Research on Vehicle Detection Algorithms in Surveillance Video Images

Linglong Tan¹ *, Wenju Zhang¹ and Yehui Chen¹

¹Electronic communication engineering college, Anhui Xinhua University, Hefei, 230088, China
*tanlinglong@axhu.edu.cn

Abstract. Intelligent traffic video surveillance system, including vehicle inspection, vehicle tracking and license plate recognition module. The video vehicle detection in this paper is divided into four steps: video image pre-processing, background modeling and updating, foreground detection and adhesion vehicle area segmentation. After the collected video sequence is subjected to mean-value down-sampling and gray-scale conversion preprocessing, the background model is built using a mixture of Gaussian models. A sliding window is set on the time axis, and the grayscale video frame image in the sliding window is median-filtered to output a background gradient map. By merging the results of the spatial background difference between the color space and the grayscale gradient, the foreground image detected by the vehicle is obtained. Finally, the background difference is obtained in the gradient domain to obtain the foreground image to achieve vehicle detection.

1. Introduction

With the development and progress of today's society, the monitoring of traffic scenes has also become a subject of concern for a person. It is common in social life and people's daily lives to monitor traffic scenes by connecting cameras to computers. However, today's science and technology are developed. It is not only inappropriate for people to analyze videos, but also to spend a lot of manpower and material resources. Due to the development of technology, the use of computers to process video images has become an increasingly common method. Its universal application scope and future commercial value have drawn great attention from academic and business circles at home and abroad.

At present, our monitoring of traffic monitoring sites is mainly achieved through computer connection monitoring systems. Based on computer and image processing technology, people have established an intelligent traffic monitoring and management system, which basically does not require human control in order to achieve the goal of saving manpower and material resources. Through the analysis of video images, the positioning and identification of vehicles can be achieved. While performing routine management tasks, it can also take the most safe and effective measures when abnormal conditions occur, thus providing a more advanced and feasible monitoring program. In order to solve these problems, the supervising department used video cameras to take pictures in some places, and then used manual analysis and management.

2. Vehicle Detection Algorithm Research
2.1. Image Preprocessing

Due to the huge difference in imaging features of the scene images, for example, the probability of discrimination is very low, the sizes are different, and the illumination is uneven. These features affect the positioning of moving objects, image cutting to character recognition and other processes. Before inputting the text image under the premise of the same scene into each module, the video image must undergo a necessary pre-processing process, which will help to improve the accuracy of positioning and recognition.

(1) Expansion and corrosion

Inflation [1,2] refers to the process of merging all the background points in contact with an object into the object and expanding the boundary line to the outside. Use this process to fill the space in the object.

Assume that A and B are sets in \( \mathbb{Z}^2 \), \( \emptyset \) is an empty set, \( \overline{B} \) is an image of B, B performs an expansion operation on A, and it is denoted as \( A \oplus B \), \( \oplus \) is an expansion operator and expands to:

\[
A \oplus B = \{ x \mid [\overline{B}], 1 \ A] \neq \emptyset \}
\]

(1)

The expansion process shown in Equation (1) is that B first maps the far point and then translates x. The expansion operation performed by B on A is that \( \overline{B} \) is translated by all X and A has at least one non-zero common element, and the definition can be changed to:

\[
A \oplus B = \{ x \mid [\overline{B}], 1 \ A] \neq \emptyset \}
\]

(2)

Corrosion [3, 4] is a process that eliminates boundary points and shrinks the boundaries to the inside. With this operation, small, non-meaningful objects can be eliminated.

Assume that A and B are collections in \( \mathbb{Z}^2 \). A is corrupted by B, and it is written as \( A \Theta B \). It is defined as:

\[
A \Theta B = \{ x \mid (B), \in A \}
\]

(3)

That is to say, the result of A being corrupted by B is the set of points x that all include A after B is translated by x.

(2) Filter enhancement of images

The purpose of the panning and sliding operation is to eliminate or minimize noise interference as much as possible and improve the quality of the image. If the additive noise is randomly distributed, the use of the average or weighted average of the neighborhoods can effectively suppress noise interference.

