A Hybrid Method of Vehicle Detection based on Computer Vision for Intelligent Transportation System

Huan Wang and Haichuan Zhang

*School of Information Science and Engineering, Dalian Polytechnic University,
116034 Dalian, China
WHuan_dlpu@126.com, zhanghc@dlpu.edu.cn*

Abstract

In this paper, a two-step approach for vehicles detection is proposed. The first step of approach is to approximate vehicles' potential locations through searching for shadow area of vehicle low-part. In order to find these shadows, Haar-like feature with Adaboost was used to train a Haar detector offline and the relearning process with hard training samples is applied to increase detection rate. Based on the previous processing, ROI (Region of interest) + HOG + SVM algorithm is used for vehicle verification. At last, K-means approach is used to combine the similar detection results. The experimental results proved that our system could be used for real-time preceding vehicle detection robustly and accurately.

Keywords: Vehicle detection, ROI, Haar-like feature, HOG feature, SVM

1. Introduction

Recently, most of the cities of the world have intelligent transport system which is equipped with electronics devices to communicate about the traffic condition with the moving vehicle and also monitor the traffic rules and regulation. Traditional technology for traffic sensing, including inductive loop detectors and video cameras, are positioned at fixed locations in the transportation network. Data related to traffic flow is currently obtained from detectors embedded in pavements or pneumatic tubes stretched across roads. Such methods do not prove to be time-efficient or cost effective. While these detectors do provide useful information and data about traffic flows at particular points, they generally do not provide useful data for traffic flows over space. It is not possible to move detectors; further, they cannot provide useful information such as vehicle trajectories, routing information, and paths through the network. The number of road accident increases day by day, on average, at least one person died at every minute due to some road accident. Auto accidents injure at least 10 million people every year, two or 3 million are seriously. It leads to the hospital bill, damaged property, and other costs will add up to 1-3 percent of the world's gross domestic product [1-3].

Video based vehicle detection technology is an integral part of Intelligent Transportation System (ITS), due to its non-intrusiveness and comprehensive vehicle behavior data collection capabilities. In transportation, vehicle detection system may be defined as a system which is capable of detecting vehicles and measure traffic parameters such as count, speed, incidents, etc. Also vehicle detection can be used for various transportation applications like: autonomous vehicle guidance, vehicle safety, etc. Vehicle detection by video cameras is one of the most promising non-intrusive technologies for large scale data collection and implementation of advanced traffic control and management schemes. Vehicle detection is also the basis for vehicle tracking. The correct vehicle detection results in better tracking.

Robust and reliable vehicle detection is an important issue with applications to driver assistance systems or autonomous, self-guided vehicles. Several factors make on-road vehicle detection very challenging including variability in scale, location, orientation, and pose. Vehicles, for example, come into view with different speeds and may vary in shape, size, and color. Vehicle appearance depends on its pose and is affected by nearby objects. In-class variability, occlusion, and lighting conditions also change the overall appearance of vehicles. Landscape along the road changes continuously while the lighting conditions depend on the time of the day and the weather. Moreover, real time constraints make this task even more challenging.

The conventional vehicle detection system was mainly used with either active sensors or passive sensors [2]. The most common active sensors for vehicle detection mostly include radar-based, laser-based, and acoustic-based. Regarding the advantage of the active sensors, they can detect preceding vehicles' distance correctly without many computing resources. However, the disadvantage of the active sensor is to be easily disturbed. To overcome this issue, we introduce passive sensors. As a representative method of passive sensors, the optical sensor (optical camera) is often applied to vehicle detection in recent researches. It has advantages of no interference, low cost, and so on. Meanwhile, visual information can be used in related works of ITS, such as lane detection, traffic sign recognition or object identification (e.g., pedestrian, obstacles) [2].

