

Research Article

Standardized Judgment Method of Shooting Training Action Based on Digital Video Technology

Guo Qing  and HuBao Hui

Physical Education Department of Gansu Agricultural University, Gansu 730070, Lanzhou, China

Correspondence should be addressed to Guo Qing; guoq@gsau.edu.cn

Received 11 October 2021; Revised 1 November 2021; Accepted 11 November 2021; Published 20 December 2021

Academic Editor: Bai Yuan Ding

Copyright © 2021 Guo Qing and HuBao Hui. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aiming at the difficulty of standardizing the action of basketball shooting training, a new method of standardizing the action of basketball shooting training is proposed based on digital video technology. The digital video signal representation, video sequence coding data structure, and video sequence compression coding method are analyzed, and the pixels of basketball shooting training action position space are sampled to collect basketball shooting training images. The time difference method is used to extract the movement target of basketball shooting training from a digital video sequence. Based on digital video technology, the initial background image is estimated, and the update rate is introduced to update the background estimation image. According to the pixel value sequence of the basketball shooting training image, the pixel model of the basketball shooting training image is defined and modified. By judging whether the defined pixel value matches the background parameter model, the standardization of shooting training can be realized. The experimental results show that the proposed method has good stability, high precision, and short time in determining the standardization of shooting movement, can correct the wrong shooting movement in real time, and can effectively guide basketball shooting training.

1. Introduction

Basketball is quite different from other sports. It is a high-intensity and comprehensive sport [1, 2]. Basketball belongs to the same field antagonistic event group dominated by technical and tactical ability, and the technical and tactical level is the decisive factor for the competitive level of basketball [3]. In the actual competition process, basketball players need to have diversified basketball qualities to ensure the victory of the competition. Among them, the more important point is the coordination and stability of athletes' physical functions. In the development of basketball, shooting is its key offensive technology. The essence of a basketball game is a shooting game, which also shows that the stability of shooting has an important relationship with the outcome of the game [4]. Among them, the factors affecting the standardization of athletes' shooting action mainly include athletes' bodies, technology, and psychology. In the process of competition, athletes need to ensure that they master the nature, time, and score of the competition [5]. In the actual basketball game, the attacking team needs to use different techniques or tactics to create more shooting opportunities and ensure

shooting scores [6, 7]. The defensive team should actively defend and prevent the other team from scoring. The accuracy and standard of shooting in basketball are directly related to the score. Therefore, it is of great significance to reasonably test and judge the shooting action of basketball.

At present, scholars in related fields have carried on the research on action judgment and obtained some research results. Reference [8] proposed a motion similarity judgment method based on motion primitives. Based on the computational model of kinematics, the similarity of motion and its performance to the human similarity judgment of the same motion are determined. By performing the action similarity task and comparing it with the computational model solving the same task, the action similarity judgment was realized by classifying the actions based on the learned kinematics primitive. The method has high reliability and provides necessary basis for human action classification. Reference [9] proposed a basketball motion image target detection method based on an improved Gaussian mixture model. Edge detection, gray processing, target capture, target recognition, image detection, and other technologies

are integrated into basketball sports video, and Gaussian probability density mixing is used to select the appropriate number of continuously updated parameters and each pixel area to achieve basketball sports image detection. The method is effective to some extent. However, the above methods are difficult to determine the standardization of basketball shooting training movements.

In view of the above problems, a method to judge the movement standardization of shooting training based on digital video technology is proposed. The innovation of the research method is to use the time difference method to extract the movement target of basketball shooting training. Based on digital video technology, the initial background image is estimated and the update rate is introduced to update the image. According to the pixel value sequence of the basketball shooting training image, the pixel model of the basketball shooting training image is defined and modified. By discriminating the matching relation between pixel value and background parameter model, the standardization judgment of shooting training movement can be realized. Compared with the previous research results, the method designed based on digital video technology has better stability, high accuracy, and short time and can correct the wrong shooting action in real time, which can effectively guide basketball shooting training.

2. Digital Video Technology

Digital video technology is to first use video capture equipment such as cameras to convert the color and brightness information of external images into electrical signals, and then record them into storage media [10–12]. Digital video is video recorded in digital form, as opposed to analog video. Digital video has different production methods, storage methods, and broadcast methods. For example, digital video signals are generated directly through digital cameras and stored on digital tape, P2 card, blue disc, or disk, so as to obtain different formats of digital video, which is then played on a PC, a specific player, etc.

2.1. Representation of Digital Video Signal. Video is described as a group of continuous images, and each image is regarded as a two-dimensional pixel array. The color representation of each pixel includes three components: red R , green G , and blue B , which is called the RGB space representation of the image. The color coordinates used for the three digital TV systems are different. For digital video capture and display, all three digital TV systems use RGB primary colors, but the definition of each primary color spectrum is slightly different. For the transmission of digital video signal, in order to reduce the required bandwidth and be compatible with monochrome digital TV system, the brightness/chroma coordinate system is adopted [13, 14]. The color coordinates used in the NTSC, PAL, and SECAM systems are all derived from the YUV coordinates used for PAL, and YUV is derived from the XYZ coordinates. According to the relationship between the RGB primary color and the YUV primary color, the value of the luminance component Y can be determined by the value of RGB . The

two chromaticity values U and V are proportional to the color differences $B - Y$ and $R - Y$, respectively, and are adjusted to the desired range. The classic conversion relationship between the YUV coordinate system and the RGB primary color value is

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (1)$$

The conversion of two color spaces is based on the characteristics of a human visual system: in RGB space, if one of the three signals R , G , and B changes, the color of the total image will change, and the human eye can easily detect this change. However, human eyes have different responses to the changes of Y , U and V signals. Among them, they are sensitive to the changes of luminance signals, but not very sensitive to the changes of chrominance signals. In this way, we can consider more luminance signals and adopt some processing methods for chrominance signals to improve the compression ratio.

