Attention-Based Traffic Sign Recognition with an Array of Weak Classifiers

Robert Kastner, Thomas Michalke, Thomas Burbach, Jannik Fritsch, Christian Goerick

2010

Preprint:

This is an accepted article published in IEEE Intelligent Vehicles Symposium (IV). The final authenticated version is available online at: https://doi.org/[DOI not available]

Attention-based traffic sign recognition with an array of weak classifiers

Robert Kastner*, Thomas Michalke[1], Thomas Burbach*, Jannik Fritsch◇, Christian Goerick◇

*Darmstadt University of Technology
Institute for Automatic Control
D-64283 Darmstadt, Germany
{robert.kastner}
@rtr.tu-darmstadt.de

◇Honda Research Institute Europe GmbH
D-63073 Offenbach, Germany
{jannik.fritsch,
christian.goerick}
@honda-ri.de

Abstract*— Currently available traffic sign recognition systems typically focus on a single class of traffic sign and therefore, the algorithms are optimized to find only this specific class. To this end, a number of approaches for real time capable classification of mostly circular signs exist. Nevertheless, to simultaneously recognize a number of classes a different way has to be taken. This paper presents a real-time capable approach, which uses a two-tiered process independent of the diameter of the sign to cope with all distances. The first stage is our attention system, parameterized to find a number of different types of sign classes. The output of our attention system is a region of interest with a potential traffic sign candidate. The second stage is an array of weak classifiers similar to the idea of Viola and Jones [1], computing a probability value for each of the sign classes. As application area we focus on inner city and therefore, evaluate on the most important traffic sign classes of *Stop* and *Give Way*. Nevertheless, the approach can also detect *Warning* signs and is easily extensible to additional sign classes. The evaluation results show the reliability and mark it as first step towards an overall traffic sign recognition.

Keywords: driver assistance, traffic sign recognition, traffic sign classification

I. INTRODUCTION

The growing importance of driver assistance systems for further decreasing the number of traffic accidents is a widely acknowledged fact. Along with that, the role of these Advanced Driver Assistance Systems grows likewise, shifting the focus from crashworthiness to crash prevention. Nevertheless, currently available traffic sign recognition systems mostly focus on circular signs in restricted application areas, like e.g. highways. Additionally, said traffic sign applications are only used as comfort functions, e.g. to warn the driver about the current speed limit. However, taking the next step towards crash prevention a number of traffic signs classes has to be recognized simultaneously. From our point of view, this has to be done by a single integrated approach, instead of an individual approach for each traffic sign class. Otherwise, the amount of required processing power is not available in the medium term. In addition, the scene complexity has to be increased, to be able to cope with various situations. Therefore, the application area should be inner-city.

In this paper, we present a two-tiered integrated approach for the classification of traffic signs. The algorithm has two stages, firstly our biologically motivated attention system, which generates regions of interest (RoI) with possible traffic sign candidates. Based on the concept of top down (TD) modulation the attention system can actively search for an object class by its use. Secondly, each RoI will be processed by a number weak classifiers, where each classifier generates a probability value for each traffic sign class. Finally, all probability values for a single traffic sign class will be multiplied, resulting in an overall probability for each traffic sign class and region. The weak classifiers were chosen as generic as possible, for being able to easily extend the number of traffic sign classes. So far the parameters of the weak classifiers were analyzed to recognize *Stop*, *Give Way* and *Warning* signs. As the evaluation will show, the algorithm reliably classifies *Stop* and *Give Way* signs in various, complex scenarios. To this end, the proposed approach is an important step towards the simultaneous detection of a number of traffic sign classes in complex environments with varying sizes.

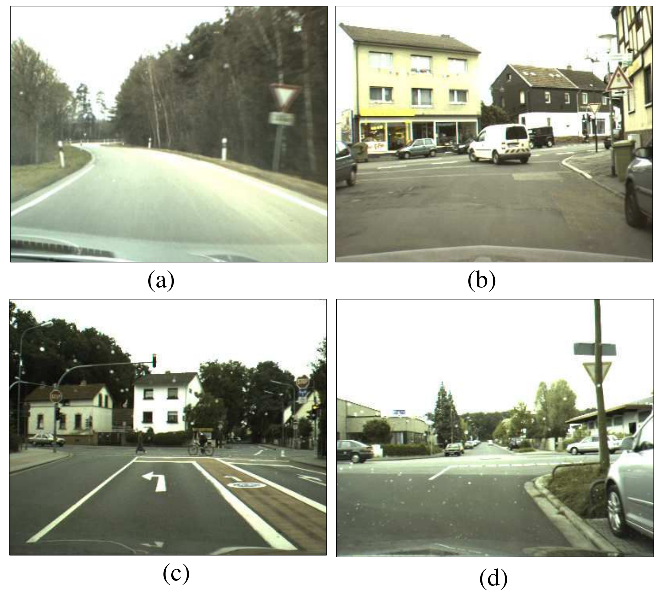

Fig. 1. Examples for traffic sign conditions and scene complexities: (a) blurred due to camera movement, (b) various textures in the surrounding, (c) similar colors in the backround, (d) partly covered by other objects.

