

The Improvement of Mean-Shift Algorithm in Target Tracking

Danping Jia, Lifeng Zhang and Chunhua Li

Shenyang University of Technology, 110870, China
winnerjia@sohu.com

Abstract

The bandwidth of kernel function is invariance in Mean-Shift tracking process which led to the problem of actual process of tracking failure. To solve this question a method is proposed to combine object contour, RGB color histogram and Mean-Shift tracking algorithm. In each frame the contour and color feature of target object are extracted to as a model. Meanwhile the size of contour is used as the bandwidth for kernel function in the next frame. Based on this Method the accuracy of Mean-Shift tracking is improved and the error of Mean-Shift tracking is reduced.

Keywords: *Bilateral filtering; Contour extraction; Mean-Shift kernel function; Tracking target*

1. Introduction

Tracking moving target is an important area of research in the field of computer. In recent years, researches have proposed many algorithm of tracking target [1]. Among them, Mean-Shift algorithm is widely used due to its fast, easy to integrate other algorithms and robust for partial occlusion. However, the bandwidth size of the kernel function, based on non-parametric kernel density estimation fast matching algorithm, play an important role in Mean-Shift algorithm. Because, it affects not only the weight of pixel, but also reflects the size and shape of target. In general, the bandwidth of kernel function is defined when selecting target in the first frame, and remains unchanged during the entire trace. However, the fixed kernel function often cause to losing target when the size of target changes. For this, scholars had proposed some more meaningful conclusions and methods. Comanicu propose the method of adopting the increment of plus or minus 10% to revise the bandwidth of kernel function respectively. This method requires that three different sizes of kernel do Mean-Shift tracking calculations about three times independently. The current frame is used in three different sizes of kernel function three times independently Mean-Shift tracking calculations. Then, the kernel window of the larger Bhattacharyya coefficient will be chosen as the best window [2-6]. However, when the size of target increases, especially when the size of target exceeds the width of kernel window, this method fails. Bradski proposed method of basically unchanged moment. After each of Mean-Shift iteration has been done to a stable position, the second central moment of pixel is calculated pixel within the window slightly larger than the Mean-Shift window, which obtains conversion factor. But this calculation is a large time consume, which seriously affected the real time of the Mean-Shift algorithm [7-8]. Meanwhile, the existing improving method of Mean-Shift algorithm are usually based on color feature extraction, it is difficult to ensure the robustness of tracking algorithm. In order to make Mean-Shift algorithm is more suitable for intelligent vehicle tracking system, this paper proposes a method for combining the Mean-Shift algorithm with object contours, RGB color and Mean-Shift.

2. The Improvement of Overall Scheme

This paper proposes the method of combining bilateral filtering and Mean Shift in order to overcome the fixed kernel function often lead to loss target. When the tracking target is determined, bilateral filtering is used to extract the outline of object and mathematical model of tracked target is established in this contour area. Next, the center position of tracked target is calculated by Mean-Shift algorithm. Using the new center position of the object and the object contour region of previously determined extracts the outline of object by using bilateral filtering. After, it replaces the object contour region of the previously determined. Finally, a mathematical model of tracked object is created in this region. By the method of adjusting according to the current situation overcomes the problem that the kernel bandwidth invariance of traditional Mean-Shift tracking process causes loss target in the actual process. The overall scheme is shown in Figure 1.

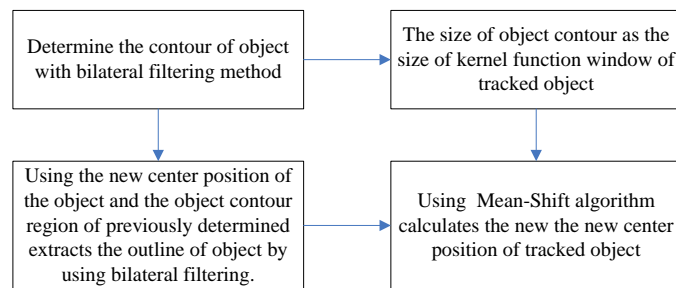


Figure 1. The Overall Scheme

3. Establish the Mathematical Model of Tracked Target

Mean Shift tracking algorithm is a semi-automatic algorithm. Tracked target is determined by manual or other identification algorithms in initial frame. Color histogram distribution of tracked target is as a mathematical model of tracked target and use the same method calculate the candidate target model in N-th frame. As the principle with the maximum similarity of two models, it makes the tracking window move along the direction of similarity fastest changing. Finally, locate the target. Therefore, establishing the mathematical model must be met two conditions: 1) The target has been detected and located in the initial frame. 2) The color characteristics of target will not suddenly change significantly.

