2016 International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration

Moving Object Tracking Method Based on Improved Camshift Algorithm

Zheng Han, Rui Zhang, Linru Wen, Xiaoyi Xie, Zhijun Li
School of Automation
Wuhan University of Technology
Wuhan 430070, China
hz287948563@whut.edu.cn(Corresponding author: Zheng Han)

***Abstract*—The traditional Camshift tracking algorithm is easy to fall into local maximum when the target is occluded. it is easy to fail to track target when similar color interference or occlusion happens, and can't recover from failure. Aiming at this problem, the Kalman predictor and the inter-frame difference are adopted to improve Camshift algorithm. Firstly Kalman predictor predicts the next location of the target frame image, and regard this position as the center of a target tracking to determine search area Camshift algorithm; then use the size of the search box of the moving target to determine whether the target is similar color interference or occluded . If not, update the Kalman predictor parameters by position Camshift algorithm obtained; If it is similar color interference, extracted the moving object position and size by inter-frame difference to update Kalman predictor parameters; If obscured, the predicted value of the Kalman predictor is to be the current position of moving target, and the size of search window is a fixed value at the same time, using the set of values to update Kalman predictor parameters. Experimental results show that the improved algorithm can accurately track the target when similar color interference or occlusion happens.**

Keywords-Camshift; inter-frame difference; kalman predictor; moving target; similar color interference; occlusion

I. INTRODUCTION

Camshift is a kind of target tracking algorithm which uses color histogram as the target mode. It is mainly used to achieve the purpose of tracking by color information of moving target. The algorithm has good effect on tracking solid object in black and white background. However, if the target's color is similar to the color of background or the target is occluded, the tracking algorithm often fails and can not recover from failure [1].

Aiming at tracking moving objects in complex background, this paper proposes an improved Camshift algorithm. This algorithm combines the inter-frame difference and Kalman predictor. When it is similar color interference, extract moving target by the use of inter-frame difference to achieve accurate tracking; When occluded, achieve continuous tracking by Kalman predictor. This algorithm improves immunity and stability of the tracking algorithm and experiments verified the effectiveness of the algorithm.

II. CAMSHIFT ALGORITHM

A. *Camshift Algorithm Principle*

Camshift algorithm [2] can be divided into three parts: Calculating color projection (back projection)、Meanshift optimization、Camshift tracking algorithm. The basic idea is to do Meanshift operation to all image frames of a video sequence [3], and the previous result (the center position and window size of the search window) as the initial value of search window of the next frame Meanshift algorithm, go on iterating.

Because the RGB color space is sensitive to light luminance change, in order to reduce the influence of varying illumination on tracking target, first, convert the image from the RGB color space to the HSV color space. Count histogram of H component, and try to find out the probability of H component's size as x or the number of pixels, and obtain the lookup table of the color probability. Then the value of each pixel in the image is replaced by the probability of the appearance of color, thereby obtaining a color probability distribution. After that, we calculate each order Matrix of the probability distribution in the target area. Get target centroid position and size by successive iterations, and pass the result to the next frame to iterate, and achieve continuous tracking.Figure 1 shows the original image and its corresponding projection.

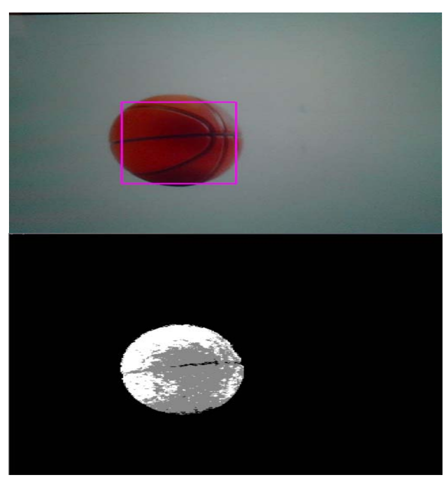


Figure 1. The original image and the corresponding projection

Camshift algorithm steps are as follows:

- Read a frame of video image, and convert it to HSV color space to extract H component.

DOI 10.1109/ICIICII.2016.55

- Initialize the search window, and add up the color histogram of the search box, and calculate the probability density of the color.
- After establishing color probability model of the target, convert the video image into a color probability distribution.
- Do the process of Meanshift algorithm, and obtain the centroid's position and size of the search window by calculating.

