

A New Improved Camshift Algorithm and its Experimental Analysis

Qiang Wang¹, Xiao Li*

¹*School of Computer and Information Science, Southwest University
Chongqing, China
¹wqwq1219@163.com, * ivy@swu.edu.cn

Chaobing Liu², Mengyao Wu³

^{2,3}School of Computer and Information Science, Southwest University
Chongqing, China
²chaobing152376@163.com, ³mengyaowu@126.com

Abstract— Camshift is an adaptive tracking algorithm based on color histogram with the advantages of translation, scale and rotation invariance, which is widely used in visual tracking. But in the moving scenes, due to the changing background, occlusion and environmental illumination, the tracking effectiveness and efficiency will be affected and even lead to loss of moving target. Combined with the idea of both bottom-up and top-down, this paper suggests an improved one by using pixel filtering, histogram improvement and occlusion judgment, as well as Kalman filter. The comparative analysis results on the traditional Camshift algorithm and combined with Kalman traditional Camshift algorithm demonstrate that the proposed algorithm enhances the accuracy and stability in complex condition in an effective way and improves the real-time performance as well.

Keywords—Visual Tracking; Camshift Algorithm; Kalman Filter; Color Histogram; Occlusion Judgment

I. INTRODUCTION

With the development of computer vision, visual tracking [1-2] has gradually become a hot research field, which is widely used in intelligent transportation, biomedical and other fields. Among so many visual tracking algorithms, Camshift algorithm has been paid much attention because of its strong adaptability and the excellent tracking effectiveness and efficiency. However, this algorithm is easily to be interfered by environment, multi-target and occlusion. Focus on changing background in moving scenes, occlusion, illumination intensity, this paper improves the Camshift algorithm and combines with Kalman filter to forecast the track, which expect to improve the algorithm's tracking effectiveness and real-time performance in moving scenes and other complex environment.

The second section of this paper introduces the related work; the third section states the proposed algorithm in detail; the fourth section gives three comparative experiments and their result analysis; the final section is the summary of this paper.

II. RELATED WORK

At present, the Camshift algorithm and its improved algorithms are shown in TABLE I[3].

The tracking effectiveness and efficiency of the traditional Camshift algorithm is not too good in the complex background. When the color histogram of the target is created, this traditional algorithm only extracts hue component H and ignores target features of saturation S.

In order to solve these problems, this paper makes use of the two-dimensional histogram of H and S components to improve the utilization of the color information. This paper also applies Kalman filter to forecast the track and judges whether there is a large occlusion, if there is not a large occlusion, Kalman filter is used to predict. Otherwise, the linear prediction is used. This will improve the target's (occlusion is included) tracking accuracy.

III. ALGORITHM IMPROVEMENT

A. Introduction of Improved Ideas

The methods to solve the problem of visual tracking can be divided into two types: the idea of bottom-up and the idea of top-down. Combined with the above two methods, pixel filtering, histogram improvement and occlusion judgment are adopted as the improvement measures. Specifically, the following aspects are improved [4]: firstly, the high pass filtering is used to filter noise pixels; then H and S components are extracted to create a two-dimensional histogram. Thirdly, the weight is added according to the different contribution to the histogram of the pixel in different regions; finally, the Kalman filter is introduced to predict the track, which achieves accurate and stable tracking of the target in moving scenes.

B. Image Filtering Processing

In HSV color space, in addition to the tonal H component (hue), component saturation S (saturation) is also an important component, which mainly describes the color's bright degree. Because the main difference between the noise area and the target area is the value of component S of the pixels in the noise area is less, the high pass filter can be used to filter the pixels of less values of component S. Usually the truncation threshold of high pass filter is set fixedly according to people's subjective judgments. If the truncation threshold of high pass filter is applied to Camshift algorithms, the following qualification (1) needs to be added to filter out pixels to generate target template [5].

$$H : 0^\circ \sim 360^\circ, S : s_{\min} \sim 255, V : v_{\min} \sim v_{\max} \quad (1)$$

