Journal of Southeast University (English Edition) Vol. 31, No. 1, pp. 67 - 73 Mar. 2015 ISSN 1003—7985

On-road vehicle verification based on VS-HOG and ELM

Fan Yanjun[1,2] Zhang Lei[1] Zhang Weigong[1]

([1]School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China)

([2]Department of Computer Science and Technology, China Jiliang University, Hangzhou 310018, China)

Abstract: A solution is proposed for the real-time vehicle verification which is an important problem for numerous on-road vehicle applications. First, based on the vertical symmetry characteristics of vehicle images, a vertical symmetrical histograms of oriented gradients (VS-HOG) descriptor is proposed for extracting the image features. In the classification stage, an extreme learning machine (ELM) is used to improve the real-time performance. Experimental data demonstrate that, compared with other classical methods, the vehicle verification algorithm based on VS-HOG and ELM achieves a better trade-off between cost and performance. The computational cost is reduced by using the algorithm, while keeping the performance loss as low as possible. Furthermore, experimental results further show that the proposed vehicle verification method is suitable for on-road vehicle applications due to its better performance both in efficiency and accuracy.

Key words: histogram of oriented gradients (HOG); vertical symmetrical histogram of oriented gradients (VS-HOG); vehicle verification; extreme learning machine (ELM)

doi: 10.3969/j.issn.1003-7985.2015.01.012

Vision-based vehicle detection systems play important roles in many applications, such as self-guided vehicles, driver-assistance systems, and automatic parking systems [1-3]. Vision-based vehicle detection systems have two main challenges: accuracy and real-time performance. To obtain high accuracy and achieve real-time processing, a typical vehicle detection system often includes two parts: 1) Hypothesis generation (HG) and 2) Hypothesis verification (HV) [4]. In the HG step, the locations of the vehicles in images are hypothesized. In the HV step, tests are performed to verify the presence of vehicles in an image.

In recent years, more and more researchers have used classification techniques in the vehicle HV step and have achieved impressive results[5-9]. In their studies, HV is regarded as a classification problem of distinguishing between vehicle and non-vehicle data, and then feature descriptors (including vehicle and non-vehicle data) are extracted from training images to generate classifier parameters. Therefore, good performance feature descriptors play very important roles in the effectiveness of the classifier. Popular methods to extract feature descriptors include Haar-like[10], Gabor filters[11-12], and histograms of oriented gradients (HOGs)[13]. HOGs are the descriptive image feature, exhibiting good detection performance in a variety of computer vision tasks, including vehicle detection[14-18], but they are generally slow to compute. In Ref. [18], by exploiting the a priori known vehicle appearance, the authors proposed three less-demanding HOG descriptors, CR-HOG, V-HOG and H-HOG, in order to lighten the computation burden. During the computation of all the three improved HOG descriptors, the number of cells is much smaller than that in the traditional HOG, thus resulting in fewer calculations and a much smaller feature vector. Furthermore, the authors in Ref. [18] pointed out that V-HOG is shown to be the most efficient among these descriptors due to its better capability for describing the intrinsic gradient content of a vehicle image. However, when extracting a V-HOG descriptor from vehicle images, the symmetry features that generally exist in vehicle images are neglected.

In this paper, using the symmetrical characteristics of vehicles, an alternative HOG descriptor called vertical symmetrical histogram of oriented gradients (VS-HOG) is presented to achieve higher accuracy and efficiency. In order to enhance classification speed, an extreme learning machine (ELM) is used as a two-class classifier to recognize vehicles. Experiments are performed on a large public database and verified the proposed method of the performance in the speed and accuracy of vehicle verification.

1 VS-HOG Vehicle Descriptor

1.1 HOG and V-HOG descriptor

HOG was proposed by Dalal and Triggs[13] for pedestrian detection, and it has also been extended to other applications such as vehicle detection. The calculation of HOG descriptors include three steps: 1) Gradient computation;

Received 2014-08-17.

Biographies: Fan Yanjun (1977—), male, graduate; Zhang Weigong (corresponding author), male, doctor, professor, zhangwg@seu.edu.cn.

Foundation items: The National Natural Science Foundation of China (No. 61203237), the Natural Science Foundation of Zhejiang Province (No. LQ12F03016), the China Postdoctoral Science Foundation (No. 2011M500836).

