An Approach to the Recognition of Informational Traffic Signs Based on 2-D Homography and SVMs

A. Vázquez-Reina, R.J. López-Sastre, P. Siegmann, S. Lafuente-Arroyo, and H. Gómez-Moreno

Department of Signal Theory and Communications,
Universidad de Alcalá,
Escuela Politécnica Superior. Campus Universitario. 28805 Alcalá de Henares, Spain
ameliovazquez@gmail.com
robertojavier.lopez@alu.uah.es
{philip.siegmann, sergio.lafuente, hilario.gomez}@uah.es

Abstract. A fast method for the recognition and classification of informational traffic signs is presented in this paper. The aim is to provide an efficient framework which could be easily used in inventory and guidance systems. The process consists of several steps which include image segmentation, sign detection and reorientation, and finally traffic sign recognition. In a first stage, a static HSI colour segmentation is performed so that possible traffic signs can be easily isolated from the rest of the scene; secondly, shape classification is carried out so as to detect square blobs from the segmented image; next, each object is reoriented through the use of a homography transformation matrix and its potential axial deformation is corrected. Finally a recursive adaptive segmentation and a SVM-based recognition framework allow us to extract each possible pictogram, icon or symbol and classify the type of the traffic sign via a voting-scheme.

1 Introduction

In this paper we handle the task of automatically detecting, recognizing, and classifying informational traffic signs. Several works have recently focused on traffic sign detection and recognition [1-9]. Some of which keep stages for sign detection and classification separated, such as [5] and [6], while some others try to address the whole process in a unique framework like [7]. Nevertheless, most of these works have only dealt with regulatory and warning traffic signs, and only a few have proposed a system to cope with guide and informational traffic signs, such as [2] and [3].

There are many challenges we must surpass in order to achieve successful results. We need to deal with some of the most common problems which usually arise in this kind of tasks, such as rotations, occlusions, variable lighting conditions of the scene, or sign deterioration. Some of these issues have been analyzed in [10]. In addition, we need to consider a great amount of different combinations of pictograms, symbols, or characters which are generally present on a typical informational traffic sign. For this

J. Blanc-Talon et al. (Eds.): ACIVS 2006, LNCS 4179, pp. 1163 – 1173, 2006.

reason, we want to bring to the reader's attention that since traffic signs can usually contain variable-size text strings, and they might be present together with a very different kind of icons or pictograms (Fig. 1-a), it would be important to be able to dynamically organize hierarchically these objects in some way so we could easily perform a traffic sign classification based on this data.

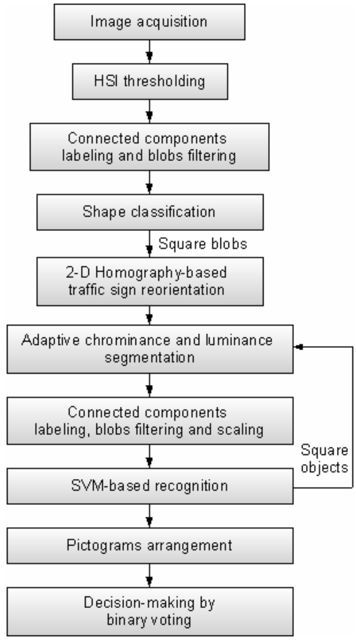
Our framework is capable to overcome all these difficulties in several steps. Firstly, we detect and reorient every possible rectangular traffic sign which might be present on the scene. Subsequently, we carry out an adaptive segmentation to discriminate each character, and symbol of candidate signs from their background. Blobs are then recognized by means of a SVM framework. Due to the nature of informational traffic signs, those which resulted to be rectangular are adaptively segmentated again recursively. Pictograms are then arranged vertically and horizontally. Finally the traffic sign is classified via a voting-scheme.

2 System Overview

Common Spanish informational traffic signs are rectangular and have a blue or white background. Foreground sign objects are designed to be clearly distinguishable from the surrounding with the help, among other things, of a high contrast to the background. Pictograms colors change generally only when they are encircled by square frames.

