ISBN 978-952-5726-02-2 (Print), 978-952-5726-03-9 (CD-ROM)
Proceedings of the 2009 International Symposium on Information Processing (ISIP'09)
Huangshan, P. R. China, August 21-23, 2009, pp. 029-032

Moving Vehicle Tracking Based on Double Difference and CAMShift

Jie Xia[1,2], Jian Wu[1,2], Haitao Zhai[1,2], and Zhiming Cui [1,2]
[1] JiangSu Province Support Software Engineering R&D Center for Modern Information Technology Application in Enterprise, Suzhou 215104, China;
[2] The Institute of Intelligent Information Processing and Application, Soochow University, Suzhou 215006, China
Email: shadow_xj@163.com, szzmcui@suda.edu.cn

***Abstract*—CAMShift is a semi-automatic tracking algorithm, which needs tracking target to be calibrated manually, and can only track one single target at a time. Aiming at these problems, this paper proposes a vehicle tracking approach based on double difference method and CAMShift algorithm. This approach can automatically calibrate moving vehicles in traffic video and achieve multi-target tracking by using a multi-tracker of CAMShift algorithm. The experimental results show that the algorithm in this paper can solve automatic multi-target tracking problem in traffic video very good along with good stability.**

***Index Terms*—-double difference, CAMShift, multi-target, automatic tracking**

I. Introduction

The research object of moving target is the video image sequence. Compared with static single-frame images, image sequences contain more information, so effectively tracking interested target in video sequences is an important issue in computer vision field [1] [2]. In traffic video surveillance, real-time detection and multi-moving vehicle tracking is the basis of intelligent monitoring system, whose main task is to detect and track moving vehicles in video sequences in real time and mark it for further operations such as identification and classification. Real-time detection of the moving vehicle requires locating vehicles accurately and quickly without manual calibration, and with as little system resources consumption as possible, whose effect has a direct impact on the subsequent part of the system. By doing research on moving target detection and tracking in video sequence of static scenes, this paper proposes a multi-vehicle detection and tracking methods. Firstly, do motion detection by means of double-difference method in video sequences contained moving objects, then regard the detected moving vehicles target as region of interest, referred to as ROI, and finally use a CAMShift multi-tracker to do multi-vehicle tracking, thereby overcoming the defect of CAMShift algorithm.

II. Double Difference Method

Moving target detection is a prerequisite of moving targets tracking, but also the basic issues in computer vision, using motion detection can obtain many kinds of information of foreground objectives. The commonly used methods of motion detection [3] are background subtraction and frame difference method. Background subtraction obtains foreground by subtracting the background image from the current video frame. The key is how to obtain the original static background model; and because the background often dynamically changes with the light, movement and target in and out the scene, it needs an efficient strategy to maintain and update the background model. The commonly used strategy are median filter method, mixed-Gaussian method and so on, but they all have some shortcomings such as imprecise initial background, maintaining and updating difficulties of background model. While frame-difference method makes the use of the difference between the two current frames in video sequences or the difference between one current frame and a certain frame prior to it to extract motion region of a image, namely directly using the (t-1)th frame of the video frame as a background frame to the tth frame, so that it doesn't require initial background, but also avoid pace-time consumption of maintaining and updating the background model.

This paper adopts an improved frame-difference method: double difference method, also known as three-difference method [4], whose basic idea is shown as follow:

Difference image is obtained by absolute value between corresponding pixels of two sequential frames. Suppose $\{F_n\}$ is input image frames, so the nth difference image is defined as:

$$D_n(i,j) = \left|F_{n-1}(i,j) - F_n(i,j)\right| \quad (1)$$

Where $F_{n-1}(i,j)$ denotes the pixel value at (i,j) in the (n-1)th frame, $F_n(i,j)$ denotes pixel value at (i,j) in the nth frame, so the nth difference $D_n(i,j)$ is the absolute value between pixel values at (i,j) in the (n-1)th frame and nth frame.

