Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

A Robust Multi-class Traffic Sign Detection and Classification System using Asymmetric and Symmetric features

Jialin Jiao, Zhong Zheng, Jungme Park, and
Yi L. Murphey, *Senior Member, IEEE*
Department of Electrical and Computer Engineering
University of Michigan-Dearborn
Dearborn, MI 48128 USA
yilu@umich.edu

Yun Luo

Driver Assistance Group
Robert Bosch LLC.
Farmington Hills, MI 49331

***Abstract*—In this paper we present our research work in traffic sign detection and classification. Specifically we present a set of asymmetric Haar-like features that will be shown to be effective in reducing false alarm rates for traffic sign detection, and a robust multi-class traffic sign detection and classification system built based upon the stage-by-stage performance analysis of individual traffic sign detectors trained using Adaboost.**

***Keywords*—asymmetric features, traffic sign detection, multi-class classification**

I. Introduction

Traffic signs are a type of traffic informative devices that use features such as shapes, colors, text, and/or symbols to communicate messages to automotive drivers. Traffic signs are usually placed at designated location by roadside or overhead above the road and carry a great deal of information necessary for successful driving. They describe current traffic situation, prohibit or permit certain directions, warn about risky factors etc. Automatic Traffic Sign Recognition (ATSR) is now considered as a critical part of the Driving Assistance System (DAS) in an intelligent vehicle [1-5]. ATSR is a field of applied computer vision research concerned with the automatic detection and classification of road signs in traffic scenes acquired from the camera of a moving car. The goal is to provide the DAS with the ability to understand its surrounding environment and so permit advanced driver support such as collision prediction and avoidance.

Although traffic signs are man-made objects regulated by government standards and they are located in specific location in traffic scenes, a number of issues make the automatic traffic sign recognition a challenging problem. Traffic signs images are acquired from the cameras mounted in cars moving on the (often uneven) road surface by considerable speed and consequently, the traffic scene images often suffer from vibrations. Traffic signs are frequently occluded partially by other vehicles, trees, pedestrians, etc. Furthermore, traffic recognition algorithms must be suitable for the real-time implementation [6] and the hardware platform must be able to process huge amount of information in video data stream.

Traffic sign recognition is one special case of outdoor object recognition with its own special requirements. First, its application, real time recognition, requires fast computational speed. Second, the system for traffic sign recognition should be accurate with respect to high recognition rate and minimal false alarms. Furthermore the traffic sign recognition system must be robust with respect to variations in lighting, partial occlusion, damage and alteration. Since color images may be affected by varying illumination [7-9] and color camera systems are more expensive, many traffic sign recognition systems need to capture gray scale images. The camera system we use in this research also captures gray scale images.

Many traffic sign detection techniques have been investigated including model based approach, classification based on similarity measure, neural networks, space-variant sensor window, kernel based classifiers, etc. [10-14]. We found that most rule based or model based approaches suffer from the problem of lack of error tolerance and are prone to fail when noise presents, while traditional machine learning based methods (neural network, SVM etc) are usually too computation-intensive and could not guarantee fast processing. Based on these observations, when building our system, we chose an appearance-based machine learning approach which has been proved to perform better.

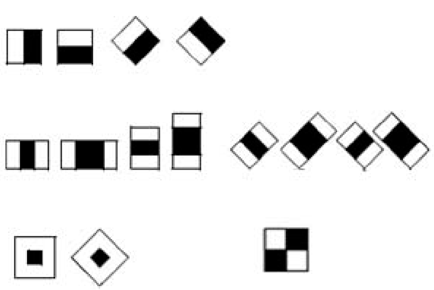
In this paper, we present a new set of asymmetric Haar-like features to be combined with the symmetric features

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

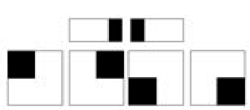
proposed by Viola and Jones [15, 16] for traffic sign detection, and a robust system for multi-class traffic sign detection and classification based on Adaboost learning. We will show that the multi-class traffic sign detection and classification system gives high detection rate with low false alarm rate and is computationally efficient. We will also show that the proposed asymmetric features combined with the symmetric features when used in traffic sign detection can significantly reduce false alarm rate while keeping the detection rate high.

