BENCHMARK EVALUATION OF HOG DESCRIPTORS AS FEATURES FOR CLASSIFICATION OF TRAFFIC SIGNS

Hasan Fleyeh[1], Janina Roch[2]

[1] *Department of Computer Engineering, School of Technology and Business Studies, Dalarna University, Sweden*

[2] *Department of Business Studies and Economic, TU Kaiserslautern, Kaiserslautern, Germany*

Abstract: The purpose of this paper is to analyze the performance of the Histograms of Oriented Gradients (HOG) as descriptors for traffic signs recognition. The test dataset consists of speed limit traffic signs because of their high inter-class similarities. HOG features of speed limit signs, which were extracted from different traffic scenes, were computed and a Gentle AdaBoost classifier was invoked to evaluate the different features. The performance of HOG was tested with a dataset consisting of 1727 Swedish speed signs images. Different numbers of HOG features per descriptor, ranging from 36 features up 396 features, were computed for each traffic sign in the benchmark testing. The results show that HOG features perform high classification rate as the Gentle AdaBoost classification rate was 99.42%, and they are suitable to real time traffic sign recognition. However, it is found that changing the number of orientation bins has insignificant effect on the classification rate. In addition to this, HOG descriptors are not robust with respect to sign orientation.

Keywords: traffic signs, HOG descriptors, Gentle AdaBoost, classification.

1. Introduction

Traffic sign recognition is an important component of Advanced Driver Assistant Systems (ADAS). Recognizing traffic signs gives important information to the driver and helps to identify changing conditions and understand complex traffic situations. Early recognition of speed limit signs can help a driver to adjust the speed of the vehicle in time which represents a potential to not only reduce pollution, but also the number of accidents which will save lives (Fleyeh, 2004).

Swedish speed signs, as one class of traffic signs, have identifying features which make them distinguishable from other objects and increase the intra-class variability with other traffic sign classes. The colors as well as the shape and the digits give important information (Fig. 1). Therefore, they represent the basis to recognize the different types of speed signs. However, all speed limit signs have a round surface with a red border, a yellow fill and black digits. This inter-class similarity makes the classification of this class of traffic signs a challenging issue.

---

[1] Corresponding author: hfl@du.se

Fig. 1.
Speed Limit Traffic Signs
Source: Wikipedia (2010)

Although speed signs are located in good visible spots on the side of the road where they can be expected by the drivers (Fleyeh, 2006), they can be found in different conditions. They could be damaged or mounted on a damaged pole. Sign colors fade by aging of the traffic sign because of environmental effects. In the dusk or dawn the colors are perceived to be much darker. Shadows can change the perception of the color of parts or the entire speed sign. Bad weather conditions such as snow, rain or fog can affect the color or sharpness of the sign. Images of traffic sign can be blurry when they are collected from a moving vehicle. The signs can be rotated or occluded by paint, trees or other obstacles. A number of examples of problematic speed sign images are shown in Fig. 2.

Fig. 2.
Examples for Critical Images of Speed Signs Included in the Dataset

Furthermore, speed signs are not the only signs with a red border, yellow fill and black pictogram. There are a number of signs which are very similar to speed limit signs. Some examples of similar looking signs are shown in Fig. 3.

Fig. 3.
Examples of Critical Images which Look Similar to Speed Sign Images

Classifying this group of traffic signs, i.e. the speed limit traffic signs, which is characterized by high inter-class similarity and low intra-class variability needs highly sophisticated feature descriptors. Among these descriptors is the Histograms of Oriented Gradients (HOG) which was introduced by Dalal and Triggs (2005) as features for pedestrian recognition. However, it has shown that it is capable of describing other objects than pedestrians (Dong et al., 2010; Zhang et al., 2007; Zhang, 2010; Bertozzi et al., 2007; Zaharescu et al., 2009; Rybski et al., 2010).

The main contribution of this paper is to investigate the suitability of this descriptor to achieve traffic sign recognition. Performance manifestation conducted by this paper include high recognition rates, real-time implementation, many traffic sign categories as recognition objects, robustness for variant environments, and feasibility under poor visibility conditions. A benchmark test was conducted by training a Gentle AdaBoost (Friedman et al., 2000) as a multiclass classifier with a set of HOG features computed for this set of traffic signs. The classifier was trained and tested with different numbers of HOG features and the behavior of the classifier should reflect the ability of HOG descriptors to represent this group of traffic signs.

The remainder of the paper is constructed so that in Section 2 the related work is presented. Section 3 demonstrates the system setup and in Section 4 the experiments conducted on the HOG descriptors are illustrated and discussed. Finally, the conclusions are drawn in Section 5.

2. Related Work

In recent years, research in object recognition has grown rapidly because of the real need for such systems in future applications.

HOG descriptors were invoked in different applications other than human detection which it was originally designed for.

Dong et al. (2010) developed a HOG-based multistage approach for object detection and object pose recognition for service robots. This approach made use of the merits of both multi-class and bi-class HOG-based detectors to form a three-stage algorithm at low computing cost.

