Real-Time Detection and Recognition of Road Traffic Signs

Jack Greenhalgh and Majid Mirmehdi, *Senior Member, IEEE*

***Abstract*—This paper proposes a novel system for the automatic detection and recognition of traffic signs. The proposed system detects candidate regions as maximally stable extremal regions (MSERs), which offers robustness to variations in lighting conditions. Recognition is based on a cascade of support vector machine (SVM) classifiers that were trained using histogram of oriented gradient (HOG) features. The training data are generated from synthetic template images that are freely available from an online database; thus, real footage road signs are not required as training data. The proposed system is accurate at high vehicle speeds, operates under a range of weather conditions, runs at an average speed of 20 frames per second, and recognizes all classes of ideogram-based (nontext) traffic symbols from an online road sign database. Comprehensive comparative results to illustrate the performance of the system are presented.**

***Index Terms*—Histogram of oriented gradient (HOG) features, maximally stable extremal regions (MSERs), support vector machines (SVMs), synthetic data, traffic sign recognition.**

I. Introduction

AUTOMATIC traffic sign detection and recognition is an important part of an advanced driver assistance system. Traffic symbols have several distinguishing features that may be used for their detection and identification. They are designed in specific colors and shapes, with the text or symbol in high contrast to the background. Because traffic signs are generally oriented upright and facing the camera, the amount of rotational and geometric distortion is limited.

Information about traffic symbols, such as shape and color, can be used to place traffic symbols into specific groups; however, there are several factors that can hinder effective detection and recognition of traffic signs. These factors include variations in perspective, variations in illumination (including variations that are caused by changing light levels, twilight, fog, and shadowing), occlusion of signs, motion blur, and weather-worn deterioration of signs. Road scenes are also generally very cluttered and contain many strong geometric shapes that could easily be misclassified as road signs. Accuracy is a key consideration, because even one misclassified or undetected sign could have an adverse impact on the driver.

Manuscript received January 13, 2012; revised May 2, 2012; accepted July 2, 2012. Date of publication August 27, 2012; date of current version November 27, 2012. This work was supported in part by the Engineering and Physical Sciences Research Council and Jaguar Cars Limited. The Associate Editor for this paper was J. Stallkamp.

The authors are with the Visual Information Laboratory, University of Bristol, BS8 1UB Bristol, U.K. (e-mail: greenhal@cs.bris.ac.uk; majid@compsci.bristol.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2012.2208909

The proposed method consists of the following two stages: 1) detection is performed using a novel application of maximally stable extremal regions (MSERs) [1], and 2) recognition is performed with histogram of oriented gradient (HOG) features, which are classified using a linear support vector machine (SVM).

Another novel aspect of this paper is the use of an online road sign database [2] that consists of synthetic graphical representations of signs. To the best of our knowledge, this is the first paper that uses the entire range of road signs in operation. Previous works, such as [3]–[5], all use hand-selected subsets. Large training sets are then generated by applying random distortions to our graphic templates, e.g., geometric distortion, blurring, and illumination variations, to capture examples of occurrences of real-scene distortions. It is essential for the classifiers to be trained on all possible signs to avoid misclassification of similar but excluded signs. Generating synthetic data in this way allows classification to be performed on all possible road signs and also avoids the tedious process of hand labeling large data sets. Examples of the proposed road sign detection system are shown in Fig. 1.

In Section II, we review previous work and state the improvements that we make. Then, in Section III, we outline the methodology used, which includes detection, recognition, and the generation of synthetic data. In Section IV, we describe comparative results to illustrate the performance of the system. Finally, conclusions are drawn in Section V.

II. Related Work

A significant number of papers that deal with the recognition of ideogram-based road signs in real road scenes have been published [4]–[15].

The most common approach, quite sensibly, consists of two main stages: detection and recognition. The detection stage identifies the regions of interest and is mostly performed using color segmentation, followed by some form of shape recognition. Detected candidates are then either identified or rejected during the recognition stage using, for example, template matching [16] or some form of classifier such as SVMs [4], [5], [11] or neural networks [3], [17].

The majority of systems make use of color information as a method for segmenting the image [9], [11], [12], [18]–[20]. The performance of color-based road sign detection is often reduced in scenes with strong illumination, poor lighting, or adverse weather conditions such as fog. Color models, such as hue–saturation–value (HSV) [8], [12], [19], YUV [21], and CIECAM97 [10], have been used in an attempt to overcome

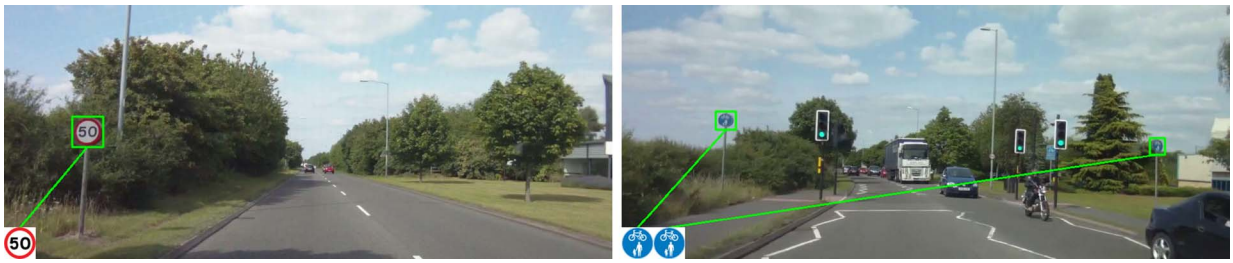

Fig. 1. Examples of the proposed road sign detection and recognition system.

these issues. For example, Shadeed *et al.* [21] performed segmentation by applying the U and V chrominance channels of the YUV space, with U being positive and V being negative for red colors. This information was used in combination with the hue channel of the HSV color space to segment red road signs. Gao *et al.* [10] applied a quad-tree histogram method to segment the image based on the hue and chroma values of the CIECAM97 color model. Malik *et al.* [12] thresholded the hue channel of the HSV color space to segment red road signs.

