

COLOR FEATURES FOR VISION-BASED TRAFFIC SIGN CANDIDATE DETECTION

Steffen Görmer, Anton Kummert

Faculty of Electrical Engineering and Media Technologies, University of Wuppertal, D-42119 Wuppertal, Germany

Stefan Müller-Schneiders

Delphi Electronics & Safety, D-42119 Wuppertal, Germany

Keywords: Traffic sign detection, Color feature extraction, Color segmentation, Color analysis.

Abstract: A common approach for traffic sign detection and recognition algorithms is to use shape based and in addition color features. Especially to distinguish between speed-limit- and end-of-speed-limit-signs the usage of color information can be helpful as the outer border of speed-signs is in a forceful red. In this paper the focus is faced on color features of speed-limit and no-overtaking signs. The apparent color in the captured image is varying very much due to illumination conditions, sign surface condition and viewing angle. Therefore the color distribution in the HSV color space of a sufficient amount of signs at different illumination conditions and aging has been collected, examined, and a matching mathematical model is developed to describe the subregion in the according color space. Once the color region of traffic signs is known, two kinds of traffic sign segmentation algorithms are developed and evaluated with the explicit focus only on color features to preselect subregions in the image where (red bordered) traffic signs are likely to be.

1 INTRODUCTION

Together with Nightvision, Automatic Headlight Control (AHC), Lane Departure Warning (LDW), Adaptive Cruise Control (ACC) and Parking Assistance (PA) the Traffic Sign Recognition System (TSR) is one of the main applications in vision-based driver assistance systems. It keeps the driver informed for instance about current speed limits, danger areas and right-of-way directives. Navigation systems already provide information about speed-limits adapted from a database which will not be absolutely up to date concerning road works and temporary speed limits. Especially variable traffic signs, upcoming more and more to handle the different traffic conditions during the day or dependent on the current weather condition, must be reported just in time. To overcome this issue, a vision based TSR-system is the proximate approach as a built-in camera for different applications becomes apparent in future vehicles anyway and driving is, for the human as well, a task based almost completely on visual information processing.

Vision based TSR is divided into two main approaches, based either on grayscale (Gavrila, 1999), (Barnes and Zelinsky, 2004) or color images

(Bahlmann et al., 2005), (Escalera et al., 2003), (Fang et al., 2003), (Siogkas and Dermatas, 2006), (Torrissen et al., 2004). Of course a color image provides additional information, but has also discredits due to the variation of colors and illumination conditions and requests more bandwidth, processing power and memory. Johansson (Johansson, 2002) gives a good overview on different approaches for either color- and shape-based approaches. Priese et al. (Priese et al., 1993), (Priese et al., 1994), (Priese and Rehrmann, 1998) developed the Color Structure Code to perform color-segmentation for TSR purposes based on region growing. Fleyeh (Fleyeh, 2006) introduces a segmentation method for TSR comparing different color spaces and figures out the advantages of the shadow and highlight invariant HSV color model.

This paper gives an introduction into colorvision followed by a detailed analysis of traffic sign color properties. Based on an adequate number of signs a color dataset is created by a labeling-tool to quantitatively evaluate the color distribution especially in different illumination conditions. This distribution is also modeled by the Covariance matrix to approximate the real world distribution. A segmentation algorithm based on the Mahalanobis Distance is im-

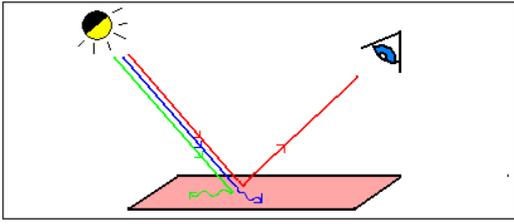


Figure 1: Principle of color perception: reflection and absorption.

plemented. At the end the performance is compared to the approach of simple thresholding presented e.g. by Fleyeh (Fleyeh, 2006) with adapted thresholds according to the results of the color distribution analysis.

