Research Article

Intelligent Recognition Method of Athlete Wrong Movement Based on Image Vision

Wang Lu[1] and JiangYuan Hou [ID][2]

[1]*The School of Physical Education of Yantai University, Yantai, Shandong 264005, China*
[2]*Sports Training Department, Hebei Sport University, Shijiazhuang 050000, Hebei, China*

Correspondence should be addressed to JiangYuan Hou; 1996004@hepec.edu.cn

Current methods of human body movement recognition neglect the depth denoising and edge restoration of movement image, which leads to great error in athletes' wrong movement recognition and poor application intelligence. Therefore, an intelligent recognition method based on image vision for sports athletes' wrong actions is proposed. The basic principle, structure, and 3D application of computer image vision technology are defined. Capturing the human body image and point cloud data, the three-dimensional dynamic model of sports athletes action is constructed. The color camera including CCD sensor and CMOS sensor is selected to collect the wrong movement image of athlete and provide image data for the recognition of wrong movement. Wavelet transform coefficient and quantization matrix threshold are introduced to denoise the wrong motion images of athletes. Based on this, the feature of sports athlete's motion contour image is extracted in spatial frequency domain, and the edge of the image is further recovered by Canny operator. Experimental results show that the proposed method can accurately identify the wrong movements of athletes, and there is no redundancy in the recognition results. Image denoising effect is good and less time-consuming and can provide a reliable basis for related fields.

1. Introduction

With the increasing importance attached to sports events, athletes need to train according to various standard movements in the training process. Referees cannot recognize the wrong movements because of the fast speed or the large number of people. With the development of computer vision technology, it is widely used in the analysis of human body structure. In order to realize the intelligent recognition of athlete's wrong action, computer vision technology is applied. It can not only improve the athlete's movement level but also quickly and accurately judge whether the athlete has the wrong movement, thus enhancing the competition fairness. Computer vision technology is one of the key researches in the field of graphics and computer vision because there are many sports items and there are some differences in recognizing athletes' wrong actions. There are a variety of organs and tissues in the human body. Different combinations of these tissues can make the human body complete

specific behavior, integrate the information of human motion, and provide a strong basis for the analysis of human behavior. Therefore, the action recognition based on image human body has very critical research value [1, 2]. At present, the commonly used methods of action recognition are usually disturbed by illumination, occlusion, or shaking, which leads to the difficulty of action feature extraction and the inability to get useful features accurately, so that the subsequent segmentation and recognition effect is not ideal.

Liu [3] proposed a method of human posture recognition based on multifeature fusion. Aiming at the problem that the existing posture recognition algorithms cannot reflect the dynamic characteristics of athletes' posture, this paper proposes a posture recognition algorithm based on multifeature fusion. Firstly, the image is captured by optical image collector, and then the image is transformed to gray level to improve the image quality. Then, the body contour and the motion region are obtained based on shadow elimination and interframe difference. Finally, posture

region and body contour are extracted based on Radon transform and discrete wavelet transform. Chen [4] proposed a moving image contour feature extraction method based on multithreshold optimization. Aiming at the problems of long extraction time and low extraction accuracy in traditional moving image contour feature extraction methods, a moving image contour feature extraction method based on multi-threshold optimization is proposed. Through moving image contour feature analysis, the membership function is obtained by using the maximum interclass variance fuzzy constraint method, and multiple thresholds of target contour in moving image are calculated by using fuzzy membership. The geometric center values of the two contour points adjacent to the center point in the image contour range are calculated by using multiple constrained thresholds. The curvature symbol is obtained by calculating the curvature angle, and the contour features of the moving image are extracted according to the curvature symbol. Shen et al. [5] proposed an action recognition method of sports athletes based on deep learning. This method processes the image data of sports athletes through dense optical flow method and extracts the characteristics of wrong actions of sports athletes by combining short-term memory neural network and convolution neural network. This method cannot obtain key frames in the process of data processing, resulting in low feature extraction rate of the method.

Aiming at the above problems, this paper proposes a new intelligent recognition method of sports athletes' wrong actions based on image vision.

2. Intelligent Recognition Method of Sports Athletes' Wrong Actions Based on Image Vision

2.1. Computer Image Vision Technology. With the development of computer technology and computer vision technology, people begin to study 3D vision gradually. Computer vision technology is mainly derived from photogrammetry and is mainly used in 2D image recognition and analysis. Today, computer vision technology is powerful enough to be used in a variety of fields [6, 7]. The main principle of computer vision system is to obtain the target image first, then extract the feature, and finally analyze, process, and calculate the feature, in order to make a reasonable decision. Figure 1 is the basic structure of computer vision system, in which the computer is the core part, needs to control the normal operation of each module, and also needs to calculate and output the results.

2.2. Establishment of a Three-Dimensional Dynamic Model of Sports Action. High realistic 3D dynamic model requires not only the deformation of skeletal joints, but also the movement of the associated skin driven by the joints, so as to produce reasonable movement. Therefore, based on the above captured and processed motion images and point cloud data, a 3D dynamic model of sports motion is built, in which the former builds the skeleton model of the sportsman and the latter adds the display appearance to the skeleton model to make the virtual human model more realistic.
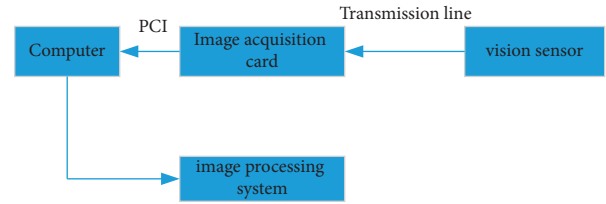

FIGURE 1: Structure of computer vision system.