In contrast, there are several approaches [7], [22] that entirely ignore color information and instead use only shape information from grayscale images. For example, Loy and Zelinksy [22] proposed a system that used local radial symmetry to highlight points of interest in each image and detect octagonal, square, and triangular road signs.

Some recent methods such as [23] and [24] use HOG features for road sign feature extraction. Creusen *et al.* [23] extended the HOG algorithm to incorporate color information using the CIELAB and YCbCr color spaces. Overett *et al.* [24] presented two variant formulations of HOG features for the detection of speed signs in New Zealand. We also use HOG features to aid our classification process and will explain later why we find they are most suited to this application.

The vast majority of existing systems consist of classifiers that were trained using hand-labeled real images, for example [3]–[5], which is a repetitive, time-consuming, and error-prone process. Our method avoids collecting and manually labeling training data, because it requires only synthetic graphical representations of signs that were obtained from an online road sign database [2]. Furthermore, although many existing systems report high classification rates, the total number of traffic sign classes recognized is generally very limited, e.g., seven classes in [4], 42 classes in [15], or 20 classes in [16], and are hence less likely to suffer mismatches against similar signs that were missing from their databases. Our proposed system uses all instances of ideogram-based traffic symbols used in the U.K. and hence performs its matching in this larger set. We expect our approach to be equally functional if applied to other countries' traffic sign databases obtained in a similar fashion.

Note that many proposed systems suffer from slow speed, making them inappropriate for application to real-time problems. Some methods report processing times of several seconds for a single frame [4], [11], [25], [26]. Our system runs at an average speed of 20 frames per second.

There are a few commercial traffic sign recognition systems on the market, including [27] and [28]. Such commercial systems also recognize a very limited set of traffic signs; for example, the system that was developed by Mobileye [28] detects only speed limit signs and no-overtaking signs. Comparison with these commercial systems is difficult, given that little information on their performance is available.

In Section IV, we first compare our proposed method with a similar road sign detection system that was proposed by Gómez-Moreno *et al.* [11]. We then evaluate the performance of our synthetically generated training data against real training data on the German Traffic Sign Recognition Benchmark (GTSRB) [29].

III. Traffic Sign Detection and Recognition System

A. Overview of the System

The proposed system consists of the following two main stages: detection and recognition. The complete set of road signs used in our training data and recognized by the system is shown in Fig. 2. Candidates for traffic symbols are detected as MSERs, as described by Matas *et al.* [1]. MSERs are regions that maintain their shape when the image is thresholded at several levels. This method of detection was selected due to its robustness to variations in contrast and lighting conditions. Rather than detecting candidates for road signs by border color, the algorithm detects candidates based on the background color of the sign, because these backgrounds persist within the MSER process. Our proposed method, as described in detail in the following section, is broadly illustrated in Fig. 3.

B. Detection of Road Signs as MSERs

For the detection of traffic symbols with white background, MSERs are found for a grayscale image. Each frame is binarized at a number of different threshold levels, and the connected components at each level are found. The connected components that maintain their shape through several threshold levels are selected as MSERs. Fig. 4 shows different thresholds for an example image with the connected components colored. It is shown that the connected component that represents the circular road symbol maintains its shape through several threshold levels. This helps ensure robustness to both variations in lighting and contrast.

Several features of the detected connected component regions are used to further reduce the number of candidates. These features are width, height, aspect ratio, region perimeter and area, and bounding-box perimeter and area. Removing the connected components that do not match the requirements helps speed up the process and improve accuracy. The parameters

Fig. 2. Full set of graphical road signs used in training the proposed system [2].

used as limits for these features were empirically determined and are shown in Table I.

We approach the detection of traffic symbols with red or blue backgrounds in a slightly different manner. Rather than detecting MSERs for a grayscale image, the frame is first transformed from red–green–blue (RGB) into a "normalized red/blue" image Ω_{RB} such that, for each pixel of the original image, values are found for the ratio of the blue channel to the sum of all channels and the ratio of the red channel to the sum of all channels. The greater of these two values is used as the pixel value of the normalized red/blue image, i.e.,

$$\Omega_{RB} = \max\left(\frac{R}{R+G+B}, \frac{B}{R+G+B}\right). \quad (1)$$

Pixel values in this image are higher for red and blue pixels and lower for other colors. MSER regions are then found for this new image. Fig. 5 shows an example image and the result of the normalized red/blue transform. Fig. 6 shows the connected components at several different thresholds of the transformed image. Again, it is shown that the red and blue road signs maintain their shape at several threshold levels, making them candidates for classification.

Although MSER offers a robust form of detection for traffic signs in complex scenes, it can be computationally expensive. Therefore, to increase the speed, we threshold only at an appropriate range of values rather than at every possible value, which is the norm in the original MSER [1]. Fig. 7 shows the number of used thresholds plotted against the processing time and accuracy of detection. The thresholds were evenly spaced between the values 70 and 190, because the MSERs that represent road signs generally appear within this range. The number of thresholds selected was 24, which, in this example, corresponds to 94.3% accuracy and 50.1-ms processing time.