In this study, we proposed a two-step approach for accelerating vehicle detection; the flowchart is shown as Figure 1. For the first step, we assume vehicles' potential locations through searching some shadows under vehicle. In order to find these shadows, Haar-like feature with Adaboost was used to train a Haar Detector offline and we use it to perform shadow detection. Meanwhile, we used Hard Sample Training method eliminate most false alarms. Based on previous detected shadow areas, we can set some extended areas to perform vehicle detection in second step. Moreover, to find vehicles in defined ROI, we have trained a HOG Detector by using HOG feature with linear SVM for vehicle detection. Finally, an improved K-means approach was used to combine the similar detection results. The experimental results proved that our system can be used for real time preceding vehicle detection robustly and accurately.
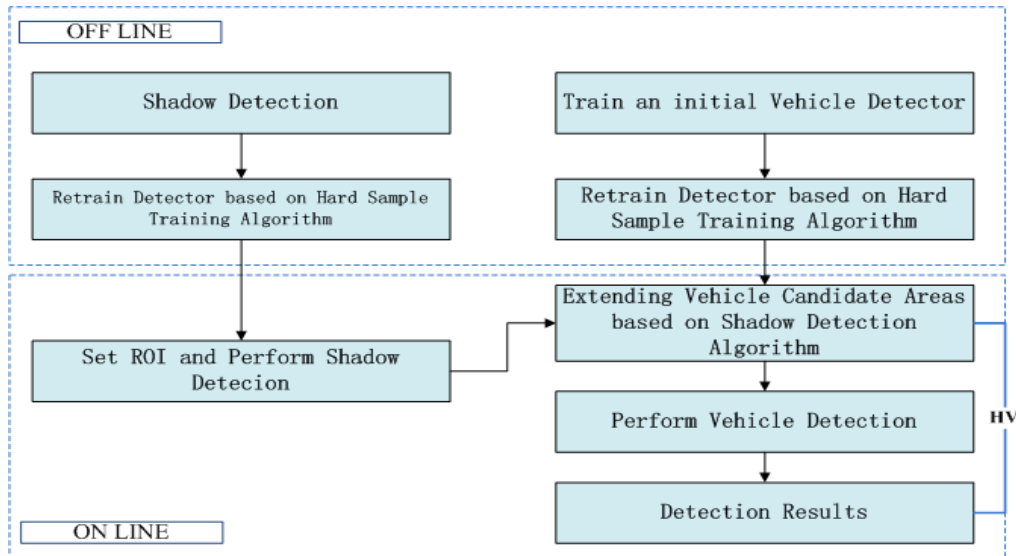

Figure 1. The Proposed Method in this Paper

2. Related Works

Until now a lot of researches have proposed vehicle detection algorithms. Most researches have been well arranged. Especially, the vehicle detection method in [4] is separated into two steps. The hypothesis generation step uses the knowledge, stereo vision and motion based methods. Knowledge-Based method uses many features (symmetry, color [5], shadow, corner, edge, and texture).

On-road vehicle detection systems have high computational requirements. Thus, they need to process the acquired images at real time or close to real time to save time for drive reaction. Generally, vehicle detection can divided into following two steps:

1) Hypothesis Generation: the locations of possible vehicles in an image are hypothesized.

2) Hypothesis Verification: the tests are performed to verify the presence of vehicles in an image.

Various HG approaches have been suggested in the literature, each of which falls into one of the following three categories: (1) knowledge-based, (2) stereo-based, and (3) motion-based. Knowledge-based methods employ knowledge about vehicle shape and color as well as general information about streets, roads, and freeways. Tzomakas et al. for example, modeled the intensity of the road and shadows under the vehicles to estimate the possible presence of vehicles [6]. Symmetry detection approaches using the intensity or edge map have also been exploited based on the observation that vehicles are symmetric about the vertical axis [7, 8].

The hypothesized locations from the HG step are the inputs of the HV step, where tests are performed to verify the correctness of the hypotheses. Approaches to HV can be classified mainly into two categories: (1) template based, and (2) appearance-based. Template-based methods use predefined patterns of the vehicle class and perform correlation between an input image and the template. Betke *et al.* [9] proposed a multiple-vehicle detection approach using deformable gray-scale template matching. In [10], a deformable model is formed from manually sampled data using Principal Component Analysis (PCA). Both the structure and pose of a vehicle can be recovered by fitting the PCA model to the image.

3. Shadow Detection Based on Haar Feature

There are many literatures, such as [11-13] which have proved that shadow between a vehicle and the road can be extracted as an important feature for vehicle detection. Meanwhile, the underneath shadow has symmetry structure since we can extract them easily. However, we didn't use shadow as feature to detect vehicle directly in our study. We only used shadow to estimate vehicles potential locations. In order to find these shadows, we used Haar detector, which was trained by Haar-like feature with Adaboost to perform shadow detection.

Haar-like feature was mainly used to describe objects' intensity difference. It has already got a significant success in face detection field. Papageorgiou *et al.* [14] proposed an alternative feature set based on Haar wavelets instead of the usual image intensities firstly. Then Viola and Jones [15] adapted the idea of using Haar wavelets and developed Haar-like features. In [16], authors made use of Haar-like features for multi-view face detection. The main advantage of Haar-like feature is that it can be calculated in short time with integral images, which is so much faster than other features. So we used Haar-like feature as our feature to perform shadow detection. Figure 2 shows us a set of Haar-like features. In our study we used edge features and center-surround features.

