ScancePG
Science Publishing Group

Research on the Method of Quickly Finding the Pedestrian Area of Interest

Yu Chun-he, Dong Cai-Fang

School of Electronics and Information Engineering, Shenyang University of Aeronautics and Astronautics, Shenyang, China

Email address:

2643909528@qq.com (Yu Chun-he), 554804690@qq.com (Dong Cai-Fang)

Abstract: In pedestrian detection based on car video, it is necessary to quickly and accurately detect pedestrians. However, in the past the method that people used is exhaustive search. The method needs to detect all the areas in the picture, which is a waste of time. Based on the purpose of reducing the area to be tested, we use the edge detection and the principle of camera imaging, and run the test in MATLAB to get the method of finding the region of interest quickly. This method can reduce the retrieval area and shorten the retrieval time. The method can meet the requirements of real-time in pedestrian detection. Compared with the exhaustive search method, the number of windows that the method requires is 1/33 of the number of windows that the exhaustive search method requires. The detection speed of this method is several times higher than that of the exhaustive search method. It can be seen that the method is effective.

Keywords: Edge Detection, Camera Imaging Principle, Vehicle Video Pedestrian Detection, Center Horizontal Line

1. Introduction

In the pedestrian detection [1] based on vehicle video, the detection speed and accuracy requirements are high, and the detection is real time. The number of frames per second of the video is greater than 24 frames (the minimum number of frames recognized by the human eye is 24 frames per second). It takes a long time to detect all the regions in each frame. As a result, the detection rate is reduced and real-time performance can not be achieved. What people need to do is find the detection area (called the region of interest) that may contain pedestrians in the figure. It can also be said to remove the detection area without pedestrians. The ultimate goal is to reduce the detection time and improve the detection efficiency [2]. Methods of finding the region of interest of the image are: background difference method [3], interframe difference method [4] and so on. These methods are not suitable for pedestrian detection based on car video because they are used for static cameras when using the background difference method and the interframe difference method to extract the areas to be detected. However, in pedestrian detection based on car video, the camera is moving, and the background of the video is changing. In order to improve the

detection rate, we propose a method to extract regions of interest quickly. The method uses edge detection and camera imaging principles. The method finds the region of interest based on the intersection of the center and the edge of the horizontal line. After finding the area of interest in the image, we only need to detect whether there is a pedestrian in the region of interest. This can improve the detection speed and achieve real-time detection.

2. Edge Detection, Aperture Imaging Principle

2.1. Edge Detection

First, we make the picture gray. Secondly, we find the edge of the picture. Because pedestrians have edges, edge points in edge detection include pedestrian edge points. In this paper, we use the Canny operator to complete the edge detection. The Canny operator needs to be smooth first and then guided. The steps are:

a. The image is handled with a Gaussian filter;

b. We use the first order difference to calculate the direction and magnitude of the edge. The convolution template is as follows:

$$H_1 = \begin{vmatrix} -1 & -1 \\ 1 & 1 \end{vmatrix} \quad (1)$$

$$H_2 = \begin{vmatrix} 1 & -1 \\ 1 & -1 \end{vmatrix} \quad (2)$$

$$\varphi_1(m,n) = f(m,n) * H_1(m,n) \quad (3)$$

$$\varphi_2(m,n) = f(m,n) * H_2(m,n) \quad (4)$$

$$\varphi(m,n) = \sqrt{\varphi_1^2(m,n) + \varphi_2^2(m,n)} \quad (5)$$

$$\theta_\varphi = \tan^{-1} \frac{\varphi_2(m,n)}{\varphi_1(m,n)} \quad (6)$$

H1 and H2 in formula are horizontal template and vertical template; The $f(m,n)$ is a function of the image pixel; The $\varphi_1(m,n)$ is the horizontal component of the edge；The $\varphi_2(m,n)$ is the vertical component of the edge; The $\varphi(m,n)$ is the amplitude of the edge; The θ_φ is the directional angles of the edge.

c. Non-maximal suppression [5]; When only finding the global gradient direction, we can not determine the edge position. In order to highlight the real edge, we need to suppress some of the gradient value that is not the largest value. In the second step, the gradient direction is 3600, which is divided into four spaces. As shown in Figure 1:
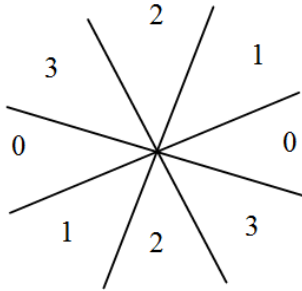

Figure 1. 4 Spatial Distributions.