C. Road Sign Classification

The recognition stage is used to confirm a candidate region as a traffic sign and classify the exact type of sign. For the classification of candidate regions, their HOG features are extracted from the image [30], which represent the occurrence of gradient orientations in the image. HOG feature vectors are calculated for each candidate region. A Sobel filter is used to find the horizontal and vertical derivatives and, hence, the magnitude and orientation for each pixel. We find the application of HOG to recognition of traffic symbols very suitable, given that traffic symbols are composed of strong geometric shapes and high-contrast edges that encompass a range of orientations. Traffic signs are generally found to be approximately upright and facing the camera, which limits rotational and geometric distortion, removing the need for rotation invariance.

The HOG features are computed on a dense grid of cells using local contrast normalization on overlapping blocks. A nine-bin histogram of unsigned pixel orientations weighted by magnitude is created for each cell. These histograms are normalized over each overlapping block. The components of the feature vector are the values from the histogram of each

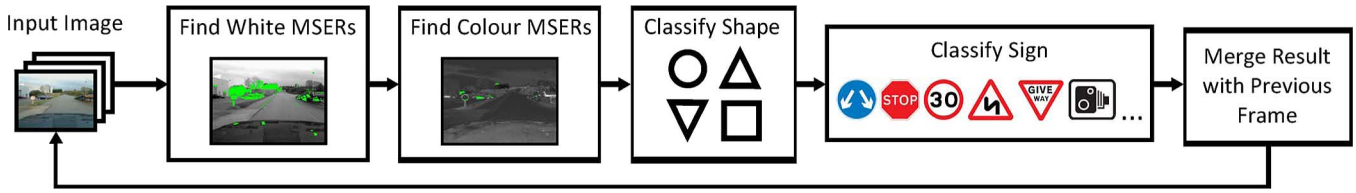

Fig. 3. Overview schematic of the proposed approach.

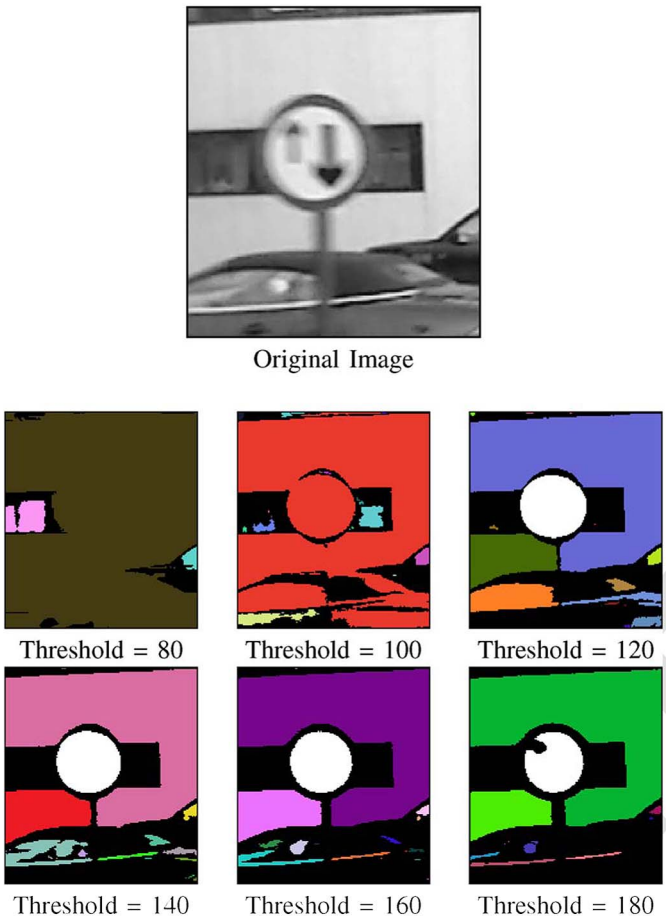

Fig. 4. (Top) Original image. (Middle and Bottom) Connected components at several threshold levels.

TABLE I
PROPERTIES USED TO SORT CONNECTED COMPONENTS

Feature	Minimum	Maximum
Width	14	100
Height	14	110
Height/Width	0.5	1.5
MSER perimeter/bounding box perimeter	0.3	1.2
MSER area/bounding box area	0.4	1

normalized cell. Fig. 8, shows an example image divided into cells and an example HOG block that consists of four cells. Although this intensive normalization produces large feature vectors (1764 dimensions for a 64×64 image), it provides high accuracy. The size N of the HOG feature vector is computed using

$$N = \left(\frac{R_{width}}{M_{width}} - 1\right) \times \left(\frac{R_{height}}{M_{height}} - 1\right) \times B \times H \qquad (2)$$

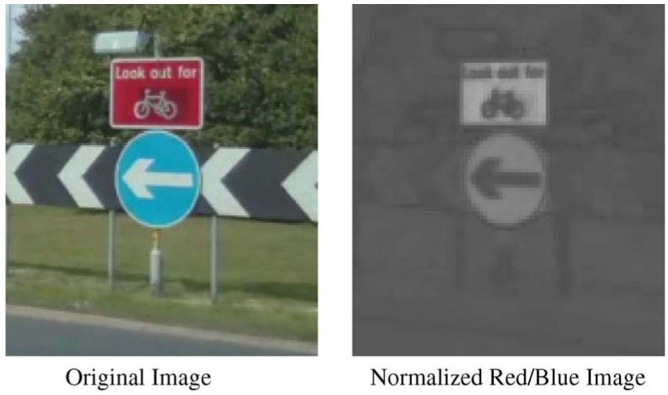

Fig. 5. Image transformed into our normalized red/blue color space Ω_{RB}.