II. EFFECTIVE FEATURES FOR TRAFFIC SIGN DETECTION AND CLASSIFICATION

We attempted to build traffic sign detection and classification systems based on the well-known Adaboost machine learning technique, initially proposed by Paul Viola and Michaels Jones [15, 16]. Adaboost techniques have been successfully applied to many different computer vision applications including face detection [15, 17, 18], pedestrian detection [19], and vehicle detection [20]. This approach selects a small number of critical features from an over-complete set of computable features using AdaBoost based learning algorithm and builds a cascade of classifiers. Such features are called Haar-like features because their templates resemble the Haar wavelet from digital signal processing in the sense that each template consists of two or three opposite “weighted” components with the same “shape”. Fig.1 (a) shows the popular Haar-like feature templates being used in various applications. Note that all those features are symmetric. Fig. 1 (b) illustrates the six asymmetric feature templates we developed for traffic sign detection. In the cascaded AdaBoost training, the most important features (i.e., those that can reject most backgrounds) are selected in the earliest stages of the training.



(a)

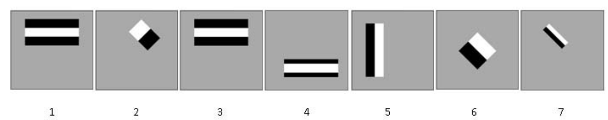


(b)

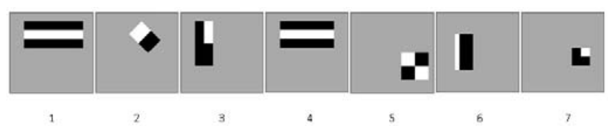
Figure 1. Haar-like feature templates used in the traffic sign detection.

We use Fig. 2 and Fig. 3 to illustrate the effects of the asymmetry features in traffic sign detection. Fig. 2 show the top 7 features selected in the first and second stage in the system trained for stop sign detection using only symmetric features (see (a)) and using both symmetric and asymmetric features (see (b)). Asymmetric feature templates were selected in the 3rd, the 6th and 7th place in Fig. 2 (b), instead of the symmetry features. In Fig. 3, the top 7 features selected by a yield sign detection system trained using only the symmetric Haar-like features or using both the symmetric and asymmetric features are shown (a) and (b) respectively. As we can see, two asymmetric features take the places of the 5th and 6th features. According to the theory of AdaBoost, as long as a new feature template takes the place of an old one, it shows that the new feature is better than the old one in minimizing the weighted error on training data, as least on the training data set.

In order to further analyze the effectiveness of the asymmetric features used for detecting traffic signs, we trained 4 sets of Adaboost systems. Each set contains two detection systems trained to detect the same traffic signs; one



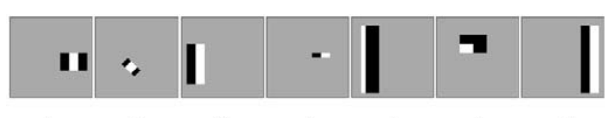
(a)



(b)

Figure 2. Top7 features selected in the first and second stage in the system trained for stop sign detection using only symmetric features (a) and using both symmetric and asymmetric features (b).

(a)



(b)

Figure 3. Top 7 features selected by a yield sign detection system trained using only symmetric Haar-like features (a) and using combined symmetric and asymmetric features (b).

is trained using only the symmetric features shown in Fig. 1 (a), and the other using both the symmetric features in Fig. 1 (a) and the asymmetric features in Fig. 1 (b).

One set of classification systems are trained for detecting traffic signs, particularly stop signs, yield signs and speed limit signs, against background, one set of classification systems are trained for detecting stop signs, one set for yield signs, and one set for speed limit signs.