Corresponding to the image, the pixel and its adjacent 8 pixels can be divided into four areas:

Figure 2. Corresponding to the Four Regions of the Image.

What we have to do now is to judge three pixels (center pixels, and two pixels in the other two gradient directions) along the gradient direction. When the gradient value of the center pixel is smaller than the gradient value of the other two pixels, the gray value of the pixel is set to zero.

d. Double threshold detection [6]; people apply the image that has been processed to two thresholds: th1 and th2. Th1 and th2 need satisfy th1 = 0.4 * th2.

The original figure is Figure 3. Figure 4 shows the result of the edge detection by using the Canny operator.

Figure 3. Original.

Figure 4. Results of the Edge detection.

2.2. The Principle of Camera Imaging

In this paper, we explain the principle of camera imaging by using a simple aperture imaging principle [7].

When the camera is parallel to the horizontal ground, there will be the following rules:

1) When the camera is parallel, the infinity of the ground is on the center of the picture

As shown in Figure 5. O is a small hole. O1 is the center of the picture. A, B and C are three points on the ground that are corresponding to A1, B1 and C1 points in the image. The horizontal distance from A to O is d1. The horizontal distance from B to O is d2. The horizontal distance from C to O is d3. And d1<d2<d3. In the image, the distance from the center to A, B and C is getting smaller and smaller. In other words, the point farther from the hole is closer to the horizontal centerline in the figure. When the distance is infinity, the point in the figure is on the horizontal centerline of the image.
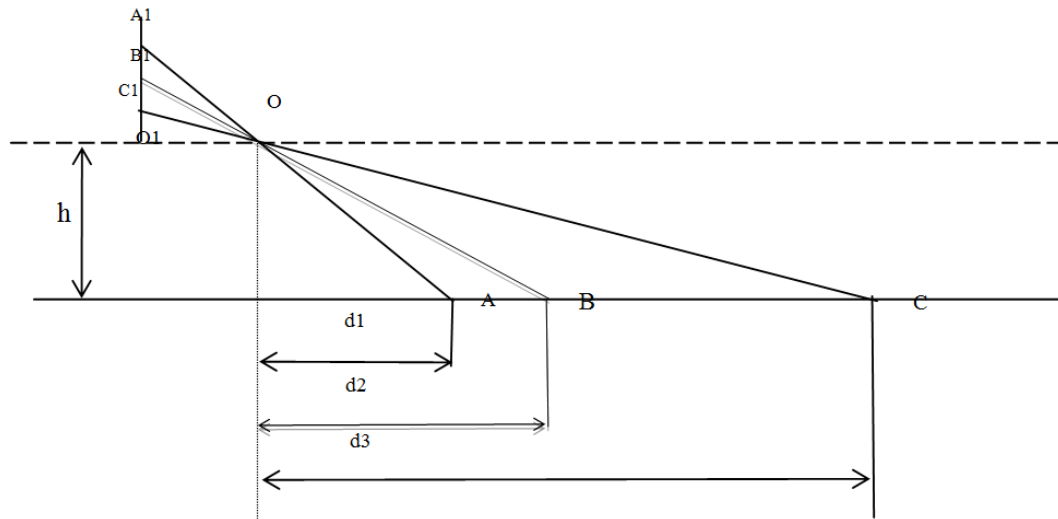
Figure 5. Pinhole Imaging Schematic (Parallel).

2) When the camera is parallel, points, the top of the object at the same height as the camera, are located on the horizontal centerline of the image.