2.2.1. Establishment of Surface Models. After the establishment of the skeleton model, in order to make the 3D human model more three-dimensional and realistic, we need to build the skin model outside the bone. Based on the point cloud data, the surface model is built by triangle mesh method. The process is as follows: first, select any point as the initial point, connect it with the nearest two points to form a triangle, and then extend along the three vertices to make the mesh grow continuously; finally all points are connected to form a triangle network. In meshing, the meshing of the parts that often produce movement should be more detailed, while the meshing of the parts that do not produce large movement should be larger.

After building the skeleton model and surface model, it is necessary to determine which segments of the skeleton affect a point on the skin mesh and then bind the two together to form a complete 3D human model. The specific expression is

$$F(n) = R_1 + R_2 + R_3. \tag{1}$$

In the formula, R_1, R_2, and R_3 are expressed as the fuzzy recognition parameters of the three-dimensional dynamic model of athletes' movements under the hybrid architecture.

2.2.2. Establishment of Skeletal Models. Skeleton is the most basic support for the human body to complete all kinds of movement. Only skeleton movement can make the whole 3D model of human body move correspondingly. The skeleton model was established according to the optical motion images. Since human motion is mainly embodied in 16 main parts, the established skeleton model is a simplified skeleton model. This is shown in Figure 2.

When describing these 16 parts, the corresponding joints connecting the bones are also included, and the rotation operation of the bones is also the operation of the corresponding joints. The bone names corresponding to the numbers are shown in Table 1.

Only when the skeleton of the 3D model is driven to move can the corresponding movement of the surface model be generated; that is, the construction of the 3D dynamic model of sports movement is realized.

2.3. Error Action Image Acquisition and Processing. This research selects the color camera which includes the CCD sensor and the CMOS sensor to collect the sports athlete wrong movement image. The camera can capture color images, depth images, and bone images of the wrong motion simultaneously [8].
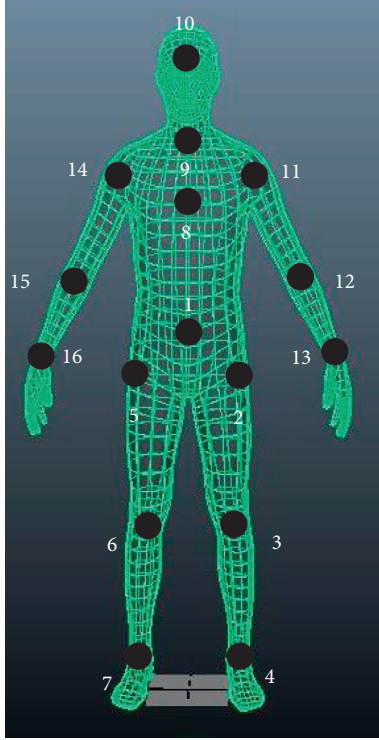
FIGURE 2: Bone model.

TABLE 1: Correspondence between bone name and serial number.

Number	Name
1	Abdomen
2	Right thigh (right hip)
3	Right lower leg (right knee)
4	Right foot (right ankle)
5	Left thigh (left hip)
6	Left lower leg (left knee)
7	Left foot (left ankle)
8	Chest
9	Neck
10	Head
11	Right upper arm (right shoulder)
12	Right forearm (right elbow)
13	Right hand (right wrist)
14	Left upper arm (left shoulder)
15	Left forearm (left elbow)
16	Left hand (left wrist)

The color image and the depth image are transmitted in the form of data stream. The color image resolution is $640 * 480$, the frame number is 30 Fps, the format is Bayer format, and the color data can be encoded as RGB-32 bits. The depth image acquisition process is consistent with color image. The effective position information is 13 bits higher and the user ID information is 3 bits lower. Skeletal images are captured from depth image data and contain 3D coordinates of 20 nodes, visually displaying skeletal maps of athletes [9].

In order to facilitate the application of error motion images, the spatial coordinates of color images, depth images, and bone images are analyzed. The color space, depth space, and skeleton space coordinates are shown in Figure 3.
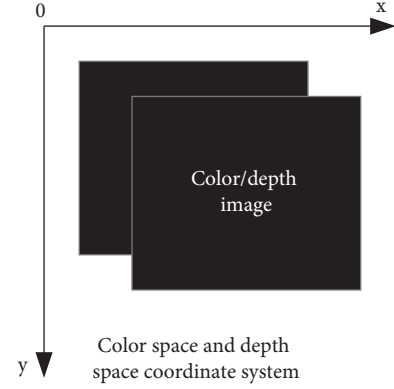

FIGURE 3: Schematic diagram of color space, depth space, and bone space coordinate system.

Set color space pixel coordinates to $(x(t), y(t), z(t))$, depth space pixel coordinates to $(x_s(t), y_s(t), z_s(t))$, and bone space pixel coordinates to $(x_d(t), y_d(t), z_d(t))$.

