

Simulation of tennis serve behavior based on video image processing and wireless sensor technology

Xiao-zhong Li

guangzhou baiyun university

Ping-lin Huang (✉ quechishao292715@163.com)

guangzhou baiyun university

Research

Keywords: Wireless sensor, Sparse representation, Best hitting point, Support vector machine, Particle swarm algorithm

Posted Date: April 24th, 2020

DOI: <https://doi.org/10.21203/rs.2.23964/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on June 30th, 2020. See the published version at <https://doi.org/10.1186/s13638-020-01746-w>.

Simulation of tennis serve behavior based on video image processing and wireless sensor technology

Xiaozhong Li, Pinglin Huang*

Faculty of Education and Sports, Guangdong Baiyun University, Guangzhou 510450, China

Email: quechishao292715@163.com

Abstract: Video-based human motion analysis is an important research direction in the field of computer vision. It detects moving objects from video sequences, extracts key parts of the human body, and obtains useful information for human movements. Further analysis and identification. In this paper, the joints of the teeing arm are first color-coded. The tennis teeing video is collected by a high-speed camera. The coordinates of the tacking points in each frame are used instead of the knuckles to study the trajectory of the teeing arm. In the process of video processing, after constructing a dictionary for a series of noise maps, the sparse representation idea was used to reconstruct an interference-free service diagram, and a mixture of Gaussian background modeling was used to extract the foreground of the motion. After obtaining the motion foreground, the marker points are extracted through the color features, and binarization operations are performed on the marker points. Next, the outline of the marker points is searched, the outline is surrounded by the minimum circle, and the returned circle center coordinates are used as the joint point coordinates. Taking the trajectory of the shoulder marking point as a research object, a tennis serving model based on an improved support vector machine was established.

Keywords: Wireless sensor; Sparse representation; Best hitting point; Support vector machine; Particle swarm algorithm

1. Introduction

With the rapid development of science and technology, the booming information has become a major feature of today's society. At the same time, modern tennis training technology has been widely disseminated. Computer technology, modern communication technology and artificial intelligence technology have been widely used in tennis training[1]. This has greatly promoted the development of the theory and practice of tennis training and optimization. Tennis training must take full advantage of modern tennis training techniques according to its own characteristics to improve the quality of training and improve the quality of athletes. Serving technique is one of the key technologies in tennis. In training, it is also one of the most difficult techniques to master. In the traditional tennis training, the traditional training mode is generally used to train, that is, coaches to explain the demonstration, the athletes to practice on the venue, the coach to conduct individual guidance, found that easy to make mistakes to correct, and then allow athletes to personally practice the exercise, so that Athletes do not have a correct comparison standard, and it is difficult to fundamentally understand their own mistakes[2]. It is difficult to form movement representations in the brain, which is not conducive to athletes' mastery of serving techniques.

Video image analysis technology is a high-end intelligent system that comprehensively uses image data synthesis technology to realize the real-time analysis of tennis tactics and tactics and simultaneously outputs the analysis results. The system can analyze real-time images of tennis match scenes or training scenes, and has many functions such as statistical techniques and tactics data, linked motion data and images, and playback of motion processes. It is a combination of computer image analysis technology and competitive sports. Video image analysis technology, as the most remarkable achievement in the training of tennis in the century and the crystallization of human wisdom, has become the main component of tennis training science and modern tennis training techniques[3]. It has a very positive significance for breaking through the traditional tennis training concept. Tennis technical training requires athletes to have a lot of observations, imitations, feedbacks, and corrections in the process of forming athletic skills. Sensory information other than proprioception is needed, especially audiovisual information. Traditional training is difficult to implement these processes, and video image analysis technology fully demonstrates technical advantages in demonstrating the breadth, integrity, vividness, and detail of practical content, and the key points and difficulties in certain training materials. The details, which are clearly presented to the athletes in words, voices, images, animations, etc., can provide effective help to the coaches. For athletes, their lively displays are more attractive.

2. Related work

Science and technology are playing an indispensable role in today's society and their importance is even more prominent in sports. In tennis day-to-day teaching and training, the effects of coaches'

visually visible training mode on athletes' competitive level have been insignificant since ancient times, and the dominant position of sports science and technology in sports training has been increasingly revealed. Therefore, how to introduce sports science and technology into sports training and continuously improve the effectiveness and scientific nature of sports training has become an important subject worth studying at present. On this basis, this paper proposes the application of video technology in the analysis of tennis serve speed and success rate. As we all know, the human eye is an important external manifestation of the visual system. It is also an important organ for people to obtain information from the outside world. Through the human visual system, it can obtain external, direct information and leave the most intuitive impression. Usually, this activity needs to rely on the static image or the dynamic video media to achieve[4].

For a long time, the path of scientific training in China's competitive training has become increasingly difficult, and there has been a clear gap compared with other countries with developed science and technology. At present, some authoritative experts and coaches at home and abroad have come to a conclusion after long-term research and practice. That is: In the process of athletes completing their actions, they record the athlete's movements with the camera and obtain the athlete's action videos through computer software. Perform processing to make the whole action more detailed. The image or data information represented by the video processing result is the most authentic portrayal of the athlete in the training process[5]. There is no other interference factor, and the technical movement of the athlete's technical movement can be directly assessed and the technical index can be evaluated to find the physical characteristics of the athlete. And expertise, and found that its training in the existing deficiencies and shortcomings, avoid weaknesses, in order to achieve the purpose of improving sports technology and inspire athletes to learn interest. At the same time, the introduction of technical videos in sports programs is not only an analysis of athlete skills, but also provides a basis for controversial penalties in competitions[6].

Video analysis systems have been introduced in tennis games: Hawkeye technology and speedometers. Among them, the Hawkeye technology records and plays back the content of the disputed ball between the player and the referee in the game, and provides evidence for the punishment. The speedometer is also calculated by the distance the ball is recorded in the unit of time. The speed of the ball, these are all accomplished using video image analysis techniques[7]. Video analysis technology is an extremely advantageous teaching tool in the process of tennis teaching. It can be applied to the acquisition of skills, the engraving of technical movements, the visualization of everything, the prevention of injuries and the training of coaches. In fact, when they see their technical moves or performances on the video, the results are very different. Obviously there is an error between ideals and reality. Video technology is an excellent helper to correct this illusion of motion[8]. For athletes, it is very important to understand what is a correct and reasonable technical movement. The video analysis technology can assist the players in the implementation of new technical movements. This study uses the sports training video analysis system in combination with the teaching of tennis special classes in China University of Mining and Technology, and uses the system's convenient and timely feedback function, video contrast playback, background similar video overlay, motion picture decomposition and other functions to serve tennis special class students. Prepare gestures, turn shoulders, bend knees, accelerate rackets, hit the ball, swing with the ball, and close the action to analyze and count the technical movements. Try to grasp the technical movements of the special class as a whole and perform intensive training to correct mistakes. The technical movements provide a reference[9].

Since tennis is really at rest only when it is serving, serving is the only action that can be fully controlled by the player. At the same time serving is also considered to be the most difficult one. It requires a series of very complex actions to be performed at the right time. The rotation of the upper arm before the impact of the racket and ball has a contribution of up to 54% to the linear swing speed at the time of impact. The athlete's exercise data in training and competition can be used to help understand the physical condition and competition skills in the competition. Therefore, these data are very important for coaches and sports workers. Video shooting was used to study high-speed serve. In the study, two synchronized high-speed cameras were used to collect data at 200 Hz. During the entire serving process, the two cameras take pictures from the front and side of the athlete respectively, and the shooting range covers the entire range of the athlete's activities. In this study, the inertial gyro method and the marker-based virtual gyro method (abbreviated as MBVG) were used for the monitoring of the upper arm rotation when serving a tennis ball, and the monitoring results were compared with the results obtained by the video shooting method. The MBVG method can also be used to confirm the peak part of the gyroscopic data. It does not need to give accurate results, only rough changes are required. can.

Girard O. et al. studied the effect of knee motion on tennis ball-shot serve by limiting the

extension of the knee during the serve. During the experiment, thirty athletes were divided into three groups: Beginner, Intermediate, and Elite athletes. A splint was placed on the thigh to extend the knee at an angle of 10° (0° indicates full knee extension), and each person performed 30 flat serves, with normal (without knee restraint) and 15 knees restricted. The experimental results show that knee exercise is an important factor affecting the efficiency of serve, and has nothing to do with the athlete's level.

3. Methods

3.1 Background Difference Method

In this study, some young elite tennis players (secondary athletes and above) were selected during the study. Actual combat data served during training and competition was used as research object. A 20-team male student from a sports college tennis special class was selected as the research object, and the students were randomly divided into the experimental group and the control group.

Table 1. Comparison of the average physical condition of experimental group students, control group students and high-level athletes

	Average height (cm)	Average age (years)	Average arm length (cm)	Average weight (kg)
Test group	177.2	21.4	73.1	65.6
Control group	176.8	21.9	72.6	64.7
Excellent athlete	175.1	15.5	72.2	64.4

As can be seen from Table 1, the average height of students in the experimental group was 177.2cm; the average age was 21.4 years; the average arm length was 73.16cm; the average weight was 65.63kg. The average age of the control group was 21.9 years; the average height was 176.8cm; the average body weight was 64.7kg; the average arm length was 72.6cm. The average age of excellent tennis players is 15.5 years old; the average height is 175.1cm; the average weight is 64.4kg; the average arm length is 73.2cm. There was no significant difference in the average height, average arm length, and average weight of experimental group students, control group students, and high-level athletes[10,11].