Citation: Fan Yanjun, Zhang Lei, Zhang Weigong. On-road vehicle verification based on VS-HOG and ELM[J]. Journal of Southeast University (English Edition), 2015, 31(1): 67 - 73. [doi: 10.3969/j.issn.1003-7985.2015.01.012]

2) Orientation binning; and 3) Histogram generation. First, all the gradients of an image are obtained by applying Sobel or Prewitt operators. Then, the image is divided into cells according to predefined size in order to accumulate the histogram of orientations. Finally, the range of each bin should be determined. For example, if the unsigned gradient is divided into eight bins with the same range, the range of each bin is $\pi/8$ (see Fig. 1(b)). In the histogram generation step, votes are weighted according to the gradient magnitude in each cell. Then, the cells are grouped into larger spatial structures, called blocks (see Fig. 1(a)), and normalization processing is carried out locally in each block according to standard measures, such as the L1 or the L2 norm. The final descriptor comprises the normalized responses from all the blocks.

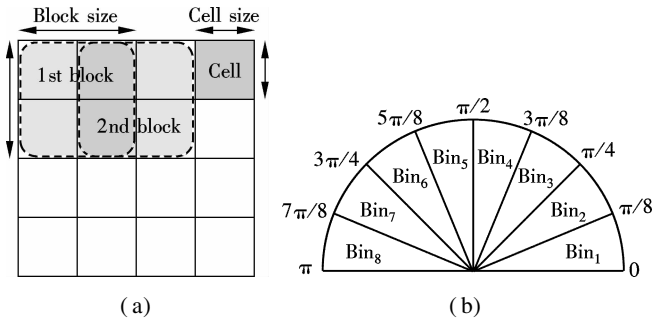


Fig. 1 Illustration of a typical HOG configuration. (a) Example of HOG grid. (b) Orientation range of each bin for 8-D HOG

HOG descriptors can provide good classification results, but feature extraction and classification require heavy computation, which greatly decreases the algorithm efficiency. When using HOG in vehicle detection, the critical real-time requirements should be satisfied for practical applications. In Ref. [18], the authors proposed three cost-effective HOG descriptors, named CR-HOG, V-HOG and H-HOG. Furthermore, the results of experiments, which were performed on the GTI vehicle database[19], shows that the V-HOG descriptor is a better alternative than the two others for real-time vehicle verification.

The V-HOG descriptor (see Fig. 2) divides an image window into several vertical stripes. Therefore, the method needs only two parameters, named the number of orientation bins in the histogram, β, and the number of cells, η. Then, the dimension of the feature vector extracted from vehicle images is $\beta \times \eta$. Moreover, these cost-effective descriptors in cells are not further grouped into blocks, and local contrast normalization is unnecessary.

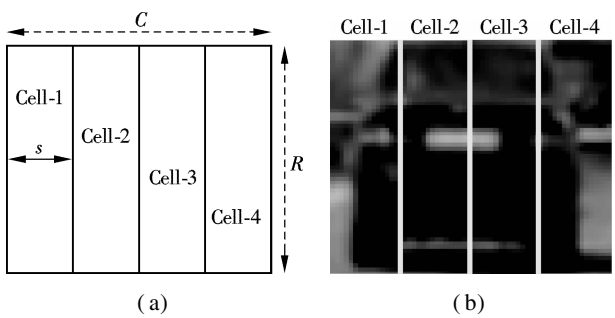


Fig. 2 The V-HOG descriptors. (a) The V-HOG descriptor with four vertical cells; (b) A vehicle image divided by four-cell V-HOG

1.2 VS-HOG descriptor

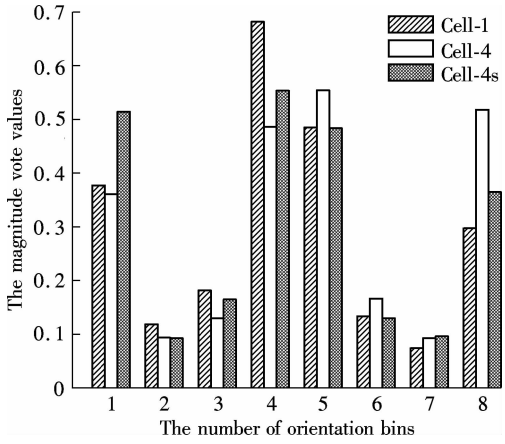
The V-HOG descriptor presented in Ref. [18] makes use of the previous knowledge of vehicle appearance. Based on the previous knowledge of the vehicles, the image window is divided into several vertical stripes instead of classical HOG grids. However, symmetry, which is a very important characteristic of the vehicle images, is still underexploited in the V-HOG method.

Symmetry is one of the most important visual characteristics of vehicles. The rear views of most vehicles are symmetric over a vertical centerline. The symmetry feature is often used to extract the symmetric regions in the image[20]. However, in this paper, symmetry is exploited to enhance the description capability of V-HOG.

