{cases} 1, & g(e, r) \\ 0, & g(e, r) \end{cases} \quad (7)$$

Where, 1 represents the characteristic target value of 3D teaching action image of physical education and 0 represents the background value of 3D teaching action of physical education.

After obtaining the threshold of 3D teaching action image of physical education, the feature extraction of 3D teaching action image of physical education is completed according to the characteristic target value and the background value as shown in equation (8).

$$p^k = \frac{\sum_{p(e,r) \leq t_k} p(e, r)}{M_k} \quad (8)$$

Where, M_k represents the characteristic target value and background value of 3D teaching action image of physical education.

After the above analysis, we get the three-dimensional teaching action image of physical education. This paper simulates it to better show the target segmentation of three-dimensional action in physical education teaching, as shown in Figure 1,



Fig.1 The image feature acquisition of 3D teaching action in physical education.

Based on the course goal and the depth of the sample video image preprocessing content, a general teaching requirement is put forward by using the directional gradient histogram feature to reflect the concrete behavior. Teaching objectives also refer to the direction of histogram features and expected results. From the point of view of the success of image feature extraction, the teaching methods and strategies of output feature selection and the arrangement of teaching process play an important guiding role. On this basis, the extraction process of sports 3D teaching action image is set, as shown in Figure 2.

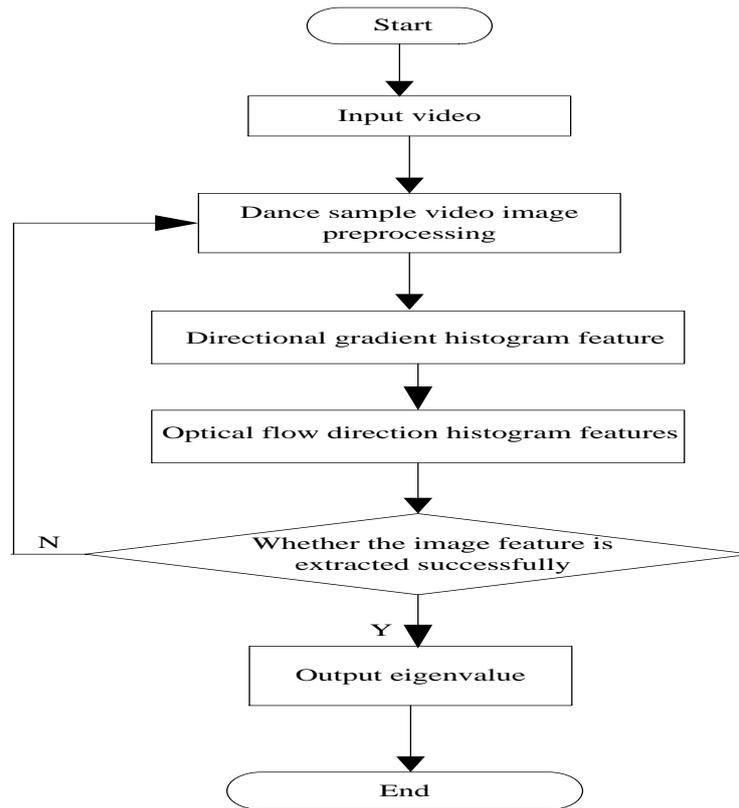


Fig. 2 Feature extraction process of 3D teaching action image of Physical Education

2.3 Pre-processing of Image Feature Data

After obtaining the above physical education 3D teaching action image data because the original image obtained does not meet the standard stipulated in the image processing technology in terms of format, pixel, brightness, clarity, etc., it will reduce the line shape feature [13] of the sports 3D teaching action image without preprocessing. Therefore, it is very important to preprocess the feature of 3D teaching action image. In the 3D teaching action image of physical education, the pre-processing of 3D teaching action image of physical education is completed by gray scaling the image lines and de-noising the image lines.

