

Research on Target Tracking Algorithm based on Improved Mean Shift and Kalman Filter

Bian Minjie¹ and Honghao Gao²

¹*School of Computer Engineering and Science Shanghai University, 200444
Shanghai, P.R. China*

²*Computing Center, Shanghai University, 200444 Shanghai, P.R. China
bianmj0302@shcu.edu.cn*

Abstract

In this paper we discuss about a target tracking algorithm based on Mean Shift and Kalman Filter, which is suitable for high speed moving target tracking. The basic Mean Shift algorithm is described in this paper as well. Although basic Mean Shift algorithm can realize target tracking without arguments or searching all the areas effectively, it has shortcomings which can limit its effectiveness under some conditions such as the target covered by something when moving or the target moving with high speed. In order to make Mean Shift target tracking algorithm satisfies the practical application, such as high speed moving target tracking, we analysis the lack of basic Mean Shift algorithm, then we design a target tracking algorithm which combine an improved Mean Shift target tracking algorithm with Kalman Filter(IMKTT). In the fourth section of the paper, we apply the algorithm to a monitoring system of Shanghai University. The result shows that the algorithm designed in this paper can effectively track high speed moving target, such as a car.

Keywords: *Intelligent video monitoring, Target tracking, Mean Shift, Kalman Filter*

1. Introduction

Currently, intelligent video monitoring involves computer vision, image processing and analysis, digital signal processing and many other fields. This technology has been applied to many industries, such as dealing with traffic violations and traffic control, etc. Intelligent video monitoring contains the key technologies such as moving target detection, identification, tracking and understanding. Moving target tracking is the processing between target detection and target understanding. Tracking moving target is proceeded by the results of target detection and thus to support the target understanding. The results of target tracking will have a direct impact on the effectiveness of the intelligent video monitoring system. Therefore, the research of target tracking is very significant.

Moving target tracking is to find the target position and related moving information in continuous image sequence. There are two major methods of moving target tracking: one is to detect directly from the video sequence and extract moving target without relying on a priori understanding of the target, then determine whether the target is need to be tracked or re-detecting; another is to reply on the prior understanding of the target which sets up the model of the tracking target at first to find the one that matches the target in the video sequence, then track the target.

2. Improved Mean Shift

Mean Shift is a gradient ascent algorithm to find the maximum probability density. Since the algorithm has fast rate of convergence and does not require a priori knowledge, it has been widely used in the cluster analysis, image segmentation, target tracking and other fields [2].

2.1. Mean Shift

Given the space of sample points, X_i ($i= 1 \dots n$), the basic form of the Mean Shift vector of the sample point X is defined as:

$$M_h(x) = \frac{1}{k} \sum_{x_i \in S_h} (x_i - x) \quad (1)$$

Where, S_h is the radius of a spherical region of high dimension h , which is the set of all points satisfying the following relationship:

$$S_h(x) = P \{ y : (y-x)^T (y-x) \leq h^2 \} \quad (2)$$

And k represents the number of the above sample points which fall within the area. As shown, the big circle delineated the scope is related to the area, the small red points are on behalf of the sample points which fall within the area, the center of the circle is on behalf of Mean Shift offset center x , arrows indicate the sample points X_i offset relative to the reference point vector.

In the formula (1), $(X_i - X)$ represents the offset vector related to the area of the sample points relative to the center. The Mean Shift vector $M_h(x)$ is the point of all the samples fall within the area with respect to the average of center offset vector. But this is based on an assumption that all of the sample points are subject to the same probability density distribution. According to the relevant gradient, gradient nonzero probability density function is always pointing to the probability density function of the direction of the fastest growing. The relevant sample points within the region will be more inclined to fall on the density gradient direction. So Mean Shift vector points to the direction, the gradient direction, along which the probability density function grows fastest. As Figure 1 shown, the Mean Shift vector points to the sample distribution with higher density [3, 4].

