Journal of Computers Vol. 34 No. 6, December 2023, pp. 107-120
doi: 10.53106/199115992023123406008

Moving Target Tracking Method Based on Improved Camshift

Qi-Jun Luo*, Zheng Li, Xin Tian, Hong-Ying Zhang

College of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China
{qjluo, 2020021050, 2020022205, hyzhang}@cauc.edu.cn

Received 30 June 2022; Revised 1 July 2022; Accepted 1 July 2022

Abstract. Aiming at the problem that target occlusion and other disturbances in complex background will reduce the tracking accuracy of moving target, and even lead to tracking failure, this paper proposes a moving target tracking algorithm based on the improved Camshift algorithm. Firstly, Gaussian background is used to model the foreground image to improve the backprojection image, and then the interference of backprojection is removed to improve the tracking effect in complex background conditions. Secondly, Kalman filtering is utilized to predict the trajectory, which further improves the tracking accuracy of Camshift algorithm in occlusion condition. A lot of experiments are processed, and the results show that the proposed algorithm could effectively improve the tracking accuracy and meet the real-time requirements.

Keywords: Camshift, Kalman filtering, target occlusion, Gaussian background

1 Introduction

Moving target tracking is an important research field in computer vision. It can be applied to many fields, including security, medical treatment [1-2], traffic, etc. At present, there are four widely used moving target tracking methods: feature-based tracking method, model-based tracking method, contour based tracking method and region based tracking method. Camshift algorithm is a feature-based tracking method, which uses the color information of the target to track. It has the advantages of small amount of calculation and good robustness under simple background [3]. The traditional Camshift algorithm needs to manually select the tracking target before tracking. The tracking effect of the moving target is good when the background is simple, but the tracking effect of the algorithm will decline and even fail when the background is complex (such as the background color is similar to the target color, there is occlusion, the tracked target moves too fast, etc.). To solve this problem, researchers have proposed various improvement methods. In 2016, Tanyan and Wangyujun [4] proposed an improved Camshift target tracking method combined with background difference. The algorithm can effectively solve the problem that the target color is similar to the background color, but the algorithm is complex and difficult to apply in practice. Tangjianfei [5] adopts the method of multi feature fusion to correct the central position of the Camshift algorithm search box by predicting the spatial position information of the target, so as to alleviate the background interference. Wei Sun [6] established a tracking template based on the three-dimensional joint histogram of the color, saturation and LBP texture features of the tracking target, and adopted an adaptive adjustment strategy to mitigate the interference of background information and improve the tracking accuracy of the algorithm. Lijianliang [7] introduced Kalman filter to predict the position of the next frame on the basis of three-dimensional histogram, which improved the robustness of the algorithm in complex background. The staple algorithm [8] proposed by bertinetto et al. Combination the discriminant scale space tracker with color histogram tracking to improves the adaptability of the algorithm to object deformation, but performs poorly in scenes with changing illumination. Gao et al. [9] proposed an updating rhythm framework based on integrated post-processing strategy to reduce the model drift of the discrimination tracker. The framework initializes a group of trackers to update the model at different intervals, and selects the tracker with the smallest deviation score as the reliable tracker for subsequent tracking. However, along with the increasing number of trackers, MTS framework performs poorly in real-time tracking. M. Nallasivam et al. [10] proposed a method with combining frame subtraction and background subtraction to support the Camshift algorithm in finding the exact target position. But the tracking accuracy of this algorithm is not enough.

In order to further mitigate the effect of complex background on the tracking effect of Camshift algorithm, this paper proposes a motion target tracking algorithm based on improved Camshift algorithm. Firstly, the Gaussian

* Corresponding Author

background is used to model the foreground image to improve the back-projection map so as to remove the interference term of the back-projection to obtain good tracking effect in the complex background condition; Secondly, the trajectory prediction is combined with Kalman filtering [11-12] to further increase the tracking accuracy of the Camshift algorithm in the occlusion condition.

2 Target Tracking based on Improved Camshift

Camshift algorithm is an improved MeanShift. Its basic idea is to perform MeanShift operation on a series of image frames in video based on color information and use the results obtained from the previous frame as the initial value of the MeanShift algorithm in the next frame. The advantage of this algorithm is that the size of the search window could be adaptively adjusted according to target size. When the size and shape of the moving target in the video change, it could also be tracked. In particularly, Camshift algorithm has good robustness when the color of tracked object does not change much during the motion. The specific steps are as follows:

Step 1: Calculating the color histogram. Typically, the input image is first converted from RGB to HSV color space, and then the tracked object is manually selected. The size of the search box is initially set to the target area and the H-channel histogram is calculated for the HSV space of the selected area.

Step 2: Back projection map generation. Based on the color histogram obtained in the first step, the original image is converted into a color probability distribution image, a process called back-projection.

