OPTOELECTRONICS AND ADVANCED MATERIALS – RAPID COMMUNICATIONS Vol. 4, No. 12, December 2010, p. 2172 - 2177

Improvement of moving object segmentation accuracy in video sequences based on adaptive symmetrical difference and background subtraction

J. LI, J. XU[a], X. SANG[*], Y. WANG[b], B. YAN, C. YU
Key Laboratory of Information Photonics and Optical Communications (Beijing University of Posts and Telecommunications), Ministry of Educations, Post box 72(BUPT), Beijing 100876, China
[a]China Civil Airport Equipment Co. Ltd, Beijing, 100621, China
[b]Institute of Forensic Science, Ministry of Public Security, Beijing, 100038, China

The segmentation accuracy of moving object in video sequences is usually influenced by illumination variation, noise, shadow, and ghost after the initially static background objects were removed. A novel algorithm combining adaptive symmetrical difference with background subtraction is proposed. The moving object mask is detected through the adaptive symmetrical difference, and points between any pairs on the contour of the mask as motion region are determined. And then, the adaptive background subtraction is carried out in the acquired region and points as motion region are labeled to extract accurate moving object. Morphological operation and shadow cancellation are adopted to refine the result. Experimental results show that the algorithm is robust and effective in improving the segmentation accuracy.

(Received April 20, 2010; accepted November 29, 2010)

Keywords: Moving object segmentation, Symmetrical difference, Background subtraction, Adaptive threshold

1. Introduction

Moving object segmentation in video sequences has attracted much attention in recent years because of its important role in various camera video applications, such as video surveillance, traffic monitoring and people tracking. It is important to detect the moving target accurately and effectively for further tracking, recognition and content analysis. Conventional segmentation algorithms can be divided into two classes. One class is based on spatial homogeneity [1,2], which result in the promising output. However, the computation is too complex and heavy to be unsuitable to real-time applications. The other class utilizes change detection as segmentation standard to extract moving object [3-11]. The typical approaches include frame difference [3-5], background subtraction [6-14] and optical flow [14]. The frame difference method based on absolute pixel-wise difference between two consecutive frames to distinguish static objects from moving objects, is very suitable to dynamic scenes and simple implementation. However, it is subject to noise, shadow and the foreground aperture problem. Background subtraction extracting moving objects through comparing current frame and background frame is effective in getting a complete contour and motion information, but it sensitive to noise, shadow and changes due to illumination. The optical flow method performs well in providing motion information, but it is unsuitable for real-time applications because of its complex and heavy computation.

Here, an algorithm combining adaptive symmetrical difference with background subtraction is presented, which select the optimal threshold for symmetrical difference and background subtraction separately. The accurate motion region is extract by marking points between arbitrary two points on the contour of the moving object mask. An optimal updating approach is adopted, which can effectively and efficiently model background and update background in time.

2. The segmentation algorithm of moving objects

As shown in Fig. 1, the segmentation algorithm mainly includes four parts: symmetrical difference, background subtraction, adaptive threshold, shadow cancellation. Some morphological operations are adopted in some procedures so as to refine the result. The performance of the symmetrical difference in excluding deterioration effects of the change of illumination and noise exceeds the background subtraction, while the latter can get more effective motion information than the former. The symmetrical difference is used to remove the noise and get the rough motion region firstly, and the motion information is obtained through background subtraction. In order to simplify the process, we introduce adaptive threshold for the two means respectively. And due to the sensitivity of the both methods to the shadow, shadow cancellation based on HSV color space is carried out to get the accurate moving objects.

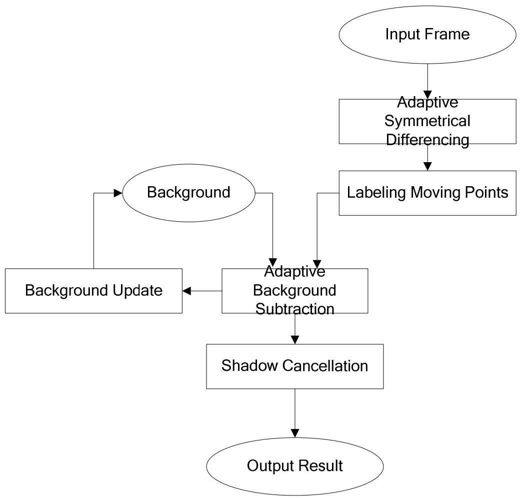


Fig. 1. The flow chart of diagram.