All these facts led us to think of dividing the process into several steps which are presented next. Initially, the original image is segmented by means of thresholding in HSI color space. This allows us to separate blue and white blobs from the context. Shape classification is then responsible for the selection of those which seem to be rectangular. Once candidate traffic signs have been extracted, we reorient them using a homography transformation matrix [11]. In the second stage we analyze the luminance and chrominance of the traffic sign in order to cope with random lighting conditions such as broad daylights, or shaded areas. Thus, we compute the colour and luminance thresholds needed to separate foreground objects from the background by way of an adaptive segmentation. This is one of the most important steps since an appropriate statistical characterization in a proper colour space may determine the success of a correct identification and recognition of every pictogram, and therefore, the right classification of the traffic sign. Afterwards, we perform connected components labeling and filter blobs in accordance with their geometrical properties such as their size or their aspect ratio. A SVM-based recognition framework classifies each blob taking as input a binary n-dimensional vector from each adaptive-segmented candidate pictogram. Blobs which are classified as square are adaptively-segmented and its pictograms classified again recursively. Objects which are succesfully identified as real pictograms are then arranged vertically and horizontally by means of simple clustering, and then sorted through an adapted version of the QuickSort algorithm. A majority voting method is finally employed to get the classification from blobs position and their recognition. The complete process is outlined in Fig. 1-b.

(a)



(b)

Fig. 1. (a) Some traffic signs with several kinds of pictograms (b) System portrayal

3 Detection and Reorientation of Informational Traffic Signs

The main goal in this stage is to detect candidate traffic signs in the original scene and to reorient them. As it was mentioned above, Spanish informational traffic signs background is usually blue or white, and therefore, the first block of the detection system consists of a blue and white segmentation stage by thresholding over a given color space. We refused direct thresholding over RGB color space because, despite it might be faster under certain circunstances, it turns out to be very sensitive to lighting changes. A combination of a fixed HSI segmentation and an achromatic decomposition was consequently chosen due to its benefits as it is explained in [1].

After segmentation stage, foreground pixels are grouped together as connected components. We then classify each blob's shape employing the method described in [12] where a comparison is made between the absolute value of the FFT applied to the signature of blobs and reference shapes. Fig. 2 shows how the signature for a reference rectangular shape and for a traffic sign sample look like. 64 samples were chosen starting at 0 radians and ending at 2*π radians, and the signature was always normalized to its energy. Blobs with rectangular shape are then sucesfully identified and reoriented.

In the following we explain why a traffic sign reorientation is considered and why we decided to fulfill it here. First of all, a reorientation would help to make the system rotation-invariant since this would allow us to deal with only one tilt and direction regardless how they actually appear in the original scene. As long as a traffic sign may contain a lot of different kinds of icons, symbols and characters, it might result also very efficient to reorient all of them together as they theoretically should share the same tilt and distortion. Furthermore, it should be noticed that it is also much easier to gather information about objects disposition from the traffic sign vertexes rather than from each of them individually. The reorientation process is done via the Direct Linear Transformation (DLT) algorithm described in [11]. We compute H, a homography transformation matrix which univocally sets the linear relation between all the points on the reoriented traffic sign P' and on the original one P. If we consider homogeneous coordinates, given the group of points $\bar{a}_i = (x_i, y_i, z_i) \in P$ and its corresponding $\bar{a}'_i = (x'_i, y'_i, z'_i) \in P'$ we can set the following relation:

$$\bar{a}_i = H \cdot \bar{a}'_i \tag{1}$$

In a general transformation case we would have nine degrees of freedom which stand for a complete projective transformation (Fig.3-a):

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \tag{2}$$

But if we consider however, a similarity transformation, results will remain practically the same as far as we suppose that traffic signs are spotted from a distance long enough when compared to their size [11]. As a result of this approximation, points coordinates from P and P' would state as $\bar{a}_i = (x_i, y_i, 1)$ and $\bar{a}'_i = (x'_i, y'_i, 1)$ respectively and we could significantly simplify our problem by reducing to four the number of variables to compute, as now, H, the homography transformation matrix, corresponds to:

$$H = \begin{bmatrix} s\cos(\theta) & -s\sin(\theta) & t_x \\ s\sin(\theta) & s\cos(\theta) & t_y \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

Where θ denotes the rotation angle, and s, t_x and t_y represent the traffic sign scale and its translation in the X and Y axes respectively. Despite all and each of these variables define the transformation between both traffic signs P and P' in the similarity transformation case, we can not easily compute them directly from segmented blobs.

