ORIGINAL ARTICLE

Adaptive strategy for sports video moving target detection and tracking technology based on mean shift algorithm

Hongquan Yu[1] · Amit Sharma[2] · Parv Sharma[3]

Abstract The continuous improvement in the level of sports competition has led to many recent research designs for providing easy and quick ways for athlete training. The aim behind this research is to present an adaptive hybrid non-rigid target tracking method by adopting (Mean-shift) and color histogram algorithm to process the characteristics of sports video. This work attempts in designing a tracking algorithm by implementing mean shift algorithm for tracking the object characteristics of sports objects. The experimental analysis presents the ideal effects of proposed approach in precision tracking. Mean shift algorithm uses the gradient method to iteratively calculate the extreme points of the probability density function using its characteristics of no parameters and fast pattern matching. In order to realize the tracking of human targets in sports videos, a tracking approach combining the mean shift process and the color histogram process is proposed. Using the statistical robustness of the mean shift process and the characteristics of rapid convergence along the direction of the density gradient, matching of the color histogram to the target shape is done. It solves the problem of variable target shape and high tracking complexity. The proposed method yields 96.04% precision and 97.10% accuracy value for tracking and recognition. The experimental outcomes obtained for the research provides the suitable evidence that the approach presented in this paper has an ideal effect.

Keywords Sports video · Target detection · Target tracking · Mean shift algorithm · Tracking complexity

1 Introduction

With the improvement of the level of sports competition, how to easily and quickly understand the training level of athletes, extract sports parameters, and provide help for athletes' daily training has become an urgent problem for sports workers. In the past, the mode of sports training only based on the coach's observation records and intuitive judgments can no longer meet the needs of today's constantly improving competitive level (Dai-Hong et al. 2019; Yao et al. 2018; Hao et al. 2019). Therefore, computer vision technology is applied to sports training, machine vision has better accuracy and memory than the human eye, can quickly capture sports targets, and can record various sports data of the targets (Huang et al. 2018; Hossain and Lee 2019). Video tracking technology is the process of detecting each independent moving target in each frame of image and locating the target in subsequent frames. In sports, due to the non-rigid characteristics of the human body, the difficulty of target tracking is increased, and the stability and accuracy of the algorithm are more demanding. Yan et al. has disclosed a paper on various approaches that implements hybrid process for the

✉ Amit Sharma
sharma.amit@chitkara.edu.in

Hongquan Yu
hongquany219@gmail.com

Parv Sharma
parvsharma560@gmail.com

1 Physical Education Department, Jingdezhen University, Jingdezhen 333000, China

2 Department of Computer Science and Engineering, Chitkara University, Chandigarh, Punjab, India

3 Department of Computer Science, JUIT Waknaghat, Solan, Himachal Pradesh, India

detection of moving target and its tracking through the video frames. It is observed from the experimentation that the hybrid process can estimate the position along with the target shape for efficient detection. Their work further describes the non-rigid shape of object, and increases the accuracy for detection to 42.8%, it offers effective solution for detection through monitoring of sports video frames (Yan 2019; Ying 2019; Hamilton et al. 2019). Hui et al. has disclosed a tracking technology on the basis of motion video of sports by implementing Mean-Shift approach. In the present disclosure, firstly, the prediction method is used to primarily identify the location of target, and then implements the Mean-Shift approach to evaluate repetitive computations for determining the real target position. Experimental outcomes presents the global search of Mean-Shift process, and the enhanced method presents the quantity of iterations, decreases the complication of computational for the process, decreases the interval of intake, and confirms the tracking process in real-time (Hui 2019). Zhang et al. demonstrates detection and tracking of moving target by implementing the approaches of Cam shift and Kalman Filter in the frame of sport video. In this paper, an improved adaptive Gaussian-Mixture framework is implemented as a background model, Camshift and Kalman filter are then implemented for tracking the ball and players (Zhang 2019a). Iqbal et al. demonstrates Computer-Vision stimulated real-time approach for the detection, tracking and locking of moving target independently. The evaluation of the processes has been confirmed on the basis of recorded video as well as considering frame of live streams. The results from the investigation reveals that the scheme has the capability for efficiently tracking the moving objects proficiently. The suggested scheme is suitable for the application of object detection and performs complex tracking in several productions like transport, sports, defense etc. (Iqbal et al. 2013).

This work contributes in providing a hybrid non-rigid target tracking approach on the basis of Mean-shift approach and the process of color histogram for exploiting the characteristics of a sports video. The mean shift approach has the characteristics of no parameters and fast pattern matching which are used to iteratively calculate the extreme points of the probability density function. Different from many algorithms in the past, the mean shift algorithm is combined with the color histogram tracking algorithm, which can not only track and estimate the location of the target object effectively. At the same time, the shape of the target can also be well described. Finally, an example is presented in this manuscript, and the experimental outcomes demonstrates that the proposed approach has good accuracy, effectiveness and adaptability (Chi et al. 2020). The statistical robustness of the mean shift approach and the characteristics of rapid convergence

along the direction of the density gradient are used for histogram matching to the target shape. It solves the problem of variable target shape and high tracking complexity. The purpose of tracking of target on the basis of sports video frames is mainly focused for building trajectory and motion information of the sports target. At present, the algorithm in this article has been widely implemented in various sports exercise video analysis system software. With the development in technology, detection of moving target and its tracking becomes more accurate and mature. The issue of correct identification of background color and athletes which may be similar requires consideration for highly accurate detection and tracking. For accurate target detection where every block is unique has brought big challenge. Therefore considering this challenge the objective behind this research is to design an efficient approach for the detection and tracking of moving target. The target is detected and tracked by implementing mean shift algorithm, digital image processing, intelligent recognition on the basis of computer vision technology.

