

Research Article

An Effective Recognition Method for Road Information Based on Mobile Terminal

Zheng Kun,¹ Wang Wenpeng,¹ and Sun Guangmin²

¹ Department of Information Engineering, Beijing University of Technology, 719 Kexuelou, 100 Pingleyuan, Chaoyangqu, Beijing 100022, China

² Department of Electrical Engineering, Beijing University of Technology, 719 Kexuelou, 100 Pingleyuan, Chaoyangqu, Beijing 100022, China

Correspondence should be addressed to Zheng Kun; zheng.kun321@outlook.com

Received 5 August 2014; Accepted 11 October 2014; Published 10 November 2014

Academic Editor: Xue Chen

Copyright © 2014 Zheng Kun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper describes a design of fast recognition of road information based on mobile terminal. Firstly, based on the HOG algorithm, we study and verify the effects of different parameters on the performance of the algorithm. Secondly, we test 800 images randomly selected from the INRIA pedestrian dataset to obtain the optimal parameters for the mobile terminal and the proportion of video resolution and detection window. Then, under the same test conditions, the time overheads of the SVMLight and the LibSVM are recorded and SVMLight training time is significantly less than LibSVM. Thirdly, we design and implement a real-time road information recognition and warning system on the Windows platform and Android platform. Features include real-time pedestrians detection, voice warning, and road signs recognition. When the vehicle speed is less than 30 km/h, the video resolution is less than 720×576 and the detection window/image ratio is less than 1:50; the system can guarantee low delay and high recognition rate (97.2%).

1. Introduction

The World Health Organization reported that there were about 1240000 people who died in road traffic around the world each year [1]. According to the statistics of National Bureau of Statistics, vehicle ownership in China reached 137.4 million in 2013. In the last decade (2001–2010), nearly nine hundred thousand people in China died in various types of road accidents. The Chinese National Highway Traffic Safety Administration said that about ninety percent of all accidents happened because of the negligence of the driver. Road information collection and warning system in real-time becomes an urgent need. On the other hand, the numbers of smart phone users in China exceeded 500 million with all kinds of mobile terminals (mobile phones, tablets, and ultrabooks) growing popularly. If mobile terminals could combine with traditional vehicle auxiliary driving system, the additional value of mobile terminals can be increased and fill

the gap to a certain extent in the vehicle without auxiliary system. The traffic accidents caused by driver inattention are shown in Figure 1.

The number of the mobile phone users in China has been more than one billion. And 47 percent of mobile phone owners have used the intelligent mobile phone. According to the statistical data published by IDC in 2012 (Table 1), the market share of the Android mobile phone has been 68.8%, especially the number of Chinese Android mobile phones has been nearly 350 million. Table 1 is the statistics of the global mobile phone shipments and the share of major operating system.

2. Previous Work

2.1. Pedestrian Detection. Pedestrian detection system (PDS) has been one of the most active research topics in the field

TABLE 1: The statistics of the global mobile phone and the operating system (2012).

	Android	iOS	BlackBerry	Symbian	Win phone	Others
Shipments (million)	497.1	135.9	32.5	23.9	17.9	15.1
Market share	68.8%	18.8%	4.5%	3.3%	2.5%	2.1%

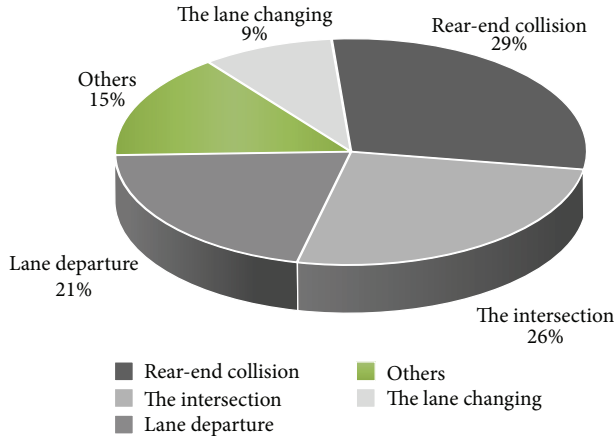


FIGURE 1: The traffic accidents caused by driver in attention.

of computer vision and intelligent vehicle. Using the camera mounted on the vehicle to detect pedestrians, it could give an early warning to drivers if a pedestrian found ahead was in some possible danger.