Zhang et al. (2007) proposed a multi-resolution framework which was based on HOG descriptors for general object detection. Different resolutions were represented using a coarse-to-fine feature hierarchy. During detection, the lower resolution features were initially used to reject the majority of negative windows at relatively low cost, leaving a relatively small number of windows to be processed in higher resolutions. The lowest resolution classifier was able to reject almost 80% of the windows with a detection rate of 98.6%. When combined, the first three low resolutions rejected more than 99.9% of the windows, with a detection rate of almost 94%.

Zhang (2010) proposed an off-line signature verification and identification method by invoking a local shape descriptors pyramid of HOG. The signature image was divided into a number of non-overlapped regions in each level of the pyramid. It starts by the whole signature image as one region, then 4 regions, and 16 regions and so on. HOG features are computed for each cell in the level and merged together to represent the whole signature. A binary SVM classifier was invoked for the purpose of verification and multi-class SVM was invoked for identification. The system achieved FRR of 4.0% and FAR 3.25% for skilled forged signatures.

Bertozzi et al. (2007) invoked HOG and SVM classifier to filter the region of interest (ROI) gathered by visible and far infrared cameras for the purposes of pedestrain detection. The system was able to classify 91% of the pedestrains in the ROI areas.
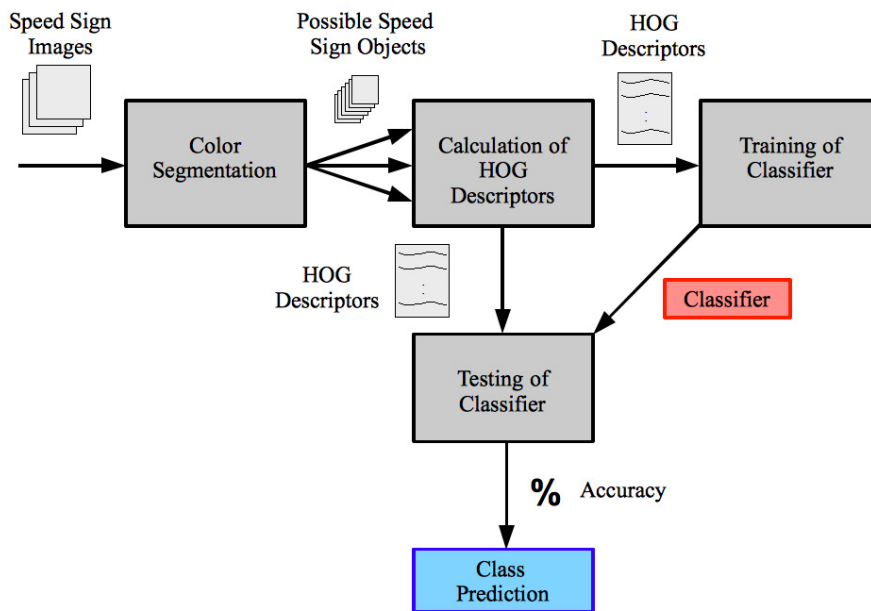
Zaharescu et al. (2009) proposed a 3D feature detector (MeshDOG) and a 3D feature descriptor (MeshHOG) for uniformly triangulated meshes, invariant to changes in rotation, translation, and scale. The descriptor was able to capture the local geometric and/or photometric properties in a succinct fashion.

Rybski et al. (2010) developed vision-based algorithms for determining vehicle orientation of vehicles in images. The proposed system was trained with a set of (HOG) descriptors to recognize different orientations of vehicles detected in imagery. The HOG orientation-specific classifiers achieved 88% classification accuracy on a test database of 284 images.

3. Experiment Setup

The block diagram of the proposed approach is illustrated in Fig. 4. The candidate speed limit traffic sign is extracted by the procedure described in subsection 3.1. The HOG descriptors of this object are computed and invoked either for training of the classifier or in the prediction of the speed limit traffic sign when the trained classifier is used for this purpose.

Fig. 4.
Block Diagram of the System Setup

3.1. Object Extraction

Extracting speed limit traffic signs starts by exploiting a shadow and highlight invariant color segmentation algorithm (Fleyeh, 2006) to extract all red objects in the image. The second column of Fig. 5 depicts the results of the segmentation of a number of traffic scenes. Binary objects in the segmented image were labeled using connected components labeling. All objects with red rims, yellow interiors, and appropriate dimensions were selected for further investigation.

The process of extracting the object under consideration from the image can be described in the following steps:

1. Fill the red region which represents a candidate sign (subimage A) with white pixels in order to produce (subimage B).

2. Extract the area corresponding to the pictogram of the sign by XOR operator between subimage A and subimage B, i.e. (subimage C = subimage A XOR subimage B). The resulting area will correspond to the pictogram of the traffic sign.