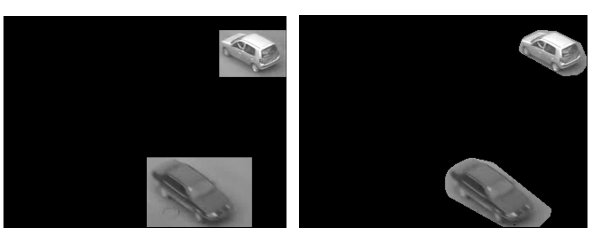
Conventional frame difference method usually mistakes noise for moving objects and suffers from the foreground aperture problem [3]. It is often replaced with symmetrical difference between three consecutive frames [4], and its acquired mask can eliminate the deterioration effect of noise. However, some holes still exist inside the movement area and the contours of the objects are not closed entirely. To get the complete motion profiles, we adopt an algorithm based on the view that the fractures of the contours can be found in the objects' edges paralleling with the motion direction, instead of exacting enclosing rectangles of moving objects. The process is as follows:

(1). Ragged edges of moving objects are extracted based on the frame difference mask, and the area of the regions surrounded by each contour is computed.

(2). If the area of one contour is lager than the given threshold, points between arbitrary two points on the contour are taken as motion region, or else the region is labeled as noise and is eliminated.

(3). Median filter operation is carried out to smooth the edge of the motion region.

The related results are shown in Fig. 2. By comparing two images, we can see the more precise motion regions are acquired compared with the enclosing rectangles.



a b

Fig. 2. The motion region detected by symmetrical difference, (a): the enclosing rectangle of moving objects (b): the motion region detected through the above algorithm.

For background subtraction, background extraction and updating are two critical procedures which influence on the accuracy of the result. The first step of background subtraction is to acquire an exact background image. The simplest method is to select a scene without any moving objects, but it is unpractical for most video applications. In recent years, two main kinds of methods have been investigated to extract an effective background image. One is based on the mathematical models, while the other one is to select appropriate pixel value from a series of preliminary frames as background according to some hypothesizes. The mixture of Gaussians method which models each pixel's intensity by a mixture of 3 to 5 Gaussians is a representative sample [6, 7] . It is able to acquire a correct simple background image, but it can not deal with the complex background and the computation is heavy. A statistical approach based on frame difference is used, which assumed that pixel-wise of moving objects changed faster than that of background [8]. A median method was proposed, which assumed that background should be observed for at least half of the time in training phase [9]. However, it is not actual in some applications. Here, we present an algorithm making use of symmetrical difference and region growth to remove moving objects and labeling the rest points as background. The result is given in Fig. 3 (e). We can see our algorithm is better than the statistical approach and the median method, which is almost equal to the result obtained through the mixture of Gaussians method. Our algorithm can be outlined as follows:

(1). Initiate the background buffer BG with -1.

(2). The region growth algorithm is adopted to the frame difference mask $\mathrm{FDM}_{i,j}(k)$ obtained among the four circumjacent pixels.

$$\mathrm{FDM}_{i,j}(k)=\begin{cases}1 & |I_{i\pm1,j\pm1}(k)-I_{i,j}(k)|<RT\\ 0 & \text{otherwise}\end{cases} \qquad (1)$$

where RT is the experimental threshold.

(3). Extract the enclosing rectangle of the frame difference mask after region growth. The pixels inside the rectangle are appointed with 1, and the rest remain 0.

(4). According to the acquired background, pixels of the current frame outside the rectangle are used to model the background.

$$\mathrm{BG}_{i,j}=\begin{cases}I_{i,j}(k) & \mathrm{BG}_{i,j}=\text{-}1\ \&\ \mathrm{FDM}_{i,j}(k)=0\\ \mathrm{BG}_{i-1,j-1} & \text{otherwise}\end{cases} \qquad (2)$$

(5). The count of the background pixels is calculated. If the count is equal to the size of the frame, the work of modeling background image is completed, otherwise return to step 2.

The background images obtained through Pixel Intensity Classification, Median Method, and our algorithm are given in Fig. 3.

a

b

c

d

e

Fig. 3. The backgrounds acquired by methods of (a) the original image, (b) mixture of Gaussians method, (c) Pixel Intensity Classification, (d) Median Method and (e) our algorithm.

After making sure the model mechanism of background, we can extract the background mask in the motion region obtained by frame difference through background subtraction for the following process.

$$BM_{i,j}=\begin{cases}1 & |F_{i,j}\text{-}BG_{i,j}|>BT\&FDM_{i,j}=1\\0 & \text{otherwise}\end{cases}\tag{3}$$

where BT is the adaptive threshold.