Step 3: Meanshift iterative process. It is the core part of the whole Camshift algorithm, and its role is to find the position of the target center in the current frame.

Assuming that the pixel point (i, j) is located in the search box, the value corresponding to this pixel point in the reverse projection diagram can be expressed as $I(i, j)$, defining the zero-order moment M_{00} and the first-order moment M_{10} and M_{01} of the search box ,and as shown in Eq. (1), Eqs. (2) and (3).

$$M_{00} = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} I(i, j). \tag{1}$$

$$M_{10} = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} iI(i, j). \tag{2}$$

$$M_{01} = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} iI(i, j). \tag{3}$$

Finally, through calculation, the centroid position of the search box is: $(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}})$, where the zero-order matrix reflects the size of the search box. The Meanshift iteration process adjusts the size of the search box according to the position of the center of mass and passes the center of the search box to the center of mass. If the distance of the motion is higher than the set threshold, the center of mass of the processed frame is calculated again and the position and size of the frame is further adjusted until the distance between the center and the center of mass does not exceed the threshold, or when the iteration terminates and the convergence condition is satisfied. The position and size of the search box are used as the input of the next frame, and then a new target search task is executed.

The traditional Camshift algorithm processes the binary image formed by the histogram back projection. If the background is more complex or the surface texture of the tracked object is richer, there is a lot of noise in the binary image formed by the back projection, which directly interferes with the determination of the object position by the Camshift algorithm. Therefore, the algorithm has a poor tracking results in complex background or the object rich in texture.

A surveillance video is selected for the experiment. This video is recorded by the terminal of the surveillance equipment. The pixel space of the moving target accounts for a small proportion. Compared with the background, the color characteristics of the moving target are unclear and the background is complex. It is inevitable that some background information will be included in the frame selection of the moving target. Therefore, in the back projection image, the tracked object will be covered by the background. The object being tracked in this video is the pedestrian on the inner side of the road near the house in this video, and the experimental results are shown in Fig. 1. From Fig. 1(b) and Fig. 1(c), it can be seen that the traditional Camshift algorithm fails.

(a) 100th frame (b) 210th frame (c) 360th frame

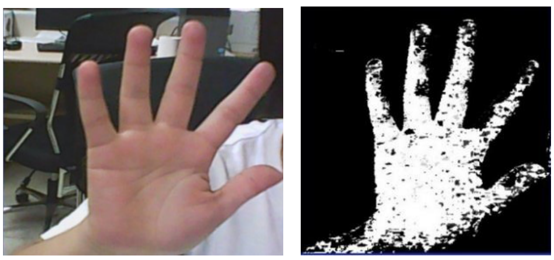
Fig. 1. Tracing effect diagram of traditional Camshift algorithm

2.1 Back Projection Graph Generation Method based on Gaussian Mixture

The back projection process in the Camshift algorithm is the process of converting the original image into a color probability distribution image based on the color histogram of the image. The specific process is to assume that the image is, the image size is, and is the value of pixel hue (Hue) of coordinate. The histogram is set to have different binvalues, and the number of pixels with different bin values is denoted by by $h(H_1)$, $h(H_2)$, ..., $h(H_k)$, respectively, and satisfies Eq. (4).

$$h(Hk) = \begin{cases} \sum_{i=1}^{m}\sum_{j=1}^{n} 1 & if \quad I(i,j) = H_k \\ 0, & else \end{cases} . \tag{4}$$

Where $h(H_k)$ is the number of pixels whose color value is H_k, which is the vertical coordinate of the histogram, and bin is the horizontal coordinate of the H_k histogram. Next, we query the value $h(H_k)$ corresponding to bin as H_k in hue histogram, and finally display the above value in the new image. Assuming that the range of bin in the Hue histogram is set to [0-255], the single channel back projection image can be directly displayed, as shown in Fig. 2.



(a) Input image (b) Output image

Fig. 2. Back projection process demonstration diagram

The function of the back projection map is to search for the area closest to the target template in the input image. Generally, during Camshift tracking, the selected target template may contain some background information if the background is complex. Therefore, in the back projection map, the tracked object will be covered by the background, which makes tracking difficult. So it is very necessary to model the background in a complex background. The modeling process is to train the background first, simulate the background in each frame using a model, and once the background is extracted, check whether the pixels match the background model. Background modeling methods include methods such as weighted average method and hybrid Gaussian background modeling method, among which hybrid Gaussian background modeling [13] is a more mature background modeling method, so this paper uses superimposed hybrid Gaussian background modeling foreground image to improve the back projection map, and then improve the tracking accuracy of Camshift algorithm under complex background interference.