3. Extract the corresponding pictogram and convert it into grey level (subimage D).

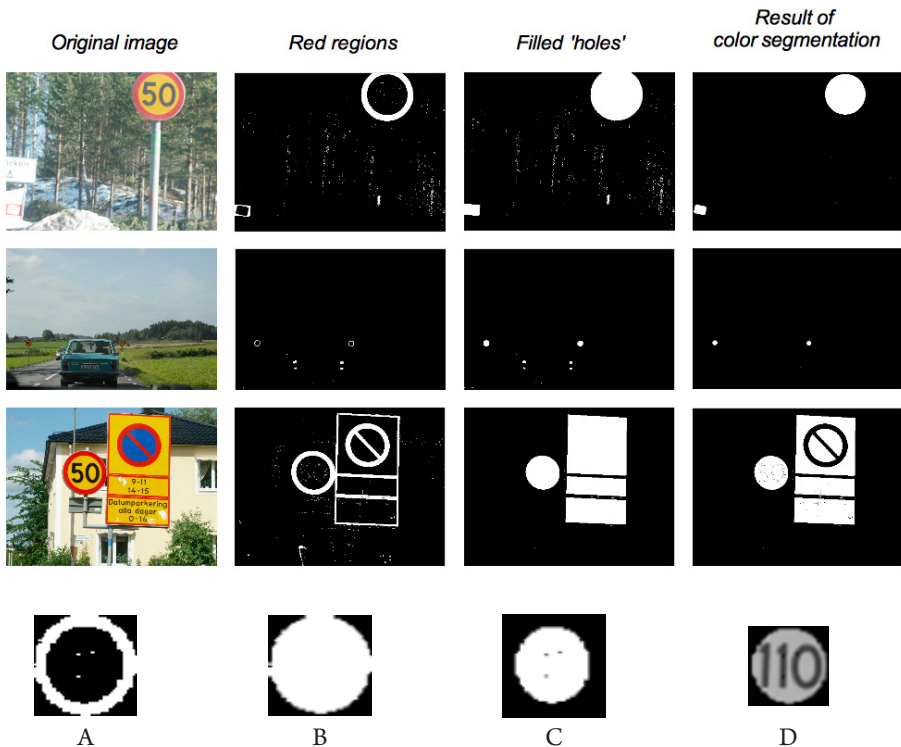4. Compute the features of this pictogram for training and testing.

Fig. 5.
Steps of Object Extraction

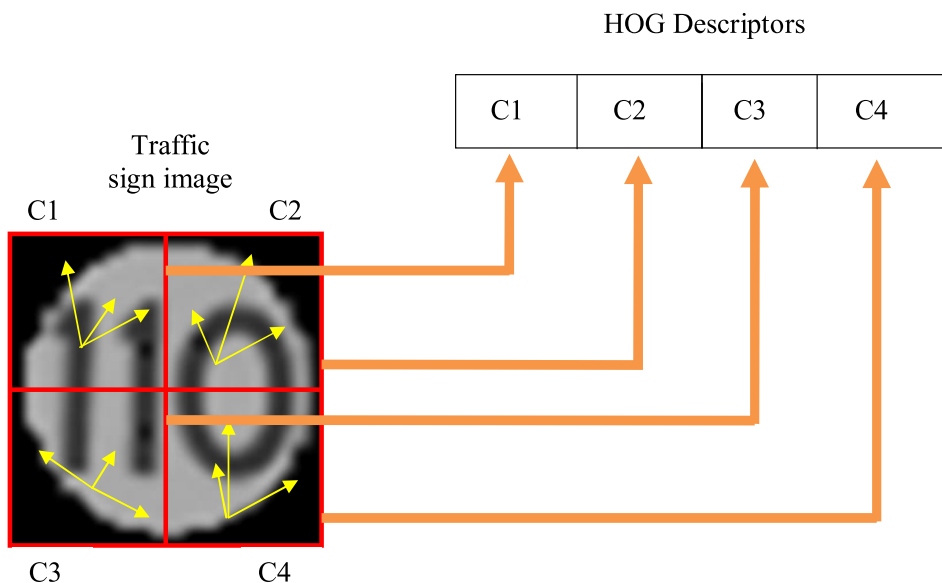## 3.2. Histogram of Oriented Gradients (HOG)

Edges play a very important role in computer vision and their orientations describe an important feature for the differentiation and recognition of objects in the scene. HOG-descriptors, which were introduced by Dalal and Triggs (2005), make use of objects' edges to create a set of features to describe the object under consideration.

To compute the HOG descriptors of any image containing the extracted object, this image is divided into a number of cells and a number of orientation bins as depicted in Fig. 6. For each cell a local 1-D histogram of the gradient directions of edge orientations over the pixels of the cell is collected. For better invariance to illumination such as

shadows, the local histogram is accumulated over a larger area called blocks. To improve the contribution of the cells in the final image descriptor, overlapping between these cells is invoked.

Edge orientations are divided into a number of bins. These bins are equally spaced over the interval 0-180° for unsigned gradient and 0-360° for signed gradients. Edge orientations should fit into one of these bins. The histograms collected for the different cells in the bins in the same block are concatenated to make the final set of features of the object under consideration.

