ELSEVIER
Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia - Social and Behavioral Sciences 53 (2012) 811 – 821
Procedia
Social and Behavioral Sciences

SIIV - 5th International Congress - Sustainability of Road Infrastructures

Improved Traffic Signal Detection and Classification via Image Processing Algorithms

Leonardo Bruno[a], Giuseppe Parla[a], Clara Celauro[a]*

[a]*DICAM – Dipartimento di Ingegneria Civile, Ambientale, Aerospaziale, dei Materiali – Università degli Studi di Palermo - Italy*

Abstract

An image analysis technique for automatic traffic sign detection and classification is proposed. It makes it possible, after proper training, to detect, recognize and classify vertical road signs from video frames acquired on a moving vehicle equipped with cameras, as well as to identify anomalies with respect to road sign regulations (positioning and visibility).
The experimental results show that this technique allows one to correctly detect and classify almost all vertical signs and, mainly in extra-urban environment, it can be considered as highly reliable, apart from being really versatile and user-friendly for road inventory and road maintenance purposes.

Keywords: road signal; automatic detection; tracking and classification; maintenance.

* Corresponding author. Tel.: +39-091-23899716; fax: +39-091-487068.
E-mail address: clara.celauro@unipa.it

 doi:10.1016/j.sbspro.2012.09.930

1. Introduction

Improving traffic quality and safety cannot be achieved without correctly applying and maintaining road traffic signs, traffic signals and road markings. In fact, when road signs are not present, visible, and adequate with respect to the road and traffic conditions, meaning they do not comply with the requirements set by the regulations currently in force, late or insufficient perception of them contributes to an increase in the accident rate, as shown by road and traffic accident statistics.

As is well known, a road sign correctly fulfills its purpose (that is, to convey correct information and guidance to drivers) only if well-designed, produced and installed; otherwise, it may also become a danger in itself, since it may lead to uncertainties in drivers' behavior which is the main cause of road accidents, even with severe consequences.

Road administrations should constantly monitor the correct positioning and maintenance of road signs. In Italy, as in other countries, it is mandatory by law that road authorities and private companies in charge of managing a road have to maintain it, including road signs, in perfect efficiency.

Therefore, it is useful to develop and perfect tools that allow one to automatically recognize and classify the existing traffic signs along a road, with a low processing time. Moreover, developing this kind of tools is also important in view of improving intelligent driver support systems (IDSSs), advances in which afford a valid opportunity to aid mobility and road safety [1] [2].

In order to make a contribution to this, as concerns vertical road traffic signs, the potential afforded by branches of Image Analysis and Computer Vision is extremely interesting, since they deal with aspects related to both extraction of useful content from bi-dimensional images, as obtained by video frames taken from cameras mounted on instrumented vehicles.

Ample bibliographic research allowed us to evaluate the most suitable techniques for analyzing video images in order to detect traffic signs from a road scene image. In particular, amongst the possible image analysis methods, the one proposed in this paper proves to be adequate for the needs specified above, i.e. detection, recognition, tracking and classification of vertical road sign, even within the limits due to partial occlusion of the detected signs, as detailed in the following sections

Nomenclature

RGB	additive colour model based on the red, green, and blue light
HSV	cylindrical colour space based on the H (hue), S (saturation) and V (brightness Value) parameters
H	Hue (ranging from 0 to 360°)
S	Saturation (ranging from 0 to 100)
V	brightness Value (ranging from 0 to 100)
H_{TR}	segmented image of the red colour
H_{TB}	segmented images of the blue colour
W	thresholded image of the white colour

2. Road traffic sign categories

Road traffic signs can be classified into several categories, since they convey information based on their visual features: color and shape, above all. The choice of adequate colors is based on human perception for perceiving colors in different weather conditions (sunny, foggy, rainy and snowy conditions may occur, depending on the season), so that an adequate chromatic contrast between the sign and the surrounding environment is obtained, in every possible weather or lighting condition.

