Moving Object Tracking Based on Camshift Algorithm

Md Shaiful Islam Babu, KH Shaikh Ahmed

***Abstract*—**Continuously adaptive Camshift is an efficient and lightweight tracking algorithm developed based on mean-shift. Camshift algorithm has the advantage of better real-time, but this algorithm is only suitable for tracking targets in simple cases, not well for tracking desired targets in complex situation. In this paper, we will present an improved method of multiple targets tracking algorithm based on the Camshift algorithm combined with Kalman filter. The tracker of the improved method was used to track each detected target. It can achieve tracking of multiple targets. A large number of experiments have proved that this algorithm has strong target recognition ability, good anti-noise performance, and fast-tracking speed.

***Keywords*—**Camshift, target tracking, Kalman, adaptive, mean-shift.

I. INTRODUCTION

Object tracking is an important part in the field of computer vision. It is widely used in the field of intelligent visual monitoring and video retrieval. Motion video is effected by light and the dynamic change of target, it is very difficult to detect and track the target. Especially for the simultaneous tracking of multiple moving targets, the occlusion, fusion and separation between the targets make the tracking become extremely difficult. Real-time performance and low computation requirements are two key features of a practical tracking algorithm in real-world applications [1]. The proposed design of multi-objectives practical filter tracking algorithm based on Meanshift, however this algorithm has not ability to resist occlusion and doesn't have adaptive function [2]. Another author proposed to use multiple camera constriction systems to deal with the problem of mutual occlusion of multiple moving targets. However, this system is expensive and the algorithm is complex. Real-time performance is not good [3].

Mean-shift is a kernel-based tracking method which uses density-based appearance models to represent targets. The method tracks targets by finding the most similar distribution pattern in a frame sequences with its sample pattern by iterative searching. It has been widely used because of its relative simplicity and low computational cost, but mean-shift would fail in changing the track window`s scale, as targets move toward or away from the camera.

After studying the color characteristic and motion laws of multi moving targets, the background difference method is used to establish the dynamic background model, and the current frame is compared with the background frame to obtain the moving target. After kalman filtering method is used to predict the initial position of each target in the next frame. Camshift is called to successively approximate the precise position of each target, and the better result are obtained through practical application.

Md Shaiful Islam Babu is with the department of Computer Science and Technology, Changchun University of Science and Technology, Changchun 130022 China.

He is with the department of Computer Science and Engineering, City University, Dhaka, Bangladesh (e-mail: mdshaifulislambabu@gmail.com).

KH Shaikh Ahmed is with the department of Computer Science and Technology, Shanghai Jiaotong University, Shanghai, China (e-mail: 158742081@qq.com).

II. DYNAMIC BACKGROUND MODEL

The background model is established by using the background and inter frame subtraction method. The motion area is found through the image of the difference current frame, and the background in motion area doesn't changed. The movement of background is updated with the current frame. After a certain number of iterations the background can be extracted. The steps are as follows:

(1) Take the first frame image I_0 of the sequence image as the initial background image.

(2) Calculate the inter-frame difference binary image BW_k of the current frame k:

$$BW_k = \begin{cases} 0 & |I_k - I_{k-1}| \leq T_k \\ 1 & otherwise \end{cases} \quad \textbf{(1)}$$

Among them, I_k and I_{k-1} are the current frame and the last frame respectively, and τ_k is the threshold to determine whether there is any change in frame k (25 is taken in the experiment).

(3) Update background B_k with binary image BW_k :

$$B_k = \begin{cases} B_k & if\ BW_k = 1 \\ \alpha . I_k + (1-\alpha) . B_{k-1} & otherwise \end{cases} \quad \textbf{(2)}$$

Where α is the renewal coefficient, which is 0.2 .

(4) $k = k+1$ Returns to step (3) and repeats again. After a certain number of repetition it will stopped, and Bk will used as the background image $B0$.

III. MOVING TARGET RECOGNITION

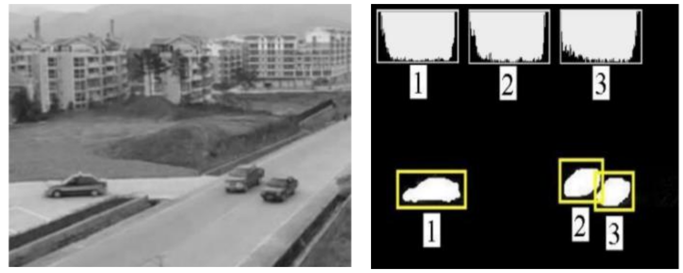
The difference between the current frame I_k and the background frame B_k is used to obtain the target image P_k , and the reverse backtracking method is used to obtain the coordinate points of the binary image corresponding to the color RGB target in the original image.