2 COLORVISION

The human idea of color is nothing more than a sensory perception. Based on the human eye design with its red-, green- and blue-sensitive receptors the human brain processes an impression containing information about the intensity, the saturation and the coloring of the incoming light. The perception of color is generally due to the spectrum of a light source as well as to the properties of the object reflecting it and of course dependent on the sensor. The spectrum of a source shows the emitted intensities dependent on the specific wavelength. Most light-sources emit many different wavelengths with approximately equal intensities which causes the perceived color to be white. An object can reflect at most the whole spectrum of the involved light source, in case it is no mirror it will actually absorb certain wavelengths and transmit and/or reflect the rest. So what we see if we call an object colored is a reflected part of the original source spectrum (for more details see (Gonzales and Woods, 2002), (Shevell, 2003)). Thus, one can conclude that the perception of color is very dependent on the incident illumination. An object will only be considered to be red, if on the one hand the illumination contains the red spectrum and on the other hand the green and blue part of the spectrum is absorbed or transmitted (see figure 1).

As RGB is the standard color-model in computer vision and all capture devices normally deliver frames in that color pattern it is easily accessible and easy to handle. But as expected the investigations from Fleyeh (Fleyeh, 2006) show that each component R, G and B are dependent on the sensor response, surface albedo, illumination intensity, surface orientation and illumination direction. This is a difficult issue for color segmentation because changing illumination

conditions like shadows and highlights can involve a vast deviation of the RGB-data affecting all three channels.

As mentioned above the human brain creates an impression of coloring, saturation and intensity. The HSV color space corresponds to that, characterizing a color by hue (H), saturation (S) and intensity (V=value). The hue-component is an angle and represents the dominant wavelength. It describes a color-circle starting from red (0°) to yellow (60°), green (120°), cyan (180°), blue (240°), magenta (300°) and back to red ($360^\circ/0^\circ$) again. Therefore this color space is represented in cylindrical coordinates, illustrated in figure 2. Saturation can be found as the radius, intensity is the height or z-component. The neutral axis (gray-scales) in this color space is the centerline of the cylinder, where saturation is very low or zero and hue is insignificant.

The transformation from 24-bit RGB to 24-bit HSV color space is done with the following equations:

$$V = \text{MAX} \quad (1)$$

$$S = \frac{V - \text{MIN}}{V} \cdot 255 \quad (2)$$

$$H = \left(0 + \frac{G - B}{\text{MAX} - \text{MIN}}\right) \cdot \frac{60^\circ}{2} \text{ for } R = \text{MAX} \quad (3)$$

$$H = \left(2 + \frac{B - R}{\text{MAX} - \text{MIN}}\right) \cdot \frac{60^\circ}{2} \text{ for } G = \text{MAX} \quad (4)$$

$$H = \left(4 + \frac{R - G}{\text{MAX} - \text{MIN}}\right) \cdot \frac{60^\circ}{2} \text{ for } B = \text{MAX} \quad (5)$$

with

$$\text{MAX} = \max(R, G, B) \quad (6)$$

$$\text{MIN} = \min(R, G, B) \quad (7)$$

Complying to the 3×8 -Bit storage model of common color images (e.g. 24-bit bitmap) saturation and intensity have a range of $\{0 \dots 255\}$, similar to RGB-channels. The hue-component is divided by two to fit

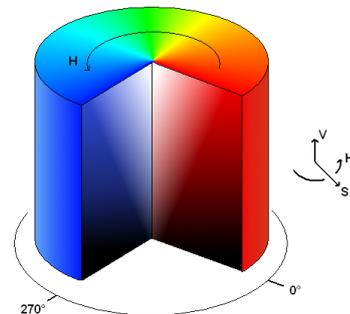


Figure 2: HSV represented by a cylinder (angle: H, radius: S, height: V).



Figure 3: Sample image from a captured sequence containing a traffic sign (originally VGA 640×480).

into 8 Bits and thus has a range of $\{0 \dots 180\}$ instead of $\{0 \dots 360\}$.