In Italy, three main categories of traffic signs are defined by law, according to their meaning: warning, prescription and direction, as depicted in Figure 1:

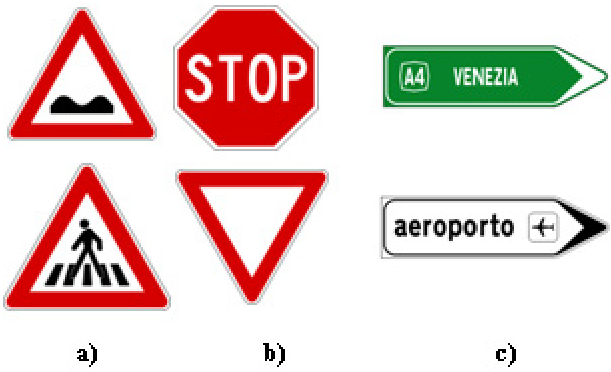


Fig. 1. Italian road signs (a) danger; (b) regulatory; (c) direction

- warning signs: they pre-advise people of a specific type of potential danger and warn drivers to adopt prudential behavior. They are usually triangular in shape (upward equilateral triangle);
- regulatory signs: they make a prescription known with which the road users must comply. They are subdivided into give-way signs; prohibition signs and obligation signs;
- signs that provide information about the location of either the driver or possible destinations.

Warning and prescription signs must be located along the right side of the road. A positioning distance is defined for placing warning and regulatory signs in order to allow drivers to make their behavior comply with what is indicated. These distances in Italy are set as in Table 1, based on the type of road.

Table 1. Positioning distances for different signals as a function of the type of road

Type of road	Warning signs	Regulatory signs
High speed roads	150 m	250 m
Low speed road	100 m	150 m
Other roads	50 m	80 m

The reason for misrecognizing traffic road signs when applying Image analysis technique is mainly due to the peculiarity of the images to be analyzed or to partial (or total) occlusion or deterioration of the sign itself, depending on the maintenance performed by the road authorities. Typical examples of occlusion of road signs are depicted in Figure 2.

Fig. 2. Typical road sign occlusion conditions

Road signs may not be efficient not only because of the presence of elements that may partially occlude them, but also because of deterioration of their reflective properties (and therefore of their colorimetric features), due to ageing or even to acts of vandalism. Furthermore, in the scene to be studied there may be many different objects whose color, similar to that of the road signs, may cause the lack of automatic detection (false negative) of an existing road sign, or also the detection of a nonexistent road sign (false positive). Other causes for erroneous detection are the following: incorrect positioning of the sign (rotation with respect to the perpendicular to the trajectory of the vehicle); unfavorable lighting conditions (low chromatic contrast with the background or insufficient lighting).

In the following section a few basic notions regarding colorimetry are mentioned and the basic principles of the proposed detection algorithm are discussed.

3. Road sign detection

The algorithm implemented is only able to detect and recognize traffic signs having the following features:

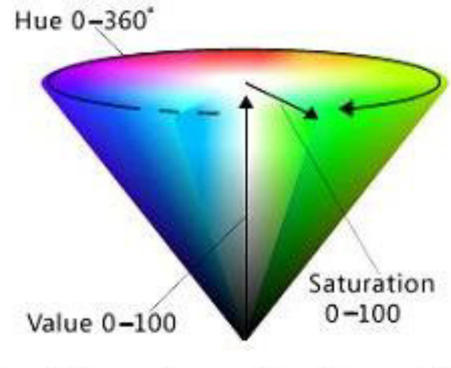
- • white areas surrounded by red areas (circular or triangular shape);
- • white areas surrounded by blue areas (circular or rectangular shape);
- • blue areas surrounded by red areas (circular or triangular shape)

By analogy with the human perceptual process, image analysis should be able to detect those areas that, taking into account some colorimetric feature, may represent a road sign [3] [4] [5] [6]. This is possible via image segmentation utilizing static or dynamic thresholding. According to Ballard [7], digital image segmentation is the process of subdividing an image into distinct regions that are homogeneous with respect to a certain quality, defined in relation to the aim of the segmentation itself. This operation represents the first step in order to distinguish the sub-regions in the image containing objects that are of interest (Region of Interest, ROI – in this case possible road sign) from the other parts of the image.

In any case, image partition has to satisfy the following criteria:

- • the detected regions should be as homogeneous as possible with respect to the chosen characteristics;
- • region boundaries should be compatible with the variation in the similarity characteristics adopted;
- • regions that are perceived as uniform should not be further divided into sub-regions;
- • small and complex regions should not be merged with adjacent ones.