2.2. Data Structure of Digital Video Sequence Coding. In the coding scheme, the video sequence is divided and multiplexed by multiple layers to establish such a data structure.

- (1) Sequence: the video sequence starts with the sequence header, including several image groups, and ends with a sequence end code.
 - (2) Group of pictures (GOP): GOP is a head followed by a series of images, which allows fast random access to the sequence, fast search, and editing. It is the smallest coding unit that can be decoded independently in the sequence [15]. The first image in the GOP is an intracoded image (I frame), followed by a forward prediction coded image (P frame) and a bidirectional prediction image (B frame). Each GOP has only one I frame, and this I frame is used as the first frame to start coding. The P frame is encoded by motion-compensated prediction relative to the previous I frame or P frame, and the P frame can be used as a reference frame for other P frame or B frame coding. B frame is encoded by motion compensation prediction of two frames, one is the past frame, and the other is the future frame. The frame arrangement of the GOP is shown in Figure 1.
- The standard does not specify the number of P and B frames in a GOP, nor their specific sequence, except for the first frame, and only one frame is I frame. Any sequence and frame number can be used to design the encoding scheme. The prediction of each P and B frame is based on the previous reference prediction frame. Too many frames in the group layer will affect the quality and compression ratio of coding. Therefore, 10–15 B frames are generally selected in each group layer, and 2–3 B frames are separated between the two P frames.
- (3) Image: image is the basic coding unit of video sequence [16]. The image is composed of three rectangular matrices representing the luminance Y and two

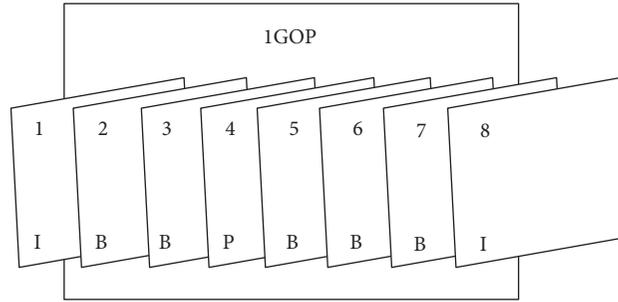


FIGURE 1: Frame arrangement of GOP.

chrominance C_b and C_r values. Each video standard divides the image into macroblock groups. H.261 and H.263 use a fixed macroblock structure, whereas MPEG1/2 allows a flexible structure, and MPEG4 arranges a variable number of macroblocks into a group.

- (4) Group of blocks (GOB): H.261 and H.263 divide the image into GOBs. Each GOB includes three macroblock lines and 11 macroblocks in each GOB line, and the GOB header defines the position of the GOB in the image.
- (5) Slice: a slice consists of several successive macroblocks into a unit. The size of the slice can be changed. The slice layer provides anti-interference ability against data errors.
- (6) Macroblock (MB): Macro block (MB): MB is a basic concept in video coding technology. It can divide the image into many blocks of different sizes and implement different compression strategies in different locations. A coded image is usually composed of several macroblocks. A macroblock is composed of a luminance pixel block and two additional chrominance pixel blocks. In general, the brightness block is 16×16 pixel block, chroma block is 8×8 -size pixel block; several macroblocks in each image are arranged in the form of slices. The video coding algorithm takes macroblocks as units, encodes

macroblocks one by one, and organizes them into a continuous video code stream.

- (7) Block: block is the smallest coding unit in the standardized video coding algorithm. It consists of 8×8 pixel composition [17].

2.3. Digital Video Sequence Compression Coding Method

2.3.1. Transform Coding. Transform coding does not directly encode the spatial image signal, but first maps and transforms the spatial signal to another orthogonal vector space to generate a batch of transform coefficients, and then encodes these transform coefficients [18–20]. In the digital video sequence image compression and coding technology, the compression performance and error of discrete cosine transform (DCT) are very close to those of K-L transform, and DCT has the characteristics of moderate computational complexity, separability, and fast algorithm. Therefore, there are many schemes using DCT coding in image data compression [21–23]. At present, DCT is used in almost all transform-based image encoders.

Assuming that the range of the spatial variables is $x = 0, 1, \dots, N-1$ and $y = 0, 1, \dots, N-1$, and the range of the frequency domain variables is $u = 0, 1, \dots, N-1$, $v = 0, 1, \dots, N-1$, the two-dimensional discrete cosine sine transform formula is

$$F(u, v) = \frac{2}{N} E(u)E(v) \left[\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left(\frac{2x+1}{2N} u\pi\right) \cos\left(\frac{2y+1}{2N} v\pi\right) \right]. \quad (2)$$

In formula (2), when $u = 0, v = 0$, $E(u), E(v) = 1/\sqrt{2}$. The two-dimensional inverse discrete cosine transform formula is

$$f(x, y) = \frac{2}{N} \left[\sum_{u=0}^{N-1} \sum_{v=0}^{N-1} E(u)E(v)F(u, v) \cos\left(\frac{2x+1}{2N} u\pi\right) \cos\left(\frac{2y+1}{2N} v\pi\right) \right]. \quad (3)$$

In formula (3), when $u = 0, v = 0$, $E(u), E(v) = 1/\sqrt{2}$. The core of the two-dimensional discrete cosine transform is separable, so both the forward and inverse transforms can

decompose the two-dimensional transform into a series of one-dimensional transforms (rows, columns) for calculation [24, 25].

