Noise Reduction Algorithm of Vehicle Detection
In Intelligent Transportation System

Mao Yanfen, Shi Pengfei,
(Institute of Image Processing & Pattern Recognition, Shanghai Jiaotong University, Shanghai, China)

Abstract—Vehicle detection is one of the key technologies in Intelligent Transportation System (ITS), and it is an important stage of vehicle tracking in visual surveillance. Due to the clutter of traffic scenes, the captured video sequence involves many big size noises. According to analyzing human vision, the non-object area and shape suppression are defined, and a new noise-removing algorithm is proposed. To reduce the computational complexity, the sequential algorithm is improved. Experimental results show the effective and efficient of the algorithm in extracting the moving vehicles in cluttered scene. The output of vehicle's contour is very complete and accurate, and it can be used in vehicle tracking systems to improve the performance.

Index Terms—area suppression, Intelligent Transportation System (ITS), improved sequential algorithm, shape compactness, shape suppression, vehicle detection

I. INTRODUCTION

Vehicle detection and tracking are important problems in extracting traffic parameters [1] in Intelligent Transportation System (ITS). The primary stage of detection and tracking is the moving objects segmentation. An accurate segmentation of the vehicles is required in order to obtain an accurate tracking.

Background subtraction is the traditional technique [2], [3] in obtaining foreground image. The detection algorithm deals with the foreground image and extracts the moving objects. The detection output, a binary image consists of moving object's contour that is labeled with value one in eight-neighbor pixels. The location of object's contour directly affects the accuracy of final binary image, but the details of object inner parts are not important to results. Due to the shadow, underexposure, dithering and so on, some edge pixels have small gray-level gradients, for example, the circled edge pixels in Fig. 1.

In this case, if we use the higher threshold value to thresholding the foreground image, the shape of detected vehicle is not integrated. Some edge pixels disappear and the segmented vehicle is not complete. If we use the lower value, it reserves the complete shape of objects but adds some noise. Therefore, there tends to find an algorithm that can not only accurately extract object's contour but also effectively eliminate noise.

There are many edge detection methods wherein Canny operator [4] seems as the most effective one. It can get the detail edges of image, but it is not appropriate for moving vehicle detection. The reasons are as follows:

1) Because the intensity difference between the edges of objects and background are variable, the double-threshold of Canny operator is difficult to detect the edge with lower intensity difference.

2) Canny operator needs to smooth the image by Gaussian filter. It makes the edge pixels which intensity values similar to background difficult to be detected.

3) Canny operator uses the local maxima of gradient values and gets some edges describing the inner details. However, the inner details of the objects are not necessary in vehicle extraction.

To resolve the above problems, a new edge-preserved and noise reduction algorithm is proposed. Considering human vision, we usually observe the whole image and perceive the approximate intensity changes. Then, we consider the closed region which intensity differs from the other regions as an object. Finally, according to the size and shape, we determine an object whether a target vehicle or not. For instance, we usually do not consider a threadlike object as a vehicle, and regard the isolated and small objects as noise. In that sense, the emphasis of the paper is rather on the contour extraction of the whole image than the inner details of objects.

 Manuscript received April 14, 2003. This work was supported in part by the National "973" Key Project on Urban Traffic Monitoring and Management System (No.2001CB711401).

Mao Yanfen is with the Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, China (Corresponding author. phone: +86-21-62932957, fax: +86-21-62932957 e-mail: sIFY@sjtu.edu.cn).

Shi Pengfei is with the Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, China. e-mail: pfshi@sjtu.edu.cn.

0-7803-8125-4/03/$17.00 © 2003 IEEE

1368
Experimental results show that the algorithm has a good performance on noise-resistance and contour detection.

The rest of this paper is organized as follows. In Section II, we describe the proposed algorithm using the area and shape suppression. Section III demonstrates the experiment results and finally, conclusions are given.