3 COLOR ANALYSIS

3.1 Traffic Sign Color Properties

Traffic signs are very important to inform drivers about prohibitions, commandments and dangers. Therefore they are designed in a very distinct way. Europe-wide, speed signs are circular in shape, the base color is white and the outer circle is painted in a bright signal red which is well silhouetted against the natural environment. In Germany this red tint is specified in the norm DIN 5381 and has the RAL-number 3001. This is a pure denotation for printing purpose and does not help very much for color segmentation tasks as the actually perceived color by the camera is dependent on so many factors. But there are some color transformation tables which give estimated RGB-values for the colors according to the RAL-standard in printing technology. RAL 3001 hence corresponds to the RGB-components $\{158^1, 21^1, 25^1\}$. The related HSV-parameters are $\{-1.75^\circ, 221^1, 158^1\}$. So these results would be expected capturing a brand new sign with a photometric color camera under well defined standard illumination conditions. Unfortunately this is very different in outdoor scenes. The color also deteriorates over time towards a bleached orange because of weather influences and UV-rays treating the sign surface.

3.2 Database for Investigation Purpose

To investigate all these issues a database of captured signs and their color attributes is necessary. Therefore a set of sequences is captured with a VGA high dynamic range automotive camera mounted on the windshield. Each sequence is saved as 24 Bit raw-data in an AVI file-container (see figure 3). Individual frames containing traffic signs are extracted

¹in the Range of $\{0 \dots 255\}$

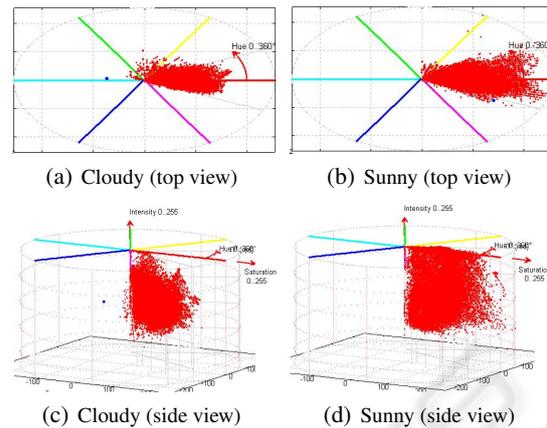


Figure 4: Color distribution at different weather conditions in HSV color space.

from the sequences and saved as bitmaps to analyze their color property. To extract a desired area with similar color in an image an extraction tool is developed. Therewith the red border of traffic signs can be marked and cut out. A region-growing algorithm is the key to this feature. The extracted region in RGB-coordinates is transformed into the HSV color space and all pixel-values are written into a txt-file. Some lines of gnuplot-commands are added to the data-files to display the data-set. Equipped with this applicatory feature different illumination conditions can be compared easily.

3.3 Analysis

The weather condition plays a significant role concerning the color appearance. With the shining sun the saturation is increasing as well as the intensity of course, whereas the opposite happens on cloudy days. In very rainy conditions the illumination is deteriorating and the color-information in the image is perceptible decreasing. Figure 4 shows the red-spectrum in the HSV color space at different weather conditions. Each dataset contains about 90000 pixels (~ 200 traffic signs). Few individual points far outside the expected region of color-values are due to inaccurate labeling and do not really belong to a traffic sign but maybe to its immediate environment.

To be able to make a statement on the actual changing of color-behavior under different weather conditions it is reasonable to have also a look on the average color-values of each data-set.

Table 1 shows the relations. The hue-value representing the most important feature does not vary too much (see figure 4). Of course saturation increases at sunny weather but the change in intensity is low and insignificant anyway. The sign color tends towards a

darker red/violett viewing against the sun (figure 5(a)) and on the contrary into a slight orange with the sun in the back (figure 5(b)).

In cloudy weather conditions there are minor differences. The colors are in a closer range due to the more or less constant illumination condition. Rainy scenes correspond to cloudy weather. There is not so much difference beside the decreasing intensity or sporadic distortions by raindrops within the image.

