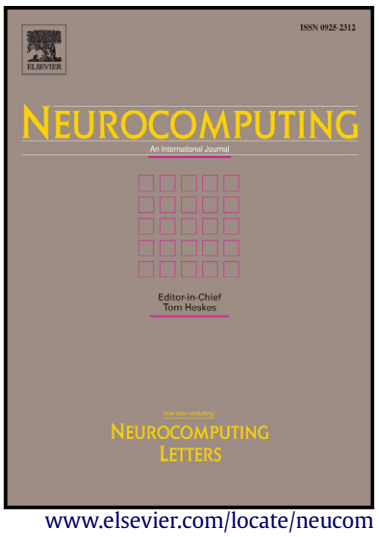
Author's Accepted Manuscript



www.elsevier.com/locate/neucom

Traffic sign segmentation and classification using statistical learning methods

J.M. Lillo-Castellano , I. Mora-Jiménez , C. Figuera-Pozuelo , J.L. Rojo-Álvarez

PII: S0925-2312(14)01559-8
DOI: http://dx.doi.org/10.1016/j.neucom.2014.11.026
Reference: NEUCOM14909

To appear in: *Neurocomputing*

Received date: 30 July 2014
Revised date: 19 October 2014
Accepted date: 14 November 2014

Cite this article as: J.M. Lillo-Castellano , I. Mora-Jiménez , C. Figuera-Pozuelo , J.L. Rojo-Álvarez , Traffic sign segmentation and classification using statistical learning methods, *Neurocomputing,* http://dx.doi.org/10.1016/j.neucom.2014.11.026

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Traffic Sign Segmentation and Classification using Statistical Learning Methods

J.M. Lillo-Castellano[a,*], I. Mora-Jiménez[a], C. Figuera-Pozuelo[a], J.L. Rojo-Álvarez[a,b]

[a] *Signal Theory and Communications Department, Rey Juan Carlos University, Camino del Molino s/n, 28943, Fuenlabrada, Madrid, Spain*
[b] *Electric and Electronic Department, Universidad de las Fuerzas Armadas ESPE, Av. Gral. Rumiñahui s/n, Sangolquí, Ecuador*

Accepted manuscript

Abstract

Traffic signs are an essential part of any circulation system, and failure detection by the driver may significantly increase the accident risk. Currently, automatic traffic sign detection systems still have some performance limitations, specially for achromatic signs and variable lighting conditions. In this work, we propose an automatic traffic-sign detection method capable of detecting both chromatic and achromatic signs, while taking into account rotations, scale changes, shifts, partial deformations, and shadows. The proposed system is divided into three stages: (1) Segmentation of chromatic and achromatic scene elements using $L^*a^*b^*$ and HSI spaces, where two machine learning techniques (k-Nearest Neighbors and Support Vector Machines), are benchmarked; (2) Post-processing in order to discard non-interest regions, to connect fragmented signs, and to separate signs located at the same post; (3) Sign-shape classification by using Fourier Descriptors, which yield significant advantage in comparison to other contour-based methods, and subsequent shape recognition with machine learning techniques. Experiments with two databases of real-world images captured with different cameras yielded a sign detection rate of about 97% with a false alarm rate between 3% and 4%, depending on the database.

*Corresponding author
Email addresses: josemaria.lillo@urjc.es (J.M. Lillo-Castellano), inmaculada.mora@urjc.es (I. Mora-Jiménez), carlos.figuera@urjc.es (C. Figuera-Pozuelo), joseluis.rojo@urjc.es (J.L. Rojo-Álvarez)

ACCEPTED MANUSCRIPT

Our method can be readily used for maintenance, inventory, or driver support system applications.

Keywords: Driver Support System; Traffic Sign Detection; Chromatic and Achromatic Segmentation; Fourier Descriptors; Classification; Machine Learning Techniques.

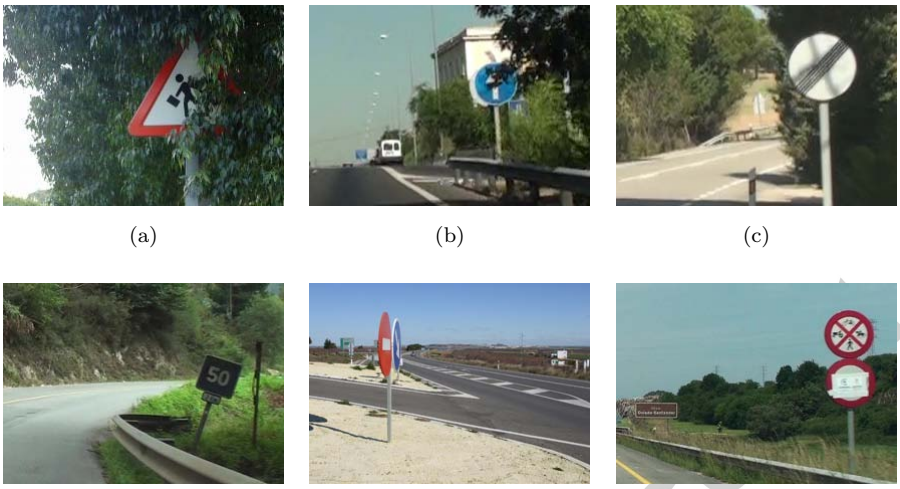
1. Introduction

Traffic signs constitute an essential part of any circulation system to control and guide traffic and to favor road safety [1], with a twofold role: conveniently regulating traffic and reporting to pedestrians and drivers on different aspects about road circulation. Nowadays, automatic traffic sign detection and recognition systems are of special interest in many applications, such as intelligent vehicles development and road maintenance. Regarding the first one, on-board automatic traffic sign detection systems aim to help users to detect and interpret traffic signs. Examples of industrial developments in this field are the traffic sign recognition module [2] used in Opel Eye® [3], or the method for detecting and recognizing traffic signs [4] used in Mercedes-Benz Traffic Sign Assist® [5]. Regarding road maintenance, appropriate positioning and maintenance of traffic signs clearly improve road safety, being supervised by authorities with regular inventory campaigns.

However, despite the vast amount of research recently conducted for automatic traffic sign detection and recognition, current traffic sign inventory and monitoring are still mainly carried out manually by an operator visualizing a video recording, and checking the presence, position and status of each traffic sign. This tedious task requires intense concentration during a long time, and errors can be originated by the operator's fatigue or by poor visibility conditions (as shown in Fig. 1). An automatic system can be designed for surpassing these difficulties, as well as for providing a significant cost reduction in the inventory process. The interest of traffic sign automatic identification is further supported by competitive challenges proposed by scientific and technical

Accepted manuscript

ACCEPTED MANUSCRIPT



(a) (b) (c)

(d) (e) (f)

Figure 1: Examples of challenging images for a traffic sign detection system: (a,b) Partial hiding; (c) Lighting variation; (d) 2D rotation; (e) 3D rotation; (f) Other factors.

societies, see e.g., [6, 7].

Recent research works usually split the identification process in two stages, namely, detection and classification (recognition), which are designed using a representative set of images or videos (training set). Image segmentation techniques are used for the detection stage, and different approaches have been proposed depending on the type of image, i.e., true-color or gray-level. Regarding true-color images, there are two approaches, either working with the standard RGB color space used by digital cameras, or performing a deeper analysis of color information using other spaces for separating color and intensity information (such as HSI, HSV, YIQ, YUV, Luv or $L^*a^*b^*$) [8]. Many authors have addressed the segmentation by thresholding on RGB images, either at pixel level [9], or with more elaborated schemes, such as preprocessing with a Simple Vector Filter algorithm before thresholding [10].

The main drawback of the RGB space is its high sensitivity to lighting changes, which hampers segmentation in scenes with excessive or insufficient light. For this reason, many works use color spaces that are theoretically more

Accepted manuscript

ACCEPTED MANUSCRIPT

robust to lighting conditions than RGB. In this way, a non linear transformation of H and S components of the HSI color space was carried out in [11], with subsequent thresholding segmentation on the transformed components. Segmentation in [12, 13] was carried out in two stages: Chromatic analysis, where components H and S were used to segment signs with predominant chromatic colors; And achromatic analysis, where RGB thresholding was performed for signs with prevalent achromatic colors. Other works resorted to other color spaces that allow an independent control of chromatic and achromatic information as well. The a^*b^* components of the $L^*a^*b^*$ space were used to extract features by using a Gabor filter in [14], which were subsequently used to detect traffic signs. In [15, 16], the YUV space was used for thresholding segmentation. Also, the $YCrCb$ space has been used for segmentation by using a dynamic thresholding scheme [17]. Whether to use the RGB space or other spaces separating chromatic and achromatic information remains controversial. On the one hand, the review in [18] evaluated and benchmarked several thresholding segmentation methods. Authors concluded that the best ones were those using normalization with respect to illumination, such as normalized RGB or *Ohta Normalized*, and that the use of HSI or YUV spaces did not provide a significant advantage. On the other hand, other authors [13, 14] proposed the use of HSI or YUV spaces using more elaborated segmentation schemes, improving the results provided by simpler thresholding-based methods.

Systems dealing with gray-level images are mainly focused on edges detection and their subsequent analysis. A shape-based approach for de-restriction signs detection was presented in [19], which used a black band detector to highlight regions of interest (ROI). A set of histograms of oriented gradients features were used in [20] to design a classifier with a boosting approach to detect pedestrians and traffic signs. A transformation for angle vertex and bisector detection was used in [21] to implement a gradient geometric model to detect triangular signs. In [22], a restricted Hough transform applied on the image contours was proposed as a traffic sign detection method. Nevertheless, detection techniques based on image-gradients and object-edges are very sensitive to noise and com-

Accepted manuscript

ACCEPTED MANUSCRIPT

putationally expensive, requiring in most cases a complex preprocessing stage. In order to improve the detection stage, several works proposed to use consecutive video frames to track traffic signs and to reduce false alarm and missing rates by using Kalman filtering [23, 24].