For that reason, we opted for computing H by considering a set of points correspondences which allow us to determine the four variables as in:

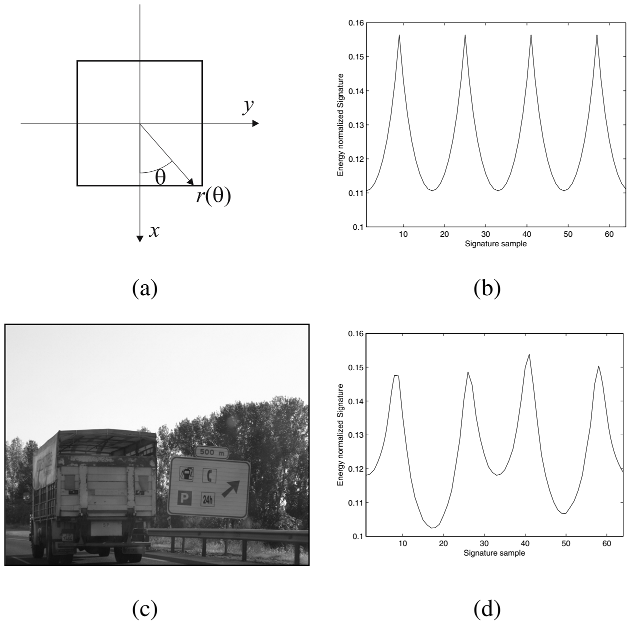


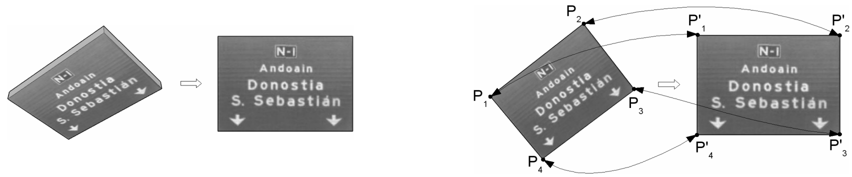
Fig. 2. Shape and traffic sign vertexes detection. (a) Ideal square blob. (b) Energy normalized signature of an ideal square blob. (c) Sample Image. (d) Energy normalized signautre of Fig2-c image.

$$H = \begin{bmatrix} R & \bar{t} \\ \bar{0}^T & 1 \end{bmatrix} = \begin{bmatrix} h'_A & -h'_B & t_x \\ h_B & h'_A & t_y \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

Theoretically we should be able to determine the four degrees of freedom of the homography matrix with two correspondences between two points each, but since this would require points coordinates to be unmistakable measured, we can rather use a greater number of correspondences so as to form an over-determined system which can be easily solved through the use of standard techniques for linear equations solving.

By reason of the former, we can use the four vertexes of the detected traffic sign already computed when calculating the blob signature to set correspondences betweenthese four vertexes named P_1, P_2, P_3 and P_4 and those of the reoriented traffic sign we are about to get (Fig. 3-b). Finally, once H is given, we can compute each pixel of the reoriented traffic sign as:

$$\bar{x}_i = H \cdot \bar{x}'_i = R \cdot \begin{bmatrix} x'_i \\ y'_i \end{bmatrix} + \bar{t} \tag{5}$$



(a) (b)

Fig. 3. (a) Projective transformation. (b) Similarity transformation and its four traffic sign vertexes correspondence.

3 Adaptive Segmentation

In order to accomplish a correct classification of a informational traffic sign, we need to know which icons and signs are actually displayed on it. These pictograms are in fact what make traffic signs different from one another, and we would like to remark that some of them may have very complex shapes. Thus, a proper segmentation of the traffic sign under test would be very convenient so that even small object details can be considered for pictograms identification.

Variable intensity conditions, the presence of noisy artifacts and possible shaded portions on a traffic sign, make very difficult to segment traffic signs in detail with fixed thresholds in a given colour space under all possible conditions. It follows that an adaptive method which might be able to extract and discriminate dynamically and accurately every object on the traffic sign would be very useful for getting fine results.