In the definition of double-difference image, a flag will be used, which is showed as follow:

$$flag = \left(D_{n+1}(i,j) > T\right) \wedge \left(D_n(i,j) > T\right) \quad (2)$$

Where T is a given threshold. $D_{n+1}(i,j) > T$ denotes that difference value of the (n+1)th frame is larger than threshold T . $D_n(i,j) > T$ denotes that difference value of the nth frame is larger than threshold T . Therefore, when both difference values at coordinate (i,j) in two consecutive difference images are larger than a given threshold T, *flag* is 1.

Thus, the double-difference image DD_n is defined as follow:

$$DD_n(i,j) = \begin{cases} 1, & flag = 1 \\ 0, & others \end{cases} \quad (3)$$

Where if $flag = 1$, $DD_n(i,j) = 1$ denotes that both difference value images has detected pixels whose difference is larger than threshold T, and these points compose of the moving target to be tracked. That is, double-difference image is obtained by logical-and between two consecutive difference image. Because noise is not easy to repeat in two consecutive difference images, double-difference method can suppress noise better than general frame-difference method.

III. Automatic Tracking of Multi-vehicles

A. CAMShift Algorithm

CAMShift (continuously adaptive mean-Shift) is an extended motion tracking algorithm of Mean Shift algorithm [5] to the continuous images sequence, which is a non-parametric method based on the gradient of dynamic distribution of probability density function. The basic idea is to do Mean Shift operation to all the frames in video sequence, use mass center and size of search window obtained in previous frame as the initial value of search window in the next frame, and achieve the target tracking by iteration.

CAMShift algorithm mainly uses information of color probability distribution of moving targets in video images to achieve the tracking purpose [6] [7], which can effectively solve objective deformation and objective shelter problems with high computational efficiency. RGB color space is quite sensitive to brightness change. In order to reduce the impact of brightness change on tracking, CAMShift algorithm converts the image from RGB space to HSV space. Whereas in HSV space only H-component can represent color information, so it must calculate one-dimensional color histogram of H-component in the HSV space, then convert the original image into a color probability distribution image based on the obtained histogram, and finally adopt Mean Shift algorithm. The calculation process [8] is shown in Fig. 1:

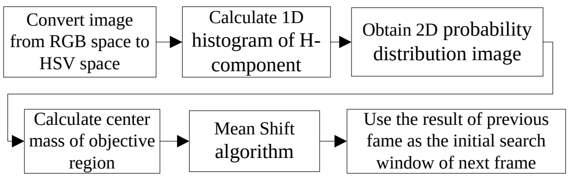


Figure 1. CAMShift algorithm calculation process

CAMShift algorithm can do real-time tracking to target of specific color, achieve very good tracking effect, and effectively remove noise and interference from image sequence.

B. Tracking Algorithm in This Paper

CAMShift algorithm use color probability models of target to track, but CAMShift is a semi-automatic tracking method, whose initial tracking objective should be manually calibrated, and just can track only one objective rather than multi-target automatically at a time, which makes its usefulness to be limited.

Aiming at these two problems, this paper firstly do real-time detection of multi-vehicle moving in video by means of double-difference method, also mark each ROI of moving vehicle to solve the manual calibration problem, and then start CAMShift tracker for each vehicles ROI separately in order to overcome the shortcoming that CAMShift algorithm can only track a single target at a time. Each tracker in this algorithm is in charge of its separate vehicle objective independently, and will no longer work until the targets disappear. Algorithm flow chart is shown in Fig. 2:

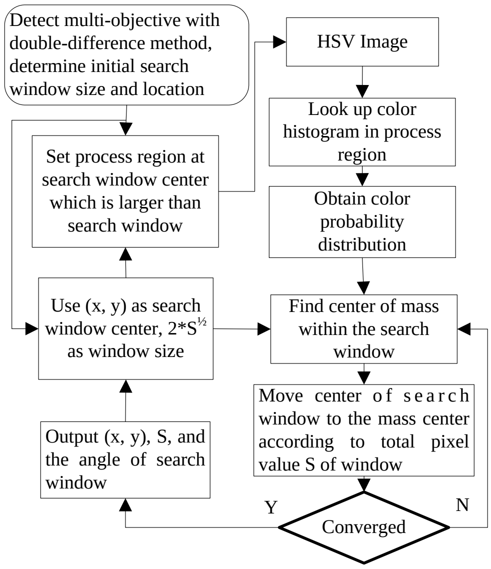


Figure 2. Flow chart of tracking algorithm in this paper

The multi-vehicle tracking process in this paper is described as follows:

1) Detect all the moving vehicles in video with double-difference method, and make respective tracking region and search window as the initial ROI and search window.

