



Research on Road Traffic Moving Target Detection Method Based on Sequential Inter Frame Difference and Optical Flow Method

Da Li^{1(✉)}, Zhi-wei Guan^{1,2}, Qiang Chen¹, He-ping Shi¹, Tao Wang¹,
and Huan-huan Yue¹

¹ Tianjin University of Technology and Education, Tianjin 300222, China
leedtute@163.com

² Tianjin Sino-German University of Applied Sciences, Tianjin 300350, China

Abstract. In order to realize the effective detection of moving objects in intelligent transportation environment, the proposed method of motion detection with higher accuracy and efficiency by combining the sequence frame difference and optical flow algorithm. Firstly, adaptive median filter and histogram equalization are used to preprocess video frames. Then, the preprocessed image is processed by pairwise sequence difference, and the sum of the difference results is calculated. Next, the optical flow method is used to process the differential video sequence to accurately detect and recognize the moving target. Finally, the proposed algorithm is compared with background difference method, optical flow method and inter frame difference method, and the detection effect and accuracy of moving objects in different scenes are analyzed. The experimental results show that the overall performance of the proposed algorithm is stable, the overall contour recognition of moving objects is high, the average true positive rate of moving objects detection is above 0.9.

Keywords: Intelligent transportation · Frame difference · Optical flow · Target detection

1 Introduction

Due to the increasingly serious traffic problems of urban roads and the urgent needs of people to travel, my country's annual investment in intelligent transportation construction has increased, and the pressure on urban intelligent transportation management systems has become increasingly serious. Using artificial intelligence technology to effectively identify vehicle information such as violations and traffic accidents during vehicle driving, and to quickly and accurately detect targets on the road ahead is the key to the development of smart road traffic monitoring systems in the future.

At present, the detection of moving objects in the traffic environment is mainly realized by traditional moving object detection algorithms and moving object detection algorithms based on deep learning. The traditional moving target detection algorithm mainly uses the pixel information of the video image to process the image frame to obtain the

relevant moving target, which is divided into background difference method [1], optical flow method [2], and frame difference method [3] etc. This type of algorithm is relatively simple and easy to implement, and has low requirements on the hardware of the device, but it is greatly interfered by dynamic factors in the background, and the detection effect based on a single motion feature is general. Moving target detection algorithms based on deep learning mainly include deep learning algorithms represented by convolutional neural networks [4] and YOLO [5]. This type of algorithm has a large amount of calculation, high requirements for computer monitoring hardware, and complex model design, which has certain shortcomings in the real-time detection.

In view of the problems that traditional algorithms are susceptible to interference from dynamic environmental factors, deep learning algorithms are too computationally expensive and require high hardware equipment. Based on the improvement of the traditional target detection algorithm, after preprocessing the collected video frames, a new moving target detection algorithm is proposed by combining the improved inter frame difference method and the optical flow method.

2 Video Frame Preprocessing

Adaptive median filter improves the defect that the median filter adopts fixed processing method for different pixels in the process of noise removal, which is easy to damage the information of some pixels. In the case of high noise, the neighborhood size of the median filter is changed according to the preset conditions, which can reduce the noise and ensure the effect of image details.

Histogram equalization determines the transformation function according to the image histogram information, and uses the transformation function to approximately evenly distribute the gray value of the input image, so as to widen the gray value with more pixels in the image and merge the gray value with less pixels, which can make the image have a larger dynamic range of gray and higher contrast, and the details of the image are clearer, Enhance the visual effect of the image.

3 Optical Flow Method

Optical flow represents the instantaneous speed of motion when a moving object in space is shot and projected onto a two-dimensional plane by a camera. The L-K optical flow method [6] is based on the premise that the brightness of the image acquisition is constant, the change of motion within a small time interval will not cause the change of the position of the imaging object, and the change of the optical flow in the neighborhood of the pixel area is constant. The basic optical flow equation in the pixel neighborhood is solved by the least square method. Using the correlation between the current frame and the previous frame of the video image, through the temporal change of the same pixel in two adjacent frames, the movement of the pixel on the image is determined according to the change in the image caused by the movement of the target object in a small time Direction and rate of movement.

