A Study of Vehicle Detection and Counting System Based on Video

Shuang XU, Lingbin PANG, Huairuo YIN
Information & Communication Engineering, Dalian Nationalities University
Liaohe West Road, 18 Dalian, China
E-mail: xushuangcong@163.com, plingb@dlnu.edu.cn, huairuosmile@163.com

Received: 24 June 2014   /Accepted: 30 September 2014   /Published: 31 October 2014

Abstract: About the video image processing's vehicle detection and counting system research, which has video vehicle detection, vehicle targets' image processing, and vehicle counting function. Vehicle detection is the use of inter-frame difference method and vehicle shadow segmentation techniques for vehicle testing. Image processing functions is the use of color image gray processing, image segmentation, mathematical morphology analysis and image fills, etc. on target detection to be processed, and then the target vehicle extraction. Counting function is to count the detected vehicle. The system is the use of inter-frame video difference method to detect vehicle and the use of the method of adding frame to vehicle and boundary comparison method to complete the counting function, with high recognition rate, fast, and easy operation. The purpose of this paper is to enhance traffic management modernization and automation levels. According to this study, it can provide a reference for the future development of related applications. Copyright © 2014 IFSA Publishing, S. L.

Keywords: Vehicle detection, Image processing, Vehicle counting, Inter-frame difference method, Image segmentation, Framed cycle approach.

1. Introduction

The current rapid development of world science and technology, to promote intelligent traffic continues to progress, for the sustainable development of the national economy has laid a solid foundation [1]. Among them, microelectronics, computer and communication technology’s large-scale use in the field of social traffic largely promote the upgrading of traffic operation and management mode [2]. Video vehicle detection and counting system is a computer analysis system of using of image processing and pattern recognition techniques to achieve traffic target detection and recognition. It has a video vehicle detection, vehicle target image processing and vehicle counting. Vehicle detection is the use of inter-frame difference method and vehicle shadow segmentation techniques for vehicle testing [3]. Use video image processing method to make vehicle detection by the image that is a change in a specific area among the pixel gray value analysis to determine whether the current road by a car [4]. The system is the use of inter-frame video difference method to detect vehicle and the use of the method of adding frame to vehicle and boundary comparison method to complete the counting function, with high recognition rate, fast, and easy operation.

Vehicle detection is to judge whether the vehicles pass by the detection zone, and the establishment of a target tracking, mainly to provide information such as vehicle flow. To simplify the complexity of vehicle detection algorithm and improve the real-time of contradictions, and resolve this conflict is very important for improving the system stability and the
accuracy of detection, the clarity of motion object background, the actual illumination variation, the existence of jitter camera, false image target and shielding phenomenon can affect the vehicle detection and segmentation accuracy, the operator must be considered in the calculation of the influence factors and removal of solutions to such factors.

2. Vehicle Shadow Segmentation Technologies

Through the above detection method, the detected image exists more or less the shadow interference phenomenon, and when the shadow area grows will cover and adjacent vehicles, lead to mistake more car detection algorithm into a car, the late recognition difficulty, therefore, need to think about shadow elimination solution in image detection.

At present, many of the literature about the shadow detection work, based on the characteristics of the detection algorithm adopted, can be divided into five categories: statistical parameters and nonparametric method [5], method of color constancy [6], deterministic models with deterministic model method [7], method of color space model [8].

These methods have their own advantages and disadvantages, needs to make choices for similarities and differences of the environment. For example based on color constancy of method applies to all shadow segmentation, based on color space model method has good stability, and nonparametric statistics method can solve the problem of different size and strength of the shadow, but the color space model to the problem of the treatment effect is good, while statistical non-parametric method in terms of the deal with fast; Deterministic model method can be applied to special environment, assuming that the more the more good effect.

3. Image Processing

Vehicle detection based on video is the study of video images, mainly using the camera image sequence. In different period, the sequence of the two or more frames images of relative movement between the covers the content and content information, this information includes the image frames and the variance of the position such as points, lines, movement area and movement direction, speed, etc.

Video image processing is the disadvantages of large amount of calculation, if use 240 x 320 pixel image frame size, frame rate of 25 frames per second, RGB image formats, 24 bits per pixel, so the image is 5.76 MB per second.

In testing hypothesis adopted many kinds of edge detection, filtering algorithm based on image region, will cause large amount of calculation, to meet the real-time and so on.