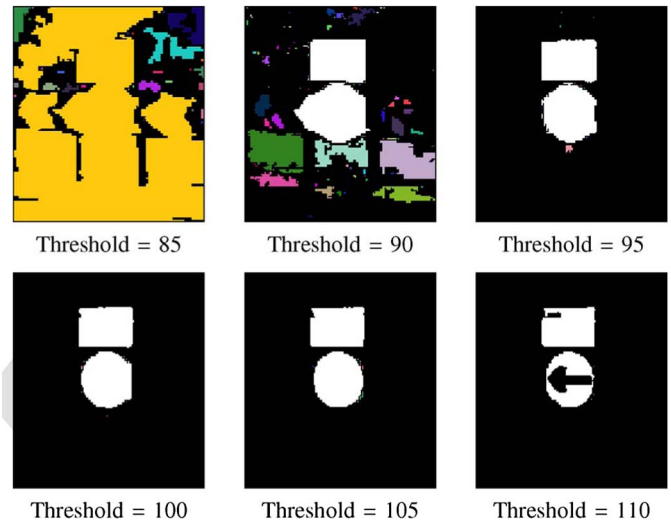

Fig. 6. Connected components at several thresholds of the normalized red–blue image.

where R is the region, M is the cell size, B is the number of cells per block, and H is the number of histograms per cell. The values used were $M = 8 \times 8$, $B = 4$, and $H = 9$.

Regions are then classified using a cascade of multiclass SVMs [31]. SVM is a supervised learning method that constructs a hyperplane to separate data into classes. The "support vectors" are data points that define the maximum margin of the hyperplane. Although SVM is primarily a binary classifier, multiclass classification can be achieved by training many one-against-one binary SVMs. SVM classification is fast, highly accurate, and less prone to overfitting compared to many other classification methods. It is also possible to very quickly train an SVM classifier, which significantly helps in our proposed method, given our large amount of training data and high

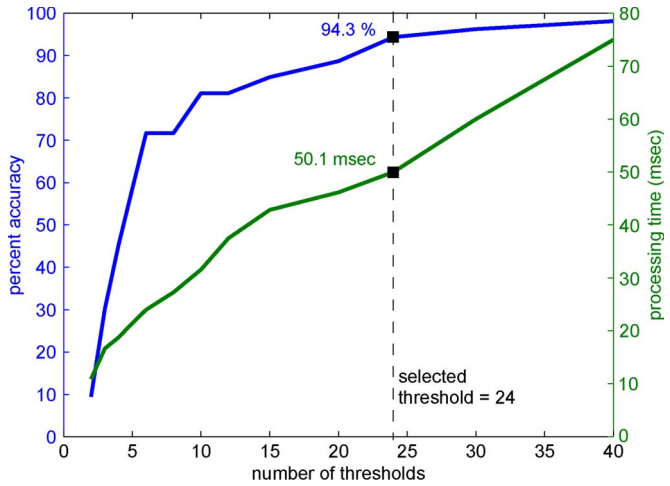

Fig. 7. Chart that shows the number of thresholds used for MSER plotted against accuracy of detection and processing time.

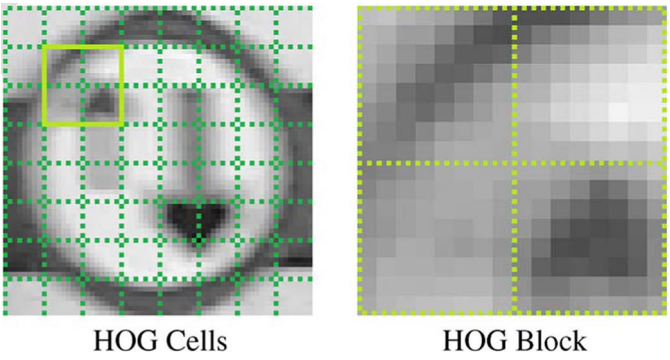

Fig. 8. Regions of HOG features.

number of classes. However, we plan to perform further comparison with other classification methods in future work.

Each region in our system is classified using a cascade of SVM classifiers, as illustrated in Fig. 9. First, the candidate region is resized to 24 × 24 pixels. A HOG feature vector with 144 dimensions is then calculated, and this feature vector is used to classify the shape of the region as a circle, triangle, upside-down triangle, rectangle, or background. Octagonal stop signs are considered to be circles. If the region is found to be background, it is rejected. If the region is found to be a shape, it is then passed on to a (symbol) subclassifier for that specific shape.

Different classifier trees are used for candidates with white background (MSERs for a grayscale image) and candidates with red or blue background (MSERs for a normalized red/blue image). Therefore, each subclassifier is specific to symbols with a certain background color and shape. Color background triangles and color background upside-down triangles are rejected as background, because no signs of these types exist in the U.K. road sign database [2].

To optimize the performance of the linear SVM classifier, an appropriate value for the cost of misclassification parameter C has to be selected. Choosing a value that is too large may result in overfitting, whereas a value that is too small may cause underfitting. Hence, a cross correlation of the training set is performed for $\log_2\ C = -5, -3, -2, \ldots, 15$, and the value of C that produces the highest cross-correlation accuracy is used.

