WCCI 2012 IEEE World Congress on Computational Intelligence
June, 10-15, 2012 - Brisbane, Australia FUZZ IEEE

A Modified Method of Vehicle Extraction Based on Background Subtraction

Jingjing Yang, Yaping Dai
School of Automation
Beijing Institute of Technology
Beijing, China,100081
jingjingyangel@gmail.com
daiyaping@bit.edu.cn

***Abstract*— A method combines Top-hat Transformation, Morphological Gradient and Background Subtraction is presented in this paper to solve the question of cast shadows and split of vehicle. The method adopts top-hat transformation both on input image and background image to remove the shadows and detect road lines respectively. Then morphological gradient is obtained by multiple structuring elements. The final modified background subtraction which performs on the binary images removes the road lines which impact the result much. Different from other methods, the proposed method eliminates the shadows and extracts the whole vehicle even if the color of the vehicle is similar to the road and the contour beside windshield is very weak. Experimental results prove that the method works well under different vehicle colors and sizes, comparing with the simple background subtraction and the background subtraction based on top-hat transformation. The error rate indicates the improvement.**

Keywords-vehicle extraction; top-hat transformation; shadow elimination; morphological gradient; background subtraction

I. INTRODUCTION

The video in intelligent transportation system (ITS) contains not only vehicles but also other information. Extracting vehicles from the abundance information is a basic job in transportation management system. The subsequent vehicle recognition or tracking will be enormously simplified if the vehicle extraction is accurate. The methods that commonly used for vehicle extraction include optical flow [1], temporal differencing [2] and background subtraction [3]. Although the optical flow computation can provide better performance, it has a heavy computational load and difficulty to realize in a real time system [4]. The temporal differencing is very adaptive to dynamic environment, but generally does a poor job of extracting all relevant feature pixels. The results will be with holes because of choosing unbefitting time interval. Compared with the two methods mentioned above, the background subtraction is more widely used in moving objects detection fields. Vehicles can be extracted by taking the difference between the background image and the input image in the prediction of stationary background. Unfortunately, the problem is not so straightforward in practice. It is sensitive to pixel variation resulting from noise and illumination changes, which frequently occur in complex natural environments. It needs to maintain an adaptive statistical background model [4]. Traditional background subtraction sometimes can't remove the stationary objects completely because of the luminance change, which causes errors when the vehicle is very close to the road lines.

Whether the method is temporal differencing or background subtraction, threshold selection of the difference image is challenging. Several adaptive threshold selection methods have been proposed, such as the Otsu's method [5], Isodata algorithm [6] and triangle algorithm [7]. The Otsu's method selects the best threshold by calculating the maximum interclass variance. It works well with bimodal histograms. The Isodata algorithm searches for the best threshold by an iterative estimation of the mean values of the foreground and background pixels. The triangle algorithm is particularly fit for unimodal histograms. Each of them works well only in certain situations.

In order to improve the accuracy of vehicle extraction, many modified algorithms have been proposed. W. W. L. Lam et al [8] proposed a texture-based background subtraction method to extract vehicle, which is more complicated than gray-based. The studies of texture analysis of road conditions and vehicles are limited [9]. J. Zhou et al [4] proposed a method with two thresholds to segment objects. If the difference between the background and the input image is less than the smaller threshold, it would be identified as the background. If the difference is greater it would be identified as the foreground. Those pixels between the two thresholds would be further checked whether it is shadow or not. However, it fails to deal with the situation that shadows are more easily detected than foreground. As a consequence, there is no universal approach for vehicle extraction that is guaranteed to work for all images.

The results of moving vehicle extraction are strongly affected by moving cast shadows especially in the case that the color of the vehicle is similar to the road in gray-scale. Moreover, the vehicle will be divided into two parts because of the weak edges with the camera position overhead. Large errors may result in the computation of the shape and other parameters. A new method is proposed in this paper which aims at classifying pixels as foreground or background accurately. It combines top-hat transformation, morphological

Identify applicable sponsor/s here. *(sponsors)*

gradient and modified background subtraction, which is called for abbreviation with "TMMB" method to eliminate the cast shadows and extract the whole vehicle. This method utilizes top-hat transformation on the input image to eliminate the shadows and on the background image to detect the road lines. Then the morphological gradient is obtained from the transformed input image. Furthermore, a modified background subtraction which performs on the binary images removes the road lines beside vehicles. At last, the subsequent morphological operations extract the final foreground that represents the moving vehicle.