To get complete motion information, we mark the points between any pairs on the contour of the BM as the moving object. To adapt to the illumination variation and the change of objects in the background, the background model should be updated in time. Blind updating and selective updating are two updating mechanisms adopted in many algorithms. Blind updating can be adapted to the change of background, but it also adds the intensity value of the noise and moving objects to the background, which may result in the misjudgments in following operation. It is critical for selective updating to decide the given pixel belongs to the foreground or the background. In our algorithm, if the counterpart of current frame is located in the inner side of region, the point in the background is not updated according to the accurate motion region obtained after background subtraction, otherwise it will be replaced with the intensity value of the corresponding point of the current frame.

The threshold is a critical to distinguish moving objects from background for both symmetrical difference and background subtraction, which should be adaptive in real applications. Several relevant threshold decision algorithms have been proposed [15-17]. By compare them, an optimal method is achieved for symmetrical difference and background subtraction separately. Ostu method based on variance is used for symmetrical difference, and the background adaptation method is utilized for background subtraction, which can eliminate the deterioration effect of illumination variation. Experimental results are shown in Fig. 4.

a

b

c

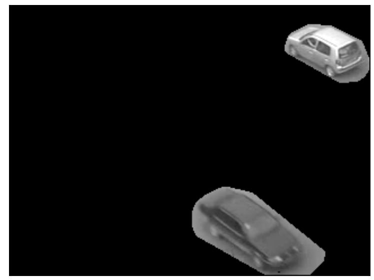
Fig. 4. The background acquired (a) Ostu Method, (b) Ostu Method based on variance, and (c) Background Adaptation Method.

Both symmetrical difference and background subtraction are not effective in distinguishing moving cast shadows from moving objects because of the roughly same visual characteristics between them, so the motion region usually include the shadows. We adopt HSV color space rather than RBG based on the assumption that the moving cast shadow has similar hue and saturation, but lower intensity with the background covered by it [18]

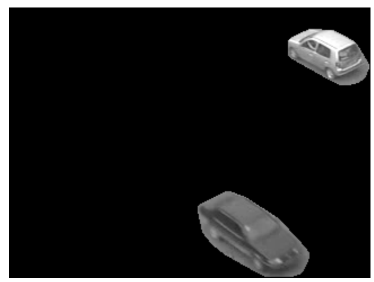
3. Experimental results

To verify the algorithms proposed above, we test them on the PC with Core™2 Duo CPU (2.1GHz), 2G DDR2, Window XP operation system and Visual Studio 2005. Experimental results of a traffic video surveillance and the hall monitor in different steps are listed in Fig. 5 and Fig. 6. In addition, the results are compared with the results obtained by frame difference in Ref. [5] and background subtraction in Ref. [6]. We can see that our algorithms can effectively eliminate the deteriorations effect of noise and shadow for both outdoor and indoor video surveillance.

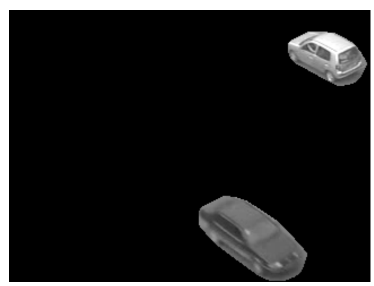
a



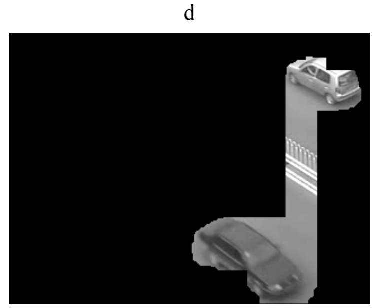
b



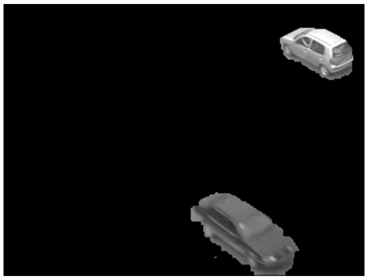
c



d



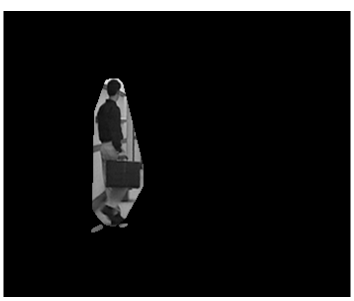
e



f

Fig. 5. The results of moving objects of traffic video surveillance in different processes and by different methods, (a): the original image (b): the result after symmetrical difference, (c): the result after background subtraction (d): the final result, (e): the result by the frame difference method in Ref. [5], (f): the result by the back ground subtraction method in Ref. [6].