The rest of the article is described as: Literature review is presented in Sect. 2 followed by principle of mean shift algorithm in Sect. 3. Section 4 demonstrates the experimental outcomes and analysis which is further trailed by conclusion and future recommendations in Sect. 5.

2 Literature review

To enhance the strength of the object tracking process in harsh environments, Li et al. proposed a scheme for detection and tracking of object based on motion through context data and closed-loop knowledge. The context section comprised of object part and the current background of object. For each edge in the video frame, the task of tracking is mostly carried out in four fragments of synchronous process: which are detection, followed by tracking, its integration and knowledge learning. Initially, the follower acquires the subsequent probability about the target location and predicts the state of target in the following frame by implementing the spatiotemporal confined statistics. At the same time, the detector combines the follower's context data to search the target in a separate frame and automatically reinitialize it when the tracker fails. Then, the integrator combines the output outcomes of the detector and follower together by optimizing the approach to achieve the best position of the target. At last, the learning phase is implemented for giving the feedback and assisting training models that are produced for updating the detector based on the outcomes of the detector and follower. Through experiments, the capabilities of this approach is validated for several latest technologies on

several standards. The outcomes illustrates that the approach performs well in terms of robustness and provides superior accuracy for tracking (Li et al. 2019). Target tracking process on the basis of joint probability for association of data meets the issues of target object loss and combination eruption in a messy and dense environment of tracking. To overcome these limitations, Shao et al. proposed a tracking process through video frames that depends on the importance of joint probabilistic information membership. The approach detects moving objects in video frames by significance calculation. The detection outcomes are categorized and the resumed classes are then are utilized as valid echoes of information association. At the same time, the confirmation matrix and joint probability in the joint probability information association approach are redefined on the basis of prominence information, and the significance-based confirmation matrix and joint probability are presented. The associated outcomes are satisfactory and corrected by implementing the color suggestion probability. The investigations in real scenarios demonstrates that the proposed approach can suppress background clutter, simplify association measures, and solve the issue of computing combination eruption (Shao et al. 2018). Cui et al. introduced tracking of objects by implementing the graphical modeling. Tracking of object is typically initiated with the methods of object detection. The basic assumption is that the pattern of the target object can be fully detached from its adjacent background. Conversely, for some target objects, such as balls in transmission football video frames, it is difficult to evaluate the valid features to identify balls in a single video frame. The approach implemented here is to classify possible candidate sections of a target object in some successive frames and then utilize custom graphics to construct relationships among the candidate sections. At last, the best path of the graph is extracted by using the Viterbi algorithm as the trajectory of the object. This method is termed as tracking at short-term. It is then implemented to initialize the Kalman filter to perform tracking of long term. During the process of tracking, tracked section is confirmed for determining whether the tracking fails, and if a failure occurs, the tracking of short-term is restarted (Cui et al. 2018).

Wang et al. demonstrates a scheme for detection and tracking by implementing the ABC shift approach. This article utilizes variance in input frame to recognize the preliminary location of moving object. It implements Adaptive Background Camshift approach for updating the background model, and decreases the prospects of identical colors which corresponds to color probability distribution, and attains exact detection and tracking of the target object. The outcomes from investigation demonstrates that that this approach can overcome limitations of deprived outcomes with Camshift approach when the target object arrives a huge region with the parallel color background, and it presents average flexibility in difficult background (Wang et al. 2011). Tian et al. presents a study for the detection of moving object by implementing mining approach which is a hybrid approach and its analysis are carried out through sports video. On the basis of motion in sports video the moving object is considered as the research system for theoretical and technical aspects, considered for layer to layer mining among motion features at low level to the motion features at high level. The approach can not only provide support for users to discover the information quickly, but also delivers decision support for the user to resolve the issues (Tian and Jun 2014).

Zhang et al. described Detection and Tracking of moving target of human through video image frames by implementing Camshift approach. The target identification stage, the improved approach screens and adjusts the target region on the evidence of growing a lesser computation. The stage of target tracking, the computational area is condensed to the standardization and following range, which increases the effectiveness of the approach and ensure the steadiness and strength of the process (Zhang 2019b). Teachabarikiti et al. demonstrates the tracking process for moving players and the detection of ball for the automatic process of tennis sports video annotation. For predicting the activities of players more accurately, their work is extended for sensing and tracking the positions of ball through frame differencing process along with few correlation approaches to eliminate false discoveries. On the basis of both patterns which includes player motion and ball position, the approach can efficiently classifies the action of players into ground stroke of backhand and forehand. The video analysis of broadcast tennis sports have been test which are collected from the Internet source. The experimental outcome demonstrates that the proposed approach is capable for accurately categorizing the player's actions with 83.8% accuracy and 82.4% recall values (Teachabarikiti et al. 2010).

Li et al. describes research on camera-based human body tracking using improved cam-shift algorithm. In this paper authors introduce some common image noise reduction algorithms. In the present disclosure, with the analysis of particle filtering and traditional Cam-shift algorithm, authors has introduced a different human body tracking way that is able to select the target automatically due to the detection result. On the basis of the detection and tracking outcomes, the algorithm of motion parameter estimation is analyzed (Li 2015). Li et al. describes Automatic text detection and tracking in digital video. This system implements a scale-space feature extractor that feeds a simulated neural processor to spot text blocks. Presented text tracking system comprises of two modules: a sum of squared difference (SSD) based module to find the

preliminary position and a contour-based module to improve the position. Experiments performed with a variety of video sources show that above system can sense and track text robustly (Li et al. 2000).