The remainder of this paper is organized as follows: session II is the design of proposed method; session III is the simulation results and analysis; session is the conclusion.

II. DESIGN OF NEW METHOD

In this paper, there are three assumptions. First, all the images have been transformed into gray-scale images. Second, the camera is overhead and stationary. Third, for better extraction, several detection windows are set in each drive. The algorithm operates on detection windows separately. The other part of the whole image is left out of account.

A. Overview

Based on these three assumptions, a new method of vehicle extraction is proposed here to classify pixels as foreground or background accurately. The new method combines the top-hat transformation, the morphological gradient and the modified background subtraction. The background image is estimated from the video sequences by running average method [10], which is shown in (1).

$$Back(k) = \frac{1}{n}\sum_{m=1}^{n} SourceImage(m), \tag{1}$$

where *SourceImage(m)* represents the current input frame.

As depicted in Fig. 1, the proposed TMMB method consists of five steps.

- Top-hat transformation on both input image and background.
- Morphological gradient evaluation.
- Threshold selection.
- Modified background subtraction.
- Morphological operations.

The top-hat transformation eliminates the cast shadows. The morphological gradient extracts weak edges and fills the gaps caused by weak edges. The modified background subtraction removes the road lines near the vehicle which can't be completely removed by traditional background subtraction and connected region labeling. The combination plays respective advantages of three different methods, which solve the question of cast shadows, split of vehicle and background noises.

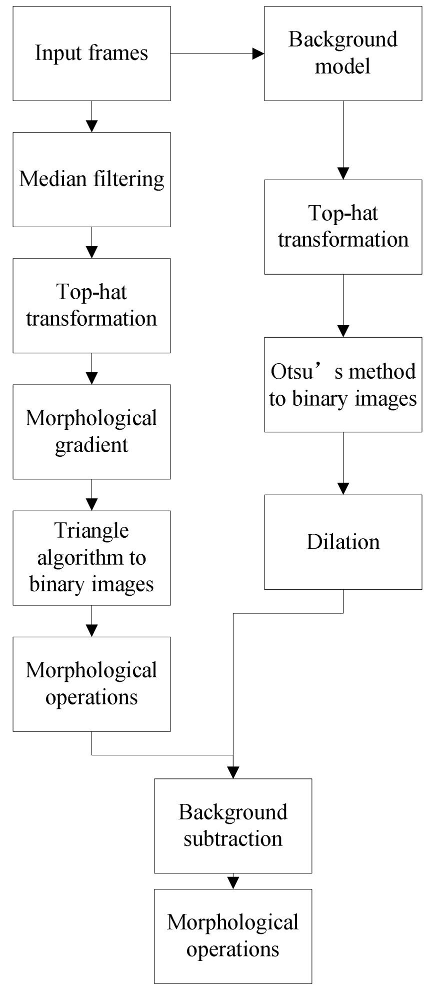


Figure 1. Architecture of proposed TMMB method

B. Top-hat transformation

Throughout the discussions that follow, *f(x, y)* and *b(x, y)* are digital image functions, where *f(x, y)* is the input image and *b(x, y)* is a structuring element.

Gray-scale dilation of f by b, denoted $f \oplus b$,is defined as

$$(f \oplus b)(s,t) = \max\left\{ f(s-x,t-y)+b(x,y) \mid (s-x),(t-y)\in D_f;(x,y)\in D_b \right\}, \tag{2}$$

gray-scale erosion, denoted $f \Theta b$, is defined as

$$(f \Theta b)(s,t) = \max\left\{ f(s+x,t+y)-b(x,y) \mid (s+x),(t+y)\in D_f;(x,y)\in D_b \right\}, \tag{3}$$

where D_f and D_b are the domains of f and b, respectively. Opening , denoted $f \circ b$,is

$$f \circ b = (f \Theta b) \oplus b, \tag{4}$$

closing, denoted $f \bullet b$, is

$$f \bullet b = (f \oplus b)\Theta b. \quad (5)$$

The so-called morphological top-hat transformation of an image, denoted h, is defined as

$$h = f - (f \circ b). \quad (6)$$

It enhances the detail in the presence of shading. The opening of the image in the transformation compensates for asymmetrical illumination on background. However, it eliminates some bright details of the image, which impacts the edges detection much.