Three conclusions arise from the above review: (1) Separating chromatic and achromatic segmentation by using HSI or $L^*a^*b^*$ color spaces seems to improve segmentation performance; (2) Achromatic segmentation is a difficult task, addressed by several works but with limited success; And (3) the potential of $L^*a^*b^*$ and HSI spaces to separately segment chromatic and achromatic elements has not been fully exploited yet. These conclusions motivate the search of advanced segmentation techniques based on these spaces.

The problem of traffic sign recognition has been often tackled with matching techniques. As an example, a distortion-invariant Fringe-adjusted joint transform correlation technique was used in [14] to find correlation peaks among segmented regions and a set of patterns extracted from different traffic signs. Also, a dissimilarity measurement was used in [25] to classify the sign by matching its color with a set of patterns. Machine learning techniques have been applied to the traffic sign recognition problem too. In this way, a combination of Convolutional Neural Networks and Multilayer Perceptron was applied in [26] on images normalized by a contrast-limited adaptive histogram equalization, which achieved good performance for German traffic signs recognition. In [17], a hybrid classifier composed of Support Vector Machines (SVM) and Naive Bayes was fed with features provided by a Gabor filter bank. Other works [10, 12, 13, 18, 27, 28, 29] also used SVM as a classifier, taking into account different features such as shape signature or Pseudo-Zernike moments. Genetic Algorithms (GA) have been less used, probably due to their high computational cost, and their application has been focused on adapting the use of certain features with other machine learning schemes. A GA was applied in [11], followed by a two-layer neural network, according to the Adaptive Resonance Theory paradigm for classification. In [30], affine transformation coefficients were used as GA parameters for sign detection. Regardless the classification al-

Accepted manuscript

ACCEPTED MANUSCRIPT

gorithm, the previously mentioned features are highly dependent on traffic sign scaling, translation, or rotation. In [31], a robust contour-based descriptor was presented, so-called *Shape Context*, which describes each contour pixel through a coarse bidimensional histogram characterizing the edge distribution in its surrounding region. The Shape Context descriptor is scale and rotation invariant, and has been successfully applied to describe cartoon characters, signs and other objects [32, 33, 34, 35]. A simpler descriptor also invariant to shift, scale and rotation is the *Fourier Descriptor* (FD), which has been successfully applied in different scenarios [36, 37]. Therefore, our aim was to design a high-quality sign shape classifier, based on the robust classification capacity of the SVM, and on the promising performance of FDs.

We present here a new automatic method to separately detect chromatic and achromatic signs in images taken in realistic scenarios. The proposed method achieves sign shape classification, and it is robust to sign rotations, scale changes, translations, shadows, and minor deformations. Our procedure is structured in three stages. First, the image is segmented using the $L^*a^*b^*$ and HSI spaces, in order to separate the chromatic and achromatic traffic sign elements. For the chromatic segmentation, two machine learning techniques (k-Nearest Neighbors -k-NN- and SVM) were benchmarked, and an additional HSI thresholding was used in the achromatic branch. Second, a post-processing stage improves the segmentation result by filtering out non-interest regions, which consists of merging fragmented signs and separating signs located at the same post (co-located signs). The later task is carried out by combining eigenvector decomposition and maxima dynamics. Third, a scheme of parallel SVMs is used to classify the shape of the segmented regions and to identify traffic signs. Because of FDs invariance with respect to scaling, shift, and rotation, they are used as the SVM input features. To sum up, the main contributions of the proposed system are: (1) A high performance segmentation scheme, based on $L^*a^*b^*$ and HSI spaces, which consists of chromatic and achromatic sub-procedures; (2) A novel algorithm for subsequent separation of co-located signs; And (3) a shift, scale, and rotation invariant shape classification procedure, based on the use of FDs

Accepted manuscript

ACCEPTED MANUSCRIPT

and SVMs, which is able to classify different complex shapes.

The paper is organized as follows. Next section presents an overview of the proposed methodology. Stages for segmentation, post-processing and detection are detailed in Sections 3, 4 and 5, respectively. Database description and experimental results are exposed in Section 6. Finally, main conclusions are drawn in Section 7.

2. System Overview

The proposed system aims to detect traffic signs with two assumptions: First, signs must be totally inside the image frame (not cut by the image borders); And second, they must be located at a suitable distance (distinguishable by the naked eye). The traffic signs in our experiments corresponded to those used by the Traffic Department of Spain, which have very particular and distinctive colors and shapes (see examples in Fig. 2). As sketched in Fig. 3, the procedure has three stages:

1. *S*egmentation: It consists of two procedures, one suitable for chromatic traffic signs and the another one for achromatic ones. Machine learning techniques and thresholding of the saturation component are applied to provide 5 binary images, one for each color of interest (blue, green, yellow and red) and another one for the white color (achromatic signs, white image). Foreground pixels correspond to ROIs, which are segmented as potential traffic signs (one pixel can only be associated to one of the 5 binary images).
2. *P*ost-processing: To reduce the number of ROIs not corresponding to traffic signs, a post-processing procedure is applied, which is different for binary images coming from the chromatic (blue, green, yellow and red) and achromatic (white) segmentation stage.
3. *C*lassification: ROIs boundary characterization is performed by FDs, which are used as input features for sign shape classification with SVMs.

Accepted manuscript

ACCEPTED MANUSCRIPT

These stages are described in detail in next sections. Outcome of every stage is illustrated with images from an inventory campaign, as described in Section 6.1.

3. Segmentation Stage

The goal of this stage is to extract those ROIs which are likely to be traffic signs. Segmentation algorithms usually consider criteria of similarity or connectivity. In our case, traffic signs usually stand out from their surroundings mainly due to their characteristic colors, and segmentation is tackled by a pixel classification process based on color criteria. Since some traffic signs are black & white, chromatic and achromatic image components are independently treated, and the $L^*a^*b^*$ and HSI color spaces are considered for this purpose. The $L^*a^*b^*$ space is used to identify traffic signs with chromatic colors, because these colors are very discernible in the a^*b^* space and color differences are better quantified using Euclidean distances [38]. Since white color is not clearly distinguishable in that space, the saturation component S of the HSI space is considered for this case too.

3.1. Chromatic Segmentation

Taking into account colors of interest, our procedure considers five traffic sign classes, which are grouped into two sets: (1) Interest-color set, which corresponds to the characteristic colors of traffic signs (blue, green, yellow and red); And (2) remaining-color set, which is associated to other colors.

Let us consider a set of uniform color patches, extracted from images with traffic signs in different lighting conditions. Patches are transformed from RGB to $L^*a^*b^*$ components, and just the a^*b^* components are considered. Each patch is labeled with its associated color class. In order to design a color classifier (segmentation machine), pixels from these patches and associated color class label are used to design a supervised color classifier (segmentation machine). Two sets are created: training set and validation set, with 80%-20% of the

Accepted manuscript

ACCEPTED MANUSCRIPT

(a) (b) (c) (d) (e) (f) (g)

Figure 2: Examples of traffic signs used by the Traffic Department of Spain: (a) Warning signs with a red rim; (b) Priority; (c) Restriction (red); (d) Regulatory (blue); (e) De-restriction (black & white); (f,g) Guide and informational.

ACCEPTED MANUSCRIPT

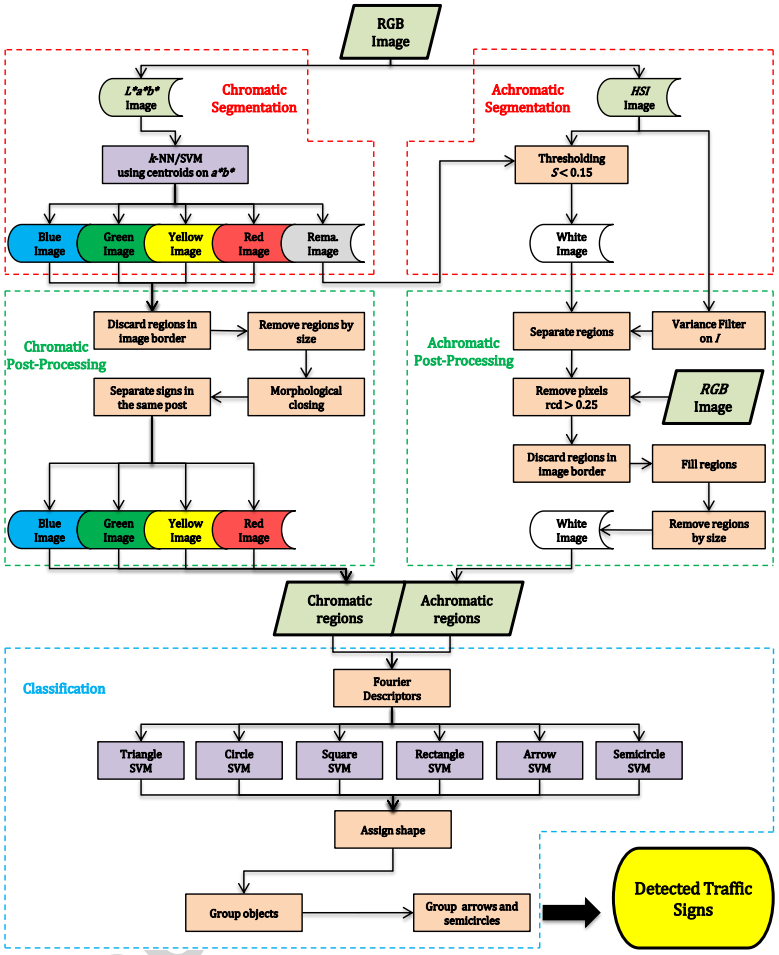


Figure 3: Diagram of the proposed procedure.

pixels in patches. The a^*b^* distribution of the training set is shown in Fig. 4a, where each point is associated with a pixel, namely, red, green, yellow, blue, and black (remaining-color set).