II. DESCRIPTION OF ALGORITHM

A. Compute Gradient

Gradient is more effective than intensity in representation of image variation. Therefore, the proposed algorithm computes the two gradients using Sobel operator in height and width directions of every pixel. Set the maximum of the absolute gradients of two directions as the gradient of this pixel.

B. Lower-threshold Processing

Take a lower threshold $T_2$, thresholding the gradient image and getting the binary image that reserved the edge details of moving vehicle.

C. Open Operation

In order to remove isolate noise, it adopts the open operator which structure element is $3 \times 3$ square. Open operation can break the connection of isolated pixels and reduce the computation load of sequential algorithm.

D. Hole Filling

There are some big holes in the binary image after open operation, which is very disadvantageous to post-processing. So it needs to fill in the holes by the improved sequential algorithm described at length in next subsection.

E. Improved sequential algorithm

In order to reduce the computational time, this paper improved the sequential algorithm [5] considering eight-neighbor pixels.

1) Mark serial number: Scan the binary image from left to right and from up to bottom. Labeling all of the one-value pixels, and marking 0 to the zero-value pixels. Fig.2 is the relevant eight-neighbor pixels of computing the serial number of pixel O. Denote $L(\cdot)$ is the serial number matrix of pixels and $E(\cdot)$ is the equivalent serial number matrix.

![Fig.2 Sketch map of computing serial number of pixel O](image)

where $L(O)$ is relevant to the serial number of pixel A, B, C and D. If there are non-zero values in $L(A)$, $L(B)$, $L(C)$ and $L(D)$, $L(O)$ is set to the minimum of them.

Otherwise, the sum of serial numbers adds 1 and evaluates to $L(O)$. The detail of algorithm is described in Fig.3. Obviously, if $L(A)$, $L(B)$, $L(C)$ and $L(D)$ all have non-zero values, their relations are as follows: $L(A) \leq L(B)$, $L(C) \leq L(D) \leq L(C)$, $L(C) \leq L(D) \leq L(B)$.

2) Equivalent relation: Creating a one-dimension matrix tail that consists of the non-zero values of equivalent matrix E and another one-dimension matrix head that consists of the corresponding pixel values in serial number matrix L. The two matrices are one-to-one. It needs to create a one-dimension matrix active that describes the validity of equivalent relations. The corresponding relation table (i.e. equivalence table) is as follows:

![Fig.3 Flow chart of serial number labeling](image)

![Fig.4 Equivalence table](image)

where active consists of one-value and zero-value. One-value describes the validity of the corresponding relation, whereas zero-value denotes invalid.

There may exist the condition that one tail corresponds to many heads in equivalence table. For example in Fig.5, the figures are serial numbers.

![Fig.5 Example of one-to-many case](image)
After sequential connection, it appears two equivalences that 3-10-2 and 3-10-1. We deal with this case as follows.

If \( i \neq j \), \( \text{active}(i) \& \text{active}(j) = 1 \) and \( \text{head}(i) \neq \text{head}(j) \), let \( \text{tail}(i) = \text{head}(j) \); otherwise, let \( \text{active}(j) = 0 \), and make the corresponding relation \( j \) invalid. Scan the equivalence table till there has no one-to-many relations.

The final equivalence should be one tail corresponding to one head that will not appear in matrix tail (define this head is root_head). That is, for every head(i), scan tail, if \( \text{tail}(j) = \text{head}(i) \), let \( \text{head}(i) = \text{head}(j) \), till every head in head table is root_head.

3) Equivalence relation maps to serial matrix \( L' \): Scan the serial matrix, and replace tail with corresponding head in the available equivalence relation, then get a new serial matrix \( L' \).

F. Non-object suppression

1) Area suppression

Denote the number of object pixels as the object's area and \( T_a \) as the area suppression threshold. Compute the sum of non-zero values in serial matrix, and get a one-dimension matrix sum1. Assume that the sum smaller than \( T_a \) is noise and eliminate the serial number in \( L' \). Therefore, many non-objects can be eliminated by area suppression. For example, there is twinkle of red-green light in the traffic video sequence. Because the target of vehicle detection is vehicle, the red-green light in binary image is considered as non-object and eliminated by area suppression.