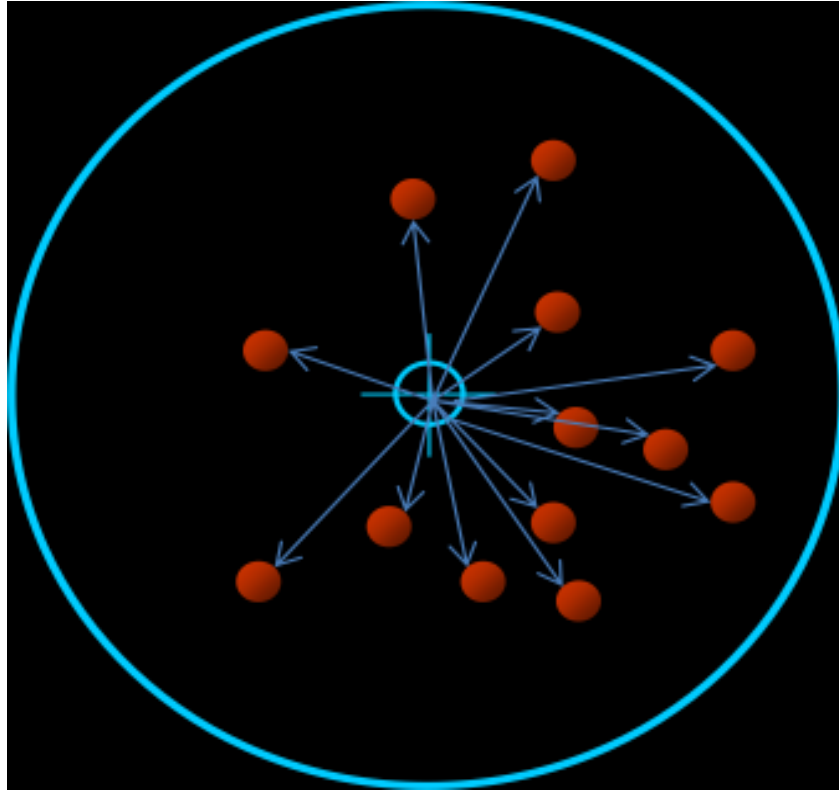


Figure 1. Mean Shift Schematic

2.2. Improved Mean Shift

In this paper, we improve Mean Shift algorithm by adding the kernel function and weight coefficient. Kernel function is just limited the importance of the sample, not exactly reflect the importance of each sample. In target tracking, the distance from the center point cannot fully determine the importance of the sample points, so the weight coefficient will be helpful. For the sample points that fall within the relevant area, their weight coefficient $\omega(x_i)$ is determined by the characteristics of each sample points.

With adding the Kernel function and weight, the improved Mean Shift algorithm became as:

$$m(x) = \frac{\sum_{i=1}^n G_H(x_i - x) \omega(x_i - x)}{\sum_{i=1}^n G_H(x_i - x) \omega(x_i)} \quad (3)$$

Where $G_H(x_i - x) = |H|^{-\frac{1}{2}} G(H^{-\frac{1}{2}}(x_i - x))$, $G(x)$ is Gaussian kernel function units, H is a positive definite matrix of $d \times d$ which calls bandwidth matrix, $\omega(x_i)$ shows the weight of sample point x_i .

Expression (3) is too complicated, and it can be simplified in practical applications. In image processing, the bandwidth matrix is often limited to a diagonal matrix or unit matrix, Expression (3) Can be rewritten to:

$$m(x) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x_i - x}{h}\right\|^2\right) \omega(x_i)}{\sum_{i=1}^n g\left(\left\|\frac{x_i - x}{h}\right\|^2\right) \omega(x_i)} \quad (4)$$

Through mathematical proof, that the use of Mean Shift vector kernel function G was proportional to the gradient in the x-point and normalized estimated from the kernel function K with the probability density function, where the normalization factor using the kernel function G at x estimate the probability density function.

$$m_{h,G}(x) = \frac{h^2 \nabla f_{h,k}(x)}{2 f_{h,G}(x)} \quad (5)$$

Where $\nabla f_{h,k}(x)$ is the gradient of probability density function and $f_{h,G}(x)$ is the estimation of density function as G is the probability kernel function.