The skeletal space and depth space coordinate system conversion formula is

$$\begin{cases} x_d(t) = x_s(t)\dfrac{320}{z_s(t)\tan(a/2)}, \\ \\ y_d(t) = y_s(t)\dfrac{240}{z_s(t)\tan(b/2)}, \\ \\ z_d(t) = z_s(t). \end{cases} \quad (2)$$

In formula (2), a represents the horizontal angle of view of the camera with a value of $57°$ and b represents the vertical angle of view of the camera with a value of $43°$.

The conversion formula between depth space and color space coordinate system is

$$\begin{cases} x_s(t) = \dfrac{\Delta l}{z(t)\tan a} + x(t), \\ \\ y_s(t) = y(t), \\ \\ z_s(t) = z(t). \end{cases} \quad (3)$$

In formula (3), Δl represents the displacement of the camera.

The color image, depth image, and skeleton image are transformed into the same coordinate system. In order to facilitate image processing and ignore z direction information, the error motion image is $I(x, y)$, which provides image data for the following error motion feature extraction.

2.3.1. Wavelet Transform of Noisy Motion Image.
An action image wavelet transform with noise can be described by

$$f(x, y) = s(x, y) + e(x, y). \quad (4)$$

In the expression, the image information can be described as $s(x, y)$, the image noise signal can be described as

$f(x, y)$, and the variance of Gaussian white noise as $\bar{\omega}^2$ is described as $e(x, y)$, subject to $N(0, \bar{\omega}^2)$. If there is multiplicative noise signal in the image, it needs to be processed by logarithmic conversion, and the multiplicative noise can be converted to Gaussian white noise in logarithmic dimension.

Image wavelet transform with noise has the following characteristics:

(1) Wavelet transform coefficients have certain spatial orientation characteristics. LH represents edge data in horizontal direction, HL represents edge data in vertical direction, and HH represents edge data in diagonal direction. These edge data will provide a strong basis for the denoising of subsequent images [10, 11].

(2) The series of white noise can be transformed by wavelet base coefficients so that it can be represented by zero mean white noise.

(3) In the wavelet transform domain of the noisy image, the signal energy is mostly near the coefficient with higher absolute value, while the noise is completely the opposite [12]. Therefore, a threshold value is set, the coefficient not exceeding the threshold value is set to 0, and the wavelet coefficients beyond the threshold value are stored. Depending on the wavelet coefficients after processing, it can be understood that this part of the coefficient is a normal signal in the image, while the residual coefficient has noise to obtain the specific location of the noise.

2.3.2. Image Denoising under Quantization Matrix Threshold.

The multiresolution of image is decomposed by decomposition algorithm, and the wavelet coefficient matrix is constructed, then the threshold of matrix is quantified, and the new wavelet coefficient matrix is obtained. Finally, under the threshold of quantization matrix, the image in the matrix is reconstructed by reconstruction algorithm to obtain the denoised image [13, 14]. The specific procedures are as follows:

(1) According to the decomposition method, at the quantization matrix threshold, the orthogonal wavelet basis is generated in the wavelet function, and the initial image is decomposed into three layers of wavelet to obtain HH, LH, HL detail component and LL smooth component of the image signal

(2) Wavelet thresholds are quantified

How to select threshold and quantize threshold is the key step in denoising. In a certain sense, at the quantization matrix threshold, this step is related to the quality of the image signal reconstruction. Therefore, this paper uses soft threshold algorithm, quantization threshold γ defined as

$$\gamma = \begin{cases} \text{sign}(\gamma), \\ 0. \end{cases} \tag{5}$$

2.4. Spatial and Frequency Domain Feature Extraction of Sports Athletes' Action Contour Image.

Based on the above notable image, the number of pixels in the recognized sports athletes action image is calculated to get the target pixel number and the number of frames in the period. Suppose a sportsman has a frame n image in his movement, the formula of the energy generated by the sportsman's movement is expressed as follows:

$$G(x, y) = \frac{\sum_n^{i=1} B_i(x, y)}{n}. \tag{6}$$

In formula (6), $G(x, y)$ represents the gray image and $B_i(x, y)$ represents the gesture parameters of the motion image.

After the above basic processing, the high-frequency part and low-frequency part of the athlete's movement are distinguished by the discrete cosine transformation method to effectively extract the frequency domain characteristics of the athlete's movement posture [15]. Five sets of motion images, in which $\{f(x, y), x, y = 0, 1, 2, \ldots, N-1\}$ is represented as $N \times N$, are represented by the DCT formula as follows:

$$F = f(x, y) \cdot \frac{(2x+1)u\pi}{2N} \cdot \frac{(2y+1)v\pi}{2N}. \tag{7}$$

In formula (7), $f(x, y)$ represents the gray value of the coordinates of pixels $f(x, y)$ in the sports athletes' action image, u and v, respectively, represent the horizontal and vertical conversion rates of pixels in the image, and N represents the DC part of the image features.

After the above calculation, the same transformation coefficient matrix as that of the original athlete can be obtained; that is, the frequency domain characteristics of the athlete's movement can be reasonably reflected. Based on the above transformation, a human posture model can be established. Its structure diagram is shown in Figure 4.