As suggested by Vitabile et al. [8], in order to simulate the way a man experiences and describes his own perception of different colors, segmentation is carried out in the HSV color space, which is a three-dimensional space that can be visualized as a single cone, based on a cylindrical-coordinate representation of points in an RGB color model: Hue, Saturation (Y axis), and brightness Value (Z axis), as depicted in Figure 3.



Color	Angle (°)	Value
Red	0	0
Yellow	60	43
Green	120	85
Cyan	180	128
Blue	240	170
Magenta	300	213
Red	360	255

Fig. 3. HSV color space

According to De La Escalera et al. [9] these properties are the ones involved in human psycho-cognitive processes in order to obtain the feeling of the color of an object. De La Escalera suggests that it is possible to distinguish between red and blue objects in an image simply using information about hue (H) and saturation (S). In this way, one can avoid considering the brightness value (V), which may be subject to immediate and uncontrolled variations.

One should bear in mind, anyway, that low brightness values (that are typical of points close to the vertex of the cone in Figure 3) do not allow one to clearly distinguish the color of different objects. On the other hand, one knows that when H > 180°, the area of the red colors is entered, which corresponds to the gradations of this color in a traffic sign. This condition alone is not enough to ensure that all possible reds are detected (from the lightest to the brightest). A more accurate selection of the red areas of the image may be obtained if thresholding of the S plane is also carried out.

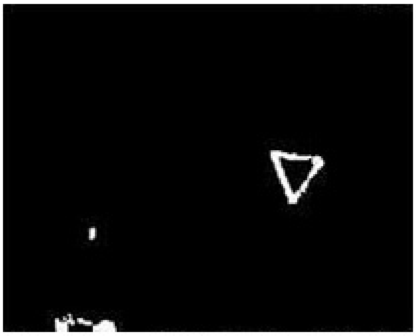
Application of the thresholding method proposed by Otsu [10] allows one to highlight those pixels whose intensity, in the S chromatic plane, is higher than a specific value (threshold) which is automatically changed with time, based on the special features of the image itself [11]. The rules used in order to identify those pixels having specific chromatic characteristics are summarized in Table 2.

Table 2. Segmentation rules

Segmented color	Color Space	Segmentation rules	
Red	HSV	H_{TR} if H>180 and S> T_{01}	
Blue	HSV	H_{TB} if 140>H>180 and S> T_{02}	
White	RGB	R_T =R>T_{03}	W=Max(R_T, G_T, B_T)
	RGB	G_T=G>T_{04}	
	RGB	B_T=B>T_{05}	

Application of the previous segmentation rules allowed us to obtain the binary images given in Figure 4, where the segmented images of the red and blue colors are denoted with H_{TR} e H_{TB}, respectively. Segmentation with respect to the white colors was also performed (see W in Figure 4), and it should be noticed that they typically represent image portions that do not belong to a traffic sign but have high contrast values with respect to the surrounding area. In order to reduce the level of noise in the image analyzed and to connect pixels that are close to each other but are not connected, a morphological mask for closure was applied, based on Minkowski summing and subtraction [12] [13]. This operation performs a logical intersection between binary matrices in order to exclude from the red and blue matrices those regions that do not belong to road signs, since they do not have inner white regions (closed and connected). A reduced area W_R* is obtained in this way, which comprises fewer areas that may not belong to a road sign than the initial matrices of the white colors, W.

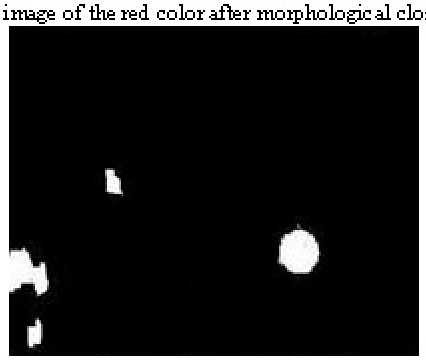
a) Initial image



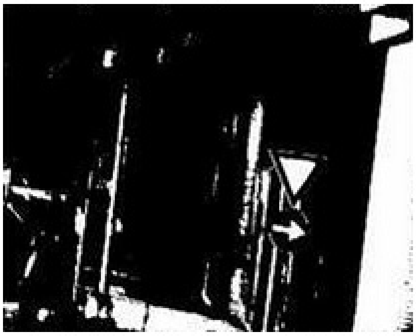
b) Thresholded image of the red color (H_{TR})

c) image of the red color after morphological closure (H_R)

d) thresholded image of the blue color (H_{TB})



e) image of the blue color after morphological closure (H_R)



f) thresholded image of the white color (W)

g) image of the white color after morphological closure (W_R*)

