

Target Detection and Tracking Algorithm based on the Improved Multi-feature Camshift

Naihui Fang^{1,2}

1. Faculty of Science and Engineering, School of Electronic Engineering and Computer Science, Queen Mary, University of London

2. Beijing University of Posts & Telecommunications
Beijing, China

e-mail: fangnaihui88@hotmail.com

Wei Fu

Department of Research and Development
Beijing Zhongliandianke Technology Co., Ltd.
Beijing, China

Abstract—In this paper, a vision-based target detection and tracking approach using improved multi-feature camshift is presented. Because the traditional camshift based on single feature is not robust when the background color, illuminate and target deformation change, it may lead to losing target or failure of tracking. In order to improve this problem, we presented target detection and tracking algorithm based on improved multi-feature camshift. The foreground image was obtained by Gaussian Mixture Model (GMM), which was used to modify the kernel function of target model. Due to the advantage of colour and texture feature in describing the target appearance, these features were used as recognition features. Additionally, EKF was integrated in tracking system to improve the accuracy of target tracking and predict the object position when the target was occluded. This paper details the architecture of the presented method and gives some experimental results to verify the effectiveness of the proposed method.

Keywords- Target Tracking; GMM; Camshift Algorithm; multi-feature; EKF

I. INTRODUCTION

With the development of computer vision technology, object detection and tracking algorithm has been an important research in the field of vision tracking [1]. The target detection and tracking technology has been developed and applied to several practical applications, such as service robotic systems for people tracking, vehicle tracking, security monitoring and so on. Generally, accuracy, real time performance and robustness are three important requirements for a target detecting and tracking system, which play an important role to the tracking. Accuracy requires that the tracking target must be recognized correctly; Robustness means that the process must be stable for a long term and has better anti-interference function; Real-time performance means that the tracking algorithm must be performed quickly and has a lower computational complexity. As video analysis needs to process large numbers of image pixels, the real-time performance problem of the system is extremely significant [2].

Except for efficiency, there are still some problems in target detection and tracking. For example, image sequences are often subject to many external factors and internal interferences from the acquisition system, including the diversity of clothing (if the target is human),

the impact of changes in illumination, the irregular movement of the target, and the impact of other objects, which brings some difficulties to the target detection and tracking research. The most important factor of these difficulties is how to extract the feature of the target accurately. Therefore, a lot of researches have been done to try to solve this problem. Generally, the tracking algorithms can be divided into two classes: one is generative models [3], which an appearance models are learned and used for searching image regions. The other one is discriminative models [4], which can separate the target from background. Collins [5] presented a successful online tracking method by selecting discriminative features. A. Yao et al. proposed an available visual tracking approaches based on appearance model using particle filtering (PF) and mean-shift (MS) algorithm. PF represents the posterior distribution by a set of samples, or particles, with associated importance weights [6]. This method has the problems such as large amount of calculation and lower real-time performance because PF algorithm usually demands a considerable number of sampled particles in order to accurately approximate the posterior probability density function of state. Fukunaga K et al. [7] and Comaniciu D et al. [8] proposed mean-shift which is a non-parametric estimation approach. The target tracking based on mean-shift algorithm represents the target as a histogram with a kernel function, and the gradient based mean shift optimization is used to iteratively seek a candidate which is the closest mode of the target model in the current frame. Jilin Tu et al. developed online updating appearance generative mixture model for mean shift tracking system [10]. In order to improve the problem of the occlusion of target tracking effectively, a new target detecting and tracking method combining sum-of-squared differences and color-based mean-shift (MS) tracker to improve the tracking failure of MS tracker in large changes of moving distance[11].

In order to eliminate the background and solve occlusion problem in the process of the target detection and tracking, this paper proposed a tracking algorithm based on the improved multi-feature camshift. Firstly, the foreground image obtained by GMM was used to modify the target model. To eliminate the background interfere for tracking, the foreground image combined color and texture back projection. This paper gives the details of the

proposed method and some experimental results verify the effectiveness of the proposed method.