The result of the HOG algorithm is a discrete amount of features which describe the input image. The number of features depends on the number of cells and orientation bins.

Fig. 6.
Computing the HOG Descriptors

3.3. Gentle AdaBoost Classifier

Gentle AdaBoost is a boosting algorithm (Schapire, 2003) which was introduced by Friedman et al. (2000). It is a robust and stable version of the Real AdaBoost and performs slightly better than the latter on regular data and considerably better on noisy data. The algorithm uses adaptive Newton steps rather than exact optimization at each step to minimize the exponential criterion in order to stabilize the learning processing.

Let X be a finite training set which is denoted by Eq. (1):

$$X = \left\{(\mathbf{x}_i, y_i) | \mathbf{x}_i \epsilon \mathbb{R}^N, y_i \epsilon \{+1, -1\}\right\}, for\ i = 1,2, \cdots, M \tag{1}$$

where $\mathbf{x}_i$ is ith training data, y_i is its corresponding target label, N is the space of the data set (number of features) and M is the number of training samples. The Gentle AdaBoost algorithm can be described by the following steps:

1. Start the weights $w_i = {}^1\!/_N$, $i = 1,2, \dots, N, F(x) = 0$.

2. Repeat for $m = 1,2, \dots, M$:

    a. Fit the regression function $f_m(x)$ by weighted least-squares of y_i and x_i with the weights w_i.

    b. Update $F(x) \leftarrow F(x) + f_m(x)$.

    c. Update $w_i \leftarrow w_i \exp\left(-y_i f_m(x_i)\right)$ and renormalize.

3. Output the classifier sign $[F(x)] = sign[\sum_{m=1}^{M} f_m(x)]$

The generalization of binary Gentle AdaBoost to its multiclass version is straightforward. The multiclass problem can be transformed into several binary classification problems. This can be accomplished by using either the One-Against-One or the One-Against-All model.

3.4. The Dataset

The dataset which was invoked in this work comprises 1727 images and it is available from Dalarna University (Fleyeh, 2009).

A total of 1710 speed limit signs and 1025 non-traffic sign objects were extracted by the objects extractor described in Subsection 3.1.
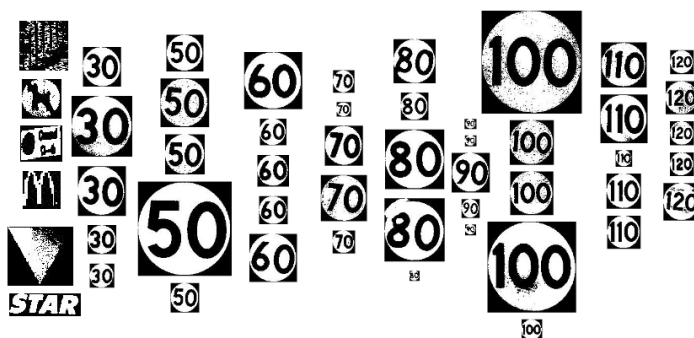
The number of speed limit traffic signs in each class is depicted in Table 1. Fig. 7 shows part of the database of the pictograms used for recognition of the speed limit signs.

Table 1
The Dataset used to Test the HOG

Class	Speed km/h	Abbreviation	Number of Extracted signs
0	-	NOSL	1025
1	30	SL30	98
2	50	SL50	213
3	60	SL60	114
4	70	SL70	380
5	80	SL80	104
6	90	SL90	415
7	100	SL100	198
8	110	SL110	116
9	120	SL120	72
	Total		**2735**

Fig. 7.
The Training Dataset

4. Experiments and Results

To analyze the performance of HOG features, region of interests (ROI) containing the traffic sign candidates were extracted from the traffic scenes. The HOG descriptors of all 2735 objects which were detected by color segmentation were computed without size normalization or any other transformation of the ROI. The HOG of the candidates were then assigned to one of the 10 classes shown in Table 1.

4.1. Effect of Number of Features

In order to evaluate the effect of number of features deduced from the ROI, a 10-fold cross validation experiment was employed. In this experiment a Gentle AdaBoost classifier was trained and tested with different numbers of features. The number of features was based on different numbers of cells per block and a constant number of orientation bins. The number of orientation bins was 9 while the number of cells varied from 2x2 to 6x6 with an increment of 1 in each direction, i.e. 2x2, 3x3... 6x6. This gives 4, 9, 16, 25, and 36 cells in each block. Table 2 depicts the classification accuracy for the different numbers of features in the different folds.