(a) Edge feature   (b) Edge feature   (c) Center-surround feature

Figure 2. Haar-like Features for Shadow Detection

Figure 3 is the training samples for training processing. Positive samples and negative samples are represented as 'plus' and 'minus', respectively. Vertical and horizontal lines represent the weak classifiers. In the first step, we can obtain the initial classifier based on the minimal error rate. Three positive samples with the circles cannot be classified successfully. Thus, three positive samples' weights will be increased in the next step to train the second classifier. Using the same method, we can obtain the third weak classifier based on a new distribution. In the end, combine all weak classifiers to construct a strong classifier and the strong classifier can classify all samples, successfully.
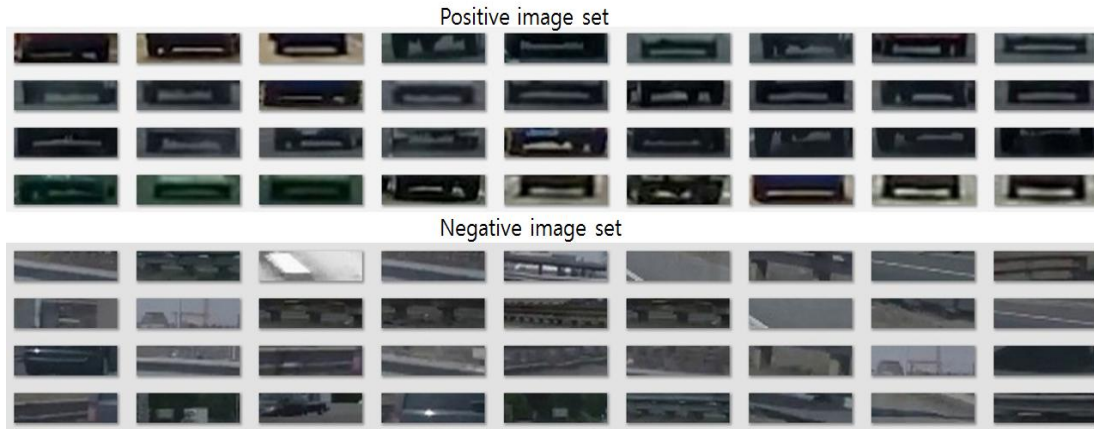

Figure 3. Positive Samples and Negative Samples for Training Processing

For scanning a high resolution image, we need much time. Generally, we consider that some areas are always without vehicles in the whole image. Thus, in order to increase the detection efficiency and speed further, we set an assigned ROI to perform scanning. Through observations, we find that the vanishing point above can be regarded as Non-vehicle area. Moreover, the vanishing point always exists in the middle of the image, and the windshield edges, and wipers always exist at the bottom of the image. Based on what was mentioned above, we can set an ROI as (0,540) to (1920, 980). The details are shown in Figure 4.

To reduce the false alarm rate and increase detector's detection performance, we used Hard Sample Training method to retrain the training dataset two times (See Figure 5). The principle of Hard Sample Training method is to use the initial Haar detector for performing detection with initial positive training set and negative training set to extract false detection parts, which were put into the initial positive and negative training set, respectively. Then update the training set for training a new Haar detector. We can obtain the final Haar detector by repeating this processing two times.

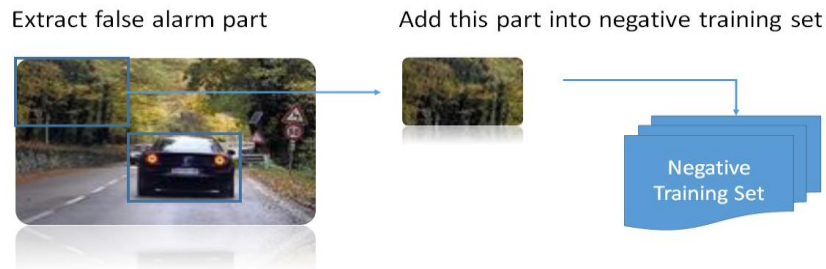**Figure 4. Region of Interest for Shadow Detection**

Figure 5. Hard Sample Training Method for

4. Vehicle Verification Using HOG and SVM Algorithm

Histogram of oriented gradient (HOG) features 17] has been used in a number of studies [18-19]. In [20], the symmetry of the HOG features extracted in a given image patch, along with the HOG features themselves, was used for vehicle detection. HOG features are descriptive image features, allowing for determination of vehicle pose [21]. The main drawback of HOG features is that they are quite slow to compute. Recent work has tackled the speed bottleneck by implementing HOG feature extraction on a graphical processing unit (GPU) [22]. And classification methods for appearance-based vehicle detection have followed the general trends in the computer vision and machine learning literature. In this section, we make use of HOG features and SVM algorithm for vehicle verification in potential position of the vehicle based on shadow detection.