In the preprocessing of physical education 3D teaching action image, although the obtained 3D teaching action image will contain more abundant edge information than the gray image, there is still no mature algorithm so far. It can be directly based on color image for image edge line feature preprocessing, and the processing is difficult [14]. Therefore, it is necessary to convert the color image lines into gray image lines in the 3D teaching action images of physical education. In the gray processing of image lines,

the three-component pixel value in the image lines is extracted as the gray value of three gray images by component method as presented by equation (9).

$$\begin{cases} \text{Gray1}(i, j) = R(i, j) \\ \text{Gray2}(i, j) = G(i, j) \\ \text{Gray3}(i, j) = B(i, j) \end{cases} \quad (9)$$

Where, *Gray1*, *Gray2* and *Gray3* represent the different gray values of the three-point line of the three-dimensional teaching action image of physical education, respectively. *R*, *G*, *B* corresponding to the three primary colors of the image lines, *R*(*i, j*), *G*(*i, j*), *B*(*i, j*) is the pixel value of the image line at the point (*i, j*) .

On the basis of this, the maximum value of-component pixel value is regarded as the gray value as shown in equation (10).

$$\text{Gray}(i, j) = \max[R(i, j), G(i, j), B(i, j)] \quad (10)$$

On this basis, the three-dimensional physical education teaching action image features are preprocessed. In this paper, the smoothing filtering technique is used to effectively suppress and eliminate the noise in the 3D teaching action image of physical education. A pixel point is randomly selected from the 3D teaching action image of physical education as the center point, and then a square field is determined around this center point, all the pixel values in the square field are counted and sorted. Select intermediate value instead of randomly selected center point pixel value. Square domain is called filter window [15], there are 3×3 filter window and 5 filter windows. Using the filter window to move up and down on the 3D teaching action image of physical education, the intermediate value obtained is replaced by the gray value of the original image, and the 3D teaching action image of physical education is filtered and de-noised as shown in equation (11).

$$G(i, j) = \text{median}\{h(k, l) | f(i, j) \in Nf(i, j)\} \quad (11)$$

Where, *G*(*i, j*) is the gray value of any pixel in the 3D teaching action image of physical education after median filtering; *h*(*k, l*) is the gray value of any pixel in the

initial 3D teaching action image; $Nf(i, j)$ represents the first order solid of the $f(i, j)$.

On this basis, it completes the preprocessing of sports 3D teaching action image data. This process is very important. It is a key step in the follow-up research and design. In this extraction, it also needs multi feature fusion of sports 3D teaching action image data. Therefore, this paper makes an in-depth study on it.

After extracting the color features, geometric features, and corner features from the 3D teaching action image of physical education, we need to integrate the features before tracking them, and then track the features using equation (12).