The set of two detectors for traffic sign detection are to detect whether a given image contains any traffic signs including stop signs, yield signs and speed limit signs. They are trained on the following data set:

$$\Omega_{training}^{joint} = \Omega_{training}^{stop} \cup \Omega_{training}^{yield} \cup \Omega_{training}^{SL} \cup \Omega_{training}^{background} \quad (1)$$

where $\Omega_{training}^{stop}$ is a set of 3257 road images containing stop signs, $\Omega_{training}^{yield}$ is a set of 1008 road images containing yield signs, $\Omega_{training}^{SL}$ is a set of 908 road images containing speed limit signs, and $\Omega_{training}^{background}$ is a set of 6059 outdoor images containing none of those traffic signs. The two stop sign detectors are trained to detect whether a given image contains any stop signs; they are trained on the data set, $\Omega_{training}^{stop} \cup \Omega_{training}^{background}$. The two yield sign detectors are trained to detect whether a given image contains any yield signs and are trained on data set $\Omega_{training}^{yield} \cup \Omega_{training}^{background}$. The two speed limit sign detectors are trained to detect whether a given image contains any speed limit signs and trained on data set, $\Omega_{training}^{SL} \cup \Omega_{training}^{background}$. The curves shown in Fig. 4 and Fig. 5 are the performances of these classifiers on validation data sets, which are separate sets from those used in the training. $\Omega_{validation}^{jo\text{int}} = \Omega_{validation}^{stop} \cup \Omega_{validation}^{yield} \cup \Omega_{validation}^{SL}$ is used by the two traffic sign detectors to generate the performance curves shown in Fig. 4 (a) and Fig. 5 (a); $\Omega_{validation}^{stop}$ by the two stop sign detectors to generate the performance curves shown in Fig. 4 (b) and Fig. 5 (b); $\Omega_{validation}^{yield}$ by the two yield sign detectors to generate the curves shown in Fig. 4 (c) and Fig.5 (c); $\Omega_{validation}^{SL}$ by the two speed limit sign detectors to generate the performance curves shown in Fig. 4 (d) and Fig. 5 (d). $\Omega_{validation}^{stop}$ has 3685 images that contain stop signs, $\Omega_{validation}^{yield}$ has 291 that contain yield signs, $\Omega_{validation}^{SL}$ has 710 images that contain speed limit signs.

Our analysis of feature effectiveness is based on three quantities: detection rate, false alarm rate, and number of classifiers in the Adaboost systems. Fig. 4 shows detection rate verses false alarm rates of these 4 sets of traffic sign detection systems, and Fig. 5 shows the false alarm rates verses the number of the classifiers contained in each system. All the curves are plotted based on the performances of the classifiers at stage 8 through stage 17 in cascaded Adaboost systems, with the dots at the farthest to the right representing the performances at stage 8.

In general the classification systems trained by using the combined symmetric and asymmetric features have better performances than the corresponding systems that use only the symmetric features as more stages been added to the classifiers. When the combined symmetric and asymmetric features are used, the system trained to detect all traffic signs and the system trained to detect stop signs, are better than their corresponding systems that use the symmetric features after stage 9. For the speed limit sign detection, the system trained on the combined symmetric and asymmetric performs better than its counterpart after stage 13. For the yield sign detection, the system trained with symmetric features outperforms the system trained using combined symmetric and asymmetric features.

When comparing false alarm rates verses the number of classifiers in each system (see Fig. 5), the systems trained using the combined symmetric and asymmetric features generate less false alarms for the same number of classifiers than their counter parts in all four detection scenarios. Through the analysis of the performances of these detection systems, we conclude that different detection systems may need to use different features and different number of stages of classification in a cascaded Adaboost system for traffic sign detection.

III. A ROBUST MULTI-CLASS TRAFFIC SIGN DETECTION AND CLASSIFICATION SYSTEM

We model the traffic sign detection and classification problem as multi-class classification problem. In [21], we discussed pros and cons of various architectures of multi-class classification systems. The traffic sign detection and classification problem has its own unique constraints.