4.1. Vehicle Verification in Region of Interest (ROI)

Before vehicle detection, we need do some extend works to find the vehicle potential location further based on previous shadow detection work. Generally, the vehicles' size in the image can be affected by the locations. For a car that is located far away, its area will be small. On the other hand, the area of car will be obvious when the car is within short distance. What is more, the distance can be transformed into 2D plane. It can be represented as the vanishing point in the middle of the image and the windshield bottom edge or wipers at the bottom of the image. So we divided the whole image into three parts for setting vehicle candidate area based on previous shadow detections. We defined ROI from (0, 540) to (1980, 690) of the area A. In this area, the width of the candidate area can be defined the same as

shadow's width. Further, we used two restricted criteria to help us filtering some non-candidate areas, because of the minimum searching window that we have already set as 64*64; so, if the shadow's width was narrower than 64, then this area was filtered out directly. Meanwhile, the shadows' width also cannot be too wide in this area, so we set threshold value as 100. In this area, the width of detected shadow must be required from 64 to 100. For the height, we need not pay too much attention on it because of previous HOG detector has already set a fixed ratio relationship (1:1) between height and width. We only need to set original coordinate of the vehicle candidate at the left-bottom of detected shadow rectangle. For area B, we defined this area as from (0, 690) to (1920, 840). The vehicle candidate area should be a little greater than area A. In this area, the width can be limited from 64 to 200. And the height can be set as the same as width. For area C, we defined this area as from (0, 840) to (1920, 980). Because this part is the closet one to us, this part can be set as great as possible. The width was limited as 100 to 200. We can see the detailed procedure as Figure 6. After we set all the three parts, we used previously trained HOG detector to detect vehicles in respective candidate areas.
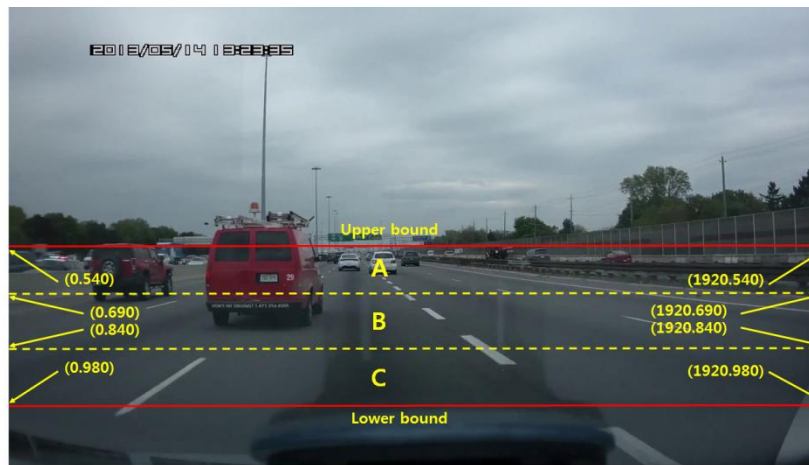

Figure 6. Defined Extending Areas of Vehicle Detection

4.2. Vehicle Features Extraction based on HOG Algorithm

It was constructed through calculating image's local gradient orientation and similarity of edge orientation histograms, scale-invariant feature transform descriptors and shape contexts. As a feature descriptor, HOG feature is used to describe object's shape better. Meanwhile, it has better invariance under different illumination and shadowing conditions because its calculation is performed in local cells. Due to the advantages mentioned above, the HOG feature has been often applied in pedestrian detection field. Generally, the HOG feature calculation can be divided into following steps:

1) Gradient computation

Apply the 1-D centered, point discrete derivative mask ($[-1,0,1]$ and $[-1,0,1]^{-1}$ are used in this study) in one or both of the horizontal and vertical directions to obtain the gradient orientation and gradient magnitude. We can use the Eq. 1 and Eq.2 to calculate each pixel point's horizontal and vertical gradient value, respectively. And use the Eq. 3 and Eq. 4 to calculate pixel point's gradient magnitude value and gradient value, respectively.

$$G_x(x, y) = H(x + 1, y) - H(x - 1, y) \tag{1}$$

$$G_y(x, y) = H(x, y + 1) - H(x, y - 1) \tag{2}$$

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

$$\theta(x, y) = \arg\tan \left. G_y(x, y) \middle/ G_x(x, y) \right. \tag{4}$$

2) Orientation binning

Each pixel within the cell casts a weighted vote for an orientation-based histogram channel based on the values found in the gradient computation. The histogram channels are evenly spread over 0 to 180 degrees or 0 to 360 degrees, depending on whether the gradient is 'unsigned' or 'signed'. In our study, we use the 'unsigned' and nine channels that evenly spread over 0 to 180 degrees to construct our histogram. As for the vote weight, pixel contribution can either be the gradient magnitude itself, or some function of the magnitude. We just use the gradient magnitude value as vote weight in our study.

3) Descriptor blocks

In order to avoid for changes in illumination and contrast, the gradient strengths must be locally normalized, which requires grouping the cells together into larger, spatially connected blocks. The HOG descriptor is then the vector of the components of the normalized cell histograms from all of the block regions. These blocks typically overlap, meaning that each cell contributes more than once to the final descriptor. Every four cells (2*2) comprise one block in our study.