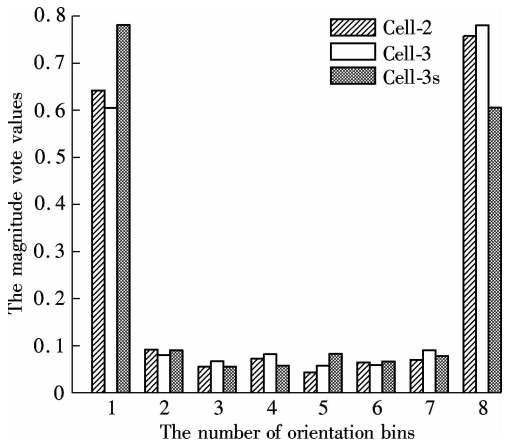
1.2.1 Cells-symmetry and bins-symmetry

The symmetry characteristics used in this paper to improve V-HOG descriptors include two categories. One is cells-symmetry, and the other is bins-symmetry. As we can see in Fig. 2, the V-HOG divides the image into vertical stripes and the dimension of V-HOG is determined by the number of cells and the number of orientation bins in the histogram. When computing the V-HOG of vehicle images, symmetry obviously exists between cell-1 and cell-4, which is denoted cells-symmetry. On the other hand, in the V-HOG computing results of vehicle images, the symmetry exists between the orientation bins in the one cell and those in the other corresponding symmetrical cell, namely the bins-symmetry. In fact, when computing the vehicle image gradients, many pixel gradients with opposite orientation and approximate equivalent magnitude value can be obtained in the mutually symmetric cells because of the symmetry of the vehicles. Fig. 1 (b) shows the symmetry among the orientation bins corresponding to each other, bin_1 is symmetric with bin_8, and bin_2 is symmetric with bin_7, and so on. Thus, these pixel gradients are used to vote according to the gradient orientation in each cell. There must be some kinds of bins-symmetry rules between the mutually symmetric cells of the vehicle image.

In Fig. 3(a), the bars with a diagonal stripe background show the vote values of eight bins in cell-1, and the bars with a white background denote those in cell-4. The bin vote values of cell-4 are interchanged according to the bin-bars interchange method based on the bins-symmetry rules (see Fig. 4), and the results are shown with the diagonal cross background bars, namely cell-4s. It can be observed that the form of the bars in cell-4s is more similar to that in cell-1 than in cell-4s. These facts prove that the bins-symmetry actually exists between the



(a)



(b)

Fig. 3 The accumulated vote results of bins in each cell shown in. (a) The comparison of bin vote values between cell-1 and cell-4; (b) The comparison of bin vote values between cell-2 and cell-3

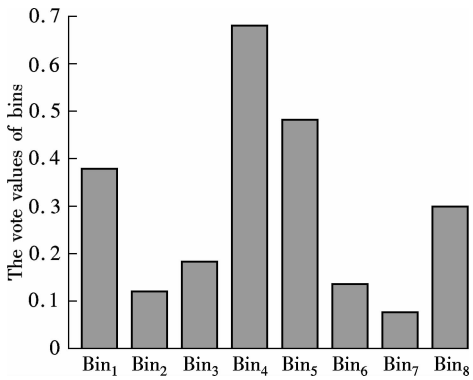
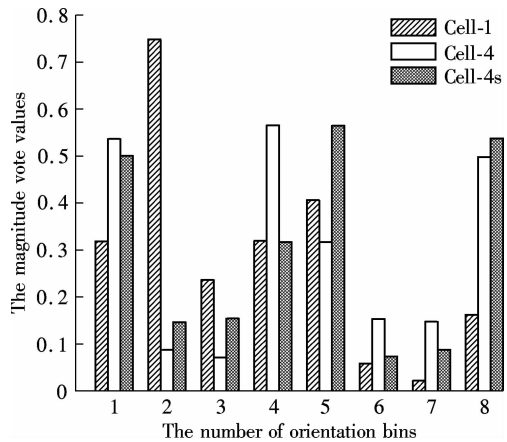


Fig. 4 The bin-bars interchange method based on the bins-symmetry rules

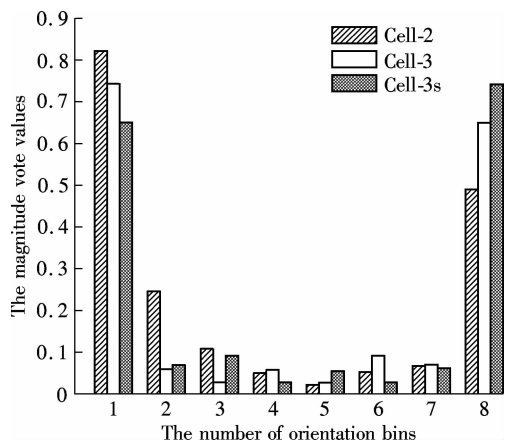
mutually symmetric cells. However, we can also find that the vote value of bin_1 or bin_8 in cell-4 is more close to that in cell-1 than in cell-4s. The main reason of this phenomenon is that a large number of pixel gradients with orientation nearby 0 or π are obtained in the gradient computation stage, and most of these pixels, such as background pixels, are not the discriminative features of the vehicle images. Therefore, it has little effect on the description results when the interchange between bin_1 and bin_8 is executed. On the other hand, the bins-symmetry rules can be clearly discovered from bin_2 to bin_7 in Fig. 3 (a). The similar phenomena discussed above also exist in cell-2 and cell-3, which are shown in Fig. 3(b).