Broggi et al. [2] establish the model of the head and shoulders on the basis of the two value image templates with different size. Contrasting the edge of the input image with the template to identify pedestrian, the method has been used in a project named ARGO about smart car in Italy, Parma University. Collins et al. [3] present a method using optical flow, which calculates the residual flow in the moving area to realize the detection of pedestrians. Papageorgiou and Poggio [4] describe an object class in terms of an overcomplete dictionary of local, oriented, multiscale intensity differences between adjacent regions, efficiently computable as a Haar wavelet transform. Viola et al. [5] use AdaBoost to take advantage of both motion and appearance information to detect a walking person. Ronfard et al. [6] build on a general “body plan” methodology and a dynamic programming approach for efficiently assembling candidate parts into “pictorial structures.” Dalal and Triggs [7] propose a histogram of oriented gradient (HOG) algorithm and have proved that the algorithm is outstanding in pedestrian detection.

The HOG features are widely used for object detection. It represents the images’ local appearance and shapes by using local gradient or the edge of the direction of density distribution. The image is divided into small cells. Histogram of oriented gradient of each cell is computed and the result is normalized. The normalized block descriptor is called the HOG descriptor. Figure 2 is the main steps of the algorithm.

Walk et al. [8] propose an improved HOG algorithm based on histograms of flow (HOF) feature and color self-similarity (CSS) feature. And it uses HIK SVM classifier.

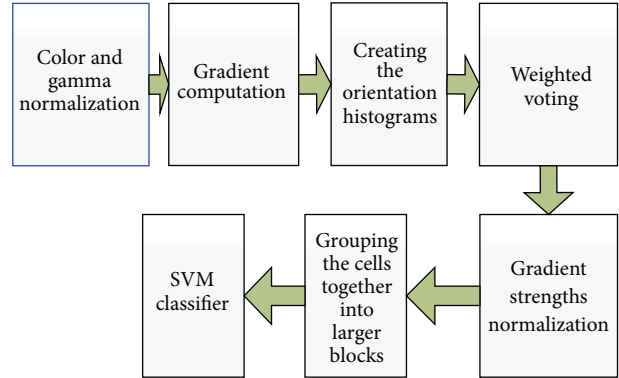


FIGURE 2: The main steps of the algorithm.

2.2. *Dataset.* Researchers in the field of pedestrian detection, in the course of the study, set up several open datasets for test. MIT pedestrian dataset, as shown in Figure 3, is relatively simple. The INRIA person dataset, as shown in Figure 4, is currently being used to test and evaluation, supporting complex background, and more posture of people.

3. Overview of the Method

We design a real-time road information recognition and warning system on the Windows platform and Android platform based on HOG algorithm. Features include real-time pedestrians detection, voice warning, and road signs recognition. Firstly, based on the HOG algorithm, we study and verify the effects of different parameters on the performance of the algorithm. Secondly, we test 800 images randomly selected from the INRIA pedestrian dataset to obtain the optimal parameters for the mobile terminal and the proportion of video resolution and detection window. Then, under the same test conditions, the time overheads of both the SVMLight and the LibSVM are recorded.

Figure 5 is the technology roadmap. The single camera, which is standard across all mobile devices, is placed in the vehicle to get road information. The road information includes pedestrian detection and road sign text extraction.

Figure 6 is the system solution. We consider that the vehicle driver’s attention should be firmly fixed on the road ahead. If the warning is just posted on screen, they might look at screen frequently, which would increase chance of drawing attention away from the task of driving. So we design the voice broadcast to reduce the risk.

Our experimental environment is in Tables 2 and 3. Table 2 is the hardware environment.

Table 3 is the development environment.



FIGURE 3: MIT pedestrian dataset.



FIGURE 4: The INRIA person dataset.