Fig. 4. Color segmentation operation

3.1. Area labeling

Usually, in order to detect an object contained in an image, a minimum bounding box is used: this is defined as the minimum rectangular box whose sides are tangent to the border of the object and fully enclose the object itself. This bounding box is easily localized by the coordinates associated with each of its vertices. Generally, the direction of two parallel sides of this rectangle is defined as coincident with the direction of the major axis. In order to complete the set of data that are necessary for defining the whole geometry and position of the object, the coordinates of the centroid (approximate barycentre of the region) as well as those of the central point of the front border of the bounding box are also used in this method. After doing this, it is verified that some geometric features of the box (height, H; width, W, area, A = H*W, ratio, R= H/W), when measured in pixels, fall within pre-defined limits. These limits are set by taking into account the optical characteristics of the camera used (projective geometry), starting from the known size of the real signs. Therefore, the analysis of the different objects found in the scene studied is carried out only if the measured sizes (W, H) of their bounding boxes are higher than a minimum value set, a priori, during the calibration of the method.

For those objects that satisfy the abovementioned condition, the ratio R = H/W is checked, since this is an invariant for road signs, as imposed by law. Thus in this method this R value was allowed to vary, due to possible perspective distortion or to a rotation of the road sign, in the range centered on the mandatory value set in the regulation. This further reduces the number of objects found in a scene that may be assimilated to a road sign. Then the remaining regions may be studied starting from their shapes, which should be similar, approximately, to those of the real signs. At this point it is possible to attribute a unique label (labeling) to the areas of interest, together with a code like the one proposed by Freeman [14], known as “chain code”. This code provides a codification of the border of a bounding box and therefore it is a compact way of representing the shape and position of each element in the image studied. After codification according to Freeman of the bounding box enclosing the element studied, a further phase of detection and tracking is carried out via a robust algorithm, based on the mutual correlation between a predefined image, saved a priori in memory, and the complete set of binary objects found in the video scene, as previously thresholded. The next section details this phase.

3.2. Correlation based approach

The proposed methodology that allows one to detect and track a predefined object in a scene is defined as object recognition based on template correlation. From a purely mathematical point of view, this process consists of searching for the existing correlation between a template and, generally speaking, a bigger image. This search is completed when the position having the highest coefficient of correlation between the two is found. In detail, when moving the template along the rows and columns of the area to be searched, a coefficient of similarity between the two images is calculated, at each position. This coefficient is calculated via the following equation:

$$c(i,j) = \frac{\sum_m \sum_n \left(A_{mn} - \overline{A}\right)\left(B_{mn} - \overline{B}\right)}{\sqrt{\sum_m \sum_n \left(A_{mn} - \overline{A}\right)^2 + \left(B_{mn} - \overline{B}\right)^2}} \tag{1}$$

where

A: matrix of the image

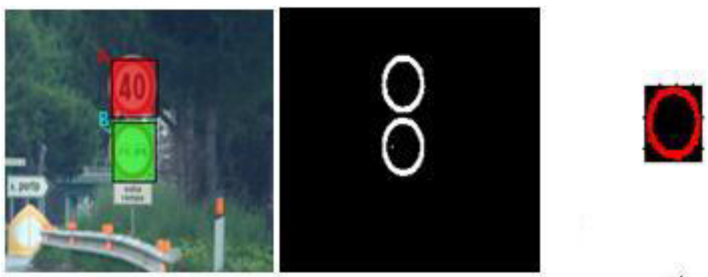
B: matrix of the template

$A_{nm:}$ average of A

B_{nm}: average of B.

Then the maximum value of c(i,j) is sought, and the corresponding position of the object having the maximum similarity to the template. This procedure, defined in the 1970s, is well-known in the technical literature [15] [16]. Nevertheless, determination of this coefficient of correlation is highly time-consuming, with a noticeable computational effort, so that it is seldom used for this kind of purpose.