Table 1: Average values for different weather conditions.

	hue	saturation	value
cloudy	-1°	75^1	145^1
sunny	0°	102^1	166^1

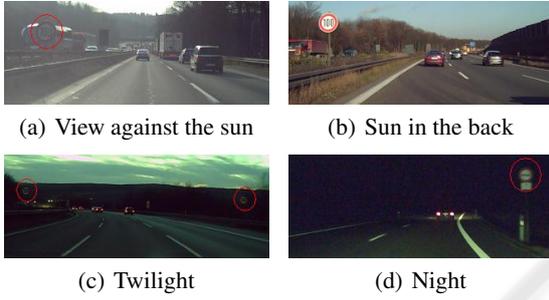


Figure 5: Traffic sign appearance at different weather conditions and daytime.

Without doubt the trait of sunlight changes during the day ((Bénallal and Meunier, 2003)). But due to constantly changing lighting conditions (changing image-background, e.g. forest, sky), the continuous automatic adjustment of the camera and the very different appearances of traffic signs, the variation of sunlight is negligible. At twilight and at night traffic sign colors appear almost equal in both conditions if not completely distorted by overexposure (highbeam reflectance) or motion blur (see figure 5(c) and 5(d)). Saturation and intensity are quite low of course and there is a drift towards dark violet as experienced with direct sunlight. Also in most cases the sensitivity of the sensor was too low to obtain suitable color-informations out of the image.

4 MODELING OF COLOR-DATA BY COVARIANCE MATRIX

Obviously the traffic sign colors form clouds around its average. These clouds can be approximated by an ellipsoid, which is defined by the Mahalanobis-

Distance γ from the average center point $\vec{\mu}$.

$$\gamma = (\vec{hsv} - \vec{\mu})^T \cdot (\text{Cov}(H, S, V))^{-1} \cdot (\vec{hsv} - \vec{\mu}) \quad (8)$$

The Mahalanobis-Distance is dependent on the orientation and describes the size of the ellipsoid and the amount of color-values of traffic signs to be included. It corresponds to the radius in a sphere. The orientation dependency which forms the ellipsoid is due to the covariance matrix (abbrev. Cov).

$$\text{Cov}(H, S, V) = \frac{1}{N} \cdot \sum_{i=0}^N (\vec{hsv}_i - \vec{\mu}) \cdot (\vec{hsv}_i - \vec{\mu})^T \quad (9)$$

with

$$\vec{hsv}_i = (H_i, S_i, V_i)^T \quad (\text{values of pixel } i) \quad (10)$$

and

$$\vec{\mu} = \frac{1}{N} \cdot \sum_{i=1}^N \vec{hsv}_i = \begin{pmatrix} \mu_H \\ \mu_S \\ \mu_V \end{pmatrix} \quad (\text{average}) \quad (11)$$

where N is the number of pixels collected in the dataset. As the ellipsoid is a special case of a sphere, it can be derived from that. A sphere is defined as follows:

$$\vec{u}^T \cdot \vec{u} = \gamma \quad (12)$$

with

$$\vec{u} = \sqrt{\gamma} \cdot \begin{pmatrix} \sin \vartheta \cdot \sin \varphi \\ \sin \vartheta \cdot \cos \varphi \\ \cos \vartheta \end{pmatrix} \quad \text{where} \quad \begin{matrix} 0 \leq \vartheta \leq \pi \\ 0 \leq \varphi \leq 2\pi \end{matrix} \quad (13)$$

So equation 8 has to be adapted to 12. Therefore the covariance matrix is split up by the Cholesky-Decomposition:

$$\text{Cov}^{-1} = \sqrt{\text{Cov}^{-1}^T} \cdot \sqrt{\text{Cov}^{-1}} \quad (14)$$