This paper assumes that color distribution of target quantified as m-dimensional. Thence, the 256 gray levels for original each color component quantified as m-dimensional, which defines the function as $b(x): R^2 \rightarrow \{1 \dots m\}$ that the pixel values of the position x in the area of tracked target corresponds the index $b(x)$ of histogram.

Since the pixel closer from regional center position is more reliable and the pixel farther from regional center position is more likely background, kernel function is used to weight for each pixel in order to provide a model of tracked target robustness.

Set the center of tracked target is x_1 . Among them, there are n pixels, represented by $\{x_i\}_{1 \dots n}$. Then, the mathematical model formula of color histogram of tracked target is :

$$q_n = C \sum_{i=1}^n k\left(\frac{|x_0 - x_i|}{h}\right)^2 \delta[b(x_i) - u] \quad (1)$$

In the formula, K is the contour function of the kernel function; h is the kernel function; δ is the function of Konekes delta; The function of b is the map from location x pixel to color;

C is the normalized function, which can be exported by according to the condition $\sum q_n = 1$. So, the formula can be got:

$$C = \frac{1}{\sum_{i=1}^n k\left(\left|\frac{x_0-x_i}{h}\right|^2\right)} \quad (2)$$

Therefore, the tracked target can be represented by the formula (1). And it is used as a measure of whether the tracked candidate target model is tracked target model in subsequent frames or not.

4. Target Tracking Algorithm of Mean-Shift based on the Variable of Kernel Function Window for Bilateral Filtering Detection

4.1. The Method for Determining the Bandwidth of Mean-Shift Kernel Function based on Bilateral Edge Detection Filtering

The window width of kernel function determines the pixel quality of candidate targets. Then, the number of pixels of candidate targets affects the operation speed of tracking targets. Therefore, the method of using the bilateral filtering method to extract the contour of target in a video sequences can remove noise and get more edge detail. Meanwhile, since the pixel closer from regional center position is more reliable in the area of tracked target, kernel function is used to weight for each pixel in order to provide a model of tracked target robustness.

(1)The extracting contours process of bilateral filtering:

Step 1: Do convolution operation between the original image f and improved bilateral filter. The algorithm is as follows:

$$h(x) = k^{-1}(x) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\varepsilon) c(\varepsilon, x) s(T(\varepsilon), T(x)) d\varepsilon \quad (3)$$

Step 2: Because the gray at the edges of object changes rapidly, a first-order partial derivatives is used to obtain extreme value in gray changing rapidly to do singular point detecting. And, Use first-order differential operators to calculate the gradient magnitude and direction.

Step 3: Do non-maxima suppression for gradient amplitude.

Step 4: Determine the high threshold and calculate the low threshold by the method of improved adaptive threshold.

Step 5: Detect and connect edge.

(2)The method of determining window width: After the target edge is detected, a list of the closed loop is gotten and the regional R of tracked target is also determined. There are n pixels within the target edge region, representing abscissa by $\{x_i\}_{1..n}$ and ordinate by $\{y_i\}_{1..n}$. Set (x_i, y_i) that is belong to the pixel coordinate within regional R and the kernel of $k(x_i, y_i)$ that indicates the size of the region.

The kernel is taken as:

$$k(x_i - x_0, y_i - y_0) = \begin{cases} 1, & |x_i - x_0| \leq \frac{R_x}{2}, |y_i - y_0| \leq \frac{R_y}{2} \\ 0, & |x_i - x_0| > \frac{R_x}{2}, |y_i - y_0| > \frac{R_y}{2} \end{cases} \quad (4)$$

Where, $|x_i - x_0| \leq \frac{R_x}{2}, |y_i - y_0| \leq \frac{R_y}{2}$ and $|x_i - x_0| > \frac{R_x}{2}, |y_i - y_0| > \frac{R_y}{2}$ is to determine whether the point (x_i, y_i) is belong to regional R.

To determine whether the point (x_i, y_i) is belonging to regional R, the method is that lead a line from the judged point to one direction and calculates the number of the intersection between line and polygon. If the number is even or zero, the point is outside the polygon. If it is odd, the point is inside the polygon.

4.2. The Target Tracking of Mean-Shift Algorithm

In a video sequence, Mean-Shift algorithm is by constantly iterative search to find candidate target that is most similar with tracked target model.