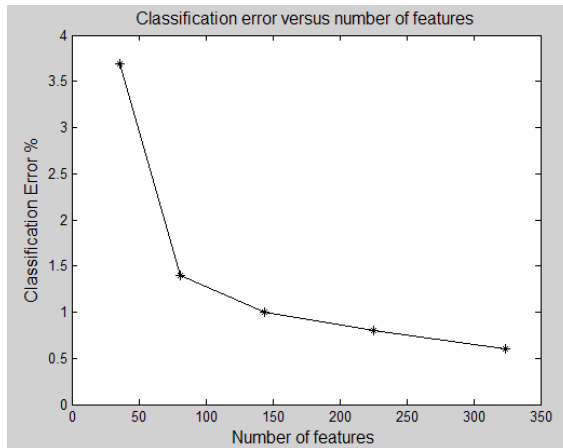A plot of the classification error versus the number of features indicates clearly that

increasing the number of features deduced per block decreases the classification error. Fig. 8 depicts the relationship between the classification error and the number of features. Since the extracted traffic signs which were exploited in this experiment were of different scales, it is obvious that HOG descriptors are scale invariant. This property is essential to traffic sign recognition because images or footage may contain traffic signs with different sizes depending on the distance between the vehicle and the traffic sign. To have a set of descriptors which performs with scale invariance means that time required for normalization can be saved for real time applications.

Table 2

Classification Accuracy versus Number of HOG Features

Fold	Classification Accuracy %				
	Features = 36 2x2 cells 9 bins	Features = 81 3x3 cells 9 bins	Features = 144 4x4 cells 9 bins	Features = 225 5x5 cells 9 bins	Features = 324 6x6 cells 9 bins
1	96.7	98.5	99.3	99.3	99.3
2	96.0	98.2	98.5	98.9	98.9
3	97.1	99.3	99.6	99.6	100
4	95.2	98.2	98.5	98.5	98.9
5	95.2	97.4	98.5	99.3	98.9
6	96.7	98.5	98.5	98.9	99.3
7	96.0	98.5	98.9	98.9	99.6
8	97.4	100	99.6	100	100
9	96.0	98.9	98.5	99.3	99.6
10	96.0	98.9	100	99.3	99.3
Ave.	**96.31**	**98.6**	**99.0**	**99.2**	**99.4**

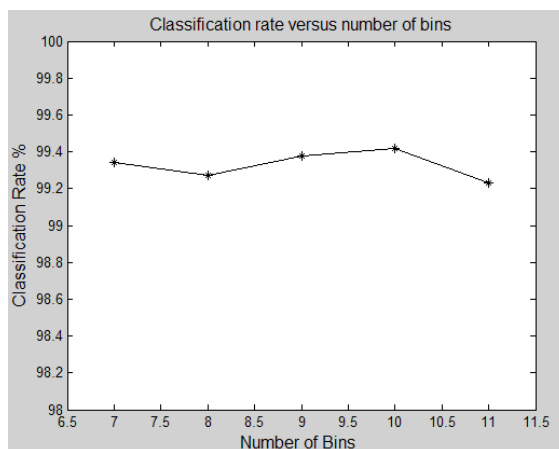**Fig. 8.**

Classification Error versus Number of Features

4.2. Effect of Number of Orientation Bins

The effect of number of orientation bins on the classification rate was also tested. The Gentle AdaBoost classifier was trained with different numbers of HOG features which were derived from the training set based on different numbers of orientation bins. In contrast to the variation of cells, the variation of the bin size did not show major differences in the accuracy. However, a slight improvement can be observed with 10 bins with a size of 18° each (Table 3). Fig. 9 depicts the effect of the number of orientation bins on the classification rate.

Table 3

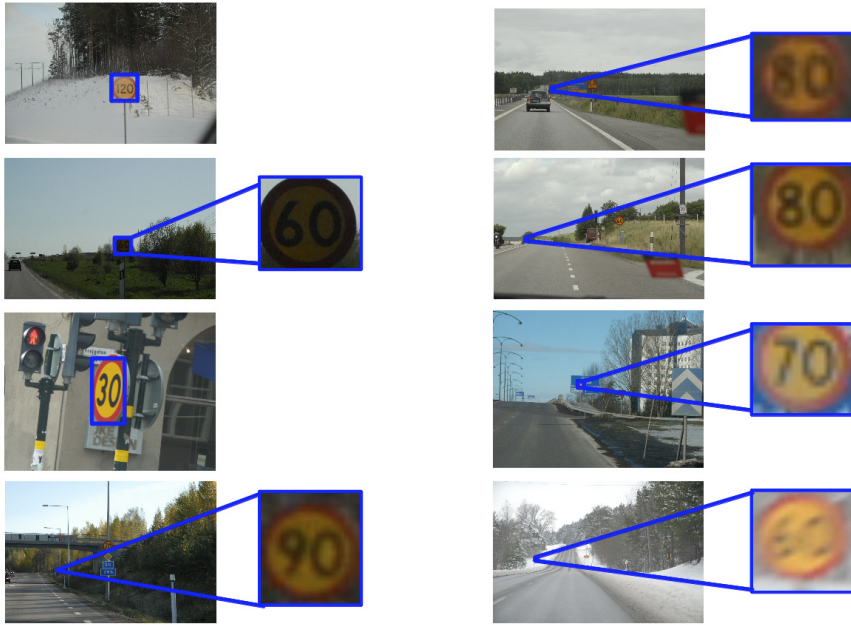*Effect of Orientation Bins on the Classification Rate*

Number of Orientation Bins	Number of HOG Features	Average Accuracy %
7	(6x6x7) = 252	99.34
8	(6x6x8) = 288	99.27
9	(6x6x9) = 324	99.38
10	(6x6x10) = 360	99.42
11	(6x6x11) = 396	99.23

Fig. 9.