In the MPEG series, since the basic unit of DCT transformation is a luminance block or a chrominance

block, the size is 8×8 , so $N = 8$ can be used in formula (2), so that

$$F(u, v) = \frac{1}{4} E(u) E(v) \left[\sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos\left(\frac{2x+1}{16} u\pi\right) \cos\left(\frac{2y+1}{16} v\pi\right) \right]. \quad (4)$$

In practical application, considering the characteristics of separable variables in formula (4), rewriting the latter part of formula (4) can obtain

$$\sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos\left(\frac{2x+1}{16} u\pi\right) \cos\left(\frac{2y+1}{16} v\pi\right) = \sum_{x=0}^7 \cos\frac{2x+1}{16} u\pi \left(\sum_{y=0}^7 f(x, y) \cos\frac{2y+1}{16} v\pi \right). \quad (5)$$

Set an intermediate variable $F(i, v)$; then,

$$F(i, v) = \sum_{j=0}^7 f(i, j) \cos\frac{2j+1}{16} v\pi. \quad (6)$$

The coefficient can be written as

$$F(u, v) = \frac{1}{4} E(u) E(v) \left[\sum_{x=0}^7 F(i, v) \cos\frac{2x+1}{16} u\pi \right]. \quad (7)$$

It can be seen that after such processing, the two-dimensional DCT transform is decomposed into one row DCT and one column DCT transform, which is easy to be realized by computer. For the inverse transformation, the variables can still be separated to facilitate the implementation of the algorithm.

2.3.2. Predictive Coding. Predictive coding is a technique to improve compression performance through statistical redundancy. Based on the previously encoded pixel values, the encoder can estimate and predict the pixel values to be encoded and decoded [26–28]. For a large number of static or slowly varying regions in the sequence image, the conditional patching method can be used to store the first frame image in the reference frame and send it to the other party. Then, the predicted value $\hat{I}_k(z)$ of the pixel sampling value $I_k(z)$ of the k frame image at $z = (x, y)$ is the restoration value $I_{k-1}'(z)$ of the pixel value at the same position of the $k-1$ frame image. The interframe difference is expressed as

$$FD_k(z) = I_k(z) - I_{k-1}'(z). \quad (8)$$

For sequence images with relatively moderate amount of motion, encode the frame difference $FD_k(z)$ of the transmitted subblock, where k is the subscript of the subblock, and use the following formula to restore the subblock:

$$I_k'(z) = I_{k-1}'(z) + FD_k(z). \quad (9)$$

3. Standardized Judgment Method of Shooting Training Action

The standardized judgment method of shooting training action based on digital video technology is mainly to collect the basketball shooting training image by sampling and characteristic analysis of the pixels in the position space of basketball shooting training action. Using the time difference method, the basketball shooting training target is found and extracted in real time in the digital video sequence. Based on digital video technology, the initial background estimation image is introduced, and the update rate is introduced to update the background estimation image. According to the pixel value sequence of the basketball shooting training image, the pixel model of the basketball shooting training image is defined, and the defined pixel value model parameters are modified. The standardized judgment of shooting training action is realized by judging whether the defined pixel value matches the background parameter model.

3.1. Collecting Basketball Shooting Training Images. Assuming that the Gaussian mixture model labels the spatial position rotation of basketball shooting action, at multiple points in the basketball shooting space, the shape coordinate of the basketball shooting action under the initial deformation is X , the width of the entire characteristic image of the basketball court is W , and the height is H . The three-dimensional spatial feature image I of basketball shooting is divided into several subblocks by using the grid model. The matching coordinate of the central point of the matching point along the gradient direction on the grid model is calculated as Xt , and then, the spherical grid model of basketball in the hands of players is calculated. The triangular partition pheromone of single-frame basketball shooting action is obtained at the j manual calibration point (x_{ij}, y_{ij}) :

$$P(i, j) = \left[X \operatorname{int}\left(\frac{W}{2}\right) - 1 \right] \left[X' \operatorname{int}\left(\frac{H}{2}\right) - 1 \right]. \quad (10)$$

The basketball shooting action sampling image has 8×8 pixels in the grid surface. The sampling point density feature is extracted, and the mean square error between the standardized feature points (x'_{ij}, y'_{ij}) of the shooting action is

$$Err_{ij} = \frac{1}{N} \sum_{i=1}^N \sqrt{(x'_{ij} - x_{ij})^2 + (y'_{ij} - y_{ij})^2}. \quad (11)$$

In the above formula, N is the total number of uniformly distributed grids of the image. Considering all the pixel feature points of n spatial positions, the difference error vector of basketball players in shooting and lifting the ball is obtained as

$$ERR = \frac{1}{n} \sum_{j=0}^{n-1} Err_{ij} = \frac{1}{n} \times \frac{1}{N} \sum_{j=0}^{n-1} \sum_{i=1}^N \sqrt{(x'_{ij} - x_{ij})^2 + (y'_{ij} - y_{ij})^2}. \quad (12)$$

$$D_k(x, y) = \begin{cases} f_k(x, y) - f_{k-1}(x, y), & \text{when } f_k(x, y) - f_{k-1}(x, y) > 0 \\ 0, & \text{when } f_k(x, y) - f_{k-1}(x, y) \leq 0 \end{cases} \quad (13)$$