Background difference method is a relatively common method for detecting moving objects. For example, most video surveillance and intelligent traffic systems are based on background difference techniques. Because there is a difference between the current frame and the background model, the part where the current frame and the background frame are similar is taken as the background, and the part with the larger difference is defined as the foreground[2,13]. The basic principle is: extract the static background from the video sequence, and then use the difference between the current frame and the background to get the motion foreground. In a static scene, the background model may be captured in advance without foreground moving objects or noise[14,15]. Differentiating the current image frame from the background reference model, determining the moving target area by statistically changing the information in the histogram, or determining the change in the grayscale feature, finally performing the threshold determination and calculation of the difference result image. Therefore, we can know the size, position, shape and other relevant information of the sports foreground target. Figure 1 shows the background difference method schematic.

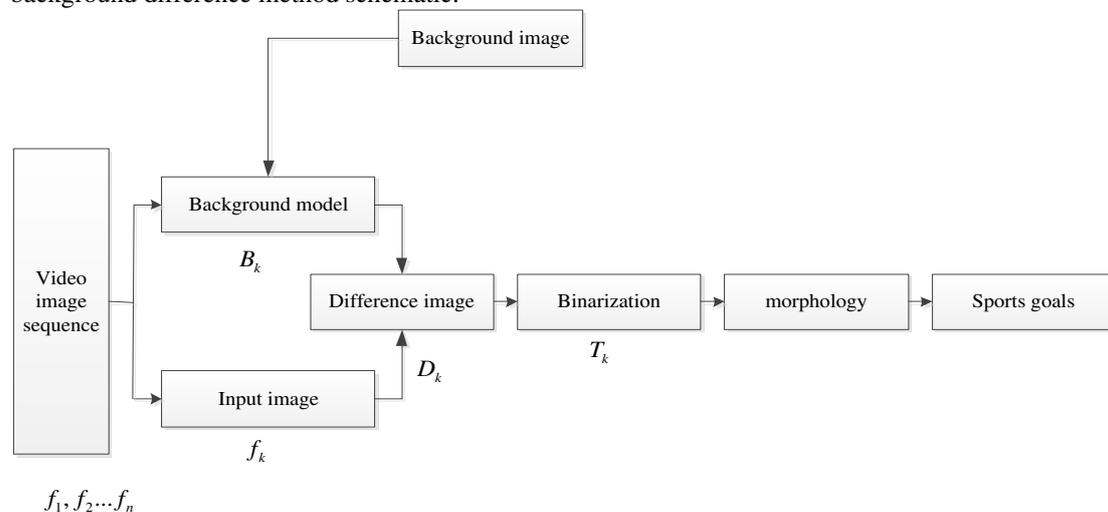


Figure. 1 Background Difference Method Schematic

The background difference method can effectively extract the target of the motion. Given the background reference model, the background difference method is an efficient detection method for moving targets. The most common way to initialize the background model is to extract a certain frame image directly from the image sequence, or to calculate the average value of the multi-frame image.

$$B(x, y, t) = \frac{1}{N} \sum_{K=t-N}^{t-1} I(x, y, k) \quad (1)$$

In which, $B(x,y,t)$ denotes the background pixel value at the position of (x,y) at time t , $I(x,y,k)$ denotes the image information of the k -th frame, and the background is taken from the previous N frame images. The advantage of this method is that it is simple and can extract complete image information features, but is more susceptible to occlusion, shadows, and changes in light. Markov random field can solve the problem of occlusion, but for obvious occlusion, the effect is still relatively poor. In order to reduce the impact of dynamic scene changes on the extraction of moving objects, the simplest and most effective time averaged image can be used to create a background model[16,17]. For example, Haritaoglu uses the minimum and maximum gray values and the largest time difference value for image sequence scenes. All the pixels in the statistics are used to create a background model that can adapt to changes in climate and light[18,19].

The background difference algorithm process is to store the background image first, and then use the difference image between the current image and the background image to perform motion detection on the target. $f_k(x, y)$ is the current frame, $B_k(x, y)$ is the background frame of the current frame, and $D_k(x, y)$ is the result of the difference between the current frame and the background frame. The principle is as shown in the following formula.

$$D_k(x, y) = |f_k(x, y) - B_k(x, y)| \quad (2)$$

$$T_k(x, y) = \begin{cases} 1, & D_k(x, y) \geq T \\ 0, & D_k(x, y) < T \end{cases} \quad (3)$$

When $T_k(x, y)$ is 0, it represents the background in the image; when $T_k(x, y)$ is 1, it represents the moving target area in the image. T is the threshold value. When the difference image point value is greater than the set T value, we think this point is a point on the moving target, otherwise it is considered as the image background point.

Since the target detection needs to perform differential processing on the image to be detected in each frame of the video and the background image model, the modeling method of the background model is very important for the accuracy of this method, and the accuracy of the model directly affects the target motion detection result[20]. The background difference method usually requires that the background model does not produce drastic changes. However, the background model is not absolutely immutable. Sometimes it is necessary to update the background model in time to ensure the correctness of the moving target detection. For example, video noise, changes in illumination, and the transition between objects' motions will require the background model to be updated in time.

3.2 Optical flow method

Gibson first proposed the concept of optical flow in 1950. Optical flow refers to the apparent motion of the image brightness pattern in an image sequence. A moving object will leave behind a series of constantly changing images on the retina of the naked eye, so objects in motion will be discovered by human eyes. The basic idea of the optical flow detection method is that according to the image information of the current frame and its subsequent frames, each pixel in the image is assigned a velocity vector to establish a two-dimensional motion field. When there are moving objects in the image, the optical flow vectors generated by the pixel points of the target area and the neighboring background in the image must be different. The target detection can be achieved by determining the optical flow vector of the pixel in the image.

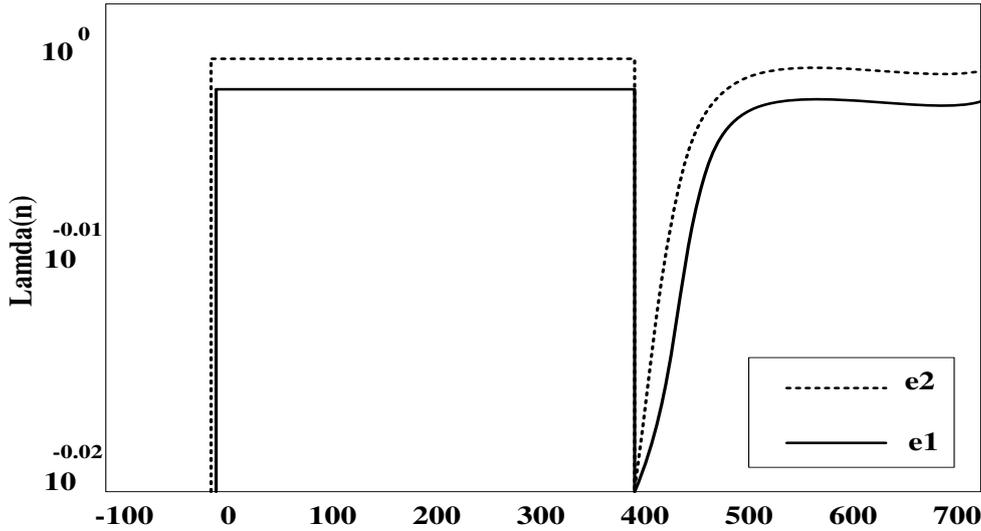


Figure. 2 Stable performance of the improved algorithm before and after the signal

Since a two-dimensional image is a projection of a three-dimensional object motion on a camera, we can use a two-dimensional image sequence to record three-dimensional motion information of an object in a real space. When there is a relative movement between the camera and the object, a corresponding change will occur. From the change between the images, the mutual motion between the object and the camera can be known.

In Figure 2, Z is the object distance between the center of the camera lens and the moving object, f is the focal length, r_i is the direct distance between the image point and the center of the lens, and r_o is the direct distance between the object point and the center of the camera lens. The basic equation of the optical flow method is based on the assumption that the image gray is constant, that is, the pixel gray value of the object in the same position remains unchanged in two adjacent frames of the video sequence[13]. In 1981, Horn and Schunck derived the basic equations of optical flow based on this assumption. The gray value of the pixel (x,y) at time t is $I(x,y,t)$. At time $t + \delta t$, the pixel moves to $(x + \delta x, y + \delta y)$ and the pixel gray value is $I(x + \delta x, y + \delta y)$. According to the previous assumptions:

$$I(x + \delta x, y + \delta y, t + \delta t) = I(x, y, t) \quad (4)$$

Formula (4) according to the Taylor formula expansion, in δt_0 limit and finishing:

$$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0 \quad (5)$$

Order $u = \frac{dx}{dt}$, $v = \frac{dy}{dt}$, $I_x = \frac{\partial I}{\partial x}$, $I_y = \frac{\partial I}{\partial y}$, $I_t = \frac{\partial I}{\partial t}$, then formula (5) can be turned into:

$$I_x u + I_y v + I_t = 0 \quad (6)$$

Equation (6) is called the basic equation of the optical flow method. Among them, I_x, I_y, I_t can be obtained directly from the image. For the aperture problems caused by u, v and two unknowns, various optical flow calculation methods such as Horn-Schunck algorithm and Lucas-Kanade algorithm are formed by adding various optical flow constraint conditions. .