For an image with size of 64*64, we assume that each cell's size is 8*8, and the size of one block is 16*16 since one block comprises four cells. Thus, we can consider that there are 49 blocks in an image since one block can be slid seven times in horizontal and vertical orientation, respectively. Meanwhile, there are nine channels in one cell and 36 features in one block. Thus, we can obtain the number of features is 1764 (36*49) in a 64*64 image that as shown as Figure 7.
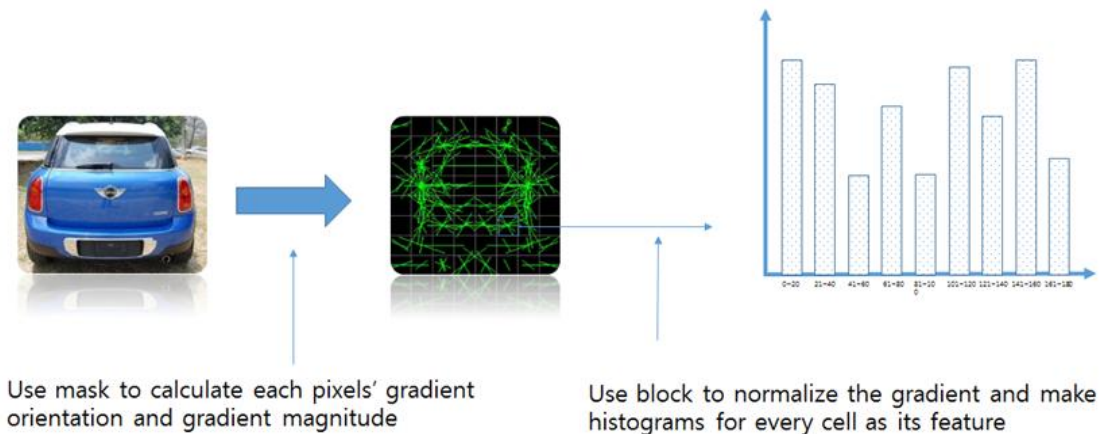

Use mask to calculate each pixels' gradient orientation and gradient magnitude

Use block to normalize the gradient and make histograms for every cell as its feature

Figure 7. HOG Features Extraction

4.3. Learning Edge Direction Based on SVM

SVM are primarily two-class classifiers that have been shown to be an attractive and more systematic approach to learning linear or non-linear decision boundaries. Given a set of points, which belong to either of two classes, SVM finds the hyper-plane leaving the largest possible fraction of points of the same class on the same side, while maximizing the distance of either class from the hyper-plane. This is equivalent to performing structural risk minimization to achieve good generalization.

Suitable kernel functions can be expressed as a dot product in some space and satisfy the Mercer's condition. By using different kernels, SVM implement a variety of learning machines (e.g., a sigmoidal kernel corresponding to a two-layer sigmoidal neural network while a Gaussian kernel corresponding to a radial basis function (RBF) neural network). In this study, we select Gaussian RBF kernel function for SVM classifier for HOG features classification.

The Gaussian radial basis kernel is given by

$$k(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \tag{5}$$

In the hypothesis verification phase, a HOG detector was used to perform vehicle detection in candidate areas. For setting parameters of the HOG detector, the size of searching window was set as same as the size of training data, which is 64*64. Also the cell size and block size were the same as training size. To avoid missing detections, we set the slide window size as 8*8 and the searching window enlargement coefficient as 1.02.

In order to get a better detection result, we used massive training set to train our classifier. For positive image set, we used 1000 different kinds of images that all images include cars and 2000 images without cars for negative image sets (see Figure 8). For calculating HOG features effectively, the histogram channels had been divided into 9 channels from 0 to 180 degree. Four cells constitute a block and each cell's size is 8*8. Through observing some vehicles' size, we set searching window's size as 64*64. So we can calculate the numbers of feature in one searching window is 1764 (4*9*7*7). Then we selected calculated results of HOG with linear SVM to train an initial HOG detector. For optimizing the detection result further, we proposed a method of merging similar detection results based on K-means algorithm to eliminate some similar detection results.
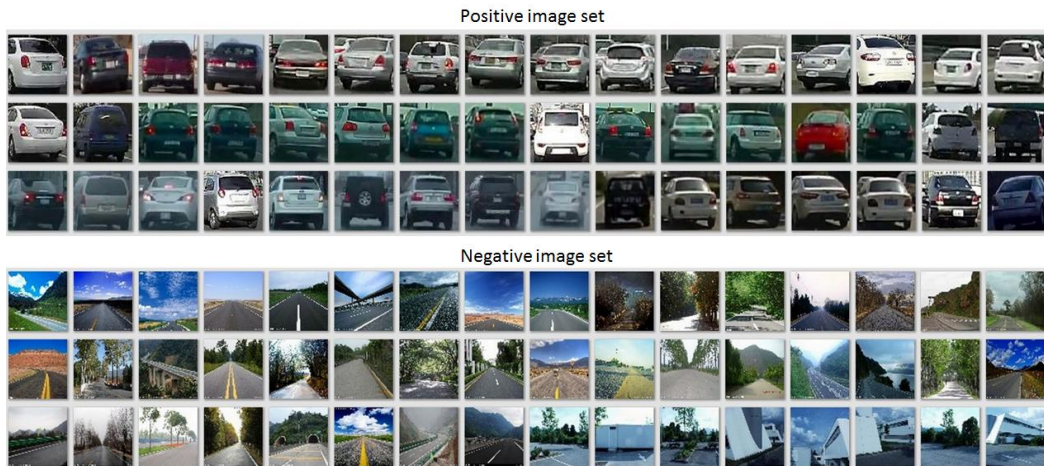