2.2 Method

In this paper, we first introduce a hybrid Gaussian background model and use normal distribution curves to accurately quantify things, decomposing things into several models based on normal distribution curves [14]. Compared with Liang's [15] moment invariant target recognition algorithm, its basic idea is to search all possible target regions in the preprocessed binary image, calculate the seven Hu moment invariant features of the region, and regard the region with high matching degree with the template as the same type of target. The algorithm is difficult to distinguish targets in complex background. The hybrid Gaussian modeling uses the weighted average of three to five normal probability density functions to characterize the distribution of each pixel in the image in the time domain. It gives a good indication of the multi-peaked state of the pixels. The hybrid Gaussian background model is shown in Eq. (5).

$$P(I(x,y,t)) = \sum_{i=1}^{k} w_i^t \eta_i (I(t), \mu_i^t, \sigma_i^t). \tag{5}$$

Where $I(x, y, t)$ denotes the pixel value of pixel $I(x, y)$ at time t, K is the number of Gaussian distributions, w_i^t is the weight of the $i-th$ Gaussian component at time t, and η is the normal probability density function. The K Gaussian distributions are sorted according to the magnitude of w/σ, and then the B (defined as equation 6 for B) Gaussian distribution with the larger value in front is selected as the background of the target, and then the $k-B$, Gaussian distributions are eventually replaced by the new Gaussian distribution. The mathematical expression is shown in Eq. (6) below.

$$B = \arg\min_b \left\{ \sum_{k=1}^{b} w_k > T \right\}. \tag{6}$$

Where T is the prior probability that the pixel being processed is classified as the background. The size of T determines the number of Gaussian distributions. The larger the value of T, the more the number of Gaussian distributions. However, the amount of calculation will increase accordingly, so T is generally taken as about 0.8. After the target pixels are processed using the Gaussian mixture background model, the reverse projection map is optimized. The optimization process is shown in Fig. 3.

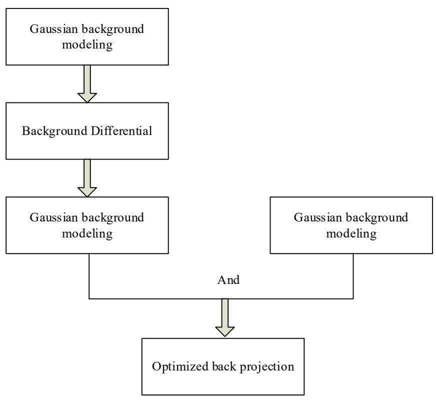


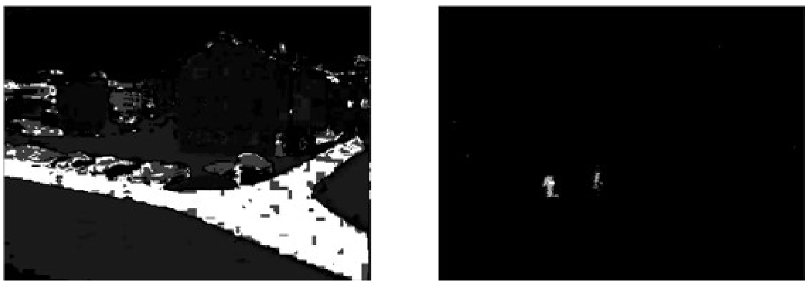
Fig. 3. Optimization process diagram of reverse projection diagram

The inherent drawback of the Camshift algorithm is that when the tracked object is framed, some background pixels around the object are included [16]. These pixels involved in the histogram statistics will definitely affect the reverse projection image of the target, and this can be done by adding a mechanism to determine whether the reverse projection image is good or bad. When the average gray value of the reverse projection image is out of the set range, Gaussian background modeling foreground image is introduced to operate with the reverse projection map to remove the interference term of the reverse projection, and only when the pixel value does not match all distributions of the hybrid Gaussian model and the H component of the pixel value is within the pre-set threshold, the pixel value is considered to belong to the tracked object and its pixel value is set to 1, otherwise it is a background pixel The mathematical equation is described in the following Eq.(7).

$$(G-H)_{i,j}^{k}=\begin{cases}1 & G_{i,j}^{k}=1\&\&H_{i,j}^{k}=1\\ 0 & else\end{cases} \tag{7}$$

Where $(G-H)_{i,j}^{k}$ is the value of the pixel point at location $p(i, j)$ of the optimized image. $G_{i,j}^{k}$ and $H_{i,j}^{k}$ are the Gaussian background modeling foreground image and the back projection image respectively. Finally, the Gaussian background modeling foreground image is operated with the back projection map to remove the interference terms of the back projection to compensate for the undetected motion target. Fig. 4 shows a frame from a surveillance video, which is reverse projected. The original back projection map is shown in Fig. 5(a), and the back projection map optimized by the algorithm in this paper is shown in Fig. 5(b).