The human contour model is represented as follows:

$$v = \{N, R, F(D)\}. \tag{8}$$

In formula (8),

$$\begin{cases} N = \dfrac{\overline{W}}{H}, \\[2mm] R = \dfrac{A_r}{W \cdot H}, \\[2mm] D = [d_1, d_2, \cdots, d_n], \\[2mm] F(D) = R \cdot D. \end{cases} \tag{9}$$

In formula (9), $\overline{W}$ represents the average width parameter of the contour of the human body, H and W, respectively, represent the height and width values of the athlete's motion image, A_r represents the area of the athlete's motion target, and R represents the characteristic value in the athlete's motion model [16].
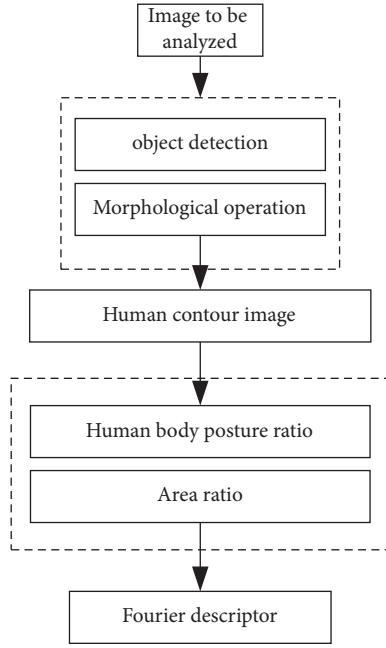
FIGURE 4: Human posture model structure.

2.5. Edge Restoration Based on Canny Operator.

The Canny method uses a suitable Gaussian function to smooth the image by columns and rows and also to convolve the image signals [17]. The Gaussian functions used are

$$G(x, y) = \frac{e^{-\left(\left(x^2 + y^2\right)/2\sigma^2\right)}}{2\pi\sigma_2}. \quad (10)$$

In formula (10), σ represents the Gaussian curve, which controls the smoothing intensity.

Canny operator is constructed on the basis of 2D convolution $\nabla G(x, y) \cdot f(x, y)$, obtains edge direction and intensity, and identifies edge features through threshold [18, 19].

The two-dimensional convolution of $\nabla G(x, y)$ is divided into two one-dimensional convolvers and the result is

$$\frac{\partial G(x, y)}{\partial x} = \frac{y^2}{20^2},$$
$$\frac{\partial G(x, y)}{\partial y} = \frac{x^2}{20^2}. \quad (11)$$

Convolution is then performed with $f(x, y)$ for each convolver to obtain

$$E_x = \frac{\partial G(x, y)}{\partial x} * f,$$
$$E_y = \frac{\partial G(x, y)}{\partial y} * f. \quad (12)$$

The following conditions are met:

$$A(i, j) = E_x^2 + E_y^2, a(i, j) = \arctan \frac{E_y(i, j)}{E_x(i, j)}, \quad (13)$$

where the edge strength is described as $A(i, j)$ and the direction perpendicular to the edge is described as $a(i, j)$.

In the process of extracting edge features by Canny operator [20, 21], the selection of threshold is very important. If the threshold is too large, the edge feature recognition will be intermittent, and the low threshold will lead to false contour of the image. In this paper, we use double threshold method to solve the problem of threshold selection. Firstly, we propose two thresholds $r1$ and $r2$, and $2^{r1} = r2$, so that we can obtain two threshold edge images $N_1[i, j]$ and $N_2[i, j]$ because $N_2[i, j]$ is obtained by using high threshold. Therefore, the double threshold algorithm needs to link the edge into the contour within $N_2[i, j]$, and when the contour is linked with each other, the method can connect the inner edge of the contour by searching for the coordinates in the 8 fields of $N_1[i, j]$, so that the algorithm can continuously collect the edges within $N_1[i, j]$ until $N_2[i, j]$ is connected. The flow of the algorithm is as follows:

(1) Calculate the derivative Gx, Gy of the image gray scale according to the reciprocal operator and simultaneously calculate the gradient direction and size of the derivative.

(2) If the gray value of an image pixel in a certain direction is low, it is necessary to set the pixel coordinate to 0, that is, nonedge pixels.

(3) The threshold is calculated based on the histogram of the image. If the gray level value exceeds the threshold, the gray level area is the edge of the image, and vice versa. Then the continuity between the point above the threshold and the previous point is found. If the gray level value is not continuous, then the neighborhood coordinates of the point are found in the stack of the low threshold, and these coordinates are connected, so that the search is iterated until the overall contour is complete [22].

2.6. Intelligent Recognition of Athletes' Wrong Actions.

Bayesian classifier is the key to intelligent recognition of sports athletes' wrong movements. Therefore, a classifier is designed according to Bayesian algorithm [23–25].

Bayesian classifier is a kind of classification method designed according to Bayesian algorithm on the premise of conditional independence assumption. For the training sample set, the joint probability distribution function of the input and output of the training set is calculated firstly, based on which, the maximum posterior probability output of the input data is calculated by Bayes algorithm [26, 27].

Suppose the training dataset is $R = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$, and the j feature of the i training sample is $x_i^{(j)}$, which is composed of a plurality of values and recorded as a_{jl}.

For input value x, the priori probabilities and conditional probabilities are calculated as follows:

$$\begin{cases} P(Y = c_k) = \sum_{i=1}^{N} I(y_i = c_k), & k = 1, 2, \ldots, \\ P(X^{(j)} = a_{jl} \mid Y = c_k) = \sum_{i=1}^{N} I(x_i^{(j)} = a_{jl}, y_i = c_k). \end{cases} \tag{14}$$

In formula (14), $Y = c_k$ represents the output space corresponding to the input space; N represents the number of training datasets; and $P(X^{(j)} = a_{jl} \mid Y = c_k)$ represents the joint probability distribution function [28, 29].