TABLE I. EXISTING CAMSHIFT ALGORITHM COMPARISON

Name	Advantage	Shortcoming
Camshift algorithm	Adaptive tracking algorithm. Easy to combine with other algorithms	Interfered by complex background. The real time is weak. lack of moving object trajectory prediction
Multi target tracking algorithm	Handling the problem of target instantly lost, as well as the occlusion between the multi-objectives.	Interfered by complex background. Ignoring the contribution of pixels in different positions on different color histogram.
Improved Camshift algorithm based on edge orientation histograms	Not sensitive to light, regardless of the color information of the object. Anti background interference.	Failed to track when the target is occluded. Do not apply to the object deformation and rotation.
Improvements of Camshift algorithm with intensity compensation	Using the light intensity compensation to solve the interference of light. Solve the problem of partial occlusion	Affecting the real-time tracking. Need to pre-judge the occlusion.

The s_{\min} is the minimum value of component S, the v_{\min} is the minimum value of component V and the v_{\max} is the maximum value of component V. Because these values are set according to people's subjective judgments, there is a large degree of uncertainty and delay, which cannot ensure the accuracy and the real-time performance of the filtered pixels. Therefore, this paper establishes a linear relationship between the search box and the filtering threshold of the component S to automatically adjust the value of s_{\min} . The specific process is as follows [6]:

- Assuming that the size of the window in the process of target initialization is S_0 , the lower bound filtering threshold of component S is $s_{\min}=30$, and the upper bound threshold is $s_{\max}=255$, the size of the frame i search box in the video sequence is S_i .
- Given: $k = S_i/S_{i-1}$. The lower bound threshold of component saturation S is updated by the following formula (2):

$$s_{\min}(i+1) = \begin{cases} s_{\min}(i) + 10, & k > 1 \\ s_{\min}(i), & k = 1 \\ s_{\min}(i) - 10, & k < 1 \end{cases} \quad (2)$$

In order to prevent the infinite increase or decrease of s_{\min} , this paper supposes the value of s_{\min} ranges from 30 to 100. Thereby the relation between the coefficient k of size change of the search box and component S is established. When the size of search boxes largen, that is, $k>1$, the lower threshold s_{\min} of S increases and the pixel filtering range is enhanced; When the size of search boxes lessen or remains unchanged, that is, $k<1$ or $k=1$, the lower threshold s_{\min} decreases and the filtering range reduces.

C. Color Histogram Improvement

In the traditional Camshift algorithm, the component H is extracted to construct color histogram [7] and the value of the

component H of each pixel in the image is calculated and added into the corresponding bin to get color histogram of component H of the tracked target as the target template of tracked target. It is shown in Fig. 1.

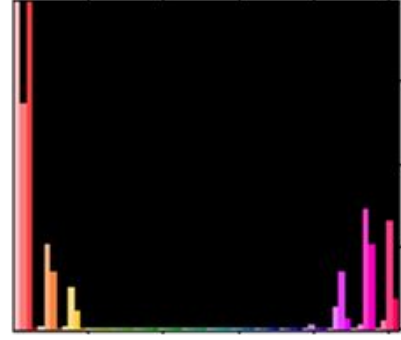


Fig. 1. The component H Histogram in HSV

In the process of extracting the component H and component S, the value range of each component in the HSV color space is as follows:

$$0^\circ < H < 360^\circ ; 0 < S < 1 ; 0 < V < 1 \quad (3)$$

In the OpenCV function library, the uchar of 8 bit cannot be used to express the component H value more than 255, so the component H value of the HSV is processed as follows:

$$h = \frac{H * 255}{360}; s = S * 255; v = V * 255 \quad (4)$$

In the above formula, h , s and v range from 0 to 255. In the process of target selection, the H-S two-dimensional histogram of the target area can be regarded as the two-dimensional coordinate function of the horizontal axis of the component H and the longitudinal axis of the component S. If the component H is divided into m bin and the component S is divided into n bin. The two-dimensional histogram of the image Lena is shown in Fig.2 below.