It is assumed that $I(x, y)$ is pixel values of (x, y) in back projection, and variation range of x and y is the size of the search window:

Zero-order matrix:

$$M_{00} = \sum_{x}\sum_{y} I(x, y) \tag{1}$$

First-order matrix of x:

$$M_{10} = \sum_{x}\sum_{y} xI(x, y) \tag{2}$$

First-order matrix of y:

$$M_{01} = \sum_{x}\sum_{y} yI(x, y) \tag{3}$$

Then calculate the central location of the search window:

$$(x_c, y_c) = (\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}}) \tag{4}$$

- Move the center of the search window to the centroid.
- It determines whether it is convergent (moving distance of centroid is less than a given threshold, or the number of iteration reaches the set value). If the convergence condition is not satisfied, repeat step 4; If convergence condition is met, then return the search results and use the current results of the center position and area size to search in the new image frame.

B. *The Deficiency of Camshift Algorithm*

The traditional Camshift algorithm has the following problems in practical applications:

- At the beginning of tracking, the traditional Camshift has to track targets artificially and to initialize the search window manually, which reduces the efficiency of the algorithm and the subjective errors are introduced.
- Calculating a color histogram and only for initialization will cause the probabilistic model of the target color not to update in time. When the lighting conditions change greatly, the statistical features of the target color have obviously changed. If you don't update probabilistic model of the target color in time, the effect of tracking will be affected.
- Camshift can achieve great tracking effect in the simple background environment, but tracking accuracy could reduce significantly when it is in the complex background, especially interfered by large area of similar color.
- Lack of prediction mechanism for target movement. It fails when the target moves too fast, the searching window of previous frame and the target of the next frame do not overlap, or the moving target is occluded temporarily in the tracking process.
- When the tracking target is lost or completely shielded, the searching window will contract and stay in a small area. Because the searching window is too small to connect with the target, it can not recover automatically even if the target reappears stably in the image sequence.

III. IMPROVED CAMSHIFT TRACKING ALGORITHM

A. *Inter-frame Difference Principle*

Inter-frame difference [4] method uses differences of two or several consecutive frame images in a video sequence to detect and extract target. The basic process of a basic inter-frame difference method is shown in Figure 2. The formulas of inter-frame difference show as the formula (1) and (2):

$$R_k(x, y) = \begin{cases} 0 & D_k(x, y) > T \\ 1 & D_k(x, y) \le T \end{cases} \tag{5}$$

$$D_k(x, y) = | f_k(x, y) - f_{k-1}(x, y) | \tag{6}$$

Where: D_k is the image after inter-frame difference; R_k is binary image; f_k is the current frame image; f_{k-1} is a previous frame image; T is the threshold value. First, use formula (5) to calculate the difference between the kth frame image and (k-1)th frame image which is used to obtain an image D_k , and then convert D_k to binary image . When a pixel value of D_k is greater than a given threshold, it believes that the pixel is a foreground pixel. Namely that the pixel might be a point on the target, otherwise it is considered to be a background pixel. Obtain R_k by converting D_k to binary image. Finally, obtain P_k by connectivity analysis to D_k . When a communication area of the region is greater than a given threshold, the moving target is detected and the area is an area which target occupies.

(a) (b) (c)

Figure 2. Inter-frame difference results

Figure 2 are experiment results of inter-frame difference. Where, (a) and (b) are two adjacent frame images, and (c) is the inter-frame difference image which is obtained by (a) and (b).

B. Kalman Predictor

In the process of tracking moving target, because time interval of two adjacent images is very short. The moving target is approximately regarded as uniform motion, and the state of the moving target is represented by two position variables and two speed variables. Define the target motion state vector, Observation state vector $Z_k = (x_k, y_k)$. x_k and y_k represent target centroid position in the image at the time of K. v_{x_k} and $v_{\mathrm{y_k}}$ represent the speed of movement of the target at the time of K.