In order to verify whether the bins-symmetry rules also exist in the non-vehicle images, a typical non-vehicle image and its bin-bars interchange results of each cell are shown in Fig. 5. As we can see from Fig. 5, the bins-symmetry rules can seldom be found both in Fig. 5(b) and Fig. 5(c). Therefore, we can use the bins-symmetry rules to make a distinction between vehicles and non-vehicles.

(a)



(b)



(c)

Fig. 5 A non-vehicle image and its bin-bars interchange results. (a) The non-vehicle image; (b) The comparison of bin vote values between cell-1 and cell-4; (c) The comparison of bin vote values between cell-2 and cell-3

1.2.2 VS-HOG computation

The image is divided into η cells, and the number of orientation bins is β. In this paper, we assume that both η and β are even numbers and the range of orientation is $[0, \pi]$. The computation process of VS-HOG is shown as below:

1) The gradient of an image is obtained.

2) According to the histogram of orientation, vote values are counted in each cell. Fig. 6(a) shows the computation result.

3) Interchange operations are executed according to the bins-symmetry rules from cell-($\eta/2$) to cell-η. Then, we obtain cell-($\eta/2$)s to cell-ηs as shown in Fig. 6(b).

4) The VS-HOG vector is organized in the sequence as shown in Fig. 6(b).

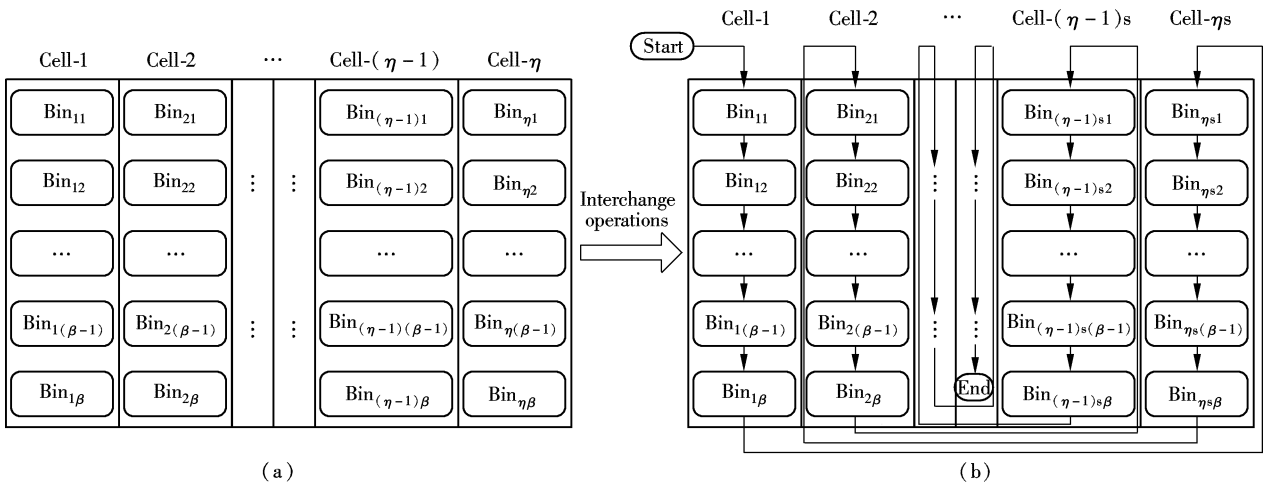


Fig. 6 The computation process of VS-HOG. (a) The computation results of bin voting in each cell; (b) The generated results of the VS-HOG descriptors

2 ELM

ELM is a single hidden layer forward network (SLFN). It has many good features, such as a fast learning speed, good generalization performance and automatically tuning hidden layer parameters[21].