II. RELATED WORK

Initial approaches for traffic sign detection date back to the 1980s (see [2] for an overview of the early approaches). To

[1]After finishing his PhD in cooperation between Honda Research Institute Europe GmbH and Darmstadt University of Technology, the author now works for Daimler AG (email: thomas_paul.michalke@daimler.com).

this end, numerous publications handle this topic in general. However, a large number of publications is caused by the variation of traffic signs between different countries. Therefore, this work focuses on German traffic signs. Recently also the first commercial systems for the recognition of speed signs are available. Nevertheless, they are restricted to circular speed signs and also targeted at highways/country roads as application area. Most of the publications focus on the recognition of a single or a few classes of traffic signs. But several try to recognize all sign classes, like [3], [4], [5]. The work of Fang et. al. [3] uses two neural nets to recognize all speed sign categories, one for color and one for edges. However, the neural nets are only trained with signs of a single size and therefore, can only recognize one size for each traffic sign. To this end, the approach [3] is not optimal in the sense that the recognition of signs must be possible for all sizes of a sign and therewith distances to a sign.

The general procedure for the majority of algorithms is similar and can be divided in two parts. First, the algorithms try to extract all regions with possible traffic sign candidates (RoI), termed as detection phase. Second, the previous detected regions are classified or recognized, if there is actually a traffic sign, referred to as classification phase. Additionally, a recognized sign can be temporally integrated to increase the recognition performance even further. But this can be done with every recognition approach and therefore, is not treated here. Because of limitations in space only the central aspects of the different approaches are raised.

A. *Detection phase*

The detection phase can be divided in two types of approaches: on the one hand color based algorithms and on the other hand shape based approaches. Starting with the former, the image is segmented by typical traffic sign colors, which can be done by color relations in the RGB color space [6], [7], [8], [9] and also subspaces of RGB as shown by [10] can be used. Even more prominent is a thresholding in the HSV color space as used by [11], [12], [13], [14], [15] to become more independent to lighting conditions. Also some exotic color spaces as the CIECAM97 are used [16]. Besides the simple thresholding a number of more complex region growing approaches can be found, as e.g. the Color Structure Code (CSC) based on a hierarchical region growing with a hexagonal topology as used by [17]. Also possible is a segmentation by fuzzy sets as introduced by [18].

The second type, the shape based approaches use e.g. a Sobel operator [4], a Canny edge detector [19] or similar on a grayscale image. Afterwards, the resulting edge image is analyzed (in some specific manner) to find regions of interest, with candidates of road signs. To this end, [4] propose a new method for the detection phase by searching for vertical symmetry axes, which should enable the detection of all types of traffic signs. Nevertheless, their example images show only highway scenes as well as their evaluation does not state the scene complexity, leaving in doubt if the approach will also show good performance on inner-city scenarios with a lot of structure.

However, a large part of publications proposes a combination of the former two, starting with a color segmentation and afterwards some kind of shape extraction. The work of Paclik et al. [13] proposes a shape based template matching after the color segmentation in the HSV-colorspace. Garcia et al. [20] use a Prewitt-operator to extract edges of the red color plane. Afterwards a maximum search for each of the image axes is done to find RoIs. Another approach by Oh et al. [21] firstly color segments the image and secondly use a number of symmetry properties to get candidate regions for traffic signs. Zhu et. al. [5] introduce the concept of the Color-Shape Pair (CSP), were a certain color segmentation is followed by a specific shape extraction depending on the type of traffic sign. To this end, the approach should be able to detect all types of traffic signs. Nevertheless, it is not stated how the color segmentation or shape extraction is done. Additionally, the evaluation is only done with emulated images and also no meaningful data (e.g. false positive and false negative rates) is provided.The work of Tsai et. al. [22] apply a RBF-network for the color segmentation and afterwards some kind of edge filtering is applied.

Nearly all approaches have one thing in common, their detection stage is targeted at mainly one type of traffic sign class and can not easily be extended to additional traffic sign classes.

B. *Classification phase*

After the initial detection phase, several regions of interest for an image exist. These RoIs could contain a traffic sign, but also regions with only similar appearance. Therefore, the task of the classification phase is the verification of the RoIs and the identification of true positive traffic signs.

Most of the publications apply well-known approaches for classification with also common assets and drawbacks. One of these is template matching ([23], [6], [11], [21]), were a cross correlation of a traffic sign template from the database with the RoI patch is applied. Additionally, the RoI patch should be previously normalized to get the same size for RoI and template. Nevertheless, the result of a normalized cross correlation is strongly dependent on the similarity of the template and the RoI patch, making the approach vulnerable to normally imperfect traffic sign patches. Another method is the Hough-transformation ([19], [24]) applicable to all kind of geometric shapes as e.g. lines, circles and so on. The method can also handle partial occlusion, but is bound to a single type of shape at a time. Also commonly used for the classification are neural networks ([14], [15], [22], [9], [7]) as already mentioned for [3]. Neural networks can cope with large variances, however the input data has to be normalized to a certain size and the independence to varying diameters of traffic signs is unclear. Also for every sign class a different network has to be trained. Furthermore, a Support Vector Machine can be used for the classification as done by [5], which is similar to a neural network. Other approaches develop special classification methods, as e.g. the comparison of a vector (containing specific extracted features) with a template from the database as [16]. Some approaches not even apply an additional classification stage, but use a combination of color and shape as described in the detection stage (e.g. [20], [18]) for the recognition of traffic signs. Nevertheless, to our knowledge there is no generic approach for the detection and classification of traffic signs, being able to handle complex inner-city scenes and also varying diameters of traffic signs.

