Dynamic badminton recognition based on binocular vision

Wei Sun[1,2*], Jianyi Kong[1,2], Huaiguang Liu[1,2], Yong Zeng[1,2], Wanting Huang[1,2]
[1]Key Laboratory of Metallurgical Equipment and Control Technology, Wuhan University of Science and Technology, 430080 Wuhan, China
[2]Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, 430080 Wuhan, China
Emails: sw35812@gmail.com (W. Sun); 15697188659@wo.com.cn (J. Kong);
lhg81219@163.com (H. Liu); 36188155@qq.com (Y. Zeng); 731864070@qq.com (W. Huang)
*Corresponding author

Abstract:

Moving targets tracking technology is the core of a potential computer vision technology. This paper aims at introducing how to identify the trajectory of dynamic badminton and capture dynamic badminton. Firstly, a contour filtering algorithm is proposed which considers contour, color as well as motion characteristics. In order to acquire the position of the badminton, the hybrid Gaussian background difference (HGBD) method is applied to extract the two-dimensional coordinates of the badminton in each camera. Then, the three-dimensional coordinates of the badminton is obtained by calibrating the binocular vision. After experiments, it indicates that the proposed method can track the dynamic badminton precisely. The recognition accuracy rate in 2D plane can get to 97.5%.

Keywords: Badminton robot; Hybrid Gaussian background difference method; Binocular vision; Moving targets tracking; Contour filtering algorithm

1. Introduction

After nearly half a century of in-depth research and development, video tracking technology has been vigorously developed and applied, which has broad application prospects in military guidance, visual navigation, security monitoring, intelligent transportation, medical diagnosis, meteorological analysis etc. Moving targets tracking technology is a challenging research topic in the field of computer vision[1]. The target recognition plays a vital role in major science and technology fields such as bio-medicine[2], human-computer interaction[3], virtual reality[4], robotics[5], multimedia contexts[6, 7], video surveillance[8], video streaming[9, 10], health-care systems[11] and smart indoor security systems[12, 13]. Due to the complexity of the natural environment (light, climate change, etc.) in a real environment, the interference of target detection and tracking will lead to inaccurate detection and low tracking efficiency. So the improvement of moving targets detection and tracking algorithm has very realistic significance and application value.

Over the past several years much work has been reported in the literature about moving targets detection and tracking. X. Lan et al.[14] proposed a new joint sparse representation model for robust feature-level fusion. The proposed method dynamically removes unreliable features to be fused for tracking by using the advantages of sparse representation. X. Lan et al. [15] proposed a novel multiple sparse representation framework for visual tracking which jointly exploits the shared and feature-specific properties of different features by decomposing multiple sparsity patterns. I. Comlekciler et al.[16] used the artificial intelligence methods (neural network and neuro-fuzzy system) to increase the accuracy of positioning the jaws during the real-time practice. Y. Qe et al.[17] proposed a human behavior recognition method based on binocular stereo vision and human face-hand features. C. Chen et al.[18] presented a robust person tracking method based on particle swarm optimization (PSO) algorithm. T. Gao et al.[19] proposed a new multiple objects tracking method based on video frames. A type of particle filtering combined with the Scale Invariant Feature Transform is applied to intelligent video surveillance system. X. Sun et al.[20] presented a novel particle filter approach to distinguish the target and background in the space effectively. F. Li and S. Liu[21] proposed an object-tracking algorithm through a cooperative appearance model. Y. Zha et al. [22] proposed a robust tracking method under intrinsic and extrinsic varieties. The object tracking problem is considered as a conversion learning problem. S. Tian et al.[23] proposed a multi-type multi-object tracking algorithm by introducing on-line inter-feedback information between the detection and tracking processes into the tracking-by-detection method. Y. Jang et al.[24] presented a target model update scheme for the mean-shift tracking with background weighted histogram. A. Li et al.[25] proposed an object tracking approach that utilizes background cues only. M. Yang et al.[26] presented a novel solution to this dilemma by considering the context of the tracking scene. J. Liu et al.[27] proposed a partial feature selection method, including coarse initial sub-block selection and refined sub-block sample particle bi-directional selection under particle filter tracking framework. H. Jiang et al.[28] presented a novel online object tracking algorithm by using multi-feature channels with adaptive weights. M. Lucena et al.[29] presented three optical flow-based observation models suitable for particle filter tracking. Y. Gao et al.[30] proposed a symbiotic tracker ensemble towards a unified tracking framework, which is only based on the output of each individual tracker without knowing its specific mechanism. The object tracking is mainly about low-speed objects. Further research is needed for tracking high speed objects.

As one of the fastest balls, badminton tracking is still a challenge task. The research object is the badminton, which has significant motion characteristics-faster movement speed, smaller size and not obvious two-dimensional shape features. That increases the difficulty of testing and tracking. In this paper, a vision system of badminton trajectory recognition based on binocular vision is developed for the badminton robot. The two-dimensional features recognition of the badminton is carried out by using the novel contour filtering algorithm. The three-dimensional coordinates of the badminton are calculated by binocular vision calibration to obtain a continuous flight trajectory of the badminton.

2. Introduction of Badminton Robot

2.1 Scene overview

Badminton robot is a machine which can catch the badminton and shuttlecock automatically. In practice, more than one badminton robot cooperates with each other to complete a game. Fig. 1 shows the layout of the badminton robot. The venue of the badminton court is a standard size. Robots capture the badminton with binocular vision systems and gain the trajectory through image processing. After obtaining the trajectory, the next badminton's position can be predicted. Therefore, robot can adjust its positions and postures to fight back the badminton.