Negative difference:

$$D_k(x, y) = \begin{cases} |f_k(x, y) - f_{k-1}(x, y)|, & \text{when } f_k(x, y) - f_{k-1}(x, y) > 0 \\ 0, & \text{when } f_k(x, y) - f_{k-1}(x, y) \geq 0 \end{cases} \quad (14)$$

Full difference:

$$D_k(x, y) = |f_k(x, y) - f_{k-1}(x, y)|. \quad (15)$$

(2) Binarize the image D_k after the difference to obtain

$$R_k(x, y) = \begin{cases} 1 & \text{foreground, } D_k(x, y) < \text{thresholding} \\ 0 & \text{background, } D_k(x, y) \leq \text{thresholding} \end{cases}. \quad (16)$$

In the above formula, D_k contains the change of the scene between two consecutive frames of images. This change is composed of many factors. It can be considered that the change of the moving target is obvious. Given a threshold, when the difference of a pixel value in the differential image is greater than a given threshold, the pixel is considered to be a foreground pixel, possibly a point on the target; otherwise, it is considered a background pixel [32–34].

(3) Postprocessing the image R_k to obtain R'_k , where the area of the moving target should be greater than the given threshold. Morphological filtering and noise removal can be used to eliminate noise in small areas.

Thus, the pixel sampling and feature analysis of three main position spaces of basketball shooting training action are realized, and the image acquisition of basketball shooting training is completed.

3.2. Extracting the Goal of Basketball Shooting Training. The time difference method mainly uses the difference of two or several consecutive frames in the digital video sequence to extract the moving target of basketball shooting training [29–31]. The basic process of the time difference method is shown in Figure 2.

(1) Calculate the difference image D_k between the k frame image f_k and the $k-1$ frame image f_{k-1} ; according to the following three methods, the difference image obtained is expressed as follows.

Positive difference:

(4) Judge the postprocessing result R'_k , mark the area larger than the given threshold as the target, and obtain its complete location information. Through the above steps, the basketball shooting training sports goal is extracted.

3.3. Judging the Standardization of Basketball Shooting Training Action

(1) Initial background estimation image: the single Gaussian distribution background model is suitable for single-modal background situations. It establishes a model $\eta(x, \mu_t, \sigma_t^2)$ represented by a single Gaussian distribution for the color of each image point, where t represents time. Let the current color value of the basketball shooting training image point be f_i ; calculate the average brightness μ_0 of each pixel and the variance δ_0^2 of pixel brightness of the basketball shooting training image in the digital video sequence, and take the image B_0 with Gaussian distribution composed of μ_0 and δ_0^2 as the initial background estimation image, which is expressed as

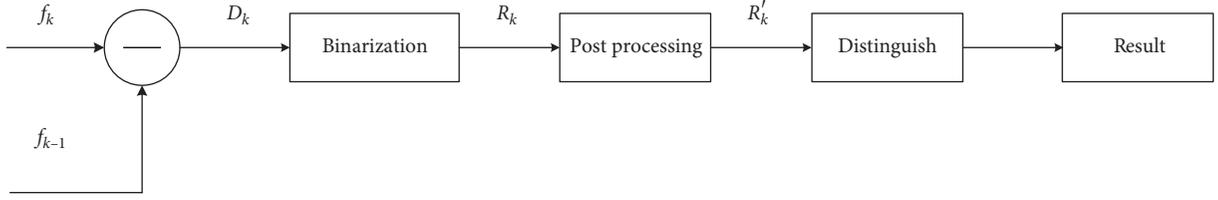


FIGURE 2: Basic process of the time difference method.

$$B_0 = [\mu_0, \delta_0^2]. \quad (17)$$

In formula (17), $\mu_0 = 1/T \sum_{i=0}^{T-1} f_i$, $\delta_0^2 = 1/T \sum_{i=0}^{T-1} (f_i - \mu_0)^2$.

- (2) Update the background estimation image: the update of the single Gaussian distribution background model refers to the update of the Gaussian distribution parameter of the basketball shooting training image. The constant update rate α that represents the update speed is introduced, and the update of the Gaussian distribution parameter of this point can be expressed as

$$B_t = [\mu_t, \delta_t^2]. \quad (18)$$

In the above formula, $\mu_t = (1 - \alpha)\mu_{t-1} + \alpha f_t$, $\delta_t^2 = (1 - \alpha)\delta_{t-1}^2 + \alpha(f_t - \mu_t)^2$.

- (3) Define the pixel model: define the distribution model for each basketball shooting training image pixel. Set the pixel value sequence of the basketball shooting training image as $\{x_{t-k}, x_{t-k+1}, \dots, x_t\}$, and on this basis, define a set of multiple single models:

$$P_t(x, y) = \{p_{i,t} | i = 1, 2, \dots, K\}. \quad (19)$$

In the above formula, $p_{i,t} = [w_{i,t}, m_{i,t}, l_{i,t}]$ is each single model, which consists of three parameters, where $w_{i,t}$ is the weight of this single model, and its size reflects the current reliability of the pixel value represented by this model; $m_{i,t}$ is the mean value of this single model, which reflects the center of each single peak distribution; and $l_{i,t}$ is the width of the unimodal distribution of this single model, and its size reflects the degree of instability of the pixel value, and its role is equivalent to that of the aforementioned single model. K is the number of single models, which reflects the number of peaks in the multiplex distribution of pixel values. Its selection depends on the pixel value distribution and also on the computing power of the system. The usual value is between 3 and 5. In order to keep the model close to the current distribution of pixel values, it is necessary to update the parameters of this model for each defined pixel value [35].