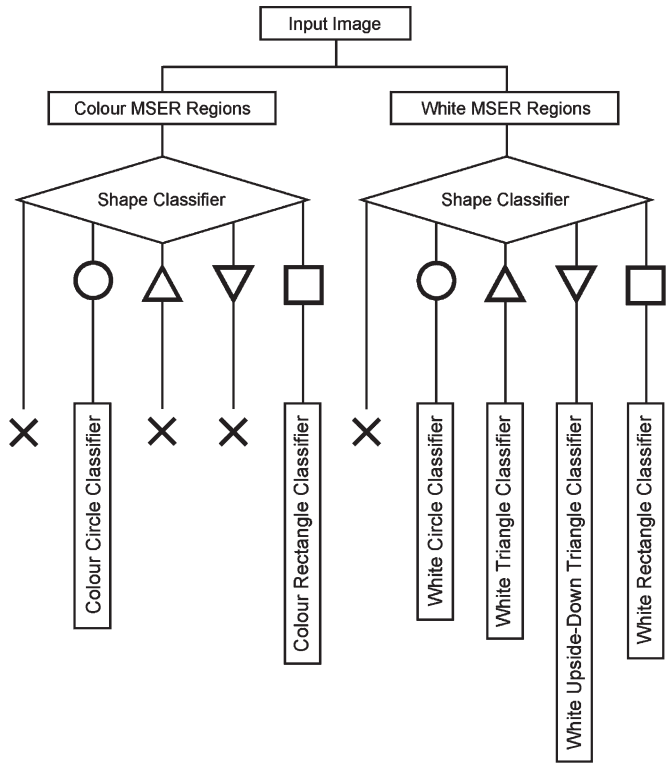

Fig. 9. Cascaded SVM classifier.

Road sign classifications from several frames are merged together to form a decision. A probabilistic SVM model is used for classification. Rather than having each classification counting as a single vote for a specific class, a vote was made for each class, weighted by its probability. The class with the highest score S was taken as the correct classification, i.e.,

$$S = \sum_{n=1}^{N} P(A_n) \tag{3}$$

where N is the total number of classifications, and A is an SVM classification. The classification is made once S exceeds the decision threshold λ.

In Section IV-B, we compare the results with and without the inclusion of this frame-merging technique.

D. Generation of Synthetic Training Data

Training the classifiers on all possible road signs is essential to avoid misclassification of unknown signs. However, gathering a sufficient amount of real data on which to train the classifiers is difficult and time consuming, given the sheer number of different existing signs and the scarcity of particular signs. This is possibly one of the reasons that most other works (and commercial systems such as Mobileye's Traffic Sign Recognition [28]) focus only on a subset of more common signs regularly found in their footage; for example, see [3]–[5]. We further suggest that using only a subset of signs also avoids misclassification against other similar but excluded signs; therefore, in many cases, the quality of the reported results can be unreliable. Our proposed solution to this problem is to use easily available graphical data and synthetically generate

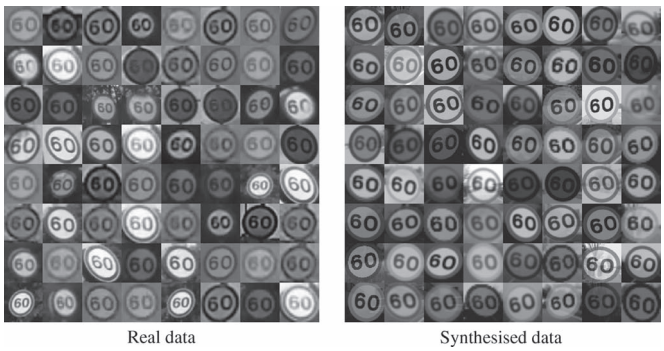

Fig. 10. Comparison of real and synthesized data.

variations and distortions of them to create training data for the classifiers. This approach allows us to use the entire range of road signs, avoid tedious manual hand labeling for training purposes, and report more reliable classification results that are a true reflection of a complete search.

The graphical base images that we use were obtained from a free online database provided by the U.K. Department for Transport [2]. Randomized geometric distortions were then applied to replicate the range of distortions likely to be seen in real data and the type of regions likely to be found during the detection stage. Each distorted example image is superimposed over a random section of background, taken from a database of typical background images. Randomized brightness, contrast, noise, blurring, and pixelation are also applied to each image.

The complete set of 131 road sign images used for training is shown in Fig. 2. For each sign, 1200 synthetic distorted images were generated. As a means of comparison, Fig. 10 shows a number of real road sign images next to a number of our generated training images.

IV. Experimental Results

The proposed system can operate at a range of vehicle speeds and was tested under a variety of lighting and weather conditions. A considerable increase in speed was gained by implementing the algorithm in parallel as a pipeline to around 20 frames per second, running on a 3.33-GHz Intel Core i5 central processing unit under OpenCV, where the frame dimensions were 640 × 480. However, the system retained a latency of around 200 ms.

We compare our proposed method (later in Section IV-B) with a road sign detection system that was proposed by Gómez-Moreno *et al.* [11], which also deals with the entire problem of detection and recognition, detects a relatively large number of road signs (encompassing a variety of different shapes and colors) compared to other methods, and uses SVMs for classification. These factors make the system in [11] a particularly suitable method for comparison purposes.

The system that was proposed by Gómez-Moreno *et al.* [11] detects candidate regions using color information and performs recognition using SVM based on a training set of between 20 and 100 images per class on an unspecified number of classes. Each frame is segmented using the hue and saturation components of a hue–saturation–intensity (HSI) image. Histograms of

Fig. 11. Example images from the test set.