The paper consists of 3 sections. Section 2 describes the architecture of the proposed improved target detection and tracking algorithm including Gaussian mixture model for eliminating the background, the improved multi-feature camshift algorithm and EKF. The experimental results are given in Section 3. Section 4 concludes the paper.

II. ARCHITECTURE OF THE IMPROVED TARGET TRACKING ALGORITHM

Fig .1 illustrates the software architecture of the improved target detection and tracking algorithm based on the improved multi-feature camshift. As described in Fig .1, the whole approach can be presented as follows: In the first frame, in order to complete the initialization, a rectangular region including the target is selected as original detecting area by hand. In this area, we extract the colour and texture of target. The foreground image obtained from GMM is used to modify the target model. Secondly, as the next frame, the colour back projection and texture back projection are calculated by colour model and texture model. The final back projection is obtained by colour back projection, texture back projection and foreground image using the Boolean AND operator. Then, the search window is located by the centroid localization algorithm. Finally, EKF is used to extract the motion feature of target, and estimate the target's position when the target is occluded.

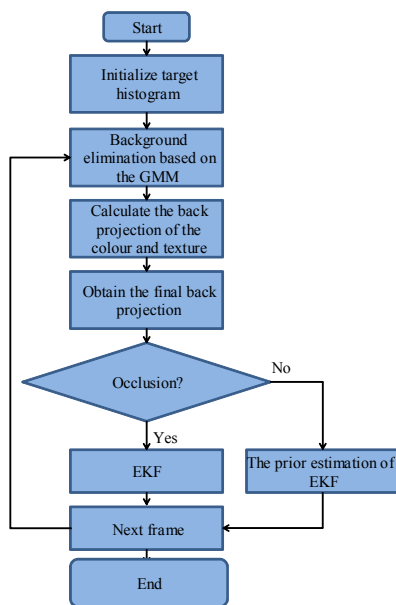


Figure 1. The flowchart of the proposed target tracking algorithm.

How to extract target feature and identify the target correctly is the most significant difficulty in target detection and tracking field. In the practical applications, it is obvious that traditional tracking algorithm based on single feature is easily influenced by the target rotation, translation, and the changes of the background colour and

illumination, which lead to lose the tracking target. Therefore, in this paper, multiple features are selected as identification features to improve the accuracy of target recognition. The foreground image is obtained by Gaussian Mixture Model (GMM), which is assembled in improved multi-feature to eliminate the background interfere. In addition, the EKF will be used to improve the accuracy of the detecting and tracking target.

A. Gaussian Mixture Model

Supposed that $X = [X_1, X_2, \dots, X_n]$ is an n-dimensional random variable and comes from a Gaussian mixture model of K components. Then, we can study the following probability density function of the Gaussian mixture model [12]:

$$\phi(x) = \sum_{i=1}^K \alpha_i q(x|\theta_i), \quad \forall x \in R^n \quad (1)$$

Where $\alpha_i \geq 0$ are the mixing proportions under the constraint $\sum_{i=1}^K \alpha_i = 1$. $q(x|\theta_i)$ is a Gaussian density with the parameters θ_i , which can be expressed as follow:

$$q(x|\theta_i) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma} e^{-\frac{1}{2}(x-m_i)^T \sum_i^{-1} (x-m_i)} \quad (2)$$

Here, $\Theta = \{\pi_1, \dots, \pi_K, \mu_1, \dots, \mu_K, \sigma_1, \dots, \sigma_K\}$ is a complete set of parameters needed to specify the Gaussian Mixture Model. We supposed there is a data set of N observations $X = \{x_1, x_2, \dots, x_n\}$, and modeled the data set with a GMM. The corresponding log-likelihood function is

$$\log p(X|\Theta) = \sum_{n=1}^N \log p(x_n|\Theta) \quad (3)$$

One approach to estimate the parameter set is maximum likelihood (ML) as following equation.

$$\hat{\Theta}_{ML} = \arg \max \{\log p(X|\Theta)\} \quad (4)$$

It is well known that the closed-form solution for equation (4) cannot be found analytically. One powerful method for finding the ML solution is the EM algorithm [13], which comprises two alternate steps: the E-step which computes the posterior probabilities [14].