3.1. Color Image Gray Level

Each pixel colors in color images are R, G, B three components to decide, and each component has 255 values available for selection, a pixel has more than 1600 (255 × 255 × 255) color range [9]. But gray image [10] is R, G, B three-component as special color images, the first item change range is 255 pixels, so in the digital image processing usually will be gray image, the image transformation to reduce the amount of calculation for the following image. At the same time, the description of the gray image is consistent with color images reflect the global and local image characteristics and distribution of chromaticity and brightness level.

Although many color information is contained in color images, but does not apply to direct the operation and the gray level, shall be converted into gray image. Color conversion formula is as shown:

\[
g = 0.299 \times R + 0.587 \times G + 0.114 \times B .
\]  

After color gray level information enough for subsequent image processing and pattern recognition requirements involved. In addition, the cost of storage and processing cost are reduced. By the gray processing image is shown in Fig. 1.

![Fig. 1. Gray scale processing of color image.](image-url)


3.2. Image Segmentation

There are many methods of image segmentation, mainly introduces two kinds of method: method of threshold segmentation [11] and edge testing technique [12]. Here are the two ways to spread.

Threshold binarization image segmentation in general needs to do to image model assumptions. Then use the image model to try to understand the area of image composition. Based on image segmentation model usually choose such a hypothesis that goal or grayscale values between adjacent pixels in the background is similar, but different target or background pixels in the differences in how gray. If the original image is \( f(x, y) \), according to certain standard in \( f(x, y) \) obtained by a characteristic value, then the characteristic value is segmentation standard in \( f(x, y) \), divides the image into two parts, after segmentation of the image formula is:

\[
g(x, y) = \begin{cases} b_0 & f(x, y) < T \\ b_1 & f(x, y) \geq T \end{cases}, \tag{2}
\]

In the presence of a variety of threshold, the image after segmentation can be expressed as:

Which \( T_0, T_1, \ldots, T_k \) is a set of threshold segmentation, \( b_0, b_1, \ldots, b_k \) is corresponding to different domain after segmentation of image gray values, \( K \) is the region or after target number, \( g(x, y) = b_i \), when \( T_i \leq f(x, y) \leq T_{i+1} \), \( i = 0, 1, 2, \ldots, K - 1 \). Whether single threshold segmentation and threshold segmentation, all is to select a reasonable threshold, in order to determine pixels in each image should be the background area or target area, thus produce the corresponding binary image.

Segmentation method based on edge detection is aimed at by extracting the characteristic value of adjacent pixels step sex to get the edge between the different areas. Edge test to find out the area of pixel intensity change, generally adopt differential method to achieve, commonly used gray first and second order derivative is the main basis of edge detection, commonly used gray first and second order derivative is the main basis of edge detection, so people design all kinds of edge detection operator. The convolution operator is:

\[
\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}, \tag{3}
\]

Two templates are \( f(x, y) \) in the \( x \) direction and \( y \) direction of the first order difference. To select a proper threshold \( T \) and make the following decision: if \( R(x, y) > T \), then \( (x, y) \) is edge point, \( \{R(x, y)\} \) is the edge image.

Roberts operator is taking the difference of two adjacent pixels in the direction of diagonal gradient amplitude approximation to detect edges. Detection is better than the effect of the horizontal and vertical edges on the edge of the slope, the positioning accuracy is higher, but is sensitive to noise of high degree.

Sobel operator [15] has two nuclear convolution calculations. At various points in the image is to use this as the convolution of two nuclei, the first nuclear generally to the vertical edge of the corresponding is the largest, the second nuclear general situation of horizontal edge is the largest accordingly. Two convolution of the maximum output value for that point, the result is a range edge image.

Sobel operator is defined as:

\[
G[f(x, y)] = |dx| + |dy|, \tag{4}
\]

\[
dx = \left[ f(x-1, y-1) + 2f(x, y-1) + f(x+1, y-1) \right] - \left[ f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1) \right], \tag{5}
\]

\[
dy = \left[ f(x-1, y-1) + 2f(x, y+1) + f(x+1, y+1) \right] - \left[ f(x-1, y-1) + 2f(x+1, y-1) + f(x+1, y+1) \right]. \tag{6}
\]

The convolution operator:

\[
H_1 = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}, H_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{7}
\]

Formula of \( dx \) and \( dy \) the \( x \) direction and \( y \) direction of first order differential, \( G[f(x, y)] \) is gradient Sobel operator, and \( f(x, y) \) of integer pixel coordinates of the input image. After calculating the gradient, and may have a constant \( T \), in \( G[f(x, y)] > T \), identified as a boundary point, the point at which the pixel size set for 255, other set as 0, then adjust the numerical constants \( T \) achieve the best effect.