With the development in technology, detection of moving target and its tracking becomes more accurate and mature. The issue of correct identification of background color and athletes which may be similar requires consideration for highly accurate detection and tracking. For accurate target detection where every block is unique has brought big challenge. The location identification is important concern for the initial and accurate estimation of target (Sharma and Singh 2021). The localization is the process of identifying the location of an unknown sensor in a field (Sharma and Singh 2020). The performance of the player closely monitored manually, therefore there is a possibility that the coaches may imbalance in concentrating on several sports players. In order to address these occlusions, Gade et al., 2015 (Gade and Moeslund 2018) had proposed an approach for the analysis of free throw followed by evaluating the performance of the individual with the help of image data.

3 Principle of mean shift algorithm

The mean shift algorithm is a non-parametric kernel density estimation theory, which uses the gradient method to iteratively calculate the extreme points of the probability density function (Wang et al. 2020; Ou et al. 2020). This algorithm has the characteristics of no parameters and fast pattern matching, and it is an effective target tracking algorithm.

The movement of the players leads to dynamic changes in the imaged scene in image sequences. Additionally, the area covered by every image varies with the camera movement. This problem limits the amount of frames that follows the common frame and hence results differences at pixel level among the images, even when the registration is sorted for compensating the background motion. These issues may presents many difficulties in calculations of target detection in real time. In order to address this issue the traditional differencing algorithm is improved for producing moving target detection and tracking algorithm through image sequences. The presented algorithm is capable of detecting the moving target on the basis of clustering and mean shift algorithm as depicted in Fig. 1.

The first stage is the image pre-processing where the input image is processed for registration and de-noising. In the next stage the processed image is partitioned for the detection of moving target and it's tracking for identifying the location. In target detection phase the initial location are grouped into clusters for the detection of location. On the basis of location information gathered from target detection and tracking phase the new target is identified. When the identified target is true the histogram of the image is computed through Kernel distribution. If the identified target is not true then the process follows the target updation and the target is updated by using the information of moving target detection. The next phase is location updation for target detection and then based on the location information and histogram information the mean shift algorithm is applied for the target tracking. The last stage is the output result which presents the information about the target.

The mean shift algorithm can "move" each point to the local maximum point of the density function along the shortest path in the probability distribution.

It is based on the theory of kernel density estimation. The principle of the kernel density estimation method is described as:d Vioic space R^d Define any n Set of points $\{X_i\}_i = 1...n$ According to the nuclear $K(x)$ And window radius h Can get a certain point x Multi-core.

Function density estimate $\hat{f}(x)$, The following calculation is expressed as Eq. (1).

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \tag{1}$$

Use the gradient of the kernel function density estimate to define the gradient estimate of the probability density as expressed in Eq. (2):

$$\overset{\wedge}{\underset{x}{\nabla}} f(x) \equiv \nabla_x \hat{f}(x) \equiv \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$$
$$\equiv \frac{1}{nh^{d+1}} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \tag{2}$$

among them

$$\nabla K(x) = \left(\frac{\partial K(x)}{\partial x_1}, \frac{\partial K(x)}{\partial x_2}, ..., \frac{\partial K(x)}{\partial x_n}\right)$$

.

∇_x its about $x_1, x_2, ..., x_n$ The gradient factor.

The gradient estimate of the probability density function is expressed as Eq. (3):

$$\overset{\wedge}{\underset{x}{\nabla}} f(x) \equiv \nabla_x \hat{f}(x) \equiv \frac{2}{nh^{d+2}} \left[\sum_{i=1}^{n} g\left(\left\|\frac{x - X_i}{h}\right\|^2\right) \right] \left[\frac{\sum_{i=1}^{n} x_i g\left(\left\|\frac{x - X_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{x - X_i}{h}\right\|^2\right)} - x \right] \tag{3}$$

where suppose $\sum_{i=1}^{n} \left[g\left(\left\|\frac{x - X_i}{h}\right\|^2\right) \right] \neq 0$.

The last term of calculation is the mean shift vector which is computed as expressed as Eq. (4):