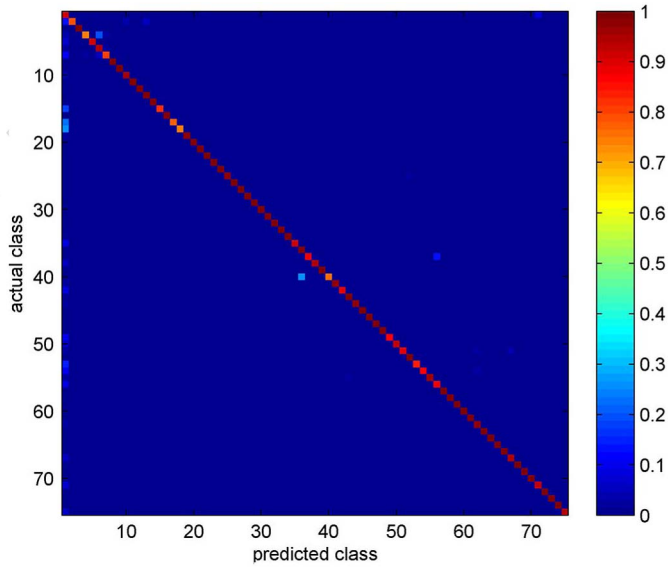

Fig. 12. Confusion matrix for a cascaded classifier with white background (accuracy = 89.2%).

hue and saturation are built for red, blue, and yellow sign colors, and created using images with a range of weather and lighting conditions. For the segmentation of white road signs, the image is binarized based on the achromaticity of each pixel, and then, each candidate blob is classified by shape. The distance from the side of the candidate blob to its bounding box is measured at each side (left, right, top, and bottom) at several points. Binary SVMs for each shape are then used to vote for each side of the blob (circle or triangle). If the blob receives four votes for a particular shape, that shape is chosen. An SVM with a Gaussian kernel is then used to classify each sign type based on shape and color. The SVM is trained using pixel values from the candidate region that falls into a template that represents the shape (circle or triangle).

A. Performance of the Proposed Method

To assess the performance of our classifiers, a test data set was collected from frames of road video footage and road sign images obtained from the Internet. This test set included many challenging images affected by geometric distortion, blurring,

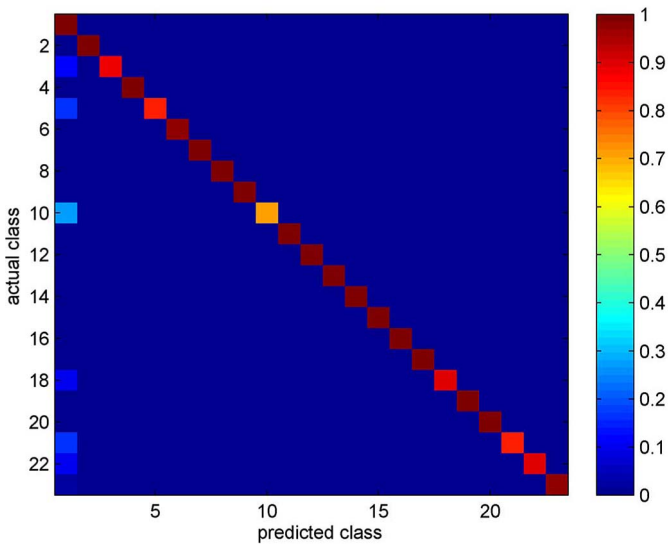

Fig. 13. Confusion matrix for a cascaded classifier with color background (accuracy = 92.1%).

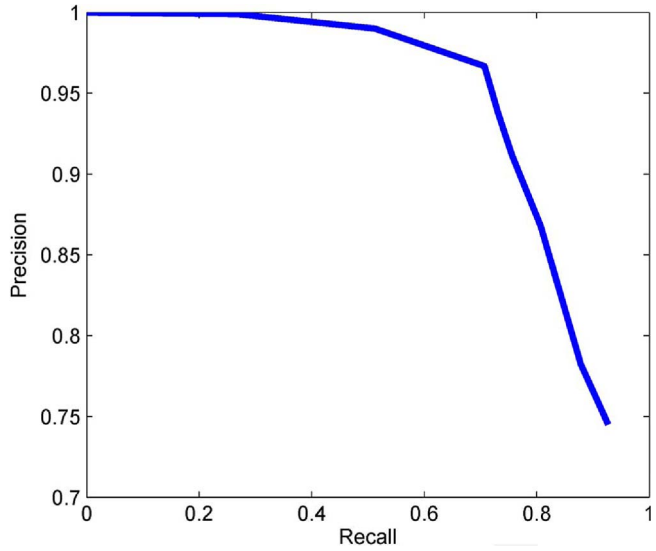

Fig. 14. Chart that shows precision against recall of system as decision threshold λ is varied.

Fig. 15. Image that shows the detection of the "Give Way" sign (taken from video 1).

Fig. 16. Image that shows the detection of the "Pedestrians in Road" sign (taken from video 2).

Fig. 17. Image that shows the detection of the speed limit sign (taken from video 3).

deterioration, and partial occlusion (see Fig. 11). The accuracy for white road signs was 89.2% and 92.1% for color signs.

The confusion matrices in Figs. 12 and 13 represent classifier results for white background and color background signs, respectively. The values on the x-axis represent the individual road sign classes, and the values on the y-axis represent the predictions made by the classifier. Column 1 of both matrices represents classification as the background (negative) class, and this is shown to have been the most common misclassification at 7.0%, with some misclassification between nonbackground classes at 3.4%. This case is preferred, because in the overall system, decisions are formed over several frames, and regions that are classified as background are simply ignored.

Fig. 14 shows the precision of the system plotted against recall as the decision threshold λ is varied. It is shown in this graph that, because λ is reduced to increase the number of detections, the precision of the system falls as the number of false positives increases.

B. Comparative Analysis

For the Gómez-Moreno *et al.* method, between 20 and 100 real training images per class were used to train the SVM classifiers for recognition, as suggested in [11]. For test data, we used several videos, filmed under a range of weather conditions, at a variety of different vehicle speeds. Video 1 was filmed in clear weather conditions, at low speeds of around 20 mi/h. Video 2 was filmed in thick fog, at high vehicle speeds, e.g., above 50 mi/h. Video 3 was filmed in clear weather conditions, at a variety of vehicle speeds, ranging from 20 to 60 mi/h. An example frame from each video, with results overlaid from our system, is shown in Figs. 15–17.