Hence, the region of zero gradient probability density distribution can be found by Mean Shift vector. So looking for the steady-state point can be obtained by the following iterative process:

- 1, Calculate the offset vector Mean Shift vector of the offset datum
- 2, Move Kernel Function to $m_{h,G}(x)$
- 3, Return 1, until convergence to a neighborhood point or the end conditions of the iteration are satisfied.

2.3 Improved Mean Shift Target Tracking Algorithm

2.3.1 Target Model: Let $\{x_i\} i = 1 \dots n$ is the normalized pixel coordinates of the target model; the kernel function of the histogram is defined as the target model:

$$q_u = c \sum_{i=1}^n k\left(\left\|\frac{x_i - x}{h}\right\|^2\right) \delta[b(x_i) - u] \quad (6)$$

Where x is the central point to the magnitude, x_i is the magnitude of the No.i point,

h is the bandwidth of the Kernel Function. In Function K, $\left\|\frac{x_i - x}{h}\right\|^2$ eliminates the effects of different size and shape of objects. C is a normalization constant which

makes $\sum_{u=1}^m q_u = 1$, where m is the number of eigenvalues, for image processing can select the number of gray levels divided. $\delta[b(x_i)-u]$ is used to determine whether the feature belongs to value u, if belongs the value is 1, if not, the value is 0.

2.3.2 The Candidate Area: Each frame after the first one which may contain the tracking target is called the candidate area, where y is the center coordinates; the histogram of the Kernel Functions is expresses as:

$$p_u(y) = c_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \delta[b(x_i)-u] \quad (7)$$

2.3.3 The Similarity Function: To calculate the degree similarity between the target model and the candidate model by using the similarity function. Bhattacharyya Coefficient is used as similarity function in this paper. B Coefficient is defined as:

$$\rho(y) = \rho[p(y), q] = \sum_{u=1}^m \sqrt{p_u(y)q_u} \quad (8)$$

2.3.4 Improved Mean Shift Target Tracking Algorithm: Matching process is the process of finding similar function maximum optimization. Make Taylor expansion with the similarity function and eliminate the former two to get:

$$\rho[p(y), q] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(y_0)q_u} + \frac{1}{2} \sum_{u=1}^m p_u(y) \sqrt{\frac{q_u}{p_u(y_0)}} \quad (9)$$

let $\omega_i = \sum_{u=1}^m \sqrt{\frac{q_u}{p_u(y_0)}} \delta[b(x_i)-u]$, $G(x) = -k'(x)$

Then $\rho[p(y), q] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(y_0)q_u} + \frac{1}{2} c_k \sum_{u=1}^m \omega_i k\left(\left\|\frac{y-x_i}{h}\right\|^2\right)$ (10)

To make a similar function to the maximum, make the search direction and gradient direction y consistent, Derivation y_0 can be lent on gradient direction:

$$\nabla \rho[p(y), q] = \rho'[p(y), q] = \frac{c_k \sum_{i=1}^n \omega_i (y_0 - x_i) G\left(\left\|\frac{y-x_i}{h}\right\|^2\right)}{h^2} = \frac{c_k}{h^2} \sum_{i=1}^n G\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \left[\frac{\sum_{i=1}^n \omega_i x_i G\left(\left\|\frac{y-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n \omega_i G\left(\left\|\frac{y-x_i}{h}\right\|^2\right)} - y_0 \right] \quad (11)$$

Then calculate the vector from center y_0 in the candidate region to y_1 in the true region

$$m(y) = y_1 - y_0 = \left[\frac{\sum_{i=1}^n \omega_i x_i G\left(\left\|\frac{y-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n \omega_i G\left(\left\|\frac{y-x_i}{h}\right\|^2\right)} - y_0 \right] \quad (12)$$

After each time the Mean Shift algorithm iterations, the center position of the target area moves from y_0 to y_1

$$y_1 = \frac{\sum_{i=1}^n \omega_i x_i G\left(\left\|\frac{y-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n \omega_i G\left(\left\|\frac{y-x_i}{h}\right\|^2\right)} \quad (13)$$

Calculate $m(y)$ vector mode, if it does not satisfy the iteration termination condition, then according to the coordinates of the center of the new y_1 , histogram candidate model calculation kernel function until the end of the iteration, the iteration end of the next frame into the calculation.