A k_m-means clustering algorithm [39] was applied on the a^*b^* space to obtain a reduced set of representative centroids for each class. The k_m-means algorithm is a well-known iterative method which clusters points into k_m groups, each one represented by one centroid, with k_m being a design parameter. The algorithm was performed 5 times (one per class). For simplicity, in this work the same

ACCEPTED MANUSCRIPT

S	Meaning (SVM context)	S	Meaning (other)
N	Number of training instances	k_m	Number of centroids for k_m-means
$\mathbf{x}_i$	i-th input instance	k	Number of neighbors for k-NN
l_i	Label for the i-th input instance	c	Interest-color variable
$\mathbf{w}$	Hyperplane coefficients	S_{TH}	Threshold for S component
b	Bias term of the hyperplane	(x, y)	Pixel coordinates
$\|\| \cdot \|\|$	Norm of a vector	N_p	Number of pixels of the ROI
C	Regularization parameter	P	ROI pixel coordinates matrix
ξ	Slack variable	C_P	Covariance matrix of P
$\phi(\cdot)$	Nonlinear mapping	θ_r	ROI tilt angle
α	Lagrange multiplier	M	Horizontal projection maximum
$\kappa(\cdot, \cdot)$	Kernel function	rcd_p	Relative color difference for pixel p
σ	Kernel width parameter	N_c	Number of boundary pixels
		C_u	Fourier Descriptors

Table 1: Summary of the notation for the most significant symbols (**S**): SVM-related (left), and other (right).

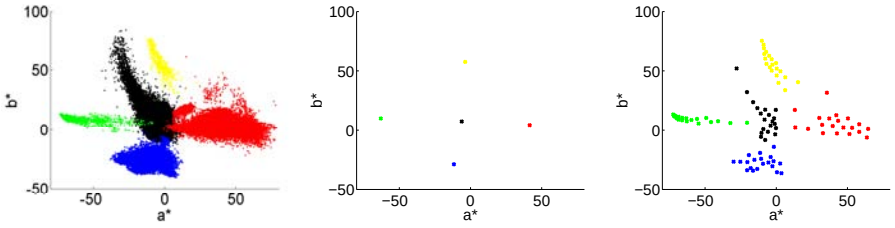
k_m value was used for each color class, and its value was searched in the range [1, 20]. Thus, a total number of centroids between 5 ($k_m = 1$ centroid per class) and 100 ($k_m = 20$ centroids per class) was used. Figs. 4b and 4c show the centroids distribution for $k_m = 1$ and $k_m = 20$, respectively. Note that high k_m values entail more computational cost but provide a more detailed representation of the a^*b^* color distribution. Since these centroids are used to design subsequent classifiers for color segmentation, a trade-off between representativity and computational cost is necessary (explained later in this section). For easier readability, Table 1 summarizes the symbol notation used in the paper.

Then, the classification process (segmentation) was benchmarked with two machine learning schemes: SVM, which has exhibited excellent performance in this kind of problems [18]; and voting k-NN, which is an extremely simple scheme with very low computational cost when the number of instances and the space dimension are low [39].

Even though the SVM basis were developed by Vapnik in 1964 [40], it was not until the 90's when they reached great popularity. In a linearly separable two-

Accepted manuscript

ACCEPTED MANUSCRIPT



(a) (b) (c)

Figure 4: Distribution of the training set in the a^*b^* space: (a) Whole training set distribution; (b)-(c) Centroids per class provided by the k_m-means algorithm with $k_m = 1$ and $k_m = 20$.

class classification process, the SVM basic idea is to find the decision boundary maximizing the *margin*, defined as the distance between the boundary and its closest instance [41]. This idea can be extended to more complex situations where classes are not linearly separable. Thus, given a set of N labeled training instances $\{(\mathbf{x}_i, l_i)\}_{i=1}^{N}$,where $\mathbf{x}_i \in \mathbb{R}^d$ and $l_i \in \{+1, -1\}$, the SVM minimizes the following *primal problem*:

$$\min_{\mathbf{w},b} \left(\frac{1}{2} \left\| \mathbf{w} \right\|^2 + C \sum_{i=1}^{N} \xi_i \right) \tag{1}$$

with constraints,

$$\begin{aligned} l_i(\phi(\mathbf{x}_i)^T \mathbf{w} + b) &\geq 1 - \xi_i \qquad \forall i = 1, \ldots, N \\ \xi_i &\geq 0 \qquad \forall i = 1, \ldots, N \end{aligned} \tag{2}$$

where $\phi(\cdot)$ is a (possibly nonlinear) function mapping the training pattern to a higher dimensional space where classes are separated by a linear hyperplane; $\mathbf{w}$ and b are parameters defining the linear classifier (hyperplane in the transformed space); the sum of slack variables ξ_i is a measure of the misclassification error; and regularization parameter C controls the trade-off between training error and margin size (given by the inverse of $||\mathbf{w}||$).

If the Karush-Kuhn-Tucker theorem [42] is used to solve the *primal problem* (1), the separating hyperplane is obtained as:

$$\mathbf{w} = \sum_{i=1}^{N} \alpha_i l_i \phi(\mathbf{x}_i) \tag{3}$$

Accepted manuscript

ACCEPTED MANUSCRIPT

where α_i are the Lagrange multipliers. Some of these multipliers may be zero, providing a sparse solution only depending on a training subset known as support vectors (instances with associated α_i greater than zero). Then, the solution is obtained by computing the Lagrange multipliers using the dual formulation of (1), that is:

$$\max_{\alpha_i} \left(\sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j l_i l_j \kappa(\mathbf{x}_i, \mathbf{x}_j) \right) \tag{4}$$

subject to the constraints,

$$\begin{aligned} & 0 \leq \alpha_i \leq C \qquad \forall\, i = 1, \ldots, N \\ & \sum_{i=1}^{N} \alpha_i l_i = 0 \end{aligned} \tag{5}$$

where $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \cdot \phi(\mathbf{x}_j)$ is called the kernel function and computes the scalar product of two instances in the transformed space as a nonlinear function in the input space [40]. In this work, the Radial Basis Function (RBF) kernel with width parameter σ is used:

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}} \tag{6}$$

Thus, for any testing instance $\mathbf{x}$, the decision function is given by:

$$f(\mathbf{x}) = sign \left(\sum_{i=1}^{N} \alpha_i l_i \kappa(\mathbf{x}, \mathbf{x}_i) + b \right) \tag{7}$$

In order to apply the SVM scheme to our problem, the original 5-class classifier was implemented with four parallel two-class SVM classifiers. Each classifier aimed to detect pixels belonging to interest-color c, with $c \in \{$blue, green, yellow, red$\}$. The training set was $\{(\mathbf{x}_i^{(c)}, l_i^{(c)})\}_{i=1}^{N}$ for the c-color classifier, with $\mathbf{x}_i \equiv (a_i^*, b_i^*)$ and a^*b^* the coordinates of each of the N centroids obtained by the k_m-means algorithm; $l_i^{(c)} \in \{+1, -1\}$ is the label of the i-th centroid for the c-color classifier, set to $+1$ if the i-th centroid color is c, and -1 otherwise. The same number of training instances (pixels) was considered to design each classifier.

Accepted manuscript

ACCEPTED MANUSCRIPT

A key factor in SVM performance is the appropriate setting of free parameters C and σ. For this purpose, a grid-search with 20 values was conducted for $C \in [1, 100]$ and $\sigma \in [1, 20]$, selecting the pair (C, σ) yielding the best accuracy on the validation set. After designing the four SVM classifiers, four outcomes were provided: (1) If just one SVM provided a $+1$ outcome, then the pattern was assigned to the color class of that SVM; (2) If more than one SVM provided a $+1$ outcome, then the instance was assigned to the color with higher occurrence rate in the training set (in order of red, blue, yellow and green); And (3) if the four outcomes were -1, then the instance was not assigned to any interest-color class. The best accuracy was obtained with $k_m = 15$, $C = 5$ and $\sigma = 2$.

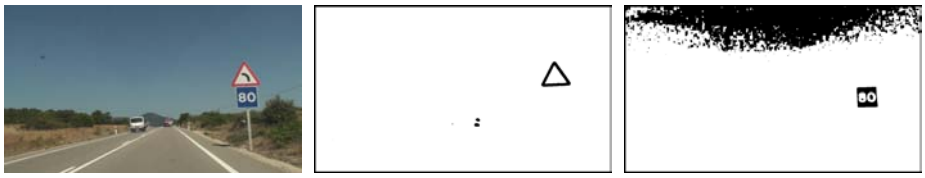
The second segmentation approach implemented a voting k-NN scheme, following the idea that neighbor instances have similar posterior probabilities. Under this assumption, any instance is assigned to the majority class among the k nearest instances (in our case, the k nearest centroids). It can be shown that when the number of instances tends to infinity, this classifier tends to the Bayes optimal classifier [39]. The best value of parameter k depends on the size and dimension of the training set [39]. In this work, the value of k was explored for each number of clusters k_m. The best accuracy was obtained for the range of $k_m \in [1, 6]$ and $k = 1$, and the configuration associated to the lowest computational burden one was selected, i.e., $k_m = 1$ and $k = 1$.