2.2 Badminton robot

As shown in Figure 2, the badminton robot usually consists of four parts: chassis, frame, sensors and batter mechanism. The chassis of the badminton robotic adopts three-wheel omnidirectional structure, which can achieve the movements in any direction on the field and zero-radius turn. Batter mechanism is made up of a plane four-bar mechanism. The racket is bound to the mechanism and driven by a single servo motor. By changing the cylinder pressure, the angle between the four-bar mechanism and the rack and the racket angle, the serve and return of the ball with different angles and different intensity will be achieved. So as to achieve the design goal of human's action simulation of reserve and return the ball.

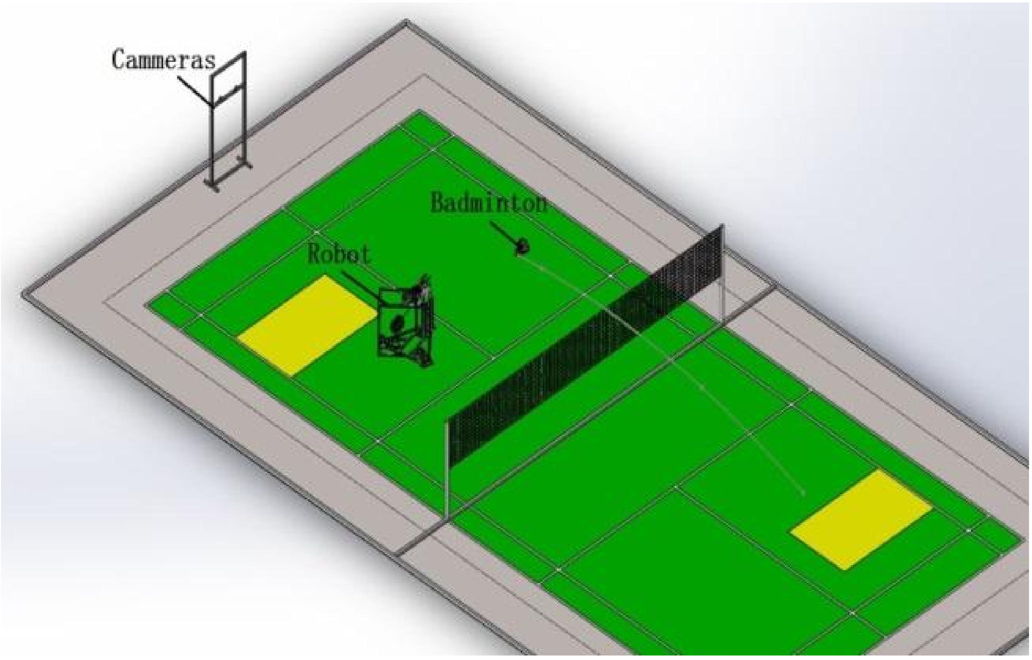


Fig. 1 global graph

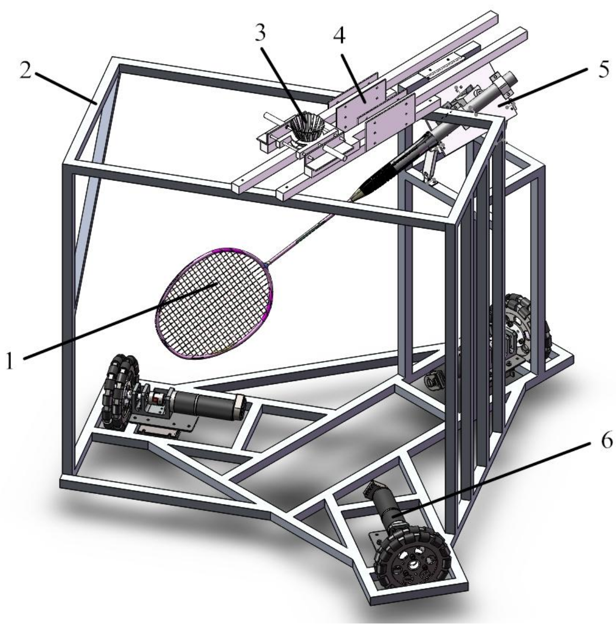


Fig. 2 Badminton robot
1. battledore 2.frame 3.badminton 4.inserting balls mechanism
5. batter mechanism 6.omnidirectional wheel chassis

2.3 Binocular hardware

Binocular vision hardware usually includes more than one identical camera to get images of different view and a computer to calculate the position of the target. In this paper, two large constant cameras are used. Two cameras are installed on a rack built by aluminum alloy. The computer is fixed to the chassis.

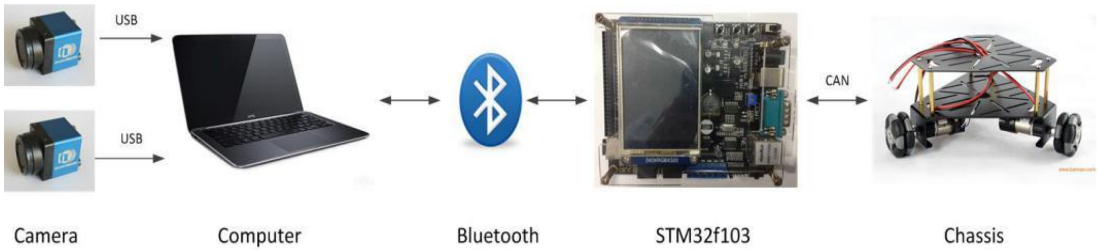


Fig. 3 hardware connection figure

Binocular vision system hardware connection is shown in Fig. 3. The two cameras capture images at the same time. And images are transferred to the computer through universal serial bus protocol. The trajectory obtained by the computer through image processing will be transmitted to the badminton robot controller through the Bluetooth. Then the robot controls the chassis and the batting mechanism.