- (4) Correct the pixel value model parameters: the parameter correction steps are as follows:

Step 1: for each new pixel value, first check whether it matches the model. The detection method is

For $i = 1$ to K

$$if |x_{i,t} - m_{i,t}| < \alpha l_{i,t} \text{ then matched} \quad (20)$$

else unmatched.

Step 2: after the detection in step 1, the weight of the single model matching the defined pixel value is corrected as

$$w_{i,t} = w_{i,t-1} + \beta w_{i,t-1}. \quad (21)$$

The parameters of the single model that matches the defined pixel values are corrected as follows:

$$m_{i,t} = \frac{1}{k+1} \sum_{j=0}^k x_{t-j} \quad (22)$$

$$l_{i,t} = \frac{1}{k+1} \sum_{j=0}^k (x_{t-j} - m_{i,t})^2.$$

Step 3: after completing the above correction, it is necessary to normalize the weight of single model in the model as follows:

$$w_{i,t} = \frac{w_{i,t}}{\sum_{j=1}^K w_{j,t}}. \quad (23)$$

- (5) Establishment of background pixel model: the above model is used to define the pixel value of the basketball shooting training image; that is, it is necessary to judge whether the defined pixel value is target pixel or background pixel, so as to realize the standardized judgment of shooting training action. Calculate each single model $w_{i,t}/l_{i,t}$, arrange each model in descending order of the single model, consider the previous model to be a background model, and obtain the model of background pixels [36] expressed as

$$B = \operatorname{argmin}_N \sum_{k=1}^N w_{i,t} > 1 - \alpha w_l. \quad (24)$$

Through the above steps, the pixel value of the basketball shooting training image is defined by using the background parameter model, and the standardized judgment of shooting training action is realized by judging whether the defined pixel value matches the background parameter model.

4. Experimental Analysis

4.1. Experimental Environment and Data. In order to verify the effectiveness of the method for determining the standardization of shooting training movements based on digital video technology, the experiments were conducted on a computer with Intel Core i7-6800K, 3.4 GHz, Nvidia GeForce GTX1080 (8G) graphics card and 24G memory. The operating system is Win10, and the software platform is Anaconda3 and Visual Studio 2015. The resolution of basketball shooting action visual image sampling is 320×240 . A group of basketball shooting action visual image simulation data express a basketball shooting action. There are 100 test sample image sets in each shooting action mode and a total of 1024×1000 test sets in basketball shooting action visual image database. The reference background neighborhood is 5×5 image blocks, that is, 20×20 pixels, and the model update parameters are $\beta_0 = 0.95, \beta_1 = 0.99, \beta_2 = 0.90$, and the threshold t_{IBSCI} is determined after many experiments. In order to ensure the absolute fairness of the experimental results, the ball selection processing in the whole experimental process is completed by the artificial intelligence robot, and the relevant participants only serve as the detection and verification personnel to supervise and investigate the ball selection operation of the robot. According to the above parameters, SolidWorks is used to establish a simplified visual analysis model of basketball shooting action, import the analysis data into ADAMS software for image processing and analysis, and make standardized judgment on basketball shooting action. The standardized action mode of basketball shooting is shown in Figure 3.

Save the basketball shooting standardized action data shown in Figure 3 as .TXT text data, load it into the image data processing software, conduct computer vision analysis, guide the actual shooting action, collect the basketball shooting training image, and obtain the original basketball shooting information. The collection results are shown in Figure 4.

Figure 4 shows the original basketball shooting information collection results. In order to realize the standardized judgment of shooting training action, it is necessary to extract the training target from the collected original basketball shooting information. The proposed standardized judgment method of shooting training action based on digital video technology is used to extract the collected original basketball shooting training target and judge the standardization of shooting training action. The results are shown in Figure 5.

It can be seen from the analysis of Figure 5 that the standardized judgment method of shooting training action based on digital video technology can effectively realize the extraction and detection of moving targets in basketball shooting training, correct shooting errors in real time, and effectively guide basketball shooting training.

In order to evaluate and compare the proposed standardized judgment method of shooting training action, the accuracy rate and recognition rate are used as evaluation indexes, and the calculation formulas of accuracy rate and recognition rate are as follows:

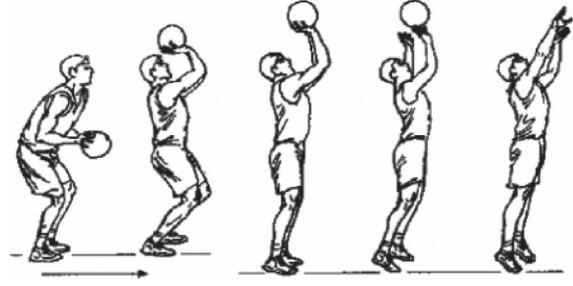


FIGURE 3: Standardized action mode of basketball shooting.

$$P = \frac{TP}{TP + FP} \quad (25)$$

$$R = \frac{TP}{TP + FN}$$

where TP is the number of foreground pixels that are correctly detected, FP is the number of pixels whose background is misjudged as foreground, and FN is the number of pixels whose foreground is misjudged as background.

4.2. Comparison of Standardized Judgment Accuracy of Shooting Training Action. In order to further verify the judgment accuracy of the proposed method, the accuracy is taken as the evaluation index. The higher the accuracy, the higher the judgment accuracy. By comparing the method of reference [8] and the method of reference [9], the standardized judgment accuracy of shooting training action of different methods is obtained, and the comparison results are shown in Figure 6.