The establishment of candidate target model is to find the location of candidate target by Mean-Shift. Then, the last time contour area of extracted by bilateral filtering and the location is used to establish the candidate target model.

$\{x_i\}_{1..n}$ is supposed the pixel in the candidate regional R for current image frame. The same kernel function is used in the center y that kernel bandwidth is h. Then, the mathematical model formula of color histogram of candidate target is :

$$p_u(y) = C \sum_{i=1}^{n_h} k \left(\left| \frac{y-x_i}{h} \right|^2 \right) \delta[b(x_i) - u] \quad (5)$$

Where, C is the normalized function, which can be exported by according to the condition $\sum q_n = 1$. So, the formula can be got:

$$C = \frac{1}{\sum_{i=1}^n k \left(\left| \frac{y-x_i}{h} \right|^2 \right)} \quad (6)$$

Should be noted that C does not depend on y, because x and y position of the pixel correspond to a given point in a given width of window. Therefore, C can be calculated by a given kernel and different width h

4.3. Similarity Metric Coefficient between Tracked Target and Candidate Target

Mean-Shift algorithm is by constantly iterative search to find candidate target that is most similar with tracked target model. In order to determine that the candidate target is tracked target, Bhattacharyya is introduced to measure the similarity between candidate target and tracked target [9].

Based on the case of Bhattacharyya, the correct position of the object to be tracked in the position of candidate target point can be effectively found. This method has a good effect for objects for different colors and textures, object obscured partially in the chaotic scene and the size of object changing from far to near.

Assuming color kernel histogram of target model $p = \{p_u\}_{u=1..n}$ is known. In the current frame, the center is in the color kernel histogram $q(y) = \{q_u(y)\}_{u=1..n}$ of target. The task of finding the target position in the current frame task is to find the most similar candidate target position y between the kernel histogram and the kernel histogram of target model. Here, the similarity between the kernel histogram of target model and the candidate target can be used to measure by Bhattacharyya coefficient. Bhattacharyya coefficient is defined as:

$$\hat{\rho}(y) = \rho[\hat{p}(y), \hat{q}] = \sum_{u=1}^m \sqrt{p_u(y)q_u} \quad (7)$$

When $\hat{\rho}(y)$ reaches its maximum, candidate target can be considered tracked target.

4.4. Position the Tracked Target in Subsequent Frames

There is one of the most important assumptions in Mean-Shift tracking algorithm that target will not change a lot in adjacent frames. Therefore, the position of tracked target in current frame must be within a local field in previous frame. Then, in the current frame, the new position y of tracked target can be searched in the area located at the coordinate y_0 .

The function (7) can be available by expanding for Taylor at the point $p_u(y_0)$:

$$\rho[\hat{p}(y), \hat{q}] = \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(y_0)q_u} + \frac{1}{2} \sum_{u=1}^m p_u(y) \sqrt{\frac{q_u}{p_u(y_0)}} \quad (8)$$

Substitute (5) into (8). Finish and get:

$$\rho[\hat{p}(y), \hat{q}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(y_0)q_u} + \frac{c}{2} \sum_{u=1}^m w_i k \left(\left| \frac{y-x_i}{h} \right|^2 \right) \quad (9)$$

In the formula:

$$w_i = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{q_u}{p_u(y_0)}} \quad (10)$$

For the second item of formula (9) on the right side, we can use Mean-Shift algorithm for optimization. In order to achieve tracking target, the second term of formula (9) must be maximized and the first item is independent of y position. The second term represents the density estimation of candidate targets at y position and the similarity of the target pattern, and use weight of the formula (10) to weight for data. Distributed object model is assumed for $\hat{q}(y) = \{q_u(y)\}_{1 \dots n}$. Target in the current frame is estimated y position. Mean-Shift tracking algorithm is as follows:

- 1) Initialize the current target frame of to y_0 .
- 2) Use bilateral filtering detection to extract the contour of tracked target.
- 3) Calculate distributed $\hat{p}(y_0) = \{p_u(y_0)\}_{u=1 \dots n}$ and estimate $\rho[\hat{p}(y), \hat{q}]$.
- 4) Calculate weight $\{w_i\}_{i=1 \dots n}$ according to the formula (10).
- 5) Based on mean shift vector, a new location y of candidate target can be calculated. Such as formula:

$$y = \frac{\sum_{i=1}^m x_i w_i k \left(\left| \frac{y_0 - x_i}{h} \right|^2 \right)}{\sum_{i=1}^m w_i k \left(\left| \frac{y_0 - x_i}{h} \right|^2 \right)} \quad (11)$$