- The images are acquired by cameras mounted on a moving vehicle and need to be processed for online detection. Therefore the computational time for classification should be minimized.
- Traffic signs usually occupy small portions of an image, background occupies the rest. This implies that many of the patches examined by the classification system are backgrounds.
- Due to the nature of application, false alarms, which are considered annoying by drivers, should be minimized, or eliminated if possible.

Based on the above design criteria, we propose the traffic sign detection and classification system illustrated in Fig. 6. The system consists of a sequence of K classifiers that are individually trained with different features and data. The classifiers are designed and trained based on the

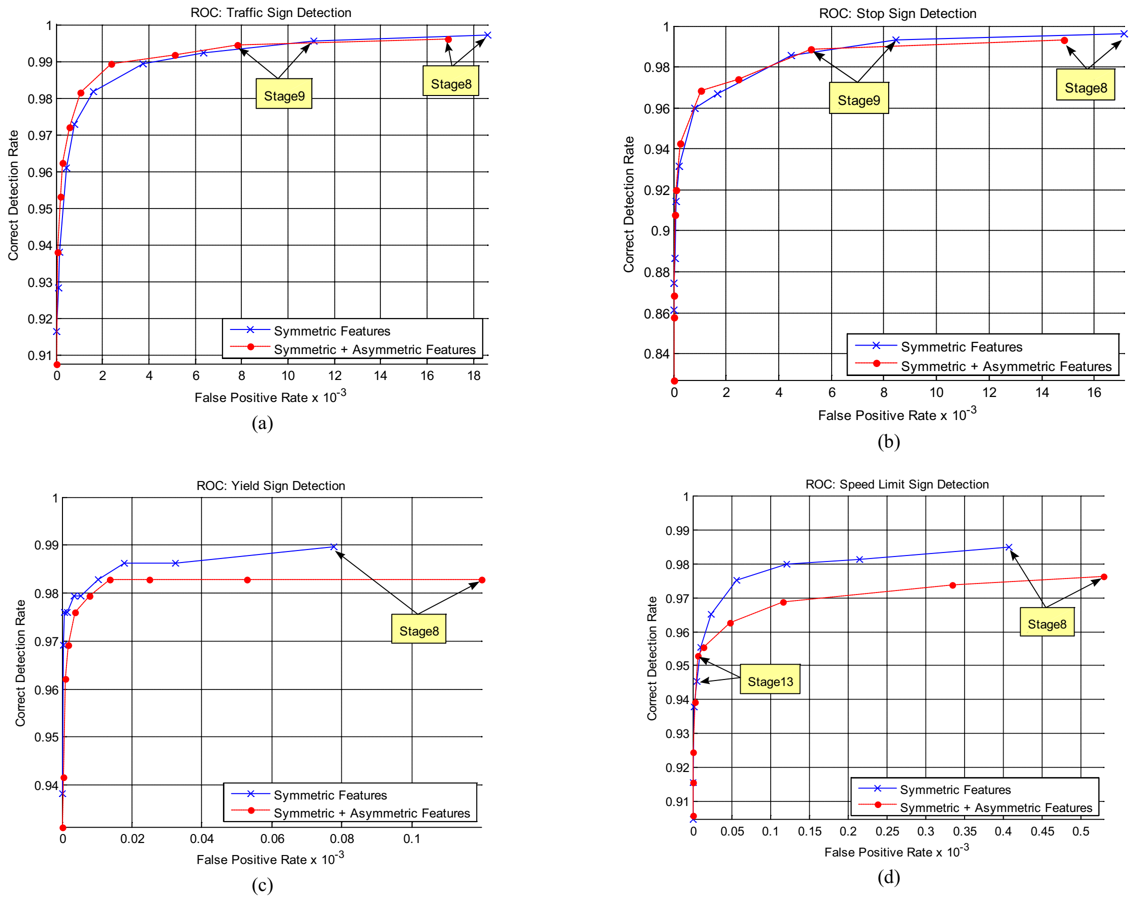


Figure 4. Detection rate verses false alarm rate for traffic sign detectors

performance curves shown in Fig. 4, which are generated from validation data mentioned earlier.