3.3 Inter-frame difference method

Similar to the background difference method, the inter-frame difference method is also one of the most commonly used algorithms in moving object detection algorithms. The principle of this algorithm is: when the gray level of the image sequence changes slightly, the difference operation is performed using the corresponding pixels of the two or three consecutive frames of the image. If the change of the pixel value of a certain point of the difference image is higher than the threshold, this is considered as The point area is caused by the motion of the target; if the change in the pixel value of a point in the difference image is lower than the threshold, the point area is

considered as the background in the image sequence. Calibrate the motion area of the target in the video and use these calibrations to lock the position of the video target. Using the inter-frame difference method directly or indirectly can remove invalid information between frames of the image sequence data, thereby obtaining the change monitoring target[14].

The two-frame difference method performs differential operations on two successive frames of an image sequence. Then, a binarization threshold decision is made on the difference result image, the static background is eliminated, and the moving target region is selected, thereby marking the moving target. The principle of this method is shown in Figure 3.

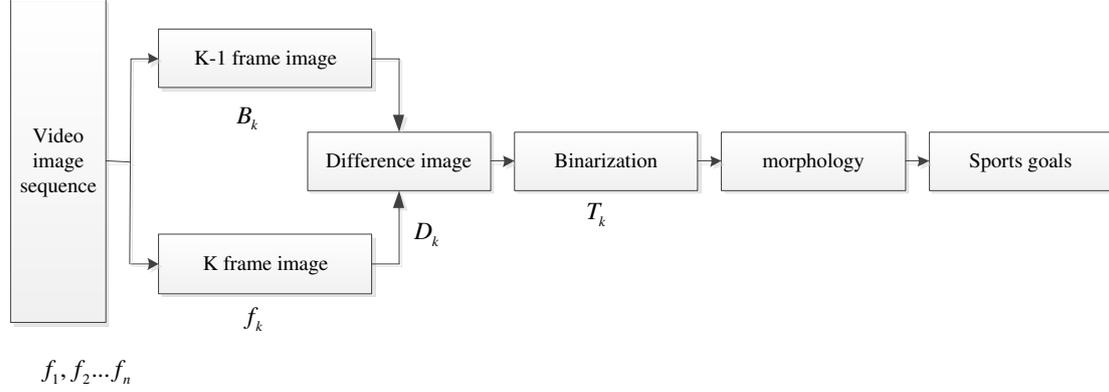


Figure. 3 Schematic diagram of the difference between frames

In order to detect an effective moving target, the two-frame difference method needs to satisfy: the target needs to have a moving speed, the background scene is still while its gray value changes little, other interference noise is small, and the target gray value changes relatively large, etc. condition. Due to the influence of noise, the influence of background brightness, etc., these factors will affect the effect of the two-frame difference method image to varying degrees. The algorithm operation process is as follows:

Set $D_k(x, y)$ as the difference result image, the gray values at points (x, y) in the $k-1$ frame image and the k -th frame image are $f_{k-1}(x, y)$ and $f_k(x, y)$, and the $k-1$ and k -frames are calculated using the following equation (7). Image difference processing, where $D_k(x, y)$ is the resulting image after difference calculation.

$$D_k(x, y) = |f_k(x, y) - f_{k-1}(x, y)| \quad (7)$$

By using the following equation (8), $D_k(x, y)$ is threshold to detect the background and the moving target.

$$T_k(x, y) = \begin{cases} 1, & D_k(x, y) \geq T \\ 0, & D_k(x, y) < T \end{cases} \quad (8)$$

In which, T is the threshold and $T_k(x, y)$ is 1, this point represents the target motion area in the image. The detection formula reflects the accuracy of the target change location depends on the threshold selection during the threshold calculation process.

The two-frame difference method is the same as the background difference method. It achieves a simple program with low complexity and good robustness. In the case of dynamic background, it is more adaptive than other algorithms. Unlike the background difference method, it does not establish a background model, which saves a lot of processing calculations and eliminates the errors generated by the model. However, this algorithm also has some disadvantages. First, when the target motion speed is too fast, the moving target is easily missed. Second, only part of the target-related motion information can usually be detected, and thus the interior of the target will appear hollow, thus the connectivity of the moving target in the image. Influencing; Finally, if a slight gray level change in the background area is misjudged as a change in the target, such interference noise will cause noise points in the detection target.

In summary, the disadvantages of the two-frame difference method include the fact that the target speed is too fast and it is easy to miss the inspection. It will cause the interior of the target to have a void affecting the connectivity, and the background region's micro-motion will increase the noise

interference; the background difference method has disadvantages that are difficult to be in a dynamic scene. Under the circumstances, it is difficult to detect the target; the disadvantage of the optical flow method is high complexity and large amount of calculation, so it is difficult to achieve. The three-frame difference rule is to improve the inter-frame difference method. In the three consecutive frames of images, first use the classic inter-frame difference method to perform motion detection on the first two frames and the last two frames respectively; secondly, two results are obtained. The image is cross-referenced to the common part, which is considered to be the target area of the second frame. The principle is shown in Figure 4.

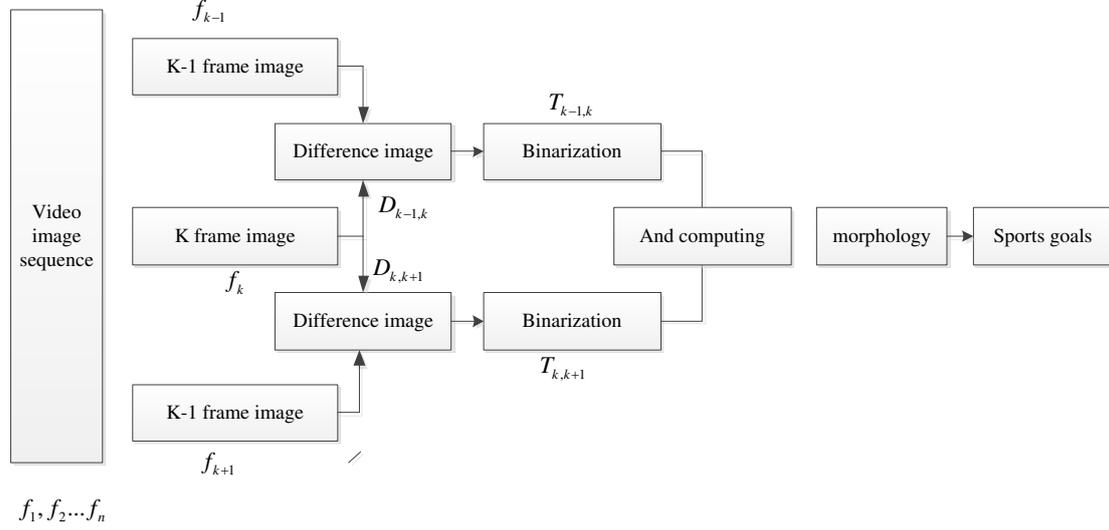


Figure. 4 Three-frame difference method schematic

Suppose three frames of image, $k-1$, k , and $k+1$ at the point (x, y) gray values are $f_{k-1}(x, y)$, $f_k(x, y)$, and $f_{k+1}(x, y)$, according to the principle of Figure 5, the difference calculation, as follows:

$$D_{k-1,k}(x, y) = |f_k(x, y) - f_{k-1}(x, y)| \quad (9)$$

$$D_{k,k+1}(x, y) = |f_{k+1}(x, y) - f_k(x, y)| \quad (10)$$

An appropriate threshold T is selected to threshold the result image to obtain a binarized image.

$$T_{k-1,k}(x, y) = \begin{cases} 1, & D_{k-1,k}(x, y) \geq T \\ 0, & D_{k-1,k}(x, y) < T \end{cases} \quad (11)$$

$$T_{k,k+1}(x, y) = \begin{cases} 1, & D_{k,k+1}(x, y) \geq T \\ 0, & D_{k,k+1}(x, y) < T \end{cases} \quad (12)$$

The above processed images $T_{k-1,k}(x, y)$ and $T_{k,k+1}(x, y)$ are logically intersected with each other to obtain the final target result set $T_k(x, y)$.

$$T_k(x, y) = \begin{cases} 1, & T_{k-1,k}(x, y) \cap T_{k,k+1}(x, y) = 1 \\ 0, & T_{k-1,k}(x, y) \cap T_{k,k+1}(x, y) \neq 1 \end{cases} \quad (13)$$

In the same way, the point at which $T_k(x, y)$ is 1 represents the target area of motion.

