Accurate Moving Target Detection Based on Background Subtraction and SUSAN

Jun Zhang, Shukui Xu, Kuihua Huang, and Tingjin Luo

***Abstract*—Moving target detection is a fundamental problem in computer vision. Most target detection methods only detect the rough area of the target. However, there are also many applications (e.g., advanced video surveillance system, hybrid video camera system, etc.) need to detect the accurate target area for analyzing the movement or behavior. These applications always product noisy video frames and will lead to high false rates using previous method. In this paper, a two-step accurate moving target detection method is proposed. This method obtains a rough target area using a color background subtraction method with a feedback to background estimation, followed by a modified SUSAN method to estimate the accurate target edge. The modified SUSAN method uses the grads magnitudes to achieve an adaptive threshold and the gray barycenter criterion for denoising. Experimental results demonstrate the proposed method is effective for accurate moving target detection in different noise level video frames compare to several competitive methods proposed in the literature.**

***Index Terms*—Moving target detection, color background subtraction, SUSAN edge detection**

I. INTRODUCTION

Moving targets detection is a fundamental problem in many computer vision applications, such as intelligent video surveillance systems [1], modern driver assistance systems [2] and hybrid motion deblurring systems [3]. Most video surveillance applications only detect the rough area of moving targets. In some advanced video surveillance systems, rough area estimation cannot suffice the requirement of activity analysis of moving targets[4]. Accurate targets need to be estimated for analyzing the activity of targets. In hybrid motion deblurring systems, accurate moving targets detection is one of the most important steps of objects motion deblurring.

Moving target detection is one of the most challenging issues in computer vision. Most frequently used target detection methods are Temporal Difference method. Average/Median Filtering, W4 method, Gaussian Mixture Model and feature based detection [5]. Like these methods, in this paper, we only consider case of stationary camera working in surveillance systems. Many methods based on background subtraction have been proposed to detect moving targets in video surveillance system. Temporal Difference method use difference between neighbor frames or between current frame and several frames before current frame[6]. Gaussian Mixture Model is used to estimate the complicated background such as outdoor background with trees [7]. But this kind of background estimation method has a high temporal complexity. Optical flow estimates the motion vector of each pixel from one image to another. However, optical flow is not an exact reflection of motion field but an explanation of illumination change. Therefore, it is not rigorous to perform the segmentation with the optical flow information only [8]. Most previous methods for moving targets detection only attempted to find rough area of targets. Noise and holes appeared in the results. For advance applications, the accurate area of moving targets must be found out. Haag Michael combined edge element and optical flow for 3D-Model-Based Vehicle Tracking in Traffic Image Sequences [9]. Due to the limitation of optical flow, in this paper we combine the background estimation and edge detection method to estimate the accurate area of the moving targets.

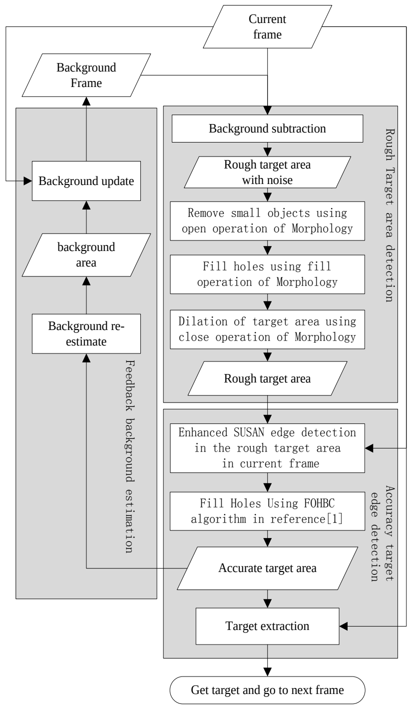


Fig. 1. Two-step accurate target detection framework

Assuming the target area has a clear edge between the background and target. Edge detection method estimates the edge of objects in the frame which contains moving targets. For fast moving targets, the exposure time of camera must be very short to freeze the target in image. A high gain has to be set to get a normal exposure image. This will make the image much noisy. Derivative based detectors such as Canny [10] and Sobel have to estimate the partial derivative which is

Manuscript received May 10, 2012; revised June 12, 2012. This work was supported partially by the National Natural Science Foundation of China under Grant 61170159, Grant 60902093.