The Traffic Sign Detector (TSD) is trained to discriminate background image patches from patches containing traffic signs including stop signs, yield signs and speed limit signs. In our application an input image has a resolution of 752x480. In the Adaboost classification process, we use a window size of 16x16 and a scale factor of 1.01 to generate image patches for detection. This typically generates a ratio of 1:20,000 with respect to patches containing signs versus no signs. Therefore it is computationally advantageous to filter out as many no-sign patches as possible at the very first step. Only those that are considered as possible sign patches are passed to the subsequent detectors in the system. Since all the images rejected by TSD will not be processed again, it is important to minimize false negative at this step. We use the combined features to train the Traffic Signs Detector for 9 stages by using the Adaboost learning since Fig. 4 (a) shows that the classifiers trained using the combined symmetric and asymmetric features give superior performances to the classifiers using only symmetric features and the classifier at stage 9 gave above 99% detection rate. There are considerable false alarms being generated by this classifier as suggested by Fig. 4 (a), but they can be rejected by the subsequent classifiers in the system shown in Fig. 6.

The Stop Sign Detector is the first process that takes the image patches of possible traffic signs as input. It accepts the images that contain stop sign and passes the rest to the Yield Sign Detector. The Stop Sign Detector was designed based on the analysis of the performance curves shown in Fig. 4 (b). We noticed that the classifiers using the combined symmetric and asymmetric features performed consistently better after stage 9 than the classifiers that use only the symmetric features. Since false alarms generated by the Stop Sign Detector can cause negative impacts, we need to choose a stage that gives not only high detection rate but also low alarm rate. Therefore we use the combined features in the Stop Sign Detector and train the classifier up to stage 12,