B. Improved multi-feature camshift

Camshift algorithm is the expansion of meanshift algorithm. In order to track target by multi-feature, the camshift algorithm is used in our system, which is based on non-parametric density estimation method [8, 15]. Although the camshift algorithm has been improved by many scholars, it can hardly solve the incompleteness of an objective description and weak robustness of background interferes. In our research, an improved convex kernel function based on the silhouette of target is

utilized to overcome the background, various illuminations and other related limitations. The colour and texture feature are extracted to describe the target.

In the camshift tracking algorithm, we select the the position of the target firstly by hand, and the selected region is the target model and its colour histogram can be calculated by equation (5).

$$q_u = C_{q_u} \sum_{i=1}^n k_u \left(\left\| \frac{x_i - x_0}{h} \right\|^2 \right) \delta[\text{bin}_u(x_i) - u] \quad (5)$$

Here, $\{x_i\}$, $i=1,2,L,n$ is the pixel inside the selected rectangle. $\text{bin}_u(x_i)$ is the bin number $(1,2,L,m_c)$ at the pixel of normalized location x_i , $\delta(x)$ is the delta function and C_{q_u} is the constant normalization. Function k_u is Epanechnikov kernel. In this paper, we introduce an improved kernel function. The binary mask that is extracted from binary image by Gaussian Mixture Model (GMM) is used to eliminate interference of the margin pixel and background with similar colour or texture. As a consequence, the value of new kernel in the background is 0, while that in the foreground are same with Epanechnikov function.

For describing the texture of the target, the Local Binary Pattern (Local binary patterns, LBP) was widely used as an effective texture description operator. LBP has advantage such as gray-scale and rotation invariant. In this paper, we implement LBP to describe the texture of target and the texture histogram of the target can be calculated by equation (6).

$$q_v = C_{q_v} \sum_{i=1}^n k_v \left(\left\| \frac{x_i - x_0}{h} \right\|^2 \right) \delta[\text{bin}_v(x_i) - v] \quad (6)$$

In our improved camshift algorithm for target detection and tracking, we integrate the target colour and texture information, a colour back projection and a texture back projection to detect and track the target in order to improve the influence of the variation of light, similar color and texture background with the target, and so on. As shown in Fig .1, the colour-texture back projection combined with GMM foreground mask, which eliminated the background interference and assembled two feature of target. The position of the search window in next frame can be obtained through the centroid localization algorithm, and the position of next search window is calculated by equation (7)

$$(x, y) = \left(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right) \quad (7)$$

Here, M_{00} is the zero moment, M_{01} and M_{10} are the first moment of the final back projection. The details of the improved multi-feature target detection and tracking algorithm is as TABLE 1.

TABLE I. THE IMPROVED MULTI-FEATURE CAMSHIFT

Algorithm 1 Improved multi-feature camshift
<ul style="list-style-type: none"> • Initialize target histogram via q_v, q_u ; $i=2,3,4,\dots,n$ • Calculate the back projection of the colour and texture via Equation(5) and (6); • Obtain the final back projection by colour back projection, texture back projection and foreground image using the Boolean AND operator; • Find the new location of search window.
Algorithm 1 END

C. Extend Kalman Filter

EKF (Extended Kalman Filter) is very powerful to estimate the future state (prediction) even when the precise nature of the model system is unknown and provides an efficient solution to prediction problem for a nonlinear system, subject to additive white Gaussian noise. In our system, we use EKF to solve the problem of target mistracking in the case that the other object or human in environment occluded the target.