For a given input x, the corresponding output space is expressed as

$$P(Y = c_k) \prod_{j=1}^{n} P(X^{(j)} = x^{(j)} \mid Y = c_k). \tag{15}$$

Determine the category of input x according to formula (15), and the determined formula is expressed as

$$y = \arg \max P(Y = c_k) \prod_{j=1}^{n} P(X^{(j)} = x^{(j)} \mid Y = c_k). \tag{16}$$

Using Bayes to estimate the conditional probability, the result is

$$P(X^{(j)} = a_{jl} \mid Y = c_k) = \frac{\sum_{i=1}^{N} I(x_i^{(j)} = a_{jl}, y_i = c_k)}{\sum_{i=1}^{N} I(y_i = c_k) \cdot S_j}. \tag{17}$$

In formula (17), S_j represents the total number of characteristic values [30, 31].

Based on formula (17), a classifier that maximizes a posteriori probability is obtained, and its expression is

$$f(x) = \arg \max P\left(\frac{Y = c_k}{X = x}\right). \tag{18}$$

The above process completes the design of Bayesian classifier, which provides a solid support for intelligent recognition of sports athletes' wrong movements.

After the above basic processing, the feature vectors and their tags are given to the classifier for action recognition. In general, not all of the data contained in the feature data is useful but also contains data that is not relevant. To avoid this, the feature dimension of the data needs to be subtracted [32] by classifying it as shown in Figure 5.

Based on the above process, SVM is used to classify the input features in advance. When there is an error classification, there is a support vector near the hyperplane of the classification [33, 34]. In classification, the test set is represented by T, T_{su} is the support vector set, and k is the number of classifiers. The process is as follows:

Step 1: support vector machine algorithm is used to calculate the corresponding support vector and to solve the coefficient and constant b.

Step 2: if T is not an empty set, take $x \in T$, and if T is an empty set, stop [35].
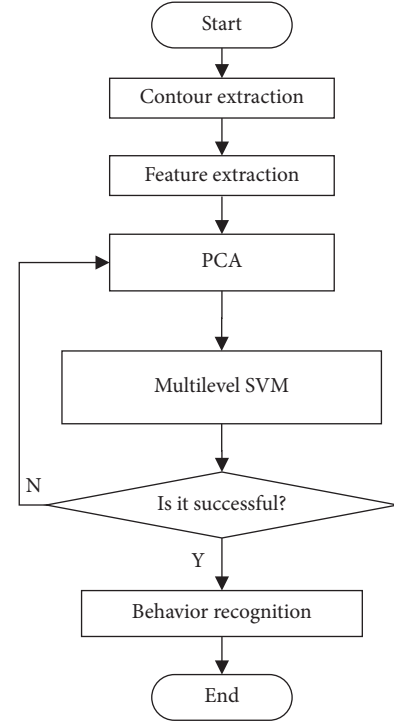

Figure 5: Flowchart of hybrid classification model.

Step 3: calculate $g(x) = \sum y_i \alpha_i K(x_i, x) - b$.

Step 4: if $g(x) > \varepsilon$, then $f(x) = \text{sgn}(g(x))$ is directly used as the output of the classifier, and if $g(x) > \varepsilon$, it is subsumed into the classifier for classification [36].

Based on the above process, we classify the support vectors and complete the recognition of the athlete's movement.

3. Simulation Experiment Design and Result Analysis

3.1. Image Acquisition of Experimental Samples. The main equipment used in the experiment is the image acquisition device, which can make the wrong movement of sports transient and fast. The experiment adopts a short time acquisition and storage system to realize the acquisition and storage of the experimental image, which is composed of the camera, the acquisition card, the cable, the computer, and the acquisition software. The parameters of the camera are shown in Table 2.

The experimental sample selects two sports videos as the experimental objects. The two videos are football match and gymnastics match, respectively. In order to verify the reliability and effectiveness of the model, the following experiments are designed. In view of the experimental object sports video frequency to carry on the athlete wrong movement recognition, the comparison method selects reference [3] proposed based on the multicharacteristic fusion athlete posture recognition method and reference [4] proposed based on the multithreshold optimization movement image outline characteristic extraction method. The recognition results of different methods are given in the

TABLE 2: Camera parameters.

Parameter name	Numerical value	Company
Image type	Mono/color	—
Pixel depth	8/10	Bit
Pixel size	5.5×5.5	μm
Data output type	Camera link	-
Maximum frame rate	340	fps
Maximum resolution	2048×1088	Pixels
Overall dimension	$63.5 \times 63.5 \times 44.1$	mm

(a)   (b)

FIGURE 6: Detection results of the algorithm studied in this paper. (a) Test sample 1. (b) Test sample 2.

(a)   (b)

FIGURE 7: Recognition results of athlete's human posture recognition method based on multifeature fusion. (a) Test sample 1. (b) Test sample 2.

experimental results. The specific recognition results are shown in Figures 6–8.

Through the analysis of the two images, it can be seen that the algorithm studied in this paper can clearly identify the key wrong actions of athletes. There are some redundant results in the recognition results of athletes' human posture recognition method based on multifeature fusion and moving image contour feature extraction method based on multithreshold optimization, and the recognition results are not unique, and the error is relatively large compared with the actual action.