Even when both SVM and voting k-NN schemes provided a low error rate, k-NN outperformed SVM with lower computational burden in our scenario. For this reason, the k-NN scheme was used in our system from now on.

After this procedure, five mutually excluding binary images were obtained: Four images were associated to interest-colors and corresponded to chromatic segmentation, while the remaining binary image was used for achromatic segmentation. As an example, Fig. 5 shows the red and blue segmentation (in black, ROIs considered as potential traffic signs) obtained with $k_m = 1$ centroids and 1-NN for an image with two traffic signs (blue rectangle and red triangle). Note that both signs have been correctly segmented, but part of the track and the

Accepted manuscript

ACCEPTED MANUSCRIPT



(a) (b) (c)

Figure 5: Chromatic segmentation with $k_m = 1$ and 1-NN: (a) Original RGB image; Segmented regions (in black) for red class (b) and blue class (c).

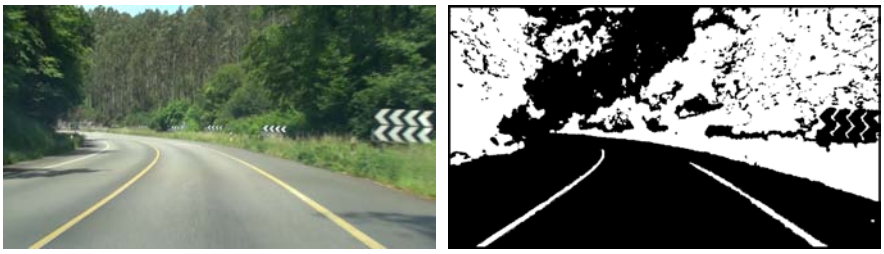
sky have been spuriously obtained too. The undesired regions will be discarded in a posterior stage.

3.2. Achromatic Segmentation

In theory, pure black & white colors do not have chromatic information and should have low a^* and b^* values. Indeed, in Fig. 4a the a^*b^* values associated to black & white colors are in the center of the a^*b^* space. However, since slight changes in lighting conditions or degradation of the traffic sign can originate significant changes in those components, we have that white traffic signs are not easily distinguishable in the a^*b^* space. In order to segment these signs, we propose to consider the saturation component S of the HSI space, since white areas present low saturation and a simple threshold can be used for segmenting them. In order to minimize the number of undetected signs, a threshold of $S_{TH} = 0.15$ (experimentally found by visual analysis of S histograms, with $S_{TH} \in [0, 1]$) was set to carry out the achromatic segmentation on the remaining-color image obtained from the chromatic segmentation process. Figure 6 shows the segmentation obtained when this process was applied to an image with a curve indication panel. Note that very large areas like the asphalt and some parts of the vegetation are also segmented, and hence a post-processing stage is further required.

Accepted manuscript

ACCEPTED MANUSCRIPT



(a) (b)

Figure 6: Example of the achromatic segmentation process: (a) Original image; (b) Segmented regions (in black) after S thresholding ($S_{TH} = 0.15$) on the remaining-color segmented image.

4. Post-Processing Stage

The aim of the post-processing stage is to refine previous segmentation by discarding certain regions and correcting undesired effects. In this setting, separate procedures for the results of chromatic and achromatic segmentations are proposed, which are next described in detail (see an overview in Figure 3).

4.1. Chromatic Post-Processing

Three concrete goals are pursued here: (O1) discarding segmented regions which are not traffic signs; (O2) merging regions belonging to the same sign but fragmented during segmentation; and (O3) separating regions belonging to different co-located signs. For a better understanding, each goal is described separately.

(O1) *Region filtering.* In the blue and green color segmentations, big regions corresponding to sky and vegetation were sometimes classified as ROIs. To tackle this problem, ROIs connected to the image border were discarded. Additional regions were eliminated by discarding very small/big ROIs. More specifically, regions with a bounding box (minimum rectangle encompassing the region) smaller than 5% or bigger than 50% of the image size were discarded.

(O2) *Fragment merging.* To deal with the problem of sign fragmentation, a morphological closing was applied to each binary image provided by the previous

Accepted manuscript

ACCEPTED MANUSCRIPT

stage. Since there is not a preferred orientation, a circular structuring element (omni-directional element) with a radius of 0.1% of the image size was used.

(O3) *Co-located signs separation.* When two or more traffic signs had the rim of the same color and were located at the same post (see Fig. 7a), they were frequently segmented in the same ROI (see Fig. 7b). Since these signs may not be in vertical position, the procedure for separating them is not trivial. For this purpose, the following method is proposed. Intermediate outcomes are illustrated in Fig. 7 for an array of two traffic signs with a red rim.

1. The bounding box of every ROI was extracted as a binary sub-image (Fig. 7b).
2. The main orientation of the ROI was estimated as the angle between the vertical axis and the largest ROI axis, and the ROI was accordingly rotated (Fig. 7c). In order to get a computational low-cost estimation of the tilt angle, an eigenvector decomposition was applied in the following way. First, the coordinates (x, y) of the N_p pixels constituting the ROI were arranged in matrix P ($N_p \times 2$). The covariance matrix of P (named C_P, 2×2) was computed and two eigenvalues were obtained. Since the eigenvector associated to the largest eigenvalue indicates the largest ROI axis, angle θ_r of this eigenvector will also indicate the ROI tilt in the xy-plane. The ROI was rotated θ_r degrees using nearest neighbor interpolation [8], resulting in a binary image where the vertical axis is the main ROI orientation.
3. In the reoriented ROI, holes were identified and filled by applying a morphological hole filling operator [8] (Fig. 7d).
4. The normalized horizontal projection of the filled ROI was computed and its maxima were obtained. For each maximum, a contrast measure (named dynamics) was computed. Using a topographic analogy, the dynamics of a maximum M is defined as the highest altitude loss that has to occur when traveling between M and the contiguous maximum (Fig. 7e). Then, if a dynamic exceeded a certain threshold (0.5, experimentally found by

Accepted manuscript

ACCEPTED MANUSCRIPT

visual analysis of some examples), the associated minimum was considered as the ROI separation point, and hence the ROI was divided in two parts (Fig. 7f).

4.2. Achromatic Post-Processing

As reported in the literature [19, 20], white traffic sign segmentation is highly dependent on the lighting conditions and environment. In fact, the result of the achromatic segmentation provides numerous regions that not correspond with achromatic traffic signs (false positives, see Fig. 6b). Besides, when the lighting level is low, these regions can be fused with others corresponding to traffic signs (see Fig. 6b). In order to tackle these problems, a post-processing stage with five steps is proposed:

1. Emphasize the contour of the image elements by a variance filter [39] with a 3×3 neighborhood mask on the intensity component of the HSI space. Edges were localized by thresholding the filtered image. Experimentally, we set the threshold to the value associated to the 95 - percentile of the values in the filtered image. Pixels in the filtered image exceeding the threshold were eliminated from the segmented image, and then traffic signs were separated from other objects (see Fig. 8b).
2. After Step 1, the relative color difference rcd among the R, G, and B components of every pixel p in the segmented region was computed as:
$$rcd_p = \frac{\max[R_p, G_p, B_p] - \min[R_p, G_p, B_p]}{\max[R_p, G_p, B_p]} \tag{8}$$
where $rcd \in [0, 1]$. It is supposed that low rcd values are associated to achromatic pixels (similar values of R, G and B components), while high values are associated with chromatic pixels. Thus, pixels with rcd greater than a threshold (experimentally set to 0.25) were removed from the result of Step 1 (see Fig. 8c).
3. As in the chromatic post-processing, regions connected to the outside of the image area (in most cases corresponding with asphalt and vegetation) were discarded (see Fig. 8d).