It is generally believed, Sobel operator that the effects on the current pixel neighborhood pixels are not equivalent, so distance is not the same pixel has different weight, so the influence of operator results also is not the same. Used to think that, the less the distance, the greater the impact.

To detect the edge of the same time reduce the noise influence, Prewitt operator since more edge
detection operator. Since the $2 \times 2$ to $3 \times 3$ is used to calculate the difference operator, so the convolution operator is:

$$H1 = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix}, H2 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix}.$$  (9)

Each pixel in the image location is used by the two templates to do convolution, Prewitt operator combine direction difference operation associated with the local average.

Measuring the Prewitt gradient, choosing the proper threshold $T$ to gradient image binarization, will obtain a binary image edges. Use Prewitt operator can detect the edge point, and can resist noise effects. These three kinds of edge detection results are shown in Fig. 2.

![Edge detection segmentation result.](image)

(a) Original image  (b) Roberts edge  
(c) Sobel edge  (d) Prewitt edge  
(e) Original image  (f) Roberts edge  
(g) Sobel edge  (h) Prewitt edge

**Fig. 2.** Edge detection segmentation result.

### 3.3. Morphology Analysis

Expansion of operator is $\oplus$, A to B expansion of writing is $A \oplus B$, are defined as:

$$A \oplus B = \{ x \mid (\hat{B} + x) \cap A \neq \emptyset \}.$$  (10)

Type, says: on the process of B expansion: B against doing about the origin of mapping the $\hat{B}$ first, and then map it translation $x$, where is A and B mapping intersection is not empty, the box B expansion set is obtained: $\hat{B}$ displacement and A has at least 1 non-zero elements of the intersection of the collection of B's origin [17].

Corrosion of operator is $\Theta$, A to B corrosion of writing is $A \Theta B$, are defined as:

$$A \Theta B = \{ x \mid (B + x) \subseteq A \}.$$  (11)

Formula above show that corrosion as A result of A with B all set $x$, B shift after $x$ will include in A. In other words, A collection is B to B corrosion include in the set of A in B's origin [18].

Inflation and the corrosion is not mutually inverse operation, so that they can be used combining. For the first image for corrosion after expanding the results (with the same structural elements here), said open operation, opens operator is $\ominus$, open A with B is $A \ominus B$, the definition for:

$$A \ominus B = (A \Theta B) \oplus B,$$  (12)

First image against doing swell after corrosion as A result, using the same structural elements (here), said the closure operation, closed operator for $\bullet$, closed A to B is writing $A \bullet B$, defined as:

$$A \bullet B = (A \oplus B) \Theta B.$$  (13)

Through combination of binary morphological basic operation, it can obtain a series of practical mathematical morphology algorithm.

$$\{[(A \Theta B) \oplus B] \oplus B\} \Theta B = (A \ominus B) \bullet B.$$  (14)

$$\beta(A) = A - (A \Theta B) \ (\beta(A) \text{ to set A boundary}),$$  (15)

$$X_k = (X_{k-1} \Theta B) \cap A^c \quad k=1,2,3,\cdots.$$  (16)

When $X_k = X_{k-1}$ stop iteration, among which, the complement of $A^c$ to $A$.

$$X_k = (X_{k-1} \Theta B) \cap A \quad k=1,2,3,\cdots.$$  (17)

When $X_k = X_{k-1}$ stop iteration, selecting $Y = X_k$, $Y$ represents a set communication component.

Set the skeleton of $S(A)$ to $A$, it can be expressed as:

$$S(A) = \bigcup_{k=0}^K S_k(A),$$  (18)

$$33$$
B belongs to the structural elements, \( \langle A \Theta k B \rangle \) says K uninterrupted B on Corrosion Behavior of A, K said the A corrosion become empty before iteration last 1 times, \( S_k(A) \) is often referred to as the skeleton subsets, writing \( S_k(A) = \langle A \Theta k B \rangle - \langle (A \Theta k B) \circ B \rangle \).