4 Sequential Frame Difference Method

The proposed method of motion detection is improved on the basis of the three frame difference method. Through continuous difference of the adjacent three frames of the moving target video, the sequence inter frame difference summation method is proposed to solve the inter frame difference method and the three-frame difference Problems with the difference method. By performing inter frame difference between three consecutive frames of the video image sequence, the difference results are added together, and the small and subtle motion results in the moving image are reflected on the final processing result. The final image processing effect is better than the three frame difference The method has a certain improvement, and a better detection effect is obtained.

Suppose that the gray values of the pixels corresponding to P_{n-1} , P_n and P_{n+1} of frames $n - 1$, n and $n + 1$ are $f_{n-1}(x, y)$, $f_n(x, y)$ and $f_{n+1}(x, y)$ respectively, and that the difference image obtained by the inter frame difference between the current frame image and the next frame image is Z_n and Z_{n+1} , then:

$$\begin{aligned} Z_n &= |f_n(x, y) - f_{(n-1)}(x, y)| \\ Z_{n+1} &= |f_{n+1}(x, y) - f_n(x, y)| \end{aligned} \quad (1)$$

In this paper, the OTSU is used for image binarization threshold segmentation. Select the appropriate threshold T to binarize the difference image, and the difference image is as follows:

$$\begin{aligned} g_n &= \begin{cases} 1, & Z_n \geq T \\ 0, & Z_n \leq T \end{cases} \\ g_{n+1} &= \begin{cases} 1, & Z_{n+1} \geq T \\ 0, & Z_{n+1} \leq T \end{cases} \end{aligned} \quad (2)$$

When the pixel value of the difference image is greater than or equal to the preset threshold value, the result of binarization is 1, which means the pixel value of the object. Otherwise, it means that the area is the background and there is no object. After the difference between two frames, the sum of the image is calculated:

$$R = g_n + g_{n+1} \quad (3)$$

The advantage of the summation operation is that under the condition that the specific position of the detection target remains unchanged, the frame difference images are added, and the moving objects on the two consecutive frame difference images are extracted respectively, which further expands the contour of the target object and improves the detection accuracy.

5 Sequential Frame Difference Method and Optical Flow Method

Because the image area extracted by the inter frame difference method is incomplete, although the position information of the target contour can be obtained, holes are prone to appear inside, and the extraction results are not accurate. In addition, when the speed of the moving object is too fast or too slow, it is easy to cause the moving object cannot

be detected, and the detection result is prone to ghosting, which affects the accuracy of the detection result.

The L-K optical flow method has a large amount of calculation and a slow running speed, resulting in poor real-time detection. In the detection, the hardware requirements of the computer are often high. Moreover, the optical flow method has many constraints, which has a greater impact on the change of light intensity, and its robustness is poor. At the same time, the inter frame difference method has good real-time performance in target detection, good adaptability to environmental changes, small amount of calculation, high detection accuracy of the optical flow method, and wide application range. The sequence inter frame difference method is combined with the optical flow method. The flow method for moving target detection can make up for the shortcomings between the two to a certain extent, and improve the detection efficiency and detection accuracy of traffic moving targets in the smart road environment.

In the algorithm processing, first carry out the sequence frame difference method adopted in this paper to roughly extract the approximate outline of the target object, and obtain the approximate area of the target object. On this basis, perform L-K optical flow calculation processing to reduce the influence of light factors on the detection result. In order to improve the progress of detection, reduce the amount of calculation of optical flow detection. The specific process is shown in Fig. 1.