For N arbitrary distinct samples $(\boldsymbol{x}_i, \boldsymbol{t}_i) \in (\mathbf{R}^d \times \mathbf{R}^m)$, $\boldsymbol{x}_i$ is the extracted feature vector and $\boldsymbol{t}_i$ is the target output label. The mathematical model of ELM with L hidden nodes is

$$\sum_{i=1}^{L} \boldsymbol{\beta}_i g_i(\boldsymbol{x}_j) = \sum_{i=1}^{L} \boldsymbol{\beta}_i G(a_i, b_i, \boldsymbol{x}_j) = \hat{\boldsymbol{t}}_j \qquad j = 1, 2, \ldots, N$$

If $N = L$, ELM can approximate the targets of the distinct N samples with zero error:

$$\sum_{i=1}^{N} \| \hat{\boldsymbol{t}}_j - \boldsymbol{t}_j \| = 0$$

That is, there are some sets of values $\boldsymbol{\beta}_i$, a_i and b_i, such that

$$\sum_{i=1}^{L} \boldsymbol{\beta}_i G(a_i, b_i, \boldsymbol{x}_j) = \boldsymbol{t}_j \qquad j = 1, 2, \ldots, N$$

which is equivalent to $\boldsymbol{H\beta} = \boldsymbol{T}$, where

$$\boldsymbol{H} = \begin{bmatrix} G(a_1, b_1, \boldsymbol{x}_1) & \cdots & G(a_L, b_L, \boldsymbol{x}_1) \\ \vdots & & \vdots \\ G(a_1, b_1, \boldsymbol{x}_N) & \cdots & G(a_L, b_L, \boldsymbol{x}_N) \end{bmatrix}_{N \times L}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{\beta}_L^{\mathrm{T}} \end{bmatrix}_{L \times m}, \quad \boldsymbol{T} = \begin{bmatrix} \boldsymbol{t}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{t}_N^{\mathrm{T}} \end{bmatrix}_{N \times m}$$

As Huang et al.[22] proved, the parameters of the hidden layer, $\{a_i, b_i\}_{i=1}^{L}$, can be randomly generated. There is $L \leqslant N$, making training error as small as possible with probability one. The training process of ELM is equivalent to solving a least squares problem.

3 Experiments

3.1 Experimental conditions

In all the experiments, the algorithm was tested on a standard laptop (Inter Core i5-3317U CPU @1.70 GHz, 4GB RAM). All the feature extraction algorithms are implemented in Matlab 7.0. Furthermore, the C-coded LIBSVM[23] package and Matlab coded ELM are used as two-class classifiers in comparison experiments.

3.2 Data set

All the experiments are carried out on the GTI vehicle database[20]. It is comprised of 3 425 images of vehicle rears taken from different points of view, and 3 900 images extracted from road sequences not containing vehicles. All the image sizes are 64×64 pixels. The images are extracted from videos acquired under different weather and illumination conditions, and the vehicles have different colors, shapes and sizes in the images. Furthermore, according to the relative distance, the images in the database are classified into four regions: front, left and right regions in the close/middle range, and far range.

3.3 Performance

In the comparison experiments, first, HOG, V-HOG and VS-HOG vectors are extracted from the images as the feature descriptors. Then, the support vector machine (SVM) and the extreme learning machine (ELM) are used as two-class classifiers to recognize vehicles. The 50% cross-validation tests, in which the training set is generated by randomly selecting half of the samples and the other half composes the testing set, are repeated 10 times for each iteration. The performance is evaluated in terms of accuracy.

HOG involves several design parameters, such as the cell size, the block size and the number of orientation bins. According to the typical configuration, the block contains 2×2 cells, and the cell is always square with size $s \times s$. Regarding the number of orientation bins, two typical values that $\beta = 8$ and 12 have been considered. As for the cell size, the experimental parameter s is set to be 4, 8, 16, and 32. For simplicity and speediness, block overlapping is ignored in our tests. Based on the HOG features and SVM classification, the results for different combinations of s and β are summarized in Tab. 1 for each of the four image regions.

Tab. 1 Accuracy results of HOG + SVM for different combinations of s and β %

Region	$\beta = 8$				$\beta = 12$			
	$s = 4$	$s = 8$	$s = 16$	$s = 32$	$s = 4$	$s = 8$	$s = 16$	$s = 32$
Front	98.70	98.07	96.99	93.73	98.64	99.10	98.48	95.01
Left	98.75	98.81	98.29	95.33	98.75	98.75	98.80	96.72
Right	96.85	96.63	95.96	93.57	97.15	97.23	97.01	95.66
Far	97.61	97.75	96.91	87.46	98.07	97.69	97.86	90.45

As shown in Tab. 1, accuracy is fairly good and stable when the cell size increases from 4 to 16, and the performance decreases abruptly when $s = 32$. In contrast, the number of orientations contributes little to improving the accuracy performance. The best accuracy results for each region are highlighted in bold in Tab. 1. It can be concluded that the best results for all four regions are obtained seemingly with $(s, \beta) = (8, 12)$.