III. SYSTEM DESCRIPTION

In the following, a rough overview of our approach for traffic sign recognition is given (see Fig. 3). Thereafter, all processing steps and their theoretical background are described in more detail.

The overall system can be divided in two main parts. The first one is the biologically motivated attention system, acting as detection stage. Therefore, the attention system searches for traffic sign classes based on a number of templates (3-4 templates are sufficient) for each class. Afterwards, each detected region is processed by an array of weak features acting as classification stage. The result is a probability value for each RoI and traffic sign class.

A. Detection stage - Attention System

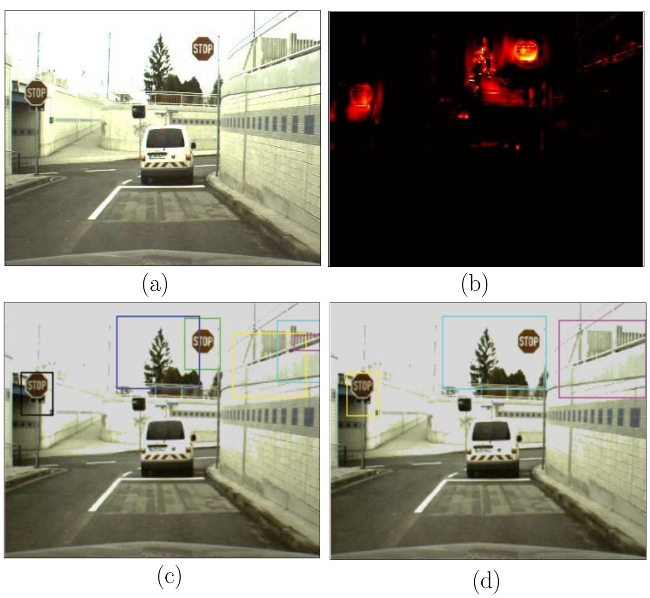

Fig. 2. Detection stage: ((a) Input image of attention system, (b) Attention map for the TD search, (c) RoIs for all sign classes, (d) Fused RoIs for classification stage.

As general detection stage the 320x240 RGB input image (recorded by our test vehicle with a stereo camera setup, see [25]) is analyzed by calculating the attention map S^{total}. Attention is a principle that was found to play an important role in the human vision processing as a mediator between the world and our actual perception [26]. Somewhat simplified, the attention map shows high activation at image positions that are visually conspicuous, i.e., that pop out (bottom-up attention) or that are important for the current system task (top-down attention). The attention map S^{total} results from a weighted combination of N biologically inspired input feature maps F_i (see Eq. (1)). More specifically, we filter the image using, among others Difference of Gaussian (DoG) and Gabor filter kernels that model the characteristics of neural receptive fields measured in the mammal brain. Furthermore, we use the RGBY color space [27] as attention feature that models the processing of photoreceptors on the retina.

The top-down (TD) attention can be tuned task-dependently to search for specific sign classes by calculating a TD weight set w_i^{TD} based on Eq. (2), where $\phi = K_{conj}Max(F_i)$ with $K_{conj} = (0, 1]$.

The bottom-up (BU) weights w_i^{BU} are set object-unspecifically in order to detect in the general case unexpected potentially dangerous scene elements. The parameter $\lambda \in [0, 1]$ (see Eq. (1)) determines the relative importance of TD and BU search and is set to one for the current task of traffic sign detection. Therefore, only the mechanism of top-down attention is used for the detection of regions with traffic signs. For more details regarding the used attention system, see [28].

$$S^{\text{total}} = \lambda \sum_{i=1}^{N} w_i^{\text{TD}} F_i + (1 - \lambda) \sum_{i=1}^{N} w_i^{\text{BU}} F_i \qquad (1)$$

$$w_i^{\text{TD}} = \begin{cases} \frac{m_{RoI,i}}{m_{rest,i}} & \forall \frac{m_{RoI,i}}{m_{rest,i}} \geq 1 \\ -\frac{m_{rest,i}}{m_{RoI,i}} & \forall \frac{m_{RoI,i}}{m_{rest,i}} < 1 \end{cases} \qquad (2)$$

$$\text{with} \quad m_{\{\text{RoI,rest}\},i} = \frac{\sum\limits_{\forall x,y \in \{\text{RoI,rest}\}} \text{pixel values in } F_i}{\text{size region } \{\text{RoI,rest}\}}$$

$$\text{and } \forall\ F_i \geq \phi$$

Each of the traffic sign classes (Stop-, Give Way-, Warning-, Prohibitive-signs, etc.) can be actively searched by using the TD attention. After the computation of the TD attention map, we detect the maximum on it and get the focus of attention (FoA). To this end, an initial segmentation on the current attention map S^{total} is carried out based on a generic region growing. In the following, the segmented FoA is treated as region of interest (RoI) for the classification stage. This procedure (attention generation, FoA segmentation and classification) models the saccadic eye movements of mammals, where a complex scene is scanned and decomposed by sequential focusing of objects in the central 2-3° foveal retina area of the visual field. Nevertheless, the different sign classes have different numbers of occurrences per image and also a visually diverse clearness. To this end, a different number of RoIs have to be extracted for a class on each image, varying from three for *Stop*-signs up to 5 for *Give Way*-signs (three for circular-, five for triangular warning signs). The size of the RoIs can vary from 15x15 pixels to 80x80 pixels, which corresponds to the minimum and maximum size of a traffic sign. The fusion of the resulting