Figure 8. Training Samples for SVM Classifier

5. Experimental Results

The vehicle detection system we designed is running on the hardware environment of Intel (R) Core (TM) i3 (3.06 GHz), an optical sensor and the software environment of Windows 7, Visual Studio 2010 using OpenCV Library.

Three videos are taken by optical camera under different weather, illuminations and environments for testing the performance of our vehicle detection system.

In order to train shadow detector, we selected an efficient implement based on the work of Viola and Jones who named it as "Haar-training" that OPENCV library has been supplied. For positive image selection, we clipped 600 underneath shadow images. Meanwhile, 1,200 images were clipped as our negative images that are without any shadows. All of the images were clipped manually. Also, Seo [23] proposed a method to generate a massive sample that uses one sample to generate five samples. In order to get a high efficiency classifier, we have set the number of cascade stage as 21, the minimum hit rate set as 0.995, the maximum false alarm rate set as 0.5, and the samples' width and height were set as 60 and 15, respectively. By using proposed shadow detector, the shadows of vehicle could be detected successfully under different conditions, the detection results are shown as Figure 9.

Figure 9. Shadow Detection under Different Conditions

And then, vehicle verification processing based on ROI+HOG+SVM will be implemented. In order to optimize the detection result further, we proposed a method of merging similar detection results based on K-means algorithm to eliminate some similar detection results. For a size of 64*64 image, we assume that each cell's size is 8*8, and the size of one block is 16*16 since one block comprises four cells. Thus, we can consider that there are 49 blocks in an image since one block can be slid seven times in horizontal and vertical orientation,

respectively. For SVM algorithm, we using Gaussian RBF as kernel function. The comparison vehicle verification results are shown as Figure 10 and Figure 11.

Although Conventional feature-based vehicle detection methods have better detection results, their drawback is much more time calculation, even though we used GPU to accelerate its processing speed. Here, we have contrasted the detection accuracy and time between original method and our method in this paper, respectively (Figure 12 and Figure 13).

The experimental results show that our method not only has better detection accuracy rate, but also faster than conventional methods without making particular efforts to optimize our programming. The proposed method is robust and satisfactory for real time preceding vehicle detection

Figure 10. Comparison of Vehicle Detection results under fine day. a) Detection result using original HOG+SVM , b) Detection result using proposed method
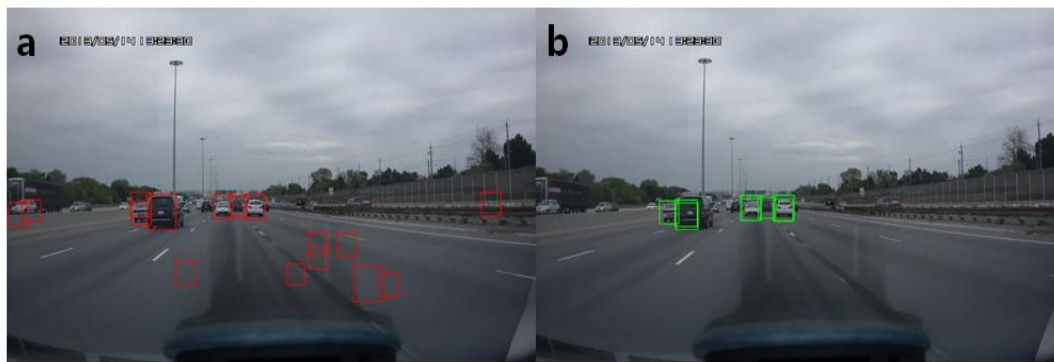