In this stage we analyze the luminance and chrominance distribution of the traffic sign in the CIE L*a*b color space (CIELAB from now on). CIELAB is based on the CIE 1931 XYZ colour space and consists of a luminance component and two chrominance components. It has been created to serve as a device independent model and it is considered one of the most complete colour model used to describe all the gamut of colours visible to the human eye [13]. Accuracy and efficiency discussions in the transformation from RGB to CIELAB can be found in [14].

We suppose that the amount of pixels which belong to an object background is always greater than the amount of those which pertain to an object. This fact can be noticed when computing the histogram of traffic sign chrominance or luminance components. Fig. 5 shows an example where it can be observed that there is always a wide range of values spread around a maximum peak which actually represents the most common background pixel value. Thereby, we can establish a frontier in both chrominance and luminance planes and consequently distinguish background from the foreground.

3.1 Luminance Segmentation

In the case of luminance there is only one component to work with. Supposing there is a high contrast difference between foreground and background, the function which may discriminate them is:

$$f(L) = \begin{cases} \text{“foreground/background“ if } L < \beta \\ \text{“background/foreground” if } \beta < L \end{cases} \tag{6}$$

Where L represents the luminance component of a pixel and β represents the threshold we need to find. Depending on where the maximum peak lies, that is, in which half of the luminance histogram most common background value is located, we can distinguish which side corrsponds to the foreground and which one to the background.

3.2 Chrominance Segmentation

CIELAB provide two chrominance components, and generally, the optimal chrominance function which could be able to separate the background from the foreground can be very complex and slow to evaluate. A convenient estimation can speed up the segmentation process whereas still offering good results. Since the most common background color of Spanish informational traffic signs and their respective frames and can be blue or white, we have chosen two functions $f_1(L)$ *and* $f_2(L)$ for evaluation which are described next.

For white backgrounds, we have $f_1(L)$ which defines a polygonal approximation of a circle C with radius r centered in the *ab* chrominance plane.

$$f_1(L) = \begin{cases} \text{“background“ if } (a,b) \in C_r \\ \text{“foreground” if } (a,b) \notin C_r \end{cases} \tag{7}$$

For blue backgrounds, we have $f_2(L)$ which defines an adequate radial portion R with a proper broadness α, centered in the *ab* chrominance plane.

$$f_2(L) = \begin{cases} \text{“background“ if } (a,b) \in R_\alpha \\ \text{“foreground” if } (a,b) \notin R_\alpha \end{cases} \tag{8}$$

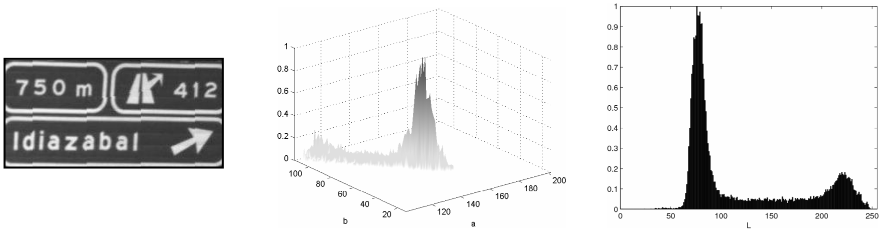
After various experimental tests we chose the parameters α, β and r which better results offered.

4 SVM Based Recognition

Once the traffic sign has been properly segmented, we group pixels into blobs by means of connected components labeling. Each pictogram contained on the traffic sign should then result in a binary blob which will be the input to the recognition system.

The recognition framework is based on a RBF (Radial Basis Function) Kernel Support Vector Machine (SVM). SVMs are a set of related supervised learning methods which can be applied to solve many pattern recognition and regression estimation problems. They were originally introduced by Vapnik [15], [16] and they are widely used nowadays to solve binary classification problems. In these cases, if both classes could be separated by a linear hyperplane (Linear-SVMs), we would have:

- The training sets $\{x_i, y_i\}$. Where $i=1,\ldots,l$, l is the number of training vectors, $y_i \in \{-1,1\}$ identifies each class and $x_i \in \{R^d\}$ are the input feature vectors.



(a) (b) (c)

Fig. 5. (a) Informational Traffic sign which has been detected from and then reoriented. (b) Traffic sign's chrominance distribution. (c) Traffic sign's luminance distribution.