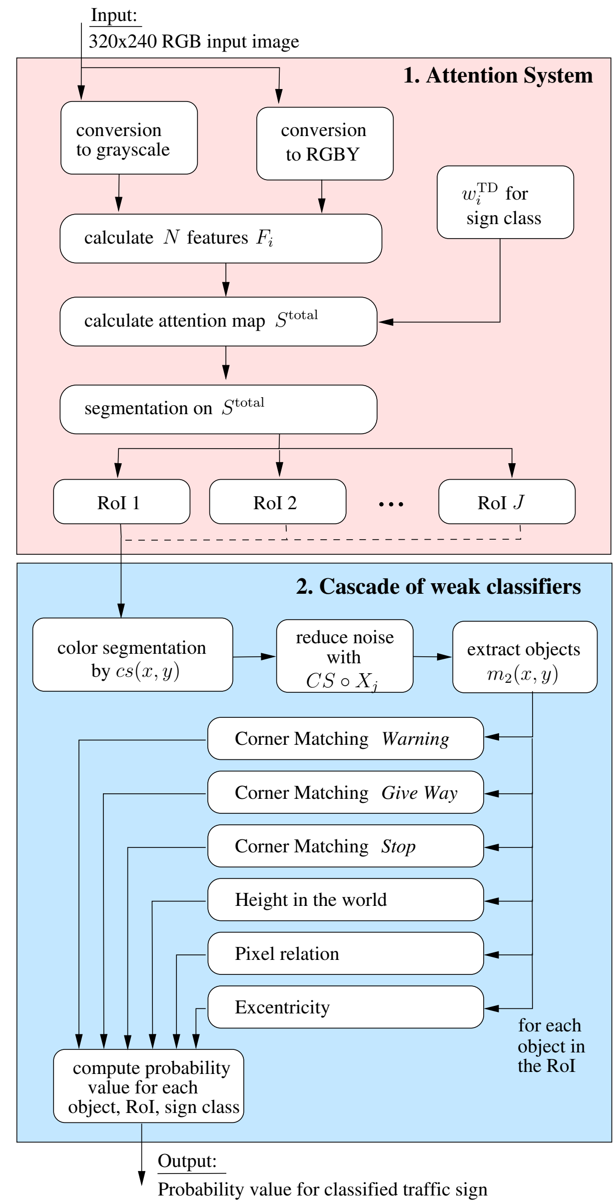

Fig. 3. System structure for traffic sign recognition

RoIs for different sign classes is realised by an overlap criteria of 30% on the smaller RoI, providing a reduced computation time by computing each image region only once. For further details on the online implementation of the attention system please have a look at [29].

B. Classification Stage - Cascade of weak classifiers

The second part of the algorithm handles the extraction of relevant information from each of the RoIs. Each RoI is handled independently, while the procedure for all of the RoIs is the same. For the sake of simplicity, we switch sometimes between the matrix notation (e.g. A) and the function notation (respectively a(x,y)), nevertheless the content of both is the same.

The result for each of the weak classifiers is a probability value, providing the correspondence of an object within a RoI to a certain sign class (the extraction of objects from a RoI is discussed later). At the end, all results for one object and class are multiplied, providing the final result for a certain traffic sign.

Since our detection stage uses a combination of color and shape features, it appeared sufficient to apply a simple color thresholding (see Eq. (3)) to extract relevant structures, similar to other approaches as described in II-A. To become more independent to ligthing conditions we use the RGBY color space (see [27]). Nevertheless, to our knowledge the usage of the RGBY colorspace for the classification of traffic signs is novel. The intervals of the color thresholds were extracted from the set of training images and therefore, show the variation within the data set. See Figure 5b for the result of the color segmentation.

$$
\begin{aligned}
cs(x,y) &= 1 \begin{cases} 0.035 \leq roi_r(x,y) \leq 0.481 \\ G_{min} \leq roi_g(x,y) \leq G_{max} \\ B_{min} \leq roi_b(x,y) \leq B_{max} \\ Y_{min} \leq roi_y(x,y) \leq Y_{max} \end{cases} \\
cs(x,y) &= 0 \text{ otherwise}
\end{aligned} \tag{3}
$$

After the color segmentation a morphological (see [30]) opening operation $CS \circ X_j$ with a number of structuring elements j is applied. The structuring elements are choosen to support the geometrical shapes of the traffic signs, while removing noise (see Fig. 5c for the result and Fig. 4 for the structuring elements). The resulting image patch NR after the noise reduction is given in Eq. (4).

$$
\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad
\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad
\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad
\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

Fig. 4. The X_j structure elements.