Fig. 1 Process flow of Target Detection Based on clustering and mean shift for target detection in image sequence
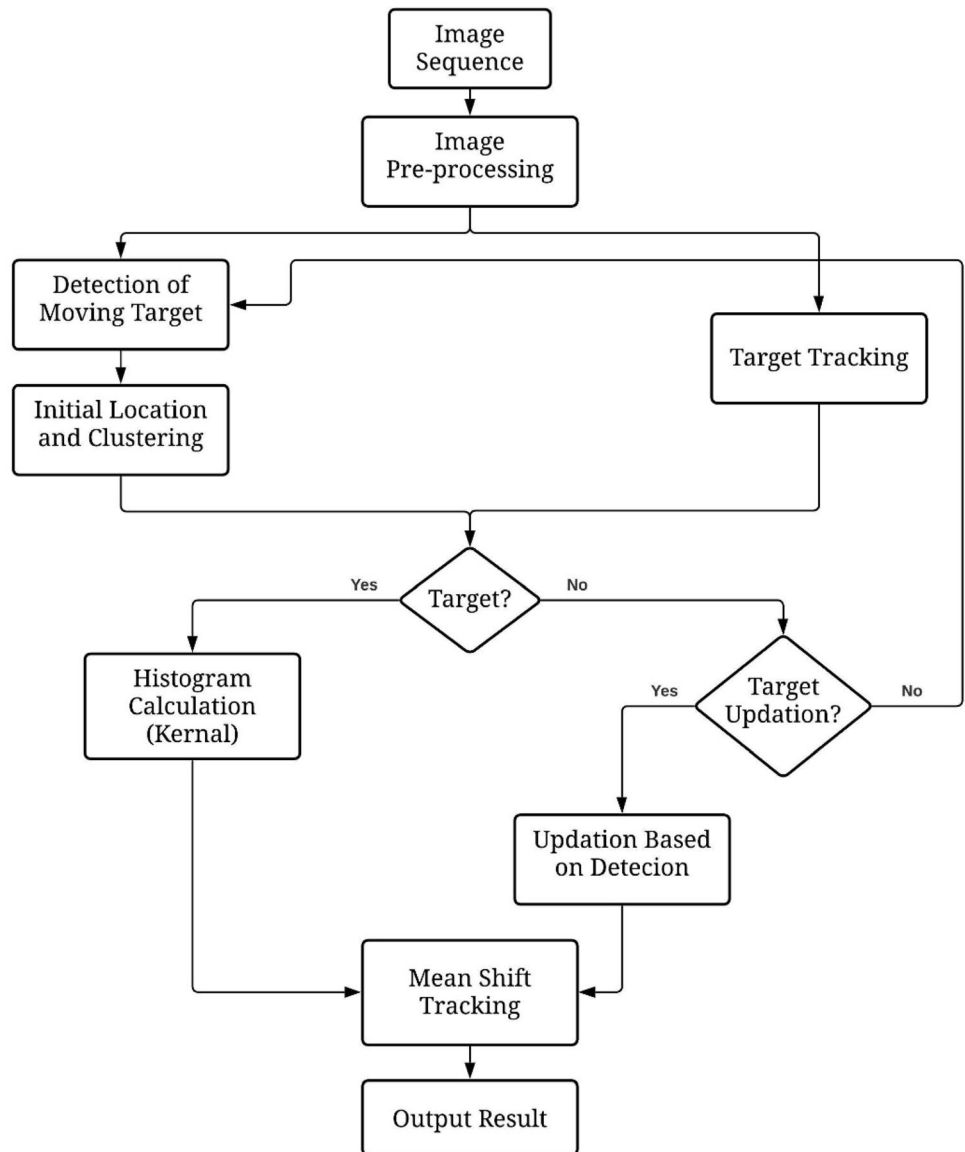
$$M_{h,G}(X) = \left[\frac{\sum_{i=1}^{n} x_i g\left(\left\|\frac{x-X_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{x-X_i}{h}\right\|^2\right)} - x\right] \qquad (4)$$

One of the major utilization of object tracking is carried through mean shift algorithm. The track objects are repaired and described through histograms. Figure 2 depicts the process object tracking through mean shift algorithm. Mean shift algorithm is not only applied for the tracking of objects but also for providing the involved visuals. This approach also provides the confidence map for other new image. Therefore the process of object tracking mainly implemented for gathering the information about color and place which influences probability density function of image. This probability density function is computed from the color histogram of the target object. To

accomplish this, the proposed algorithm initially identifies the confidence map at high points which are located exactly in the old position of an object. The other corresponding steps are implemented for exploring the object that can be applied for the searching process and to calculate the beginning position along with their mean position which are computed in previous process.

3.1 Color histogram tracking and positioning method

The color distribution in the image is the most reliable feature in the image, and it can be described by the color histogram. The color histogram method has good adaptability to dynamic video, image rotation and changes of
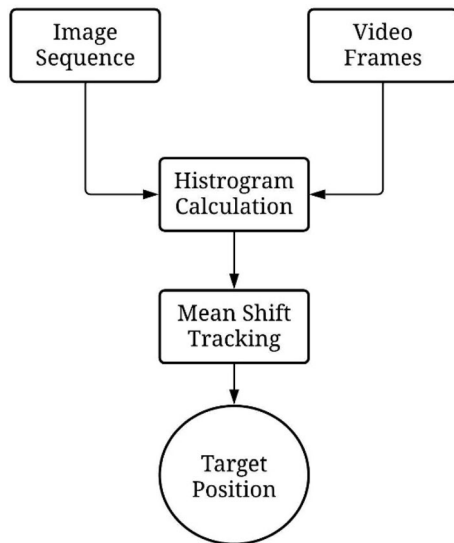
Fig. 2 Process of tracking through mean shift algorithm

observation viewpoint, so it is widely used in video tracking. The specific principle is as follows.

Define data collection $X = \{x_1, ..., x_N\}$ Contains N Independent samples, probability density function $p(x) = N(x; \theta, v)$, θ Is the mean vector, V Is the covariance matrix.

Assuming that the shape of the non-rigid human body target is an ellipse, first initialize the target selection of the current frame image and define x_i Is the position of the pixel contained in the target, θ_0 Is the initial position of the target center. Then the shape of the target can be approximated by the calculation of Eq. (5).

$$V_0 = \sum (x_i - \theta_0)(x_i - \theta_0)^T \qquad (5)$$

The color histogram divides the color space into M discrete color subintervals, function $b(xi) : R2 \rightarrow 1, ..., M$ Defines the color subrange x_i. The pixel chroma value of the position.