In our target detecting and tracking system, we define that the state vector including the position and velocity of the target is $X = [x, y, \& \&]$. $[X_r, Y_r]$ is the 3D position of the target, and $[\&, \&]$ is the velocity in the horizontal plane. The state equation is given by equation (7) and the observation equation is defined by equation (8).

$$X_r = F(\Delta t)X_r + W_r \quad (7)$$

$$Y_r = H_t X_r + p_t \quad (8)$$

Where $F(\Delta t_r)$ and W_r is the system process white Gaussian noise. By using EKF, the target detecting and tracking accuracy can be improved even if the target was occluded.

III. EXPERIMENTAL RESULTS

To verify the effectiveness of the improved multi-feature target detection and tracking algorithm, a laptop equipped with USB camera is taken as experimental platform. Fig .2 and Fig .3 show some comparison experimental results of different camshift algorithms. According to Fig .2, the traditional camshift cannot precisely track given target due to background color interference (such as frame number 121 and 246). In comparison with the traditional mean-shift algorithms, the proposed method is more accurate and robust because we used the improved convex kernel function and background elimination based on GMM.

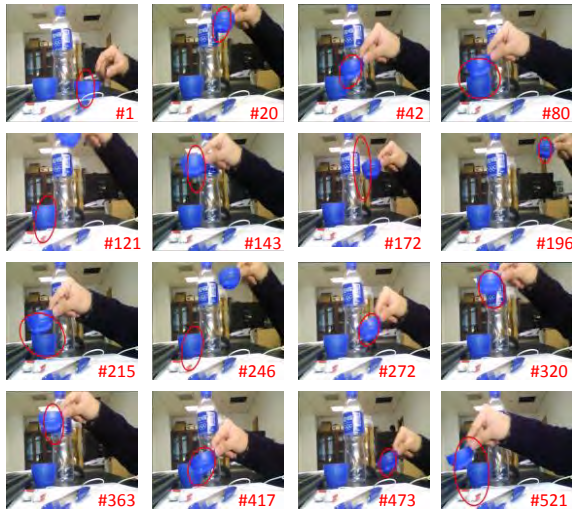


Figure 2. The tracking result using the traditional camshift

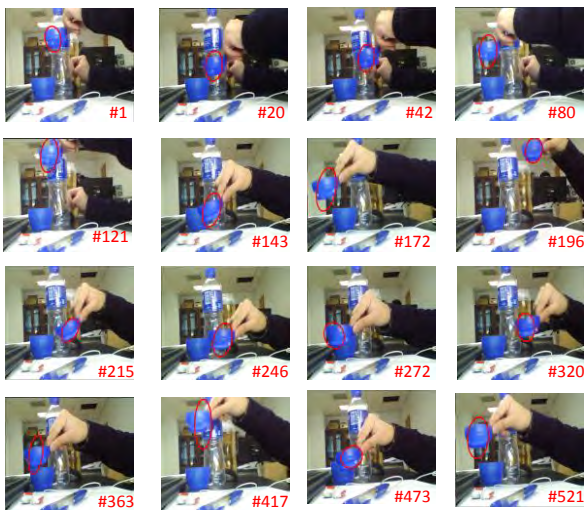


Figure 3. The tracking result using the improved camshift

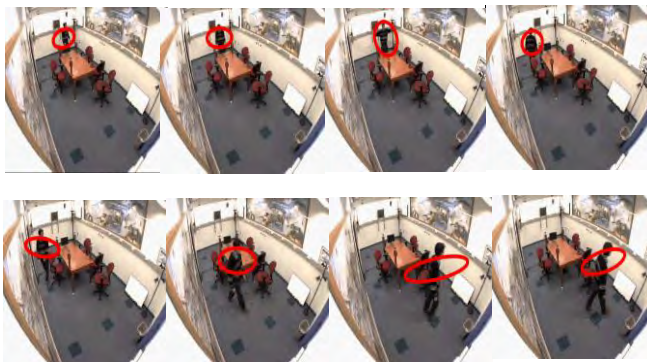


Figure 4. The experimental results of indoor human tracking using the improved camshift algorithm

In Fig .4, the experiment for indoor human detecting and tracking were demonstrated using real video (ViSOR, [16]). Red circle is the location results based on the improved multi-feature camshaft. From the Fig .4, it can

be seen that the presented method in this paper is successful to detect and track target human by EKF.