The specific steps of the algorithm are as follows:

1. The video image is preprocessed, including the adaptive median filter of image sequence frame to reduce the overall noise of the image, and then histogram equalization to make the image have a larger dynamic range of gray and higher contrast;
2. Three consecutive image sequences P_{n-1} , P_n and P_{n+1} are obtained by frame extraction;
3. The obtained images P_n and P_{n-1} are differentiated to obtain a differential image g_n ; The P_{n+1} and P_n images are differentiated to get g_{n+1} ;
4. For the difference images g_n and g_{n+1} , the threshold T corresponding to the maximum inter class variance is obtained by using the OTSU, and the binary threshold segmentation is performed to obtain the moving region I of the target image;
5. Taylor expansion is performed on the pixels in the moving region of the target image, and
$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} \end{bmatrix}^{-1} \begin{bmatrix} -\frac{\partial I}{\partial t} \Delta t \end{bmatrix}$$
 is performed on each pixel to obtain the optical flow $(\Delta x, \Delta y)$ of each pixel in the moving region, and then the optical flow field of the moving region is obtained;
6. The morphological closed operation is performed on the image that detects the optical flow field of the image, so as to bridge the discontinuity of the image in the narrow area, the whole image is incomplete, and fill the small image hole.
7. According to the binary image obtained after processing, it is judged that the position with large amplitude is the foreground, that is, the moving target, and the area with small amplitude is the background, that is, the static background;

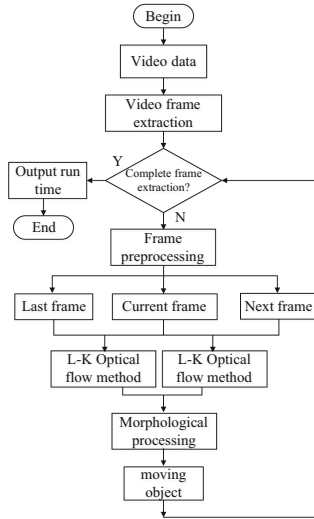


Fig. 1. The algorithm flow chart of this paper

6 Experimental Analysis

The hardware environment of this paper is based on Windows platform, the processor is Intel Core i5-7500, frequency is 3.40 GHz, and the software environment is matlab. Video image is mainly video image in MATLAB software, and the camera installed at high place to take driving image of road vehicles.

Figure 2, 3, 4 shows the typical detection results obtained from three video images using different algorithms. According to the detection results, it can be seen that errors in the comparative analysis of background model and foreground in the process of vehicle detection by background difference method are easy to cause misjudgment of foreground detection results, and the virtual image of vehicle under illumination is judged as a moving target. The detection results of inter frame difference method are incomplete and can not reflect the overall contour of the vehicle. In optical flow detection, the road contour is determined as the foreground, which has a deviation from the actual results. Under the same conditions, the proposed algorithm has good detection effect, the extracted target contour is relatively complete, and the internal cavity is well filled.

In order to further verify the overall performance of the proposed algorithm in target detection, the correct detected foreground pixels (TP), the correct detected background pixels (TN), the wrong detected foreground pixels (FP) and the wrong detected background pixels (FN) are counted respectively [7]. Using the formula:

$$TPR = \frac{TP}{TP + FN} \tag{4}$$

$$FPR = \frac{FP}{FP + TN} \tag{5}$$

The true positive rate (TPR) represents the ratio of all pixels detected as foreground that are correctly judged as foreground. The false positive rate (FPR) represents the ratio