The V-HOG method involves two design parameters, namely the number of orientation bins in the histogram, β, and the number of cells, η. In this case, the number of cells takes values $\eta = 2$, 4, 8 and 16, and the number of orientation bins is set to be $\beta = 8$, 12, 16 and 24. In this test, ELM is used as a two-class classifier in order to compare the speed with the SVM method. The results for different combinations of η and β are shown in Tab. 2.

As can be seen in Tab. 2, for the front close/middle range, the best performance (97.96%) is obtained when $\beta = 16$ and $\eta = 8$, and the best accuracy results for all four regions are highlighted in bold in Tab. 2. It can be found that the better performance can be achieved for $\eta = 4$ or $\eta = 8$ than other cases, regardless of the number of orientation bins.

Tab. 2 Accuracy results of V-HOG + ELM for different combinations of η and β %

Region	$\beta = 8$				$\beta = 12$				$\beta = 16$				$\beta = 24$			
	$\eta = 2$	$\eta = 4$	$\eta = 8$	$\eta = 16$	$\eta = 2$	$\eta = 4$	$\eta = 8$	$\eta = 16$	$\eta = 2$	$\eta = 4$	$\eta = 8$	$\eta = 16$	$\eta = 2$	$\eta = 4$	$\eta = 8$	$\eta = 16$
Front	91.68	97.22	96.62	96.81	94.02	97.23	**97.78**	97.23	94.67	97.65	**97.96**	97.14	96.03	97.65	**97.92**	97.19
Left	94.61	96.45	**96.48**	96.29	95.21	96.65	**97.07**	96.02	95.41	**97.26**	96.43	96.14	95.97	**97.08**	96.43	94.86
Right	93.39	**94.24**	94.23	94.15	94.64	**94.86**	94.81	93.15	94.47	**95.09**	94.69	92.88	94.72	**94.75**	93.69	93.20
Far	90.04	**94.95**	94.71	94.90	90.53	95.01	**95.40**	94.70	90.72	95.01	**95.07**	94.08	90.35	**95.27**	93.80	93.58

For the VS-HOG descriptors proposed in this paper, there are also two design parameters, β and η. The experiments are performed with the same configurations $\eta =$ 2, 4, 8 and 16, $\beta = 8$, 12, 16 and 24. ELM is also used as a two-class classifier. Tab. 3 shows the results of our experiments.

Tab. 3 shows the best results in bold font. We can clearly draw a conclusion that the better accuracy performance is obtained using the VS-HOG method than the V-HOG method. That is to say, the correct detection rate is increased by about 1%. However, we can also find that the result is not so well when the number of cells is assigned $\eta = 2$ although it is better than V-HOG. The reason is that the symmetry characteristics cannot be described quite well in detail when the image window is divided merely into only two vertical symmetrical parts. Furthermore, as we can see in Tab. 3, the performance increases gradually with the growth of the number of orientation bins, and the better result can be obtained when $\beta = 16$. The same as V-HOG, the better performance can be achieved for $\eta = 4$ or $\eta = 8$ than other cases, if β is set to be a fixed value.

In Tab. 4, the best results of all three series experiments are summarized. In this table, the parameters for HOG, V-HOG and VS-HOG are summarized and their respective accuracies or correct detection rates (DR) are compared. The dimension of HOG vector (DHV) denotes the dimension of the generated HOG vector according to

Tab. 3 Accuracy results of VS-HOG + ELM for different combinations of η and β %

Region	$\beta=8$				$\beta=12$				$\beta=16$				$\beta=24$			
	$\eta=2$	$\eta=4$	$\eta=8$	$\eta=16$	$\eta=2$	$\eta=4$	$\eta=8$	$\eta=16$	$\eta=2$	$\eta=4$	$\eta=8$	$\eta=16$	$\eta=2$	$\eta=4$	$\eta=8$	$\eta=16$
Front	93.62	**97.85**	97.50	97.16	94.56	**97.88**	98.18	97.64	95.44	98.21	**98.64**	97.76	96.71	98.05	**98.58**	97.20
Left	95.50	**97.04**	96.97	96.64	95.94	**97.30**	96.93	96.37	96.07	**97.79**	96.85	96.53	96.35	**97.54**	96.84	95.98
Right	94.46	**95.64**	95.61	95.43	95.22	**96.26**	96.06	95.18	95.68	**96.34**	95.75	95.44	95.92	**96.12**	95.38	94.43
Far	90.99	**95.50**	95.04	95.09	91.84	**95.54**	95.34	95.22	92.64	**95.72**	95.50	95.09	92.37	**95.63**	95.25	94.92

the parameters, such as s, η and β. For simplicity, there is no fold block overlapping during HOG computation in our experiments. The feature extraction time (FET) is the time required to generate the feature vector for a given sample. The classification time (CT) is the time required by the classifier to make a final decision on whether it is a vehicle or not. The total time (TT) required to process a sample is the sum of the feature extraction and the classification time. The time in Tab. 4 is the average time measured on the database.