- The optimized hyperplane *{**w**,b}* computed from the training sets which separates the two classes.
- The decision function given by:

$$f(x) = sgn(x \cdot w^T + b) \tag{9}$$

which determines on which side of the former hyperplane a given test vector x lies. Our case differs from the above one in two aspects. Firstly, we need to identify more than only two classes, so several one-vs-all SVMs classifiers have been actually used. Secondly, data to be classified can not usually be separated by a linear function, so we resorted to what is commonly known as the "kernel trick". This solution consists in:

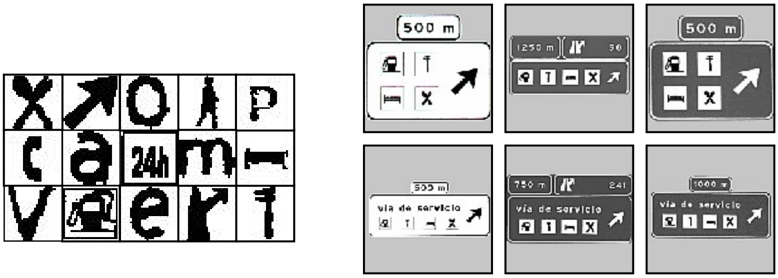
- Map the input data into a different space $\Phi(\mathbf{x})$ by means of the kernel function K which let us to use non-linear hyperplanes that may fit better to our problem in question.
- Build the new decision function $f(x)$ in which the scalar product of Eq .9, results in $<\Phi(\mathbf{x}), \Phi(\mathbf{w})>$, also labeled as $K(\mathbf{x},\mathbf{w})$.

$$f(x) = sgn(K(w,x) + b)) \tag{10}$$

The Kernel K we chose was the RBF since it was the one which better results offered. The RBF kernel can be defined as:

$$K(x_i, x_j) = e^{\frac{-\|x_i - x_j\|^2}{2\sigma^2}} \tag{11}$$

where σ is defined as the RBF width, and x_i and x_j represent sample vectors. The SVM input vector consists of a block of 31x31 binary pixels for every candidate blob, so the interior of the bounding-box of the blob is normalized to these dimensions. σ was optimized heuristically and σ = 1e-04 was the one which better results offered. Some examples of these vectors can be seen in Fig. 6-a.



(a) (b)

Fig. 6. (a) Sample blobs used as input vectors in the SVM recognition system. (b) Spanish Informational Traffic signs S-261, S-263, S-263a, S-264, S-266, S-266a.

5 Pictograms Arrangement

The way symbols and characters are positioned in traffic signs is not random at all, and they actually follow some fixed patterns. Pictograms relative position provide thus important information about how succesful have been the traffic sign detection and pictograms recognition. Furthermore, we can in fact gather very useful information about the type of the traffic sign to be classified from pictograms position. According to this, blobs which were succesful recognized by the SVM framework are clustered in rows and columns and sorted by means of an adapted version of QuickSort. Under normal circunstances no complex clustering techniques are needed since traffic signs reorientation provides enough alignment. Grid spacing is selected based on the average size of identified blobs.

7 Binary Voting

Spanish informational traffic signs differs from one another in their pictograms, characters and color schemes. Moreover, depending on the type of traffic sign, they present some specific properties which can be taken into account in a classification framework. Pictograms are usually placed following a fixed pattern which can be easily noticeable and used for identification purposes. There is also usually some redundant information which can be very useful for avoiding false alarms and making the identification more robust. Some of these examples can bee seen in Fig. 6-b where Spanish traffic signs named S-261, S-263, S-263a, S-264, S-266 and S-266a [17] are showed. They all share some common properties such as an indication arrow and they differ from one another in the highway-exit pictogram and the text string "vía de servicio". Our framework makes the most of these facts. Blobs position and identification are taken as input to a binary voting system and several conditions are setted so as to determine which traffic sign best suits to the information gathered from blobs.

8 Experimental Results

Images used for testing were compressed in JPEG. The sample set is composed of an average of 60 images for each informational traffic sign of types S-261, S-263, S-

263a, S-264, S-266 and S-266a under very different ligthining conditions and environments. Tests were done in a conventional PC desktop.

Table 1 represents experimental results obtained from the above mentioned test set. A traffic sign is considered to be detected when it was properly segmented, its shape correctly classified, its blob succesfully reoriented and the binary voting system recognized it as a valid informational traffic sign. False alarms occur when an image blob is wrongly considered to be a traffic sign and it is classified as one valid type of informational traffic sign. An average of 33% of false alarms was obtained from the total sample set.