The process flow diagram consists of five processing steps as shown in Fig. 4. In the initialization step, two cameras capture images at the same time and images are transferred to the computer through Universal Serial Bus protocol. In the second step, the motion objects are detected by background subtraction method of Gaussian mixture model algorithms. Badminton trajectory with multiple characteristics in the 2D coordinates is tracked in the third step. Fourth, binocular stereo vision measures the three-dimensional coordinates of the badminton. Then, the coordinates can be sent to the moving robot in the fifth step. If the three-dimensional coordinates do not meet the conditions, go to step 1. Meanwhile, the robot adjusts the coordinates.

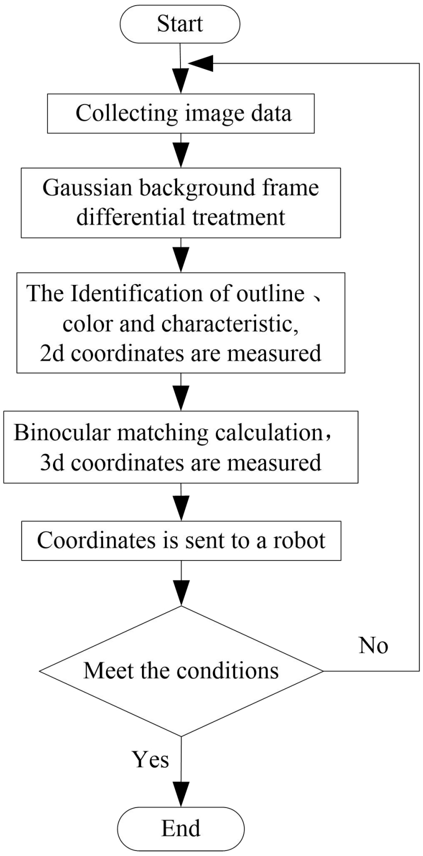


Fig. 4 process flow diagram

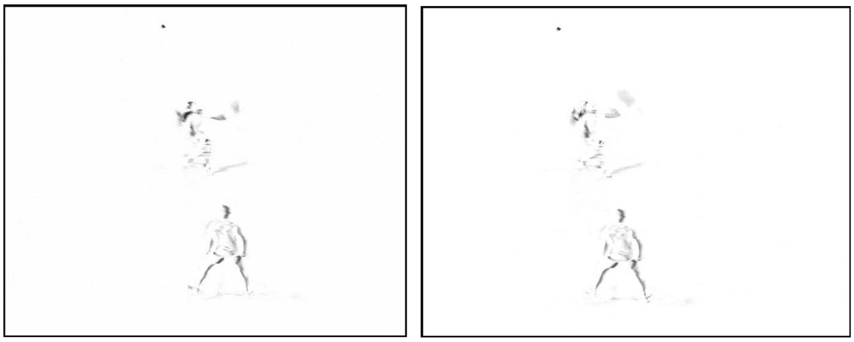
3. Two-dimensional badminton image processing

3.1 Hybrid Gaussian Background Difference Method

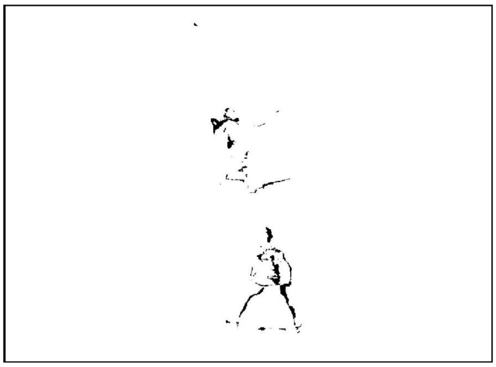
At present, there are many methods for detecting moving objects, in which frame difference method has been widely used because it is simple and effective. Three-frame difference method is an improved two-frame difference algorithm. Its main idea is to take three consecutive frames of images and make frame differences for the adjacent two frames.Then do the AND operation on the results of the two frame differences. The algorithm can get the region of the moving object better when the moving object's speed and frame frequency are suitable.

As can be seen from the Fig. 5, the badminton shows out clearly by the three-frame difference method. But the players also exist because they move at the same time, which will cause big interference to the detection of badminton. Moreover, the human in the difference image is not a monoblock but divided into several small blocks. That will disturb the badminton detection and influence the accuracy as well as the stability of detection. Except the method of inter-frame difference, there are background difference method and optical flow method commonly used for target detection [31]. Optical flow method uses the optical flow field formed by the change target gray level to detect the target, but its real-time is poor. Therefore, the background difference method with higher efficiency will be used in this paper. Background difference method is also known as background subtraction method. Its principle is similar to the inter-frame difference method, which uses the current frame and the obtained background image for the difference. The specific steps are as follows:

The 142th frame The 143th frame The 144th frame



Frame different between 142th frame and 143th frame Frame different between 143th frame and 144th frame



Three-frame difference picture

Fig. 5 Frame Difference picture

(1) Assuming that the current frame and previous frame of the image sequence are $F_k(x,y)$ and $B(x,y)$ respectively, the two frames are differentiated:

$$D_k(x,y)=|F_k(x,y)-B(x,y)| \tag{1}$$

(2) The difference image is binarized as follows:

$$B_k(x,y)=\begin{cases}255 & D_k(x,y)>T\\ 0 & else\end{cases} \tag{2}$$

Where, the value T is the threshold of binarization (generally between 140 and 160).