$$
NR = \sum_{j=1}^{4} CS \circ X_j \tag{4}
$$

Finally, a mask is generated by evaluation of the local neighbourhood (see [30]) for each pixel. To this end, k_i operators (given through Eq. (6) to Eq. (10)) are applied in two stages on the result of the noise reduction stage NR. The computation of the first stage (with result $m_1(x,y)$) is started at the upper left and continued to the right in a row-wise manner. While, the computation of the second part (with overall result $m_2(x,y)$) is started at the upper right part and continued to the left, also in a row-wise manner. The resulting mask is depicted in Fig. 5d.

$$
m_1(x,y) = \begin{cases} 1 & : k_1(x,y) = 1 \ \vee \ k_2(x,y) = 1 \ \vee \\ & \quad nr(x,y) = 1 \\ 0 & : otherwise \end{cases} \tag{5}
$$

$$k_1(x,y) = 0.25 \cdot [nr(x-1,y-1) + nr(x-1,y) + nr(x-1,y+1) + nr(x,y-1)] \cdot (1 - nr(x,y)) \tag{6}$$

$$k_2(x,y) = 0.25 \cdot [nr(x-1,y-1) + nr(x-1,y) + nr(x+1,y) + nr(x,y-1)] \cdot (1 - nr(x,y)) \tag{7}$$

$$m_2(x,y) = \begin{cases} 1 & : k_3(x,y) = 1 \ \vee \ k_4(x,y) = 1 \ \vee \\ & \quad m_1(x,y) = 1 \\ 0 & : otherwise \end{cases} \tag{8}$$

$$k_3(x,y) = 0.25 \cdot [m_1(x-1,y-1) + m_1(x-1,y) + m_1(x-1,y+1) + m_1(x,y-1)] \cdot (1 - m_1(x,y)) \tag{9}$$

$$k_4(x,y) = 0.25 \cdot [m_1(x-1,y-1) + m_1(x-1,y) + m_1(x+1,y) + m_1(x,y-1)] \cdot (1 - m_1(x,y)) \tag{10}$$

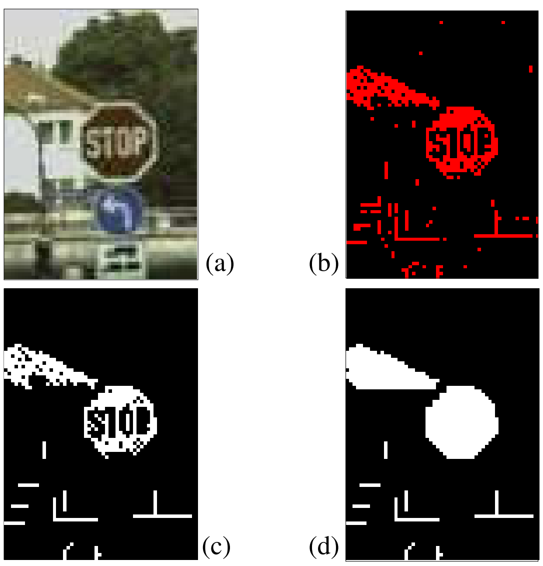

Fig. 5. Classification stage: ((a) Input RoI of classification system, (b) Result of the color segmentation, (c) After opening operation, (d) Masks for each object within the RoI.

The procedure described so far is generic, and therefore work with all kinds of sign shapes (octagonal, rectangular, triangular, circular).

In the following, the weak classifiers (also called features) will be described. Due to space limitations the pixel relation feature shows exemplarily the overall concept, whereas similar processing steps of the other features will be left out.

To extract all relevant objects from the RoI, the Neumann neighborhood is applied on M_2 defining all independent objects. The order of the objects is defined by the results of the corner matching feature. Hence, starting with the object having the highest result of the corner matching, followed by the second highest and so on. Therefore, the mask of each object is multiplied with the initial RoI providing only the relevant pixels.

The pixel relation is defined by the ratio of red to white pixels. Therefore, the amount of pixels for each of the classes has to be determined. This has already been done for the class of red pixels with the initial color segmentation. An additional color segmentation (similar to the initial one) for white pixels is applied to extract all white pixels of the RoI. Afterwards, the mask $m_2(x,y)$ separates the independent objects, facilitating the counting of red and white pixels to allow the easy use for the computation of the pixel relation. As ground truth data the red and white area sizes of the signs have been estimated from the specifications of German road traffic regulations (see Eq. (11) for *Give Way* and Eq. (12) for *Stop*).

$$\left.\frac{p_{red}}{p_{white}}\right|_{\text{GW,ideal}} = \frac{A_{red}}{A_{white}} = 0,98. \tag{11}$$

$$\left.\frac{p_{red}}{p_{white}}\right|_{\text{STOP,ideal}} = 5,04. \tag{12}$$

Due to a number of noise sources (e.g. slant, blurred, light conditions, etc.) the ideal pixel relation will hardly be the outcome. To this end, the pixel relation is transformed to a kind of fuzzy set (see Figure 6 for *Give Way* signs) providing the membership function of the pixel relation to a certain traffic sign class. The membership function was acquired from the image test set and provides in our case the probability value P_{pix} for a certain sign class.