The color histogram model of the target contains the histogram $o = [o_1, ..., o_M]^T$ middle M. The value of each color subrange. Zori m The value of each color subrange are calculated by Eq. (6):

$$o_m = \sum_{i=1}^{N_{v_0}} N(x_i; \theta; V_0)\delta[b(x_i) - m] \qquad (6)$$

where δ is the Kronecker function. For the candidate target area of the new frame, the position vector of the area θ and covariance matrix V Determine the shape of the area, use the color histogram function $r = (\theta, V)$ Describe the color characteristics of the area, m the value of each color subrange is expressed in Eq. (7).

$$r_m(\theta, V) = \sum_{i=1}^{N_v} N(x_i; \theta; V)\delta[b(x_i) - m] \qquad (7)$$

The similarity between the current frame candidate target and the target model can be obtained by calculating the similarity of the histogram. This article uses the Bhattacharyya coefficient to measure the similarity of two histograms. The following calculation is expressed in Eq. (8):

$$\rho[r(\theta, V), o] = \sum_{m=1}^{M} \sqrt{r_m(\theta, V)}\sqrt{o_m} \qquad (8)$$

Current estimate $r\left(\theta^{(k)}, V^{(k)}\right)$ Perform a first-order Taylor series expansion to get the coefficients and expressed as Eq. (9):

$$\rho[r(\theta, V), o] \approx c_1 + c_2 \sum_{i=1}^{N_v} \omega_i N(x_i; \theta; V) \qquad (9)$$

among them c_1c_2 is a constant, ω_i The following expression can be calculated as Eq. (10):

$$\omega_i = \sum_{m=1}^{M} \sqrt{\frac{o_m}{r_m\left(\theta^{(k)}, V^{(k)}\right)}}\delta[b(x_i) - m] \qquad (10)$$

To maximize the Bhattacharyya coefficient, that is, the candidate target is most similar to the target simulation, the second term in Eq. (9) should be maximized. The mean shift algorithm can be used to achieve this process.

3.2 Realization of tracking and positioning

Step 1: Customize to obtain the tracking target model of the current frame.

Step 2: Initialize the candidate target window in the new frame.

Initialize the candidate target and set the target center position θ_0 and target pixel position x_i According to Eq. (5), (6) and (7), the shape and color characteristics of the initial candidate targets are calculated respectively.

Step 3: Calculate the Bhattacharyya coefficient of the current candidate target and the target model $\rho_i°$

Step 4: in case $\rho_{i-1} < \rho_i$ Then update the candidate target position according to the mean shift algorithm θ, Update the bit shape of the candidate target according to the color histogram algorithm $V°$

Step 5: Repeat steps 2 and 3 until $\rho_i < \rho_{i-1}°$

Finally, a new position of the tracking target that maximizes the Bhattacharyya coefficient is obtained.

4 Experimental results and analysis

The video sequence image resolution used in the experiment is 324 × 244, the frame rate is 16 frames per second, and there are 10 frames in total. Experimental results obtained by tracking different targets in a video sequence. Four frames of images extracted from the video tracking results of the team players (Yi et al. 2018; Zhang et al. 2019). Tracking results of no team player. Since the moving human body is a non-rigid target, it can be seen from the example diagram that the position of the target moves abruptly and the shape is constantly changing, so the tracking is more difficult. It can be seen from the tracking results that the tracking method proposed in this paper can accurately track the position of the target, and can adjust the tracking window to adapt to the shape of the human target in the current frame (Lu and An 2020).

This paper has improved the hybrid tracking algorithm. The purpose is to more accurately capture the movement of moving targets in different scenes, and avoid the blur or disappearance of moving target features caused by factors such as appearance changes and occlusion (Sharma and Kumar 2019; Sharma et al. 2019; Dhiman et al. 2020a, b; Margarat and Sivasubramanian 2019). In this paper, we use this algorithm to track the moving target in combination with the football game video. The specific process is as follows.

4.1 Using the mean shift tracking algorithm to delineate the target

The first step is to determine a certain athlete in the video as the tracking object. In order to improve the tracking accuracy, this paper uses the mouse to select the tracking target in the first frame of the video and mark the target with a red circle. Figure 3 is a diagram of a tracked moving target.

The second step is to use the mean shift tracking algorithm for the moving target, first select the feature space of the moving target circled by the red circle, and then use the kernel density estimation method to establish the corresponding probability density distribution for the target feature, and describe the target feature. The third step is to measure the probability density distribution of the target model, compare it with the probability density distribution of the candidate target, and select the similarity measure as the target coefficient.

The third step is to use the iterative search method of the mean shift vector to determine the position of the target in the current frame of the image, so as to realize the tracking of the target. It should be noted that in the process of using this method to track the target, because the clutter background of the sports field itself will affect the gray characteristics of the target, the target model must be continuously updated throughout the tracking process to enhance its robustness.. The mean shift algorithm is more efficient, can achieve the purpose of fast pattern matching, and the algorithm is more effective. But the algorithm will not be able to accurately match the target when there is severe occlusion. Therefore, a color histogram is used for further tracking.