3. Target Tracking Algorithm based on Improved Mean Shift and Kalman Filter

3.1. Deficiencies of Mean Shift

Although Mean Shift algorithm can realize better target tracking without arguments and searching no exhaustive area or other features, it has shortcomings which can limit its effectiveness under certain tracking conditions at the same time. When the background is comparatively complicated, it is easy to fall into local extreme with a non-ideal tracking result. If we select grayscale and color feature when the color of tracking target is similar to the grayscale, the result will be unsatisfied. Mean Shift algorithm selection of traditional fixed-size window size on the track, which is defined bandwidth of kernel function H is a diagonal matrix or matrix. It is easy to fail to track when the target moves too fast or the target size changes largely. When there is occlusion occurs, there will be a certain size and a new gray area and resulting in target tracking failure [8-12].

3.2. Kalman Filter

Kalman Filter is a recursive filter which was proposed for time-varying linear systems by Kalman. This system contains the model equations can be used to describe the orthogonal state variables. Kalman Filter is to estimate past measurements to estimate the future merge into a new error in measurement error. Kalman Filter is an efficient recursive filter which can estimate the state of dynamic systems from a series of incomplete and contain noise measurements.

3.3. IMKTT-Target Tracking Algorithm based on Improved Mean Shift and Kalman Filter

The basic Mean Shift algorithm is relatively simple and with low computational

complexity and can achieve good tracking result by improving and compatible with other tracking methods. Since Mean Shift algorithm is based on characteristic tracking without any prediction, so when there is something covering the target or the target moves fast, the algorithm may lose efficacy and the tracking target will be lost. In order to overcome the problem, a combination of Kalman Filter and improved Mean Shift target tracking algorithm mentioned above is designed in this paper to achieve moving target tracking. The algorithm IMKTT is as follows and the schematic is shown in Figure 2.

- 1-Track the moving target based on the improved Mean Shift target tracking algorithm and record the target position in current frame and other related information.
- 2-If the target is not covered or missed, go back to the step (1) to continue tracking. Otherwise, go to the step (3).
- 3-Analyze the motion parameters by Kalman Filter to predict the moving target status and then jump to the step (1) continue tracking.