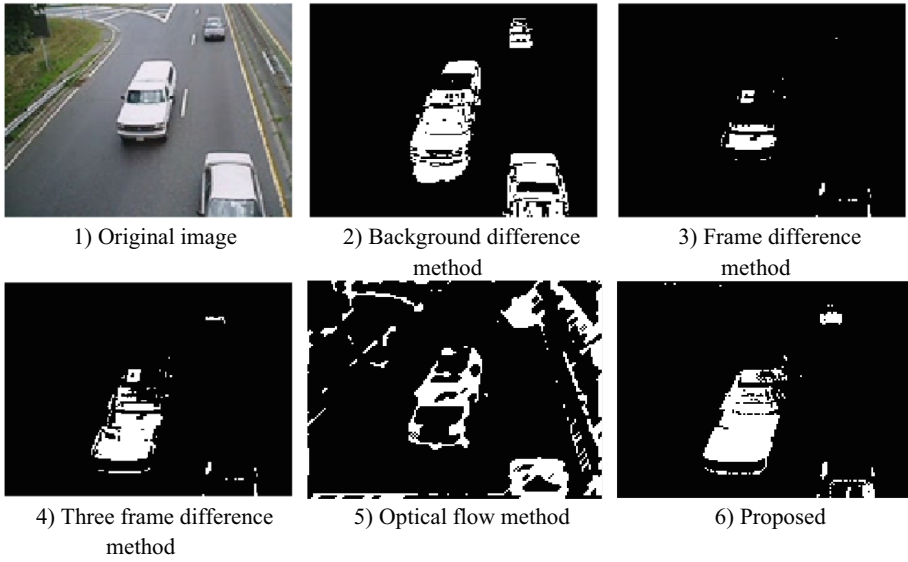


Fig. 2. Multi-vehicle detection

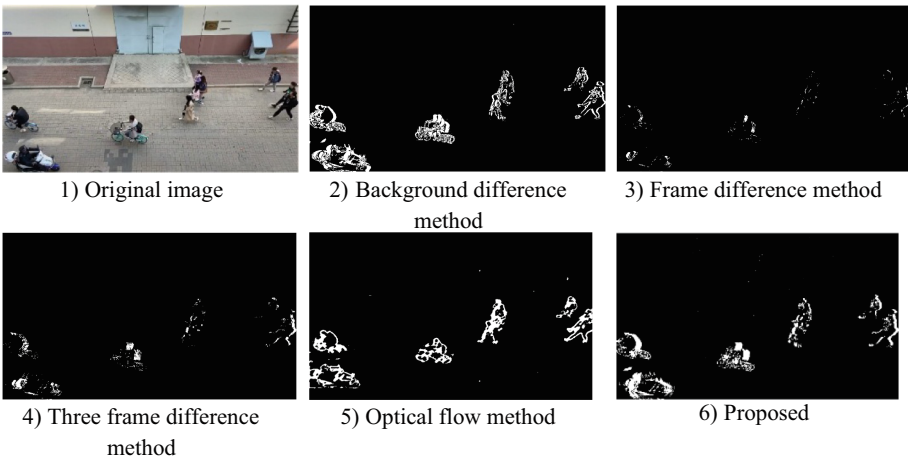


Fig. 3. Pedestrian target detection

of all detected background pixels that are correctly judged as background. The final statistical results are shown in Tables 1, 2:

In the comparison results of the detection performance of each algorithm in different traffic scenarios, the algorithm proposed in this paper has higher overall recognition than other detection algorithms, good detection stability, and the pixel values of the target foreground pixels that can be correctly detected are all greater than 0.88, The detection

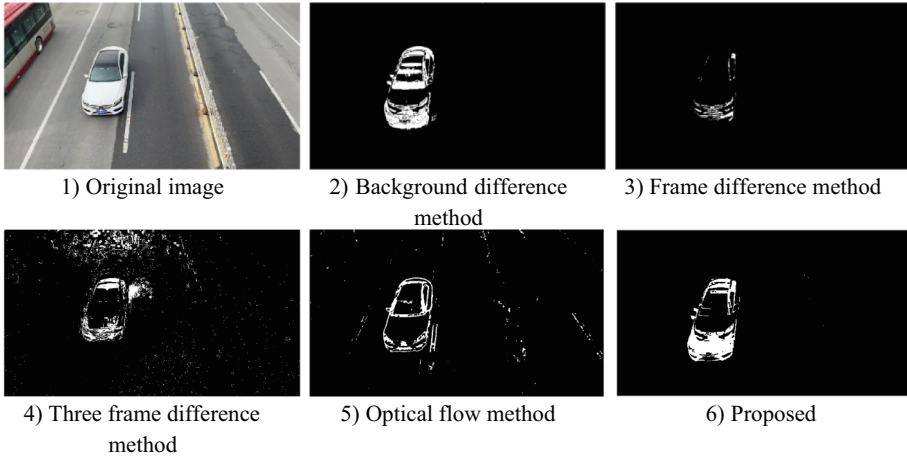


Fig. 4. Single vehicle detection

Table 1. True positive rate of different detection algorithms in each video (TPR)