Fig. 4. Input image



(a) Back projection diagram before optimization (b) Optimized back projection map

Fig. 5. Optimization process diagram of reverse projection diagram

From Fig. 4, it can be seen that the background of the original image is complex, and the backgrounds of vehicles, houses, and roads interfere a lot with the extraction of motion targets. From Fig. 5(a) and Fig. 5(b), it can be seen that the average gray value of the back projection map before optimization is large. Extracting the tracked object is difficult, and the background is removed after optimization, leaving only the foreground. Because the motion target is manually framed, there is no motion target detection previously.

3 Target Tracking based on Improved Camshift

The algorithm in this paper optimizes the reverse projection map, which improves the tracking accuracy of the algorithm in this paper in complex backgrounds. However, in general surveillance video, the moving objects will be obscured from each other or the moving objects will be obscured due to the cumbersome motion information. If the tracked object is occluded in the current frame, the Camshift algorithm cannot obtain the center-of-mass position of the object at this time. Thus, the search window for the next frame cannot be initialized, resulting in tracking failure. The reference [19] detects and analyzes targets through the improved Camshift tracking algorithm, and detects moving objects through mixed Gaussian background detection for searching and tracking targets. However, this algorithm is not good for detecting the occlusion of the target. Therefore, it is necessary to predict the motion trajectory of the tracked object and consider the predicted position as the target position obtained by the Camshift algorithm in order to continue tracking. Given that the Kalman filtering algorithm has good stability, is easy to initialize, and can accurately predict the real position of the target after multiple filtering. Therefore, this paper combines the Kalman filter for predicting the position information of the tracked object to improve the occlusion processing capability by reasonably predicting the motion trend of the tracked object.

In motion target tracking, the ideal state is to use the measurements to estimate the motion state of the tracked target, so the accumulation of a certain number of measurements allows us to obtain the trajectory of which is not affected by noise. This process of estimating the motion state can be divided into two parts: the first part is the prediction phase, where the model is obtained empirically and where the motion target is expected to appear at the next moment. The Kalman filter is the most widely used and common predictor, which is suitable for single-target tracking and has a small number of operations. The mathematical model of the Kalman filter is as follows:

Equation of state:

$$X_k = \Phi_{k,k-1} X_{k-1} + \Gamma_{k,k-1} w_{k-1}. \tag{8}$$

Observation equation:

$$Z_k = H_k X_k + v_k, \tag{9}$$

where $\Phi_{k,k-1}$ is the system state transfer matrix, $\Gamma_{k,k-1}$ is the noise input matrix, H_k is the observation matrix, X_k, X_{k-1} are the system state vectors at moments k and $k-1$, Z_k is the system observation vector at moment k, w_k is the system state noise vector, v_k is the system observation noise vector, and both w_k and v_k are considered to be mutually independent normal white noise. Firstly, the model of Kalman filtering is introduced, and the method used is as follows: let T be the time interval of sampling, then the state transfer matrix can be expressed as shown in Eq. (10):

$$\Phi_{k,k-1} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{10}$$

Noise input matrix:

$$\Gamma_{k,k-1} = \left[T^2/2 \quad T^2/2 \quad T \quad T \right]^T . \tag{11}$$

Observation matrix:

$$H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{12}$$

Substitute the above noise input matrix and observation matrix into Eq. (8) and Eq. (9) to obtain the best estimate of X_k , and then take the central position output by Camshift algorithm as the observation value. And use Z_k to correct the predicted value and finally iterate. The tracking block diagram combined with Kalman is shown in Fig. 6:

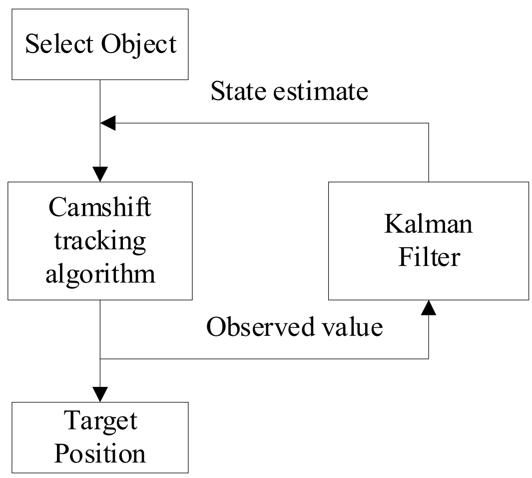


Fig. 6. Tracking block diagram combined with Kalman

Combining Kalman's tracking block diagram the tracking process can be divided into two parts: Firstly, the Kalman filter is used to predict the position of the tracked object in the current image based on the observations of the previous frame. Then, based on the results of the Camshift algorithm, the target that best matches the template of the tracked object is found in the expected region, and finally, the Kalman filter state is updated. And so on to anticipate the position of the target in the next image frame again. For static occlusion, a mechanism is added to determine whether the target is lost or not, that is, to determine whether the tracked object is completely occluded. Judge whether the tracking is lost according to the size of the prediction selection area and the distance between the Kalman prediction point and the actual point. If the tracking is lost, take the Kalman prediction point as the center of the lost area and search until the target reappears in the picture and continue tracking. Based on the above analysis, the overall flow chart of the improved algorithm is shown in Fig. 7.