4.2 Using color histogram tracking algorithm to segment the target

Because the mean shift vector algorithm may not accurately match the target in the case of severe occlusion or continuous tracking, the color histogram method is used to accurately segment the target. This method uses the target color feature as the basis for segmentation, and segmentation is based on the color appearance probability, which has strong stability. The histogram corresponding to the tracking moving target is shown in Fig. 4 where x-axis
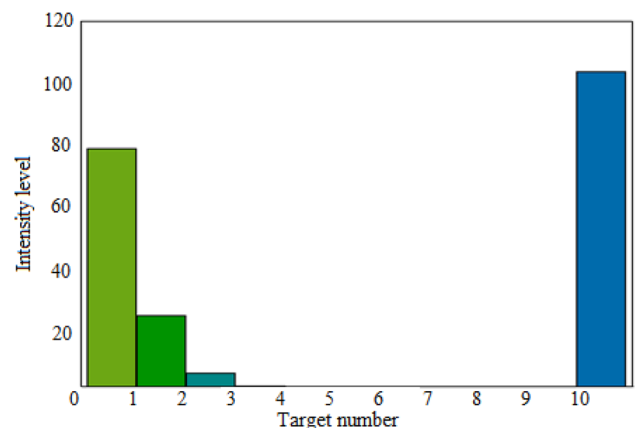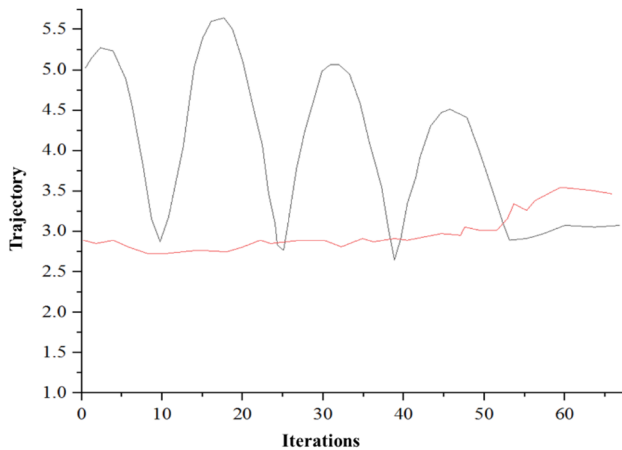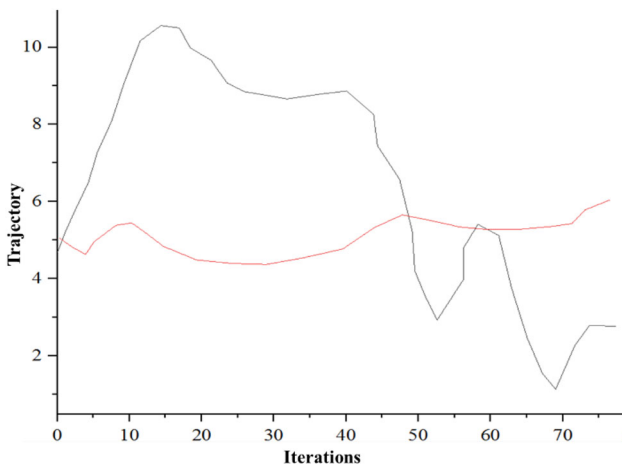

Fig. 3 Tracked moving target

Fig. 4 Histogram corresponding to tracking moving target

Table 1 Technical parameters during the test

Test video	Mean shift average number of iterations	Calculation time/ms
Pingpong	6	3542
Horizontal bar	4	5074
Football	5	7253
Weightlifting	4	12,931

Fig. 5 Table tennis tracking trajectory diagram

Fig. 6 Barbell track diagram

Table 2 Results for table tennis tracking and barbell tracking

TP	TN	FN	FP	Precision	Recall	Accuracy
Table tennis tracking						
98	96	2	4	96.08%	98%	97.81%
Barbell tracking						
95	97	5	3	96.94%	97%	97.10%

represents the target number and y-axis represents intensity level of the target.

In this paper, based on the mean shift vector algorithm to calculate the density function to delineate the target object, the color histogram tracking algorithm is used to separate the moving target and realize the tracking of the moving target. The specific method is to set the shape of the moving target in the sports video to be an ellipse, select the moving target in the current frame of image, perform the initialization operation, and define the position of the tracking moving target containing the pixel as x_i. The initial position of the moving target center is θ_0. Through Eqs. (5–7) the shape of the moving target can be obtained.

After getting the shape of the moving target, it is necessary to initialize the candidate target window in a new frame of image, and then initialize it until the shape and color characteristics of the moving target are finally obtained, so as to realize the dynamic and accurate tracking of the moving target.

### 4.3 Hybrid tracking algorithm

Based on the characteristics of fast target movement and rapid change of target speed in sports video, this paper proposes four motion models: uniform speed, uniform acceleration, static and collision. Since only one motion mode is often considered, the effect is often poor when tracking targets whose motion modes often change. Although it can be solved by increasing the number of sample points and expanding the range of sample point distribution. The computational complexity will also increase significantly, the error will increase, and many useless sample points will be generated. In this paper, by introducing a variety of motion models, the motion model is dynamically updated according to the characteristics of the target motion, thereby using a small number of samples to achieve better results, and reducing the computational complexity, while combining with the mean shift algorithm to improve the tracking performance accuracy. The advantage of the mean shift algorithm is that it can obtain a local optimal solution and can obtain higher accuracy. The disadvantage is that it requires a higher initial predicted position. If the initial predicted position differs greatly from the actual position of the target, it often fails to converge to the local optimal solution. In order to overcome the above shortcomings, this paper calculates and predicts the target o_m Similarity to the target model $\rho[r(\theta, V), o]$, when $\rho[r(\theta, V), o] > \theta$. When the target motion model has not changed, when $\rho[r(\theta, V), o] < \theta$. When, change the motion model to predict the initial position again, until $\rho[r(\theta, V), o] > \theta$. Condition $\rho[r(\theta, V), o] > \theta$. The distance between the initial predicted position and the actual position of the target is small,

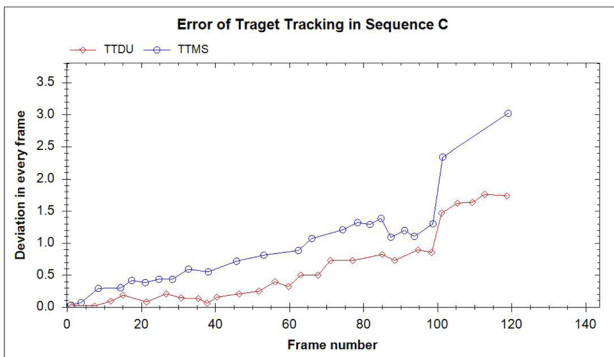**Table 3** Performance comparison with the state-of-the-art methods