$$U(i, j) = \alpha p(x) \sum_{i=1}^n (pv) \quad (12)$$

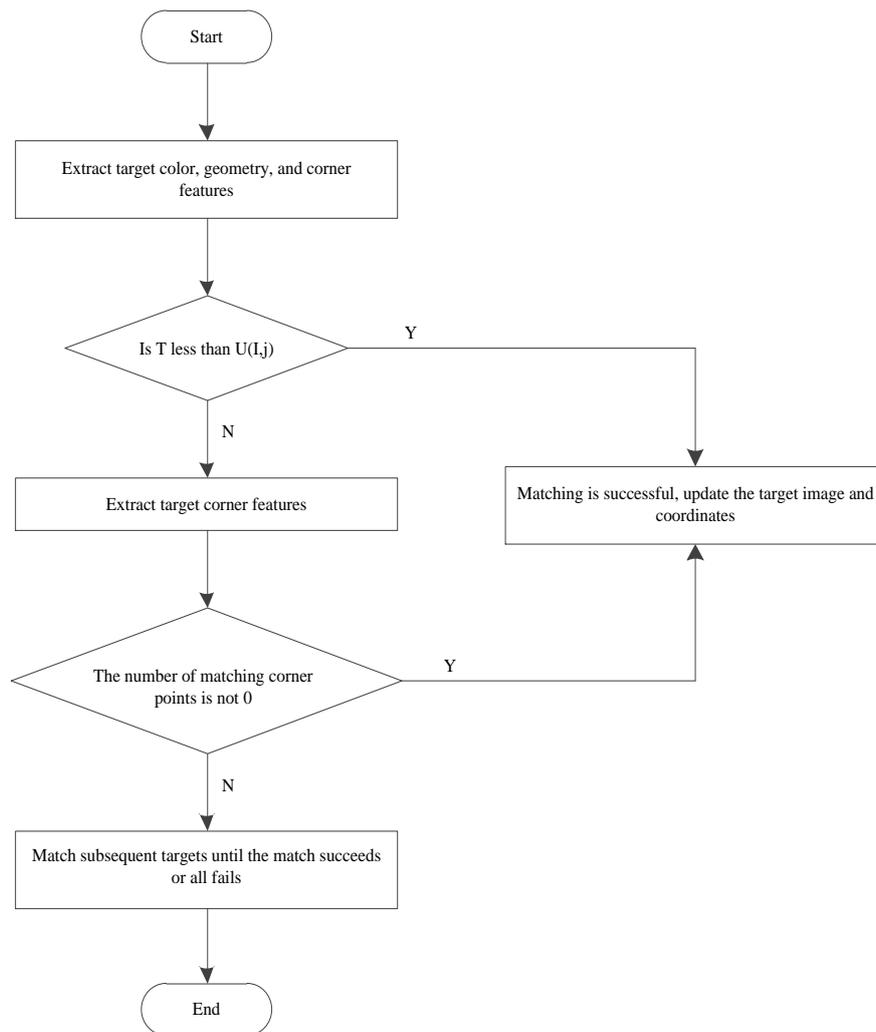


Fig. 3. The process of feature fusion and matching of sports 3D action.

The process of feature fusion and matching of sports 3D action is shown in Figure 3.

3. Physical Education Using 3-D Teaching Action Positioning

On the basis of the three-dimensional teaching action of physical education, the three-dimensional teaching action of physical education is positioned. This paper uses the convolution neural network in the intelligent neural network to locate the three-dimensional teaching action of physical education.

Deep convolution neural network is an artificial neural network connected by a fixed neuron, which has a certain hierarchical [16]. The structure of deep convolution neural network is receptive field, channel, filter bank, feature map, activation function, convolution layer and pool layer. The convolution layer of deep convolution neural network is the range of local connection, that is, the receptive field of convolution layer. The principle is shown in Figure 4:

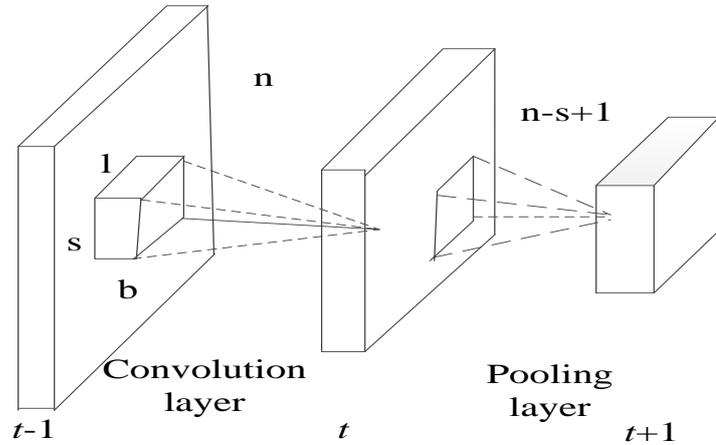


Fig. 4. A schematic diagram of convolution and pooling of deep convolution neural networks.