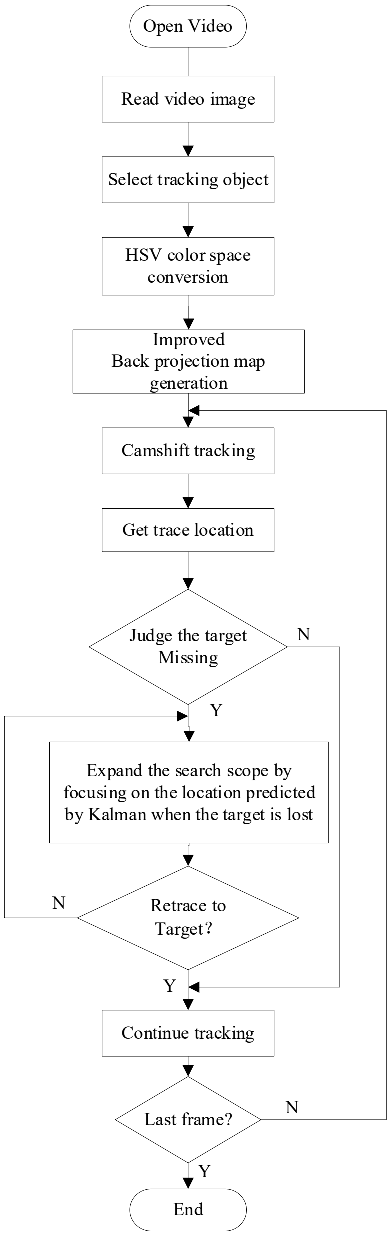


Fig. 7. System tracking flow chart

4 Experimental Results and Analysis

The hardware configuration of the computer used for the algorithms in this paper: Win8, CPU Intel(R) CORE(TM) i5-5200U, 4 GB of RAM, 500 GB of hard disk. the software involved in this experiment mainly includes Microsoft Visual Studio 2013 and opencv computer open source visual library version 2.4.9 installed on a personal computer. In order to verify the tracking effect of the improved algorithm in this paper, this paper selects four videos [17]. The background color of the first and second videos is similar to the tracked object and the background is complex. The third video has dynamic occlusion and the fourth video has static occlusion. The video parameters are shown in Table 1.

Table 1. Table of selected video parameters

Video sequence	Image size	Frame rate (fps)	Video length (Frame)
Video1	768x576	25	412
Video2	352x288	25	175
Video3	541x310	30	242
Video4	541x310	25	325

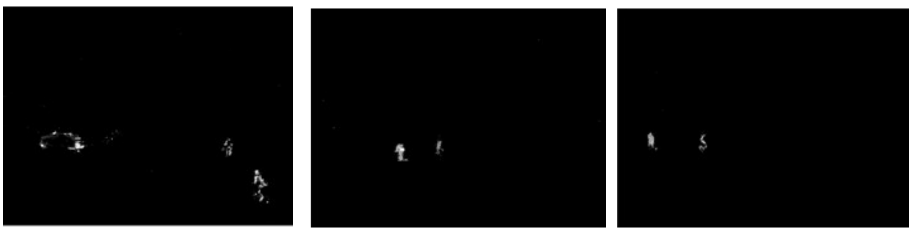
In this paper, we choose the traditional Camshift algorithm and the algorithm of this paper to conduct comparison experiments on the above four video segments. video1 is a surveillance video. The selected tracking object is small, and the pixel space of the moving object accounts for a small proportion. Compared with the complex background, the color characteristics of the tracked object are not obvious. The tracking results of the traditional Camshift algorithm and the algorithm in this paper for pedestrians are shown in Fig. 8 and Fig. 9, respectively.