2) Shape suppression

The objects (e.g. vehicles) usually have basic shapes and these pixels have common arrangement. In order to describe this compact extent of one-value pixels, a scalar quantity, shape compactness is proposed. The compactness is more relevant to the relatively distribution of one-value pixels than their pixel numbers. Refer to Fig. 6 (c), \( O \) is a center point, and pixels that the distance from \( O \) smaller than chess distance \( m + 1 \) form an array. The shape arranged as Fig. 6 (c) is considered as the highest compact extent and the pixel numbers satisfied the function: \( f(m) = 2m^2 + 2m + 1 \).

Assume that the sum of one-value pixels in \( L' \) is \( \text{sum2}, (\text{sum2} > 1) \), it must exist a \( m \) meeting the inequation: \( f(m) < \text{sum2} \leq f(m + 1) \). If the sum of the pixels arranged as Fig.6(c) is \( \text{num} \), the ratio of \( \text{num} \) to \( \text{sum2} \) is defined as shape compactness \( \text{den_ratio} \). Predefined the threshold of shape compactness \( T_c \), it can eliminate the regions which \( \text{den_ratio} \) is smaller than \( T_c \).

G. Object contour extraction

After the above processing, the non-object and noise are eliminated. The final object contour extraction is comparatively simple. In binary image, if the eight-neighbor of one pixel are all one-value, it does not belong to the contour. Otherwise, it is a contour pixel and needs to be preserved.

III. EXPERIMENTAL RESULTS

The proposed algorithm is tested by the real data. The traffic sequence is captured from the Yan'an Xi Road in Shanghai, China. The parameters are setting as: \( T_i = 25 \), \( T_o = 200 \), \( T_r = 0.8 \). Fig.7 and Fig.8 show some results of frame 180 and 295, respectively. From Fig.7 (b) and Fig.8 (b), it can be seen that there are many noise after gradient thresholding. Moreover, the uniform gray-value of objects results in many holes in the binary images. However, after hole filling and area and shape suppression, many non-object regions are removed and eventually the integrated shapes of the moving vehicles are obtained. Fig.8 (f) and Fig.9 (f) demonstrate that the final results of contour detection are satisfactory, and the detected moving vehicles' outlines are very accuracy and integrity.
The area suppression removes about 200–300 non-object regions, whereas the shape suppression eliminates comparatively fewer regions.

However, as the above mentioned, the foreground image is obtained from background subtraction technique [6], extraction of background image is a key stage in detection algorithm. Because of the variety of illumination and the clutter of traffic scene, the background image needs to update over time. Although the dynamic background image can decrease the number of non-object and noise to some extent, it requires heavy computational cost and inappropriate to real-time applications. Therefore, there are a trade-off between the accuracy and the computational efficiency.

### IV. CONCLUSION

A new noise-removing algorithm is proposed in this paper. Analyzing the human vision in detecting the vehicles, this paper defines the area and shape suppression and improved the sequential algorithm. The algorithm makes the finally detected vehicle contour very integrated and the computation very efficient. Due to the effectiveness in extracting the integrated contour, the future work will focus on the feature-based tracking algorithm using the vehicle's contour, and to improve the performance of traffic surveillance [7] system.

### ACKNOWLEDGMENT

The authors would like to thank the reviewers for their valuable comments that improved the presentation of the paper.

### REFERENCES


Mao Yanfen received the B.S and M.S degrees, both in Control Theory and Control Engineering, from Northwest Polytechnical University (NPU), Xi'an, China, in 1999 and 2001, respectively. She joined the Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, in 2001 and is working toward the Ph.D degree. Her research interests include image and video processing and Intelligent Transportation System.

Prof. Shi is a doctoral supervisor in Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, China. His research interests include image processing, pattern recognition, artificial intelligence, computer vision and data mining.