2) Use CAMShift tracker to each tracking area (search window) as follows:

a) Calculate zeroth moment and the first moment in order to get mass center of search window

$$M_{00} = \sum_x \sum_y I(x,y) \quad (4)$$

$$M_{01} = \sum_x \sum_y yI(x,y) \quad (5)$$

$$M_{10} = \sum_x \sum_y xI(x,y) \quad (6)$$

Where $I(x,y)$ is pixel probability value at coordinate (x,y), the variation range of x and y is in the scope of search window. The zeroth moment denotes the area of objective in the color probability distribution images, so the search window size can be set as a function of zeroth moment to adaptively adjust search window size in the process of tracking.

b) Calculate the mass center of search window. Suppose the coordinate of mass center is $\left(x_{c,}y_{c}\right)$ *:*

$$x_c = \frac{Z_{10}}{Z_{00}} \text{ , } y_c = \frac{Z_{01}}{Z_{00}} \tag{7}$$

At the same time, calculate secondary moment on the basis of *a),b)*, obtain long-axis, short-axis and direction of the target. The secondary moment is:

$$M_{20} = \sum_x \sum_y x^2 I(x,y) \tag{8}$$

$$M_{02} = \sum_x \sum_y y^2 I(x,y) \tag{9}$$

$$M_{11} = \sum_x \sum_y xyI(x,y) \tag{10}$$

The direction of target is:

$$\theta = \frac{1}{2}\arctan\left[\frac{2\left(\frac{Z_{11}}{Z_{00}} - x_c y_c\right)}{\left(\frac{Z_{20}}{Z_{00}} - x_c^2\right) - \left(\frac{Z_{02}}{Z_{20}} - y_c^2\right)}\right] \tag{11}$$

Where suppose

$$a = \frac{Z_{20}}{Z_{00}} - x_c^2, b = \frac{Z_{11}}{Z_{00}} - x_c y_c, c = \frac{Z_{02}}{Z_{00}} - y_c^2 \tag{12}$$

So the long axis l , short-axis w of tracking objective can be presented as:

$$l = \sqrt{\frac{(a+c) + \sqrt{b^2 + (a-c)^2}}{2}} \tag{13}$$

$$w = \sqrt{\frac{(a+c) - \sqrt{b^2 + (a-c)^2}}{2}} \tag{14}$$

c) Reset the size S of search window as the function of color probability distribution of the above search window.

d) Repeat a, b, c steps until convergence (change of mass center is less than a given threshold).