Accepted manuscript

ACCEPTED MANUSCRIPT

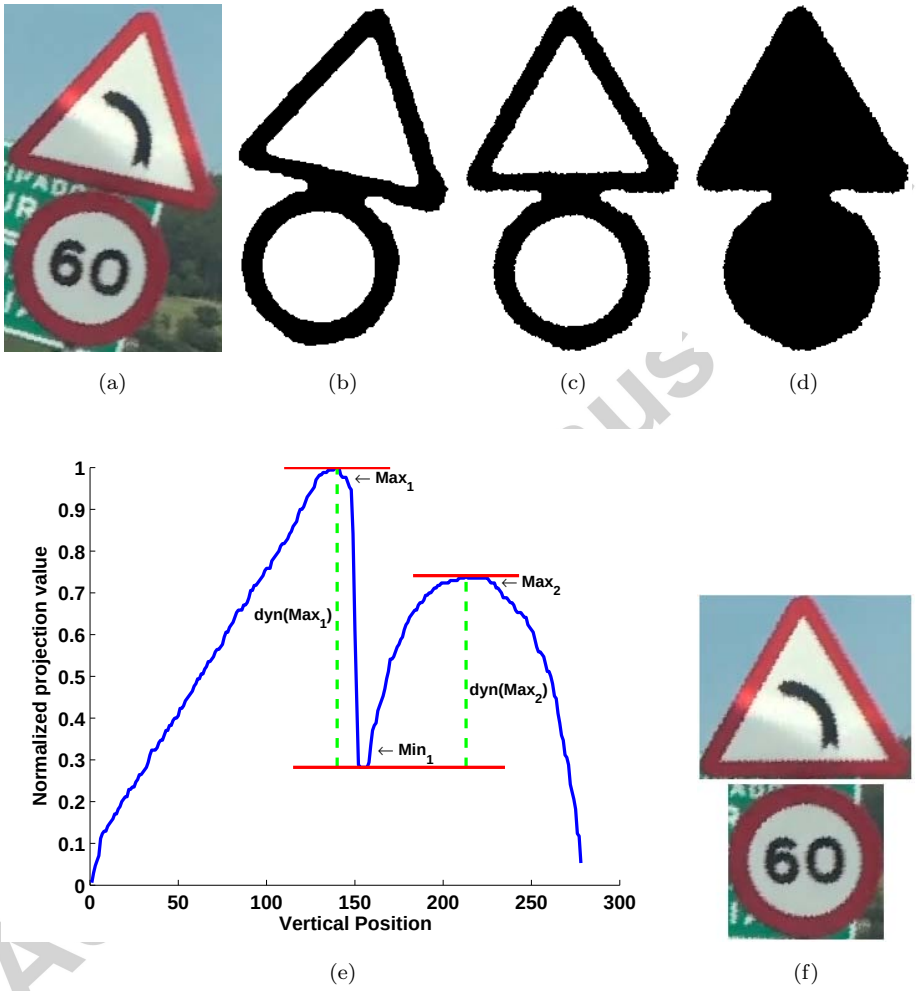


Figure 7: Separation of two co-located signs: (a) Original *RGB* image in the ROI bounding box; (b) Segmented ROI; (c) Correction of ROI orientation; (d) Filling of ROI holes; (e) Normalized horizontal projection and corresponding dynamics of maxima; (f) Bounding box of separated signals.

ACCEPTED MANUSCRIPT

4. To facilitate the contours extraction in the classification stage, regions corresponding to holes were filled by applying a morphological hole filling operator (see Fig. 8e).
5. Finally, regions with a bounding box smaller than 5% or bigger than 50% of the image size were discarded (see Fig. 8f).

After the post-processing stage, the non-discarded regions were considered as potential traffic signs. As illustrated in Figs. 8f and 9b, there may be regions not corresponding with traffic signs, and they may be fragmented in several regions. In these examples, the curve indication panel was fragmented in 8 arrow-shaped regions, and the de-restriction sign was fragmented in 2 quasi-semicircular regions. These issues are further addressed in the classification stage.

5. Shape Classification Stage

Regions provided by the post-processing stage are now classified into six geometric shapes, namely, circle, triangle, square, rectangle, arrow, and semicircle (though the procedure can be easily generalized to any other shape). Two different situations are taken into account, namely, signs segmented in a single ROI (see example in Fig. 5b) and signs fragmented in several ROIs (see examples in Figs. 8f and 9b). The shape classification has two objectives: First, to provide a way of filtering (discarding) regions which are false positives; And second, to simplify a subsequent sign identification stage (since each shape characterizes a reduced subset of signs). This stage has two phases: (1) Shape characterization, using spectral features of the region boundary, specifically the FDs [8]; (2) Shape recognition, using a set of parallel SVMs with the FDs as input features.

5.1. Feature Extraction

Several methods have been proposed for region shape characterization, which can be classified in region-based and contour-based methods [8]. While region-based methods consider all pixels in a region, contour-based methods only take

Accepted manuscript

ACCEPTED MANUSCRIPT

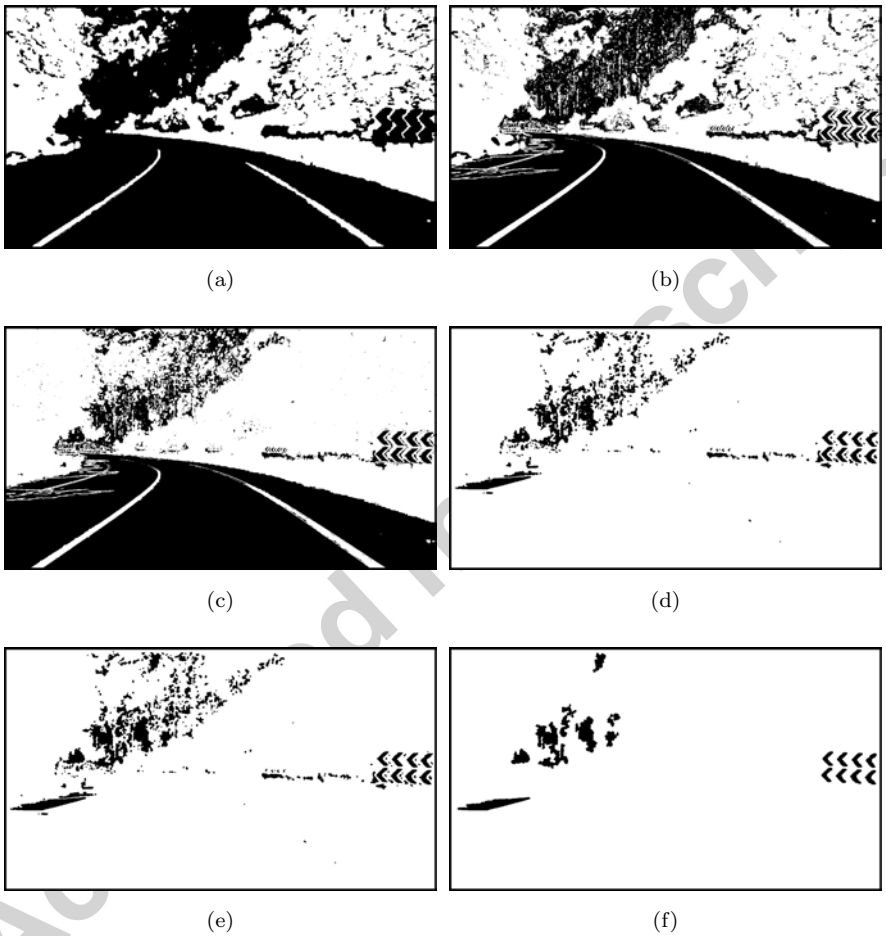
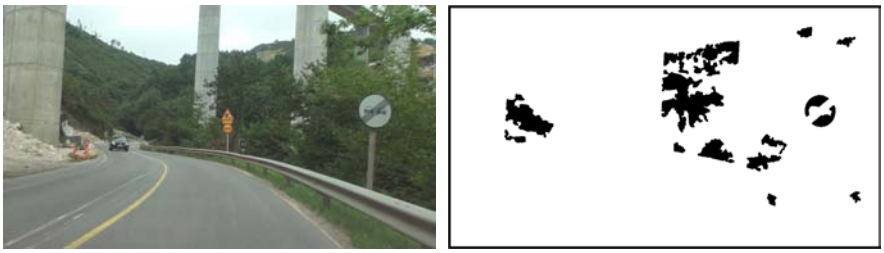


Figure 8: Example of the achromatic post-processing process for original image in Fig. 6a: (a) Outcome of the achromatic segmentation stage; (b) Separation of the regions associated to traffic signs; (c) Regions of (b) with $rcd < 0.25$; (d) Removal of regions connected to the image border; (e) Holes filling; (f) Regions filtering by bounding box size thresholding.

ACCEPTED MANUSCRIPT



(a) (b)

Figure 9: Example of the achromatic segmentation and post-processing process: (a) Original *RGB* image; (b) Final result (in black).

into account the boundary ones. There are three main types of contour-based descriptors: (1) Global descriptors (e.g., circularity and eccentricity), which can only discriminate shapes with large dissimilarities; (2) Shape signatures, like centroid distance function, tangent angle or chord length function, which represent the boundary as a one dimensional function derived from the boundary pixels and are sensitive to noise; And (3) spectral descriptors as the Fourier Descriptors, obtained by applying the Fourier Transform on the pixel coordinates of the region contour in a polar-raster way. In the context of traffic sign characterization, another contour-based descriptor named "distance to borders" (DtB) has been used in [12]. This method computes the distance between every pixel of the region contour and the side of the bounding box encompassing it, and it is quite sensitive to rotation. As mentioned in Section 1, also Shape Context has been successfully used as a descriptor in complex object matching [35], and [33] applied it for signs recognition in blue traffic signs. However, since FDs are computationally more efficient than Shape Context descriptors, we propose to use FDs for shape characterization.

The first step to compute the FDs is representing every region boundary as a sequence of Cartesian coordinates. Since every region has a particular size, corresponding boundaries may have a different number of pixels. Hence, to characterize every region with the same number of descriptors, every boundary was resampled to N_c points, what might bring a shape smoothing. Then, starting

Accepted manuscript

ACCEPTED MANUSCRIPT

at an arbitrary pixel (x_0, y_0), the (x, y) coordinates of the N_c boundary pixels are read in a polar-raster way, transformed to complex numbers $c_n = x_n + jy_n$ with $n = 0, \ldots, N_c - 1$ and collected in the set $\{c_n\}$. The Discrete Fourier Transform (DFT) of the sequence $\{c_n\}$ is computed as

$$C_u = \frac{1}{N_c} \sum_{n=0}^{N_c-1} c_n e^{\frac{-j2\pi un}{N_c}}, u = -N_c/2, \ldots, N_c/2 - 1 \quad (9)$$

where complex coefficients C_u are the shape FDs. The rotation information of the boundary is described by the FDs phase, so rotation invariance is achieved by considering only the magnitude of the coefficients. We chose $N_c = 128$ in order to accelerate the DFT computation.