Tab. 4 Experimental results comparison among HOG + SVM, V-HOG + ELM and VS-HOG + ELM

Method	Parameter	Front	Left	Right	Far	Mean
HOG + SVM	s	8	8	8	4	
	β	12	8	12	12	
	DHV	768	512	768	3 072	
	FET/ms	2.615	1.974	2.517	5.836	3.236
	CT/ms	0.309	0.539	0.325	4.518	1.423
	TT/ms	2.924	2.513	2.842	10.354	4.658
	DR/%	99.10	98.81	97.23	96.53	97.92
V-HOG + ELM	η	8	4	4	8	
	β	16	16	16	12	
	DHV	128	64	64	72	
	FET/ms	2.368	1.684	1.527	1.856	1.859
	CT/ms	0.014	0.012	0.011	0.014	0.013
	TT/ms	2.382	1.696	1.538	1.870	1.872
	DR/%	97.96	97.26	95.09	95.40	96.43
VS-HOG + ELM	η	8	4	4	4	
	β	16	16	16	16	
	DHV	128	64	64	64	
	FET/ms	2.542	1.546	1.851	1.684	1.906
	CT/ms	0.013	0.011	0.012	0.011	0.012
	TT/ms	2.555	1.557	1.863	1.695	1.917
	DR/%	98.64	97.79	96.34	95.72	97.12

First, it can be concluded that VS-HOG has better detection rate results than V-HOG, and its mean detection rate is 97.12% which is close to classic HOG descriptors. Then, we can also find that the mean value of the detection rate of HOG is slightly higher than that of VS-HOG. The reason is that different initial parameter values are selected, which determine the dimension of HOG vectors and the HOG capability of describing local image features. On the other hand, HOG vectors can slowly be extracted from the images and the long HOG vector will also make the classification slow. As Tab. 4 shows, compared with the HOG-SVM method, the computation time is reduced by more than half when the method proposed in this paper is used. Therefore, as for the applications of on-road vehicle verification, VS-HOG achieves a better trade-off between cost and performance because of shorter descriptor vectors. Finally, by observing Tab. 2 and Tab. 3, an important conclusion can be drawn that VS-HOG can obtain better accuracy results than V-HOG when $\eta = 4$; that is to say, VS-HOG descriptor can gain a fairly good performance when the images are divided into a few cells.

As we can see in Tab. 4, although FET of VS-HOG is longer than that of V-HOG due to the feature vector recombination, all the computation total time of VS-HOG is still less than 3 ms, which can satisfy the requirements of most real-time vehicle detection systems. Using VS-HOG and ELM as the vehicle verification tools, there remains much time for other tasks, such as hypothesis generation and vehicle tracking. From Tab. 4, it can also be concluded that the ELM classification speed is more than 10 times that of the SVM method, the mean of which is less than 0.02 ms.

Through a great deal of experiments, the parameter values of VS-HOG are determined to achieve a better trade-off between performance and computational cost. The number of cells η should be set to be 4 in order to relieve the computational requirements. Another parameter β should be set to be 12 or 16, in which case the accuracy rate can be above 95%.

4 Conclusion

In this study, a variation of HOG, namely VS-HOG, is proposed, which reduces the computational cost by using the symmetry characteristics of vehicles, and the accuracy performance is improved steadily at the same time. The main idea of VS-HOG is that the symmetry characteristics of vehicle structure are exploited to improve the efficiency and accuracy of vehicle verification. In the classification stage, ELM is used as a two-class classifier to achieve greater efficiency. The experimental results indicate that the efficiency and accuracy of the classifier are both greatly enhanced by using VS-HOG descriptors. Furthermore, in the tests, the presented method for the vehicle verification by using VS-HOG and ELM is proved to be an excellent alternative method for real-time vehicle verification.

Based on the VS-HOG and ELM vehicle verification methodology, further work will be concentrated on the realization of on-road vehicle detection systems, in which the performance of the method proposed in this paper will be proved in practice.

References

[1] Bertozzi M, Broggi A, Fascioli A. Vision-based intelligent vehicles: state of the art and perspectives [J]. *Robotics and Autonomous Systems*, 2000, **32**(1): 1 - 16.

[2] Bishop R. A survey of intelligent vehicle applications worldwide [C]//*Proceedings of the IEEE Intelligent Vehicles Symposium* 2000. Dearborn, USA, 2000: 25 - 30.

[3] Darms M S, Rybski P E, Baker C, et al. Obstacle detection and tracking for the urban challenge [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2009, **10**(3): 475 - 85.