Here a weighted sum of the transformed image and the original input image is chosen to solve the problem, which is shown in (7).

$$SumImage = \alpha \times TransImage + (1-\alpha) \times SourceImage, \quad (7)$$

where *SourceImage* is the original input image，*TransImage* is the top-hat transformed image and *SumImage* is the input for morphological gradient. α is the weight, which is determined by experiments. It is selected as 0.8 here. α can't be so small, or the shadows will not be eliminated completely, which may result in errors.

C. *Morphological gradient*

Dilation and erosion are often used to compute the morphological gradient of an image, denoted g:

$$g = (f \oplus b) - (f \Theta b), \quad (8)$$

where f and b are mentioned as before. g is the gradient result. It highlights sharp gray-level transitions in the input image. As opposed to gradients obtained using the methods such as Canny, Sobel, Roberts and Laplacian, morphological gradients obtained using symmetrical structuring elements tend to depend less on edge directionality and the result is more consecutive.

Here the equation in [11] is used to obtain the morphological gradient on transformed input image, which is shown in (9).

$$g_i = (f \circ b_i) \oplus b_i - (f \circ b_i) \bullet b_i \quad (9)$$

Here the details of edges detection are shown in (10).

$$OR = \alpha \times g_1 + (1-\alpha) \times g_2 \quad (10)$$

Where b_1 is a 3×3 square with elements of one, b_2 is a 5×5 square with elements of one. α is the weight, which is determined by experiments. It is selected as 0.8 here. *OR* represents the final result of gradient.

D. *Threshhold selection*

Since the distribution of the morphological gradient is unimodal, triangle algorithm is the most suitable to apply [12]. The triangle algorithm works as follows. Given a histogram function *h(v),* a line is constructed between the peak of the histogram *h(v) (at v = v1)* and the largest nonzero value of v (at $v = v2$). Then, the perpendicular distance between this line and the histogram is evaluated. The value of v that corresponds to the maximum distance is taken as the threshold value τ.

The threshold of the transforming of background to binary image is selected by the Otsu's method.

TABLE I. SUBTRACTION RULES

Input Image	**Background Image**	**Result**
1	1	0
1	0	1
0	1	0
0	0	0

1 represents the foreground and *0* represents the background.

E. *Modified background subtraction*

Traditional background subtraction sometimes can't remove the stationary objects completely, which causes errors when the vehicle is very close to the road lines. There are two reasons. First, the difference between the input image and the background image still exists though the background has been updated, which is usually caused by luminance change and respective top-hat transformation. Second, the method of morphological gradient enhances the errors caused by the first reason after traditional subtraction while it enhances the weak edges.

Here the modified subtraction performs on two binary images instead of performing on two gray-scale images. Usually, the image after opening with right structuring element is used as background for objects detection. As a result, the road lines can be easily detected after top-hat transformation. Compared with the traditional background subtraction, the modified method removes the road lines which are very close to vehicles and can't be removed by connected component labeling. The binary subtraction rules are shown in TABLE I.

The steps are shown as fallows.

- Top-hat transformation on estimated background.
- The Otsu's method for threshold selection.
- Dilation with a 5×5 square on binary background.
- Subtraction between binary processed input image and background.
- Morphological operations.

III. SIMULATION AND RESULTS ANALYSIS

Some typical outdoor traffic image sequences in daytime have been captured with the camera position overhead and stationary to test the performance of the proposed algorithm in this paper. The proposed method TMMB was tested under different vehicle colors and sizes. Simulation is implemented in Matlab7.11.0. The b of top-hat transformation is selected as a ten-pixel disk involving elements of one, where ten-pixel is the radius. All the top-hat transformations are performed according to (7). The first morphological operations include closing, fill and opening in turn. b_2 is for the closing, b_1 is for the opening. The second morphological operations include connected component labeling, closing, fill and opening in turn. The structuring element of top-hat transformation is both for closing and opening.

In the first sample, an input frame of a tint color car is shown in Fig. 2(a). Fig. 2(b) is the reference image which is got by hand. Fig. 2(c) shows the result of simple background subtraction (SB). Fig. 2(d) shows the result of background subtraction based top-hat transformation (TB). It does top-hat transformation on both the input image and background before traditional background subtraction. Fig. 2(e) shows the result of the method which combines top-hat transformation, morphological gradient and traditional background subtraction (TMTB). The subtraction performs after top-hat transformation. Fig. 2(f) shows the result of proposed method TMMB in this paper. Fig. 2(g) shows the result of TMMB in gray-scale image.