In this connection, let I be the image to be analyzed and T the template to be searched. The required number of operations increases linearly with the number of templates as well as with the number of pixels in I and T (and therefore with I and T squared). Furthermore, in order properly to manage the colour variations in the image, it is important to operate on the edges of the image rather than on image intensity, and this increases the computational effort for performing this procedure. The main advantages of the template correlation method are its high reliability and precision: the correlation coefficient is independent of possible offset between the image and the template, as well as of each linear transformation of the latter. Another peculiarity of this method is precise knowledge of the object to be identified in the video frame to be analyzed. Most tracking methods are based on preliminary detection in a series of frames of the object to be tracked, based on segmentation of each image at the generic time t, and after doing this, detection of this object only depends on the indications obtained by binary mapping of the image after segmentation. This approach, on one hand, makes the search very easy to perform but, on the other hand, implies reduced precision and reliability of the algorithm that allows one to define a Boolean image of the object to be detected. Another aspect to be considered is that, since we are dealing with objects whose size, as shot by the camera, progressively reduces when approaching the vanishing point of the image, the method to be used at each time interval should adapt the initial shape of the object to the progressively deformed one as resulting from perspective distortion. Further, if tracking of the object is only based on the Boolean indication, possible fusions with binary objects spatially adjacent to the one searched may also affect the final result. This may happen not only in the first frame when the object enters the area studied, but also in each intermediate frame and, as can easily be understood, it may cause one to lose all trace of the object to be tracked. The proposed methodology is able to overcome these problems via introduction of adequate threshold values on the size of the objects (as clarified in the previous section), before their detection in the tracking phase. Besides, the proposed procedure is innovative since it performs template correlation between a binary template and a binary object which was segmented a priori (see Figure 5).



a) b) c)

Fig. 5. Phases of the tracking algorithm (a) initial image; (b) Boolean image; (c) template for comparison and correlation

4. Road traffic sign categories

In order to validate the proposed procedure, three different routes were analysed: two in an urban environment, the other one in a suburban area: scenes were collected on a moving vehicle, using a CCD camera. During the measuring campaign, weather conditions were good, in order to avoid alteration or degradation of contrast and color of traffic signs, due to weather effects. Each route was covered twice. In the urban areas, the maximum speed was set equal to 35 km/h; in the suburban area a limit of 70 km/h was chosen. The latter was set in order to acquire a sufficient amount of consecutive frames containing road signs, although with different sizes and positions. As far as the urban routes is concerned, the data regarding the total number of road signs, divided into visible, occluded and detected via the proposed algorithm, are summarized in Table 3. A preliminary survey on the site allowed us to distinguish, based on direct visual examination, between perfectly undamaged and visible signs and undetectable signs (due to occlusion, deformation, etc). Among the occluded signs were those that were perceived for too short a time – a few fractions of a second – during the normal progress of the vehicle at 35 km/h. Some road sings, totally or partially occluded, are actually perceivable (visible) only if the distance between the vehicle and the road signs is really small. It is to be considered that these signs may be recognized if the minimum number of frames set for complete detection is reduced.

Table 3. Details of the results obtained in urban areas

	Urban route A		Urban route B	
	Pass #1	Pass #2	Pass #1	Pass #2
Total signs	166	166	150	149
Visible signs	131	131	115	114
Detected signs	129	129	113	112
Detected (%)	98.47	98.47	98.26	98.25
Average (%)	98.47		98.26	
Occluded signs	35	35	35	35

The perfect agreement between the results of the two different passes of each route proves the good repeatability of the procedure as well as the good quality of the results obtained via image analysis.

The same procedure applied in suburban area is even more reliable, due to both the lower complexity of the images to be analyzed and the higher contrast between the object searched and its background. This can be clearly observed from the data summarized in Table 4.

Table 4 Details of the results obtained in the suburban area

	Suburban route	
	Pass #1	Pass #2
Total signs	243	243
Visible signs	235	235
Detected signs	230	230
Detected (%)	97.87	97.87
Average (%)	97.87	
Occluded signs	8	8

As expected, it can be noticed that the percentage of the partially or totally occluded signs is, in the suburban area, much lower than that obtained in the urban areas. The percentage of correctly detected signs is also higher, when working in the suburban area.