- 6) When $\rho[\hat{p}(y), \hat{q}] < \rho[\hat{p}(y_0), \hat{q}]$, make $y \leftarrow (y_0 + y_1)/2$.
- 7) If $\|y_0 + y_1\| < \varepsilon$, Stop the cycle. Otherwise, make $y_0 \leftarrow y$, and back to step one. When the final cycle stops, y is the position of tracking target. The process of tracking is to apply to each frame of image sequences by above optimizing algorithm. Based on Bhattacharyya coefficient, find the target location of the best match. The program flowchart of the entire improved Mean-Shift algorithm is shown as Figure 2:

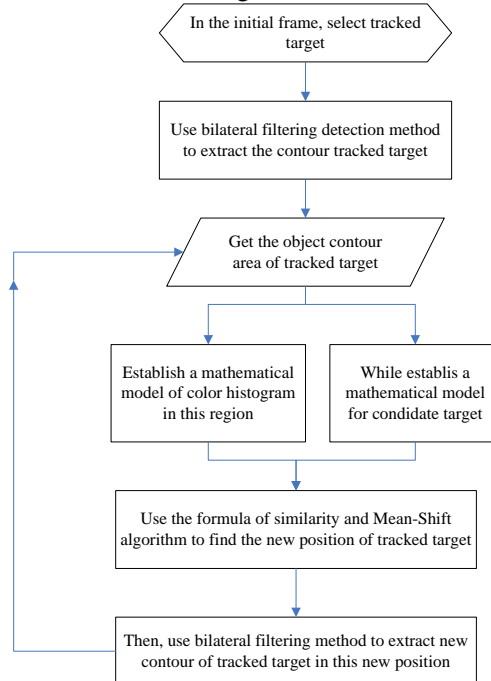


Figure 2. The Program Flowchart of the Entire Improved Mean-Shift Algorithm

5. Experimental Result

Experiment that a car in a real shot video had been used by Mean-Shift target tracking based on bilateral filtering detection for extracting edge is to verify the feasibility of this scheme. And it had been compared with the traditional Mean-Shift track. The entire experiment uses matlab to verify the effect.

The same car in the same video had been used these two kinds of methods to tracking. Then, the third frame, thirteen frame and forty-three frame from the video had been picked to compare the actual results (shown in Figure 3). Wherein, (a) used the method of improved Mean-Shift. The red region is the position of tracked target calculated by algorithm. (b) used the method of traditional Mean-Shift. The white frame is the position of tracked target calculated by algorithm.

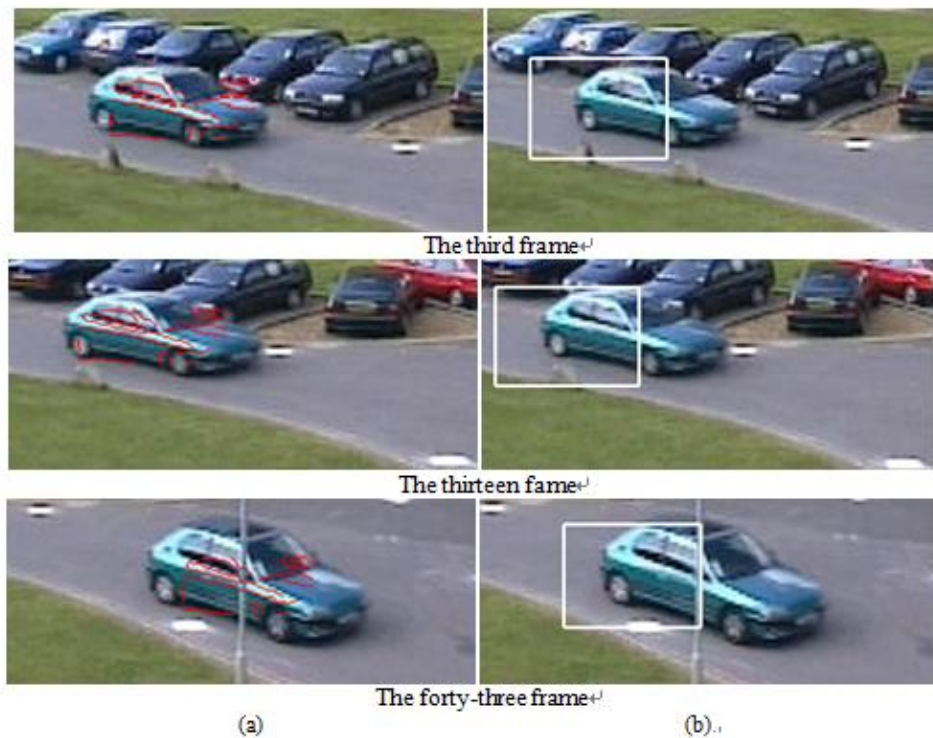


Figure 3. The Schematic of Comparison of Algorithms

As can be seen, improved Mean-Shift not only improves the accuracy of tracking, but also reduces the unnecessary region of pixel.