In figure 4, the receptive field size of the convolution layer of the deep convolution neural network is $\xi \times l$, the corresponding shared weights are convolution kernels, also known as filters. The movement of different convolutional kernels on spatial dimensions will generate multiple feature maps. Figure 4 generates feature maps of different sizes after convolution operation, the number of which is the number of channels, which is determined by the number of convolution kernels in the convolution layer. Thus, after passing through a convolution layer, k feature map of the T layer has the following elements at the location (i, j) of the space as presented by equation (13).

$$x_{i,j,k}^t = \varphi\left(\sum_{c=1}^b \sum_{m=1}^{\xi} \sum_{n=1}^l x_{i+m,j+n,c}^{(t-1)} \times \varpi_{m,n,c,k}^{(t)} + y_k^{(t)}\right) \quad (13)$$

In equation (13), $\varphi(\bullet)$ represents the activation function of deep convolutional neural networks, ξ, l represent the size of the receptive field space of the deep convolution neural network, b represents the number of input layer channels, y is the bias term of the deep convolution neural network, ϖ represents the size of the weights. For deep convolution neural network, there are no learning parameters in the pool layer. In essence, the convolution layer and the pool layer need linear filtering and multi-feature transformation [17]. Assuming that the $v^{(1)} = a$ is the input layer of the deep convolution neural network, then the v^n represents the response of layer N as presented in equation (14).

$$\begin{aligned} e^{(n+1)} &= Q^{(n)}v^{(n)} + \tau^{(n)} \\ v^{(n+1)} &= \varphi^{(t+1)}e^{(n+1)} \end{aligned} \quad (14)$$

Where, e , v and τ vectors representing response and bias, Q matrices representing weights.

Firstly, every candidate frame of 3D teaching action image of physical education is divided into two parts in the 3D teaching action orientation of physical education $M \times N$, then pool each region is pooled and the candidate box of different size in the 3D teaching action image of physical education is transformed into the same size data input to the next layer of convolution neural network [18-120], Suppose the nodes of the input layer are x_i , node of the output layer i is y_i to get the cost function as shown in equation (15).

$$\frac{\partial L}{\partial x_i} = \begin{cases} 0 & \delta(i, j) = F \\ \frac{\partial L}{\partial x_i} & \delta(i, j) = T \end{cases} \quad (15)$$

Where, $\delta(i, j)$ indicates whether the decision function is the maximum output

value of the candidate box at the i node. Cost for the x_i gradient and the sum of the correlation gradient of the latter layer.

It is assumed that the original gray scale range of the three-dimensional teaching action image of physical education is $[0, L-1]$. The maximum value obtained is the threshold of the 3D teaching action image of physical education, and the location of the action point is completed as given in equation (16) to (18).

$$P_1(x) = \sum_{i=0}^{n-1} -(\mathcal{A}[i]/\sigma_0(x) \times \log(\mathcal{A}[i]/\sigma_0(x))) \quad (16)$$

$$P_2(x) = \sum_{i=1}^{L-1} -(\mathcal{A}[i]/\sigma_1(x) \times \log(\mathcal{A}[i]/\sigma_1(x))) \quad (17)$$

$$P = \text{Max}(P_1(x) + P_2(x)) \quad (18)$$

Where, σ_0 shows the proportion of the number of foreground pixels to the whole picture, σ_1 indicates the proportion of the number of background pixels to the whole picture, $\mathcal{A}[i]$ represents the probability of i gray values, $P_1(x)$ representative foreground entropy, $P_2(x)$ represents the background entropy value.

4. Experimental analysis

To verify the effectiveness of this method, the experiment uses the method of comparing this paper, the method of literature [5] and the method of literature [6], set up a certain experimental environment, and the experimental index to verify the effectiveness of this method.

4.1 Experimental environment

The experiment is carried out on the MATLAB 7.2 platform. The experimental operating system is a Windows XP system with a running memory of 8 GB, CPU and 36. GHz, in order to ensure the accuracy of experimental data analysis, the SPSS 13.0 software is used to analyze the experimental data. Through the image acquisition of basketball players in a sports school, three methods are used to locate them.