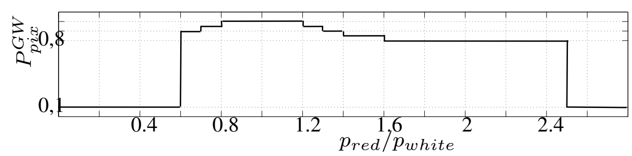

Fig. 6. Fuzzy membership function of the pixel relation for the *Give Way* sign class.

The next feature is the corner matching, which is similar for all sign classes, but differs in the number and type of corner templates that are used (see Fig. 7 for used corners of different sign classes).

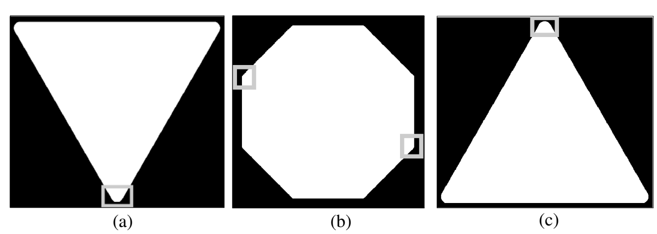

Fig. 7. Used corners for matching: (a) *Give Way*, (b) *Stop*, (c) *Warning*.

Therefore, only the *Give Way* sign will be explained in detail. The corner matching is a template matching (see [30]) with an ideal template of certain characteristic corners for a traffic sign class. As example the template of the lower corner from a *Give Way* sign is given in Fig. 8. Every template W_l will be cross-correlated with each mask $m_2(s,t)$ from an RoI by use of Eq. (13) (N and O define the size of $m_2(s,t)$).

$$
\begin{aligned}
cr_l(x,y) = \sum_s \sum_t & m_2(s,t) w_l(x+s, y+t) \\
\text{for } \ & x = 0,1,2,...,N-1, \\
& y = 0,1,2,...,O-1
\end{aligned} \tag{13}
$$

The result $cr_l(x,y)$ contains a complete RoI and therefore, the highest result determines the object to start with (as already mentioned). Additionally, for each object the highest correlation result is transformed by a fuzzy set (each corner w_l has its own fuzzy set) to a probability value (similar to the pixel relation). To this end, each object has a certain probability P_{w_l} that the characteristic corner w_l is present. If there is more than one corner for a certain sign class (e.g. *Stop* sign), not only the single probabilities P_{w_l} will be computed but also a combined probability P_{rs}, which evaluates the spatial relation of the single corners to each other. This is also transformed by a fuzzy set, which evaluates that a certain corner has to be higher in the image than the other one.

2	1	1	1	1	1	2
-1	4	2	2	2	4	-1
-2	-1	4	4	4	-1	-2
-4	-2	-1	6	-1	-2	-4
-6	-4	-2	-1	-2	-4	-6

Fig. 8. Template W_1 for *Give Way* sign.

The following feature is the excentricity, which defines the relation between length to width of an object. Therefore, the mask of each object is used the determine the maximum and minimum row and column numbers of the object. Given this information the excentricity of an object can be easily computed. Afterwards, the excentricity will also be transformed by a fuzzy set to a probability P_{exc}, describing the presence of a certain traffic sign corresponding to the choosen fuzzy set. The ground truth for most of the traffic signs concerning the excentricity is about one, nevertheless we use independent fuzzy sets for the different sign classes since their noise sensitivity differs.

The final weak classifier is the height in the world. Due to the stereo camera system in our test vehicle we are able to estimate the world position of all pixels. For further details about the camera and matrix transformations please refer to [31]. Hence, for each object the median of the single pixel positions is computed and again transformed by a fuzzy set to a probability value P_{height}.

Finally, the overall probability for each of the traffic sign classes has to be evaluated. Therefore, the Eq. (14), Eq. (15) and Eq. (16) provide the overall probability for the traffic sign classes of *Stop*, *Give Way* and *Warning*.

$$P_{GW} = P_{w_1} \cdot (1 - P_{w_4}) \cdot P_{pix}^{GW} \cdot P_{exc}^{GW} \cdot P_{height} \tag{14}$$

$$P_{STOP} = P_{w_2} \cdot P_{w_3} \cdot P_{pix}^{STOP} \cdot P_{exc}^{STOP} \cdot P_{height} \cdot P_{rs}^{STOP} \tag{15}$$

$$P_{WARN} = P_{w_4} \cdot (1 - P_{w_1}) \cdot P_{pix}^{WARN} \cdot P_{exc}^{WARN} \cdot P_{height} \tag{16}$$

Traffic signs	# number signs	Correct-ness	Comple-teness	Quality
Stop	64	100%	97%	97%
Give way	53	96.2%	98.1%	94.4%
Both	**117**	**98.3%**	**89.8%**	**88.5%**

TABLE I

RESULTS OF THE EVALUATION

IV. RESULTS

In this section, we evaluate the performance of our system with a total of 820 images, taken from two image streams. The 820 images show 117 relevant traffic signs on 93 images. Therefore, a number of images contain two relevant traffic signs. Nevertheless, the approach and also the evaluation measure treats each traffic sign independently. To this end, each of the traffic signs on an image has to be classified. No temporal integration (usage of sign information from previous images) of a traffic sign is done here to show the single image performance.

The images were manually labelled, providing the exact position and type of each traffic sign. The images show various inner-city scenes, with different scene complexities (see Fig. 1).