$I.Rk, I.Gk, I.Bk, \text{B.Rk}, \text{B.Gk}, \text{B.Bk}$ Are used to represent the values of R, G and B channels of target image I_k and background model B_k respectively.

$$P_k = \begin{cases} 0 & \delta_R < \tau_R\ and\ \delta_R < \tau_R\ and\ \delta_R < \tau_R \\ 1 & otherwise \end{cases} \quad \textbf{(3)}$$

Here, $\delta_R = |I.R_K - B.R_K|$, $\delta_G = |I.G_K - B.G_K|$, $\delta_B = |I.B_K - B.B_K|$ Represents R, G and B Three channel chromatic aberration difference between the current frame and the background frame respectively. τ_R, τ_G, τ_B For R, G, B Threshold value (used for expriment $\tau_R = \tau_G = \tau_B = 20$).

After binarization, there will be many outliers, holes, and gaps in the target image. The noise should be removed by filtering techniques such as expansion and corrosion in mathematical morphology, and then connected or segmented by the $\delta - 4$ connected region segmentation method. The external rectangle and numerical number indicate the sequence, position, height and width of the target block. For isolated areas where the area is less than a threshold (Ta=30 is set for the experiment), the separate area should be removed. What remains in the final foreground area is the eventually acquired moving target group $\{\text{p}_k(\text{i})\}_{i=1..n}$. Where n is the number of targets. Fig. 1 is background modeling and target extraction screenshots of moving vehicle video. Fig. 1 (b) is the three-vehicle targets obtained by mathematical morphology after making the difference between the original fig. 1 (a) and the background frame.



(A)Original image (B)Moving object histogram

Figure 1: Experimental results of multi-target object tracking

IV. MOVING OBJECT PREDICTION BASED ON KALMAN FILTER

The moving target group is a continually changing group, how to determine the position of the different target in the next frame. After review some previous research, this paper proposes to use the Kalman filter to predict the position of multi-target. Then the improved CamShift algorithm is used to approximate the precise position of the target further, to realize the fast location and tracking of multi-target. Kalman filter is an algorithm to consider the linear minimum variance error of the state sequence of the dynamic system. It represents the system by the equation of state and the equation of observation. Based on the extended Kalman filter [5], the area where the target i may appear in the next frame:

$$p_{k+1}(\text{i}) + \Delta\, \text{p}_{k+1}(i) \quad (4)$$

$$\Delta p_{k+1}(i) = p_k(i) + v_k(i)\Delta t + a_k(i)\Delta t^2 / 2 \quad (5)$$

$$\Delta p_{k+1}(i) = \Delta p_k(i) + \Delta v_k(i)\Delta t + \Delta a_k(i)\Delta t^2 / 2 \quad (6)$$

$p_k(i)$ Is the central and i is the target position. $p_{k+1}(\text{i})$ is the predicted position of the object i in the next frame. $\Delta p_{k+1}(i)$ Is the time interval of two frames.

V. IMPROVED CAMSHIFT ALGORITHM

In order to get color information to achieve accurate target location, CamShift is needed to check. First, RGB color space is transformed into HSV color space to adapt to CamShift. Then the chromaticity histogram is constructed with each i .

By standardizing the range of values corresponding to each chroma level in the histogram to the interval of [0.255], the gray value of each target in the sequence is obtained. In gray projection, the brightest pixel may be the more likely the target is [6]. After using Kalman filter, the predicted target position $p_{k+1}(i)$ is taken as the starting point, and $s_0(i) = \Delta\, \text{p}_{k+1}(\text{i}) / 255$ as the initial size of the search window, $i = 0$, start CamShift tracking:

(1) In the current frame search window, let $I(x,y)_i$ as the pixel position of target i . $I(x,y)_i$ is the pixel value at $(x,y)i$ in the projection. In the formula, $I(x,y)$ is the pixel value of (x,y) . The range of X and Y varies in the search window. Then the zero-order moments M_{00} and M_{10} of the I search window are as follows:

$$M_{00}^i = \sum_x \sum_y I(x,y)_i, M_{10}^i = \sum_x \sum_y xI(x,y),$$

$$M_{01}^i = \sum_x \sum_y yI(x,y)_i \qquad (7)$$

So the center position in the search window of i is:

$$(x_c, y_c)_i = (\frac{M_{10}^i}{M_{00}^i}, \frac{M_{01}^i}{M_{00}^i}) \qquad (8)$$

(2) Search window update $s = 2\sqrt{M_{00}^i / 256}$

(3) Repeat (1) and (2) processes, move the center of the search window to the center position. If the moving distance is larger than the preset fixed threshold, the center position of the window is recalculated, and then the position and size of the window are adjusted. When the moving distance between the center of the window and the center point is less than the exact threshold, or the number of cycles reach to maximum, the convergence condition is considered to be satisfied. At this point, $(x_c, y_c) = i$ is the centroid coordinate of the i .

(4) $i = i+1;$; Return to (1) and continue searching for the centroid of other targets until all the targets in this frame are searched. New frame number image search is carried out.