According to the analysis of Figure 6, when the number of experiments is 30, the average standardized judgment accuracy of the shooting training action of the method of reference [8] is 84.6%, the average standardized judgment accuracy of the shooting training action of the method of reference [9] is 70.3%, and the average standardized judgment accuracy of shooting training action of the proposed method is as high as 95.2%. Therefore, compared with the method of reference [8] and the method of reference [9], the proposed method has a higher accuracy of the standardized judgment of shooting training action and can effectively improve the accuracy of standardized judgment of shooting training action.

4.3. Comparison of Standardized Judgment and Stability of Shooting Training Action. Further verify the judgment stability of the proposed method, and take the recognition rate as the evaluation index. The higher the recognition rate, the better the judgment stability of the method. By comparing the method of reference [8], the method of reference [9], and the proposed methods, we get the comparison results of the standardized judgment stability of shooting training actions of different methods, as shown in Figure 7.

According to the analysis of Figure 7, when the number of experiments is 30, the average standard judgment

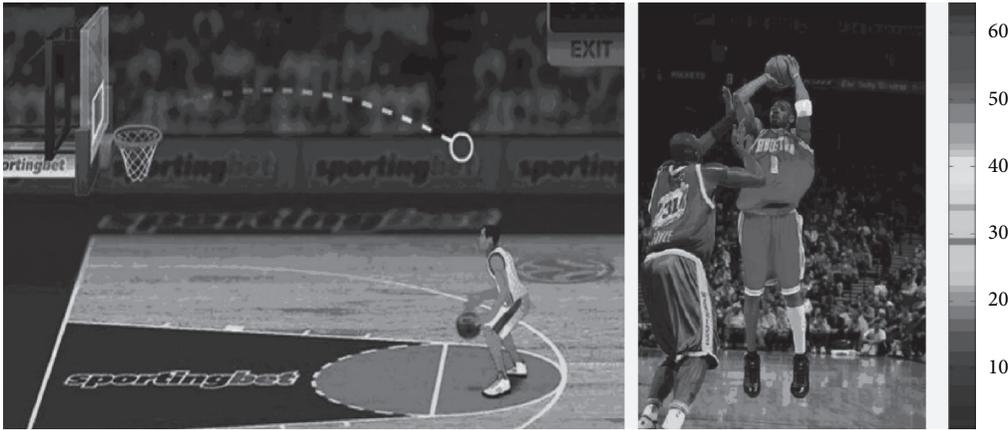


FIGURE 4: Original basketball shooting information collection results.

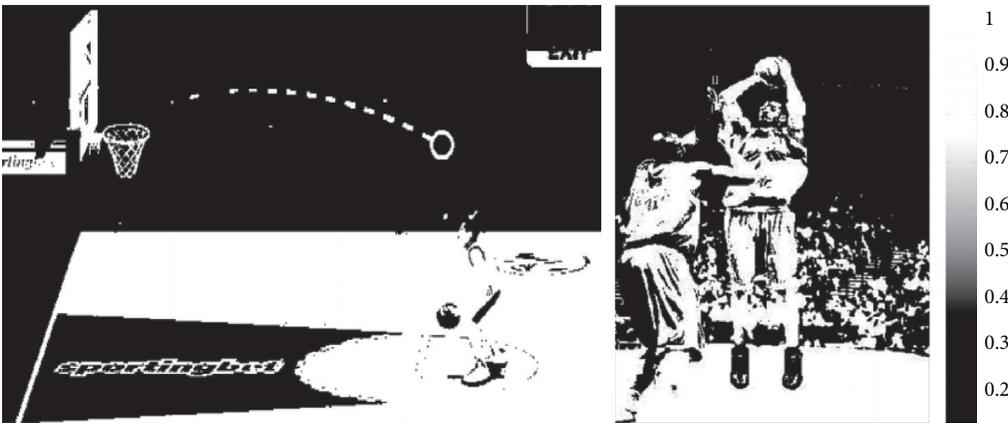


FIGURE 5: Extract the target results of basketball shooting training.

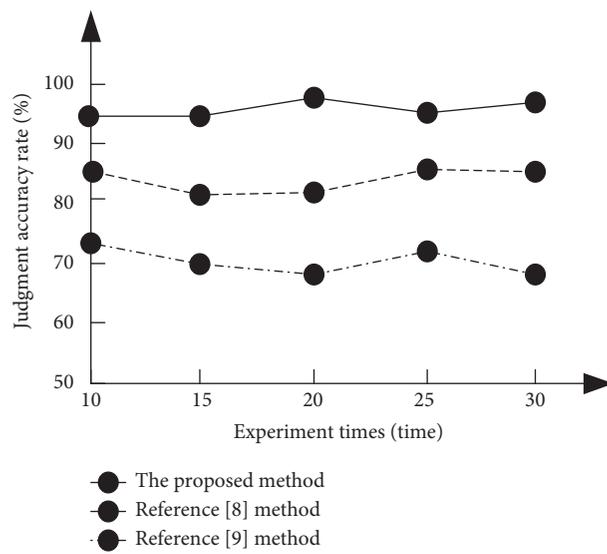


FIGURE 6: Comparison results of the standardized judgment accuracy of shooting training actions of different methods.