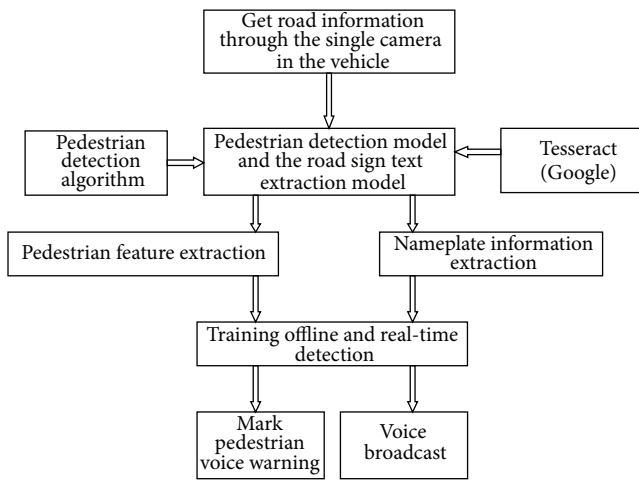


FIGURE 5: The technology roadmap.

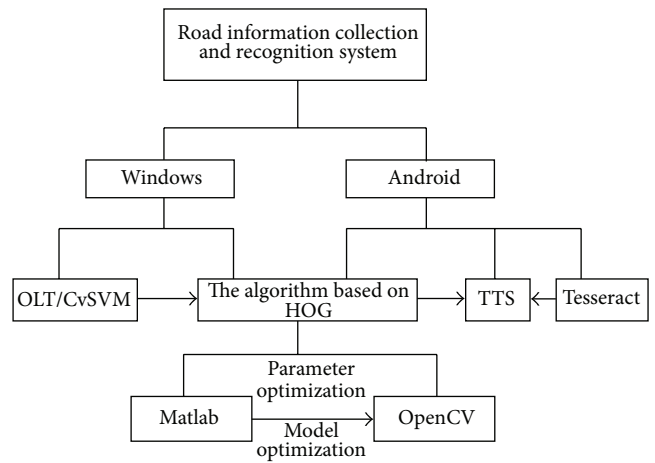


FIGURE 6: The system solution.

4. Kernel Algorithm for Pedestrian Detection

Figure 7 is the main results of pedestrian detection algorithm based on HOG.

4.1. *Gamma Correction.* Gamma correction can suppress noise to improve the robustness to partial shade. Gamma normalized as

$$I(x, y) = I(x, y)^\gamma. \quad (1)$$

We selected 100 images to test the image recognition rate for different values of the gamma. Figure 8 is a sample of gamma correction (normalized).

Table 4 shows the mistake rate and missing rate with different gamma value.

Experiment shows that when the sample size is 100 and gamma value is from 0.5 to 2, gamma normalization did not improve the performance of the algorithm.

4.2. *Gradient Calculation and the Template Selection.* The gradient of pixel (x, y) is defined as

$$G_x(x, y) = I(x + 1, y) - I(x - 1, y), \quad (2a)$$

$$G_y(x, y) = I(x, y + 1) - I(x, y - 1). \quad (2b)$$

TABLE 2: Hardware environment.

	Computer		Mobile
CPU	Intel(R) Core(TM) i7-2670@2.20 GHz		Exynos 4412@1.6 GHz
RAM	8.00 GB		2.00 GB
OS	Windows 8.1 64 bit	Ubuntu 12.04 64 bit	Android 4.1.1

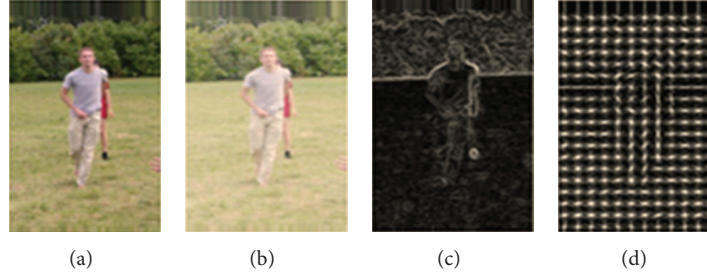


FIGURE 7: The main results of pedestrian detection algorithm based on HOG (64 * 128). (a) Input image; (b) gamma correction; (c) gradient image; (d) hog features.

TABLE 3: The development environment.

Windows	Android
Matlab R2013a	android-ndk-r9b-windows-x86
Visual Studio 2013 (Sp1)	OpenCV-2.4.9-android-sdk
OpenCV 2.4.9	Android Developer Tools

TABLE 4: The mistake rate and missing rate sample size (100).