Classification Accuracy versus Number of Orientation Bins

Deep analysis of the objects which were misclassified showed that the classifier committed 16 misclassifications. Among them there was one 90 km/h speed limit sign which was classified as 80 km/h and all other errors were false negatives. However, the classifier did not generate any false positives. Common to all false negatives was the fact that they were blurry because of bad weather conditions or because they were far away in the background. A number of examples which were misclassified are depicted Fig. 10.
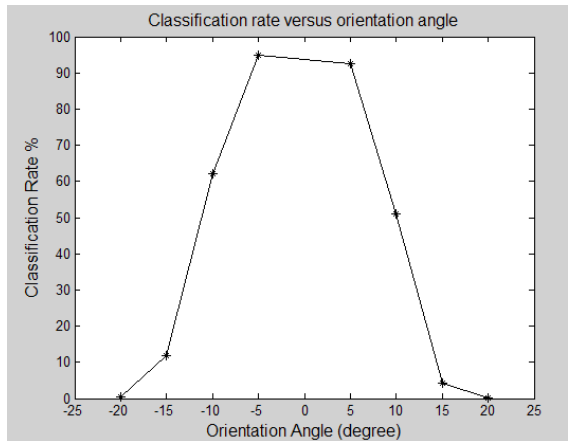
Fig. 10.
Signs Misclassified by the Classifier

4.3. Classifying Disoriented Traffic Signs

Traffic signs are usually installed on poles which are always vertical with respect to the ground level. However, for many reasons such as the nature of the ground or environmental effects, these traffic signs depart from the vertical situation. Two of the speed limit traffic signs shown in Fig. 2 belong to this category. In this experiment, the effect of disorientation (rotation) of traffic signs is studied. The Gentle AdaBoost classifier was trained with 360 HOG features (6x6 cells and 10 bins). All of these features were derived from vertically oriented speed limit signs.

The classifier was tested with HOG features derived from speed limit images which were rotated by different angles in clockwise and counter clockwise directions. Fig. 11 depicts a plot of the classification rate versus angle of rotation.
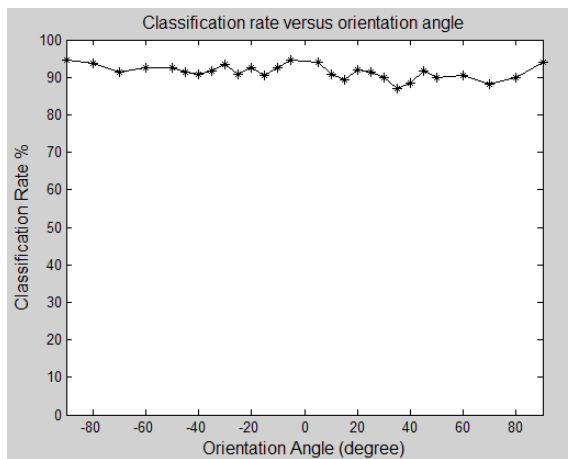
The plot shows a large drop in the classification rate when disoriented traffic signs were classified with the trained classifier, which means that HOG descriptors are rotation variant and cannot be utilized in these situations. As this is a crucial issue as far as traffic signs are concerned, a proper solution is essential in this case to avoid this kind of invariance.

Fig. 11.

Classification Rate of the Gentle AdaBoost Trained with Vertically Aligned Signs versus Angle of Rotation of Traffic Signs

This experiment was repeated by training the Gentle AdaBoost with a set of HOG descriptors which was derived from speed limit sign rotated by different angles between -90° to 90°. These HOG descriptors were derived in the same manner described in the first part of this experiment and the number of features was the same. The classification rate did not drop as traffic signs rotate. Although there was a slight variation in the classification rate, the average classification rate was 92%. The plot of the classification rate versus the angle of rotation when Gentle AdaBoost was trained with rotated signs is depicted in Fig. 12.

Fig. 12.

Classification Rate of the Gentle AdaBoost Trained with Rotated Signs versus Angle of Rotation of Traffic Signs

4.4. The Overall Performance of Classification

To test the effect of HOG descriptors as discrimination features among speed limit signs and non-speed limit signs, the set of images illustrated in Table 1 is invoked. The set was divided by 66.6% for training and 33.3% for testing. The Gentle AdaBoost was trained with HOG descriptors of the training subset and tested with the remainder of the images. This experiment was repeated for different numbers of HOG descriptors and each time the rejection ratio was computed. Table 4 depicts the ratio of rejection for the non-speed limit signs for different values of HOG descriptors.