Method	Background difference method	Frame difference method	Three frame difference method	Optical flow method	Proposed
Video1	0.92603	0.63243	0.84407	0.76074	0.88762
Video2	0.89598	0.64879	0.73412	0.90393	0.92810
Video3	0.86958	0.60368	0.78457	0.89990	0.93962

Table 2. False positive rate of different detection algorithms in each video (FPR)

Method	Background difference method	Frame difference method	Three frame difference method	Optical flow method	Proposed
Video1	0.00547	0.00344	0.00957	0.13522	0.00398
Video2	0.01471	0.06008	0.03721	0.01143	0.02010
Video3	0.02127	0.04326	0.09117	0.03073	0.01361

effect of the target is good. The accurate judgments for background pixels are all below 0.02, which can clearly distinguish the foreground and the background.

In this paper, the moving target detection results based on the combination of sequence frame difference and optical flow have a certain improvement in the overall effect of target detection. In terms of the recognition effect of moving targets, the foreground and background processing is more prominent. Combined with the real-time

overall analysis, the algorithm is superior to the comparison algorithm, which has good robustness and further improves the integrity of the target.

7 Conclusion

In this paper, the moving target detection combining the difference between the sequence frames and the optical flow is adopted. After the road traffic video image is pre-processed, the moving target is accurately extracted by combining the difference between the sequence frames and the optical flow method. Before the optical flow detection, the image processing of the frame difference is performed, which reduces the amount of calculation for the subsequent optical flow detection. To a certain extent, the efficiency of video-based target detection is improved. The experimental results show that the detection accuracy of the proposed algorithm is improved to a certain extent compared with the traditional algorithm. The overall detection effect of moving targets is good. The average true positive rate of the detection results is above 0.9. The accuracy of the algorithm has been improved.

Acknowledgements. This work is supported by the Tinajin Research Innovation Project for Postgraduate Students: Research on multi-sensor fusion vehicle detection algorithm in complex weather conditions (2020YJSS086); Tianjin Artificial Intelligence Science and technology major project (17ZXRGX00070); Key projects of scientific research plan of Tianjin Municipal Education Commission (2019ZD20); Tianjin Science and Technology Plan (20KPH-DRC00030); The University Foundation of the Tianjin University of Technology and Education (KJ1903).

References

1. Zeng, W., Xie, C., Yang, Z., et al.: A universal sample-based background subtraction method for traffic surveillance videos. *Multimed. Tools Appl.* **79**(31), 22211–22234 (2020)
2. Peng, Y., Chen, Z., Wu, Q., et al.: Traffic flow detection and statistics via improved optical flow and connected region analysis. *SIViP* **12**(1), 99–105 (2018)
3. Jia, R.D., Chen, Y.G., Zhang, Y.E.: Fusion algorithm of infrared and visible images based on frame difference detection technology and area feature. *Int. J. Comput. Appl.* **6**, 1–6 (2018)
4. Qin, Z., et al.: How convolutional neural networks see the world—a survey of convolutional neural network visualization methods. *Math. Found. Comput.* **1**(2) (2018)
5. Du, J.: Understanding of object detection based on CNN Family and YOLO. *J. Phys. Conf. Ser.* **1004**(1) (2018)
6. Bai, J., et al.: Research on LK optical flow algorithm with Gaussian pyramid model based on OpenCV for single target tracking (2018)
7. Fan, P., Lang, G., Yan, B., et al.: A method of segmenting apples based on gray-centered RGB color space. *Remote Sens.* **13**(6), 1211 (2021)