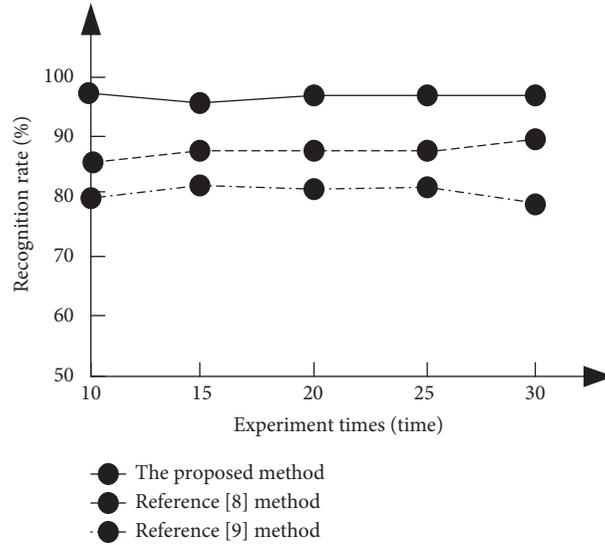


FIGURE 7: Comparison results of the standardized judgment stability of shooting training actions of different methods.

TABLE 1: Comparison results of the standardized judgment time of the shooting training action of different methods.

Number of experiments	The proposed method	The method of reference [8]	The method of reference [9]
10	4.36	7.98	9.76
15	7.12	11.8	13.8
20	9.23	15.7	17.6
25	12.2	19.5	23.9
30	15.3	23.9	26.5

recognition rate of shooting training action of the method of reference [8] is 88.4%, the average standard judgment recognition rate of shooting training action of the method of reference [9] is 80.2%, and the average standard judgment recognition rate of the shooting training action of the proposed method is as high as 96%. It can be seen that compared with the method of reference [8] and the method of reference [9], the proposed method has better stability in judging the standardization of shooting training action.

4.4. Comparison of Standardized Judgment Time of Shooting Training Action. On this basis, the judgment time of the proposed method is verified, and the method of reference [8], the method of reference [9], and the proposed method are compared. The standardized judgment time of the shooting training action of different methods is compared, and the comparison results are shown in Table 1.

According to the data in Table 1, with the increase in the number of experiments, the standardized judgment time of shooting training actions of different methods increases. When the number of experiments reaches 30, the standardized judgment time of the shooting training action of the method of reference [8] is 23.9 s, the standardized judgment time of the shooting training action of the method of reference [9] is 26.5 s, whereas the standardized judgment time of the shooting training action of the proposed method is only 15.3 s. Therefore, compared with the method of

reference [8] and the method of reference [9], the standardized judgment time of shooting training action of the proposed method is shorter.

5. Conclusion

- (1) The proposed standardized judgment method of shooting training action based on digital video technology gives full play to the advantages of digital video technology
- (2) The standardized judgment of shooting training action is high, which can effectively shorten the judgment time and has good judgment stability
- (3) Correct shooting mistakes in real time, and effectively guide basketball shooting training

However, in the process of standardized judgment of shooting training action, dimension reduction is not considered to deal with the characteristics of shooting training action, so as to reduce the amount of calculation. Therefore, in the next research, the dimension of shooting training action characteristics is reduced to further reduce the judgment 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.

Acknowledgments

This work was supported by key topics in 2020 of the 13th Five-Year Plan of Educational Science in Gansu Province: Research on the implementation path of physical education to improve college students' physical quality.