4.2 Experimental parameters

In order to ensure the reliability of the experiment, the relevant parameters used in the experiment are designed, as shown in Table 1.

Table 1. Data sheets for experiments

Project	Value
Image Size of Basketball Training/dpi	246*246
Number of noise points per unit	1000
Area of noise area/cm ²	2
Number of positioning/times	100
Number of iterations/times	200
Sampling interval/ s	0.2
Number of actual operations/times	100

4.3 Result Analysis

In order to verify the advantages of this method, the experiment compares the method of this paper, the method of literature [5] and the method of literature [6] and takes the accuracy and time-consuming of the sample image action positioning as the experimental indicators to complete the experimental analysis.

4.3.1 Accuracy analysis

To verify the effectiveness of this method, the experimental analysis of this method, literature [5] method and literature [6] method in the sample image location accuracy, the experiment carried out a total of 100 iterations, the value is more objective. The results are shown in Figure 5.

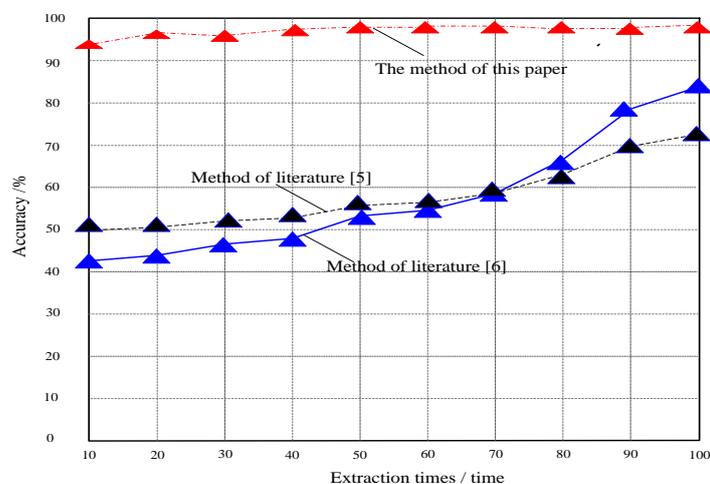


Fig .5 Accuracy Analysis of Three-Dimensional Teaching Action Positioning of Physical Education by

Different Methods

From the test results, with the change of test times, there is a certain gap in the positioning accuracy of the sample action image between the method in this paper, the method in literature [5] and the method in literature [6]. When the number of tests is 50, the location accuracy of reference [5] method and reference [6] method is about 55% and 52% respectively, and the location accuracy of this method is about 95%; when the number of tests is 100, the extraction accuracy of reference [5] method and reference [6] method is about 71% and 83% respectively, and the location accuracy of this method is about 97%. Through the comparison of the test results, the positioning accuracy of this method is higher. This is because this method obtains the 3D teaching action image of physical education through machine vision, determines the world coordinate system of the image; uses MRF algorithm to extract the features of the 3D teaching action image of physical education, determines the spatial correlation of the 3D teaching action data of physical education, and effectively suppresses and eliminates the noise in the 3D teaching action image of physical education with the help of smooth filtering technology, so as to complete the 3D teaching of physical education. The preprocessing of 3D teaching action image improves the positioning accuracy of this method.