The background difference image frames obtained as shown in Fig. 3, the image are derived from the use of corrosion expansion algorithm, algorithm, and open and closed algorithm for image processing results of vehicle.

After the above treatment can be seen, often could not extract a fully connected target vehicle, vehicle target will be the next track brings certain influence, should fill image. The system on the vehicle region is to adopt the following methods of filling process:

Scan two value image \( g(x, y) \) of each line from left to right, assuming scan line I to obtain the first target pixel gray value is j, 0, marking the coordinates of the point is \( g(I, J) \); then in the same line or I line from right to left. Gets the first target pixel gray value is k 255, gray value of 0, marking the coordinates of the point \( G(I, K) \), then the I of each pixel in the two target point into target and background, as shown below:

\[
g(x, y) = 0, \quad y = j, j+1, \cdots, k-1, k, \quad (19)
\]

Transverse after filling the image \( g(x, y) \) of each column scanning, hypothesis from top to bottom scanning to the j column for the first target pixel gray value is I, 0, marking the coordinates of the point \( G(I, J) \); then the same column J column was the first gray value is the target pixel K 0 from top to bottom, the gray scale value of 255, marking the coordinates of the point \( G(k, J) \), then each pixel in the j column of the two target points into the background of the target.

After filling, the target vehicle can complete to be detected, as shown in Fig. 4.

\[\text{Fig. 3. Mathematical morphology algorithm results.}\]

\[\text{Fig. 4. Fill image processing results.}\]
Among them, Fig. 4(a), Fig. 4(c) are vehicle area map, unfilled graph Fig. 4(b), Fig. 4(d) is the image after filling.

4. Vehicle Detection and Counting Method Based on Inter frame Difference Method

As shown in Fig. 5 for vehicle target detection processing.

4.1. Inter Frame Difference Method

The inter frame difference method for image sequence is adjacent to two or three frame image to make differential treatment, after thresholding to obtain motion changed region, then take the edge the edge of animal body, finally detect the target. Is the basic principle of the method: the specific location of the image sequence as target in different frames of the change, the position corresponding to different frames in the gray will also change, but the object without changing position, the gray value is not changed, after thresholding to obtain the target object. Usually includes the adjacent frame difference and multi frame difference, this paper mainly adopts the method of the first category.

The adjacent frame difference is in the adjacent frame difference, form detection target formula is as follows. \( T \) is the two value image threshold.

\[
d(x, y) = |f_n(x, y) - f_{n-1}(x, y)|,
\]

\[
M(x, y) = \begin{cases} 
1, & d_n \geq T \\
0, & d_n < T
\end{cases}
\]

The inter frame difference method to detect moving target is simple, fast and has strong adaptability.

4.2. Analysis of Object and Counting

After the movement of vehicles in the video are extracted, usually able to obtain the object extraction, and then can make the vehicle counting function design of the system. In the algorithm, the two features of image contour are extracted after obtaining the rectangular frame data.

Vehicle detection and counting the overall flow chart is as follows Fig. 6.

![Design flow chart](image)

The two values and morphological filtering results run the following Fig. 7.

![Design of interface](image)

The video frame marking and counting results showed as follows Fig. 8.
4.3. Results and Analysis

In the algorithm, the cycle of each rectangular frame processing. For a rectangular box, first it calculates the value of its center point. In order to prevent the repeat count, can use the boundary do count, we can according to the numerical center point judging whether the vehicle is in the boundary, the boundary is not entirely if picture, need to continuously update the boundary, and based on the image boundary on a frame to determine whether the car enters the picture. Do count, for just entering the vehicle so that, even if the vehicle on the screen pause motion will not repeat count.

5. Conclusions

Moving vehicle detection and real-time tracking and counting are the basic behavior of the vehicle traffic flow analysis, the core of intelligent transportation. This paper around the video image moving target detection, based on the analysis and integration of previous research achievements, to grasp the characteristics of vehicle detection, preliminary design, realizes the real-time vehicle detection system, proved the feasibility and validity of the method used in this paper is realized, and on this basis the counting function, satisfies the general traffic flow control demand. In this paper, the frame difference method is used to extract moving vehicle, can effectively do vehicle detection and counting of the given video sequence.

Acknowledgment

This works are partly supported by the Science and Technology Research Project of Liaoning Provincial Department of Education (L2012479), Research Project of State Ethnic Affairs Commission (12DLZ018) and the Fundamental Research Funds for the Central Universities (DC120101132).

References