The deviation from the average is substituted by \vec{v} :

$$\vec{v} = \vec{hsv} - \vec{\mu} \quad (15)$$

$$\gamma = (\vec{hsv} - \vec{\mu})^T \cdot (\text{Cov}(H, S, V))^{-1} \cdot (\vec{hsv} - \vec{\mu}) = \quad (16)$$

$$= \vec{v}^T \cdot \sqrt{\text{Cov}^{-1}^T} \cdot \sqrt{\text{Cov}^{-1}} \cdot \vec{v} \quad (17)$$

equating 12 and 16:

$$\vec{u} = \vec{v} \cdot \sqrt{\text{Cov}^{-1}} \quad (18)$$

Finally, the ellipsoid in the HSV color space is specified as follows:

$$\begin{pmatrix} H \\ S \\ V \end{pmatrix} = (\sqrt{\text{Cov}^{-1}})^{-1} \cdot \sqrt{\gamma} \cdot \begin{pmatrix} \sin \vartheta \cdot \sin \varphi \\ \sin \vartheta \cdot \cos \varphi \\ \cos \vartheta \end{pmatrix} + \vec{\mu} \quad (19)$$

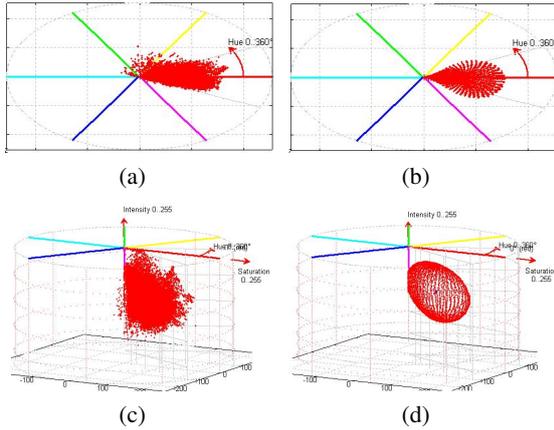


Figure 6: Color distribution and the corresponding covariance ellipsoid.

with

$$0 \leq \vartheta \leq \pi \quad \wedge \quad 0 \leq \varphi \leq 2\pi \quad (20)$$

In figure 6 a comparison between the real color distribution and the modeled covariance ellipsoid with a Mahalanobis-Distance of $\gamma = 7$ is shown.

5 IMPLEMENTATION OF A SEGMENTATION ALGORITHM

Now that the facts about the appearance of traffic signs are pointed out, different segmentation methods are investigated. As traffic signs should only appear in a certain region in the image (at the border of the road), a region of interest (ROI) is defined to narrow down the image processing task. The OpenCV-library supplies very useful functions to handle images and access individual pixels. Easy capturing of frames from a camera adapted to the video-for-windows interface is embedded as well as grabbing them from an existing AVI-file. So the captured sequences can be used as input for testing purpose. Since the HSV color space is the most promising approach due to comparative illumination independence, evaluation is presented in this matter. Of course it would be applicable in the other color spaces as well.

The first step is to convert the grabbed image into the HSV space. This is done with the transformation equations (1) to (7). Now two segmentation methods are considered. On the one hand the segmentation can be based on the Mahalanobis Distance as shown in 4, on the other hand a simple thresholding can be applied by cutting out a subspace in the shape of a slice of pie.

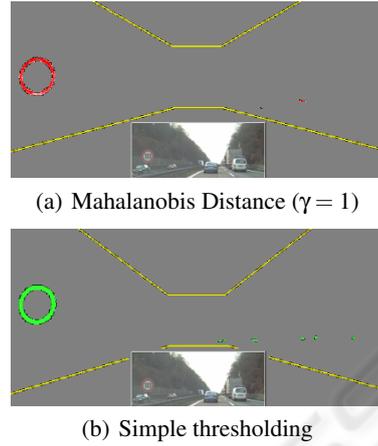


Figure 7: Segmented image with different methods.

5.1 Mahalanobis Distance

As the traffic sign color-data can be approximated by the according covariance matrix and Mahalanobis Distance, this method can also be used for segmentation purpose. The covariance of a training data set can be calculated and printed into a text-file together with the average value. This text-file is loaded into the segmentation algorithm and the inverse of the covariance matrix is calculated. Therewith the Mahalanobis Distance can be processed for each pixel of an image according to equation (8). All pixels within a certain γ -range from the average are segmented. This approach allows highlighting the segmented pixels with different weighing so that closer color-values are displayed in darker red and colors being further away from the average in a lighter red (see figure 7(a)).