3.4 Mixed gaussian background modeling

Mixed Gaussian background modeling is a background description method based on statistical information of pixel samples. It uses statistical information such as the probability density of pixels (such as the expected and standard deviation of each mode, the number of modes, etc.) Long time sample values represent the background. Then, the target pixel is determined by statistical difference method (such as 3δ principle, δ is standard deviation). Complex dynamic backgrounds can also be modeled, but the amount of computation may be large.

In the mixed Gaussian background model, the color information between the pixels is defined as not related to each other, and the processing of each pixel point is also independent of each other. In the video image, the change of the value of each pixel in the image sequence can be seen as a random

process of continuously generating pixel values, that is, a Gaussian distribution network can be used to represent the regularity of the color of each pixel. For the multimodal Gaussian distribution model, first assign different weights to each pixel in the image, then make multiple Gaussian distributions superimposed by different weights to build the model. Each Gaussian distribution and one pixel may generate The rendered color state corresponds. Over time, the weights and distribution parameters of each Gaussian distribution will be continuously updated. When processing a color image, it is assumed that the RGB three color channels of the pixel point are independent of each other and the variance is the same. Mixing Gaussian Background Modeling In order to describe the state of a pixel at a certain moment, K Gaussian models are created for this pixel. The mixed Gaussian distribution probability function is as follows:

$$p(X_t) = \sum_{i=1}^K \omega_{i,t} \times \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (14)$$

$$\eta(X_t, \mu_{i,t}, \Sigma_{i,t}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_{i,t}|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu_{i,t})^T \Sigma_{i,t}^{-1} (X_t - \mu_{i,t})} \quad (15)$$

In which, X_t is the pixel value of a pixel at time t , K is the number of Gaussian models, and $\mu_{i,t}$ and $\omega_{i,t}$ represent the mean and weight of the i -th Gaussian model at time t . $\eta(X_t, \mu_{i,t}, \omega_{i,t})$ represents the probability density function, $\Sigma_{i,t}$ is the covariance matrix of the Gaussian model, where $i = 1, 2, \dots, k$.

The detailed algorithm flow is as follows:

(1) According to equation (16), each pixel value X_t is compared with the current K models until a distribution model matching the new pixel value is found, ie, the expected deviation from the model is within 2.5δ .

$$|X_t - \mu_{i,t-1}| \leq 2.5\sigma_{i,t-1} \quad (16)$$

(2) If the matched pattern matches the background requirement, the pixel belongs to the background, otherwise it will belong to the foreground.

(3) The weight of each mode is updated according to equation (17), where α is the learning rate, and for pattern $M_{k,t} = 1$ that matches successfully, otherwise $M_{k,t} = 0$, then the weights of the modes are normalized.

$$w_{k,t} = (1 - \alpha) * w_{k,t-1} + \alpha * M_{k,t} \quad (17)$$

(4) Mean and standard deviations that do not match the pattern of success remain unchanged, and the parameters for the matching pattern are updated according to the following formula.

$$\rho = \alpha * \eta(X_t | \mu_k, \sigma_k) \quad (18)$$

$$\mu_t = (1 - \rho) * \mu_{t-1} + \rho * X_t \quad (19)$$

$$\sigma_t^2 = (1 - \rho) * \sigma_{t-1}^2 + \rho * (X_t - \mu_t)^T (X_t - \mu_t) \quad (20)$$

(5) If no pattern matching is successful in (1), the pattern with the least weight will be updated. The mean of the pattern is the current pixel value, the standard deviation is replaced by the initial large value, and the weight is updated to a smaller value.

(6) The patterns are arranged in descending order, with the pattern of heavy weight and small standard deviation placed first.

(7) Select the first B patterns as the background, B satisfy the formula (21), and T represents the proportion of the background.

$$B = \arg(\min(\sum_{k=1}^b w_k > T)) \quad (21)$$

In which, T represents the proportion of the background, and by setting the value of T, the best background pixel can be selected.

Among these methods, the background difference method is suitable for situations where light and shadows do not change significantly, and the background modeling and updating process is more complicated; the optical flow method is extremely sensitive to noise and has a large amount of calculation, but the camera is in the process of shooting. The motion and sloshing do not affect the

detection results, and the robustness is strong. The inter-frame difference method has a simple operation and strong anti-interference. However, when the foreground part of the scene stops moving, it cannot detect the complete foreground, and the sampling frequency and prospect The speed of the target will also affect the detection effect. Mixed Gaussian background modeling has strong robustness to the dynamic changes of the scene, but is sensitive to changes in illumination.

4. Experiment

Variability is an inherent feature of human motion. It cannot be accurately repeated twice by human eyes. This chapter marks the joints of the delivery arm with marked points and obtains the tennis serve video through a high-speed camera. In the process of video processing, the video image sequence is first denoised, and then the foreground of the movement is acquired through the target detection algorithm. Perform coordinate extraction and analyze the trajectory of the marker points. With the analysis of the movement trajectory and serving data of the marked points, the best hitting point of the tennis is predicted. The use of video image processing technology to improve the quality of training will help athletes better master motor skills and improve training efficiency.

The collected tennis ball leveling serve video was analyzed by video image technology. When the noise interference is filtered out, median filtering, wavelet denoising and sparse denoising are used for comparative analysis. Then, the mixture of Gaussian background modeling is used to extract the foreground of motion, and the foreground is further analyzed and processed to obtain the coordinate information of the three markers. Three stages of the tennis ball serve are selected: throwing the ball, backward swinging and hitting the ball. Through data analysis, a range of best hitting points was obtained, and hitting the ball within this range can improve the accuracy of serving.

The various joints of athletes in the actual movement will be effectively captured, so that to have the corresponding joint movement process. Figure 6 is example of utilizing this method to position the joint node of athletes. During the intense movement, each joint of the player will be effectively captured, without leaking any point.

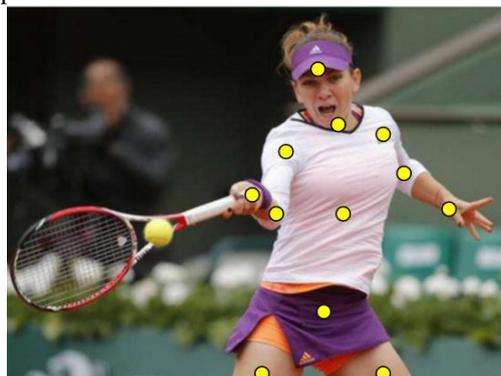


Figure. 5 Athlete Joint Point Capture

It can be seen that net ball and competition have become world competitions and sports events. In China, the main tennis events are the Chinese Open Tennis, Shanghai Masters, and so on. At present, competitive tennis has been treated as an important project in the development of competitive sports in China. Therefore, it is an important way to improve the competitive level of tennis in China by strengthening the scientific research on tennis competition, probing into the winning factors and rules of tennis competition, and providing scientific and technological support for competitive tennis. It is a subject worthy of further study as shown in Figure.6.

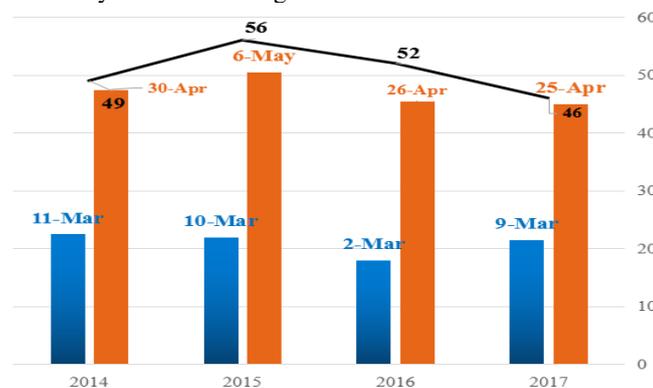


Figure. 6 Results of tennis singles

The corresponding coordinate system is established with the left elbow joint point as the origin, and then the tracking of the joint point can be realized. In the actual application process, the coordinate transformation of the joint point can be obtained by combining the coordinate transformation of the joint point. On this basis, we can get the corresponding joint point coordinates, even the effective analysis of the whole joints, and thus draw the corresponding conclusions.

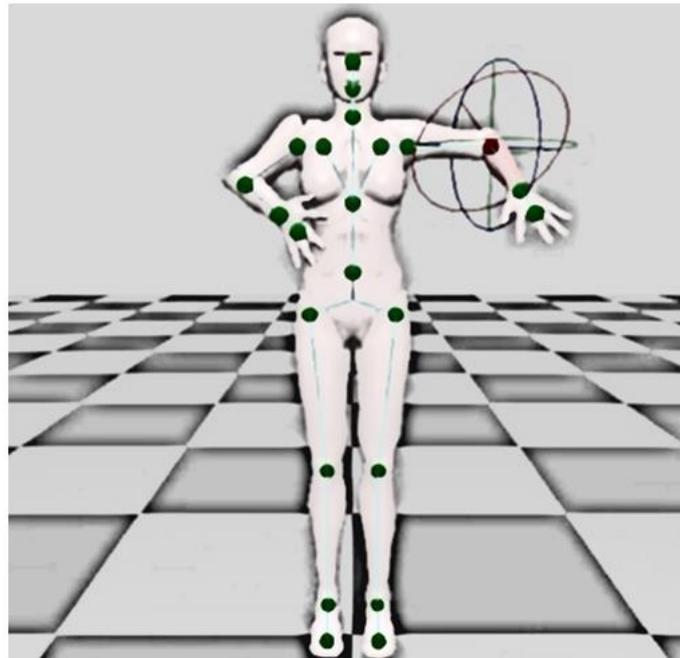


Figure. 7 Example of Dynamic Tracking of Human Joint Node Model

The main function of the tennis player information capture system is to collect the information of the athletes during the actual movement and carry out the corresponding data analysis. It can be observed from the figure 8 that the system will transmit the corresponding information of the athletes to the computer information processing center module in the actual analysis. Then the information can be integrated processed. During the actual analysis, it is necessary to combine the training process of athletes. That's the reason for the structure has facilitated the terminal display module for athletes to have simulating exercises, which could capture the movement process through the high-definition camera equipment and draw on the corresponding conclusions. In this regard, then cycled information collection system of the process of athletes can be achieved, promoting the athletes' skills. The simulation results are shown in Figure 8.