4.3.2 Time-consuming analysis

In order to further verify the scientific effectiveness of the proposed method, the experiment analyzes the time-consuming of the three methods to locate the sample teaching actions

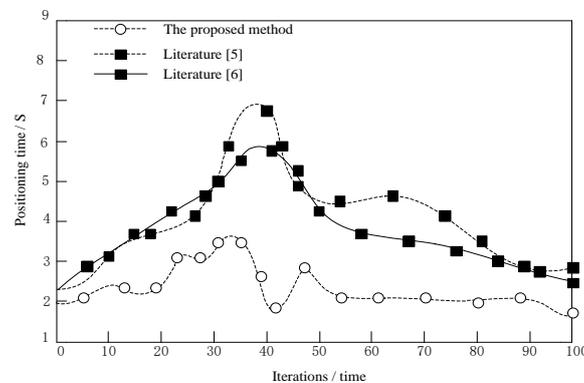
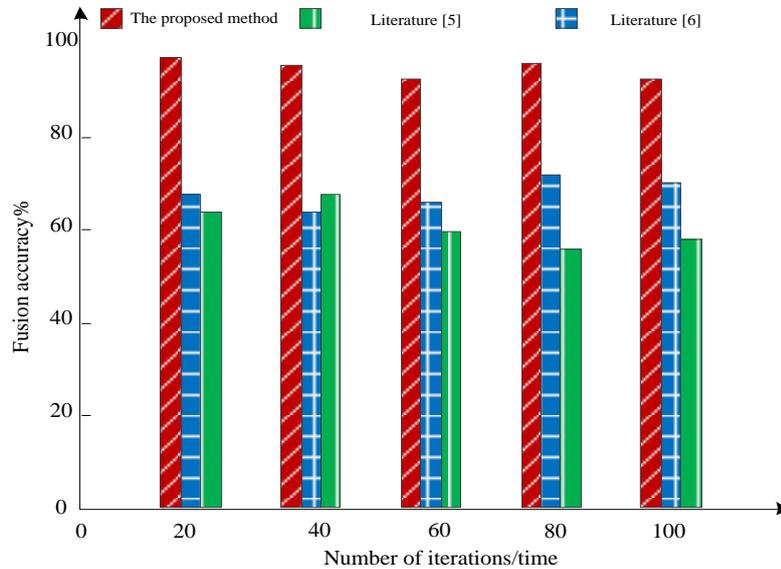


Fig. 6. Time-consuming comparison of three-dimensional physical education movement localization in different methods.

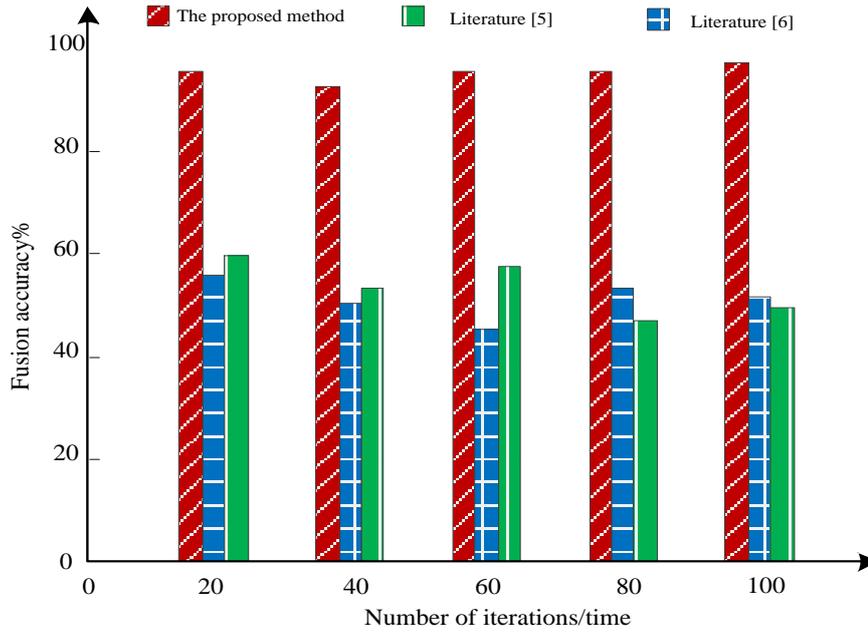
By analyzing the data in Figure 6, we can see that in the same experimental environment, the positioning time of the three methods is different. Among them, the time-consuming of this method is always lower than that of the other two traditional methods, and the minimum is about 1.9 s, while the time-consuming of the other two methods is higher than that of this method. This is because this method uses convolution neural network to input and output the processed three-dimensional motion image and achieves fast positioning.