Table 1. Percentage results

	S-261	S-263	S-263a	S-264	S-266	S-266a
Detection	78,00%	82,35%	86,67%	75,00%	77,78%	83,19%
Classification	72.73%	77.53%	80.31%	69.34%	75.36%	75.82%

9 Conclusions and Future Work

This paper describes a complete method to detect and recognize informational traffic signs. It is able to classify traffic signs according to their color schemes and symbols displayed on them.

The overall performance of the classifier depends mainly on how well foreground objects are extracted from the background. Chrominance and luminance analysis characterization of traffic signs and their square frames are essential, and the overall performance depends in a great extent on setting proper thresholds.

Future lines of work can include video tracking, and improvements in traffic sign detection in difficult enviroments. Video tracking would give more reliability to the system since more frames would be given for each traffic sign, and possible misses could be compensated with hits in other frames. Improvements in detection with shape reconstruction techniques can make the system to be able to cope with big occlusions and camera distortions.

Acknowledgment

This work was supported by the project of the Ministerio de Educación y Ciencia de España number TEC TEC2004/03511/TCM.

References

1. 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 Transactions on Intelligent Transportation Systems, (Submitted).
2. W. Wu, X. Chen, and J. Yang, Detection of Text on Road Signs From Video, IEEE Trans. Intelligent Transportation Systems, Vol. 6, No. 4, (2005) 378-390.

3. X. Chen, J. Yang, J. Zhang, and A. Waibel, Automatic Detection and Recognition of Signs From Natural Scenes, IEEE Trans. Image Processing, Vol.13 no.1, (2004) 87-89.
4. E. D. Haritaoglu and I. Haritaoglu, Real time image enchancement and segmentation for sign/text detection, in Proc. Int. Conf.Image Processing (ICIP), Barcelona, Spain, vol. III, 993-996.
5. P. Paclik, J. Novovicova, P. Somol, and P. Pudill, Road sign classification using the laplace kernel classifier, Pattern Recognition Letters, vol. 21, (2000) 1165-1173.
6. J.Miura, T. Kanda, and Y. Shirai, An active vision system for real-time traffic sign recognition Proc. 2000 Int Vehicles Symposium, (2000) 52-57.
7. M. V. Shirvaikar, Automatic detection and interpretation of road signs, Proc. of the Thirty-Sixth Southeastern Symposium on System Theory, (2004) 413 - 416.
8. Zin, T.T.; Hama, H. Robust road sign recognition using standard deviation; 7th International IEEE Conference on Intelligent Trasportation Systems, (2004) 429 - 434.
9. A. de la Escalera; J.M. Armingol, J.M. Pastor, F.J. Rodriguez; Intelligent Transportation Systems, IEEE Transactions on Volume 5, Issue 2, (2004) 57 - 68
10. S. Lafuente-Arroyo, P. Gil-Jiménez, R. Maldonado-Bascón, Traffic sign shape classification evaluation evaluation I: SVM using distance to borders, Proc. IEEE Intelligent Vehicles Symposium, Las Vegas, USA, (2005).
11. Hartley, R.I. and Zisserman, A. Multiple View Geometry in Computer Vision, Cambridge University Press, (2004).
12. P. Gil-Jiménez, S. Lafuente-Arroyo, H. Gomez-Moreno, F.López Ferreras and S. Maldonado-Bascón, Traffic sign shape classification evaluation II: FFT applied to the signature of blobs, Proc IEEE Int Vehicles Symposium, Las Vegas, USA, (2005).
13. G.M. Johnson and M.D. Fairchild, A top down description of S-CIELAB and CIEDE2000, Color Research and Application, 28, (2003) 425-435.
14. Connolly,C.;Fleiss,T. Image Processing, IEEE Transactions on Image Processing, Volume 6, Issue 7, (1997) 1046-1048.
15. V. Vapnik, The nature of Statistical Learning Theory. Springer-Verlog. New York, (1995).
16. V. Vapnik, Statistical Learning Theory, John Wiley and Sons, Inc., New york, (1998).
17. Boletín Oficial del Estado Español, Real Decreto 1428/2003, núm 306.