Low frequency FDs contain information about the general shape features, while finer boundary details are described by the higher frequencies. Since these details are not helpful for shape discrimination, the number of FDs can be reduced by discarding the high-frequency coefficients. After preliminary experimental results, we found that a number of 21 coefficients was enough to reconstruct any sign boundary in the spatial domain while allowing to discriminate among different shapes. According to the FDs properties, translation and scale invariance were achieved by setting $C_0 = 0$ and normalizing the FDs by the magnitude of C_{-1}. Fig. 10 shows an example of three triangular signs with different size (scale) and orientation, together with their associated normalized FDs sequence. Note that the FDs sequence is very similar for all of them.

Since coefficients $|C_0|$ and $|C_{-1}|$ were always equal to 0 and 1, respectively, both were removed from the set of features, resulting in 19 coefficients.

5.2. Traffic Sign Shape Classification

Three schemes were proposed for shape classification, namely, voting k-NN, linear SVM, and non-linear SVM. To design the classifiers, the normalized FDs sequence was computed for 300 objects of the training database presented in Section 6.1. Each sequence was manually labeled with its shape: 40 objects for each shape class (a total of 240 objects) and 60 for the non-shape class (i.e. objects not belonging to any of the classes of interest).

Accepted manuscript

ACCEPTED MANUSCRIPT

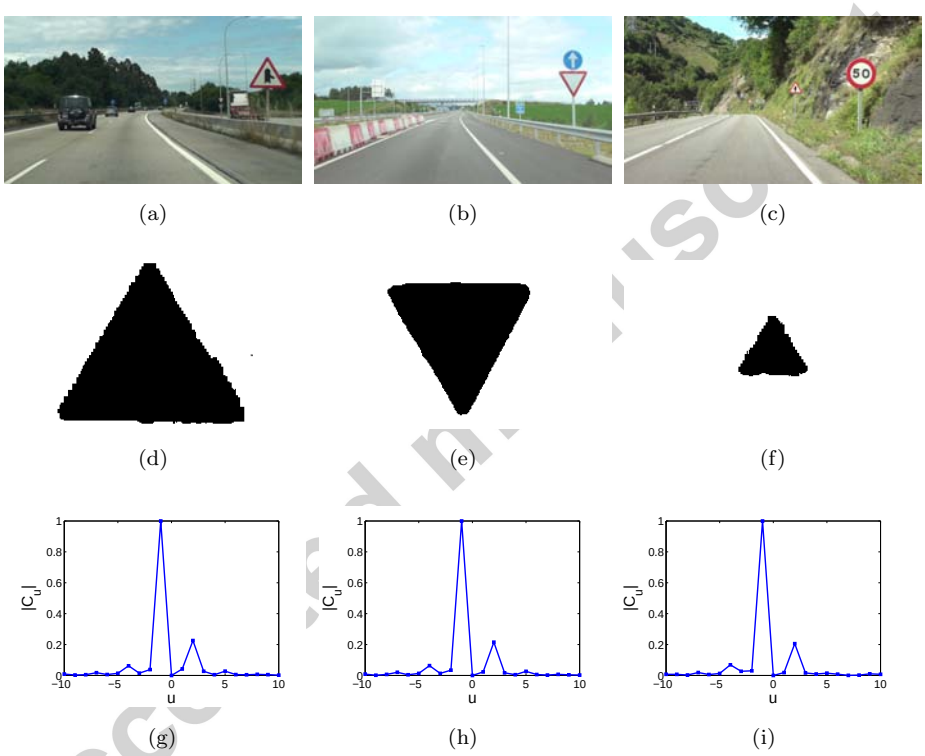


Figure 10: Normalized magnitude of the FDs of interest for three triangular traffic signs with different size and orientation: (a-c) Original images; (d-f) Segmented ROIs; (g-i) Normalized magnitude of the FDs of interest.

ACCEPTED MANUSCRIPT

For the voting k-NN scheme, the best accuracy was achieved for $k = 1$. However, a high rate of semicircular and arrow-shaped regions were incorrectly classified (assigned to the non-shape class).

With respect to the SVM schemes (linear and nonlinear), six parallel two-class SVM classifiers were designed, each one to identify a shape of interest. In order to reduce the computational burden, free parameters were set to the same values for the six classifiers by conducting a grid-search for $C \in [1, 100]$ and $\sigma \in [0.001, 10]$, choosing the pair (C, σ) yielding the best accuracy on the validation set. Every FD sequence was processed by the six two-class SVMs, performing classification as follows: (1) If none of the six SVMs assigned the sequence to a shape of interest, then the ROI was discarded; (2) If just one SVM assigned the sequence to a shape of interest, then the ROI was assigned to it; (3) If more than one SVM determined that the sequence was a shape of interest, then the shape occurrence rate in the training set was considered for classification (in order of triangle, circle, square, rectangle, arrow, and semicircle).

Best results were provided by the nonlinear SVM scheme ($C = 60, \sigma = 0.01$) in comparison to linear SVM ($C = 50$). Semicircular and arrow-shaped objects were assigned to the non-interest shape class by the linear SVM approach (the same as in the k-NN scheme). However, the nonlinear-SVM scheme only misclassified some non-interest shape objects, which were classified as one of the six shapes of interest. Hence, the non-linear SVM scheme was chosen for the shapes classification stage due to its low missing rate.

5.3. Final Grouping Stage

As mentioned in Section 3, the segmentation procedure may fragment a traffic sign into several regions. The last phase of the proposed method (see Fig. 3) addressed this problem by identifying regions associated to the same traffic sign. For this purpose, two sequential steps were carried out: (1) Embedded regions were grouped together; (2) Nearby semicircles were grouped to form a unique circular sign, and nearby arrow-shaped regions were combined to form a unique panel. These situations are common for de-restriction signs and curve indication

Accepted manuscript

ACCEPTED MANUSCRIPT

panels, respectively. To determine whether two semicircles/arrow-shaped ROIs were in the same traffic sign, the mass center of every ROI was obtained and distance between mass centers was computed. ROIs with mass centers close enough (distance lower than 10% of the image size) were combined in the same traffic sign.

6. Experimental Results

In this section, the global performance of the proposed system is evaluated. First, the database used in the system design and evaluation is presented. Second, the procedure is revisited and results are illustrated with two complete examples. Finally, the merit figures used for the assessment and the statistical results are described.

6.1. Experimental Configuration

For designing the system, the Spanish company IPS-Vial [43] supported our group providing us with an enough number of representative images with traffic signs, specifically, a set of 10 videos recorded with a high definition video camera Sony® Handycam HDR-SR11E fixed to the windshield of a vehicle. These videos, in *m2ts* format, were recorded in the public roads of the Spanish national network as part of the inventory campaign performed by IPS-Vial. A software, also provided by the company, was used to manually select and extract a set of 700 true color images, in *bmp* format and with 1440×810 pixels size. A number of 650 images had at least one traffic sign, registered in different conditions, including luminosity variations, rotations, and scale changes (see examples in Fig. 1). The whole set was further divided into two subsets: (1) A training set with 500 images for system design (see Section 2); (2) A test set with 200 images for independent evaluation of the system performance. Matlab® software was used for the experiments.

Accepted manuscript

ACCEPTED MANUSCRIPT

6.2. Proposed Method for Traffic Sign Detection

In order to provide with an overview of the method and visually illustrate results after main stages, two representative examples are next presented. Example 1 corresponds to identification of traffic signs in Fig. 11a. In order to get some insight into the segmentation and post-processing stages, Figs. 11d - 11f show the segmented regions for the three colors of interest here, i.e. red, yellow, and white (no regions of blue and green colors are segmented in this case). Images after the subsequent post-processing stages are presented in Figs. 11g - 11i, and the final result is presented in Fig. 11c, where circular and triangular co-located signs are separated, and the three signs are correctly identified. This image illustrates that regions are filtered according to their shape and correctly detected as circle and triangle signs. Note that although the small circular sign in the right part of the image is far from the camera, it complies with the size requirements and is still properly detected, whereas the blue circular sign on the left is too small and it is not detected. The final grouping stage allows to detect the two triangular regions as a single traffic sign, since the yellow region is embedded into the red one.

Example 2 shows the sign identification process in Fig. 12a, which conveys a circular sign and a black & white rotated curve indication panel. Images in Figs. 12d-12f correspond with the output of the segmentation stage for red, yellow, and white colors, and those in Figs. 12g-12i are the outputs of the subsequent post-processing stage. Note that segmented regions corresponding to sky and vegetation are discarded by the classification stage (see Fig. 12b). Finally, the four arrow-shaped regions are correctly grouped together into the same traffic sign by the final grouping stage, as shown in Fig. 12c.

6.3. Performance Evaluation

Three merit figures were considered for performance assessment:

- The *False Alarm Rate* (FAR), defined as the percentage of regions identified as traffic signs when they are not traffic signs. False alarms come from regions with shapes and colors that are similar to traffic signs.