(3) During the procedure of image acquisition, some noise may be brought to the images. There are usually two types of operations: filtering and morphological processing for image denoising. However, the separate background image is hard to be extracted directly in a dynamic background. So background modeling is usually executed alternately.

Gaussian Mixture Model (GMM) is a kind of common background modeling method. Gaussian Mixture Model uses K (generally between 3 and 5) Gaussian models to represent the characteristics of each pixel in image. When an image acquired, the mixed Gaussian model is then updated according to the formulas change. Each pixel in current image matches each Gaussian mixture model. If it succeeds, it is judged as the background point. Otherwise it is the foreground point.

The idea of single Gaussian background modeling is to assume that n observations

$X = [x_1, x_2, \cdots, x_n]$ of a point X in the image obey the Gaussian distribution. Then the probability distribution of the point can be expressed by the Gaussian probability density function:

$$\eta(x, \mu, \delta) = \frac{1}{\delta\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\delta^2}\right\} \tag{3}$$

Where, μ represents the expectations of normal distribution. δ represents the standard deviation of normal distribution . Pixels will be classified to the background or the foreground by comparing the pixel value with the preset threshold. Because the tree branches swing, light changes etc. may lead to that there are multiple pixel values in some locations in the background image, so the mixed Gaussian Background modeling method is used in practical application. That is, taking K Gaussian probability density functions to represent the probability distribution of the point. Let X_t denotes the pixel value at time t . Then the probability density function of the point is:

$$f(X_t) = \sum_{i=1}^{K} w_{i,t} \cdot \eta(X_t, \mu_{i,t}, \delta_{i,t}) \tag{4}$$

Where, $\eta(X_t, \mu_{i,t}, \delta_{i,t})$ represents the ith Gaussian model at time t . $w_{i,t}$ is its weighting coefficient. And K is the number of Gaussian distribution. When a new video frame is received, the parameters of the Gaussian distribution are updated as follows:

$$w_{i,t} = (1-\alpha) w_{i,t-1} + \alpha \tag{5}$$

$$\mu_{i,t} = (1-\rho) \mu_{i,t-1} + \rho X_t \tag{6}$$

$$\delta_{i,t}^2 = (1-\rho)\delta_{i,t-1}^2 + \rho(X_t - \mu_{i,t})^2 \tag{7}$$

where, α is defined as learning efficiency $(\alpha \in [0,1])$, generally adopt 0.03. ρ is the learning rate of the mean, and δ represents variance of Gaussian distribution. It can be calculated as $\rho = \partial / \omega_{i,t}$, and also can be treated as the rate of current frame integrated into the background.

When the variance and mean of whole Gaussian model are learning according to equations (5) ~ (7), the stability, accuracy and convergence of the model will be affected directly when adopting different learning mechanisms.

From Fig. 6, the noise of HGBD image is increased when compared to the inter-frame difference image. But the game players are more obvious as a monoblock, not fragments in inter-frame difference in Fig. 5. In general, the HGBD is more sensitive to the movement of the objects and can detach objects more completely.

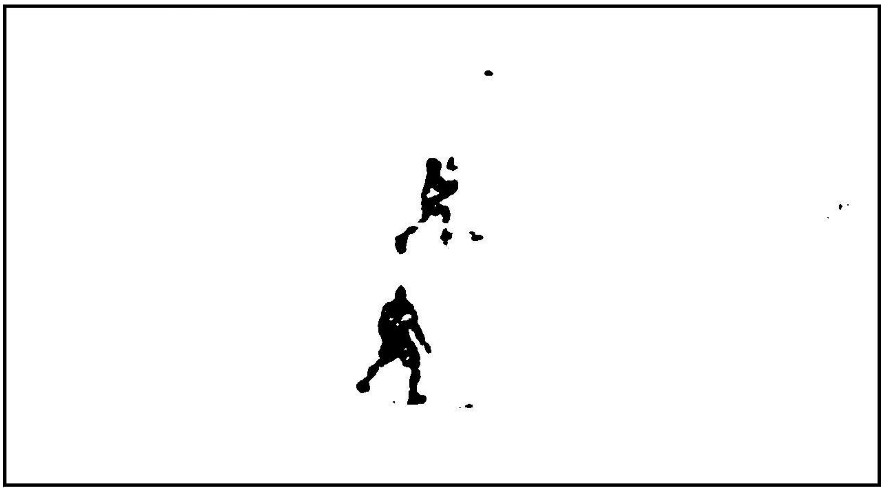


Fig. 6 the HGBD result

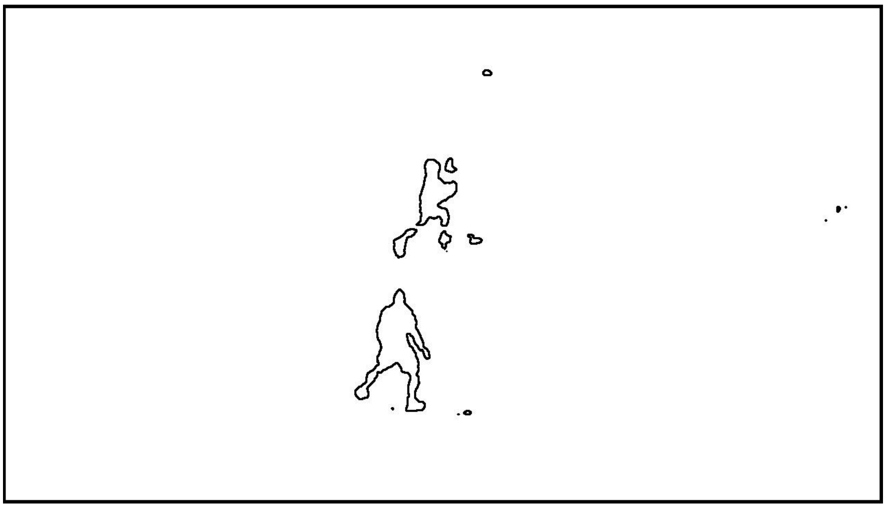


Fig. 7 the contour result

The contours of human gained by Gaussian background difference method will be larger and easier to be filtered out.