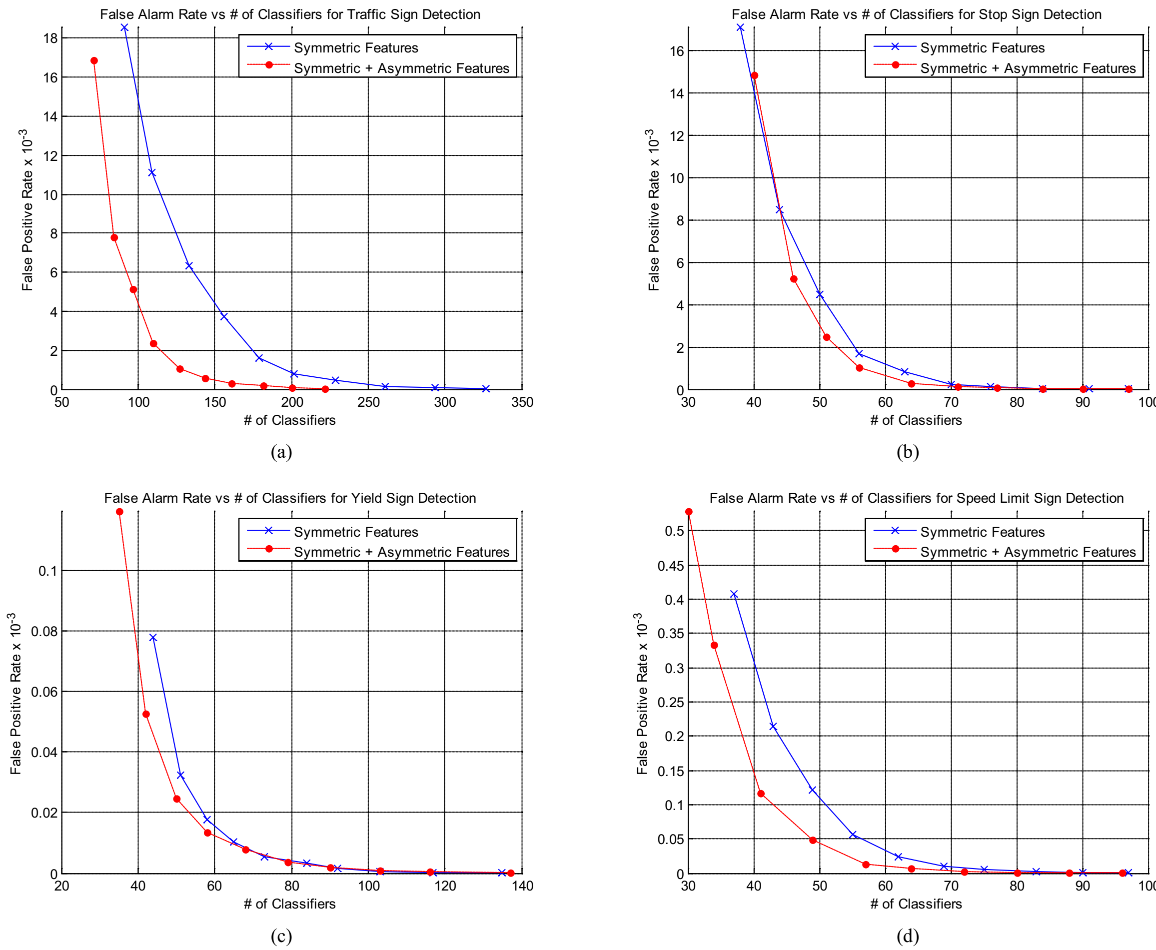


Figure 5. False alarm rates verses number of classifiers

which gives more than 94% detection rate with a false alarm rate about 0.0002% according to Fig. 4 (b).

We designed the Yield Sign Detector by analyzing the performance curves shown in Fig. 4 (c). For the yield sign detection, the classifiers trained with the symmetric features only gave better performances at all stages than the classifiers using the combined features. Therefore, the Yield Sign Detector is trained using the symmetric features and 16 stages of Adaboost learning, which gives a performance of 97% detection rate with almost 0% false alarm rate.

For the detection of speed limit signs, performance curves in Fig. 4 (d) show that after stage 13, the classifiers using the combined features are generating less false alarms than the classifiers trained with the symmetric features only. Therefore, the Speed Limit Detector is trained using the combined features and the Adaboost learning with 13 stages. Although, for the illustration purpose, we show only three specific traffic signs, the design methodology shown above can be applied to many more traffic signs, and the sequential system architecture can be scaled to detecting more traffic signs.

One more important characteristic of the proposed sequential system architecture is that, an input image is only scanned once to generate various image patches by the Traffic Signs Detector. For an image of 752x480, we will need to generate up to 200,000 image patches for traffic sign detection. Most of these image patches are processed only by the Traffic Sign Detectors, and only a small number of image patches are processed by the subsequent detectors. Therefore the proposed system architecture is far more computational efficient than a parallel system architecture, which sends the input image to all detectors [21].

We tested the multi-class traffic sign detection and classification system on 600 images, each has a resolution of

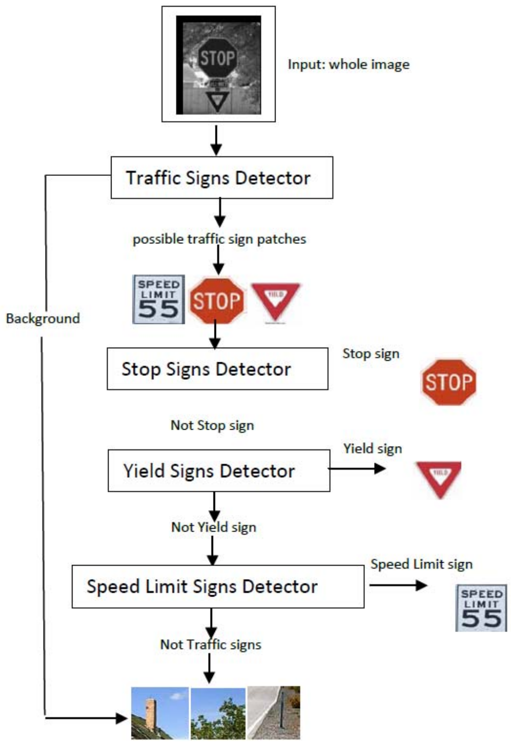


Figure 6. A robust system for multi-class traffic sign detection and classification.