Gamma value	Miss rate	Mistake rate
$\gamma = 0.0$	100%	Object not found
$\gamma = 0.5$	1%	3%
$\gamma = 1.0$	1%	3%
$\gamma = 1.5$	1%	3%
$\gamma = 2.0$	1%	3%

Gradient magnitude is computed using the following:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2}. \quad (3)$$

Gradient direction is computed using the following:

$$\theta(x, y) = \arctan\left(\frac{G_y(x, y)}{G_x(x, y)}\right). \quad (4)$$

$I(x, y)$ is the pixel value of pixel dot (x, y) . $G_x(x, y)$ is the value of horizontal gradient of pixel dot (x, y) . $G_y(x, y)$ is the value of vertical gradient of pixel dot (x, y) . $G(x, y)$ is the gradient magnitude of pixel dot (x, y) . $\theta(x, y)$ is the gradient direction of pixel dot (x, y) .

We use different gradient template processing the same image. And Figure 9 is the result of the edges of the image.

Processed with a simple template $[-1, 1]$, (b) does not fully display the outline of the man. (d) and (g), processed with $[1, -8, 0, 8, -1]$ and Sobel, respectively, are blurred severely. With the diagonal I template, the left of (e) is not

clear. And with the diagonal II template, the right of (f) is not clear. With a simple template $[-1, 0, 1]$, (c) is the best one because we can distinguish clearly between human contour and the background.

4.3. *Bin*. Using the INRIA person training set, we test the influence of the gradient space and the number of bins on recognition rate. Figure 10 shows that the number of bins in different directions influences pedestrian detection.

4.4. *Normalization and Descriptor Block*. Because of changing levels of local light changing levels of local light and changing contrast between the background and the foreground, the gradient changes in the very large scope. Local contrast normalization is used to deal with this question. The cells are taken into the larger blocks with contrast normalization to each block. The final descriptor is the histogram vector of cells in all blocks in the detection window. In order to improve the algorithm recognition rate, overlap is between blocks of 0.5.

Figure 11 shows that the influence of different block on pedestrian detection.

Cell size is the number of pixels of a cell. Block is composed of cell. The z -axis is the identification rate of pedestrian detection. Experiments show that pedestrian recognition rate is lower as cell becomes larger.

The performance of 2×2 or 3×3 block size is better than that of 1×1 or 4×4 block size. When block size is 2×2 and cell size is 8×8 ; the pedestrian recognition rate is the best.

4.5. *Detection Window*. The experiment involves 100 groups of images (64×128 detection window) and we perform detection in different video resolution. The time overhead of calculation of HOG features is recorded in Table 5.

By using the multiple threads, frame cache, and GPU technology, the system delay is not more than 453.1 ms.

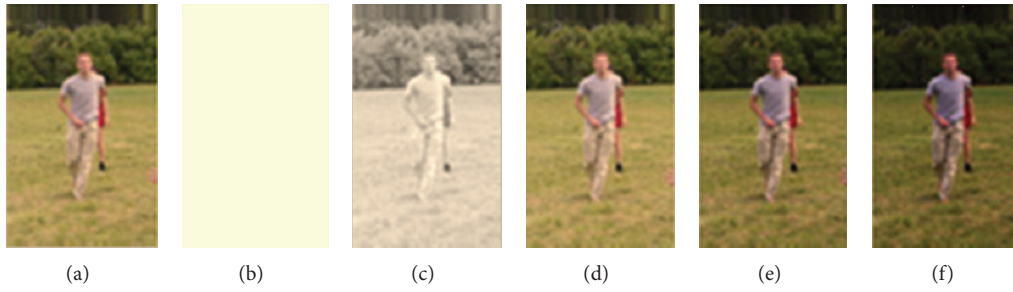


FIGURE 8: The correction results of different gamma value ($64 * 128$). (a) The original image; (b) $\gamma = 0$; (c) $\gamma = 0.5$; (d) $\gamma = 1.0$; (e) $\gamma = 1.5$; (f) $\gamma = 2.0$.