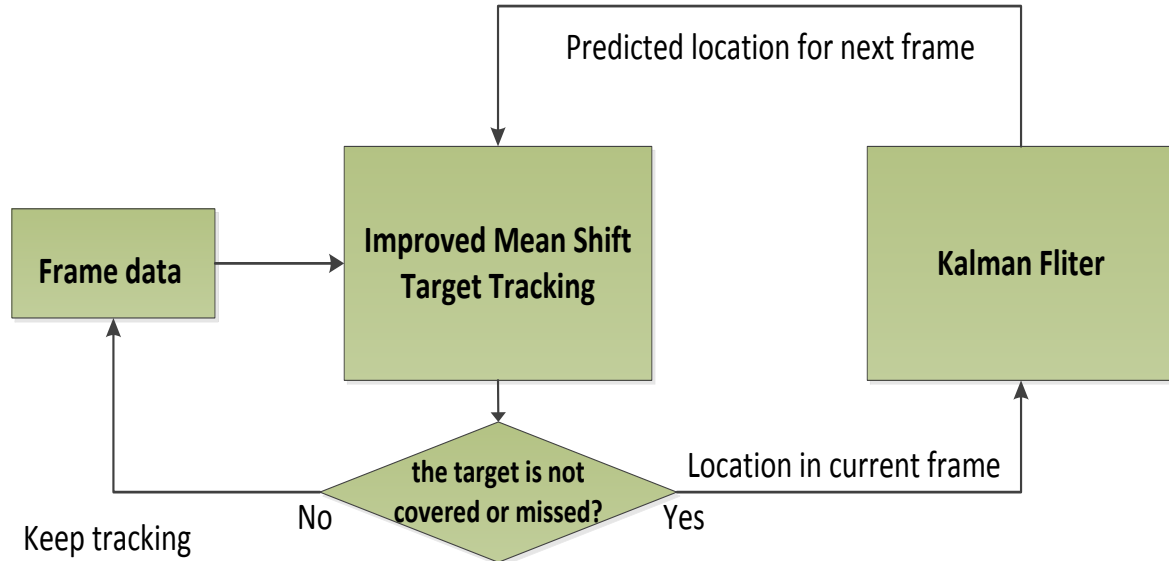


Figure 2. Schematic of IMKTT

4. Application of IMKTT

In this paper, we apply IMKTT to the monitoring system of Shanghai University. The improved Mean Shift algorithm is mainly used to track moving target and record its location in each frame and related information. Then Kalman Filter is used to predict the location of the moving target to ensure the continuity and accuracy of the moving target tracking. Figure 3 is the screenshot of the application developed based on IMKTT. As Figure 4 is the experiment screenshots, we can see that the car as a moving target can be tracked successfully which proved IMKTT has good effectiveness on high speed moving target tracking.