TABLE II
COMPARATIVE RESULTS FOR GÓMEZ-MORENO *ET AL.*'S SYSTEM [11] AND THE PROPOSED METHOD. THE TOTAL NUMBER OF SIGNS WAS 14 IN VIDEO 1, 5 IN VIDEO 2, AND 38 IN VIDEO 3

Method	Gómez-Moreno et al.[11]				Proposed			
	video 1	**video 2**	**video 3**	**total**	**video 1**	**video 2**	**video 3**	**total**
Signs Correctly Classified	8	1	10	19	12	4	30	46
Signs Misclassified	4	0	4	8	3	0	2	5
Background Misclassified as Sign	7	2	8	17	0	0	2	2
Signs Undetected	6	4	24	34	2	1	6	9
Total Signs in Video	14	5	38	57	14	5	38	57
Precision	**42.1%**	**33.3%**	**45.5%**	**43.2%**	**80.0%**	**100%**	**88.2%**	**86.8%**
Recall	**57.1%**	**20.0%**	**26.3%**	**33.3%**	**85.7%**	**80.0%**	**79.0%**	**80.7%**
F-Measure	**0.48**	**0.25**	**0.33**	**0.38**	**0.83**	**0.89**	**0.83**	**0.84**

The results in Table II show that our proposed method outperformed the method used in [11]. Although their detection method reasonably classified in clear weather conditions, scenes that suffer from poor lighting conditions and strong illumination caused it to fail. Our MSER detection system provides robustness by searching for candidate regions at a range of thresholds rather than using a single fixed value. The recognition method that was proposed by Gómez-Moreno *et al.* [11] also produced a large number of false positives. Our approach of using the HOG feature descriptor with SVM performed better than directly using pixel values.

Our method was also tested without the use of frame merging, which was described in Section III-C. Removing this part of the system reduced the total precision to 67.7%, which is a considerable drop from the 86.8% achieved with the use of frame merging. Although the reported results are still too low for use in practice, the performance is high, given the large number of classes recognized.

C. Performance of the Synthetically Generated Test Set

To assess the relevance of the concept of synthetically generated training data, a comparison was made between a classifier that was trained on real data from the GTSRB [29] and a classifier that was trained on synthetic training data.

A training data set of German road signs was generated from graphics-based images, with a single example for each class, resulting in a total of 43 classes. A total of 1200 synthetic images were generated for each class, and HOG features were calculated for each image. A data set of HOG features was also created from the images contained in the GTSRB training data set. A linear SVM classifier was trained for both the real data set and the generated synthetic data set. Both classifiers were then tested using the GTSRB test data set.

The classifier that was trained on the synthetic data gave an accuracy of 85.7%, and the classifier that was trained on real data gave an accuracy of 85.9%. Based on these results, it is shown that the synthetic data set produced results comparable to a data set of hand-labeled real images. Although the results for real training data were slightly higher than for the synthetic data, the use of synthetic data allowed the tedious time-consuming process of manually hand labeling a large data set of real images to be avoided.

To show that the features learned by the classifier relate only to the road signs and not to background information, the classifier was also tested using a data set that comprises synthetically generated images, but with different backgrounds from that in the training set. The accuracy achieved for this experiment was 97.6%, which verified the claim.

To more thoroughly validate the system, another classifier was trained, with a data set that contains real images and synthetically generated interpolations, created using randomly distorted version of the real images. The total number of images in this data set was 43 509. This classifier had an overall accuracy of 89.2%, which was greater than either the fully synthetic or the fully real data set.

V. CONCLUSION

We have proposed a novel real-time system for the automatic detection and recognition of traffic symbols. Candidate regions are detected as MSERs. This detection method is significantly insensitive to variations in illumination and lighting conditions. Traffic symbols are recognized using HOG features and a cascade of linear SVM classifiers. A method for the synthetic generation of training data has been proposed, which allows large data sets to be generated from template images, removing the need for hand labeled data sets. Our system can identify signs from the whole range of ideographic traffic symbols currently in use in the U.K. [2], which form the basis of our training data. The system retains a high accuracy at a variety of vehicle speeds.

ACKNOWLEDGMENT

The authors would like to thank Dr. A. Dunoyer and Dr. T. Popham of Jaguar Research for their support.

REFERENCES

[1] J. Matas, "Robust wide-baseline stereo from maximally stable extremal regions," *Image Vis. Comput.*, vol. 22, no. 10, pp. 761–767, Sep. 2004.
[2] Dept. Transp., London, U.K., Traffic Signs Image Database, 2011.
[3] M. Hossain, M. Hasan, M. Ali, M. Kabir, and A. Ali, "Automatic detection and recognition of traffic signs," in *Proc. RAM*, 2010, pp. 286–291.
[4] K. Ohgushi and N. Hamada, "Traffic sign recognition by bags of features," in *Proc. TENCON*, 2009, pp. 1–6.
[5] H. Pazhoumand-Dar and M. Yaghobi, "DTBSVMs: A new approach for road sign recognition," in *Proc. ICCICSN*, Jul. 2010, pp. 314–319.
[6] D. Kang, N. Griswold, and N. Kehtarnavaz, "An invariant traffic sign recognition system based on sequential color processing and geometrical transformation," in *Proc. IAI*, 1994, pp. 88–93.
[7] P. Paclik and J. Novovicová, "Road sign classification without color information," in *Proc. ASIC*, Lommel, Belgium, 2000.