Figure 11. Comparison of Vehicle Detection Results Under Cloudy Day, a) Detection result using original HOG+SVM, b) Detection result using proposed method
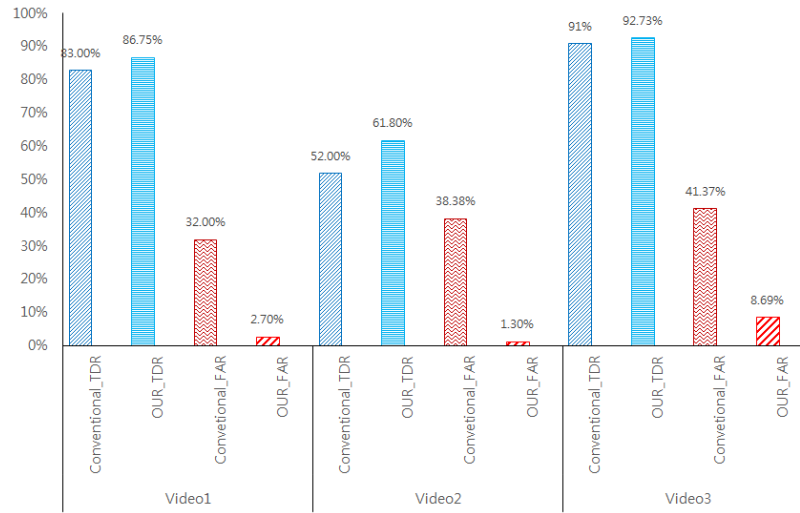
Figure 12. Comparison of Vehicle Detection Rate and False Alarm Rate with Conventional Method Based on Three Test Video (TDR=True Detection Rate, FAR=False Alarm Rate)
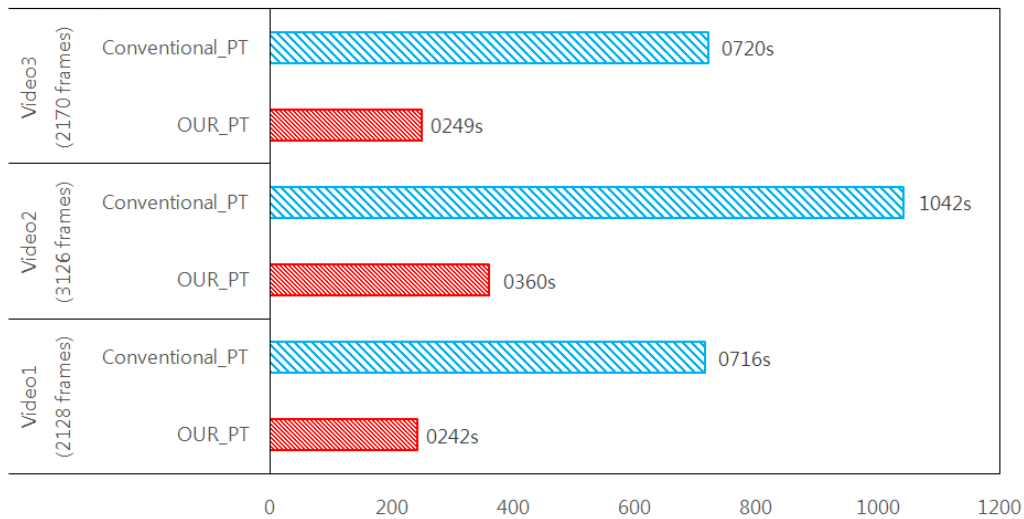

Figure 13. Comparison of Time Computation for Tree Test Video (PT=Processing Time)

6. Conclusion

In our study, we proposed a two-step method to perform preceding vehicle detection. Based on shadow's horizontal and vertical structure, we can extract shadows easily from the visual information. To detect shadows, we trained an initial Haar detector by using Haar-like feature with Adaboost algorithm. In order to increase detection accuracy further, we used Hard Sample Training method to retrain the initial detector two times to obtain the final Haar detector. And then, final Haar detector was used to detect shadows by scanning images. After that we utilized the final shadow detection results to extend vehicle's probable areas. At last, well-known HOG feature and Gaussian RBF SVM algorithm were used to train a HOG detector for detecting vehicles

in candidate areas. Also to increase detection accuracy, we used Hard Sample Training method to retrain the HOG detector two times to obtain the final HOG detector. The experimental results expressed that our method is robust for real time preceding vehicle detection. In future works, we will focus on calculating the distance from preceding vehicles and other traffic sign detections to construct a satisfactory Intelligent Transportation System.