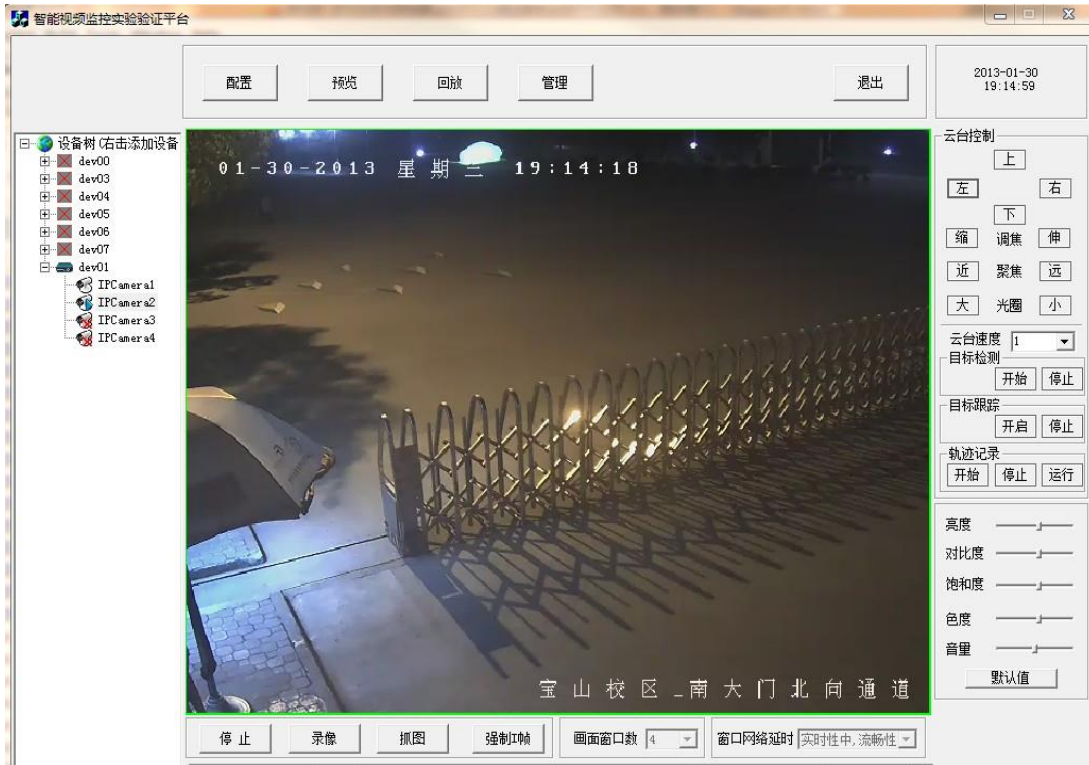


Figure 3. Application of IMKTT Screenshot



Figure 4-1. Experiment Screenshots



Figure 4-2. Experiment Screenshot



Figure 4-3. Experiment Screenshot



Figure 4-4. Experiment Screenshots

5. Conclusions

In this paper, IMKTT, an improved Mean Shift target tracking algorithm combined with Kalman Filter is designed according to the shortcomings of basic Mean Shift in practical application. We defined the bandwidth of the kernel function no longer as a fixed diagonal matrix. We detect the contour of the current location frame and then calculate its size at first. The new bandwidth will be defined by this size and then iterate with the next frame data. The bandwidth of the window will change with the size change of the target which overcomes the problem of size change of the moving target. Then Kalman Filter provides the algorithm the ability to predict the location of the moving target which ensure the continuity and accuracy of high speed moving target tracking. The Application of this paper proved that the algorithm IMKTT combining improved Mean Shift algorithm with Kalman Filter, which makes use of the two advantages, can achieve good effectiveness on tracking high speed moving target.

Acknowledgements

This work was supported by Young University Teachers Training Plan of Shanghai Municipality under Grant No. ZZSD13008. We gratefully acknowledge and thank those who provided comments and suggestions. The anonymous reviewers and the editor of this paper are also acknowledged for their constructive comments and suggestions. I'd like to express my sincere thanks to all those who have lent me hands in the course of my writing this paper. First of all, I'd like to take this opportunity to show my sincere gratitude to my supervisor, Prof.Xu HuaHu, who has given me so much useful advices on my writing, and has tried his best to improve my paper. Secondly, I'd like to express my gratitude to my colleague, Dr.Gao HongHao who offered me references and information on time. Without their help, it would be much harder for me to finish my study and this paper.

References

- [1] H. Zhou, Y. Yuan and C. Shi, "Object tracking using SIFT features and mean shift", CVIU, New York, NY, USA, vol. 113, no. 3, (2009) March, pp. 345-352.
- [2] A. Yilmaz, "Object tracking by asymmetric kernel mean shift with automatic scale and orientation selection", Proc. IEEE Conf. Comp. Vision Pattern Recognition, (2007), pp. 1-6.
- [3] C. Yang, R. Duraiswami and L. Davis, "Efficient mean-shift tracking via a new similarity measure", IEEE Computer Society Conference on Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, (2005), pp. 176-183.
- [4] C. Xia, Y. Li, W. Gu and P. Zhang, "Canny edge detection based on Mean-Shift tracking method for determining nuclear window width", 3rd edition of Journal of Naval Aeronautical Engineering Institute, (2008) May 15, pp. 45-49.
- [5] X. Han, "Video target tracking algorithm and its application", Master Thesis, University of Science and Technology of China, (2010) May 20.
- [6] Y. Gong, "Research based on Image segmentation and support vector machine semi-automatic image tagging system", Master Thesis, Shandong University, (2008) April 5.
- [7] H. Zhang, J. Zhang, H. Yue and S. Haung, 67-71, No. 11, 27 Target Tracking Algorithm based on Camshift. Computer Engineer and Design (2006) June.
- [8] X. Yun and G. Xiao, "Adaptive Multi-feature Fusion Vessel Tracking Algorithm based on Camshift", Optical Engineer, vol. 38, no. 5, (2011) May, pp. 52-56.
- [9] H. Zhang and H. Li, "Target Tracking Algorithm based on Monte Carlo Method", Chinese Journal of Image and Graphics, vol. 13, no. 5, (2008), pp. 937-94.
- [10] X. Liu, F. Chang and H. Wang, "Video Objective Tracking Method based on improving Camshift Algorithm", Microcomputer Information, vol. 23, no. 7-3, (2007), pp. 297-298.
- [11] L. Wang, W. Hu and Y. Tan, "Summary of visual analysis of human movement", Journal of Computer, vol. 25, no. 3, (2002), pp. 225-23.
- [12] L. Wang, W. Hu and Y. Tan, "Identity-based Gait", Journal of Computer, vol. 26, no. 3, (2003), pp. 353-360.

Author



Bian MinJie, the author is a PhD candidate in graduate students in School of Computer Engineering and Science Shanghai University. His main research is about computer vision. He works at the Information Technology Office of Shanghai University during the study of PhD.