In Fig.2, the left one is a $120*120$ two-dimensional histogram, and the right one is a $256*256$ two-dimensional histogram. Considering the real-time requirements, the component H is divided into 30 bin and the component S is divided into 20 bin in the verification experiment, 600 bin in total, which reduce the amount of calculation. This improved method only involves the statistics of color feature of the histogram with simple operation and high efficiency. This method meets the requirements of real-time performance as well as improves the accuracy.

D. Occlusion Judgment

Kalman filter is an optimal estimation method under the criterion of minimum error covariance with less calculation and high real-time performance, which can use real-time motion parameters to constantly revise the estimation value of prospective motion status and improve the estimates' precision [8]. Therefore, the Kalman filter can be used to predict the target position to deal with the situation of the occlusion.

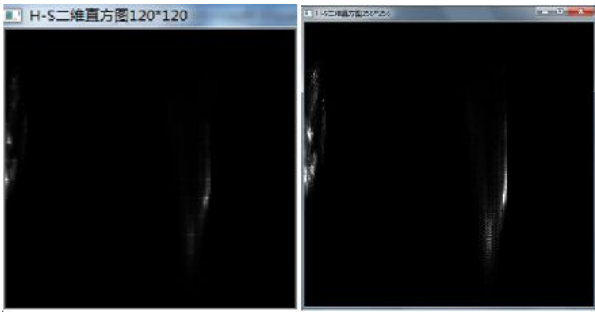


Fig. 2. The two-dimensional histogram of the image Lena

The tracking process combined with Kalman filter is shown in Fig.3:

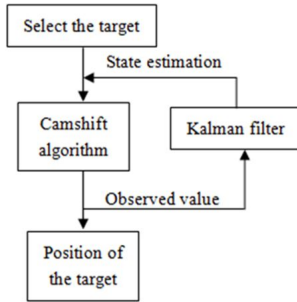


Fig. 3. Flow chart of the tracking system

The tracking process is divided into three main parts: firstly, the possible position of the target in the current frame is predicted according to the observation of the target position information in the last frame. Then, the Camshift algorithm is used to search the current target in the predicted area according to the color feature of the target. Thirdly, the position of the current target is regarded as the observation value of the Kalman filter to update the Kalman filter status, which estimates the position of the next frame.

To a certain extent, Kalman filter is used to solve the problems of tracking accuracy reducing and loss of the target because of the occlusion. [9] But if the target area is occluded for a long time, the color information cannot be extracted normally and the deviation of the observation value from the Camshift algorithm will occur. If the estimated value of the target is still predicted by the Kalman filter according to the observation value, the deviations will largen. Thus, this paper proposes that the large occlusion of the target is judged according to the distance between the observed value and the estimated value. The specific process is as follows:

- Assuming the size of a search window in the current frame is S , the window's width is w , the estimated value of the target is $X(k) = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ and the observed

value is $Z(k) = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$. The distance between the two

$$is d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

- In the process of the initialization of tracking target, a threshold is set according to the size of the selected window. The process of judging occlusion is shown in Fig.4 below:

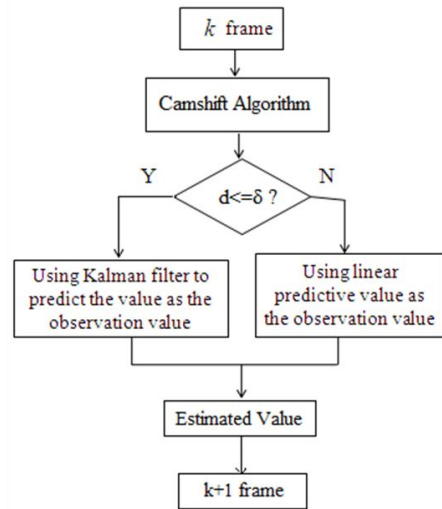


Fig. 4. The process of Occlusion judgment

When $d \leq \delta$, Kalman filter is used to predict the trajectory; if $d > \delta$, there is a large occlusion and the credibility of the predicted value from the Kalman filter reduces. The Kalman filter is stopped to forecast updates and the previous target position is used to predict linearly the position of the next frame. [10] The complete process of the tracking system is shown in Fig.5.