3.2 Badminton trajectory tracking with multiple characteristics

Contour feature is a stable character commonly used in target tracking. The contour is the edge of an object as shown in Fig. 7, which is the sequence of pixels at the edge of the connected region in the image. The contour of each frame image acquired probably more than one, and there are still disturbances even if the denoising process is performed.

In order to obtain the coordinates of the badminton, it is the beat situation that there should be only one contour. For this reason, the contour filtering is needed. Fig. 7 shows that the noise is usually small and small in size. The noise contour can be detected by the contour length and the radius of the smallest enclosing circle. Finally, obtain the outline of the badminton. Meanwhile, the two-dimensional coordinates of the badminton can be obtained.

(1) Filtering contour by the contour length

Contour length can be measured by the number of points in the contour. The non-target contours are all treated as noise in the contour image. Compared to the contour of badminton, the contour of the player is large enough and the other noise contours are very small. So it can filter contour through the contour size (that is the contour points) directly.

(2) Filtering contour by the size of the tilted bounding

The tilted bounding of the contour is the smallest enclosing of the contour which can present the approximate geometrical size and rotation characteristics. Because of the lack of frames or the possibility of stretching by ghosting in high speed, the size of the badminton is essentially elliptical. So the shortest side length of the tilted bounding of the badminton is basically the same. The minimum edge length of all contours can be calculated. In this case, the contour with the smallest side length greater than a certain value can be deleted. Some large-size contours can continue to be filtered out.

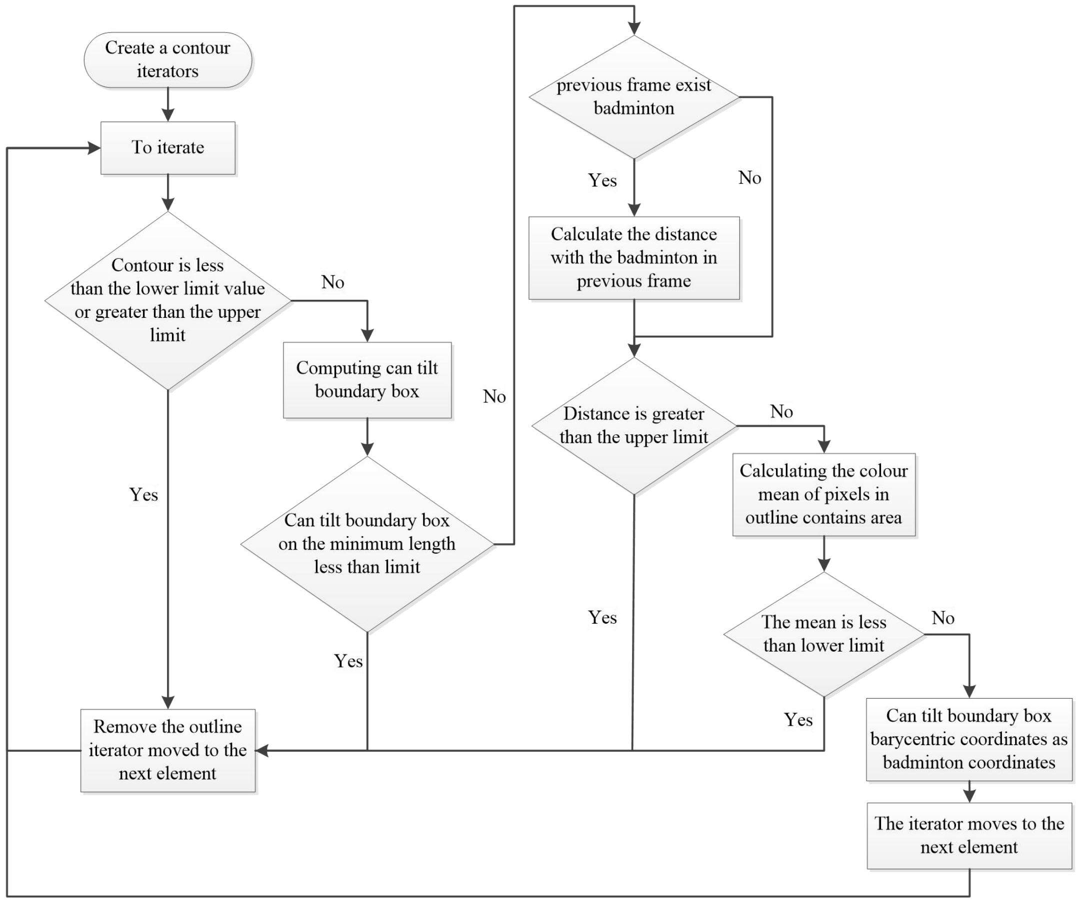


Fig. 8 flow chart of contour filter algorithm

(3) Filtering contours by color features

There are only some small noise contours left after the above steps. In general, the color of the badminton is obvious as the white color differs greatly with the surrounding environment. As a result, the color of the pixels in each contour contained can be counted and the contours that are far away from the white will be filtered out.