If S is the neighborhood of pixel \((x_0, y_0)\) (including \((x_0, y_0)\)), where \((x, y)\) represents the element in S, \(r(x, y)\) represents the gray value of point \((x, y)\), and \(a(x, y)\) represents the weight of each point, smoothing \((x_0, y_0)\) can be expressed as:

\[
r' (x_0, y_0) = \frac{1}{\sum_{x,y} a(x, y)} \left[ \sum_{x,y} a(x, y) r(x, y) \right]
\]

(4)

In general, the weights are symmetrical with respect to the middle point. For the following 3*3-size templates, their weights are equal:

\[
T = \frac{1}{5} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}
\]

(5)

The corresponding function expression for this template is:

\[
r'(x, y) = \frac{1}{5} [r(x, y - 1) + r(x - 1, y) + r(x, y) + r(x + 1, y) + r(x, y + 1)]
\]

(6)

Can be expressed as:

\[
r'(x, y) = \frac{1}{5} [1 \times r(x, y - 1) + 1 \times r(x - 1, y) + 1 \times r(x, y + 1) + 1 \times r(x, y + 1)] = T \times r
\]

(7)

In other words, the neighborhood operation can be obtained by the convolution of the neighborhood and the template, which is also very convenient for calculation.
2.2. Detection Algorithm for Moving Vehicles

(1) Inter-frame difference method

The inter-frame difference method [5, 6, 7] is a method for obtaining a target (vehicle) surface in motion by performing a differential operation on a moving target between two adjacent frames in a video surveillance image. This shows that if the difference between each frame is relatively large, then it can explain the existence of the moving target; conversely, if the difference between each frame is relatively small, then it can indicate that the moving target does not exist. The difference here mainly refers to the continuity and independence between images and images. The principle of this calculation method is shown in Figure 1. That is, the extracted image is first subjected to image enhancement processing, ie, filtering, to remove various noise disturbances in the image, and then the image is subjected to binary processing and differential processing. In order to detect the goal in motion.

Figure 1. Schematic diagram of difference between frames

The inter-frame difference method has two frame methods and three frame methods. The two-frame method is to process dynamic objects, and the binary and differential images must be processed first. The two-frame difference method works as follows:

The original image is denoted as $f_i(x, y)$. Here, $i=1, 2, ..., n$, which represents the number of image frames, denotes the image after the difference and the binarized image as $D_i(x, y)$ and $T_i(x, y)$, respectively. The two-frame difference method can be expressed as:

$$D_i(x, y) = f_i(x, y) - f_{i-1}(x, y)$$

$$T_i(x, y) = \begin{cases} 1, & D_i(x, y) \geq T \\ 0, & D_i(x, y) \leq T \end{cases}$$

(8)

(9)

Although this method captures the target in motion, it has a significant effect, but its requirements are also more demanding. It requires the target to be in a state of motion, and its brightness changes must be very obvious. If the target is stationary, it is very easy to catch the target information. This will make the moving target unable to guarantee its continuity.

The three-frame difference algorithm is used in three-frame images. This method is used if two consecutive frames are overlapped in the detection target. Its specific operations are as follows: First, each frame image is processed with various grayscale processing methods to obtain a corresponding grayscale image; then, the processed image is binarized to obtain a corresponding binary image;

Finally, the target region is extracted on the binarized image. The expression is as follows:

$$D_{i-1}(x, y) = (k - 1, k) \begin{cases} 1, & f_i(x, y) - f_{i-1}(x, y) \geq T \\ 0, & \text{otherwise} \end{cases}$$

(10)

$$D_i(x, y) = (k, k + 1) \begin{cases} 1, & f_i(x, y) - f_{i-1}(x, y) \geq T \\ 0, & \text{otherwise} \end{cases}$$

(11)

$$D_{i+1}(x, y) = (k, k + 1) \begin{cases} 1, & f_i(x, y) - f_{i-1}(x, y) \geq T \\ 0, & \text{otherwise} \end{cases}$$

(12)

In the formula, $f_i(x, y)$, $f_{i-1}(x, y)$, and $f_{i+1}(x, y)$ respectively represent consecutive three-frame images, $D_{i-1}(x, y)$ and $D_i(x, y)$ represent grayscale images of three consecutive frames of images, and T is the binarization threshold. In the formula, K and D represent the intersection of two frames of images, ie coincident portions, which represent the moving target.

(2) Background Difference Method

The background difference method [8, 9] is a method for testing a moving object by comparing the current frame with a background reference model in a video surveillance image, and it relies on a background modeling technique.
In the target detection of the motion based on the background difference method, the accuracy of the model image and the simulated background image will have a great influence on the test result. However, any moving object detection algorithm must meet the requirements of any image scene processing. However, it is sometimes difficult to achieve background modeling because of large motion scenes, unpredictability, external environment interference, and different noise. Therefore, the research in this paper is based on the background difference algorithm in the background. Let \((m, n)\) be the plane coordinates of the two-dimensional digital image. The binary mathematics based on the background subtraction is described as:

\[
(D_m, n) = |I_m - B_m|
\]

\[
M_m = \begin{cases} 
1, & D_m \geq T \\
0, & D_m < T 
\end{cases}
\]

\(I_m\) represents the grayscale image after the current frame transition in the video surveillance image, \(B_m\) represents the monitored background grayscale image of the current frame in the video surveillance image, \(M_m\) represents the subtracted binarization result, and \(T\) represents the grayscale image threshold.