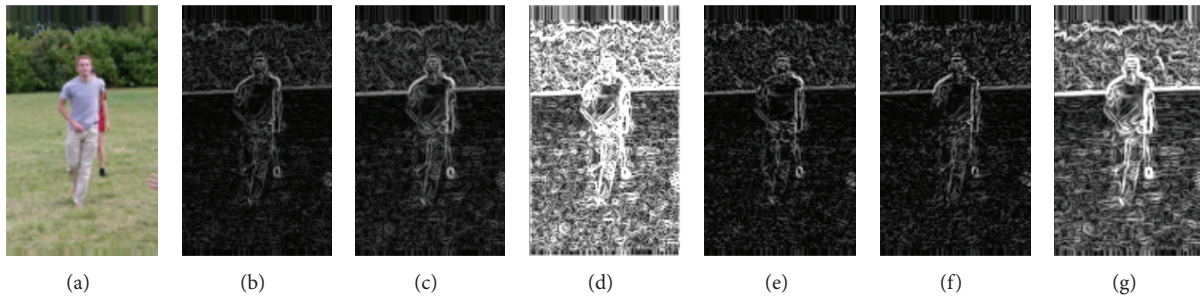


FIGURE 9: The edge after processing with the template. (a) The original image; (b) $[-1, 1]$; (c) $[-1, 0, 1]$; (d) $[1, -8, 0, 8, -1]$; (e) the diagonal I; (f) the diagonal II; (g) Sobel.

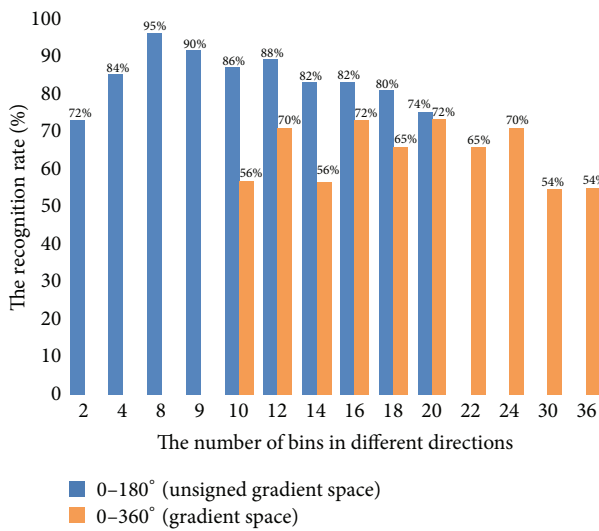


FIGURE 10: The number of bins in different directions influence on pedestrian detection.

Based on the above results, the conclusion is that as long as the video resolution is less than 720×576 and the detection window/image ratio is less than 1: 50, the system can guarantee low delay and high recognition rate.

4.6. SVM. Training 20 groups of HOG features under the same testing conditions, we get the time overhead statistical table about SVMLight and LibSVM as in Table 6.

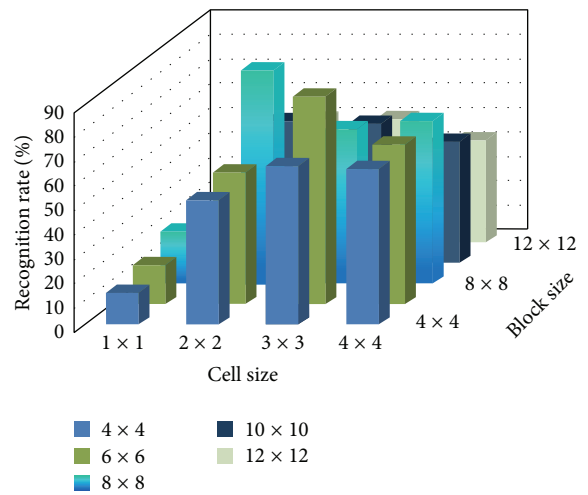


FIGURE 11: The influence of different block on pedestrian detection.

Experiments show that, in the condition that the number of the positive samples is equal to the negative and equal resolution, SVMLight training time is significantly less than LibSVM.

Some positive samples are in Figure 12.

Some negative samples are in Figure 13.

Some examples of the hard examples are in Figure 14.