Accepted manuscript

ACCEPTED MANUSCRIPT

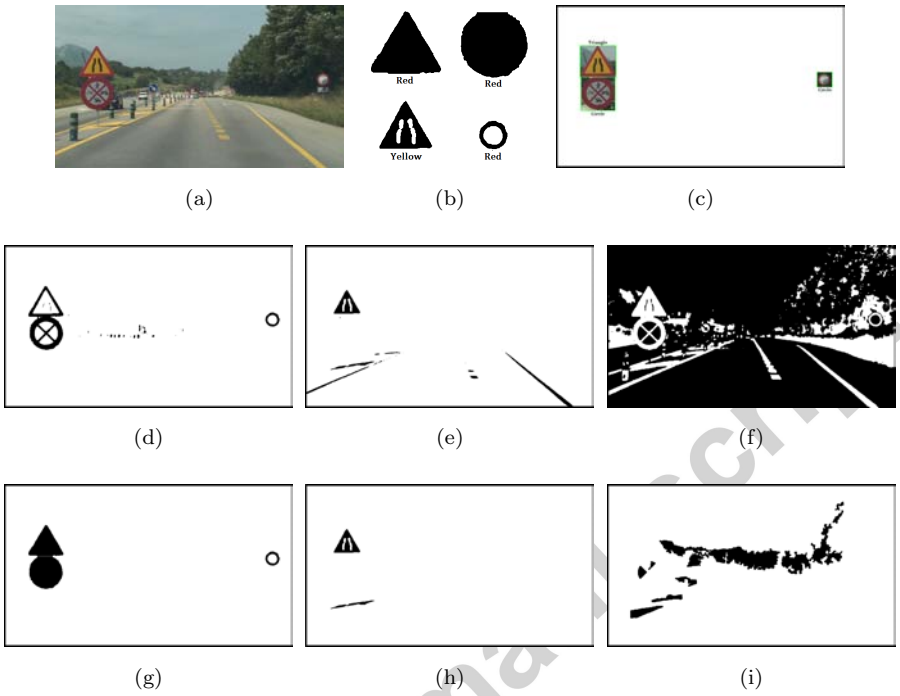


Figure 11: Example 1: (a) Original image; (b) Segmented regions (in black); (c) Detected traffic signs inside original bounding box. Segmentation results (in black) for red (d), yellow (e), and white (f) colors of interest. Post-processing results for red (g), yellow (h), and white (i) colors.

- The *Detection Rate* (DR), defined as the percentage of signs that are correctly identified.
- The *F1-Score Rate* (FSR), a measure that combines both FAR and DR, providing a more generic measure of the overall system performance [44].

As mentioned in Section 6.1, the inventory test set consisted of 200 images. A total number of 207 traffic signs (175 images conveyed at least one traffic sign) distributed into six shapes were present. Note that curve indication signs are a composition of arrow shapes, and de-restriction signs are a composition of two semicircle shapes. The occurrence probability of each kind of signal in the database corresponds to that in a real road network scenario.

ACCEPTED MANUSCRIPT

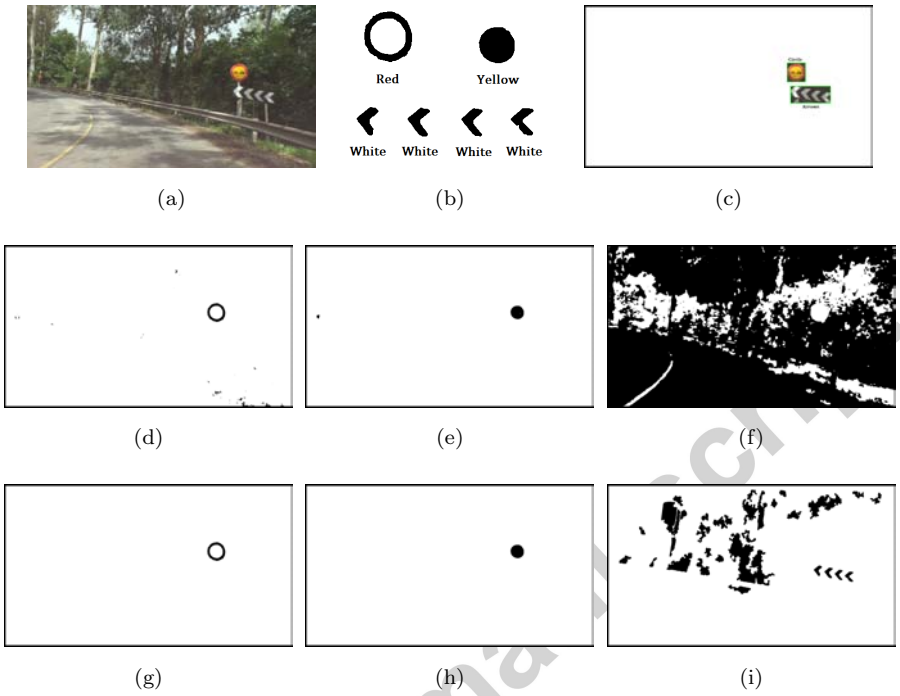


Figure 12: Example 2: (a) Original image; (b) Segmented regions (in black); (c) Detected traffic signs inside original bounding box. Segmentation results (in black) for red (d), yellow (e), and white (f) colors of interest. Post-processing results for red (g), yellow (h), and white (i) colors.

Table 2 presents the performance for each sign shape. All merit figures shown that the best performance was obtained for triangle and circle signs, because these signs have a color and shape that is quite discernible from the rest of the scene. Regarding worst results, they corresponded to rectangle, square, and de-restriction signs. A deeper analysis of these results (in missing and false alarm cases) showed that they corresponded to achromatic signs, which are the most difficult signs to identify, and more, they are occasionally over-segmented. The highest FAR corresponded to rectangular signs, since they can be confused with similar small regions in the scene, such as advertising panels or the back of the trailer trucks. As expected, the FSR provided with an overall system performance evaluation which was consistent with the joint consideration of

ACCEPTED MANUSCRIPT

	Signs	FAR	DR	FSR
Triangle	55	0	100	100
Circle	59	1.7	100	99.2
Square	31	6.1	100	96.9
Rectangle	27	10	92.6	90.9
Curve indication	20	4.8	100	97.6
De-restriction	15	6.3	80	85.7

Table 2: Performance after segmentation and shape classification when considering the inventory test images, in terms of FAR, DR, and FSR measures. The number of traffic signs is indicated for each shape.

DR and FAR, which was excellent for triangle and circle signs, good for curve indication and square signs, and reduced for rectangle and de-restriction signs. When evaluating the performance considering all signs in a real world scenario, a very good global DR (97.6%) with a reasonable overall FAR (4.2%) were obtained.

In order to test the proposed procedure in other scenarios, a new database was collected by the authors from the public Google Street View application [45]. Through a manual selection process with screen-shots, authors collected a set of 600 images, in *bmp* format and with 1440×810 size, from public roads close to Madrid (A-42 and N-301). Each image conveyed at least one traffic sign, registered at different rotations, scales, and illumination conditions, and a number of 100 signs for each shape were collected. Table 3 presents the performance for this new Google dataset and for each sign shape. In this set, as the images were not taken in a so controlled scenario as the inventory one, the difficulties in detection were more evident for signs of achromatic nature. The worst results were obtained for curve indication panels and de-restriction signs, where the FSR yielded 94.9% and 95.4%, respectively. In this case, the performance with FSR was slightly reduced for triangle, circle, and square signs. It was noticeably reduced for curve indication signs, and it significantly increased for rectangle and de-restriction signals (recall that they were the lower performance cases in

Accepted manuscript

ACCEPTED MANUSCRIPT

	Signs	FAR	DR	FSR
Triangle	100	0	97	98.5
Circle	100	2.9	98	97.5
Square	100	4.9	97	96.1
Rectangle	100	2.9	98	97.5
Curve indication	100	4.1	94	94.9
De-restriction	100	3.1	94	95.4

Table 3: Performance after segmentation and shape classification when considering Google images for FAR, DR, and FSR measures.

the design database). As in the inventory sign dataset, our method provided a high overall DR (96.7%) with a reasonable FAR (3.02%). On the other hand, performance remained similar when increasing the representativity of each kind of sign, even when evaluating images with different recording conditions from those used in the design stage, which proves the robustness and generalization capabilities of the proposed system.

7. Conclusions

This paper presented a complete procedure for chromatic and achromatic traffic sign detection and shape classification. A segmentation stage based on $L^*a^*b^*$ and HSI spaces represents the first contribution of this work. This procedure has shown a high performance with color and black & white signs. The second contribution is the post-processing stage, where the algorithm for separating co-located signs provides excellent results. The description of sign shapes by means of Fourier Descriptors, and their use together with a SVM-based classification algorithm represents the third contribution. The proposed methodology has proven to be robust against rotations, scale variations, partial deformations, and shadows. The sign-shape classification has provided a really high accuracy.

The final contribution of this work is the whole procedure, which provides

Accepted manuscript

ACCEPTED MANUSCRIPT

with good performance for different kinds of traffic signs in a variety of circumstances (including non-uniform lighting, rotations, and co-located signs), even when images are taken with different cameras. The results showed a very high DR, which was our primary objective, with a low FAR. The highest FAR was for the black & white signs (specially for rectangle, square, and de-restriction signs), which are specially difficult to discriminate from other achromatic objects.

Finally, the procedure could be used as part of a traffic sign recognition system, since it is capable of detecting both chromatic and achromatic signs with different geometric shapes. Obviously, a tracking algorithm taking into account the spatio-temporal correlation between consecutive video frames would improve the results of this work, which is ongoing future work.

8. Acknowledgments

The authors would like to thank IPS-Vial [43], Spanish company specialist in inventory and traffic signs projects, for its help providing the image database, collected as part of their inventory process.

This work has been partly supported by TSI-020100-2009-735, TEC2010-19263 and TEC2013-48439-C4-1-R projects from Spanish Government, and by Prometeo Project from the Secretariat for Higher Education, Science, Technology and Innovation of the Republic of Ecuador. JMLC is supported by the Spanish FPU grant FPU13/03134.

References

[1] The United Kingdom Department for Transport, Guidance. Traffic signs manual, 8th Edition (2009).

[2] L.-P. Becker. Method and device for traffic sign recognition [online] (Oct 2014). Patent, US 2010/0283855 A1.

[3] Opel-International. Opel innovation for your safety [online] (Oct 2014). http://www.opel.com/experience_opel/innovation/safety.html.