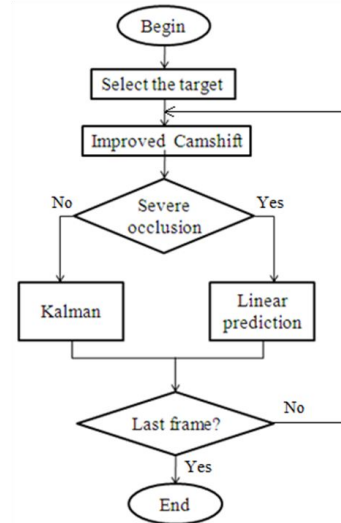


Fig. 5. The flow chart of the system tracking

The improved algorithm, the traditional Camshift algorithm and the combined with Kalman traditional Camshift algorithm will be compared and analyzed by using experiments and result data in the next section.

IV. EXPERIMENTAL VERIFICATION AND ANALYSIS

To validate the tracking accuracy and real-time performance of the improved algorithm, three videos filmed in moving scenes are selected as subjects in the experiment. The experiment is realized by using VC++ 6.0 and OpenCV [11] computer vision library to program in the win7.

The selected three videos in the experiment are filmed in moving scenes and video parameters are shown in TABLE II.

TABLE II. VIDEO PARAMETER TABLE

Video Sequence	Image Size	Frame Rate(fps)	Video Length(Frame)
Video 1	1280*720	25	510
Video 2	640*480	30	400
Video 3	720*480	30	600

At the same time, in order to verify and analyze the tracking performance and real-time performance, the traditional Camshift algorithm (hereinafter collectively referred to as algorithm 1), combined with Kalman traditional Camshift algorithm (hereinafter collectively referred to as algorithm 2) and the improved algorithm in this paper respectively are selected to track the three videos. Reference [7] uses the rectangle inscribed ellipse to show the target area. Reference [8] and the algorithm in this paper directly use rectangular box to display the target area.

A. Experimental Platform

The platform of verifications and analysis is composed of hardware and software. The hardware part of the experimental platform is a desktop of Acer Verition m275. The CPU is Pentium dual core CPU E5800 @ 3.20GHZ, the memory capacity is 4GB and the hard disk capacity is 500GB.

The software mainly includes the Windows 7 flagship operating system installed on PC, Visual Studio Microsoft 2010, OpenCV computer vision library and Matlab R2010b.

B. Background Interference Experiment

Among the three videos, Video_1 is a football match. The accuracy of the algorithms is verified by tracking the referee (the people wearing yellow clothes). The target tracking results of three algorithms are shown in Fig.6, Fig.7 and Fig.8.

By contrasting the tracking result of Video_1, it can be found that the tracking effect of the traditional Camshift algorithm for this video is poor (Fig.6). When the position of the target changes, the tracking frame gradually deviates from the target, finally completely lost; The target can be tracked normally by using the combined with Kalman traditional Camshift algorithm for this video in the beginning (Fig.7), but if the target is partially occluded, the target gets interference and the target is lost at last; From the Fig.8, it can be found that the target can be normally tracked by using the proposed algorithm in this paper and the tracking results are very accurate when the background gets interference or partially occluded.



Fig. 6. The results of the traditional Camshift algorithm



Fig. 7. The results of traditional Camshift algorithm based on Kalman



Fig. 8. The results of the new algorithm proposed in this paper

C. Occlusion Interference Experiment

In order to test the tracking effectiveness and efficiency of the algorithms under the condition of occlusion, the Video_2 that filmed in moving scenes is selected as a subject, which is a badminton teaching video. The video contains not only the target occlusion, but also the global motion, camera jitter and other interference factors. Above three algorithms are used to track the head of the coach in this video. The results are shown in Fig.9, Fig.10 and Fig.11[3].



Fig. 9. The results of the traditional Camshift algorithm



Fig. 10. Results of the combined with Kalman traditional Camshift algorithm



Fig. 11. The results of this improved algorithm

By contrasting the tracking results of this video, it can be found that the target can be tracked accurately and effectively by using algorithm 1 without any occlusion in the simple background. But if the tracking box stays in black background

areas, that is, the target is interfered by the complex background (Fig.9), the tracking target will be lost; The Fig.10 is tracking results of algorithm 2.