5.2 Thresholding

Vitabile et al. (Vitabile et al., 2002) defined three sub-spaces in the HSV color space:

- achromatic area with $S \leq 60 \vee V \leq 50$
- unstable area with $60 \leq S \leq 130 \wedge V \geq 50$
- chromatic area with $S \geq 130 \wedge V \geq 50$

The achromatic area must be factored out in color segmentation tasks as the hue-value is very unstable and the color has no dominant wavelength. In other words it is even visually gray-scale. Of course the boundary values are closely dependent on the capture device. In this case a minimum saturation of $S = 50$ has proved to be still suitable. So in segmentation process all pixels are cut out which do not comply with the following constraint:

$$-25^\circ \leq H \leq +25^\circ \quad \wedge \quad S \geq 50 \quad \wedge \quad V \geq 50 \quad (21)$$

All other pixels considered to be red are highlighted in green. The result can be seen in figure 7(b).

6 EVALUATION AND COMPARISON

To get an impression which method is superior both are tested with specially selected individual frames which are taken out of the whole captured data and include all kinds of conditions. These frames are labeled by hand to mark the red traffic sign areas which should be segmented by the algorithm to create Ground Truth. Now these frames are segmented with both Mahalanobis Distance and simple thresholding and compared to the hand-labeled images. The number of correctly segmented pixels is called True Positives (TP). The number of missed pixels is assigned as False Negatives (FN). Pixels which are segmented but do not belong to a traffic sign are False Positives (FP). All other pixels not belonging to a sign and not segmented by the algorithm are True Negatives (TN) (see (Lazarevic-McManus et al., 2006)).

Table 2: Evaluation scheme.

		Ground Truth		FPR
		sign	non-sign	
segm. result	sign	TP	FP	
	non-sign	FN	TN	
		DTR		

In table 2 this is shown visually for a better understanding. The interesting results giving an impression how good the segmentation algorithm works is on the one hand the detection rate (DTR) and on the other hand the percentage of false positives (FPR) with:

$$DTR = \frac{TP}{TP + FN} \quad \text{and} \quad FPR = \frac{FP}{TP + FP} \quad (22)$$

One may not forget to see that the detection-rate refers to the number of pixels which are recognized, not on the actual number of traffic signs. A detection-rate of 80% means that 80% of all pixels belonging to a traffic sign are recognized which is completely enough for further processing. The 20% missed pixels are usually dispersed over the very border of the sign where color is deteriorating.

Table 3 shows the difference between the two presented methods. Of course the simple thresholding algorithm delivers a high percentage of all sign pixels due to a wide range of color-values covered. But that comes along with a relatively high false-positive rate (FPR). As the Covariance matrix describes the distribution of the sign pixels in the training data set much

better, the FPR is decreasing. However the detection-rate is comparably low due to many traffic sign pixels deteriorating in sunny conditions as mentioned in section 3.3.

Table 3: Performance evaluation results.

segmentation results (HSV color space)		
segmentation	DTR	FPR
simple thresholding	83.2%	78.9%
Mahalanobis Distance	51.0%	48.4%

Limited to cloudy scenes only the results would be different. Of course the γ -range can be increased to catch more sign pixels, anyhow this leads to a defined color subregion very similar to the section of the simple thresholding method as the modeled elliptically shaped subregion exceeds the color space limits and, within the color space, forms a body very alike the slice of pie.