The approach is implemented with Matlab and was evaluated on a 2 GHz Intel Core2 Duo, having 2 GB Ram and running Windows Vista. Only one of the CPU cores was used for the computation.

In order to evaluate our algorithm, we adopt the Equations (17), (18), and (19) (with True positive traffic signs (TP), False negative traffic signs (FN), and False positive traffic signs (FP)). The equations define different ground truth based measures, which were taken from [32].

$$\text{Completeness} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{17}$$

$$\text{Correctness} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{18}$$

$$\text{Quality} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \tag{19}$$

On a descriptive level the Completeness states, based on given ground truth data, how many of the traffic signs were actually detected. The Correctness states how many of the detected regions were actually relevant traffic signs. The Quality combines both measures. Its computation is appropriate, since a trade-off between the Completeness and Correctness is possible. Based on this, the Quality measure should be used for a comparison, since it weights the FP and FN signs equally. For a more detailed analysis the Completeness and Correctness state what exactly caused a difference in Quality. A traffic sign is counted as true positive detection if the corresponding traffic sign class as well as position (with a range of five pixel to the ground truth center) on the image is detected.

The three measures were calculated on the detected traffic signs over all images of the two inner-city streams. The gathered results are depicted in Tab. I.

So far no optimisation of the classification is done concerning the speed of the implementation (all computations are done in the image domain), nevertheless the average computation time in Matlab per image is only 9.6sec. Therefore, with modifications concerning speed and the experience that a factor of 100 can be gained with a native C implementation, the approach should be capable of a 25Hz frame rate.

An in depth comparison with other approaches is difficult to realise. Because, there is no commonly accessible database with image streams, providing the same input images for all approaches. Therefore, some publications use only images that always contain traffic signs, others use only images of traffic signs with a certain size, some use only images of highway scenes and a few generate synthetic images. In general, the class of recognized traffic signs varies and also the appearance of signs varies with countries. Additionally, most of the publications do not state their classification rates, to name only a few of the differences. Due to the mentioned difficulties, a direct comparison with other approaches was not carried out. Nevertheless, an overall Quality of nearly 90% shows the reliability of our approach.

V. SUMMARY

This paper describes a generic method for traffic sign detection and classification. As far as we know, is the combination of attentation based detection and an array of weak classifiers for classification a novel approach for generic traffic sign classification. Based on the proposed approach, results are obtained that allow the building of safety relevant algorithms, like e.g. active collision prevention when violating a *Stop* sign in inner-city situations. In general, information on traffic signs is important for the understanding of complex scenarios in future high level applications. Currently, we are extending the approach with new weak classifiers which support circular signs being the last missing traffic sign class. Nevertheless, to differentiate within a traffic sign class and therefore, between the different pictograms of e.g. warning signs an additional classification step is necessary.

REFERENCES

[1] P. Viola and M. Jones, "Robust Real-time Object Detection," in *Internat. Journal of Computer Vision*, 2001.

[2] P. Paclik, "Road Sign Recogniton Survey," http://euler.fd.cvut.cz/research/rs2/files/skoda-rs-survey.html, 1999.

[3] C. Fang, S. Chen, and C. Fuh, "Road-Sign Detection and Tracking," *IEEE Transactions on Vehicular Technology*, vol. 52, 2003.

[4] C. Nunn, A. Kummert, and S. Müller-Schneiders, "A two stage Detection Module for Traffic Signs," in *Proc. IEEE Internat. Conf. on Vehicular Electronics and Safety (ICVES)*, 2008, pp. 248–252.

[5] S. Zhu, L. Liu, and X. Lu, "Color-Geometric Model for Traffic Sign Recognition," in *Proc. IMACS Multiconference on Computational Engineering in Systems Applications*, Oct. 2006, pp. 2028–2032.

[6] V. Andrey and K. H. Jo, "Automatic detection and recognition of traffic signs using geometric structure analysis," in *Proc. Internat. Joint Conference SICE-ICASE*, 18–21 Oct. 2006, pp. 1451–1456.

[7] A. d. l. Escalera, L. E. Moreno, M. A. Salichs, and J. M. Armingol, "Road Traffic Sign Detection and Classification," *IEEE Transactions on Industrial Electronics*, vol. 44, no. 6, pp. 848–859, Dec. 1997.

[8] H.-M. Yang, C.-L. Liu, K.-H. Liu, and S.-M. Huang, "Traffic Sign Recognition in Disturbing Environments," in *ISMIS 2003, LNAI 2871*. Springer Verlag, 2003, pp. 252–261.

[9] S. Kantawong, "Road Traffic Signs Detection and Classification for Blind Man Navigation System," in *Proc. Internat. Conf. on Control, Automation and Systems (ICCAS)*, 17–20 Oct. 2007, pp. 847–852.

[10] M. Zadeh, T. Kasvand, and C. Suen, "Localization and Recognition of Traffic Signs for Automated Vehicle Control Systems," in *Conf. on Intelligent Transportation Systems, part of SPIE's Intelligent System and Automated Manufacturing*, 1997, pp. 272–282.

[11] B. Cyganek, "Road Signs Recognition by the Scale-Space Template Matching in the Log-Polar Domain," in *IbPRIA 2007, Part I, LNCS 4477*. Springer Verlag, 2007, pp. 330–337.