VI. Multi-Target Search window initialization

Due to the complexity of background image, the target may enter, exit, merge and separate at any time in the image. For this reason, the search strategy adopted in this paper is as follows: each target in the frame k is matched with each target in the frame $k+1$ through kalman filtering, and then Camshift tracking matching is carried out. If the matching is just successful, it means that the target has no variation. If there are remaining targets in frame k or frame $k+1$, it indicates that the target has variation. At the same time, the remaining targets are repeated and start to combine similar comparison and track again. If there are still redundant targets in the frame k after repeated comparison and calculation, it indicates that some targets are still remaining. At this time, the information of these targets in the frame k is retained for use when they appear again. If there are redundant targets in the frame $k+1$, it means that a new target enters, and then a new number is added to the new target, and update the existing model.

VII. Experimental results

In order to prove the improvement of our approach over traditional CAMSHIF, we implemented our algorithm with C++ code using OpenCV library and applied our method on various video sequence where the video frame rate is 25 f/s. figure 2 is a partial screenshot of multi-target video tracking in the park. In figure 2(1), there are two groups of targets, which merge into one target group in figure 2(4) and separated into three target groups in figure 2(5). However, this algorithm can still track the target well and automatically group the number.

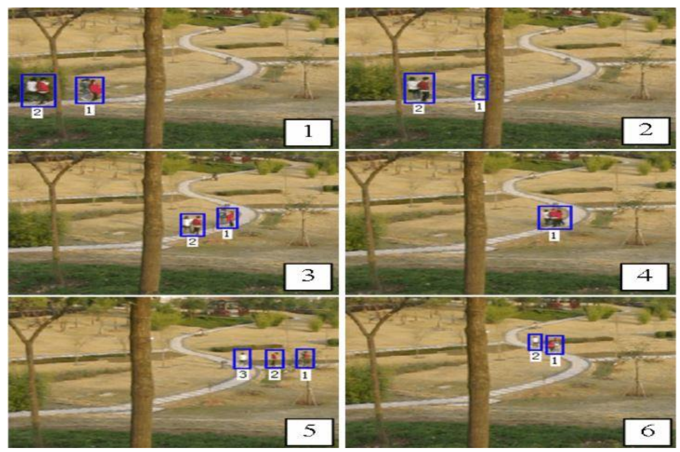


Figure 2: Experimental results of multi-target video tracking.

VIII. Conclusion

We proposed an efficient color based CAMSHIFT algorithm for target tracking in this paper. Combined with low cost motion segmentation techniques, we improved the traditional CAMSHIFT performance and showed how our approach can solve its drawbacks. Our improvements on CAMSHIFT are computationally attractive. The algorithm is use for dynamic background modeling technology combined with the Kalman filter to predict the target first and then use the CamShift algorithm to locate the target accurately. The dynamic tracking of the target is found by multi-objective search strategy and template updating. For motion segmentation we used simple background differencing in our implementation. A large number of experimental results show that the algorithm has strong target recognition ability, good anti-noise performance, and fast-tracking speed.

REFERENCES

[1]. J A. Yilmaz, O. Javed, and M. Shah, “Object Tracking: A Survey,” ACM Computing Surveys, vol. 38, p. 13 . 2006.

[2]. Wang Zhaowen, Yang Xiaokang, Xu Yi, et al., “CamShift guided particle filter for visual tracking,” Pattern Recognition Letters, v 30, n 4, pp.407-413, 2009.

[3]. Sun Hongguang,Zhang Jin,Liu Yantao,et al., “Optimized Particle Filter Tracking by CamShift Based on Multi-feature,” Opto-Electronic Engineering, v37, n2, pp 1-6,31,2010.

[4]. Shen Xuanjing, and Zhang Bo, “CamShift tracker based on image moments,” Journal of Beijing University of Technology, v 38, n 1, p 105-109, 2012.

[5]. W. Xiangyu, and L. Xiujuan, "The study of moving target tracking based on Kalman-CAMShift in the video," 2nd International Conference on Information Science and Engineering (ICISE), pp.1-4, 2010.

[6]. J. Yin, Y. Han, J. Li, and A. Cao, "Research on Real-Time Object Tracking by Improved CAMShift," International Symposium on Computer Network and Multimedia Technology, pp.1-4, 2009.

[7]. G. J. Allen, Y. D. Richard Xu and S. Jin Jesse, “Object Tracking Using CAMShift Algorithm and Multiple Quantized Feature Spaces”, Inc. Australian Computer Society, vol.36, 2004.

[8]. Li Chao, Liu Tiegen, Liu Hongli, et al., “Face tracking based on Haar detection and improved Camshift algorithm,” Journal of Optoelectronics Laser, v 22, n 12, p 1852-1856, 2011.

[9]. Comaniciu Dorin, Ramesh Visvanathan, and Meer Peter, “Kernel-based object tracking,” IEEE Trans. on Pattern Analysis and Machine Intelligence, v25, n5, pp.564-577, 2003.