(a) 100th frame (b) 210th frame (c) 360th frame

(d) 100th frame back projection (e) 210th frame back projection (f) 360th frame back projection

Fig. 8. Traditional Camshift algorithm tracking results

(a) 100th frame (b) 210th frame (c) 360th frame



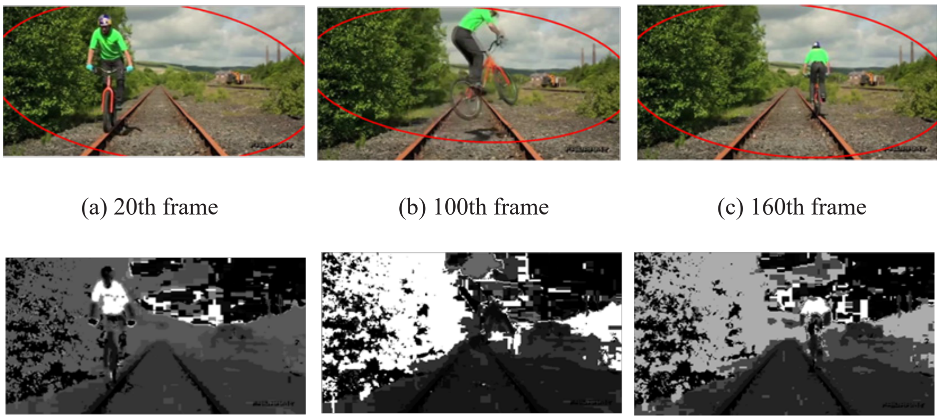
(d) 100th frame back projection (e) 210th frame back projection (f) 360th frame back projection

Fig. 9. Algorithm tracking results in this paper

Fig. 8(a) to Fig. 8(c) show the tracking result plots of the conventional Camshift algorithm. Fig. 8(d) to Fig. 8(f) show their reverse projection plots in order from left to right. Fig. 9(a) to Fig. 9(c) show the tracking results of the improved Camshift algorithm from left to right. Fig. 9(d) to Fig. 9(f) show their improved back-projection maps.

The color features of the tracked pedestrians in Video1 are not obvious, and the background of the video is complicated. From the tracking results of the two algorithms for the pedestrians in Video1, the traditional Camshift algorithm can not track the pedestrians. The reason is that the traditional Camshift algorithm fails to track the pedestrians because the back-projection effect in the complex background environment is not good enough to effectively extract the information of the tracked objects. From Fig. 9(a) to Fig. 9(c), we can see that the improved algorithm in this paper can track well from the beginning to the end, and the tracking effect is good. In summary, the improved back-projection map is pivotal to the tracking results of the Camshift algorithm.

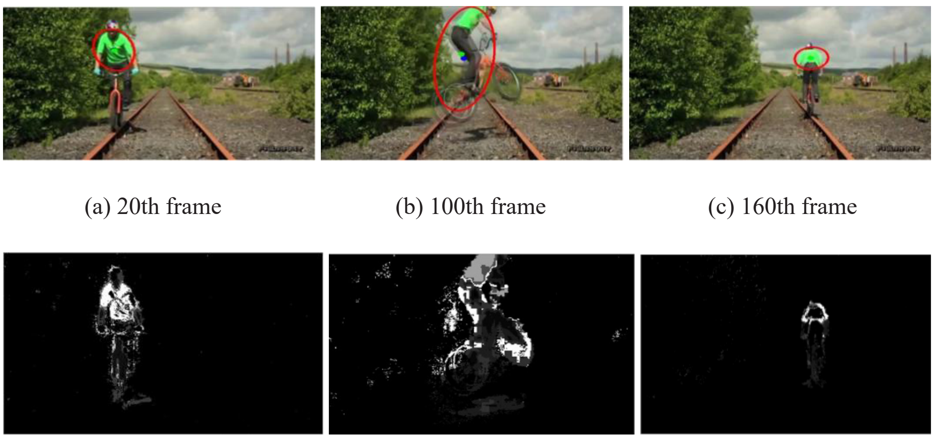
Video 2 is a video taken by a fixed camera, which is overexposed and has a complex background. The tracking results of the traditional Camshift algorithm and the algorithm in this paper for the target are shown in Fig. 10 and Fig. 11 respectively.



(a) 20th frame (b) 100th frame (c) 160th frame

(d) 20th frame back projection (e) 100th frame back projection (f) 160th frame back projection

Fig. 10. Traditional Camshift algorithm tracking results



(a) 20th frame (b) 100th frame (c) 160th frame

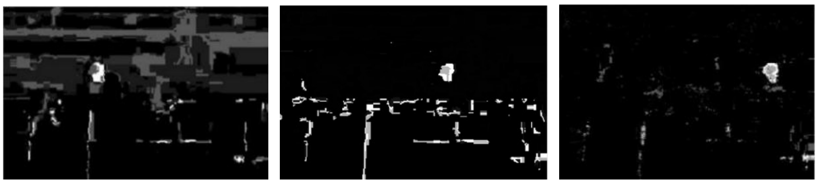
(d) 20th frame back projection (e) 100th frame back projection (f) 160th frame back projection

Fig. 11. Algorithm tracking results of this paper

From Fig. 11(a) to Fig. 11(c), we can see that the algorithm in this paper can still track the tracked object accurately in the complex background environment. However, the traditional Camshift algorithm mistakenly uses the background as the tracking object and the tracking fails, which illustrates the superiority of the improved algorithm in this paper.