(4) Filtering contours by motion features

Except the contour and color features, the badminton also has significant movement characteristics which can be expressed by the change in coordinates of the badminton in the front and rear frames. The center coordinates of the sloped boundary will be the badminton coordinates in the current frame. The motion distance filtering can be added to enhance the stability of the overall filtering algorithm with this motion feature.

The specific algorithm is: Firstly, calculating the boundary for the tilted contour. If the badminton of the previous frame exists, the distance of badminton coordinates between the current frame and the previous one is calculated with the centroid of the tilted bounding as coordinates. If the distance is greater than the specified value, delete the contour.

Fig. 9 video tracking results

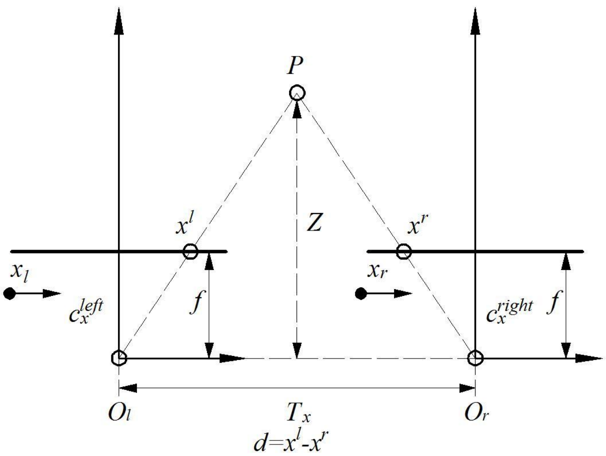


Fig. 10 the principle diagram of the binocular

A badminton match video is selected as shown in Fig. 9. It can then be concluded that this video has a total of 280 frames, including 7 frame tracking errors.

4 Three-dimensional coordinates of badminton

4.1 Binocular calibration

Binocular Stereo Vision is an important aspect of the machine vision. It obtains two images of the measuring object from the parallax of different positions and the principle of the imaging device. The object three-dimensional geometric information is obtained by calculating the corresponding point of position deviation of the image. The two eyes obtain the images and observe the difference between them, so that a clear sense of depth can be computed.

Stereo matching mainly obtains the disparity map according to the principle of triangulation by finding the correspondence between each pair of images. After obtaining the parallax information, the depth information and the three-dimensional information of the original image can be easily obtained according to the projection model. In the previous section, the features of badminton are extracted and the corresponding features are matched to convert the two-dimensional coordinates into three-dimensional coordinates.

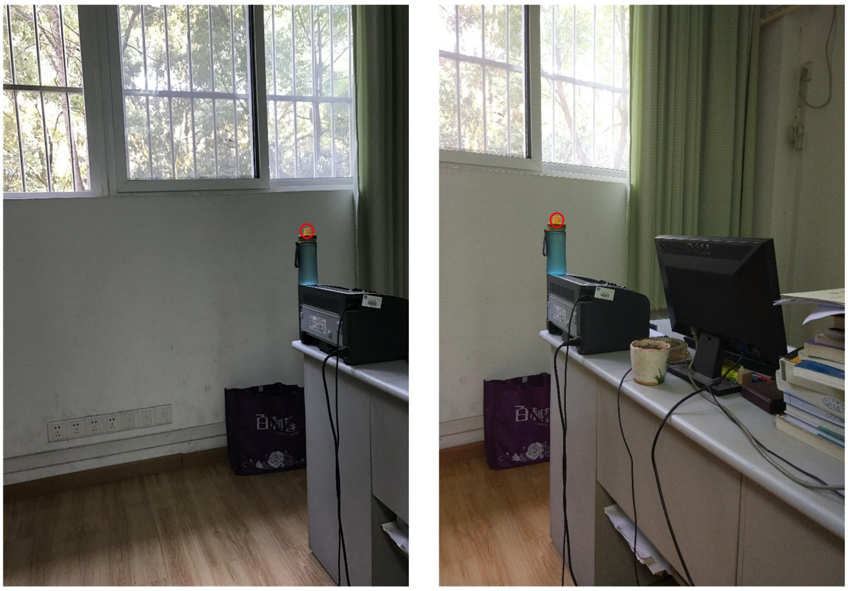
First, calibrate each camera in the binocular system to confirm the camera's distortion coefficient and internal reference matrix. After loading images, the left and right images are processed by corner detection algorithm respectively. Then, the projection matrix, the internal parameter matrix and the external parameter matrix are solved by two-step method. Because the image distortion and parameters in the lens are different in different camera, the selected camera calibration parameters need to be confirmed. After obtaining the parameters of a single camera, a known world coordinate system (usually with high accuracy binocular calibration plate) is a reference. Binocular cameras take the calibration plate from different angles. Due to different angles, two cameras will inevitably produce two different images when shooting the same object. Calculate the image coordinates of the two images respectively and then the coordinates of the actual calibration board can be matched to the current position of the camera parameters.

$$\frac{T_x-\left(x^l-x^r\right)}{Z-f}=\frac{T_x}{Z}\ \Rightarrow\ Z=\frac{fT_x}{x^l-x^r} \tag{8}$$

Assuming that the coordinates of the target point in the left view are (x, y), the parallax formed in the left and right views is d. And the coordinates of the target point in the world coordinate system are (x, y, z) when the left camera center as the origin. There is a transformation matrix Q, that satisfies $Q\cdot[x,y,d,1]'=[X,Y,Z,W]'$. If you also need to obtain the X coordinate and Y coordinate, you also need to know the offset of the left and right image plane coordinate system c_x and the offset of the origin in the three-dimensional coordinate system c_y . Among them, f , T_x , c_x , and c_y can be initialized and optimized by stereoscopic calibration. So that the two cameras are mathematically perfectly parallel. And the c_x , c_y and f of the left and right cameras are the same which is the ideal form of the left and right view completely parallel alignment in Fig. 10.