Acknowledgements

References

[1] W. Jones, "Keeping Cars from Crashing", IEEE Spectrum, vol. 38, no. 9, **(2001)**, pp. 40-45.
[2] Z. Sun, G. Bebis and R. Miller, "On-road vehicle detection: a review", Pattern Analysis and Machine Intelligence, vol. 28, no. 5, **(2006)**, pp. 694-711.
[3] W. Jones, "Building Safer Cars", IEEE Spectrum, vol. 39, no. 1, **(2002)**, pp. 82-85.
[4] R. Miller, Z. Sun and G. Bebis, "On Road vehicle Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, **(2004)**, pp. 694-711.
[5] J. Rojas and J. Crisman, "Vehicle Detection in Color Images", Proceeding of IEEE Conference Intelligent Transportation, **(1997)** November 01-12, Boston, MA.
[6] C. Tzomakas and W. V. Seelen, "Vehicle detection in traffic scenes using shadows", Proceeding of Internal Report 98-06, Institute fur Neuroinformatik **(1998)**
[7] A. Kuehnle, "Symmetry-based recognition for vehicle rears", Pattern Recognition Letters, vol. 12, **(1991)**, pp. 249-258.
[8] T. Zielke, M. Brauckmann and W. V. Seelen, "Intensity and edge-based symmetry detection with an application to car following", Image Understanding, vol. 58, no. 2, **(1993)**, pp. 177-190.
[9] M. Betke, E. Haritaglu and L. Davis, "Multiple Vehicle Detection and Tracking in Hard Real Time", Proceeding of IEEE Intelligent Vehicles Symposium, **(1996)** September 19-20, pp. 351-356, Tokyo, Japan.
[10] J. Ferryman, A. Worrall, G. Sullivan and K. Baker, "A generic deformable model for vehicle recognition", Proceedings of British Machine Vision Conference, **(1995)** September 11-14, Birmingham, UK.
[11] S. Kim, S.-y. Oh, JK. Kang and YW. Ryu, "Front and rear vehicle detection and tracking in the day and night times using vision and sonar sensor fusion", Proceeding of Intelligent Robots and Systems, **(2005)** August 2-6, pp. 2173-2178, Alberta, Canada.
[12] C. Hoffman, T. Dang and C. Stiller, "Vehicle detection fusing 2D visual features", Proceeding of Intelligent Vehicles Symposium, **(2004)** June 14-17, pp. 280-285.
[13] B. Aytekin, V. Isra and E. Altug, "Increasing driving safety with a multiple vehicle detection and tracking system using ongoing vehicle shadow information", Proceeding of Systems Man and Cybernetics (**2010**), pp. 3650-3656, Istanbul, Turkey.
[14] C. P. Papageorgiou, M. Oren and T. Poggio, "A general framework for object detection", Proceeding of 6th Computer Vision, **(1998)** January 04-07, pp. 555-562, Bombay, India.
[15] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", Proceeding of Computer Vision and Pattern Recognition, **(2001)** December 8-14, pp. 511-518, Kauai, USA.
[16] W. Xu and E. J. Lee, "A Novel Multi-view Face Detection Method Based on Improved Real Adaoost Algorithm", Transactions on Internet and Information Systems, vol. 7, no. 11, **(2013)**, pp. 2720-2736.
[17] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection", Proceeding of Computer Vision and Pattern Recognition, **(2005)** June 25, pp. 886-893, San Diego, USA.
[18] P. Geismann, and G. Schneider, "A two-staged approach to vision-based pedestrian recognition using Haar and HOG features", Proceeding of Intelligent Vehicles Symposium, **(2008)** June 4-6, pp. 554-559, Eindhoven, Netherlands.
[19] J. Shotton, A. Blake and R. Cipolla, "Contour-based learning for object detection", Proceeding of Computer Vision, **(2005)** October 17-21, pp. 503-510, Beijing, China.
[20] D. Santos, and P. L. Correia, "Car recognition based on back lights and rear view features", Proceeding of Image Analysis for Multimedia Interactive Services, **(2009)** May 06-08, pp. 137-140, London, UK.
[21] B. Alefs, "Embedded Vehicle Detection by Boosting", Proceeding of Intelligent Transportation Systems Conference, **(2006)** September 17-20, pp. 536-541, Toronto, Canada.

[22] L. Mao, M. Xie, Y. Huang and Y. Zhang, "Preceding vehicle detection using Histograms of Oriented Gradients", Communications, Circuits and Systems, **(2010)** July 28-30, pp. 354-358, Chengdu, China.

[23] N. Seo, "Tutorial: OpenCV Haartraining (Rapid Object Detection with a Cascade of Boosted Classifiers Based on Haar-like Features)", (Online) (**2008**) http://note.sonots.com/SciSoftware/haartraining.html

Authors
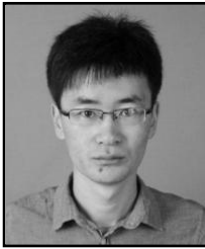
Huan Wang, he received his B. S. in Dalian University of Technology in 2010. Currently, he is the graduate student of Dalian Polytechnic University，Dalian，China. His main research fields are the intelligent controls and process control.

Haichuan Zhang, he received his B. S. and Sc. D. in Dalian University of technology in 2004 and 2010; currently, he is a lecture of School of Information Science and Engineering, Dalian Polytechnic University, Dalian, China. His main research fields are the intelligent controls and mechanical and electrical integration.