The authors are with National University of Defense Technology, College of Information System and Management, ChangSha, Hunan, China (e-mail: zhangjun1975@nudt.edu.cn, xskgfkd@163.com).

very sensitive to the noisy images. In the case of high gain images, derivative based detectors could not detect the edge effectively. Non derivative based detectors, SUSAN detector [11] use the gray-scale features of images to detect the edge of target. SUSAN principle has a prominent advantage that it is not sensitive to local noise, and especially it has a very good inhibitory effect for Gaussian noise [12].

Aiming at detecting the accurate moving target in noisy video frames, we propose a two-step detection method. In the first step, we use the color background subtraction method and morphology method to detect the rough area of target. In the second step, a modified SUSAN method is proposed to detect the accurate edge of target in the rough area. In the case of noisy video frames, some pixels are recognized as background, holes appear in the estimation results. Holes need to be filled to get the whole target area. After filling the holes, we get the accurate area of moving target. The background area is returned to background estimation as a feedback to update the background. The framework of the proposed method in this paper is shown in Fig. 1.

II. Rough Target Area Detection Based on Background Subtraction

As most direct method used for moving target detection, background subtraction uses the difference between the background frame and current frame to detect the target area. Target area needs to meet the following formula:

$$\left| I_t(x,y) - I_b(x,y) \right| > T_a \tag{1}$$

where It(x,y) is the current frame and Ib(x,y) is the background frame. Ta is a threshold defined in advance. Background frame estimation is a key point in the background subtraction. If the background is not estimated accurately, the result of background subtraction will contain an overlapped target area (figure 2-e) or separate target area (figure 2-f). These situations cause high level process can't work by rule and line. We assume that the first frame does not contain a target; the initialization of the background estimation can be initialized as the first frame. At the end of accurate target detection, we use the background area as a feedback to update the background estimation as shown in Fig. 1.

To get the real background and denoise the background, we propose a feedback background estimation framework which returns the background area at the end of accurate target detection and updates the background frame using the following formula:

$$I_b^{new}(x,y) = \alpha * I_b^{old}(x,y) + (1-\alpha) * I_t(x,y) * (1 - M(x,y)) \tag{2}$$

a. Real background b. background with target c. foreground

d. background subtraction e. overlapped target area f. separate target area

Fig. 2. Target detection based on background subtraction

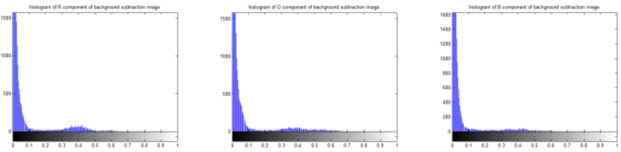
With α is a feedback coefficient which determines the updating speed of background frame. $M(x,y)$ is a binary mask which comes from the accurate target detection result.

$$M(x,y) = \begin{cases} 0 & (x,y) \in \text{background area} \\ 1 & (x,y) \in \text{target area} \end{cases} \tag{3}$$

Most background subtraction converts the true color image RGB to the gray scale intensity image by forming a weighted sum of the R, G, and B components, for example,

$$I = 0.29 * R + 0.58 * G + 0.13 * B \tag{4}$$

A uniform threshold is set to get the target area in gray scale image. In the process of RGB to gray scale conversion, the energy in each component is not distributed equally. The difference in certain component may be attenuated such as B component in formula (4). To avoid this situation, we use three thresholds for each component to detect the target area respectively. The threshold for each component is computed by finding the first inflexion of energy distribution graph.