3.2. Image Denoising Experiment. To verify the performance of the proposed method, within 250 frames, two frames are extracted: as shown in Figures 9 and 10, there is uneven noise distribution in the two frames, the contrast is too high, the image is denoised by the proposed method, and the result is shown in Figures 9 and 10.

It can be seen from Figures 9 and 10 that after the image denoising is completed, the noise elimination effect in the initial image is good, which provides a strong basis for subsequent action image recognition.

(a)

(b)

FIGURE 8: Recognition results of moving image contour feature extraction method based on multithreshold optimization. (a) Test sample 1. (b) Test sample 2.
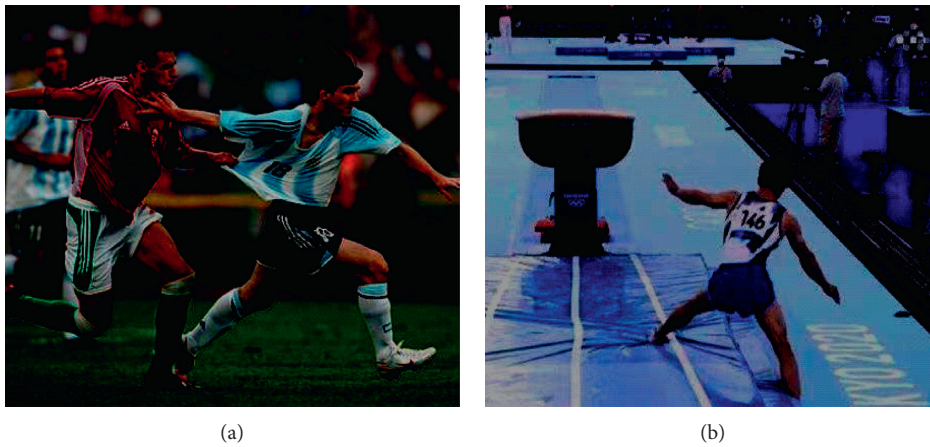

(a)

(b)

FIGURE 9: Takeoff image with noise and back high jump image. (a) Test sample 1. (b) Test sample 2.

(a)

(b)

FIGURE 10: Two-frame image denoising effect. (a) Test sample 1. (b) Test sample 2.

3.3. Accuracy Test of Athletes' Wrong Movement Recognition. In addition, in order to further verify the better accuracy of the proposed algorithm, we compare it with the traditional algorithm, establish a 3D visual inspection model, and then test two different methods for many times, so as to get the results shown in Figure 11. As can be seen from Figure 5, the accuracy is above 90%, while the progress of the traditional algorithm is between 70% and 77%. So the algorithm studied in this paper has better accuracy and can control the error in a reasonable range.
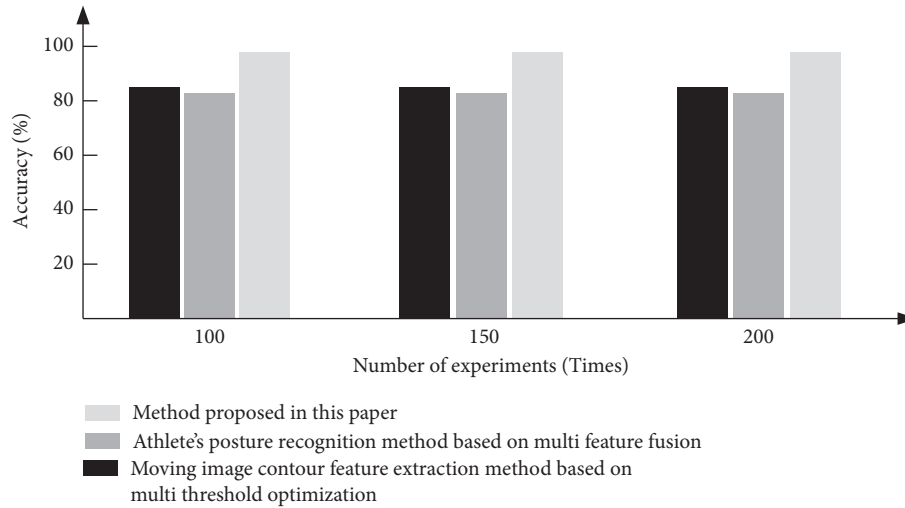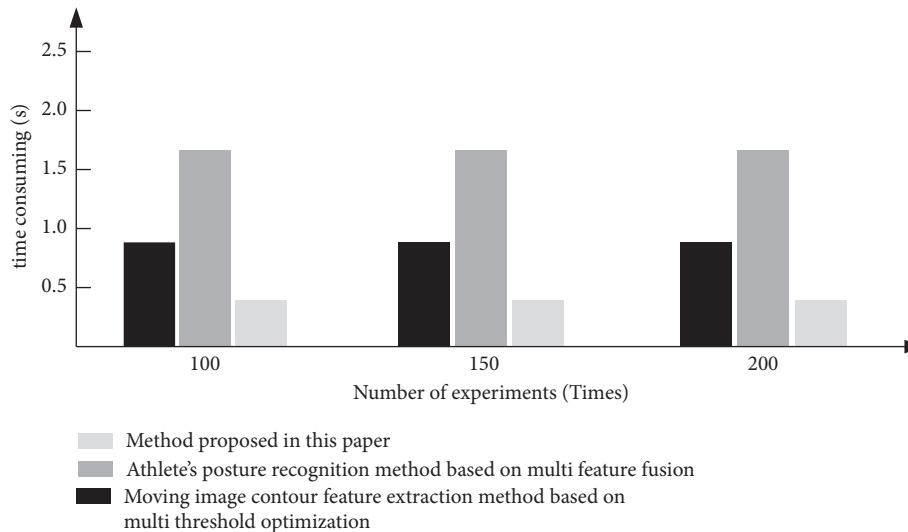
FIGURE 11: Comparison of recognition accuracy.