Methods	Precision	Recall	Accuracy
Teachabarikiti, et al. (Teachabarikiti et al. 2010)	83.7%	82%	—
Margarat and Sivasubramanian (Sharma et al. 2019)	93.6%	—	—
Proposed algorithm	96.04%	97%	97.10%
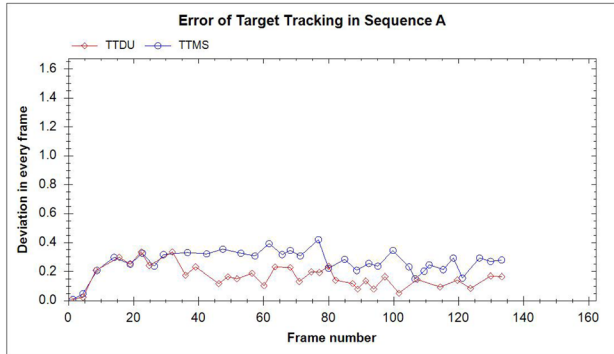
(a)
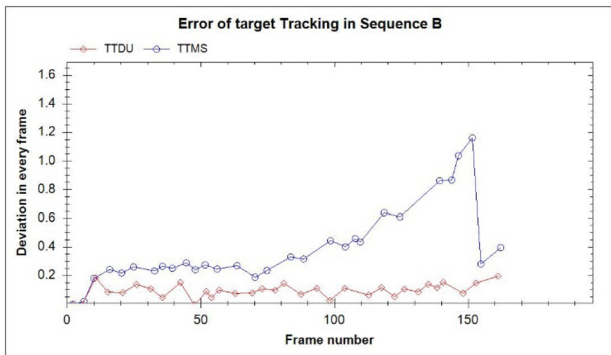
(b)

(c)

Fig. 7 Plots of error in target tracking (**a** to **c** respectively), comparative analysis in terms of tracking error generated from TTDU and TTMS through **a**, **b** and **c** sequences

which ensures that the mean shift algorithm converges to the local optimal solution. In the experiment, we take $\theta = 0.8$.

In order to verify the effectiveness of this algorithm, we tested a variety of sports video sequences with different types of sports and different resolutions, and all achieved good results. Table 1 shows the technical parameters during the experiment in this article, and Figs. 5 and 6 are the track diagrams of sports video tracking.

Figures 3 and 4 presents the trajectory diagrams of table tennis tracking and Barbell track respectively, for the sports video tracking. It is visible from the graphical representation that the proposed method does not have and achieves the expected results. The research results of table tennis tracking and barbell tracking are presented in Table 2.

The results obtained in terms of precision, accuracy and recall reveals that the presented algorithm performs better providing effective results for sports video detection and tracking. The results are also compared with the state-of-the-art methods in Table 3.

As depicted in Fig. 7, the observed tracking error for TTDU (target tracking based on detection updates) is 0.27 on average which is comparatively smaller than TTMS (target tracking mean shift) is 0.57. In this figure TTMS error presents increasing value for every sequence and sharp rise-fall are observed when the object is disappeared from the image. In case of contrast, when the accuracy of target detection is efficient, more specifically when the detection probability is greater than 92%, TTDU can efficiently controls the accumulation of error and attain good target tracking through tracking model updation for the sequence of A and B.

Additionally, with the increase in height of image in range from 450 to 750 m and then to 1050 m the observed tracking error for TTMS is increased from 0.28 to 0.42 and then 1.03. This is due to the reason that with the increase in height the resolution of image decreases, and leads to less contrast and also huge overlap in characteristics of grayscale for target and its background. However, TTDU reduces the ill influence of image height for tracking target and presents efficient tracking outcomes.

The proposed algorithm outperforms the state-of-the-art methods in terms of precision, recall and accuracy. This comparison justifies the effectiveness of the sports video detection and tracking system as the method used does not have tracking loss, and provides robust detection approach.

5 Conclusion

The algorithm has the characteristics of no parameters and fast pattern matching. Combined with the color histogram algorithm, not only can the position of the target be effectively tracked and estimated, but the shape of the target can also be depicted well, which solves the problem of complex and difficult-to-track non-rigid target shapes in sports videos. In order to test the effectiveness of this algorithm this work attempts to build a sports video detection and tracking system on the MATLAB platform. In this article an image and video analysis based for target tracking by implementing clustering and mean shift algorithm is proposed. This article presents the discussion of earlier investigated various methods of tracking. The implementation of mean-shift algorithm for tracking the target is evaluated by computing the histograms. The test results show that the method used does not have tracking loss, and achieves the expected results in terms of 96.04% precision, 97% recall and 97.10% accuracy value for tracking and recognition.

Although this article has achieved certain results, due to time and conditions, there are still many possibilities for further exploration. The color histogram algorithm proposed in this article can have a good tracking effect on moving targets. The disadvantage of relying too much on the color of the moving target, once the background color is close to the moving target or there are objects of similar color in the background, it is very easy to interfere with the tracking results. The proposed algorithm presents the improvement in tracking of objects and its accuracy is tested with simulation example. The proposed process presents good performance of tracking however, new features can be mined as tracking conditions to improve the tracking effect. In the later stage, based on this algorithm, it can be applied to the PTZ (Pan-tilt-zoom) camera for security monitoring, using proposed algorithm to control the PTZ movement to achieve the effect of target following. The recorded video can be transmitted to the cloud and mobile devices for monitoring and notification for enhancing the robustness.