The state equation:

$$X_k = F_k X_{k-1} + BU_k + W_k \tag{7}$$

Where, X_k is 4×1 dimensional vector which shows a state vector at discrete time k and is unpredictable; U_k is the control amount of the corresponding system; F_k is 4×4 dimensional vector which shows a state transition matrix at the time of t_k .

$$F_k = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{8}$$

In the formula, $\Delta \mathrm{t}$ represents interval of two consecutive frame images, and the paper takes 1 frame; W_k is 4×1 dimensional vector which is random interference of the state at the time of $\mathrm{t_k}$. This paper adopts uncorrelated zero mean standard normal white noise sequence.

Systematic observation equation:

$$Z_k = H_k X_k + V_k \tag{9}$$

Where, H_k is 4×4 dimensional vector which shows an observation matrix at the time of $\mathrm{t_k}$. V_k is 4×1 dimensional vector which shows an observation noise vector at the time of $\mathrm{t_k}$.

$$H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{10}$$

According to the formula (7), obtain systematic status prediction equation:

$$\hat{X}(k\,|\,k\text{-}1) = F_k\,\hat{X}(k\text{-}1\,|\,k-1) + BU_k \tag{11}$$

Where, U_k is the control amount of the current state, $\hat{X}(k \mid k-1)$ is the prediction result at the time of t_k , $\hat{X}(k\text{-}1 \mid k-1)$ is the result of the system state at time of $k-1$.

The current state of the system's update equation:

$$\hat{X}(k|k) = \hat{X}(k|k-1) + K_k[Z_k - H_k\,\hat{X}(k|k-1)] \tag{12}$$

K_k is Kalman gain coefficient:

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \tag{13}$$

$$P_k^- = F_k P_{k-1}^+ F_k^T + Q_k \tag{14}$$

$$P_k^+ = (1 - K_k H_k) P_k^- \tag{15}$$

Where, $\hat{X}(k \mid k-1)$ is the priori estimate of X_k , $\hat{X}(k \mid k)$ is the updated value which is obtain by the correction of Z_k to $\hat{X}(k \mid k-1)$; In addition, the random interference W_K and observation noise V_k are independent white noise sequences, and therefore their covariance matrixes are as follows:

$$Q_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad R_k = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{16}$$

Kalman predictor predicts the position of the target, and effectively improves the tracking failure caused by the short-term occultation of the target.

C. The Improvement of Camshift Algorithm

First of all, choose the tracking target, and then enter the Camshift algorithm to get the size and location of the target area. To overcome the problem of the serious occultation and interference of the background color, this paper introduces an improved algorithm which combines Kalman predictor [5] and the frame difference. Kalman predictor is used to estimate the target location in the process of occlusion [6], and the frame difference is used to overcome similar color interference. The concrete implementing process which is to use Kalman predictor to predict the center location of the searching window of the kth frame (k>2) in Camshift algorithm is: Firstly, initialize Kalman predictor. When the frame number K>2, predict the search area of Camshift algorithm by using Kalman predictor [7]; secondly, calculate the optimal location and the size of the target by using the Camshift algorithm; thirdly, decide whether or not there are similar colors in the background that interferes the system or occluded objects ; if not, use the target position which is calculated by Camshift algorithm as the observed value to update Kalman predictor directly; if it is only interfered by similar colors, use the location of the moving target captured by using the frame difference as the observed value to update Kalman predictor so as to realize the real-time tracking; if it is interfered by occultation, use the predictive value of Kalman's predictor as the current location of the moving target, and target size is a fixed value, then update Kalman predictor.

To solve the question about how to determine whether or not the system is interfered by similar color in the background or by occlusion [8], this article uses α—the change rate of the relative size of CurrentRect (the current frame of the tracking target area) and PreviousRect (the previous frame of the tracking target area) to determine whether there is interference by similar colors in the background, and uses β—the change rate of CurrentRect relative to OriginRect (initialized target area) to determine whether there is interference by occultation. The calculation formulas of α and β are as follows:

$$\alpha = CurrentRect / PreviousRect \tag{17}$$

$$\beta = CurrentRect / OriginRect \tag{18}$$

CurrentRect is the current frame of the tracking target area; PreviousRect is the previous frame of the tracking target area; OriginRect is the initialized target area. If $\alpha>2$, It is considered that the similar color interference is happening in the background; if $\beta<0.3$, It is considered that occultation interference is happening. The improved Camshift algorithm process is shown in figure 3.