As shown in Figure 6, O is a small hole. A, B, C, D and E are the top of objects that is the same height as the camera. In the picture, A, B, C, D and E correspond to A', B', C', D' and E'. A', B', C', D' and E 'are on the same horizontal line in the picture. Assuming that P is the point on the ground at infinity, corresponding to the point P'. The P' on the picture is also on this horizontal line. According to Principle 1, in the picture, P ' is located at the center level. So, when the camera is parallel, points, the top of the object at the same height as the camera, are located on the horizontal centerline of the image.
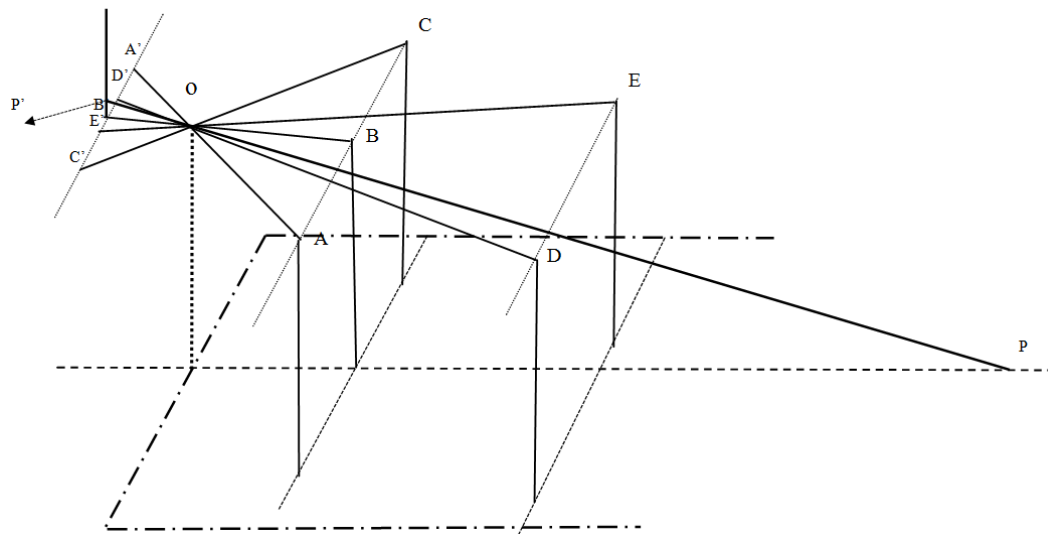

Figure 6. Imaging of the Same Height Object.

3. Reduce Pedestrian Detection Range

According to the above rules: When the camera is parallel, Point, the top of the object and the height of the same camera, located in the image of the horizontal center line. In pedestrian detection based on car video, the camera is placed horizontally. The camera height is set to 1 meter. When recording, objects over 1 meter can pass the horizontal centerline of the picture. The pedestrians on the road are more than 1 meter (less than 1 meter are childrens, which are accompanied by an adult). When the camera is parallel, the infinity road is located at the horizontal centerline of the picture, and the feet of the pedestrian is not possible to above the center line of the picture. So the pedestrian in the picture must intersect the horizontal centerline of the photo. When looking for the area of interest, what needs to be done is that you first need to perform edge detection on the image and get the edge point; then find the common point between the edge point and the horizontal centerline. These common points include not only the intersection of the pedestrian's edge and the horizontal centerline, but also the intersection of the other objects of 1 meter above the horizontal center line. Finally,

we use these intersections to find the area to be detected. Finding the detection area can exclude other areas that do not need to be detected and, to a large extent, reduce the detection range. Ultimately, we detect the area of interest and detect pedestrians, while effectively improving the detection efficiency.

The process of seeking a region of interest is:

a) The image gray processing;

b) Edge detection by Canny operator;

c) Find the intersection of the center horizontal and edge detection points, and find the intersection of the upper and lower horizontal and edge detection points of the center. Recorde the position of the intersection (X, Y);

d) According to the position recorded in step 3, we draw the area to be detected.
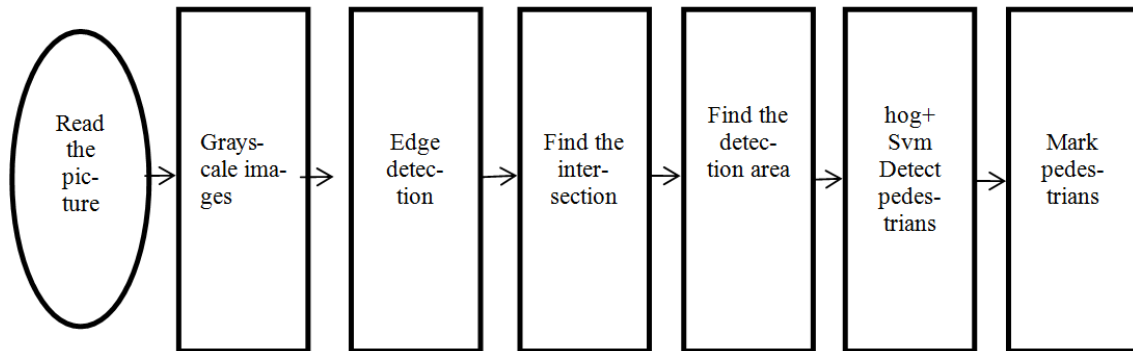
4. Detection Flow Chart

Figure 7. Flow Chart.

