

Adaptive traffic road sign panels text extraction

A. VÁZQUEZ REINA, R. J. LÓPEZ SASTRE, S. LAFUENTE ARROYO, P.GIL JIMÉNEZ.

Departamento de Teoría de la Señal y Comunicaciones
Universidad de Alcalá
Alcalá de Henares, Madrid
SPAIN

Abstract: - In this paper we present an approach to the detection and extraction of text in road sign panels. Text strings, indicators and signs extraction is efficiently performed so OCR algorithms can recognize different characters that may be present on the traffic plane. In a first step, basic color segmentation and shape classification is done for the purpose of detecting possible rectangular planes. Every detected plane is extracted from the original image and then reoriented. Chrominance and luminance histogram analysis and adaptive segmentation is carried out, and connected components labeling and position clustering is finally done for the arrangement of the different characters on the panel. Special emphasis has been placed on the adaptive segmentation. Experimental results have showed that following steps strongly depends on correct separation between the background and foreground objects of the panel. Moreover, OCR systems are highly sensitive to noise, and we have put special attention into it in order that the OCR system could be able to recognize characters properly.

Key-Words: - Road-sign, detection, classification, image segmentation.

1 Introduction

Many works have been developed in the field of traffic sign detection and recognition [1], [2], [3], [4], [5]. Automatic classification of traffic road panels can be very useful for inventory purposes. Road panels provide drivers important information about the route by means of text strings and different traffic signs. Most of common road panels are rectangular, and the text on them is designed with high contrast to their background. These foreground and background colors are not randomly distributed and they are distinguishable from the surrounding environment. The font and the size of the text of a panel are always chosen in order to be easily readable, and every object and sign displayed with it are usually big and clear enough to be visualized from a long distance. Fig.1 shows four examples of road sign panels where these facts can be noticed.

There are however many problems involved in the traffic road sign and panels detection and recognition. Some of the most important difficulties are variable lighting conditions, partial occlusion and rotations. They have been conscientiously analyzed in [6]. Many research projects have tried to address these issues, but most of them are limited to symbol recognition [7], [8]. They are usually based on segmentation through static color thresholding, region detection and shape analysis, [2].



Fig.1. Some road sign panels examples.

It is important to understand that text strings and other remarks make each road panel different to another, and it would be useful an adaptive method to extract and discriminate dynamically every object displayed on them. In fact this is what our proposed approach does. It overcomes the task in two successive steps. First, we find, extract, and reorient possible panels on the scene, and then we analyze them so as to try to separate each object which may appear on them.

2 System description

As it has been said above, the framework is divided in two stages. In the first one we need to be able to detect every potential panel on the input captured image. Common Spanish road panels are rectangular and their background is white or blue. This allow us to extract candidate blobs from the input image by thresholding using HSI color space and computing shape classification. Once each potential panel is located and properly separated, we need to reorient it, what is done through a homography transformation matrix. In the second stage we extract every potential character on the panel. It is known that the color of the text in typical Spanish road panels depends on the color of the panel. The most common combinations are showed in Fig. 2.

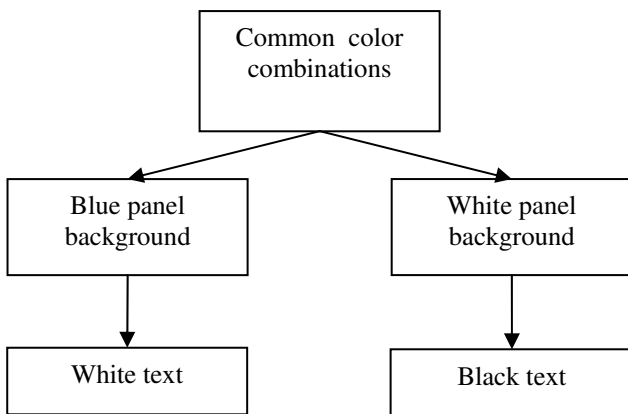


Fig.2. Common traffic panels color combinations

This fact could help us to discriminate text on the panel, but the system we propose on this paper can detect and separate characters regardless of their nature of color as long as there is always some color contrast between the panel background and foreground. It follows that the key idea of this stage is to analyze chrominance and luminance components in LAB color space so we can apply a suitable and adaptive segmentation. This let us to label connected components and then filter only those blobs with a character appearance in accordance with their geometrical properties. The ones that surpass the filter are arranged vertically and horizontally by means of clustering, and then sorted, that is, each character is placed in a row and a word. The complete process is described in Fig. 3.

3 Panel detection and reorientation

The first step here is to segment the original image. HSI color space is used because the main difficulties that we encounter are related to intensity changes and we believe Hue and Saturation components of the HSI space are invariant enough to intensity conditions. Blue pixels are easily detected trough appropriate fixed thresholds, whereas white ones need an achromatic decomposition as explained in [9]. The goal is the creation of a mask where pixels of image that may belong to a traffic panel are marked as object, whereas pixels that may not belong to a traffic panel are marked as background. Afterwards connected components labeling is computed and then all candidate blobs are analyzed in a selection process where some of them are discarded according to their size or aspect ratio.