IV. CONCLUSIONS AND FUTURE WORK

In this paper, the target detection and tracking approach based on the improved multi-feature camshift was proposed. In order to solve the mistracking problem caused by background color, deformation and so on, the proposed tracking method integrated color with texture, which can make use of the advantages of color and texture to improve robustness of tracking algorithm. Additionally, the EKF was integrated in tracking system, which was used to improve the accuracy of tracking and predicted the object position when the target was occluded. The experimental results in the real environment verified the effectiveness of the proposed target detecting and tracking method.

For future research, improving our current algorithm in order to track the target robust in more complex environment will be main topic.

REFERENCES

- [1] J. Wang, "Integrating Color and Shape-Texture Features for Adaptive Real-Time Object Tracking". *IEEE Transactions on Image Processing, Correspondence*, 2008, 17(2): 235-240.
- [2] C. Liang, "Weighted modified Hu moment in human behavior recognition". *Advanced Information and Computer Technology in Engineering and Manufacturing Environmental Engineering*, Xiamen :AMSMT, 2013, 2194-2198.
- [3] M. Black and A. Jepson, "Eigen tracking: Robust matching and tracking of articulated objects using a view-based representation," In Proc. Eur. Conf. Comput. Vis, pp. 329-342, 1996.
- [4] H. Grabner, C. Leistner, and H. Bischof, "Semi-supervised on-line boosting for robust tracking," In Proc. Eur. Conf. Comput. Vis. pp. 234-247, 2008.
- [5] S. Avidan, "Support vector tracking," *IEEE Trans. Pattern Anal. Mach. Intel.*, vol. 29, no. 8, pp. 1064-1072, 2004.
- [6] A. Yao, XG. Lin, et al, "A compact association of particle filtering and kernel based object tracking," *Pattern Recognition.*, 45, (2012) 2584-2597.
- [7] Fukunaga K, Hostetler L, "The estimation of the gradient of a density-function with applications in pattern recognition," *IEEE Trans on Information Theory*. 21(1) (1975) 32-40.
- [8] Comaniciu D, Ramesh V, Meer P, "Kernel-based object tracking" *IEEE Trans on Pattern Analysis and Machine Intelligence*. 25(5) (2003) 564-577.
- [9] A. Yilmaz, O. Javed, M. Shah, "Object tracking: a survey," *ACM Computing Surveys*. 38(4) (2006): 1-45.
- [10] Jilin Tu, Hai Tao, Thomas Huang. Online updating appearance generative mixture model for meanshift tracking[J]. *Machine Vision and Applications*. 2009, 20 :163 - 173.
- [11] Babu.R, S. Suresh, Anamitra Makur, "Online adaptive radial basis function networks for robust object tracking," *Computer Vision and Image Understanding*. 2010, 14(3): 297-310.
- [12] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer Press, New York, 2006.
- [13] G.J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, Wiley, New York, 1997.
- [14] D.T. Nguyen, L. Chen, and C. K. Chan, "An outlier-aware data clustering algorithm in mixture model," in Proc. 7th Int. Conf. Information, Communication and Signal Processing, Macau, China, 2009: 1-5.
- [15] Cheng Yizong, "MeanShift, mode seeking, and clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1995, 17(8): 790-799.
- [16] http://imagelab.ing.unimore.it/visor_test/index.asp.