5. Comparison of Experimental Results

When using the methods in this article, we run the program in Matlab, and we find out the area to be detected. At this time, there are 55 detection areas, as shown in Figure 8.
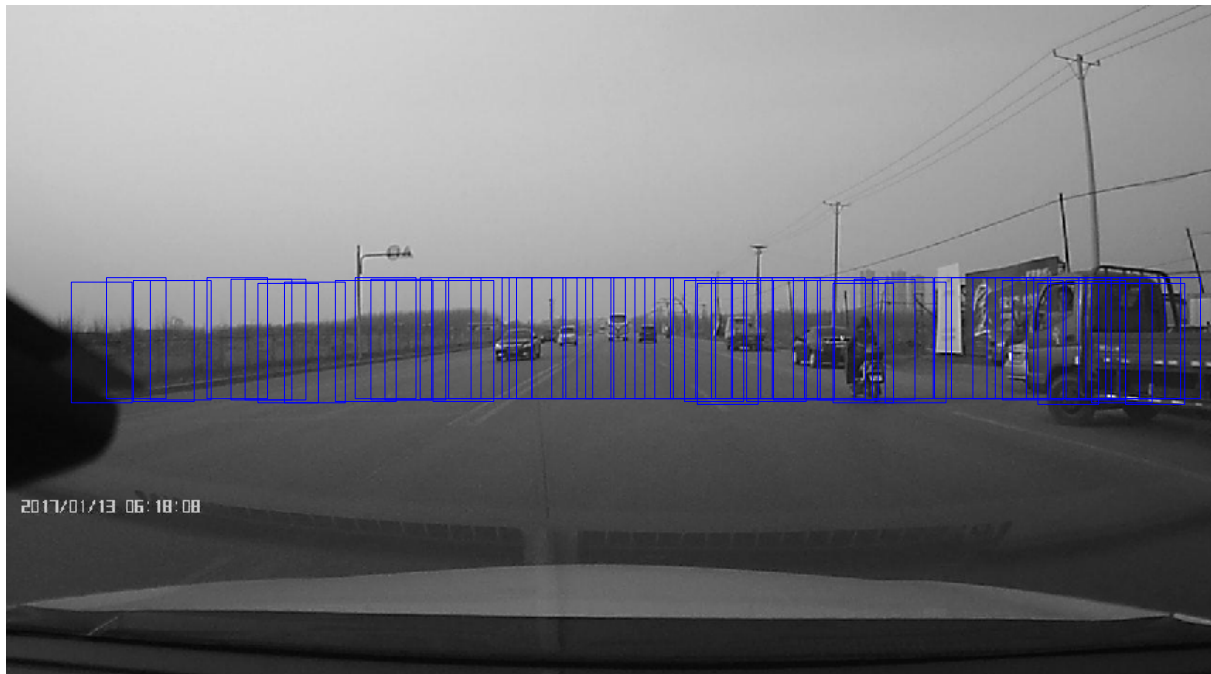

Figure 8. The Area to be Detected is Obtained by Edge Detection.

When using the detailed search method, we run the program in Matlab. We set the step size 20 pixels, and find out the area to be detected. At this time, there are 1830 detection areas, as shown in Figure 9.
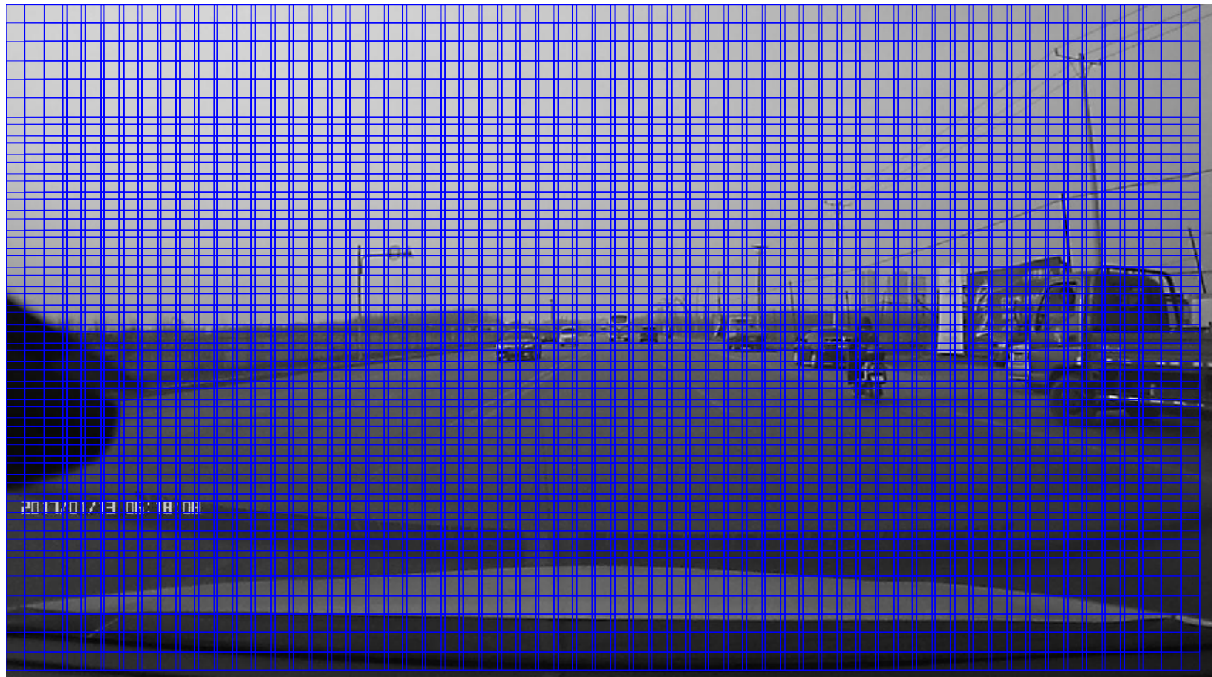
Figure 9. *The Window to be Detected is Obtained by Exhausting the Search.*