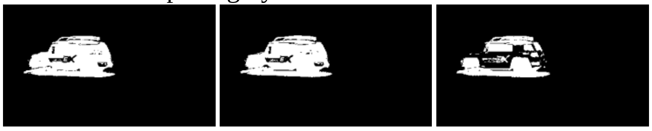
Energy distribution in R component Energy distribution in G component Energy distribution in B component

Fig. 3. Threshold in each component by finding the first inflexion in energy distribution graph

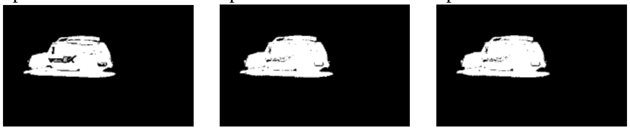
We apply the threshold to formula (1) to get background subtraction results in each component shown in Fig. 4-a to figure 4-c. We use a binary OR operator to combine the three target areas in each component shown in figure 4-e. The result covers more area than the result of using gray scale image. Isolated points in the result are possible noise so we use open operator to remove the isolated points from the background subtraction result. Holes in the result are filled using fill operator. After that, close operator are used to make the target area overlay the whole target. Results are shown in Fig. 4-f to Fig. 4-h.

III. Accurate Target Detection Based on Enhanced SUSAN Edge Detector

SUSAN stands for "Smallest Univalue Segment Assimilating Nucleus" which is a non-derivative based detector due to S.M. Smith [11]. The idea behind SUSAN detector is to use a pixel's similarity to its neighbors gray values as the classification criteria. SUSAN detector uses a quasi-circular template which moves in the image. The gray value of each image pixel within the template is compared with the center pixel gray value.



a. Background subtraction in R component | b. Background subtraction in G component | c. Background subtraction in B component



d. Background subtraction in grayscale image | e. Result combined by RGB components' result | f. Remove small object using open operator

g. Fill holes in target area | h. Dilate target area using close operator | i. Target

Fig. 4. Background subtraction results in different component, grayscale image, combination in three components and morphology operation results.

$$c(r,r_0)=\begin{cases}1 & |I(r)-I(r_0)|\le t\\ 0 & |I(r)-I(r_0)|>t\end{cases} \tag{5}$$

With t is a threshold measuring the gray similarity between center pixel and other pixels around. If $c(r,r_0)=1$, we consider the point pixel and nuclear pixel (center pixel) with the same (or similar) gray, so the region by the composition of the pixel value which meet above conditions is known as Univalue Segment Assimilating Nucleus, USAN, shown as Fig. 5.

At each pixel, the weight of USAN is defined as:

$$n(r_0)=\sum_r c(r,r_0) \tag{6}$$

A "smoother" version of formula (5) is defined as the following formula to make it more dependable and more effective [11].

$$c(r,r_0)=\exp\left\{-\left[\frac{I(r)-I(r_0)}{t}\right]^6\right\} \tag{7}$$

The edge strength at each pixel defined as:

$$R(r_0)=\begin{cases}g-n(r_0) & n(r_0)<g\\ 0 & otherwise\end{cases} \tag{8}$$

With g a geometric threshold which is set to $\frac{3}{4}n_{\max}$ (proof of optimality of this threshold can be found in [11]). $n_{\max}$ is the max value of the USAN weight, 36 for a 37 pixels quasi-circular template for example.

A. Auto Threshold Estimation

SUSAN detector uses a fixed threshold but it does not suitable for all images. Iterative method [14] is used to find an adaptive threshold but it needs more computation time. We use an adaptive threshold method in reference [15]. Grads magnitude of a pixel is defined as:

$$g(i,j)=\frac{|f(i-1,j)+f(i+1,j)+f(i,j-1)+f(i,j+1)-4f(i,j)|}{4} \tag{9}$$

Experiments show that the grads magnitude has a value no more than p=5 in smooth area. We use an average value of grads magnitudes which have a value greater than p at the rough area of target detected in section II.

$$t=\frac{1}{n}\sum_{G_n(i,j)>p} g_n(i,j) \tag{10}$$

With n is the number of grads magnitudes with a greater value than p.

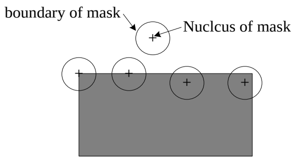


Fig. 5. Univalue segment assimilating nucleus (USAN). quasi-circular masks placed at different locations of an image containing a single rectangle. The USAN of each mask is marked in dark color. The nucleus of the masks are marked with a +.