The proposed method in this paper also works well with vehicle color similar to the road, as shown in Fig. 3(a), which is the second sample. The final extracted vehicle is depicted in Fig. 3(f). Figs. 3(b) through Figs. 3(e) represent the same methods as Figs. 2(b) through Figs. 2(e) respectively.

The third sample shows the comparison results of a bigger size.

The fourth sample shows the comparison results of a car which is very close to the road lines. Fig. 5(d) doesn't be performed with connected region labeling, or the vehicle will be divided into two parts.

The fifth sample shows the comparison results of a car with very weak edges beside windshield. Two vehicles are extracted as one in Fig. 6(c) resulting from a big structuring element . If the structuring element of morphological operation is smaller, the vehicle will be divided into two parts due to the weak edges beside windshield.

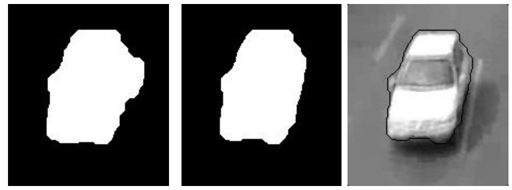
The simulation results depict that simple subtraction between image sequences and background reference frames contains a large amount of errors due to the shadow. The method of background subtraction based on top-hat transformation eliminates the moving cast shadows under different vehicle colors and sizes. Fig. 5(e) proves that after traditional background subtraction, road lines sometimes still be detected as foreground, which is resolved by using the proposed method.

The error rate is calculated in (11).

$$E = \frac{R \otimes I}{R} \times 100\%, \qquad (11)$$

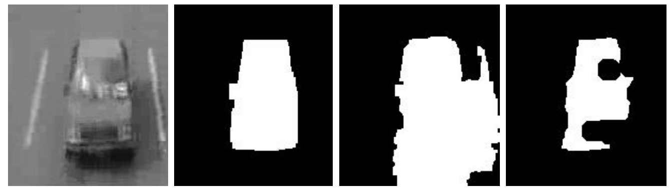
where R represents the reference image, I represents the result, $\otimes$ means XOR.

(a) (b) (c) (d)

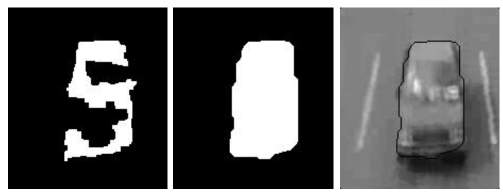


(e) (f) (g)

Figure 2. Sample 1— tint color.(a) The input image. (b) The reference image. (c) The result of SB. (d) The result of TB. (e) The result of TMTB. (f) The result of TMMB. (g) The result of TMMB in gray-scale image.



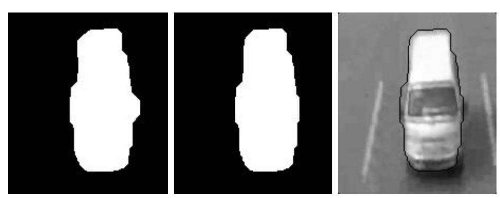
(a) (b) (c) (d)



(e) (f) (g)

Figure 3. Sample 2—dark color. (a) The input image. (b) The reference image. (c) The result of SB. (d) The result of TB. (e) The result of TMTB. (f) The result of TMMB.(g) The result of TMMB in gray-scale image.

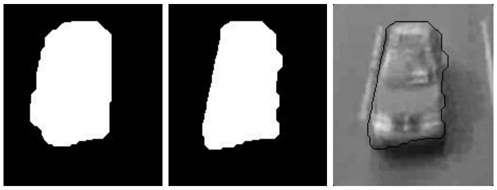
(a) (b) (c) (d)



(e) (f) (g)

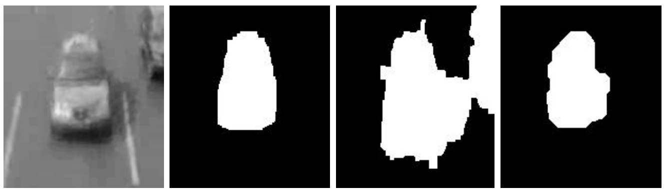
Figure 4. Sample 3—big size. (a) The input image. (b) The reference image. (c) The result of SB. (d) The result of TB. (e) The result of TMTB. (f) The result of TMMB.(g) The result of TMMB in gray-scale image.