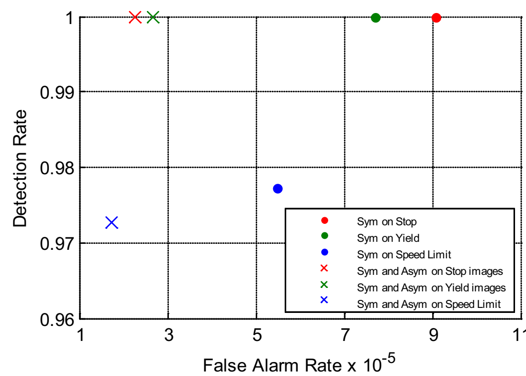


Figure 7. Performances of two traffic sign detection and classification systems using two different sets of features

752x480. Out of these images, 200 contain stop signs, 200 yield signs and 200 speed limit signs. These 600 images are not contained in the training set or in the validation set. All images are captured by a gray scale camera mounted on a moving vehicle in urban streets. The performance of the system on these test images are shown in Fig. 7 with cross (x) symbols. The proposed system's detection rate for the stop signs in these images is 100% with a false alarm rate of 2.28×10^{-5}, for the yield signs the detection rate is 100% with a false alarm rate of 2.68×10^{-5}, for the speed limit signs the detection rate is 97.27% with a false alarm rate of 1.75×10^{-5}.

In order to evaluate the effectiveness of the combined features asymmetric and symmetric features, we implemented the same system but using the symmetric features in every detector and then train the detectors using the same number of stages as discussed earlier. The performance of this system is shown in Fig. 7 with circle symbols. Its detection rates are close to the system that uses the combined features, but its false alarms rates are much higher: 9.09×10^{-5} for the stop signs detection, 7.71×10^{-5} for the yield signs, and 5.49×10^{-5} for the speed limit signs. This shows that the asymmetric features help to reduce the false alarm rates, which is important this application.

IV. Conclusion

We have presented a robust system for multi-class traffic sign detection and classification system based on Adaboost learning. We also proposed a set of asymmetric Haar-like feature templates to be combined with the symmetric features proposed by Viola and Jones for traffic sign detection. Our experimental results show that the proposed system gives high recognition rates with few false alarms, and the asymmetric features combined with the symmetric features give the traffic sign detection system with considerably reduced false alarm rates without decreasing the detection rate.

Acknowledgment

This work was supported in part by a research grant from TRW Education Foundation.

References

[1] C. Bahlmann, Y. Zhu, V. Ramesh, M. Pellkofer, and T. Koehler, A system for traffic sign detection, tracking, and recognition using color, shape, and motion information, in Intelligent Vehicles Symposium, 2005. Proceedings. IEEE, 2005, pp. 255-260.

[2] Escalera A, Moreno L.E., Salichs M.A., and Armingol J. M., Road Traffic Sign Detection and classification, 1997. IEEE Transactions on Industrial Electronics, Vol. 44, No.6, pp.848-859.

[3] D. Gavrila, Traffic sign recognition revisited, in Proc. 21st DAGM Symp. Mustererkenn., 1999, pp. 86–93.

[4] D. Ghica, S.W. Lu, X. Yuan, Recognition of traffic signs by artificial neural network, Proc. IEEE Int. Conf. on Neural Network, Perth, Australia, vol 3, Volume: 3, pages 1444-1449, Nov, 1995.