[12] N. Kehtarnavaz, N. C. Griswold, and D. S. Kang, "Stop-Sign Recognition based on Color and Shape Processing," *Machine Vision and Applications*, vol. 6, pp. 206–208, 1993.

[13] P. Paclk, J. Novovicov, P. Pudil, and P. Somol, "Road Sign Classification Using Laplace Kernel Classifier," *Pattern Recognition Letters*, vol. 21, pp. 1165–1173, December 2000.

[14] S.-D. Zhu, Y. Zhang, and X.-F. Lu, "Detection for Triangle Traffic Sign Based on Neural Network," in *ISNN 2006, LNCS 3973*. Springer Verlag, 2006, pp. 40–45.

[15] Y. Wang, J. Dang, and Z. Zhu, "Traffic Signs Detection and Recognition by Improved RBFNN," in *Proc. Internat. Conf. on Computational Intelligence and Security*, 15–19 Dec. 2007, pp. 433–437.

[16] X. Gao, L. Podladchikova, and D. Shaposhnikov, "Application of Vision Models to Traffic Sign Recognition," in *ICANN/ICONIP 2003, LNCS 2714*. Springer Verlag, 2003, pp. 1100–1105.

[17] L. Priese, J. Klieber, R. Lakmann, V. Rehrmann, and R. Schian, "New Results on Traffic Sign Recognition," in *Proc. IEEE Intelligent Vehicles Symposium*, 1994, pp. 249–254.

[18] H. Fleyeh, "Traffic Sign Recognition by Fuzzy Sets," in *Proc. IEEE Intelligent Vehicles Symposium*, 4–6 June 2008, pp. 422–427.

[19] M. A. Garca-Garrido, M. A. Sotelo, and E. Martn-Gorostiza, "Fast Road Sign Detection Using Hough Transform for Assisted Driving of Road Vehicles," in *EUROCAST 2005, LNCS 3643*. Springer Verlag, 2005, pp. 543–548.

[20] M. A. Garcia, M. A. Sotelo, and E. M. Gorostiza, "Traffic Sign Detection in Static Images using Matlab," in *Proc. IEEE Conf. on Emerging Technologies and Factory Automation (ETFA)*, vol. 2, 16–19 Sept. 2003, pp. 212–215.

[21] J.-T. Oh, H.-W. Kwak, Y.-H. Sohn, and W.-H. Kim, "Segmentation and Recognition of Traffic Signs Using Shape Information," in *ISVC 2005, LNCS 3804*. Springer Verlag, 2005, pp. 519–526.

[22] L.-W. Tsai, Y.-J. Tseng, J.-W. Hsieh, K.-C. Fan, and J.-J. Li, "Road Sign Detection Using Eigen Color," in *ACCV 2007, Part I, LNCS 4843*. Springer Verlag, 2007, ch. 169-179.

[23] A. d. l. Escalera, J. M. Armingol, J. M. Pastor, and F. J. Rodriguez, "Visual Sign Information Extraction and Identification by Deformable Models for Intelligent Vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 2, pp. 57–68, June 2004.

[24] J. Turan, M. Fifik, and L. Ovsenik, "Transform Based System for TrafficSign Recognition," in *Proc. 15th Internat. Conf. on Systems, Signals and Image Processing (IWSSIP)*, June 2008, pp. 441–444.

[25] T. Michalke, A. Gepperth, M. Schneider, J. Fritsch, and C. Goerick, "Towards a human-like vision system for resource-constrained intelligent cars," in *Proc. 5th Internat. Conf. on Computer Vision Systems (ICVS)*, 2007.

[26] D. Simons and C. Chabris, "Gorillas in our midst: Sustained inattentional blindness for dynamic events," *British Journal of Developmental Psychology*, vol. 13, pp. 113–142, 1995.

[27] S. Frintrop, "VOCUS: A visual attention system for object detection and goal-directed search," Ph.D. dissertation, University Bonn.

[28] T. Michalke, J. Fritsch, and C. Goerick, "Enhancing robustness of a saliency-based attention system for driver assistance," in *Proc. 6th Internat. Conf. on Computer Vision Systems (ICVS), Santorini, Greece. Lecture Notes in Computer Science, Springer*.

[29] J. Fritsch, T. Michalke, A. Gepperth, S. Bone, F. Waibel, M. Kleinehagenbrock, J. Gayko, and C. Goerick, "Towards a human-like vision system for driver assistance," in *Proc. IEEE Intelligent Vehicles Symposium*, 4–6 June 2008, pp. 275–282.

[30] B. Jähne, *Digitale Bildverarbeitung*. Springer Verlag, 2005.

[31] T. Michalke, R. Kastner, J. Adamy, S. Bone, F. Waibel, M. Klainhagenbrock, J. Gayko, A. Gepperth, J. Fritsch, and C. Goerick, "An Attention-based System Approach for Scene Analysis in Driver Assistance," *Automatisierungstechnik*, vol. 56, pp. 575–584, 2008.

[32] P. Lombardi, M. Zanin, and S. Messelodi, "Unified stereovision for ground, road and obstacle detection," in *Proc. IEEE Intelligent Vehicles Sysposium*, 2005.