a

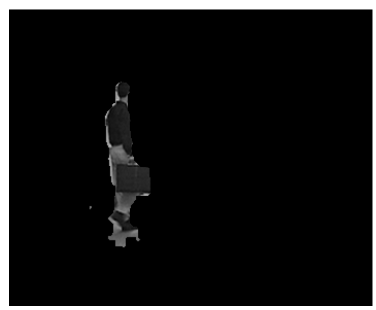


b

c

d

e



f

Fig. 6. The results of moving objects of hall monitor in different processes and by different methods. (a): the original image, (b): the result after symmetrical difference, (c): the result after background subtraction, (d): the final result, (e): the result by the frame difference method in Ref. [5], (f): the result by the back ground subtraction method in Ref. [6].

4. Conclusion

We propose a comprehensive algorithm combining adaptive symmetrical difference with background subtraction. It detects complete motion region by drawing line between all pairs on the contour of the mask acquired through adaptive symmetrical difference. The adaptive background subtraction in the last step is used to extract the accurate movement area. Morphological operation and shadow cancellation are adopted to refine the result. Experimental results show that t our algorithms can effectively eliminate the deterioration effects of noise and shadow for both outdoor and indoor video surveillance.

Acknowledgement

The research was supported by the Fundamental Research Funds for the Central Universities under no.2009RC0414.

References

[1] A. Prati, I. Mikic, M. Trivedi, R. Cucchiara, IEEE Transactions on Pattern Analysis and Machine Intelligence **25**, 918 (2003).
[2] S. Nadimi, B. Bhanu, IEEE Transactions on Pattern Analysis and Machine Intelligence **26**, 1079 (2004).
[3] X. Zheng, Y. Zhao, N. Li, H. Wu, in Proceedings of the 2009 International Conference on Embedded Software and Systems, pp. 541-543, July 13-16, 2009, Las Vegas Nevada, USA.
[4] S. Wei, M. Li, International Workshop on Intelligent Systems and Applications, pp. 1-4, 23-24 May, 2009, Wuhan, China.
[5] Y. Zhao, Z. Zhang, Z. Gao, Proceedings Elmar-2004 46th International Symposium Electronics in Marine, pp. 585-590, 16- June 18, 2004, Zadar, Croatia.
[6] H. Han, Z. Wang, J. Liu et al. Proceedings of IEEE Intelligent Transportation Systems, pp. 729-724, 12-15 Oct. 2003, Shanghai, China.
[7] C. Stauffer, W. E. L. Grimson. IEEE Transaction on Pattern Analysis and Machine Intelligence **22**, 747 (2000).
[8] S. Chien, S. Ma, L. Chen. IEEE Transactions on Circuits and Systems for Video Technology **12**, 577 (2002).
[9] B. Gloyer, H. K. Aghajan, K. Y. Siu, T. Kailath, Proc. SPIE **2421**, 173−180 (1995).
[10] W. Wang, J. Yang, W. Gao, IEEE Transactions on Circuits and Systems for Video Technology **18**, 670 (2008).
[11] R. Zhang, S. Zhang, S. Yu, IEEE Transactions on Consumer Electronics **53**, 1177 (2007).
[12] C. R. Jung, IEEE Transaction on Multimedia **11**, 571 (2009).
[13] S. Chien, Y. Huang et al. IEEE Transactions on Multimedia **6**, 732 (2004).
[14] E. Memin, P. Dence, IEEE Transaction on Image Processing **7**, 703 (1998).
[15] N. Ostu, IEEE Transaction on Systems, Man and Cybement **9**, 62 (1979).
[16] L. Chen, B. Zou, Application Research of Computers **25**, 488 (2008).
[17] Y. Gao, M. Leung, S. Hui, IEEE Transactions on System, Man and Cybernetics **33**, 407 (2003).
[18] P. Kumar, K. Sengupta, A. Lee, Proceedings of IEEE International Conference on Intelligent Transportation Systems, pp. 100-105, 3-6 Sept. 2002, Singapore.

*Corresponding author: xzsang@bupt.edu.cn