4.2 Binocular distance measurement

The two large constant cameras are installed on a rack built by aluminum alloy. Because the target object has obvious color difference distinct from the other surrounding objects. The top of the glass is yellow as shown in Fig. 11. The images are color recognition processing that can directly find the target object.



The left camera　　　　The right camera

Fig. 11 the picture of binocular distance measurement

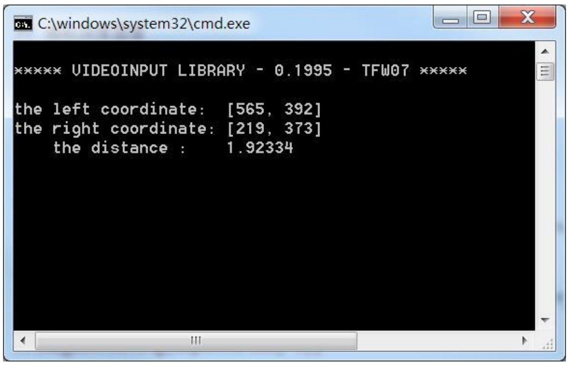


Fig. 12 Result of distance measure

With the extracted feature points in the left and right camera field *X*, *Y* coordinates, achieve the same name point match (the same name points: objects on the same feature points in the left and right camera point) . The *X*, *Y* and *Z* coordinates of these feature points can be calculated by rotating the matrix parameters according to the previously calibrated translation vectors. As shown in figure 12, the distance from the coordinates of the badminton is 1.92334 *m*.

5. Experiments

A complete binocular camera was brought to the badminton stadium for experiment. The binocular vision had been calibrated in the previous section without repeating calibration. The camera was placed on the side of the badminton field to monitor the dynamic badminton real-timely .

The badminton recognition software was written based on MFC as shown in Figure 13. In the software, the left and right camera pictures display in real time. The badminton that is monitored in the picture will be marked with a red circle. At this moment the distance from the coordinates of the badminton is 2704.09 *mm*.

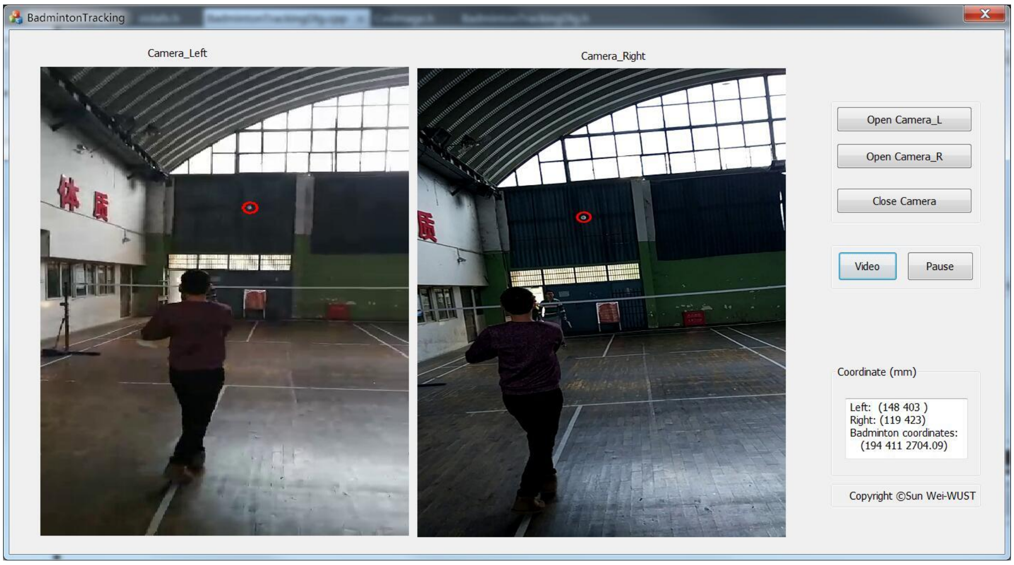


Fig. 13 the picture of Badminton trajectory identification software

6. Conclusion

In this paper, a target tracking method based on hybrid Gaussian background difference (HGBD) and multiple characteristics is proposed. HGBD is effective for the moving object detection. And multiple characteristics include contour, color and motion features are used to contour filtering. The proposed original tracking algorithm can achieve better results and the tracking results are more accurate. The three-dimensional coordinate of badminton is extracted. Meanwhile, the badminton speed test and trajectory prediction can be resolved.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grant (No. 51875418).
The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