Video 3 is a monitoring video near the subway station. The background of the video is simple, the color characteristics of the tracked object are obvious, there are many pedestrians in the video and pedestrians block each other. In order to ensure the accuracy of the experiment, the pedestrian in red in the video is selected as the tracking object. The tracking of pedestrians by the two algorithms are shown in Fig. 12 and Fig. 13 respectively.

(a) Frame 75 original (b) Frame 120 original (c) Frame 201original



(d) Frame 75 back projection (e) Frame 120 back projection (f) Frame 201 back projection

Fig. 12. Traditional Camshift algorithm tracking results

(a) Frame 75 original (b) Frame 130 original (c) Frame 223 original

(d) Frame 75 reverse projection (e) Frame 130 reverse projection (f) Frame 223 reverse projection

Fig. 13. Algorithm tracking results in this paper

It can be seen from the third video that there are more pedestrians and the selected tracked object (the red-clad pedestrian in the video) is obscured by other pedestrians. From Fig. 12(b) to Fig. 12(c), it can be seen that when the tracked object is obscured by other pedestrians, the elliptical box no longer converges; instead, it gradually diverges to other moving targets and backgrounds, which eventually leads to tracking failure. And from Fig. 13(a) to Fig. 13(c), it can be seen that the algorithm in this paper can continue to track after the tracked object is obscured by other pedestrians respectively, and the tracking effect is better.

The interference present in Video 3 above is dynamic occlusion, i.e., the occlusion encountered by the tracked object is dynamic. Another type of occlusion interference is static occlusion, which is caused by static objects (e.g., houses or trees) on the tracked object. In Video4, there is static occlusion where the blob is completely occluded by the box and then appears in the frame again, and the tracking of the blob by the two algorithms is shown in Fig. 14 and Fig. 15, respectively.

(a) 20th frame (b) 164th frame (c) 242th frame

Fig. 14. Traditional Camshift algorithm tracking results

(a) 20th frame (b) 164th frame (c) 242th frame

Fig. 15. Tracking results of the algorithm in this paper

In the fourth video, as shown in Fig. 14(a), the traditional Camshift algorithm can track well before the target disappears. After the target disappears, as shown in Fig. 14(b), the traditional Camshift algorithm mistakenly takes the box as the target, and then takes the hand of the tester in the video as the tracked object as shown in Fig. 14(c), resulting in tracking failure. Fig. 15(a) to Fig. 15(c) show that the algorithm in this paper can track the moving target well, because the algorithm in this paper takes the prediction point of Kalman as the center of the lost area, and expands the search range until the target reappears in the picture, and then can track it immediately.

Through the above experimental results, it can be concluded that this algorithm effectively improves the accuracy of tracking under the occlusion interference of complex background. The average time per frame [18] index is selected as the reference to show the tracking velocity. It is an important indicator of the real-time performance of the detection algorithm and has a great relationship with the selected hardware platform. The results are shown in Table. 2. According to the parameters in Table 1, the average time per frame can be obtained by dividing the total time by the number of frames. The results are shown in Table. 3.

Table 2. Total tracking time (s)

Algorithm	Video1	Video2	Video3	Video4
Our algorithm	45	21	20	31

Table 3. Average time per frame (ms)

Algorithm	Video1	Video2	Video3	Video4
Our algorithm	109	120	82	95

It can be seen from Table. 3, since the algorithm in this paper adds a mechanism to judge the effect of the back projection image, when the selected video background is simple, it does not need to use the Gaussian mixture model. Therefore, for the algorithm in this paper, the average time per frame of video3 is the least. Kalman filtering increases the operation time. It is worthwhile to improve the accuracy of the algorithm at the cost of increasing the average time per frame, and the average processing speed of 12 frames per second on the hardware platform of this experiment can basically meet the real-time requirements. Although reference [19] uses the improved CAMSHIFT target tracking algorithm to track and identify the real trajectory of basketball targets and achieves good accuracy, its robustness is not ideal. However, the algorithm combined with Kalman filter and

CAMSHIFT target tracking adopted in this paper has good robustness, and the accuracy of the algorithm is improved.

5 Conclusion

In order to improve the Camshift tracking algorithm, we add a mechanism to determine whether the reverse projection map is good or bad. If the reverse projection was not good, we remove the interference items of the reverse projection by Gaussian background modeling of the foreground image and the reverse projection map. Combined with Kalman filtering, the tracking accuracy of Camshift algorithm is improved when occlusion. Taking the actual tracking effect and average time per frame as indicators, the traditional Camshift algorithm and the algorithm in this paper are compared. The experimental results show that the proposed algorithm could effectively improve the tracking accuracy and meet the real-time requirements. In the following work, we would research the moving targets in the dynamic scene. The envisioned method is to estimate the global motion by establishing a multi-parameter affine model to achieve more complex tracking.