From Fig.10, it can be found that the target is tracked accurately by using algorithm 2 without any occlusion, but the target is lost when it is occluded completely. And Fig.11 shows that the target can be tracked accurately by using the algorithm proposed in this paper with stronger robustness in complex background or occlusion environment.

D. Illumination Change Experiment

Light is also an important interference factor for the field of target tracking. In order to verify the tracking effectiveness and efficiency of the algorithm improved in the paper in the environment of different optical intensities, the video_3 is selected to experiment, which is a character moving video filmed in a close range (target frame is relatively large).

The light intensity of the close range changes from weak to strong. The results of three algorithms are shown in Fig.12, Fig.13 and Fig.14.



Fig. 12. The results of traditional Camshift algorithm



Fig. 13. Results of the combined with Kalman traditional Camshift algorithm



Fig. 14. The results of this improved algorithm in this paper

By analyzing the tracking results, it can be known that the target can be accurately tracked (as shown in Fig.12) by using the algorithm 1 in the dark environment (that is the initial tracking stage). But the tracking window gradually deviates from the moving target with the gradual increasing of the light intensity; Fig.13 shows the tracking result of algorithm 2. It can be found that the light intensity has a great influence on algorithm 2 from the Fig.13. The tracking error is bigger and bigger with more and more obvious difference between the current and initial environment; The Fig.14 indicates that the algorithm proposed in this paper is not sensitive to the light intensity. And the tracking frame doesn't deviate from the target with the change of the light intensity. Experimental results show that the target can be still effectively tracked by using the improved algorithm proposed in this paper in the process of light intensity changing with stronger robustness.

By analyzing the above three groups of comparative experiments, it is known that the improved algorithm proposed in this paper, which based on Camshift algorithm, can effectively solve problems of lost of the tracking target in moving scenes by using pixel filtering, histogram improvement and occlusion judgment, as well as Kalman filter. At the same time, this algorithm achieves the accurate and stable tracking in moving scenes with stronger adaptability to the background interference, occlusion and illumination intensity.

E. Real-time Analysis

The above experimental results show that the improved algorithm proposed in this paper realizes the stable target tracking in moving scenes. The real time performance of the improved algorithm will be analyzed by contrasting iterations and time-consuming in the convergence of each frame next.

Due to the different video shooting environment, background interference, occlusion, illumination and the size of a target window will affect iterations of each frame in the process of convergence, this paper extracts the iterations statistics of the first 30 frame to compare according to tracking results of the three algorithms for the video_1. The result is shown in Fig.15.

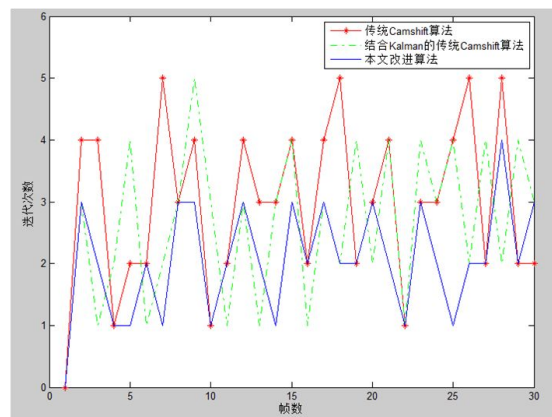


Fig. 15. The comparative chart of iteration number

In Fig.15, the star red line indicates the iterations of the traditional Camshift algorithm, the dashed green line shows the iterations of combined with Kalman traditional Camshift algorithm, the solid blue line on behalf of the iterations of the improved algorithm proposed in this paper. Obviously, after the first frame of the target is selected (the iteration value is 0), the number of iterations of the blue solid line is substantially lower than the star red line and the dashed green line. And the iterations of the improved algorithm are significantly less by contrast, that is, the improved algorithm is better than the traditional Camshift algorithm and combined with Kalman traditional Camshift algorithm in the aspect of convergence.