FIGURE 12: Time-consuming intelligent recognition of wrong actions.

3.4. Time-Consuming Comparative Test of Different Methods. Comparing the three methods for the intelligent recognition of erroneous actions of 200 successful samples of erroneous actions takes time, and the comparison results are shown in Figure 12.

As shown in Figure 12, the intelligent recognition time of reference [3] method for each wrong action is 0.25 s on average, the intelligent recognition time of reference [4] method for each wrong action is 0.39 s on average, and the intelligent recognition time of each wrong action is 0.08 s on average. Therefore, the proposed method has high speed and accuracy of intelligent recognition of athlete's wrong movement, which fully shows that the model has good intelligent recognition performance.

To sum up, the intelligent recognition method of sports athletes' wrong actions based on image vision has good effect and high recognition accuracy. It can complete the recognition of sports athletes' wrong actions in a shorter time, and the recognition effect is ideal.

4. Conclusion

(1) This paper puts forward an intelligent recognition method of athletes' wrong action based on image vision, which improves the application disadvantages of traditional athletes' wrong action recognition methods

(2) The intelligent recognition method of sports athletes' wrong actions based on image vision has ideal recognition accuracy

(3) The image noise is low, which can complete the recognition of sports athletes' wrong actions in a shorter time

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

[1] E. Chen and J. fan, "Method of human action feature extraction and recognition based on MEM-LBP [J]," *Application Research of Computers*, vol. 35, no. 4, pp. 1277–1280, 2018.

[2] X. Luo, H. Li, X. Yang, Y. Yu, and D. Cao, "Capturing and understanding workers' activities in far-field surveillance videos with deep action recognition and bayesian nonparametric learning," *Computer-Aided Civil and Infrastructure Engineering*, vol. 34, no. 4, pp. 333–351, 2019.

[3] S. Liu, "Player's posture recognition algorithm based on multi-feature fusion," *Information & Technology*, vol. 23, no. 8, pp. 17–19, 2019.

[4] C. Chen, "Extraction method of contour features by multi-threshold optimization for motion images," *Journal of Shenyang University of Technology*, vol. 41, no. 3, pp. 315–319, 2019.

[5] Y. P. Li, T. T. Liu, and L. Zhang, "Human action recognition based on deep learning," *Application Research of Computers*, vol. 41, no. 4, pp. 261–265, 2020.

[6] S. Cagnoni, H. Al-Sahaf, Y. Sun, B. Xue, and M. Zhang, "Special issue on evolutionary computer vision, image processing and pattern recognition," *Applied Soft Computing*, vol. 97, no. 12, Article ID 106675, 2020.

[7] L. Antoniomeira, L. E. Techiopereira, M. E. Rozalinosantos et al., "USPLeaf: automatic leaf area determination using a computer vision system1," *Revista de Ciencias Agronomicas*, vol. 51, no. 4, pp. 1–10, 2020.

[8] Z. Liao, H. Hu, and Y. Liu, "Action recognition with multiple relative descriptors of trajectories," *Neural Processing Letters*, vol. 51, no. 1, pp. 287–302, 2020.

[9] C. Cao, Y. Zhang, C. Zhang, and H. Lu, "Body joint guided 3-D deep convolutional descriptors for action recognition," *IEEE Transactions on Cybernetics*, vol. 48, no. 3, pp. 1095–1108, 2018.

[10] S. Chakraborty, S. H. Shaikh, A. Chakrabarti, and R. Ghosh, "An image denoising technique using quantum wavelet transform," *International Journal of Theoretical Physics*, vol. 59, no. 11, pp. 3348–3371, 2020.

[11] A. Muthukrishnan, J. Charles Rajesh kumar, D. Vinod Kumar, and M. Kanagaraj, "Internet of image things-discrete wavelet transform and Gabor wavelet transform based image enhancement resolution technique for IoT satellite applications," *Cognitive Systems Research*, vol. 57, no. 10, pp. 46–53, 2019.

[12] C.-J. Zhang, X.-Y. Huang, and M.-C. Fang, "MRI denoising by NeighShrink based on chi-square unbiased risk estimation," *Artificial Intelligence in Medicine*, vol. 97, no. 6, pp. 131–142, 2019.

[13] C. Rodrigues, Z. M. Assis Peixoto, and F. Magalhaes Freitas Ferreira, "Ultrasound image denoising using wavelet thresholding methods in association with the bilateral filter," *IEEE Latin America Transactions*, vol. 17, no. 11, pp. 1800–1807, 2019.

[14] C. Vimala and P. A. Priya, "Artificial neural network based wavelet transform technique for image quality enhancement," *Computers & Electrical Engineering*, vol. 76, no. 6, pp. 258–267, 2019.