3) Repeat 1), 2) steps until the end of image sequence

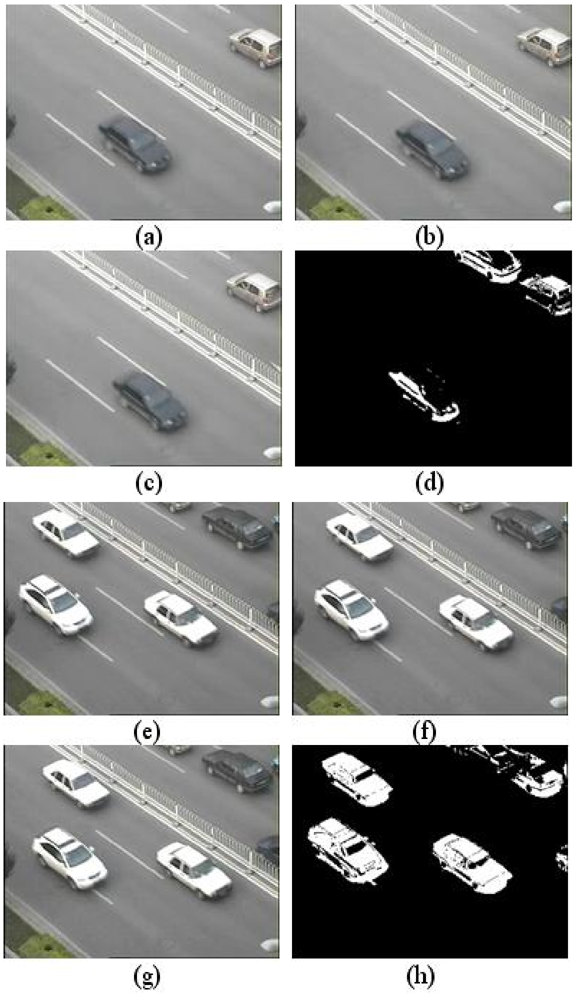
IV. Experimental Results and Analysis

A. Moving Vehicle Detection

Fig.3 shows the contrast effect of double difference method and background subtraction.

In Fig. 3, (d) and (h) are the effect drawings by means of double-difference method, showing that the double-difference method has a good effect of detecting moving target. We can find two vehicles appear in the upper right corner in (d), that's because before the 90th frame, there is a car stopping for a while and start again immediately, and result in a ghost appeared in foreground image.

Double difference method indeed has the limitation to solve this problem, can't eliminate ghost in real time; but background subtraction method even with good background model, it can't achieve real-time ghost elimination due to the restriction of model updating coefficient (complete updating the current template needs a certain amount of time and space consumption). Therefore, the double-difference method with less time and space of consumption has an advantage over background subtraction method. In addition, double-difference method use only three sequential frames to determine the foreground, while background subtraction method use background model gained from a lot of frames to determine foreground image, with slower updating speed(due to different model updating coefficient). In contrast, double-difference method has a faster real-time updating speed.



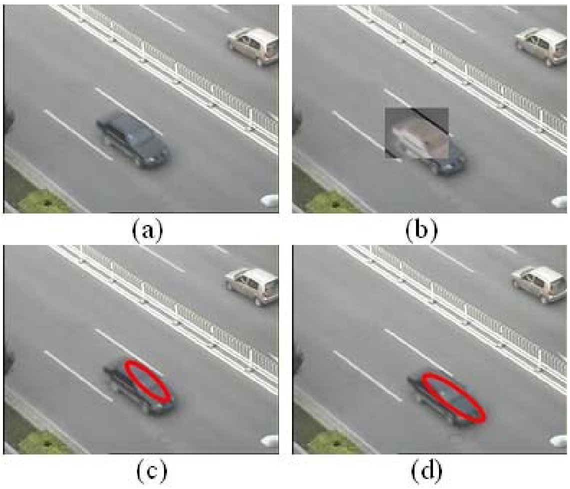
(a)-(c) are respectively the 89th, 90th, 91st (d) is the 90th foreground gained from the 89th, 90th, 91st by using double-difference method (e)-(g) are respectively the 239th, 240th, 241st frame of video (h) is the 240th foreground gained from 239th,240th,241st by using double-difference method

Figure 3. Contrast effect of double differenceMethod and background subtraction

B. Moving Vehicle Tracking

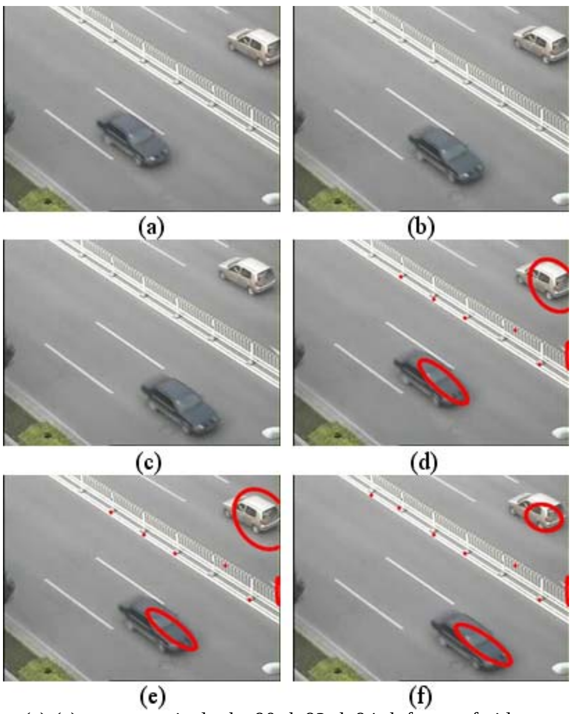
General CAMShift can only track a single moving target at a time, and must pre-calibrate a moving target manually, initialize the search window of tracking target, and then track the selected target. The process is shown in Fig. 4:

After real-time detection of multi-target, the algorithm start a vehicle CAMShift tracker to each detected moving vehicle respectively, which can track every moving vehicle independently, so that all moving vehicles can be tracked well. Figure 5 shows the tracking process:



(a) is the 90-th frame in the video sequence, (b) is manual calibration of tracking target, (c) (d) are respectively the subsequent frames of the calibrated target in tracking process

Figure 4. Single target tracking with single CAMShift Tracker



(a)-(c) are respectively the 90-th,92-th,94-th frame of video sequence (d)-(f) are the corresponding effect drawings of the aforesaid frames.

Figure 5 Effect drawings of multi-target tracking with CAMShift multi-tracker

In the tracking process of (d) - (f), the shape of mark ellipses change continuously according to the search window's changing, but never leave the tracking vehicle. With double-difference method and multi-tracker CAMShift, this paper can achieve good tracking effect of real-time multi-vehicle tracking without manual calibration.

V. Conclusion and Future Work

Real-time moving vehicle tracking is the basic subject of Intelligent Transportation Surveillance and Control System and requires auto-calibration, real-time tracking, and minimal system consumption in practical application. In this paper, a moving vehicle tracking method based on double difference and CAMShift was proposed, to overcome the shortcoming of algorithm CAMShift such as manual calibration of tracking target and tracking a single target at a time, realizing real-time multi-vehicles tracking in traffic video. Experiments on the research of video traffic have a good scalability, which is applied to other occasions of multi-target tracking. In addition, how to differentiate tracking targets with very similar color histogram distribution, it will be the focus of future research.

Acknowledgment

This research was partially supported by the National Natural Science Foundation of China(60673092), the Project of Jiangsu Key Laboratory of Computer Information Processing Technology and the Beforehand Research Foundation of Soochow University.

References

[1] F. J. Huang and T. Chen, "Tracking of Multiple Faces for Human-Computer Interfaces and Virtual Environments", IEEE Intl. Conf. on Multimedia and Expo, New York ,2000.

[2] J Li, CS Chua, YK Ho. "Color based multiple people tracking". In:Seventh International Conference onControl, Automation, Robotics And Vision(ICARCV'02), Singapore, 2002, pp.309 –314.

[3] DAI Ke-xue， LI Guo-hui， TU Dan， YUAN Jian. "Prospects and Current Studies on Background Subtraction Techniques for Moving 0bjects Detection from Surveillance Video". Journal of Image and Graphics. 2006, Vol.11, pp.919–927

[4] Collins R, Lipton A J, Kanade T, et a1. "A system for video surveillance and monitoring: VSAM final report" . Technical Report: CMU-RI-TR-00,Carnegie Melon University, Pittsburgh, Peen, America, 2000.

[5] D Comaniciu, V Ramesh, P Meer. "Kernel-Based Object Tracking". IEEE Transaction on PAMI, 2003, Vol.25, pp.564–577

[6] MA Lei, TIAN Yuan, SU Hong qi. "A Small Target Tracking Algorithm Based on Monte Carlo Method". Journal of Image and Graphics.2008, Vol.13, pp.445–449

[7] ZHAO Yi， MU Zhi Chun, LIU Ke. "Real-time Ear Tracking Based on Both Skin-color and Contour Information". Journal of Image and Graphics. 2006,Vol.7, pp.949–953.

[8] Bradski, G.R., "Computer Vision Face Tracking for Use in a Perceptual User Interface," IEEE Workshop on Applications of Computer Vision, Princeton, NJ, 1998, pp.214-219