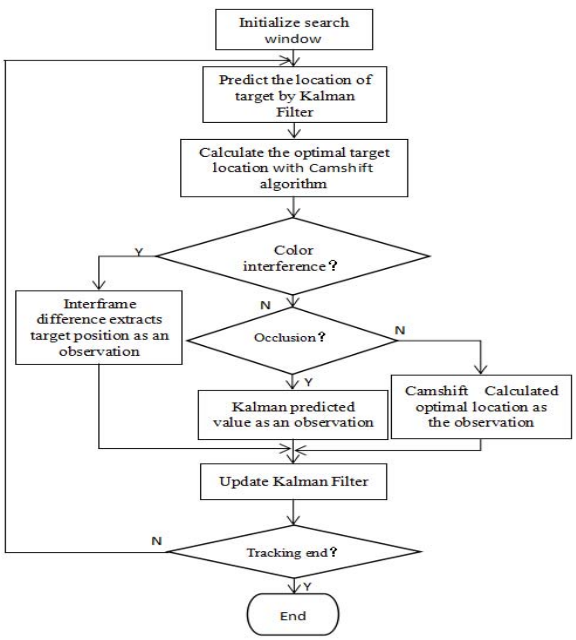


Figure 3. Improved algorithm processes

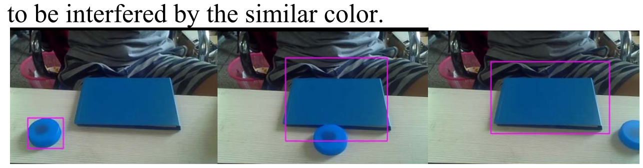
IV. RESULTS

In this paper, the experimental source code is developed in Visual Studio 2010, and the experiment is aiming at the occluded target and the similar color interference.

A. *Occlusion and Similar Color Interference Experiment*

In order to verify the effectiveness of the improved algorithm, the traditional Camshift algorithm is compared with the improved Camshift algorithm in terms of the tracking effect. As shown in figure 4 and figure 5, the experimental object is a set of video sequences, and the tracking results are showed in the red box.

In figure 4, a circular object(the blue cap in the original image) tracked is moving to the right and partially exposed to the similar color stationary object. (a) is the result of the original algorithm, and (b) is the result of the improved algorithm. As shown in (a), when the target is in contact with the interfering object, the search box becomes larger, and the search box stays on the interfering object when they separate, and is failed in tracking. In figure (b), When the target is in contact with the interfering object, the similar color interference is detected, and the frame difference is calculated in the next frame, and the moving target is extracted again and the normal tracking is maintained. In the experiment, the threshold value of the similar color interference is set to the change rate of the searching window, when the current search window is more than two times of the size of the previous search window , it is judged to be interfered by the similar color.

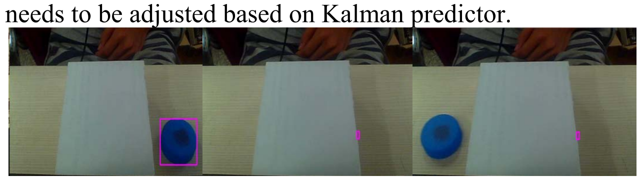


(a) Raditional Camshift algorithm

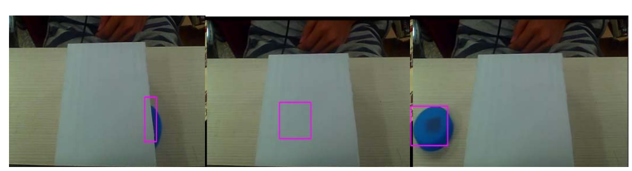
(b) Improved Camshift algorithm

Figure 4. Similar color interference experiment

As shown in figure 5, a circular object(the blue cap in the original image) tracked is occluded by a rectangular stationary object (the white in the image) in the left motion, and then it appears on the other side after a few frames. The result of the original algorithm and the improved algorithm are shown in the figure (a) and (b). When the window area is less than 100, which we set as the threshold for the loss of tracking, it is judged to fail in tracking, and at this time, it needs to be adjusted based on Kalman predictor.