References

Chi YH, Hu LH, Gao X, Li SF (2020) Research on infrared passive ranging algorithm based on unscented Kalman filter and modified spherical coordinates. J Phys: Conf Ser. https://doi.org/10.1088/1742-6596/1629/1/012066

Cui Z, Zhang J, Zhang X, Xu L (2018) Research on object tracking based on graph model in sports video. J Inf Technol Res (JITR) 11(3):1–14

Dai-Hong J, Lei D, Dan L, San-You Z (2019) Moving-object tracking algorithm based on PCA-SIFT and optimization for underground coal mines. IEEE Access 7:35556–35563

Dhiman G, Oliva D, Kaur A, Singh KK, Vimal S, Sharma A, Cengiz K (2020) BEPO: a novel binary emperor penguin optimizer for automatic feature selection. Knowl-Based Syst 211:106560

Dhiman G, Singh KK, Soni M, Nagar A, Dehghani M, Slowik A, Cengiz K (2020) MOSOA: A new multi-objective seagull optimization algorithm. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2020.114150

Gade R, Moeslund TB (2018) Constrained multi-target tracking for team sports activities. IPSJ Trans Comput Vis Appl 10(1):1–11

Hamilton JA, Lambert G, Suss J, Biggs AT (2019) Can cognitive training improve shoot/don't-shoot performance? Evidence from live fire exercises. Am J Psychol 132(2):179–194

Hao X, Liu J, Yuan X, Tang X, Li Z (2019) A moving block sequence-based evolutionary algorithm for resource-constrained project scheduling problems. Int J Bio-Inspir Comput 14(2):85–102

Hossain S, Lee DJ (2019) Deep learning-based real-time multiple-object detection and tracking from aerial imagery via a flying robot with GPU-based embedded devices. Sensors 19(15):3371

Huang X, Zhang T, Deng Z, Li Z (2018) Design of moving target detection and tracking system based on cortex-A7 and openCV. Traitement du Signal 35(1):61

Hui Q (2019) Motion video tracking technology in sports training based on mean-shift algorithm. J Supercomput 75(9):6021–6037

Iqbal J, Pasha SM, Baizid K, Khan AA, Iqbal J (2013) Computer vision inspired real-time autonomous moving target detection, tracking and locking. Life Sci J 10(4):3338–3345

Li J (2015) Research on camera-based human body tracking using improved cam-shift algorithm. Int J Smart Sens Intell Syst 8(2):1104–1122. https://doi.org/10.21307/ijssis-2017-798

Li H, Doermann D, Kia O (2000) Automatic text detection and tracking in digital video. IEEE Trans Image Process 9(1):147–156

Li J, Wang J, Liu W (2019) Moving target detection and tracking algorithm based on context information. IEEE Access 7:70966–70974

Lu Y, An S (2020) Research on sports video detection technology motion 3D reconstruction based on hidden Markov model. Clust Comput 23(3):1899–1909

Margarat GS, Sivasubramanian S (2019) basketball tracking using mean shift algorithm. J Adv Res Dyn Control Syst 11(7):158–171

Ou M, Wang J, Wu Y, Yi J (2020) Research and implementation of the HD video real-time edge detection system based on FPGA. J Phys: Conf Ser. https://doi.org/10.1088/1742-6596/1646/1/012144

Shao H, Shen J, Zhang Z, Liu H (2018) Research and analysis of video image target tracking algorithm based on significance. Int J High Perform Syst Archit 8(1–2):82–93

Sharma A, Kumar R (2019) Service-level agreement—energy cooperative quickest ambulance routing for critical healthcare services. Arab J Sci Eng 44(4):3831–3848

Sharma A, Singh PK (2020) Taxonomy on localization issues and challenges in wireless sensor networks. Recent Adv Electr Electron Eng (Former Recent Patents Electr Electron Eng) 13(2):193–202

Sharma A, Singh PK (2021) Localization in wireless sensor networks for accurate event detection. Int J Healthc Inf Syst Inform (IJHISI) 16(3):74–88

Sharma A, Kumar R, Talib MWA, Srivastava S, Iqbal R (2019) Network modelling and computation of quickest path for service-level agreements using bi-objective optimization. Int J Distrib Sens Netw 15(10):1550147719881116

Teachabarikiti, K., Chalidabhongse, T. H., & Thammano, A. (2010). Players tracking and ball detection for an automatic tennis video annotation. In 2010 11th International Conference on Control Automation Robotics & Vision (pp. 2461–2494). IEEE.

Tian H, Jun WX (2014) Study on the detection of moving target in the mining method based on hybrid algorithm for sports video analysis. Sens Transducers 181(10):16

Wang D, Xie SR, Li C, Li QM, Luo J (2011) Target detection and tracking based on abcshift algorithm. Comput Eng 37(4):203–205

Wang T, Wang X, Shi W, Zhao Z, He Z, Xia T (2020) Target localization and tracking based on improved Bayesian enhanced least-squares algorithm in wireless sensor networks. Comput Netw 167:106968

Yan R (2019) Researches on hybrid algorithm for moving target detection and tracking in sports video. Clust Comput 22(2):3543–3552

Yao T, Luo Y, Chen Y, Yang D, Zhao L (2018) Single-image super-resolution: A survey. International Conference in Communications Signal Processing and Systems. Springer, Singapore, pp 119–125

Yi H, Zhengmin L, Lejun C, Jinxiu L (2018) Research on community detection algorithm based on site topic similarity and topology. High Technol Lett 24(2):75–81

Ying Y (2019) Research on college students' information literacy based on big data. Clust Comput 22(2):3463–3470

Zhang B (2019a) Moving target detection and trackingbased on camshift algorithm and kalmanfilter in sport video. Int J Perform Eng 15(1):288

Zhang Y (2019b) Detection and tracking of human motion targets in video images based on camshift algorithms. IEEE Sens J 20(20):11887–11893

Zhang B, Jiao D, Lv X (2019) A target detection algorithm for SAR images based on regional probability statistics and saliency analysis. Int J Remote Sens 40(4):1394–1410