Also, the same car in the same video carried out 100 times and 10 times the target tracking experiments by using two methods respectively. If the target area calculated by algorithm beyond the actual area of tracked target, the tracking failed. Finally, algorithm computation time is obtained by the average computation time of 100 times successful tracking. And, the improved algorithm not only improves the accuracy, but also increases the speed as shown in Table 1.

Table 1. Performance Comparison Table

Method	100 times tracking accuracy (%)	150 times tracking accuracy (%)	Algorithm computation time(s)
--------	---------------------------------	---------------------------------	-------------------------------

Traditional algorithm	91.30	87.70	0.5338
Improvement algorithm	97.70	95.20	0.4046

6. Conclusions

In the smart car tracking system, Mean-Shift algorithm has completed the center of target tracking by optimizing the performance. The width of tracking window is an important factor to the performance of tracking. Since edge of the image focuses on the most information of the image, the width of Mean-Shift tracking window can be determined according to the edge of image. Correctly detecting the edge of target is premise of determining the width of tracking window accurately. Bilateral edge detection algorithm has a higher edge positioning accuracy. According to the simulation results, Bilateral edge detection algorithm combining with the Mean-Shift algorithm is a good way to achieve the goal of complex motion tracking status.

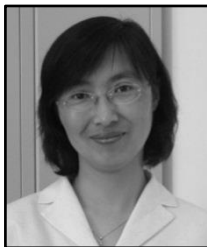
Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant No.61100091; the Technology Research project of Liaoning Province Education Administration under Grant No.L2013059; Research Fund for the Doctoral Program of Shenyang University of Technology. The authors would like to express the greatest appreciation and thanks to the reviewers for their details reviews and constructive comment, which have helped improve the quality of this paper.

References

- [1] Y. Cheng, "Pattern Analysis and Machine Intelligence", vol. 8, no. 17, (1995).
- [2] D. Comaniciu and V. Ramesh, "Mean Shift and Optimal Prediction for Efficient Object Tracking", Proceedings of the IEEE Int. Conf. on Image Processing, (2000) March 70-73; Canada.
- [3] D. Comaniciu, V. Ramesh and P. Meer, "Real-Time Tracking of Non-rigid Objects using Mean-Shift", Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition, (2000) October 142-149; Kyoto, Japan.
- [4] D. Comaniciu and P. Meer, "Pattern Analysis and Machine Intelligence", vol. 5, no. 24, (2002).
- [5] D. Comaniciu, V. Ramesh and P. Meer, "Pattern Analysis and Machine Intelligence", vol. 5, no. 25, (2003).
- [6] W. Ma, J. Cheng and M. Zhang, "Journal of Tsinghua University Science & Technology", vol. 1, no. 52, (2012).
- [7] G. R. Bradski, "Intel Technology Journal", 2nd Quarter, (1998).
- [8] J. Ning, L. Zhang and D. Zhang, "IET Computer Vision", vol. 1, no. 6, (2012).
- [9] T. Kailath, "IEEE Trans on Communication Technology", vol. 15, (1967).

Authors



Danping Jia, She received her M.Sc. in Instrument Science and Technology (2002) and PhD in Electrical Engineering (2009) from Shenyang University of Technology. Now she is the dean of electronic information engineering department of Shenyang University of Technology, national temperature measurement committee. Her current research interests include different aspects of photoelectric detection and Automatic control.



Lifeng Zhang, He received his B.Sc in Electronic and Information Engineering (2012) and will receive his M.Sc in Control Engineering (2015) from Shenyang University of Technology. Now he is a postgraduate student. Meanwhile he is an iOS engineer in Tapatak. His current research interests include image processing and intelligent terminal software.



Chunhua Li, She received her M.Sc. in Control Engineering (2004) and PhD in Instrument Science and Technology (2012) from Shenyang University of Technology. Now she is an associate professor of Shenyang university of technology, her current research interests include different aspects of motor controlling, nondestructive testing.