Accepted manuscript

ACCEPTED MANUSCRIPT

[4] G. P. Stein, O. Shachar, Y. Taieb, U. Wolfovitz. Detecting and recognizing traffic signs [online] (Oct 2014). Patent, US 8064643.

[5] Mercedes-Benz. Techcenter: Traffic sign assist [online] (Oct 2014). http://techcenter.mercedes-benz.com/_en/traffic_sign_assist/detail.html.

[6] J. Stallkamp, M. Schlipsing, J. Salmen, C. Igel, Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition, Neural Networks 32 (2012) 323–332. doi:10.1016/j.neunet.2012.02.016.

[7] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, C. Igel, Detection of traffic signs in real-world images: The german traffic sign detection benchmark, in: Intl Joint Conf on Neural Networks, 2013, pp. 1–8. doi:10.1109/IJCNN.2013.6706807.

[8] R. González, R. Woods, Digital Image Processing, 3rd Edition, Pearson Prentice Hall, 2008.

[9] J. Abukhait, I. Abdel-Qader, J. Oh, O. Abudayyeh, Occlusion-invariant tilt angle computation for automated road sign condition assessment, in: IEEE Intl Conf on Electro/Information Technology, 2012, pp. 1–6. doi:10.1109/EIT.2012.6220740.

[10] Z. Chen, J. Yang, B. Kong, A robust traffic sign recognition system for intelligent vehicles, in: Intl Conf on Image and Graphics, 2011, pp. 975–980. doi:10.1109/ICIG.2011.58.

[11] A. de la Escalera, J. Armingol, M. Mata, Traffic sign recognition and analysis for intelligent vehicles, Image and Vision Computing 21 (3) (2003) 247–258. doi:10.1016/S0262-8856(02)00156-7.

[12] S. Maldonado-Bascón, S. Lafuente-Arroyo, P. Gil-Jiménez, H. Gómez-Moreno, F. López-Ferreras, Road-sign detection and recognition based on support vector machines, IEEE Trans on Intelligent Transportation Systems 8 (2) (2007) 264–278. doi:10.1109/TITS.2007.895311.

Accepted manuscript

ACCEPTED MANUSCRIPT

[13] S. Maldonado-Bascón, J. Acevedo-Rodríguez, S. Lafuente-Arroyo, A. Fernández-Caballero, F. López-Ferreras, An optimization on pictogram identification for the road-sign recognition task using svms, Computer Vision and Image Understanding 114 (3) (2010) 373–383. doi:10.1016/j.cviu.2009.12.002.

[14] J. Khan, S. Bhuiyan, R. Adhami, Image segmentation and shape analysis for road-sign detection, IEEE Trans on Intelligent Transportation Systems 12 (1) (2011) 83–96. doi:10.1109/TITS.2010.2073466.

[15] J. Miura, T. Kanda, Y. Shirai, An active vision system for real-time traffic sign recognition, in: IEEE Conf on Intelligent Transportation Systems, 2000, pp. 52–57. doi:10.1109/ITSC.2000.881017.

[16] P. Sermanet, Y. LeCun, Traffic sign recognition with multi-scale convolutional networks, in: Intl Joint Conf on Neural Networks, 2011, pp. 2809–2813. doi:10.1109/IJCNN.2011.6033589.

[17] Y. Fatmehsan, A. Ghahari, R. Zoroofi, Gabor wavelet for road sign detection and recognition using a hybrid classifier, in: Intl Conf on Multimedia Computing and Information Technology, 2010, pp. 25–28. doi:10.1109/MCIT.2010.5444860.

[18] H. Gómez-Moreno, S. Maldonado-Bascón, P. Gil-Jiménez, S. Lafuente-Arroyo, Goal evaluation of segmentation algorithms for traffic sign recognition, IEEE Trans on Intelligent Transportation Systems 11 (4) (2010) 917–930. doi:10.1109/TITS.2010.2054084.

[19] C. Caraffi, E. Cardarelli, P. Medici, P. Porta, G. Ghisio, G. Monchiero, An algorithm for italian de-restriction signs detection, in: IEEE Intelligent Vehicles Symposium, 2008, pp. 834–840. doi:10.1109/IVS.2008.4621306.

[20] G. Overett, L. Petersson, L. Andersson, N. Pettersson, Boosting a heterogeneous pool of fast hog features for pedestrian and sign detection, in:

Accepted manuscript

ACCEPTED MANUSCRIPT

Intelligent Vehicles Symposium, 2009, pp. 584–590. doi:10.1109/IVS.2009.5164343.

[21] R. Belaroussi, J. Tarel, Angle vertex and bisector geometric model for triangular road sign detection, in: Workshop on Applications of Computer Vision, 2009, pp. 1–7. doi:10.1109/WACV.2009.5403030.

[22] M. García-Garrido, M. Ocaña, D. Llorca, E. Arroyo, J. Pozuelo, M. Gavilán, Complete vision-based traffic sign recognition supported by an I2V communication system, Sensors 12 (2) (2012) 1148–1169. doi:10.3390/s120201148.

[23] M. Boumediene, J.-P. Lauffenburger, J. Daniel, C. Cudel, A. Ouamri, Multi-roi association and tracking with belief functions: Application to traffic sign recognition, IEEE Trans on Intelligent Transportation Systems PP (99) (2014) 1–10. doi:10.1109/TITS.2014.2320536.

[24] M. Boumediene, J.-P. Lauffenburger, J. Daniel, C. Cudel, Coupled detection, association and tracking for traffic sign recognition, in: IEEE Symp on Intelligent Vehicles, 2014, pp. 1402–1407. doi:10.1109/IVS.2014.6856492.

[25] R. Andrzej, L. Yongmin, L. Xiaohui, Real-time traffic sign recognition from video by class-specific discriminative features, Pattern Recognition 43 (1) (2010) 416–430. doi:10.1016/j.patcog.2009.05.018.

[26] D. Ciresan, U. Meier, J. Masci, J. Schmidhuber, A committee of neural networks for traffic sign classification, in: Intl Joint Conf on Neural Networks, 2011, pp. 1918–1921. doi:10.1109/IJCNN.2011.6033458.

[27] F. Boi, L. Gagliardini, A support vector machines network for traffic sign recognition, in: Intl Joint Conf on Neural Networks, 2011, pp. 2210–2216. doi:10.1109/IJCNN.2011.6033503.

Accepted manuscript

ACCEPTED MANUSCRIPT

[28] S. Sathiya, M. Balasubramanian, S. Palanivel, Pattern recognition based detection recognition of traffic sign using svm, International Journal of Engineering and Technology (IJET) 6 (2) (2014) 1147–1157.

[29] F. Zaklouta, S. B., Real-time traffic sign recognition in three stages, Robotics and Autonomous Systems 62 (1) (2014) 16–24.

[30] K. Kaplan, C. Kurtul, H. Levent Akin, Real-time traffic sign detection and classification method for intelligent vehicles, in: Intl Conf on Vehicular Electronics and Safety, 2012, pp. 448–453. doi:10.1109/ICVES.2012.6294316.

[31] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, IEEE Trans on Pattern Analysis and Machine Intelligence 24 (4) (2002) 509–522. doi:10.1109/34.993558.

[32] J. Yu, D. Liu, D. Tao, H. Seah, Complex object correspondence construction in two-dimensional animation, IEEE Trans on Image Processing 20 (11) (2011) 3257–3269. doi:10.1109/TIP.2011.2158225.

[33] K. Li, W. Lan, Traffic indication symbols recognition with shape context, in: 6th Intl Conf on Computer Science Education, 2011, pp. 852–855. doi:10.1109/ICCSE.2011.6028771.

[34] J. Yu, D. Tao, M. Wang, Adaptive hypergraph learning and its application in image classification, IEEE Trans on Image Processing 21 (7) (2012) 3262–3272. doi:10.1109/TIP.2012.2190083.

[35] J. Yu, D. Liu, D. Tao, H. Seah, On combining multiple features for cartoon character retrieval and clip synthesis, IEEE Trans on Systems, Man, and Cybernetics, Part B: Cybernetics 42 (5) (2012) 1413–1427. doi:10.1109/TSMCB.2012.2192108.

[36] L. P. E., Fourier Descriptors and Their Applications in Biology, 1st Edition, Cambridge University Press, 1997.

Accepted manuscript

ACCEPTED MANUSCRIPT

[37] C. Sari, C. Akgul, B. Sankur, Combination of gross shape features, fourier descriptors and multiscale distance matrix for leaf recognition, in: 55th International Symposium ELMAR, 2013, pp. 23–26.

[38] R. Hunt, M. Pointer, Measuring Colour, 4th Edition, Wiley, 2011.

[39] R. Duda, P. Hart, D. Stork, Pattern Classification, 2nd Edition, Wiley-Interscience, 2001.

[40] V. Vapnik, The Nature of Statistical Learning Theory, 2nd Edition, Springer Verlag, 1999.

[41] C. Bishop, Pattern Recognition and Machine Learning, 8th Edition, Springer, 2009.

[42] R. Fletcher, Practical Methods of Optimization, 2nd Edition, Wiley-Interscience, 1987.

[43] IPS-Vial. Inventarios y proyectos de señalización vial [online] (Oct 2014). http://www.ipsvial.es/.

[44] M. Sokolova, N. Japkowicz, S. Szpakowicz, Beyond accuracy, f-score and roc: A family of discriminant measures for performance evaluation, in: Advances in Artificial Intelligence, Vol. 4304 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2006, pp. 1015–1021. doi:10.1007/11941439_114.

[45] Google. Street view - google maps [online] (Oct 2014). https://www.google.com/maps/views/streetview/.

Accepted manuscript