(a) Original Algorithm



(b) Improved Algorithm

Figure 5. Occlusion interference experiment

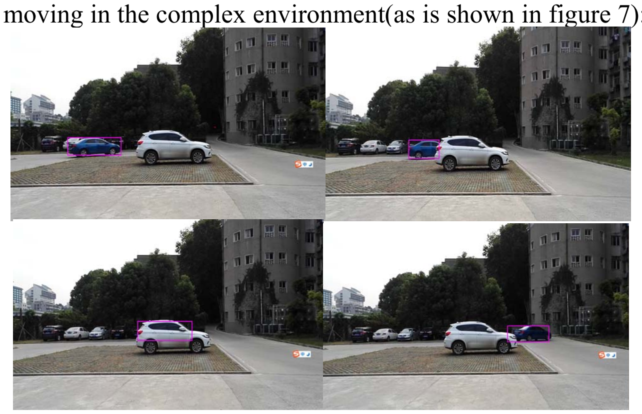
B. *Comprehensive Interference Experiment*

In figure 6, a set of video sequences of an electric car which runs from the right side of the scene to the left side are shown. The electric car goes through the similar color interference object and the occlusion on the way, as the result, the improved algorithm can track the moving target accurately and effectively.

Figure 6. Comprehensive interference experiment

C. *Vehicle Tracking Experiment*

There is a set of video sequences that a blue car is moving in the complex environment(as is shown in figure 7):



(a) Occlusion experiment

(b) Similar color interference experiment

Figure 7. Vehicle tracking experiment

In figure (a), the car moves through Occluded object. When detecting the occlusion interference, Kalman predictor predicts the next position where the target appears by the law of target motion, and target occluded is still able to be tracked accurately. In figure (b), when the car moves through a blue car, similar color interference is detected, and calculate inter-frame difference in the next frame, in this way, the moving target is extracted again and the normal tracking is maintained.The experiment shows that the improved Camshift algorithm can track moving target accurately.As can be seen from the video sequences, In the case of complicated background of the moving target, the algorithm effectively completes tracking.

V. CONCLUSION

This paper combines inter-frame difference, Kalman predictor and Camshift algorithm. Experiment shows that The new algorithm improves the stability and the accuracy of Camshift algorithm to tracking moving target. The algorithm has the following characteristics: When detecting the similar color interference, the inter-frame difference helps to solve the problem that the traditional Camshift has failed to tracking; In the meanwhile, when detecting the occlusion interference, the Kalman predictor helps to solve the problem that the traditional Camshift has failed to track; the new algorithm improves real-time tracking in complex background.

VI. ACKNOWLEDGEMENT

In this paper, the research was supported by Nature Science and Technology Major Project of China (2011ZX4002-21) and the fundamental research funds for the central universities (WUT:2013-IV-056).

REFERENCES

[1] DI Nan, ZHU Ming. “An Algorithm for Real-time Tracking Target in Complex Environment”. Journal of Image and Graphics, 2009, 13 (9):1758-1765.

[2] Ran Wang. “Reserch on Improved Moving Target Tracking Torecast Algorithm Based on Camshift”. JiNan:Master's degree thesis of Shandong University, 2012.

[3] Comaniciu D, Meer P. “Mean shift analysis and applicatioin”. Proceedings of the Seventh IEEE International ConferenceComputer Vison, Kerkyra, Greece, September 20-27, 1999. Piscataway, NJ, USA: IEEE, 1999: 1197-1203.

[4] XUE Li-xia, LUO Yan-li, WANG Zuo-cheng. “Detection algorithm of adaptive moving objects based on frame difference method”. Application Research of Computers, 2011, 28 (4):1551-1552.

[5] YAN Jun-hua, CHEN Shao-hua, AI Shu-fang, etc. “Target tracking with improved CAMShift based on Kalman predictor”. Journal of Chinese Inertial Technology, 2014(4):0536-07.

[6] Mazinan A H, Amir-Latifi A. “Applying mean shift, motion information and Kalman filtering approaches to object tracking”. ISA Transactions, 2012, 51(5): 485- 497.

[7] NIU Chun-feng, LIU Shi-ping, WANG Zhong-yuan. “Approach of Kalman filtering estimation of guided projectile attitude”. Journal of Chinese Inertial Technology, 2012, 20 (5):510-514.

[8] CHU Jun, DU Li-Hui, WANG Ling-Feng, etc. “Local Background-aware Target Tracking”. Acta Automatica Sinica, 2012, 38(6):1985-1995.