[8] S. Vitabile, G. Pollaccia, G. Pilato, and E. Sorbello, "Road signs recognition using a dynamic pixel aggregation technique in the HSV color space," in *Proc. IAP*, 2001, pp. 572–577.
[9] 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 *Proc. IVS*, 2005, pp. 255–260.
[10] X. Gao, L. Podladchikova, D. Shaposhnikov, K. Hong, and N. Shevtsova, "Recognition of traffic signs based on their color and shape features extracted using human vision models," *J. Vis. Commun. Imag. Represent.*, vol. 17, no. 4, pp. 675–685, Aug. 2006.
[11] S. Maldonado-Bascón, S. Lafuente-Arroyo, P. Gil-Jimenez, H. Gomez-Moreno, and F. Lopez-Ferreras, "Road-sign detection and recognition based on support vector machines," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 264–278, Jun. 2007.
[12] R. Malik, J. Khurshid, and S. Ahmad, "Road sign detection and recognition using color segmentation, shape analysis and template matching," in *Proc. ICMLC*, Aug. 2007, vol. 6, pp. 3556–3560.
[13] M. Rahman, F. Mousumi, E. Scavino, A. Hussain, and H. Basri, "Real-time road sign recognition system using artificial neural networks for Bengali textual information box," in *Proc. ITSim*, 2008, vol. 2, pp. 1–8.
[14] Y. Wang, M. Shi, and T. Wu, "A method of fast and robust for traffic sign recognition," in *Proc. ICIG*, 2009, pp. 891–895.
[15] A. Ruta, Y. Li, and X. Liu, "Real-time traffic sign recognition from video by class-specific discriminative features," *Pattern Recognit.*, vol. 43, no. 1, pp. 416–430, Jan. 2010.
[16] Y. Gu, T. Yendo, M. Tehrani, T. Fujii, and M. Tanimoto, "A new vision system for traffic sign recognition," in *Proc. IV*, 2010, pp. 7–12.
[17] E. Cardarelli, P. Medici, P. P. Porta, and G. Ghisio, "Road sign shapes detection based on Sobel phase analysis," in *Proc. IEEE IVS*, 2009, pp. 376–381.
[18] H. Huang, C. Chen, Y. Jia, and S. Tang, "Automatic detection and recognition of road sign," in *Proc. ICMESA*, 2008, pp. 626–630.
[19] P. Wanitchai and S. Phiphobmongkol, "Traffic warning signs detection and recognition based on fuzzy logic and chain code analysis," in *Proc. IITA*, 2008, pp. 508–512.
[20] C. G. Kiran, L. V. Prabhu, R. V. Abdu, and K. Rajeev, "Traffic sign detection and pattern recognition using support vector machine," in *Proc. ICAPR*, 2009, pp. 87–90.
[21] W. Shadeed, D. Abu-Al-Nadi, and M. Mismar, "Road traffic sign detection in color images," in *Proc. ICECS*, 2003, vol. 2, pp. 890–893.
[22] G. Loy, "Fast shape-based road sign detection for a driver assistance system," in *Proc. IROS*, 2004, pp. 70–75.
[23] I. Creusen, R. Wijnhoven, and E. Herbschleb, "Color exploitation in HOG-based traffic sign detection," in *Proc. ICIP*, 2010, pp. 2669–2672.
[24] G. Overett, "Large-scale sign detection using HOG feature variants," in *Proc. IV*, 2011, pp. 326–331.
[25] A. Reiterer, "Automated traffic sign detection for modern driver assistance systems," in *Proc. FCBC*, Sydney, Australia, 2010, pp. 11–16.
[26] J. F. Khan, S. M. A. Bhuiyan, and R. R. Adhami, "Image segmentation and shape analysis for road-sign detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 1, pp. 83–96, Mar. 2011.
[27] Volkswagen Media Services, (2012, Jan. 9), Phaeton debuts with new design and new technologies. [Online]. Available: https://www.volkswagen-media-services.com/medias_publish/ms/content/en/pressemitteilungen/2010/04/22/phaeton_debuts_with.standard.gid-oeffentlichkeit.html
[28] Mobileye, (Oct. 26, 2011). Traffic Sign Detection, [Online]. Available: http://mobileye.com/technology/applications/traffic-sign-detection/
[29] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "German traffic sign recognition benchmark: A multiclass classification competition," in *Proc. IJCNN*, 2011, pp. 1453–1460.
[30] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. CVPR*, 2005, pp. 886–893.
[31] C. Cortes and V. Vapnik, "Support vector networks," *J. Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.

Jack Greenhalgh received the M.Eng. degree in computer systems engineering from the University of Sussex, Brighton, U.K., in 2010. He is currently working toward the Ph.D. degree, with a focus on driver assistance using automated sign and text recognition, in the Visual Information Laboratory, University of Bristol, Bristol, U.K.

His research interests include image processing, computer vision, machine learning, and intelligent transportation systems.

Majid Mirmehdi (SM'10) received the B.Sc. (Hons.) and Ph.D. degrees in computer science from the City University, London, U.K., in 1985 and 1991, respectively.

He is currently a Professor of computer vision with the Visual Information Laboratory, University of Bristol, Bristol, U.K., where he is also the Graduate Dean and the Head of the Graduate School of Engineering. He is the Editor-in-Chief of the *IET Computer Vision Journal* and an Associate Editor for *Pattern Analysis and Applications*. His research interests include natural-scene analysis and medical imaging, and he has published more than 150 refereed journal publications and conference proceedings in these and other areas.

Dr. Mirmehdi is a Fellow of the International Association for Pattern Recognition (IAPR) and a Member of the Institution of Engineering and Technology (IET). He serves on the Executive Committee of the British Machine Vision Association.