4.7. Programme. Figure 15 shows the details of training and recognition. (a) is the drawing of training process. (b) is the schematic diagram of recognition.



FIGURE 12: Some positive samples.



FIGURE 13: Some negative samples.

TABLE 5: A statistical table of the cost of calculation of HOG features (64×128 detection window).

Resolution (pixel)	Moving blocks	Detection window	Time overhead (ms) (Matlab)	Time overhead (ms) (OpenCV)
176×144	105	3780	408.5	9.5
320×240	105	3780	814.9	62.6
352×288	105	3780	1013.4	91.1
640×480	105	3780	2808.5	330.0
704×576	105	3780	3622.9	427.9
720×480	105	3780	3081.1	359.0
720×576	105	3780	4233.1	453.1
1280×720	105	3780	8048.8	1028.2
1920×1080	105	3780	17962.7	2352.6



FIGURE 14: Some examples of the hard examples.

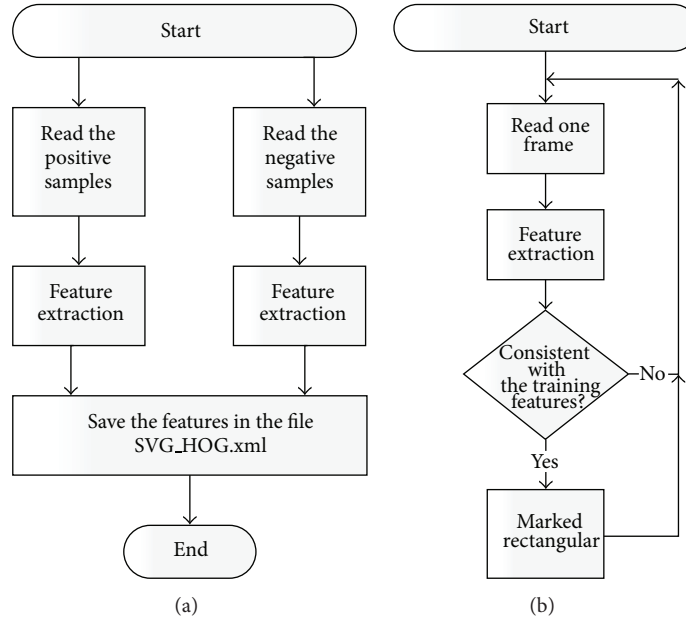


FIGURE 15: (a) Training; (b) recognition.

TABLE 6: Under the same testing conditions, the time overhead about SVMLight and LibSVM.

SVM training tool	Positive samples (frame)	Negative samples (frame)	Resolution (pixel)	Training time (S) required
SVMLight	2415	1210	96 × 160	104.3965
LibSVM	2415	1210	96 × 160	199.5718

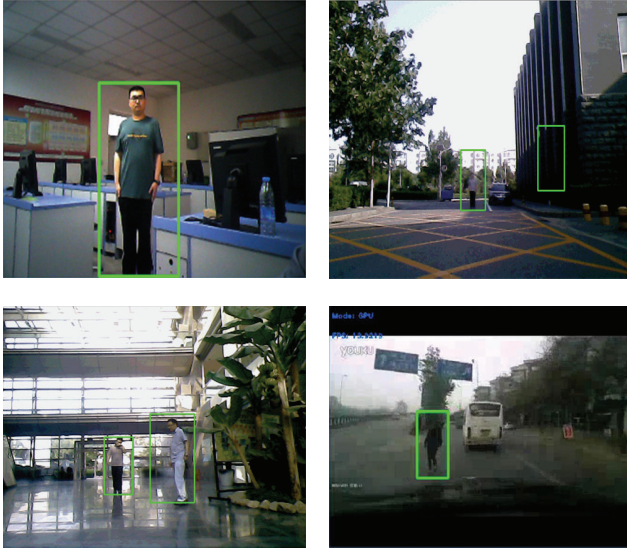


FIGURE 16: The actually measuring results.