5. Conclusions

An algorithm for detection, classification and recognition of road sings is presented, specifically developed for defining an automated procedure useful for both road inventory purposes and maintenance of existing road networks. This procedure, when the user is appropriately instructed, is able to highlight differences in the signs with respect to the requirements in terms of position and size, as set by the regulation.

The results obtained show that more than 20% of the signs detected do not comply with the requirements of the regulation and that more than 98% of the signs that are correctly positioned are actually detected via image analysis. In particular, when operating in an urban area the proposed algorithm allows one to detect, and automatically classify, almost all those road signs that are visible and correctly positioned (no rotation or occlusion). In any case, occluded, damaged and uncoloured signs are of concern. A possible improvement of the performances offered by the proposed algorithm may deal with detection of those roads signs that are not perfectly visible. This problem is less serious when operating in a suburban area, where the signs are more visible and comply with the legal requirements.

Currently, the development of the proposed algorithm is in progress and deals with classification of detected signs as well as with determination of size via template correlation for comparison with the requirements.

References

[1] Gil-Jiménez P., Lafuente-Arroyo S., Gomez-Moreno H., López Ferreras F. & Maldonado-Bascón S. (2005) Traffic sign shape classification evaluation II: FFT applied to the signature of blobs. Proceedings of IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems., Las Vegas, NV, 607 – 612.

[2] Maldonado-Bascón M., Lafuente-Arroyo S., Gil-Jiménez P., Gomez-Moreno H. & López Ferreras F. (2007) Road-Sign Detection and Recognition Based on Support Vector Machines, IEEE Transactions on Intelligent Transportation Systems,. 8:2, 264 - 278.

[3] Benallal M. & Meunier J. (2003). Real-time color segmentation of road signs. Proceedings of CCECE 2003 Canadian Conference on Electrical and Computer Engineering. IEEE, 3, 1823 - 1826 .

[4] Arthur W. R &, Eric H. G. (1997). Color segmentation in the HSI color space using the k-means algorithm. In Nonlinear Image Processing VIII, E. R. Dougherty & J. T. Astola Eds, Proceedings of SPIE, 3026, (pp. 143 - 154).

[5] Wu J., Hong Y. & Andrew N. (1994).Color image segmentation using fuzzy clustering and supervised learning, Journal of electronic imaging, 3:4, 397- 403.

[6] Mehmet C. & Uijt de Haag, M. (1998), Optimal thresholding for color images. In Nonlinear Image Processing IX, E. R. Dougherty & J. T. Astola Eds, Proceedings of SPIE, 3304, (pp. 250-259).

[7] Ballard D. H. & Brown C. M. (1982) Computer Vision, Prentice-Hall Inc., New Jersey.

[8] Vitabile. S., Pollacia. G, Pilato. G. & Sorbello. F. (2001) Road signs recognition using a dynamic pixel aggregation technique in the HSV color space. Proceedings of the 11th International Conference on image analysis and processing, Palermo, Italy, 572 - 577.

[9] De La Escalera A., Armingol j. M A. & Mata M. (2003). Traffic sign recognition and analysis for intelligent vehicles. Image and vision computing, 21, 247 - 258.

[10] Otsu N. (1979). A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics, 9:1, 62 - 66.

[11] Bruno L., Parla G. & Celauro C. (2012) Image analysis for detecting aggregate gradation in asphalt mixture from planar images. Construction and Building Materials. 28:1, 21 – 30.

[12] Ghosh P. K. & Deguchi K. (2009) Mathematics of shape description: a morphological approach to image processing and computer graphics, Wiley-Blackwell, Singapore.

[13] Serra J. (1982). Image analysis and mathematical morphology, Academic Press, Inc. Orlando, FL, US.

[14] Freeman H, (1974) Computer Processing of Line-Drawing Images, ACM Computing Surveys, 6:1, 57 - 97.

[15] Rosenfeld A. (1969). Picture processing by computer, New York: Academic Press.

[16] Pratt W, K. (2001). Digital Image Processing, John Wiley & Sons, Inc., New York, New Jersey.