B. Gray Barycenter Criterion for Denoising in [15]

For most noise points, the gray barycenter of USAN is the center of temple. The edge points have a gray barycenter not at the center of temple. This criterion is used to eliminate the false edge points caused by noises in the result of SUSAN edge detection.

C. Fill Holes Using a FOHBC Algorithm in [13]

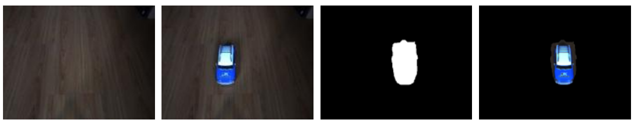
The edge detection detects the accurate edge of the target, smooth area without edge in the target area bring holes. We fill these holes using a FOHBC (Filling Object Holes Based on CBIBP) algorithm which is a fast hole-filling method proposed in [13]. CBIBP is a clustering method of binary image based on path. Results are shown in Fig. 8.

After filling holes in accurate target edge, we complete the accurate target detection. The target is sent to next level process and the background is feedback to background estimation for next frame detection.

IV. Experiments

We use a PointGrey® Bumblebee2 camera which has a resolution 640*480@48fps to capture a car and a train toy with different gains control. The two-step accurate target detection framework is implemented in Matlab 2010.

Background, foreground, rough target area detection results are shown in Fig. 6. The rough target area contains shadow and effects of the reflected light.



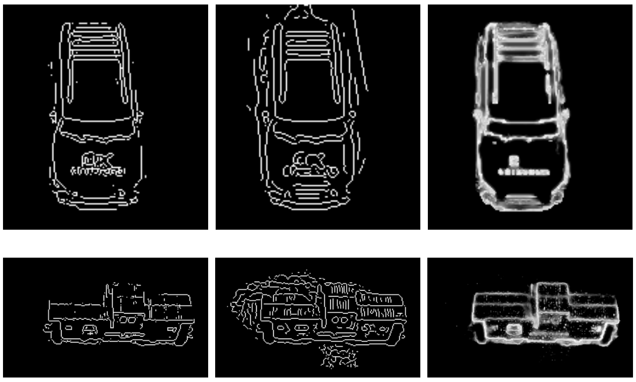
background | foreground | Rough area | Rough target

Fig. 6. Rough target area detection results.

We compare our method with canny edge method, Sobel edge method as shown in Fig. 7.

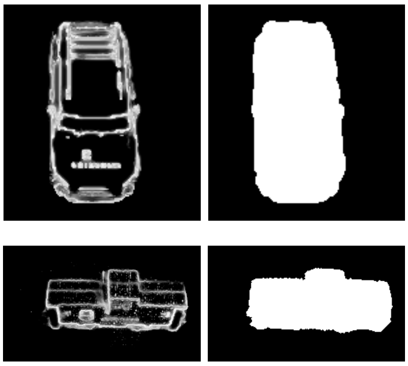
Fig. 8 shows the results of filling holes by FOHBC algorithm.

Fig. 9 shows the target detection method working in different gain cases. Our methods can detect the target accurately in video frames up to 15db gain.



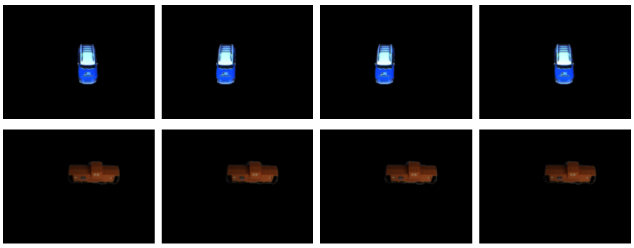
a. Sobel detector b. Canny detector c. Enhanced SUSAN

Fig. 7. Compare of different edge detectors. sobel and canny detector use the implements in matlab 2010. all of the detectors only find edges in the rough target area detected by algorithm in section II. our method detects more detail of the edge than others.