3. vehicle detection algorithm simulation

3.1. Realization of the Difference Between Frames

Use \(mmread()\) function to read video files, and \(nframes = get(videoObj, 'NumberOfFrames')\) function is to get the number of video file frames. After reading the video file, use \(currentFrame = read(videoObj, k)\) to read any frame video image, and then use function \(grayFrame = rgb2gray(currentFrame)\) to grayscale it to obtain a grayscale image. Subtract the two adjacent images to obtain the result of the difference between the frames.

The following are the effect diagrams of the 19th and 20th frames, 79th and 80th frames, and 139th and 140th frames of the selected video, respectively, as shown in Figure 2.

![Figure 2. Experimental results of inter-frame difference](image-url)
3.2. Background Difference Method Implementation

The background difference method is to obtain the background, and then perform the difference operation on the background. Read the video file, use \texttt{m=rgb2gray(m)} to grayscale the read video file \cite{10}, obtain the grayscale image and then perform the differential operation on the adjacent two frames before and after to get the differential effect, as shown in Figure 3. For the aspects of operation, a vehicle detection interactive interface is designed in the MATLAB GUI, and the running effect is shown in Figure 4.

![Figure 3. Moving target detection image with background difference method](image)

4. Conclusion

The vehicle video detection in this paper is processed from video image preprocessing, background modeling and updating, foreground detection, and segmentation of stuck vehicle areas. The collected video image is subjected to mean sampling and grayscale conversion, and the extracted image is subjected to image filter processing to remove various noise interferences in the image. Then use the binarization method and differential processing method to detect the moving target in the image. In the design, the inter-frame difference method, background difference method and edge detection method were analyzed. Among them, the inter-frame difference method and the background difference method were implemented by subtraction. However, they are also different. The interframe difference method is a method for obtaining the target contour in motion by performing a differential operation on the moving target between two adjacent frames in the video surveillance image. The background difference rule is a comparison of the current frame and the background reference model in the video surveillance image.
Figure 4. vehicle detection interface

Acknowledgments
This work was supported by the Scientific Research Project of Anhui Xinhua College No. 2017zr003, Anhui Natural Science Research Project No.KJ2017A628 and the domestic outstanding young backbone talents in universities of visiting program No. gxgnfx2018070.

References
[1] Fan Yao, Chu Yanli. Study of a new vehicle detection algorithm based on linear CCD images[J]. Optik - International Journal for Light and Electron Optics, 2015, 126(24).
[2] Bao Xu, Li Haijian, Xu Dongwei, Jia Limin, Ran Bin, Rong Jian. Traffic Vehicle Counting in Jam Flow Conditions Using Low-Cost and Energy-Efficient Wireless Magnetic Sensors.[J]. Sensors (Basel, Switzerland), 2016, 16(11).
[3] Meng Xiaoli, Wang Heng, Liu Bingbing. A Robust Vehicle Localization Approach Based on GNSS/IMU/DMI/LiDAR Sensor Fusion for Autonomous Vehicles.[J]. Sensors (Basel, Switzerland), 2017, 17(9).
[4] Fan Yao, Chu Yanli. Study of a new vehicle detection algorithm based on linear CCD images[J]. Optik - International Journal for Light and Electron Optics, 2015, 126(24).
[5] Xiaorong Gao, Yingyong Wu, Kai Yang, Jinlong Li. Vehicle bottom anomaly detection algorithm based on SIFT[J]. Optik - International Journal for Light and Electron Optics, 2015, 126(23).
[6] Zhonghua Zhang, Xuecai Yu, Feng You, George Siedel, Wenqiang He, Lifang Yang. A Front Vehicle Detection Algorithm for Intelligent Vehicle Based on Improved Gabor Filter and SVM[J]. Recent Patents on Computer Science, 2015, 8(1).
[7] Hong He, Rui Zheng, Zhi Hong Zhang. Vehicle Detector Based on the AMR Sensor[J]. Applied Mechanics and Materials, 2014, 3082(543).
[8] Y H Lin, R Bai, Z H Qian. The design method and research status of vehicle detection system based on geomagnetic detection principle[J]. Journal of Physics: Conference Series, 2018, 986(1).
[9] Zi Yang, Lilian S.C. Pun-Cheng. Vehicle detection in intelligent transportation systems and its applications under varying environments: A review[J]. Image and Vision Computing, 2018, 69.
[10] Yingfeng Cai, Ze Liu, Xiaqiang Sun, Long Chen, Hai Wang, Yong Zhang, Mi Zhang. Vehicle Detection Based on Deep Dual-Vehicle Deformable Part Models[J]. Journal of Sensors, 2017, 2017.