References

- [1] Z. Pan and C. Li, "Robust basketball sports recognition by leveraging motion block estimation," *Signal Processing: Image Communication*, vol. 83, no. 10, Article ID 115784, 2020.
- [2] H. Dong, "Evaluation of the value of basketball players based on wireless network and improved Bayesian algorithm," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, no. 1, pp. 1–11, 2020.
- [3] J. Li and Y. Yang, "Study on the influential factors of ordinary college elite men's basketball teams tactic execution," *Bulletin of Sport Science & Technology*, vol. 27, no. 2, pp. 125–127, 2019.
- [4] Y. Chen, Y. Qiu, and W. Ren, "A normalized score-based weighted PageRank algorithm on ranking prediction of basketball games," *Modern Physics Letters B*, vol. 35, no. 18, Article ID 2150302, 2021.
- [5] D. Castillo, J. Raya-González, A. T. Scanlan, S. Sánchez-Díaz, and Y. Javier, "The influence of physical fitness attributes on external demands during simulated basketball matches in youth players according to age category," *Physiology & Behavior*, vol. 233, no. 1, Article ID 113354, 2021.
- [6] J. Xu and C. Yi, "The scoring mechanism of players after game based on cluster regression analysis model," *Mathematical Problems in Engineering*, vol. 2021, no. 3, pp. 1–7, 2021.
- [7] J. Vera, R. Molina, D. Cárdenas, B. Redondo, and R. Jiménez, "Basketball free-throws performance depends on the integrity of binocular vision," *European Journal of Sport Science*, vol. 20, no. 3, pp. 407–414, 2020.
- [8] V. Nair, P. Hemeren, A. Vignolo, N. Noceti, and G. Sandini, "Action similarity judgment based on kinematic primitives," *Robotics*, vol. 30, no. 8, Article ID 13176, 2020.
- [9] H. Lv and X. Dong, "Target detection algorithm for basketball moving images based on improved Gaussian mixture model," *Microprocessors and Microsystems*, vol. 83, no. 6, Article ID 104010, 2021.
- [10] N. A. Shelke and S. S. Kasana, "Multiple forgeries identification in digital video based on correlation consistency between entropy coded frames," *Multimedia Systems*, vol. 24, no. 7, pp. 1–14, 2021.
- [11] W. E. Bruehs and D. Stout, "Quantifying and ranking quality for acquired recordings on digital video recorders," *Journal of Forensic Sciences*, vol. 65, no. 4, pp. 1155–1168, 2020.
- [12] L. O'Donnell, R. Mander, M. Denton et al., "Portable digital video camera configured for remote image acquisition control and viewing," *Driving ip forward*, vol. 2, no. 20, Article ID 10356304, 2019.
- [13] S. Dinmore, "Beyond lecture capture: creating digital video content for online learning – a case study," *Journal of University Teaching and Learning Practice*, vol. 16, no. 1, pp. 1–10, 2019.
- [14] D. Chi and J. Zhou, "Deep learning-based luma and chroma fractional interpolation in video coding," *IEEE Access*, vol. 7, no. 8, pp. 112535–112543, 2019.
- [15] S. Sowmyayani, V. Murugan, and J. Kavitha, "Fall detection in elderly care system based on group of pictures," *Vietnam Journal of Computer Science*, vol. 8, no. 2, pp. 199–214, 2020.
- [16] A. M. Atto, A. Benoit, and P. Lambert, "Timed-image based deep learning for action recognition in video sequences," *Pattern Recognition*, vol. 104, no. 11, Article ID 107353, 2020.
- [17] M. Wang, J. Lin, J. Zhang, and W. Xie, "Fine-grained region adaptive loop filter for super-block video coding," *IEEE Access*, vol. 8, no. 12, pp. 445–454, 2020.
- [18] A. Nakagawa and K. Kato, "Quantitative understanding of VAE by interpreting ELBO as rate distortion cost of transform coding," *Machine Learning*, vol. 30, no. 7, Article ID 15190, 2020.
- [19] N. Li, Y. Zhang, and C. Kuo, "Explainable machine learning based transform coding for high efficiency intra prediction," *Image and Video Processing*, vol. 21, no. 12, Article ID 11152, 2020.
- [20] S. Milani, E. Polo, and S. Limuti, "A transform coding strategy for dynamic point clouds," *IEEE Transactions on Image Processing*, vol. 29, no. 3, pp. 8213–8225, 2020.
- [21] S. Kansal and R. K. Tripathi, "Adaptive geometric filtering based on average brightness of the image and discrete cosine transform coefficient adjustment for gray and color image enhancement," *Arabian Journal for Science and Engineering*, vol. 45, no. 3, pp. 1655–1668, 2020.
- [22] L. Sun, S. Liang, P. Chen, and Y. Chen, "Encrypted digital watermarking algorithm for quick response code using discrete cosine transform and singular value decomposition," *Multimedia Tools and Applications*, vol. 80, no. 2, pp. 1–16, 2020.
- [23] K. Ramadan, M. I. Dessouky, and F. El-Samie, "Equalization and blind CFO estimation for performance enhancement of OFDM communication systems using discrete cosine transform," *International Journal of Communication Systems*, vol. 33, no. 3, Article ID e3984, 2020.
- [24] S. Agha, U. A. Gulzari, F. Shaheen, and F. Jan, "A high throughput two-dimensional discrete cosine transform and MPEG4 motion estimation using vector coprocessor," *Journal of Real-Time Image Processing*, vol. 17, no. 5, pp. 1319–1330, 2020.
- [25] Z. Yuan, D. Liu, X. Zhang, and Q. Su, "New image blind watermarking method based on two-dimensional Discrete Cosine Transform," *Optik - International Journal for Light and Electron Optics*, vol. 204, no. 2, Article ID 164152, 2019.
- [26] S. Hovsepyan, I. Olasagasti, and A.-L. Giraud, "Combining predictive coding and neural oscillations enables online syllable recognition in natural speech," *Nature Communications*, vol. 11, no. 1, pp. 3117–3124, 2020.
- [27] L. Annabi, A. Pitti, and M. Quoy, "Bidirectional interaction between visual and motor generative models using Predictive Coding and Active Inference," *Neural Networks*, vol. 143, no. 11, pp. 638–656, 2021.
- [28] T. Hueber, E. Tatulli, L. Girin, and J. L. Schwartz, "Evaluating the potential gain of auditory and audiovisual speech-predictive coding using deep learning," *Neural Computation*, vol. 32, no. 3, pp. 1–30, 2020.
- [29] L. Zhang, T. Zhang, H. S. Shin, and X. Xu, "An efficient underwater acoustical localization method based on time difference and bearing measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, no. 12, Article ID 8501316, 2020.

- [30] B. Yazdanshenasshad and M. S. Safizadeh, "Reducing the additional error caused by the time-difference method in transit-time UFM's," *Science, Measurement & Technology, IET*, vol. 51, no. 6, pp. 1049–1057, 2019.
- [31] B. Zhang, W. Bu, and A. Xiao, "Efficient difference method for time-space fractional diffusion equation with Robin fractional derivative boundary condition," *Numerical Algorithms*, vol. 244, no. 4, pp. 1–24, 2021.
- [32] Y. Q. Chen, Z. L. Sun, and K. Lam, "An effective sub-superpixel-based approach for background subtraction," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 1, pp. 601–609, 2019.
- [33] S. M. Roy and A. Ghosh, "Foreground segmentation using adaptive 3 phase background model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 6, pp. 2287–2296, 2019.
- [34] A. Kushwaha, A. Khare, O. Prakash, and M. Khare, "Dense optical flow based background subtraction technique for object segmentation in moving camera environment," *IET Image Processing*, vol. 14, no. 5, pp. 2532–2538, 2020.
- [35] S. Wang, "Simulation of visual saliency evaluation method for multimedia human-computer interaction interface," *Computer Simulation*, vol. 37, no. 03, pp. 161–164, 2020.
- [36] P. Guo, X. Zhu, H. Zhang, and X. Zhang, "Adaptive background mixture model with spatio-temporal samples," *Optik*, vol. 183, no. 4, pp. 433–440, 2019.