Edge detected after filling holes

Fig. 8. Holes filled using FOHBC



No gain 5db gain 10db gain 15db gain

Fig. 9. Detection in different gain video frames

V. Conclusions

Aiming at accurate target detection mission, we propose a two-step accurate target detection framework which contains a rough target detection step using background subtraction and an accurate target detection step using an enhanced SUSAN edge detection method. After the accurate target detection, background in current frame is feedback to background estimation. Experiments show that this method can detect accurate target area in different noise level video frames robustly. The two-step method proposed in this paper can be used in the applications which need to detect the accurate moving target.

References

[1] C. Li, J. Si, and G. P. Abousleman, "Low False Alarm Target Detection and Tracking within Strong Clutters in Outdoor Infrared Videos," *Optical Engineering*, vol. 49, pp. 086401, 2010.

[2] A. Behrad, Ali Shahrokni and Seyed Ahmad Motamedi. “A Robust Vision-based Moving Target Detection and Tracking System,” *In the Proceeding of Image and Vision Computing Conference, University of Otago, Dunedin*, New Zealand, November 2001.

[3] L. F, Y. J. Y, and C. J. X, “A Hybrid Camera for Motion Deblurring and Depth Map Super-resolution,” *2008 IEEE* Conference *on Computer Vision and Pattern Recognition*, vol. 1-12, 2008, pp.1803-1810.

[4] R. I. Hammoud, “Augmented Vision Perception in Infrared,” *Springer Verlag London Limited*, UK. 2009.

[5] M. Camplani and L. Salgado, "Adaptive Background Modeling in Multicamera System for Real-time Object Detection," *Optical Engineering*, vol. 50, 2011.

[6] A. Lipton, H. Fujiyoshi, R. Patil, “Moving Target Classification and Tracking from Real-Time Video,” in *Proceedings of IEEE* Workshop *on Applications of Computer Vision, Princeton*, NJ, USA, 1998, pp. 8-14.

[7] C. Stauffer and W. E. L. Grimson, “Adaptive Background Mixture Models for Real-Time Tracking,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Fort Collins, Colorado, USA, 1999, vol. 2, pp. 246-252.

[8] W. Yua, X. Yu, P. Zhang, and J. Zhou, “A New Framework of Moving Target Dectection and Tracking for UAV Video Application,” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVII. Part B3b, Beijing 2008, pp. 609-614.

[9] H. Michael and H. Nagel, “Combination of Edge Element and Optical Flow Estimates for 3D-Model-Based Vehicle Tracking in Traffic Image Sequences,” *International Journal of Computer Vision*, 1999, vol.35, pp. 295-319.

[10] C. John, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine* Intelligence, vol. PAMI-8, no. 6, 1986, pp. 679-698.

[11] Smith S M, Brady J M, “SUSAN-a new Approach to Low Level Image Processing,” *International Journal of Computer* Vision, 1997,vol.23, pp.45-78.

[12] Y. C. S. Zhang and Y. H. Y. Wang, “Edge Detection Algorithm based on SUSAN Operation on Auto Hub Image,” in *Proc. 9th WSEAS Internantional Conference on Multimedia Systems and Signal Processing*, Stevens Point, Wisconsin, USA, 2009, pp.63-66.

[13] J. Bo, “Research on Moving Object Detection and Tracking Technology for Intelligent Video Surveillance,” *Dissertation of* Graduate *School of National University of Defense Technology*, Changsha, Hunan, P. R. China, 2009.

[14] M. M. Petez and T. J. Dennis, “An Adaptive Implementation of the SUSAN Method for Image Edge and Feature Detection,” in *Proceedings of International Conference on Image Processing, Santa Barbara*, 1997, pp. 397-397.

[15] Y. Qifeng, L. Hongwei, and L. Xiaolin, “Accurate Measurement and Motion Measurement Based on Image,” *Scinece Press*, BeiJing, China, 2002.

Jun Zhang received her BS degree in application mathematics from Information Engineering University, Zhengzhou, China, in July 1996. She then received MS degree in management science and engineering and PhD in system engineering from National University of Defense Technology , Changsha, China, in 2002 and 2008, respectively. Her research interests include image processing, image coding, compressive sensing, image fusion, target detection and data mining.

Jun Zhang now is an associate professor in the college of Information System and Management at National University of Defense Technology (NUDT).