[15] A. Ullah, J. Ahmad, K. Muhammad, M. Sajjad, and S. W. Baik, "Action recognition in video sequences using deep Bi-directional LSTM with CNN features," *IEEE Access*, vol. 6, no. 99, pp. 1155–1166, 2018.

[16] H. Wu and Z. Cheng, "Action recognition algorithm based on complexity measure and multi-scale motion coding," *Optical Technique*, vol. 44, no. 04, pp. 427–434, 2018.

[17] I. Al-Nahhal, O. A. Dobre, E. Basar, C. Moloney, and S. Ikki, "A fast, accurate, and separable method for fitting a Gaussian function [tips & tricks]," *IEEE Signal Processing Magazine*, vol. 36, no. 6, pp. 157–163, 2019.

[18] X. Lu and Y. Zhang, "Human body flexibility fitness test based on image edge detection and feature point extraction," *Soft Computing*, vol. 24, no. 12, pp. 8673–8683, 2020.

[19] F. Yuan, G. Li, X. Xia, B. Lei, and J. Shi, "Fusing texture, edge and line features for smoke recognition," *IET Image Processing*, vol. 13, no. 14, pp. 2805–2812, 2019.

[20] Z. Wu and Z. Zheng, "Action recognition algorithm based on deep learning and motion information," *Computer Engineering and Design*, vol. 39, no. 08, pp. 2668–2674, 2018.

[21] Y. Shi, M. Sun, Z. Li, J. Luo, and M. Yang, "Action recognition based on motion history image and convolution neural network," *Natural Science Journal of Xiangtan University*, vol. 041, no. 02, pp. 109–117, 2019.

[22] N. Slavkovic and M. Bjelica, "Risk prediction algorithm based on image texture extraction using mobile vehicle road scanning system as support for autonomous driving," *Journal of Electronic Imaging*, vol. 28, no. 3, pp. 1–033034.13, Article ID 033034, 2019.

[23] D. Ludl, T. Gulde, and C. Curio, "Enhancing data-driven algorithms for human pose estimation and action recognition through simulation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3990–3999, 2020.

[24] M. Hajizadeh and H. Ebrahimnezhad, "Eigenspace compression: dynamic 3D mesh compression by restoring fine geometry to deformed coarse models," *Multimedia Tools and Applications*, vol. 77, no. 15, pp. 19347–19375, 2018.

[25] P. X. Fuchs, A. Fusco, J. W. Bell, S. P. von Duvillard, C. Cortis, and H. Wagner, "Movement characteristics of volleyball spike jump performance in females," *Journal of Science and Medicine in Sport*, vol. 22, no. 7, pp. 833–837, 2019.

[26] K. Hirokatsu, S. Yutaka, A. Yoshimitsu, O. Shoko, and M. Yasuhiro, "Temporal and fine-grained pedestrian action recognition on driving recorder database," *Sensors*, vol. 18, no. 2, pp. 627–641, 2018.

[27] F. Meng, B. Xu, T. Zhang, B. A. Muthu, and C. B. Sivaparthipan, "Application of AI in image recognition technology for power line inspection," *Energy Systems*, vol. 248, no. 1, pp. 1–23, 2021.

[28] U. B. Angadi, A. Rai, and G. Uma, "MBFerns: classification and extraction of actionable knowledge using Multi-Branch Ferns-based Naive Bayesian classifier[J]," *Soft Computing*, vol. 25, no. 6, pp. 1–13, 2021.

[29] T. T. T. Nguyen, T. T. Nguyen, R. Sharma, and A. W.-C. Liew, "A lossless online bayesian classifier," *Information Sciences*, vol. 489, no. 7, pp. 1–17, 2019.

[30] H. Zhang, B. Wen, J. Liu, and Y. Zeng, "The prediction and error correction of physiological sign during exercise using bayesian combined predictor and naive bayesian classifier," *IEEE Systems Journal*, vol. 13, no. 12, pp. 4410–4420, 2019.

[31] S. Wang, S. Zhang, T. Wu, Y. Duan, L. Zhou, and H. Lei, "FMDBN: a first-order Markov dynamic Bayesian network classifier with continuous attributes," *Knowledge-Based Systems*, vol. 195, no. 3, Article ID 105638, 2020.

[32] A. Maa, B. Ac, C. Sv, and K. Abdessamad, "A single Bayesian network classifier for monitoring with unknown classes," *Engineering Applications of Artificial Intelligence*, vol. 85, no. 10, pp. 681–690, 2019.

[33] C. Mao, L. Lu, and B. Hu, "Local probabilistic model for Bayesian classification: a generalized local classification model [J]," *Applied Soft Computing*, vol. 93, Article ID 106379, 2020.

[34] Z. Shen, Z. Man, Z. Cao, and J. Zheng, "A new intelligent pattern classifier based on structured sparse representation," *Computers & Electrical Engineering*, vol. 84, no. 5, Article ID 106641, 2020.

[35] J. Liu, H. Rahmani, N. Akhtar, and M. Ajmal, "Learning human pose models from synthesized data for robust RGB-D action recognition," *International Journal of Computer Vision*, vol. 127, no. 8, pp. 1545–1564, 2019.

[36] L. Dong, D. Moc, E. Kmc et al., "A computer vision-based method for spatial-temporal action recognition of tail-biting behaviour in group-housed pigs," *Biosystems Engineering*, vol. 195, no. 7, pp. 27–41, 2020.