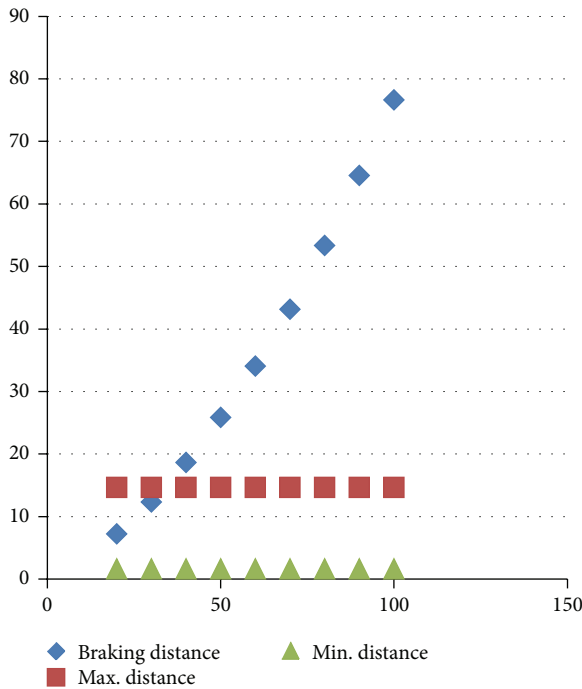


FIGURE 17: The corresponding relationship between speed and braking distance.

5. Performance Comparison of Training Tools

CvSVM is a training class of OpenCV. OLT (Object Detection and Localization Toolkit) is a HOG algorithm Toolkit, supporting Windows and Linux operating system.

We use the OLT to generate the final training file named Model_4BiSVMLight.alt. In the Ubuntu environment, training lasts for 1 hour and 13 minutes and lasts for 47 minutes in Windows. Using the data file trained by OLT to test the INIRA pedestrian dataset, we can take the correct rate increase to 97.2% (280/288).

6. Result

The actual test results show that the stability range of the pedestrian detection system is 2–12 meters; the valid range is 2–15 meters. The actually measuring results are in Figure 16.

Combined with the corresponding relationship between speed and braking distance (Figure 17), we can get the conclusion that when speed is less than 30 km/h, the system can support the driver effectively.

7. Conclusion

When block size is 2×2 and cell size is 8×8 , with a simple template $[-1, 0, 1]$, the pedestrian recognition rate is the best. When the vehicle speed is less than 30 km/h, the video resolution is less than 720×576 and the detection window/image ratio is less than 1:50; the system can guarantee low delay (483 ms) and high recognition rate (97.2%).

The recognition rate is low yet as soon as the object is partly occluded. And the supported speed is lower than average. So we still need further research to solve above questions.

Conflict of Interests

The authors declare that there is no conflict of interests regarding to the publication of this paper.

References

- [1] *Global Status Report on Road Safety 2013*, World Health Organization, Geneva, Switzerland, 2013, http://www.who.int/violence_injury_prevention/road_safety_status/2013.
- [2] A. Broggi, M. Bertozzi, A. Fascioli, and M. Sechi, "Shape-based pedestrian detection," in *Proceedings of IEEE Intelligent Vehicles Symposium*, pp. 215–220, Dearborn, Mich, USA, 2000.
- [3] R. T. Collins, A. J. Lipton, and T. Kanade, "A system for video surveillance and monitor-ring: VSAM final report," Tech. Rep. CMU-RI-TR-00-12, Carnegie Mellon University (cmu), 2000.
- [4] C. Papageorgiou and T. Poggio, "Trainable system for object detection," *International Journal of Computer Vision*, vol. 38, no. 1, pp. 15–33, 2000.
- [5] P. Viola, M. J. Jones, and M. Snow, "Detecting pedestrians using patterns of motion and appearance," in *Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV '03)*, vol. 2, pp. 734–741, Nice, France, October 2003.
- [6] R. Ronfard, C. Schmid, and B. Triggs, "Learning to parse pictures of people," in *Proceedings of the 7th European Conference on Computer Vision (ECCV'07)*, vol. 4, pp. 700–714, Copenhagen, Denmark, May 2002.
- [7] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 886–893, San Diego, Calif, USA, June 2005.
- [8] S. Walk, N. Majer, K. Schindler, and B. Schiele, "New features and insights for pedestrian detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'10)*, pp. 1030–1037, IEEE, San Francisco, Calif, USA, June 2010.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