(a) (b) (c) (d)

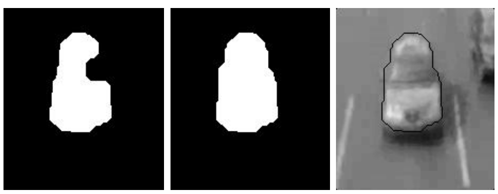


(e) (f) (g)

Figure 5. Sample 4—near road lines. (a) The input image. (b) The reference image. (c) The result of SB. (d) The result of TB. (e) The result of TMTB. (f) The result of TMMB.(g) The result of TMMB in gray-scale image.



(a) (b) (c) (d)



(e) (f) (g)

Figure 6. Sample 5—weak edges. (a) The input image. (b) The reference image. (c) The result of SB. (d) The result of TB. (e) The result of TMTB. (f) The result of TMMB.(g) The result of TMMB in gray-scale image.

TABLE II. ERROR RATE

Error Rate(%)	**SB**	**TB**	**TMTB**	**TMMB**
Sample 1	74.80	14.60	19.07	11.52
Sample 2	93.97	33.65	49.51	10.25
Sample 3	84.35	15.20	16.61	14.63
Sample 4	87.93	69.35	19.05	10.62
Sample 5	132.19	15.06	16.00	8.97

The error rate is shown in TABLE II. It shows that the error rate of the proposed method TMMB is the minimum. The average error rate of these samples is 9.2%.

The TMMB algorithm is implemented in Matlab and the average runtime to process a 150×170 gray-scale image is 0.7 seconds. If this algorithm is implemented in C or C++ with code optimization, the runtime will be reduced. Furthermore, several consecutive frames contain the same vehicles, not each frame should be processed. As a result, the algorithm is potential to run in real time.

IV. CONCLUSIONS

A vehicle extraction method TMMB has been proposed in this paper. This method combined top-hat transformation, morphological gradient and modified background subtraction together. The method utilized top-hat transformation both on input image and background image to remove the shadows and detect road lines respectively. Then morphological gradient was obtained on the transformed input image to enhance the weak edges to fill the gaps. The final modified subtraction removed the road lines which impact the result much. The simulation results have proved that the TMMB method eliminated the shadows and extracted the whole vehicle with an average error rate 9.2%.

REFERENCES

[1] B.K.P. Horn and B.G. Schunck, Determing optical flow, Artifical Intelligence,vol.17, Aug.1981, pp.185-203.

[2] S. Murali and R. Girisha, Segmentation of motion objects from surveillance video sequences using temporal differenceing combined with multiple correlation, IEEE International Conference on Advanced Video and Signal Based Surveillance, 2009,pp.472-477 .

[3] W. Liu and N. He, Moving object detection method based on background subtraction,vol.47. Computer Engineering and Applications , vol.47,2011,pp.175-179.

[4] J. Zhou and J. Hoang, Real time robust human detection and tracking system, in Proc. IEEE Comput. Vis. Pattern Recog. Workshop, San Diego, CA, 2005, pp. 149.

[5] N. Otsu, A threshold selection method from gray-level histograms,IEEE Trans. Syst., Man, Cybern., vol. SMC-9, no. 1, Jan. 1979,pp. 62–66.

[6] T. W. Ridler and S. Calvard, Picture thresholding using an iterative selection method, IEEE Trans. Syst., Man, Cybern., vol. SMC-8, no. 8, Aug.1978,pp. 630–632.

[7] G. W. Zack, Automatic measurement of sister chromatid exchange frequency, J. Histochem. Cytochem., vol. 25, no. 7, 1977,pp. 741–753.

[8] W. W. L. Lam, C. C. C. Pang and N. H. C. Yung, A highly accurate texture-based vehicle segmentation method,*Opt. Eng.*, vol. 43, no. 3, 2003,pp. 591–603,.

[9] M. Yamada, K. Ueda, I. Horiba and N. Sugie, Discrimination of the road condition toward understanding of vehicle driving environments, IEEE Trans. Intell. Transportation Syst, vol.2, 2001,pp.26–31.

[10] H.Zhu, Research on the moving detection and target tracking based on video sequences, Southwest Jiaotong University Master Degree Thesis,China,2008.

[11] J. S. Yang, C. Huang and Y. J. Fan, Edge detection of medical image based on multi-structure element morphology, Opto-Electronic Engeering, vol.35, no.3, 2008,pp.112–116

[12] Lu Wang, Nelson H.C. Yung, Extraction of moving objects from their background based on multiple adaptive thresholds and boundary evaluation ,IEEE Transactions on Intelligent Transportation Systems,vol.11, no.1,2010,pp. 40–51.