[4] Sun Z H, Bebis G, Miller R. On-road vehicle detection: a review [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, **28**(5): 694 - 711.

[5] Sun Z H, Bebis G. Monocular precrash vehicle detection: Features and classifiers [J]. *IEEE Transactions on Image Processing*, 2006, **15**(7): 2019 - 2034.

[6] Liu W, Wen X Z, Duan B B, et al. Rear vehicle detection and tracking for lane change assist [C]//*Proceedings of the IEEE Intelligent Vehicle Symposium* . Istanbul, Turkey, 2007: 252 - 257.

[7] Chang W C, Cho C W. Online boosting for vehicle detection [J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*: *Cybernetics*, 2010, **40**(3): 892 - 902.

[8] Sivaraman S, Trivedi M M. A general active-learning framework for on-road vehicle recognition and tracking [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2010, **11**(2): 267 - 276.

[9] Teoh S S, Bräunl T. Symmetry-based monocular vehicle detection system [J]. *Machine Vision and Applications*, 2012, **23**(5): 831 - 842.

[10] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Kauai, USA, 2001, **1**: I-511 - I-518.

[11] Sun Z H, Bebis G, Miller R. On-road vehicle detection using evolutionary Gabor filter optimization [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2005, **6**(2): 125 - 137.

[12] Arróspide J, Salgado L. Log-Gabor filters for image-based vehicle verification [J]. *IEEE Transactions on Image Processing*, 2013, **22**(6): 2286 - 2295.

[13] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Los Alamitos, USA, 2005: 886 - 93.

[14] Negri P, Clady X, Hanif S M, et al. A cascade of boosted generative and discriminative classifiers for vehicle detection [J]. *EURASIP Journal on Advances in Signal Processing*, 2008, **2008**: 782432.

[15] Gandhi T, Trivedi M M. Video based surround vehicle detection, classification and logging from moving platforms: issues and approaches [C]//*Proceedings of the IEEE Intelligent Vehicle Symposium* 2007. New York, 2007: 1067 - 1071.

[16] Southall B, Bansal M, Eledath J. Real-time vehicle detection for highway driving [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Miami, USA, 2009: 541 - 548.

[17] Cheon M, Lee W, Yoon C, et al. Vision-based vehicle detection system with consideration of the detecting location [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2012, **13**(3): 1243 - 1252.

[18] Arróspide J, Salgado L, Camplani M. Image-based on-road vehicle detection using cost-effective histograms of oriented gradients [J]. *Journal of Visual Communication and Image Representation*, 2013, **24**(7): 1182 - 1190.

[19] GTI. Vehicle image database [EB/OL]. (2012) [2014-07-06]. http://www.gti.ssr.upm.es/data/Vehicle_database.html.

[20] Dai B, Fu Y J, Wu T, et al. A vehicle detection method via symmetry in multi-scale windows [C]//*Proceedings of the* 2*nd IEEE Conference on Industrial Electronics and Applications*. Harbin, China, 2007: 1827 - 1831.

[21] Huang G B, Wang D H, Lan Y. Extreme learning machines: a survey [J]. *International Journal of Machine Learning and Cybernetics*, 2011, **2**: 107 - 122.

[22] Huang G B, Zhou H M, Ding X J, et al. Extreme learning machine for regression and multiclass classification [J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*: *Cybernetics*, 2012, **42**(2): 513 - 529.

[23] Chang C C, Lin C J. LIBSVM: a library for support vector machines [J]. *ACM Transactions on Intelligent Systems and Technology*, 2011, **2**(3):27.

基于垂直对称 HOG 和极限学习机的在线车辆验证方法

范延军[1,2]　张　雷[1]　张为公[1]

([1] 东南大学仪器科学与工程学院, 南京 210096)

([2] 中国计量学院计算机科学与技术系, 杭州 310018)

摘要:为了满足车辆在线应用所需的实时性和准确性,提出了一种在线实时车辆验证的解决方法. 首先,基于对车辆图像对称特性的分析,提出了垂直对称 HOG 描述子,用来提取图像的特征. 在车辆分类阶段,为了提高算法的实时性,使用极限学习机作为分类器. 与其他经典算法的实验数据进行比较,结果表明基于垂直对称 HOG 和极限学习机的车辆验证方法能够在算法的运行效果与计算代价方面取得较好的折中,并且能够在尽可能保证算法效果的同时降低计算开销. 实验结果进一步表明,提出的车辆验证方法在执行效率和准确性方面均能取得较好的效果,能够满足车辆的在线实时应用要求.

关键词:梯度方向直方图;垂直对称梯度方向直方图;车辆验证;极限学习机

中图分类号:TP391. 4