Acknowledgement

This research was funded by the Chinese Fundamental Research Funds for the Central Universities (3122018D002).

References

[1] A. Khamparia, S. Bharati, P. Podder, D. Gupta, A. Khanna, T.K. Phung, D.N.H. Thanh, Diagnosis of Breast Cancer Based on Modern Mammography Using Hybrid Transfer Learning, Multidimensional Systems and Signal Processing 32(2)(2021) 747-765.

[2] S. Bharati, P. Podde, M. Mondal, V. B. Prasath, Co-Resnet: Optimized Resnet Model for Covid-19 Diagnosis from X-Ray Images, International Journal of Hybrid Intelligent Systems 17(1-2)(2021) 71-85.

[3] X.-B. Li, G.-Q. Xu, X.-Y. Yang, G.-Y. Zhao, A Multi Feature Tracking Algorithm Based on Camshift, Computer and Digital Engineering 48(1)(2020) 73-77.

[4] Y. Tan, Y.-J. Wang, On a Combination of Background Difference Improved Camshift Target Tracking Method, Journal of Southwest China Normal University (Natural Science Edition) 41(9)(2016) 120-125.

[5] J.-F. Tang, X.-S. Zhuo, X.-C. Qi, W. Wang, Research on Target Tracking of Monitoring Rats Invasion Around Substation, Automation and Instrumentation 37(1)(2022) 60-65+80.

[6] S.-Y. Shen, Z.-X. Li, W. Sun, Research on Target Tracking of UAV Based on Improved Camshift Algorithm, Microelectronics and Computer 36(11)(2019) 76-83.

[7] J.-L. Li, T.-T. Zhang, Z.-F. Tao, S.-Q. Li, Q.-R. Guo, Pilot Vehicle Tracking Algorithm Based on Improved Camshift and Kalman Filter Fusion, Journal of Electronic Measurement and Instrumentation 35(6)(2021) 131-139.

[8] L. Bertinetto, J. Valmadre, S. Golodetz, O. Miksik, P.H.S. Torr, Staple: Complementary Learners for Real-Time Tracking, In: Proc. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[9] Y.-F. Gao, Z.-X. Hu, H.W.F. Yeung, Y.Y. Chung, X.-H. Tian, L. Lin, Unifying Temporal Context and Multi-Feature With Update-Pacing Framework for Visual Tracking, IEEE Transactions on Circuits and Systems for Video Technology 30(4)(2020) 1078-1091

[10] M. Nallasivam, V. Senniappan, Moving Human Target Detection and Tracking In Video Frames, Studies In Informatics and Control 30(1)(2021) 119-129.

[11] Y.-R. Shao, Fault Location Method of Satellite Network Node Based on Kalman Filter and APNN, Journal of Ordnance Equipment Engineering 43(2)(2022) 191-196.

[12] H. Li, J. Zhu, Target Tracking Algorithm Based on Meanshift and Kalman Filter, Journal of Beijing Institute of Technology 28(2)(2019) 365-370.

[13] Y. Jung, J. Park, Scalable Inference for Hybrid Bayesian Hidden Markov Model Using Gaussian Process Emission, Journal of Computational and Graphical Statistics 31(3)(2022) 666-683.

[14] X.-Y. Wang, L.-G. Wang, J.-W. Wang, K.-P. Sun, Q.-M. Wang, Hyperspectral Anomaly Detection Via Background Purification and Spatial Difference Enhancement, IEEE Geoscience and Remote Sensing Letters 19(2022) 1-5.

[15] S. Liang, Y. Li, Using Camshift and Kalman Algorithm to Trajectory Characteristic Matching of Basketball Players, Complexity 2021(2021) 4728814

[16] T. Guo, F.-L. Luo, L.-Y. Fang, B. Zhang, Meta-Pixel-Driven Embeddable Discriminative Target and Background Dictionary Pair Learning for Hyperspectral Target Detection, Remote Sensing 14(3)(2022) 481.
[17] M. Kiran, L.T. Nguyen-Meidine, R. Sahay, R.M.O.E. Cruz, L.-A. Blais-Morin, E. Granger, Generative Target Update for Adaptive Siamese Tracking, in: Proc. International Conference on Pattern Recognition and Artificial Intelligence, ICPRAI 2022, 2022.
[18] Y.-F. Liu, C.-H. Shen, C.-Q. Yu, J.-D. Wang, Efficient Video Segmentation Models with Per-frame Inference. <https://arxiv.org/abs/2202.12427>, 2022 (accessed 02.05.2023).
[19] N. Liu, P. Liu, Goaling Recognition Based on Intelligent Analysis of Real-Time Basketball Image of Internet of Things, The Journal of Supercomputing 78(1)(2022) 123-143.