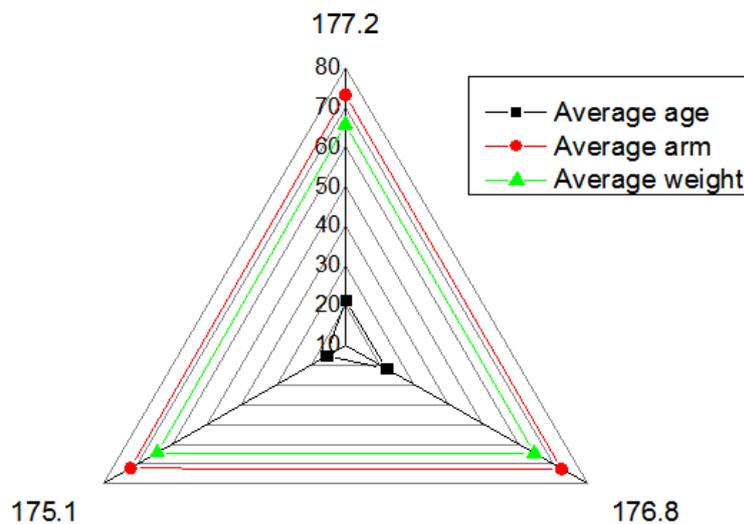


Figure. 8 Simulation results

5. Results and Discussion

Sports analysis based on video image processing is a research hotspot and difficult point in the field of computer vision. It detects moving objects from video sequences, extracts key parts of the human body, and obtains useful information for human movements to achieve human movements, postures, etc. The essay made predictions on the best hitting point for tennis. We color-coded the joints of the serving arm, collected the tennis ball video by a high-speed camera, and used the coordinate of the point in each frame instead of the joint point coordinate to study the trajectory of the arm when the ball was served. During the video processing, in the process, the interference of the environment with the mark color should be avoided. Blending Gaussian background modeling for motion foreground extraction. After obtaining the motion foreground, the marker points are extracted by the color features, then the binarization operation is performed on the marker points, then the contour search is performed, the outline is surrounded by the minimum circle, and the returned center coordinates are the joint point coordinates. Through the analysis of the trajectory, we find that the arms have periodical mathematical characteristics in the process of serving and find the inherent characteristics of the best ball hitting point in the trajectory analysis of each set of serving movements. For the prediction of the best shot point, the efficiency of serving can be improved and the purpose of auxiliary training can be achieved.

References

- [1] Feng, Ren, Jiang. Generic framework for content-based stereo image/video retrieval (Image and vision processing and display technology). *Electronics Letters*, 2011, 47(2):97-98.
- [2] Hockaday S. Evaluation of image processing technology for applications in highway operations: final report. *Physical Review B*, 1991, 14(3):1323-1326.
- [3] Nachtergaele L, Catthoor F, Balasa F, et al. Optimization of memory organisation and partitioning for decreased size and power in Vvideo and image processing systems. *Leukemia Research*, 1995, 30(11):1377-1384.
- [4] Aizawa K, Sakaue K, Suenaga Y. Image processing technologies. Algorithms, sensors, and applications. *Embo Journal*, 2004, 18(12):3419-3430.
- [5] Camana P. Image processing: Video-bandwidth compression: A study in tradeoffs: In this fledgling field, designers are matching criteria to applications and developing `first-cut` hardware. *IEEE Spectrum*, 1979, 16(6):24-29.
- [6] Zhenhua Huang, Xin Xu, Juan Ni, Honghao Zhu, Cheng Wang. Multimodal representation learning for recommendation in Internet of Things. *IEEE Internet of Things Journal*, 2019, 6(6): 10675-10685.
- [7] Borie B, Gautier Y. An autonomous automatic incident detection equipment using video image processing: EVA. *Environmental Science & Technology*, 1995, 29(4):74-76.
- [8] Secker A, Taubman D. Lifting-based invertible motion adaptive transform (LIMAT) framework for highly scalable video compression. *IEEE Transactions on Image Processing*, 2004, 12(12):1530-1542.
- [9] Hanna G, Cuschieri A. Image display technology and image processing. *World Journal of Surgery*, 2001, 25(11):1419-1427.
- [10] Li Y N, Lu Z M, Niu X M. Fast video shot boundary detection framework employing pre-processing techniques. *Iet Image Processing*, 2009, 3(3):121-134.
- [11] Mudde R F, Schulte H B M, Akker H E A V D. Analysis of a bubbling 2-D gas-fluidized bed using image processing. *Powder Technology*, 1994, 81(2):149-159.
- [12] Yokoo A, Taniguchi H. Application of video image processing to detect volcanic pressure waves: A case study on archived images of Aso Volcano, Japan. *Geophysical Research Letters*, 2004, 31(23):275-295.
- [13] Nishikata K, Kimura Y, Takai Y, et al. Real-time lock-in imaging by a newly developed high-speed image-processing charge coupled device video camera. *Review of Scientific Instruments*, 2003, 74(3):1393-1396.
- [14] Cohen D B, Mont M A, Campbell K R, et al. Upper extremity physical factors affecting tennis serve velocity. *American Journal of Sports Medicine*, 1994, 22(6):746-750.
- [15] Z. Huang, J. Tang, G. Shan, J. Ni, Y. Chen and C. Wang. An Efficient Passenger-Hunting Recommendation Framework with Multi-Task Deep Learning. *IEEE Internet of Things Journal*. 2019, DOI: 10.1109/JIOT.2019.2901759.
- [16] Zhang, M., Zhang, D., Goerlandt, F., Yan, X., & Kujala, P. Use of HFACS and fault tree model for collision risk factors analysis of icebreaker assistance in ice-covered waters. *Safety science*,

2019,111(1): 128-143.

[17]Zhou, T., Zhang, J. Analysis of commercial truck drivers' potentially dangerous driving behaviors based on 11-month digital tachograph data and multilevel modeling approach. *Accident Analysis and Prevention*, 2019,132, 105256.

[18]Wei, W., Song, H., Li, W., Shen, P., & Vasilakos, A. Gradient-driven parking navigation using a continuous information potential field based on wireless sensor network. *Information Sciences*, 2017, 408(2):100-114.

[19]Dong L, Wu W, Guo Q, et al. Reliability-Aware Offloading and Allocation in Multilevel Edge Computing System. *IEEE Transactions on Reliability*, 2019, DOI: 10.1109/TR.2019.2909279.

[20] Xiaojuan Sun, Hong Zhang, Wenjun Meng, Ronghui Zhang, Kening Li, and Tao Peng.

Primary resonance analysis and vibration suppression for the harmonically excited nonlinear suspension system using a pair of symmetric viscoelastic buffers. *Nonlinear Dynamics*, 2018, 94(2):1243-1265.

List of abbreviations

MBV

(minimum breakdown voltage)

Declarations

Availability of data and material

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

declares that he has no conflict of interest.

* Research involving human participants and/or animals

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

* Informed consent

All authors agree to submit this version and claim that no part of this manuscript has been published or submitted elsewhere.