7 CONCLUSIONS AND FUTURE WORKS

A preselection of subregions in the image containing traffic signs combined with an a-priori region of interest is a very effective approach to lower down processing time and hardware requirements for traffic sign recognition. In this paper the main issues about color features and how traffic signs in the real world look alike for a color imager were examined. The modeled color distribution proved to be a good approximation in cloudy and therefore consistent illumination conditions. Autumn leaves in the environment showed to be the hardest challenge to separate them from sign pixels as they partially reside in the same color subregion as traffic signs. Considered all illumination conditions, the higher processing costs for the Mahalanobis Distance of each pixel in the region of interest and the achieved detection rate a color based sign detection system can rely on the simple thresholding algorithm for an initial candidate detection. The larger amount of false positives can be tackled with further processing steps like gaussian filtering and tracking over several frames, as false positives usually do not appear similarly in every frame. Anyhow the lower sensitivity of color imagers compared to greyscale cameras, which is due to the RGB color filters on the imager chip, is still a limiting factor for the usage of color information at nighttime.

REFERENCES

- Gavrila, D.M. (1999). "Traffic Sign Recognition Revisited". In: Proc. of the 21st DAGM Symposium für Mustererkennung, pp. 86-93, Bonn, Germany
- Barnes, N., Zelinsky, A. (2004). "Real-time Radial Symmetry for Speed Sign Detection". In: IEEE Proc. of the Intelligent Vehicles Symposium 04, pp. 566-571, Parma, Italy
- Bahlmann, C., Zhu, Y., Ramesh, V., Pellkofer, M., Koehler, T. (2005). "A System for Traffic Sign Detection, Tracking, and Recognition Using Color, Shape, and Motion Information". In: IEEE Proc. of the Intelligent Vehicles Symposium 05, pp. 255-260
- de la Escalera, A., Armingol, J., Mata, M. (2003). "Traffic sign recognition and analysis for intelligent vehicles". In: Image and Vision Comput. Vol. 21, pp. 247-258
- Fang, C., Chen, S., Fuh, C. (2003). "Road-sign detection and tracking", In: IEEE Transactions on Vehicular Technology vol. 52, No. 5, pp. 1329-1341
- Siogkas, G. K., Dermatas, E. S. (2006). "Detection, Tracking and Classification of Road Signs in Adverse Conditions". In: IEEE MELECON 2006, Malaga, Spain
- Torresen, J., Bakke, J. W., Sekanina, L. (2004). "Efficient Recognition of Speed Limit Signs". In: IEEE Intelligent Transportation Systems Conference, Washington D.C., USA.
- Johansson, B. (2002) "Road Sign Recognition from a Moving Vehicle". Master's Thesis Report No. 56, University of Uppsala, Sweden
- Priese, L., Rehrmann, V., Schian, R., Lakmann, R. (1993). "Traffic Sign Recognition Based on Color Image Evaluation". In: IEEE Proc. of the Intelligent Vehicles Symposium, pp. 95-100
- Priese, L., Klieber, J., Lakmann, R., Rehrmann, V., Schian, R. (1994). "New Results on Traffic Sign Recognition". In: IEEE Proc. of the Intelligent Vehicles Symposium, pp. 249-254, Paris
- Priese, L., Rehrmann, V. (1998). "Fast and Robust Segmentation of Natural Color Scenes". In: 3rd Asian Conference on Computer Vision, pp. 598-606
- Fleyeh, H. (2006). "Shadow And Highlight Invariant Colour Segmentation For Traffic Signs". In: IEEE Conf. on Cybernetics and Intelligent Systems, Thailand
- Gonzales, R. C., Woods, R. E. (2002) "Digital Image Processing", 2nd ed.: Prentice Hall
- Shevell, S. K. (2003). "The Science of Color". 2nd ed.: Optical Society of America
- Bénallal, M., Meunier, J. (2003). "Real-time color segmentation of road signs". In: IEEE Canadian Conf. on Electrical and Computer Engineering, Montréal
- Vitabile, S., Gentile, A., Sorbello, F. (2002). "A neural network based automatic road sign recognizer". In: Intern. Joint Conf. on Neural Networks, Honolulu
- Lazarevic-McManus, N., Renno, J., Jones, G.A. (2006). "Performance evaluation in visual surveillance using the F-measure". In: Proc. of the 4th International Workshop On Video Surveillance And Sensor Networks, pp. 45-52, ACM Press, New York