4.3.3 Accuracy analysis

After the above two indexes are determined, the accuracy of the method in literature [5] and literature [6] in the fusion of three-dimensional sports action features is verified in the experiment. The resolution of the image is $720 * 400$ and $384 * 288$ respectively, and the experimental results are obtained by using three methods for 100 iterations as shown in Figure 7:



(a) Fusion accuracy of different methods at $720 * 400$ resolution.



(b) Fusion accuracy of different methods at 384 * 288 resolution

Fig. 7. Comparison of feature fusion accuracy of 3D sports action image.

By analyzing Fig. 7 (a), it can be seen that when the resolution of 3D sports action image is 720 * 400, the fusion accuracy of the method in this paper, the method in reference [5] and the method in reference [6] changes with the number of iterations. When the number of iterations is 20, the fusion accuracy of our method is 95%, that of reference [5] is 62%, and that of reference [6] is 64%; when the number of iterations is 40, the fusion accuracy of our method is 94%, that of reference [5] is 66%, and that of reference [6] is 64%; when the number of iterations is 80, the fusion accuracy of our method is 94% The fusion accuracy of this method is 97%, that of reference [5] is about 70%, and that of reference [6] is about 75%, When the resolution of three-dimensional sports action image is 384 * 288, the fusion accuracy of this method, literature [5] method and literature [6] method changes with the number of iterations. When the number of iterations is 20, the fusion accuracy of our method is 96%, that of reference [5] is 47%, and that of reference [6] is 59%; when the number of iterations is 40, the fusion accuracy of our method is 93%, that of reference [5] is 52%, and that of reference [6] is 45%; when the number of iterations is 80, the fusion accuracy of our method is 93% The

fusion accuracy of this method is 95%, that of reference [5] is about 62%, and that of reference [6] is about 54%. Compared with the comparison results of two images, it is not difficult to find that the fusion accuracy of this method is always higher than that of reference [5] and reference [6], which verifies the effectiveness of this method.

5. Conclusion

Aiming at the problems of low positioning accuracy and long positioning time in traditional 3D teaching of sports, a 3D teaching motion positioning method based on machine vision was proposed. The world coordinate system of the image is determined by machine vision. MRF algorithm is used to extract the characteristics of 3D teaching action images and determine the spatial correlation of 3D teaching action data. By means of smoothing filter technology, the noise in 3D teaching action image is effectively suppressed and eliminated, and on this basis, the reconstruction of 3D teaching action image is completed. The convolutional neural network is used to determine the entropy value of 3D motion teaching action image and realize the positioning of 3D motion teaching action. Compared with traditional methods, the proposed method has the following advantages :(a) the highest accuracy of the proposed method is about 95%, and the positioning accuracy is high; (b) The proposed method takes about 1.9s, and the time is relatively short. In the future, further research on PE 3D teaching is needed so as to improve it further.

Conflicts of interest/Competing interests: Authors declare no conflict of interest.

References

1. Y. Lin, H. Zhao, X. Ma, Y. Tu and M. Wang, "Adversarial Attacks in Modulation Recognition with Convolutional Neural Networks," in IEEE Transactions on Reliability, doi: 10.1109/TR.2020.3032744.
2. Qin Y, Na Q, Liu F. Strain gauges position based on machine vision positioning[J]. Integrated Ferroelectrics, 2019, 20(1):191-198.