Funding

None

Authors' contributions

Zhong Li Xiaowrote the entire article.Part of the experiment was with Ping-lin Huang

Acknowledgements

None

References

- [1] Feng, Ren, Jiang. Generic framework for content-based stereo image/video retrieval (Image and vision processing and display technology). *Electronics Letters*, 2011, 47(2):97-98.
- [2] Hockaday S. Evaluation of image processing technology for applications in highway operations: final report. *Physical Review B*, 1991, 14(3):1323-1326.
- [3] Nachtergaele L, Catthoor F, Balasa F, et al. Optimization of memory organisation and partitioning for decreased size and power in Vvideo and image processing systems. *Leukemia Research*, 1995, 30(11):1377-1384.
- [4] Aizawa K, Sakaue K, Suenaga Y. Image processing technologies. Algorithms, sensors, and applications. *Embo Journal*, 2004, 18(12):3419–3430.
- [5] Camana P. Image processing: Video-bandwidth compression: A study in tradeoffs: In this fledgling field, designers are matching criteria to applications and developing `first-cut` hardware. *IEEE Spectrum*, 1979, 16(6):24-29.
- [6] Zhenhua Huang, Xin Xu, Juan Ni, Honghao Zhu, Cheng Wang. Multimodal representation learning for recommendation in Internet of Things. *IEEE Internet of Things Journal*, 2019, 6(6): 10675-10685.
- [7] Borie B, Gautier Y. An autonomous automatic incident detection equipment using video image processing: EVA. *Environmental Science & Technology*, 1995, 29(4):74–76.
- [8] Secker A, Taubman D. Lifting-based invertible motion adaptive transform (LIMAT) framework for highly scalable video compression. *IEEE Transactions on Image Processing*, 2004, 12(12):1530-1542.
- [9] Hanna G, Cuschieri A. Image display technology and image processing. *World Journal of Surgery*, 2001, 25(11):1419-1427.
- [10] Li Y N, Lu Z M, Niu X M. Fast video shot boundary detection framework employing pre-processing techniques. *Iet Image Processing*, 2009, 3(3):121-134.
- [11] Mudde R F, Schulte H B M, Akker H E A V D. Analysis of a bubbling 2-D gas-fluidized bed using image processing. *Powder Technology*, 1994, 81(2):149–159.
- [12] Yokoo A, Taniguchi H. Application of video image processing to detect volcanic pressure waves: A case study on archived images of Aso Volcano, Japan. *Geophysical Research Letters*, 2004, 31(23):275-295.
- [13] Nishikata K, Kimura Y, Takai Y, et al. Real-time lock-in imaging by a newly developed high-speed image-processing charge coupled device video camera. *Review of Scientific Instruments*, 2003, 74(3):1393-1396.
- [14] Cohen D B, Mont M A, Campbell K R, et al. Upper extremity physical factors affecting tennis serve velocity.. *American Journal of Sports Medicine*, 1994, 22(6):746-750.
- [15] Z. Huang, J. Tang, G. Shan, J. Ni, Y. Chen and C. Wang. An Efficient Passenger-Hunting Recommendation Framework with Multi-Task Deep Learning. *IEEE Internet of Things Journal*. 2019, DOI: 10.1109/JIOT.2019.2901759.
- [16] Zhang, M., Zhang, D., Goerlandt, F., Yan, X., & Kujala, P. Use of HFACS and fault tree model for collision risk factors analysis of icebreaker assistance in ice-covered waters. *Safety science*, 2019,111(1): 128-143.
- [17]Zhou, T., Zhang, J. Analysis of commercial truck drivers' potentially dangerous driving behaviors based on 11-month digital tachograph data and multilevel modeling approach. *Accident Analysis and Prevention*, 2019,132, 105256.
- [18]Wei, W., Song, H., Li, W., Shen, P., & Vasilakos, A. Gradient-driven parking navigation using a

continuous information potential field based on wireless sensor network. Information Sciences, 2017, 408(2):100-114.

[19] Dong L, Wu W, Guo Q, et al. Reliability-Aware Offloading and Allocation in Multilevel Edge Computing System. IEEE Transactions on Reliability, 2019, DOI: 10.1109/TR.2019.2909279.

[20] Xiaojuan Sun, Hong Zhang, Wenjun Meng, Ronghui Zhang, Kening Li, and Tao Peng. Primary resonance analysis and vibration suppression for the harmonically excited nonlinear suspension system using a pair of symmetric viscoelastic buffers. Nonlinear Dynamics, 2018, 94(2):1243-1265.

Figure 1:

Title: Background Difference Method Schematic

Legend: Figure 1 show Background Difference Method Schematic

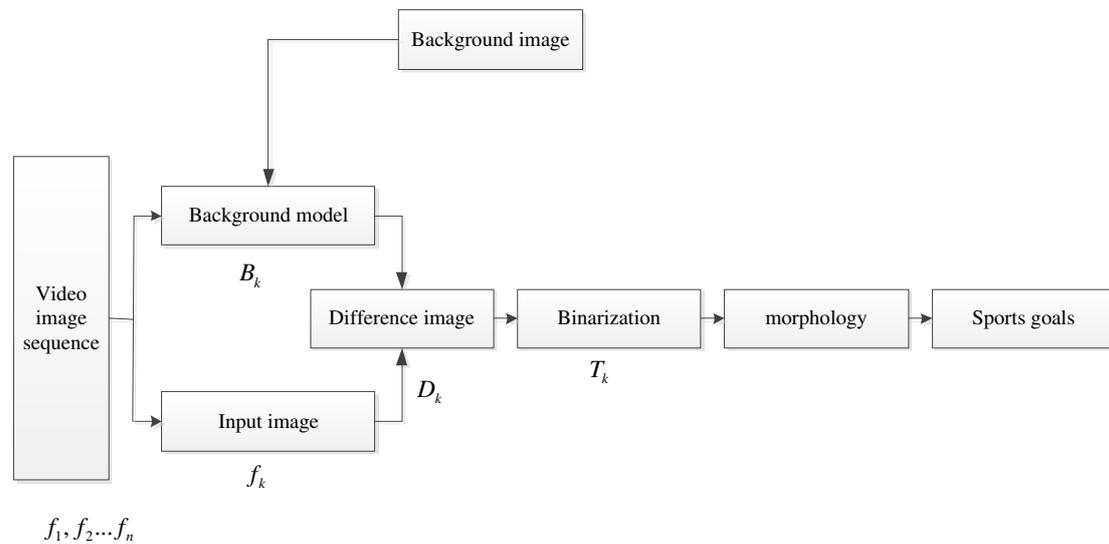


Figure 2:

Title: Stable performance of the improved algorithm before and after the signal

Legend: Figure 2 show Stable performance of the improved algorithm before and after the signal

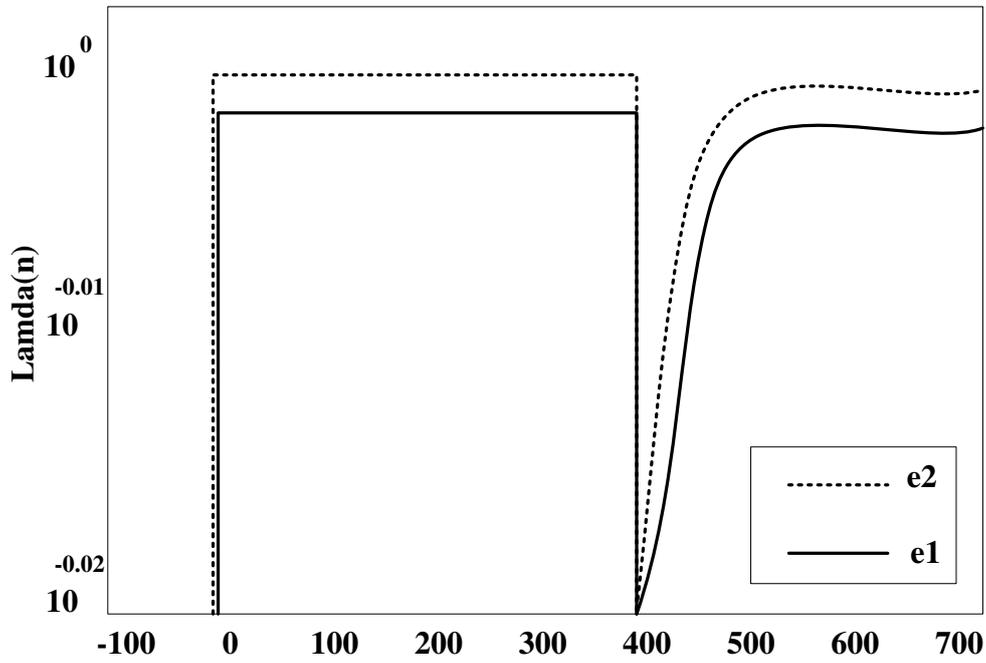


Figure 3:

Title: Schematic diagram of the difference between frames

Legend: Figure 3 show Schematic diagram of the difference between frames

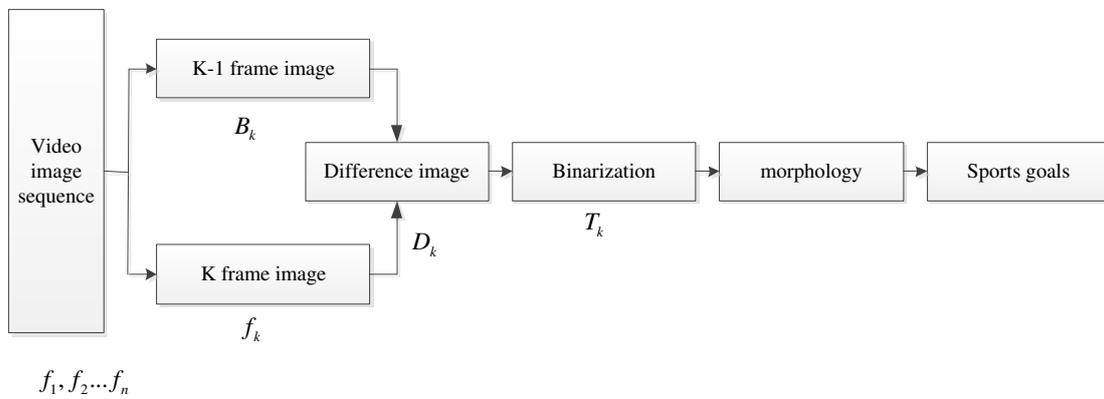


Figure 4:

Title: Three-frame difference method schematic

Legend: Figure 4 show Three-frame difference method schematic

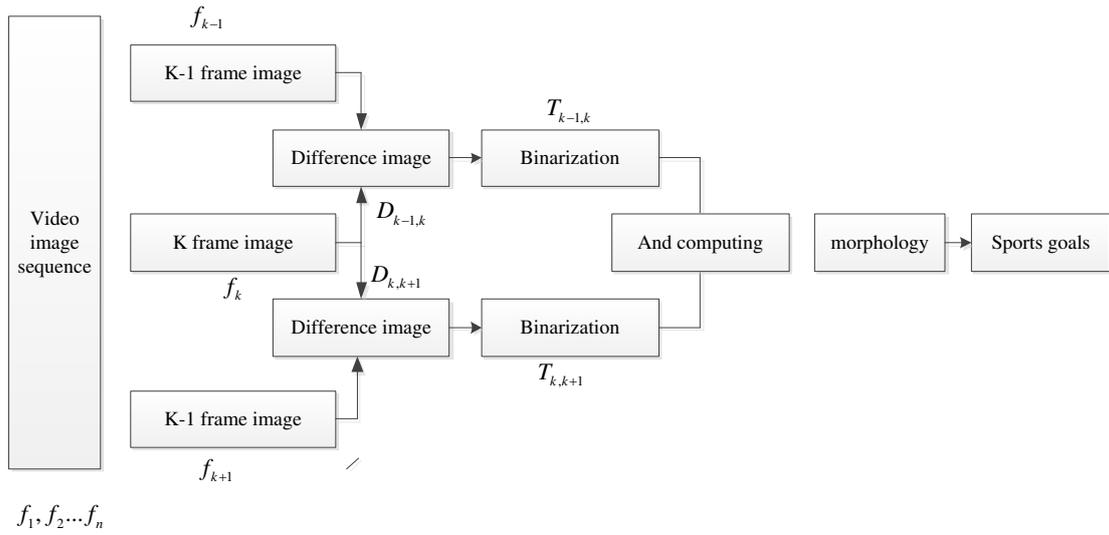


Figure 5:

Title: Athlete Joint Point Capture

Legend: Figure 5 show Athlete Joint Point Capture

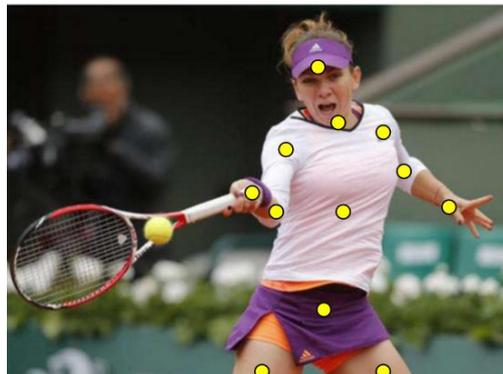


Figure 6:

Title: Results of tennis singles

Legend: Figure 6 show Results of tennis singles

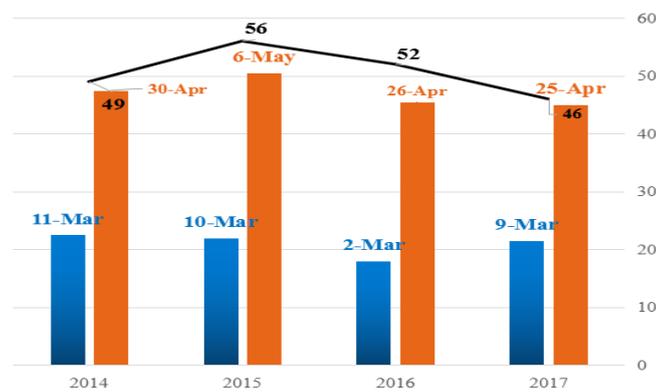


Figure 7:

Title: Example of Dynamic Tracking of Human Joint Node Model

Legend: Figure 7 show Example of Dynamic Tracking of Human Joint Node Model

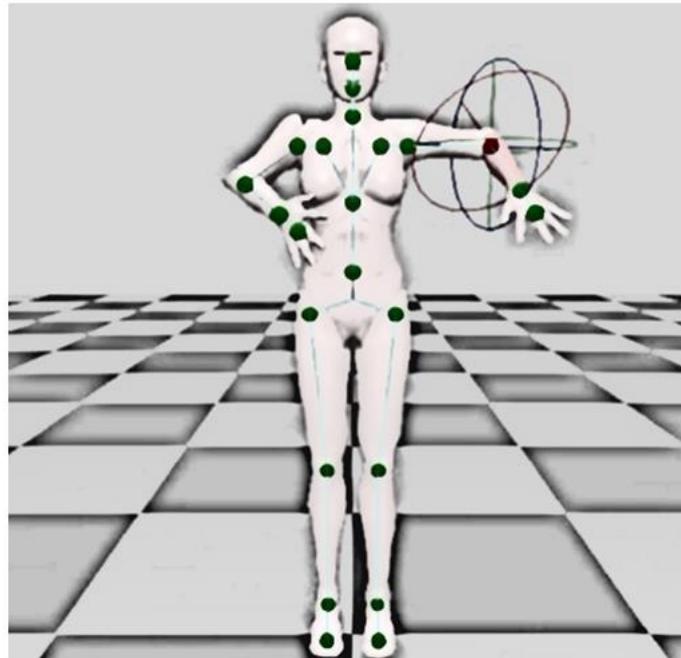
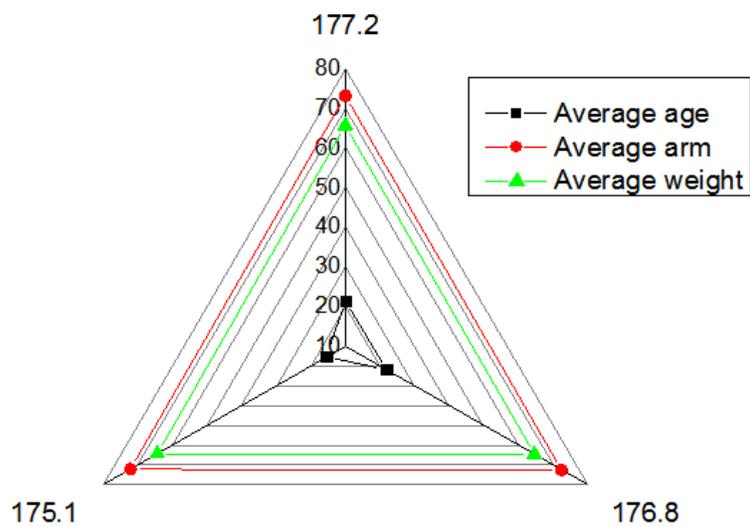


Figure 8:

Title: Simulation results

Legend: Figure 8 show Simulation results



Figures

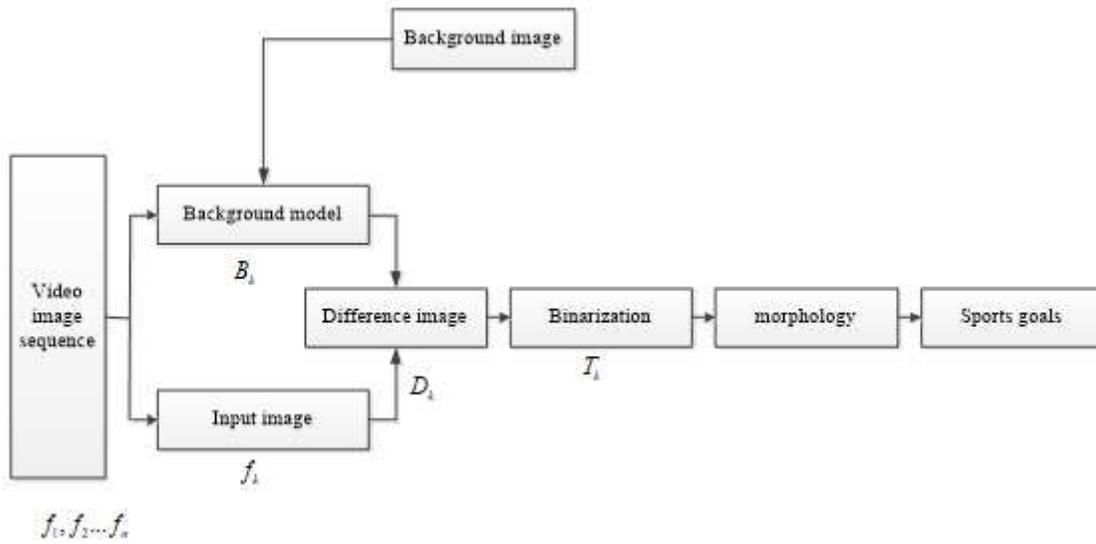


Figure 1

Background Difference Method Schematic

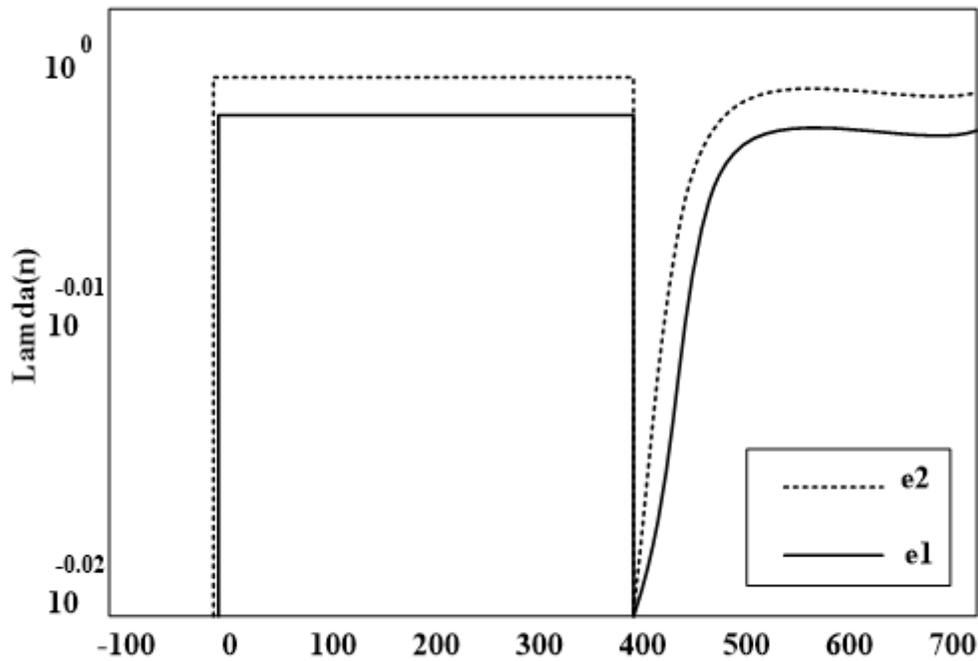


Figure 2

Stable performance of the improved algorithm before and after the signal

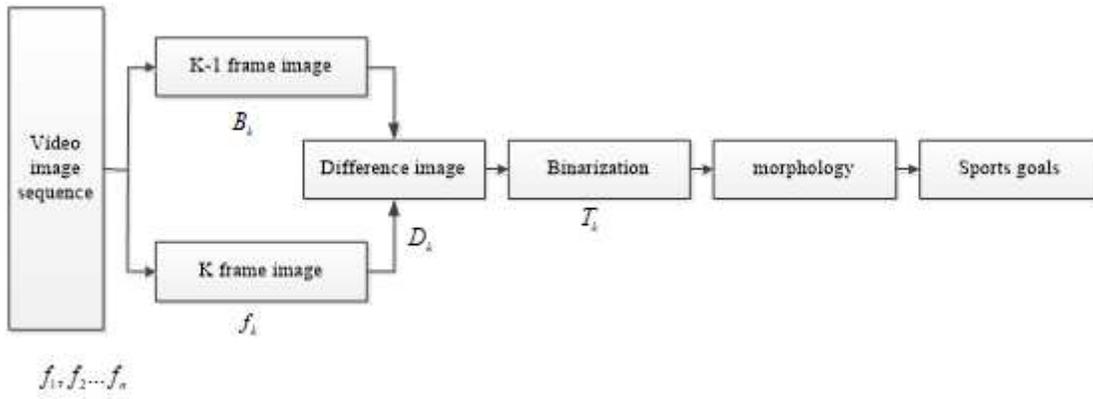


Figure 3

Schematic diagram of the difference between frames

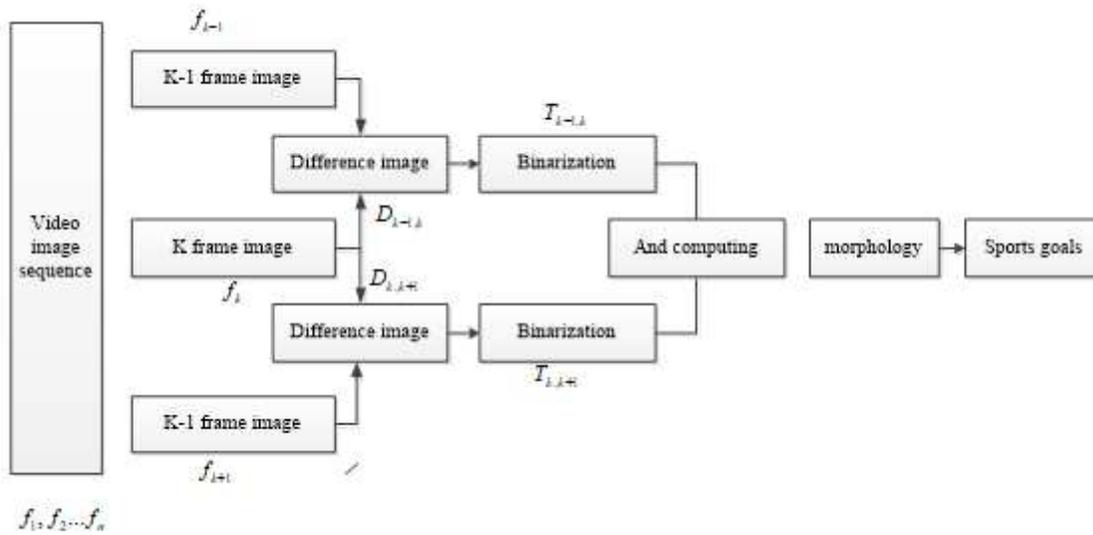


Figure 4

Three-frame difference method schematic



Figure 5

Athlete Joint Point Capture

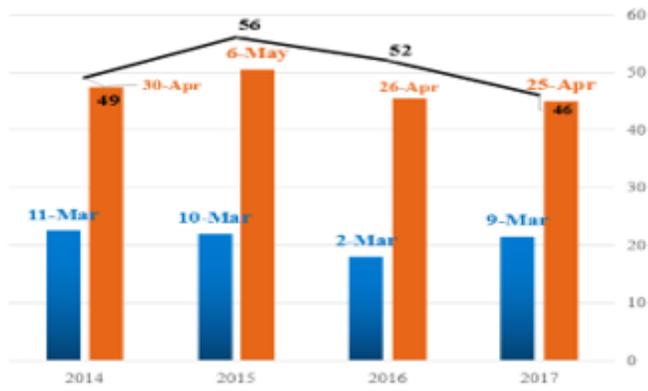


Figure 6

Results of tennis singles

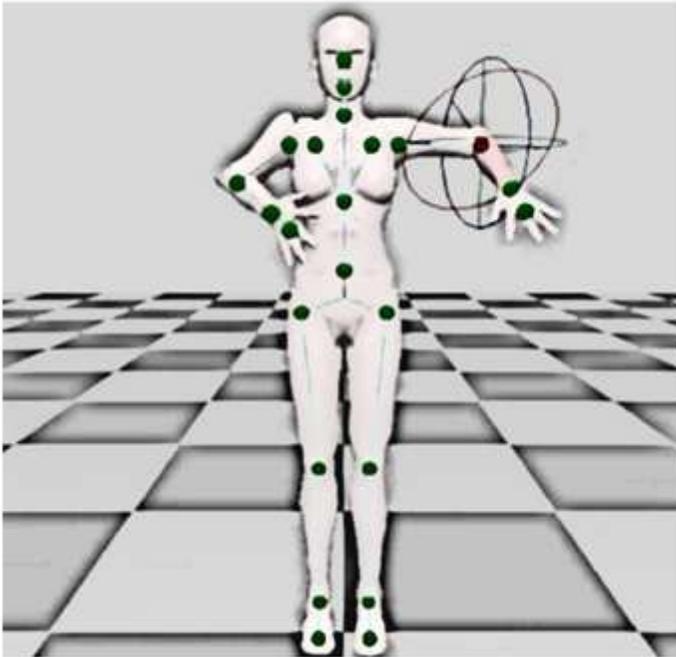


Figure 7

Example of Dynamic Tracking of Human Joint Node Model

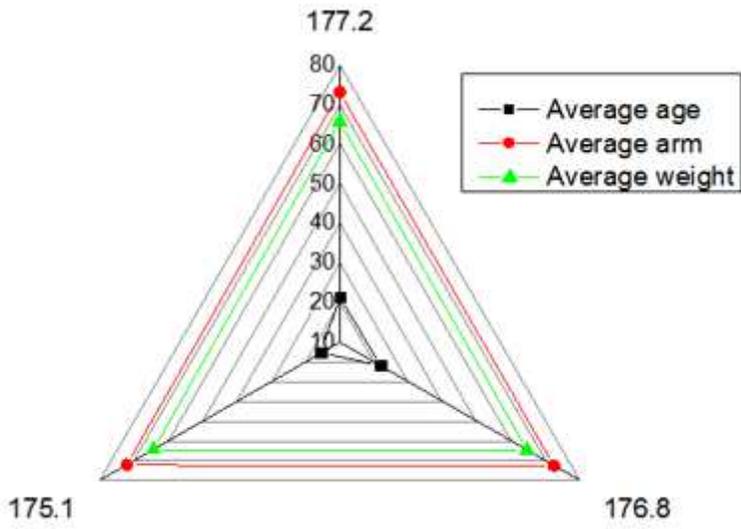


Figure 8

Simulation results