Table 4

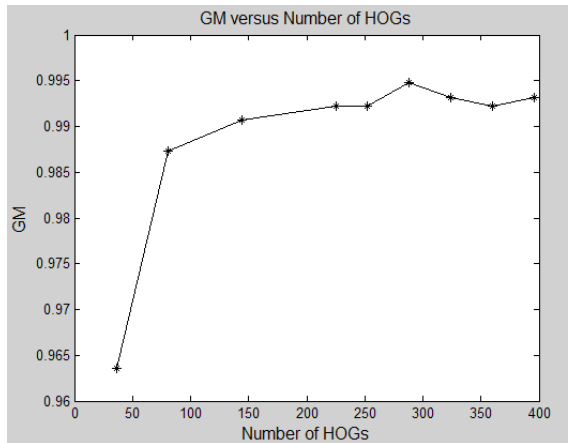*Performance Analysis of the Gentle AdaBoost based on Different Numbers of HOG Descriptors*

HOG Features	TP	TN	FP	FN	Misclassified
(2x2x9) = 36	543	328	1	16	24
(3x3x9) = 81	570	328	1	7	6
(4x4x9) = 144	574	328	1	7	2
(5x5x9) = 225	574	329	0	6	3
(6x6x7) = 252	574	329	0	6	3
(6x6x8) = 288	577	329	0	5	1
(6x6x9) = 324	575	329	0	6	2
(6x6x10) = 360	574	329	0	8	1
(6x6x11) = 396	575	329	0	5	3

In order to evaluate the overall performance of the classifier versus the number of features involved in each case, the geometric mean was calculated for each number of features as follows (Eq. (2)):
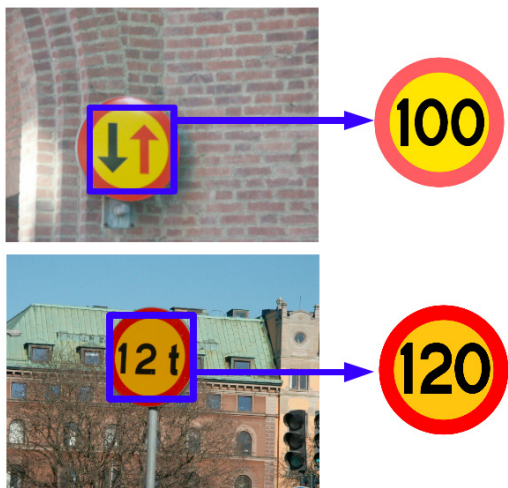
$$GM = \sqrt{\left(\frac{TP}{P}\right) \times \left(\frac{TN}{N}\right)} \qquad (2)$$

Where TP is the number of the true positives (speed limit signs classified as speed limit signs), P is the number of speed limits sign involved in the experiment, TN is the number of true negatives (non-speed limit signs classified as non-speed limit signs), and N is the number of non-speed limits signs. Fig. 13 depicts a plot of the geometric means versus the number of features. The plot indicates that increasing the number of features beyond 225 HOG features is not so beneficial because it burdens the classifier and increases the computational time without obtaining better classification performance. Each of the best classifications generated the same misclassification. Two examples of the false positives and their predicted classes are shown in Fig. 14.
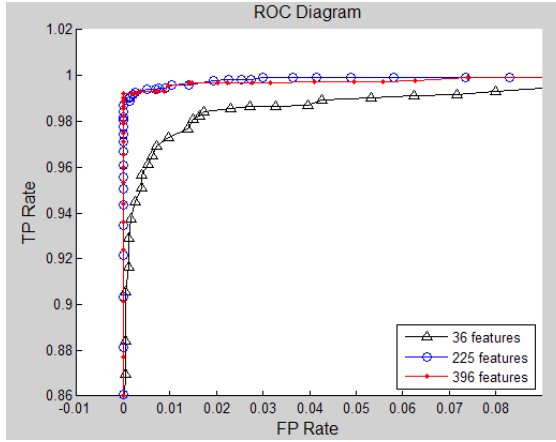
Fig. 13.
Geometric Mean versus Number of HOG Descriptors

Fig. 14.
False Positives Generated by Mixed up HOG Descriptors

In order to validate this issue, the receiver operating characteristic (ROC) diagram of the lowest, the optimum, and the highest number of features was generated and illustrated in Fig. 15. From this plot, it is clear that the optimum features which are 225 descriptors perform as well as the maximum number of features. This result is in full agreement with that obtained from the geometric mean plot.