References

[1] X. Sun, J. Zhang, Z. Xie, J. Gao, L. Wang, P. Heidingsfelder, Active-matting-based object tracking with color cues, Signal Image & Video Processing, 8 (2014) 85-94.
[2] B. Zhong, S. Pan, C. Wang, T. Wang, J. Du, D. Chen, L. Cao, Robust Individual-Cell/Object Tracking via PCANet Deep Network in Biomedicine and Computer Vision, BioMed research international, 2016 (2016) 1-15.
[3] P. Guo, Clarifying human-computer interaction, Communications of the Acm, 57 (2014) 10-11.
[4] M. Meißner, J. Pfeiffer, T. Pfeiffer, H. Oppewal, M. Meißner, J. Pfeiffer, T. Pfeiffer, H. Oppewal, Combining virtual reality and mobile eye tracking to provide a naturalistic experimental environment for shopper research, Journal of Business Research, (2018).
[5] M. Azizian, M. Khoshnam, N. Najmaei, R.V. Patel, Visual servoing in medical robotics: a survey. Part I: endoscopic and direct vision imaging - techniques and applications, International Journal of Medical Robotics + Computer Assisted Surgery Mrcas, 10 (2013) 263-274.
[6] A.Y. Yang, S. Iyengar, P. Kuryloski, R. Jafari, Distributed segmentation and classification of human actions using a wearable motion sensor network, Computer Vision and Pattern

Recognition Workshops, 2008. CVPRW '08. IEEE Computer Society Conference on, 2008, pp. 1-8.
[7] A. Jalal, S. Kamal, Real-time life logging via a depth silhouette-based human activity recognition system for smart home services, IEEE International Conference on Advanced Video and Signal Based Surveillance, 2014, pp. 74-80.
[8] Y. Song, J. Tang, F. Liu, S. Yan, Body Surface Context: A New Robust Feature for Action Recognition From Depth Videos, IEEE Transactions on Circuits & Systems for Video Technology, 24 (2014) 952-964.
[9] A. Jalal, S. Kamal, D. Kim, A depth video sensor-based life-logging human activity recognition system for elderly care in smart indoor environments, Sensors, 14 (2014) 11735-11759.
[10] F. Farooq, J. Ahmed, L. Zheng, Facial expression recognition using hybrid features and self-organizing maps, IEEE International Conference on Multimedia and Expo, 2017, pp. 409-414.
[11] A. Jalal, S. Kim, Global Security Using Human Face Understanding under Vision Ubiquitous Architecture System, Enformatika, (2011).
[12] A. Jalal, IjazUddin, Security Architecture for Third Generation (3G) using GMHS Cellular Network, International Conference on Emerging Technologies, 2008, pp. 74-79.
[13] A. Jalal, S. Kim, Advanced Performance Achievement using Multi-Algorithmic Approach of Video Transcoder for Low Bitrate Wireless Communication, 5 (2004).
[14] X. Lan, A. Ma, P.C. Yuen, R. Chellappa, Joint Sparse Representation and Robust Feature-Level Fusion for Multi-Cue Visual Tracking, IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 24 (2015) 5826.
[15] X. Lan, S. Zhang, P.C. Yuen, R. Chellappa, Learning Common and Feature-Specific Patterns: A Novel Multiple-Sparse-Representation-Based Tracker, IEEE Transactions on Image Processing, 27 (2018) 2022-2037.
[16] I.T. Comlekciler, S. Gunes, C. Irgin, Artificial 3-D contactless measurement in orthognathic surgery with binocular stereo vision, Elsevier Science Publishers B. V., 2016.
[17] Q. Ye, J. Dong, Y. Zhang, 3D Human Behavior Recognition based on Binocular Vision and Face-hand Feature, Optik - International Journal for Light and Electron Optics, 126 (2015) 4712-4717.
[18] C.H. Chen, C.C. Wang, M.C. Yan, Robust tracking of multiple persons in real-time video, Multimedia Tools & Applications, 75 (2016) 1-15.
[19] T. Gao, G. Li, S. Lian, J. Zhang, Tracking video objects with feature points based particle filtering, Multimedia Tools & Applications, 58 (2012) 1-21.
[20] X. Sun, H. Yao, X. Lu, Dynamic multi-cue tracking using particle filter, Signal Image & Video Processing, 8 (2014) 95-101.
[21] F. Li, S. Liu, Object tracking via a cooperative appearance model, Knowledge-Based Systems, 129 (2017) 61-78.
[22] Y. Zha, Y. Yang, D. Bi, Graph-based transductive learning for robust visual tracking, Pattern Recognition, 43 (2010) 187-196.
[23] S. Tian, F. Yuan, G.S. Xia, Multi-object tracking with inter-feedback between detection and tracking, Elsevier Science Publishers B. V., 2016.
[24] Y.H. Jang, J.K. Suh, K.J. Kim, Y.J. Choi, Robust Target Model Update for Mean-shift

Tracking with Background Weighted Histogram, Ksii Transactions on Internet & Information Systems, 10 (2016) 181-207.
[25] A. Li, S. Yan, Object Tracking With Only Background Cues, Circuits & Systems for Video Technology IEEE Transactions on, 24 (2014) 1911-1919.
[26] M. Yang, Y. Wu, G. Hua, Context-aware visual tracking, IEEE Trans. PAMI, 2008, pp. 1195-1209.
[27] J.W. Liu, W.P. Sun, T. Xia, Adaptive structured sub-blocks tracking, Elsevier Science Publishers B. V., 2016.
[28] H. Jiang, J. Li, D. Wang, H. Lu, Multi-feature tracking via adaptive weights, Neurocomputing, 207 (2016) 189-201.
[29] M. Lucena, J.M. Fuertes, N.P.D.L. Blanca, Optical flow-based observation models for particle filter tracking, Pattern Analysis & Applications, 18 (2015) 135-143.
[30] Y. Gao, R. Ji, L. Zhang, A. Hauptmann, Symbiotic Tracker Ensemble Toward A Unified Tracking Framework, IEEE Transactions on Circuits & Systems for Video Technology, 24 (2014) 1122-1131.
[31] M. Liu, N. Liang, Detection of Moving Target Using Improved Optical Flow Method, Software Engineering, 2013, pp. 311-315.