Average time per frame is one of the important criteria for the real-time of the tracking algorithm, which is easily affected by the hardware, software and experimental subjects. Therefore the time-consuming of the above three videos is recorded in the same experimental environment and the results as shown in TABLE III:

TABLE III. TOTAL TRACKING TIME

Name	Video 1	Video 2	Video 3
Algorithm 1	16s	11.8s	26s
Algorithm 2	28.8s	18.1s	31s
Algorithm improved in the paper	18s	12.9s	28s

According to the TABLE II of video parameters, the average time per frame of different videos is shown in TABLE IV:

TABLE IV. AVERAGE TIME OF EACH FRAME

Name	Video 1	Video 2	Video 3
Algorithm 1	31.4ms	29.4ms	43.7ms
Algorithm 2	56.4ms	45.2ms	51.4ms
Algorithm improved in the paper	35.4ms	32.2ms	46.1ms

From the above table, it can be seen that the time-consuming per frame of the algorithm in this paper is smaller by contrasting with algorithm 2, which improves the real-time performance. But the time-consuming per frame of the algorithm in this paper is slightly longer compared with the traditional algorithm. This is because the Kalman filter is used by the improved algorithm in the process of tracking and the Kalman filter takes some time. With respect to the stable and accurate tracking, it is also acceptable to slightly increase the time-consuming of the algorithm. Thus, the algorithm in the paper effectively improves the real time performance of the tracking system.

To sum up, it is proven that the real-time performance of the algorithm in this paper has certain promotion under the premise of stable and accurate tracking by analyzing the iterations and time consuming of each frame with above three algorithms.

V. CONCLUSIONS

In complex environment, the traditional Camshift algorithm is easy to be influenced by illumination and occlusion. In this paper, the traditional Camshift algorithm is

improved by pixel filtering, H-S two-dimensional color histogram and occlusion judgment, which can improve the tracking stability of the moving target in complex environment. Experimental results show that the improvement measures proposed in this paper can effectively solve the target loss problem led by environmental change, light intensity change and occlusion in complex environment. And these measures can improve the tracking stability of the system as well.

ACKNOWLEDGMENT

This work was funded by the National Key Technology R&D Program of China (2013BAD15B00).

REFERENCES

- [1] S. Afzal, N. Vaswani, A. Tannenbaum, A. Yezzi. Object Tracking: A Survey. IEEE, No.6,4244-4249,2009.
- [2] Bradski G R. Computer vision face tracking for use in a perceptual user interface[J]. 1998.
- [3] Chaobing Liu, Congcong Chen, Xiao Li. Object Tracking System in Dynamic Scene based on Improved Camshift Algorithm Kalman Filter. International Conference on Precision Mechanical Instruments and Measurement Technology, 2014.
- [4] Yizong Cheng. Mean shift, mode seeking, and clustering [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995.
- [5] Wenig M, Pang K. Arbitrarily colored ball detection using the structure tensor technique [J]. Mechatronics, 2011, 21(2): 367-372.
- [6] Fashing M, Tomasi C. Mean shift is a bound optimization [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(3) : 471-474.
- [7] Hsia K H, Lien S F, Su J P. Moving target tracking based on Camshift approach and Kalman filter[J]. Int J Appl Math Inf Sci, 2013, 7(1): 193-200.
- [8] J. Wang and F. He and X. Zhang and G. Yun . Tracking Object through Occlusions Using Improved Kalman Filter . IEEE 2010, No. 6, pp. 223-228 .
- [9] E. David, B. Erich, K. Daniel, S. Anselm. Fast and Robust Camshift Tracking [J]. IEEE Transaction, No.8, 1-8.2010.
- [10] X. Song, R. Nevatia. Camshift Tracker Design Experiments with Intel OpenCV and Sai. International Conference on Pattern Recognition and IEEE Computer Society, No.4, 1-4.2004.
- [11] Pulli K, Baksheev A, Korniyakov K, et al. Real-time computer vision with OpenCV[J]. Communications of the ACM, 2012, 55(6): 61-69.