We used HOG [8] + SVM [9] to perform pedestrian detection on the detection window in Figure 8 and Figure 9. The results obtained are shown in Figure 10. The detection time of Figure 8 is 16.488694 seconds, and The detection time of Figure 9 is 2240.336819 seconds.

Figure 10. *shows the result graph by HOG + SVM.*

We get the results by comparing the two methods, as shown in Table 1:

Table 1. *Comparison Results.*

	Windows (number)	Detection time (seconds)
Edge detection method	55	16.488694
Exhaustive search method	1 830	2240.336 819

6. Conclusion

By comparing the results of the detailed search method and the results of the proposed method, we find that the edge detection method is better. This method has few detection windows and less detection time. The method is more conducive to pedestrian detection. The method can be used for self-driving driving, auxiliary driving [10] and so on. However, the study of this paper is imperfect, we should continue to learn. In the follow-up study, we would like to further reduce the number of windows to be detected by the body aspect ratio. We need to find the pedestrian zone and need to reduce the detection time. Our aim is to enable real-time detection.

References

[1] Xi Yao. Pedestrian Detection and Tracking Based on On-board Vision Systems [D]. Beijing: Beijing Institute of Technology. 2015.

[2] Jin Pei-fei, Zhou Li, Liu Jian, Ge Zhi-wei, Chen Jie. Pedestrian Detdction Based on Region of Interest Extracted by Support Vector Machine [J]. Computer Engineering and Desing, 2017, 38(4): 1 098-1 102.

[3] Zhang Yin－hui, Liu Yang－shuo. Moving Object Detection Based on Method of Frame Difference and Background Subtraction [J]. Computer Technology and Development, 2017, 27(2): 25-28.

[4] Li Liang, Luo Yi. Application Research of Inter-frame Difference in the Video Monitoring [J]. Journal of Sichuan University of Science & Engineering (Social Sciences Edition), 2015, 28 (6): 58-62.

[5] Zhang Fan, Peng Zhong-wei, Meng Shui-jin. Improved Canny edge detection method based on self-adaptive threshold [J]. Journal of Computer Applications, 2012, 32(8): 2296－2298.

[6] Duan Jun, Gao Xiang. Adaptive Statistical Filtering Double Threshholds Based On Improved Canny Operator Edge Detection Algorithm [J]. Laser Journal, 2015, 36(1): 10-12.

[7] Wang Xiao-kun, Jia Qing-xuan, Tan Sheng. Study on a monocular vision ranging method [J]. Science & Technology Information, 96-98.

[8] Dalal N, Triggs B. Histograms of Oriented Gradients for Human Detection [C] //Proc. of IEEE Conference on Computer Vision and Pattern Recognition. San Diego, USA: [s. n.], 2005: 886-893.

[9] Chen Yu-yuan, Guo Shu-qin, Lu Jun-jie, Zhang Biao. Pedestrian Detection in Video Monitoring Systems [J]. Journal of Hangzhou Dianzi University, 2014, 34 (1): 95-98.

[10] Zhnag Yang, Liu Wei-ming, Wu Yi-hu. A Fast Pedestrian Detection Method for Vehicle Auxiliary Driving System [J]. Journal of Highway and Transportation Research and Development, 2013, 30(11): 131-138.