[5] G. Loy and N. Barnes, Fast shape-based road sign detection for a driver assistance system, in Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on, vol. 1, 2004, pp. 70-75 vol.1.

[6] Estevez, L. and Kehtarnavaz, N. A., Real-Time Histographic Approach to Road Sign Recognition. Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation, 95-100. 1996. San Antonio, TX.

[7] Mohamed Benallal, Jean Meunier, Real-Time Color Segmentation of Road Signs, IEEE Canadian Conference on Electrical and Computer Engineering, 2003, Volume 3, Issue , 4-7 May 2003 Page(s): 1823 - 1826 vol.3.

[8] S Vitabile, G Pollaccia, G Pilato, E Sorbello,Road signs recognition using a dynamic pixel aggregation technique in the HSV color space, Image Analysis and Processing, 2001. Proceedings. 11th International Conference on (2001), pp. 572-577.

[9] Wen-Yen Wu, Tsung-Cheng Hsieh, and Ching-Sung Lai , Extracting Road Signs using the Color Information. International Journal of Computer Science and Engineering, Volumn 2, No. 1, 2007.

[10] Y. B. Lauziere, D. Gingras, and F. P. Ferrie, A model-based road sign identification system, in Proceedings of the 2001 IEEE Computer Society Conference, vol.1, Hawaii, USA, 2001, pp.1163-1170.

[11] H. Ohara, I. Nishikawa, S. Miki, and N. Yabuki, Detection and recognition of road signs using simple layered neural networks,in Proceedings of the 9th International Conference on Neural Information Processing, vol.2, Singapore, 2002, pp.626-630.

[12] Paclik, P.; Novovicova, J.; Duin, R.P.W. Building Road-Sign Classifiers Using a Trainable Similarity Measure, Intelligent Transportation Systems, IEEE Transactions on Volume 7, Issue 3, Sept. 2006 Page(s): 309 – 321.

[13] P. Paclík, J. Novoviˇcová, P. Pudil, and P. Somol, Road sign classification using laplace kernel classifier, Pattern Recognition Letter., vol. 21, no. 13/14,pp. 1165–1173, 2000.

[14] Shaposhnikov D.G., Podladchikova L.N., Golovan A.V., Shevtsova N.A.,Hong K., Gao X. Road sign recognition by single positioning of space-variant sensor window. Proceeding of the 15-th International Conference on Vision Interface, Canada, Calgary, 2002, 213-217.

[15] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Computer Vision and Pattern Recognition*, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, Vol. 1 (2001), pp. I-511-I-518 vol.1.

[16] P. Viola and M. Jones, "Robust real-time object detection," *International Journal of Computer Vision*, 57(2): 137-154, 2004.

[17] C. Demirkir, B. Sankur, "Face Detection Using Look-Up Table Based Gentle AdaBoost," *Audio- and Video-Based Biometric Person Authentication, 5th International Conference*, AVBPA 2005, Hilton Rye Town, NY, USA, July 20-22, 2005, Proceedings. Lecture Notes in Computer Science 3546 Springer 2005, ISBN 3-540-27887-7: page 339-345.

[18] B. Wu, H. Ai, C. Huang, and S. Lao, "Fast rotation invariant multi-view face detection based on real adaboost," in *Automatic Face and Gesture Recognition*, 2004. Proceedings. Sixth IEEE International Conference on, 2004, pp. 79-84.

[19] P. Viola, M. Jones, and D. Snow, "Detecting Pedestrians Using Patterns of Motion and Appearance," *International Journal of Computer Vision* 63(2): 153-161 (2005).

[20] P. Negri, X. Clady, S. M. Hanif and L. Prevost, "A Cascade of Boosted Generative and Discriminative Classifiers for Vehicle Detection," *EURASIP Journal on Advances in Signal Processing*, Volume 2008.

[21] G. Ou and Y. L. Murphey, "Multi-class Pattern Classification Using Neural Networks," *Journal of Pattern Recognition*, Vol. 40, Issue 1, January 2007, Pages 4-18.