Fig. 15.
ROC Plot of the Gentle AdaBoost Trained with the Lowest Number of Features (36 HOGs), Highest Number of Features (396 HOGs), and Best Number of Features (225 HOGs)
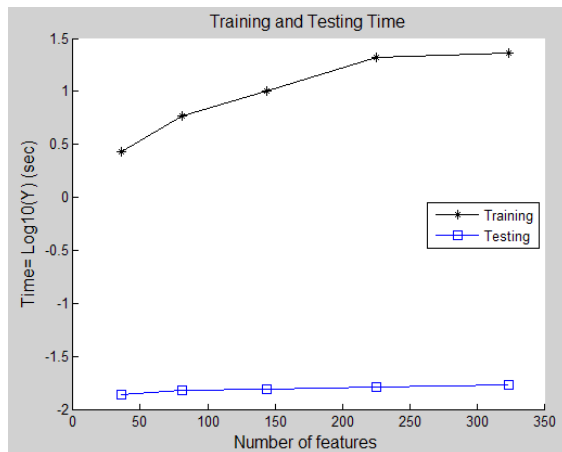
4.5. Training and Testing Time

This experiment is accomplished using Macbook 6.1 with Intel Core 2 Duo, 2.26 GHz processor and the code was implemented using Matlab R2012a (maci64) version. Timings of training the Gentle AdaBoost classifier with 2735 descriptors together with the classification time of a speed limit sign using different numbers of HOG descriptors are depicted in Table 5. While training time increases almost linearly, the testing time is almost constant regardless of the number of features. The increment in the testing time between a low number of features and a high number of features is not crucial. Fig. 16 depicts a plot of training and testing times versus number of features.

Table 5
Time Required for Training and Testing of Different Numbers of HOG Descriptors

Number of HOG Features	Time of Training (sec.)	Time of classification (msec.)
36	2.67	13.8
81	5.83	15.2
144	10.16	15.4
225	21.02	16.2
324	22.77	16.8

Fig. 16.
Training and Testing Time of the Classifier with HOG Descriptors

5. Conclusion

Histograms of Oriented Gradients (HOG) are descriptors which were developed for human detection. However, they have been exploited in other fields of computer vision and object detection. This paper investigated the suitability of these descriptors for traffic sign recognition and classification. For this purpose, a dataset of speed limit traffic signs were invoked. The major outcomes of this benchmark test are that HOG descriptors are suitable for traffic sign recognition because they are easy to implement, characterized with high robustness against intra-class variability, and fast enough for real time detection and classification. The descriptors are scale invariant which means that traffic signs can be extracted and classified without any need for size normalization. In addition, the number of descriptors needed for classification is not more than 225 which also support real time applications.

According to the experiments a slight change in the number of orientation bins is not effective and does not improve the classification rate. Finally it is important to mention that HOG descriptors are variant against angle of rotation. Disoriented traffic signs might not be correctly classified. Unless the classifier is trained with rotated signs, the classifier trained with HOG fails to classify those disoriented traffic signs.

References

Bertozzi, M.; Broggi, A.; Rose, M.D.; Felisa, M.; Rakotomamonjy, A.; Suard, F. 2007. A pedestrian detector using histograms of oriented gradients and a support vector machine classifier. In *Proceedings of the Intelligent Transportation Systems Conference*, ITSC 2007. IEEE. 143-148.

Dalal, N.; Triggs, B. 2005. Histograms of oriented gradients for human detection. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR 2005. 886-893.

Dong, L.; Yu, X.; Li, L.; Hoe, J. 2010. HOG based multi-stage object detection and pose recognition for service robot. In *Proceedings of the 11th International Conference on Control Automation Robotics & Vision (ICARCV)*. 2495-2500.

Fleyeh, H. 2009. *Traffic Sign Database*. Available from Internet: <http://users.du.se/~hfl/traffic_signs/>.

Fleyeh, H. 2004. Color detection and segmentation for road and traffic signs. In *Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems*, Singapore. 808-813.

Fleyeh, H. 2006. Shadow and Highlight Invariant Colour Segmentation Algorithm for Traffic Signs. In *Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems*, Bangkok, Thailand. 108-114.

Friedman, J.; Hastie, T.; Tibshirani, R. 2000. Additive logistic regression: a statistical view of boosting, *Annuals of Statistics*. DOI: http://dx.doi.org/10.1214/aos/1016218223, 28(2): 337-407.

Rybski, P.; Huber, D.; Morris, D.; Hoffman, R. 2010. Visual classification of coarse vehicle orientation using histogram of oriented gradients features. Presented at the *IEEE Intelligent Vehicles Symposium (IV)*. 921-928.

Schapire, R. 2003. *The boosting approach to machine learning: An overview*. Lecture Notes in Statistics-New York-Springer Verlag. 149-172.

Wikipedia. 2010. Road signs in Sweden. Available from Internet: <http://en.wikipedia.org/wiki/Road_signs_in_Sweden>.

Zaharescu, A.; Boyer, E.; Varanasi, K.; Horaud, R. 2009. Surface feature detection and description with applications to mesh matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, CVPR 2009. 373-380.

Zhang, B. 2010. Off-line signature verification and identification by pyramid histogram of oriented gradients, *International Journal of Intelligent Computing and Cybernetics*. DOI: http://dx.doi.org/10.1108/17563781011094197, 3(4): 611-630.

Zhang, W.; Zelinsky, G.; Samaras, D. 2007. Real-time accurate object detection using multiple resolutions. In *Proceedings of the IEEE 11th International Conference on Computer Vision*, ICCV 2007. 1-8.