3. Wang Q, Cheng M, Noureldin A. Research on the improved method for dual foot-mounted Inertial/Magnetometer pedestrian positioning based on adaptive inequality constraints Kalman Filter algorithm[J]. *Measurement*, 2019, 15(12):189-198.
4. Wang Ke , Wu JunZhou , Tianxiang , Li Ruifeng . An action recognition method based on global spatial-temporal feature convolutional neural networks[J]. *Journal of Huazhong University of Science and Technology (Nature Science Edition)*, 2018, 46(12): 36-41.
5. Gadekallu TR, et al. Deep neural networks to predict diabetic retinopathy. *Journal of Ambient Intelligence and Humanized Computing*. 2020, 24:1-4.
6. Belhadi A, et al. Deep learning for pedestrian collective behavior analysis in smart cities: A model of group trajectory outlier detection. *Information Fusion*. 2021, 65:13-20.
7. Bai L, Yang X, Gao H. Corner Point-Based Coarse–Fine Method for Surface-Mount Component Positioning[J]. *IEEE Transactions on Industrial Informatics*, 2018, 14(3):877-886.
8. Tu Y, Lin Y, Wang J, et al. Semi-Supervised Learning with Generative Adversarial Networks on Digital Signal Modulation Classification[J]. *CMC-Computers Materials & Continua*, 2018, 55(2): 243-254.
9. Wang Q, Cheng M , Noureldin A , et al. Research on the improved method for dual foot-mounted Inertial/Magnetometer pedestrian positioning based on adaptive inequality constraints Kalman Filter algorithm[J]. *Measurement*, 2019, 135(20):189-198.
10. Yun Lin, Ya Tu, Zheng Dou, et al, "Contour stella image and deep learning for signal recognition in the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, DOI: 10.1109/TCCN.2020.3024610.
11. Shuai Liu, Shuai Wang, Xinyu Liu, et al, Human Memory Update Strategy: A Multi-Layer Template Update Mechanism for Remote Visual Monitoring, *IEEE Transactions on Multimedia*, 2021, online first, 10.1109/TMM.2021.3065580

12. Jh C, Jj Z, Rj G. Research on Modified Algorithms of Cylindrical External Thread Profile Based on Machine Vision[J]. Measurement Science Review, 2020, 19(15):11-21.
13. Huang X, Wang F, Zhang J, et al. A Posture Recognition Method Based on Indoor Positioning Technology[J]. Sensors, 2019, 19(6): 259-264.
14. Liu Shuai, Wang Shuai, Liu Xinyu, et al. Fuzzy Detection aided Real-time and Robust Visual Tracking under Complex Environments. IEEE Transactions on Fuzzy Systems, 2021, 29(1), 90-102
15. Wu X, Cao Q, Li Y. A research on wireless sensor networks' node positioning mechanism based on Narrowband Internet of Things data linking[J]. International Journal of Distributed Sensor Networks, 2018, 14(12):36-42.
16. Ortiz S D F, Amestoy M E, Carrero-Blanco J. Positioning measurement using a new artificial vision algorithm in LabVIEW based on the analysis of images on an LCD screen[J]. The International Journal of Advanced Manufacturing Technology, 2020, 10(5):1-16.
17. Shuai Liu, Chunli Guo, Fadi Al-Turjman, et al. Reliability of response region: A novel mechanism in visual tracking by edge computing for IIoT environments, Mechanical Systems and Signal Processing, 2020, 138, 106537
18. Ge L, Dan D, Li H. An accurate and robust monitoring method of full-bridge traffic load distribution based on YOLO-v3 machine vision[J]. Structural Control and Health Monitoring, 2020, 27(01):17-25.
19. Shi C, Qian J, Han S. Developing a machine vision system for simultaneous prediction of freshness indicators based on tilapia (*Oreochromis niloticus*) pupil and gill color during storage at 4°C.[J]. Food Chemistry, 2018, 243(15):134-140.
20. Liu S, Dongye Liu, Gautam Srivastava, et al. Overview and methods of correlation filter algorithms in object tracking. Complex & Intelligent Systems, 2020, online first, 10.1007/s40747-020-00161-4