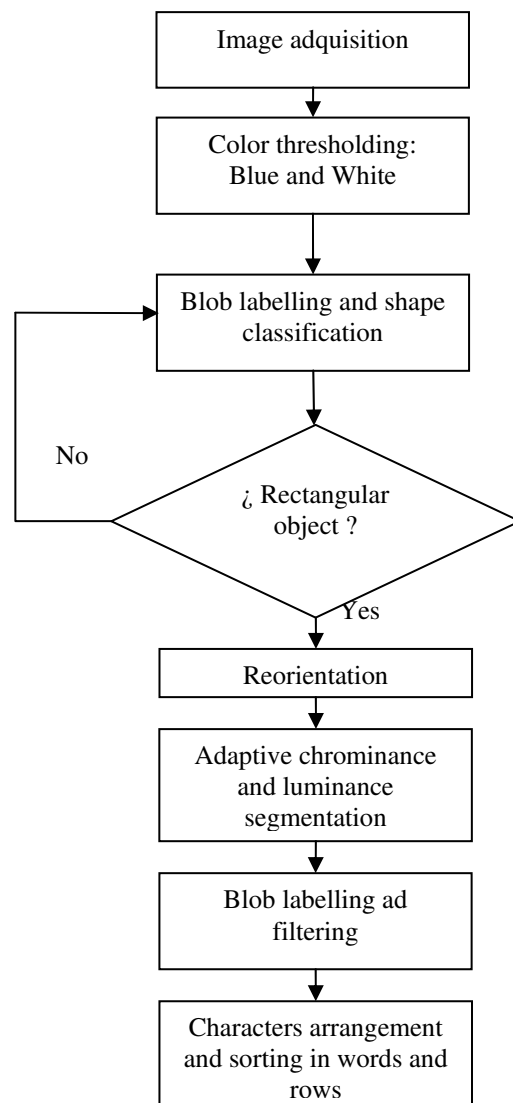


Fig.3. Algorithm description

The second step is to classify each blob's shape. This is done through the method described in [10], where FFT is applied to the signature of blobs. Rectangular blobs should have a signature like the one showed in Fig. 4 and the classification technique basically consists of a series of comparisons between the FFT of the signature of a blob and the FFT of the signature of a rectangular shape reference. It's important to point out that this method successfully deals with object rotation and deformation and camera projection distortions.

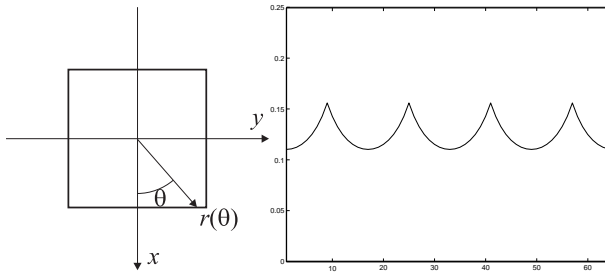


Fig.4. Rectangular shape reference and its signature.-

The third and last step in this first stage is to reorient the extracted blob. What we do is to compute each of the four panel vertexes on the original image from the blob signature. It is easy to see that the four maximum-peaks of the signature correspond to the four vertexes of the panel. We just need to set a correspondence between these four vertexes named P_1, P_2, P_3 and P_4 and those of the reoriented panel we want to obtain. P_1 is the first vertex we find beyond a 180° angle measured from the x axis in both images as long as we suppose there have not been strong rotations, so it is a good reference point.

From this correspondence and applying the DLT algorithm described in [11], we can compute H , the homography transformation matrix, which definitely sets the linear transformation between all the points on the original panel X and those on the reoriented one X' :

$$X' = H \cdot X \tag{1}$$

A graphical view of the reorientation process can be seen on Fig. 5.

4 Panel segmentation in CIELAB color space

The main objective in this step is to separate the foreground objects from the background of the panel. The panel background lacks of information, and we actually need to distinguish and separate every object from it.

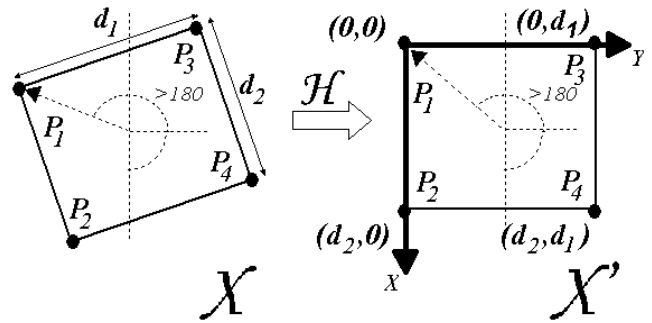


Fig.5. Reorientation of the panel

Traffic road signs can be of several colors, and they are present in different illumination conditions. Image resolution is not always high enough, and JPEG artifacts can make it difficult to separate the objects from the panel background. Furthermore, characters need to be well extracted if we want to perform subsequently efficient recognition. All these difficulties drive us to put special attention into the image segmentation. Luminance and chrominance analysis can help us to discriminate noisy regions and to avoid they could adversely affect following steps. Several color spaces have been tried, and CIE $L^*a^*b^*$ (CIELAB) presented the best performance over all. CIELAB is based on the CIE 1931 XYZ color space and it is considered one of the most complete color model used to describe all the gamut of colors visible to the human eye. It has been created to serve as a device independent model and the non linear relations for L^* , a^* , and b^* are intended to mimic the logarithmic response of the eye. It is inherently parameterized correctly, and thereby it always defines an exact color in contrast to RGB where an ICC (International Color Consortium) profile is needed to obtain the same result.

LAB consists of a luminance component and two chrominance components and it allow us to perform independent chrominance and luminance analysis. In this way we can overcome different intensity conditions easily.

4.1 Chrominance analysis and segmentation

Different signs, indicators, symbols and characters are usually placed on the panel with a color that let the driver to see, interpret and read them easily. Along the whole process is supposed that the amount of pixels which belong to the panel background is always greater than the amount of those which pertain to an object. This fact has been true for the bench of samples we have used in our tests, and we have not found any panel where the proportion of foreground pixels is greater than the background one.

The panel background color is not homogeneous, and it actually consists of a wide range of colors usually spread around a unique value. When working in the LAB color space, we can compute a histogram of the two chrominance components in a 3D plot which allows us to analyze graphically how pixels on the background image are distributed. One example of this can be seen on Fig. 6.

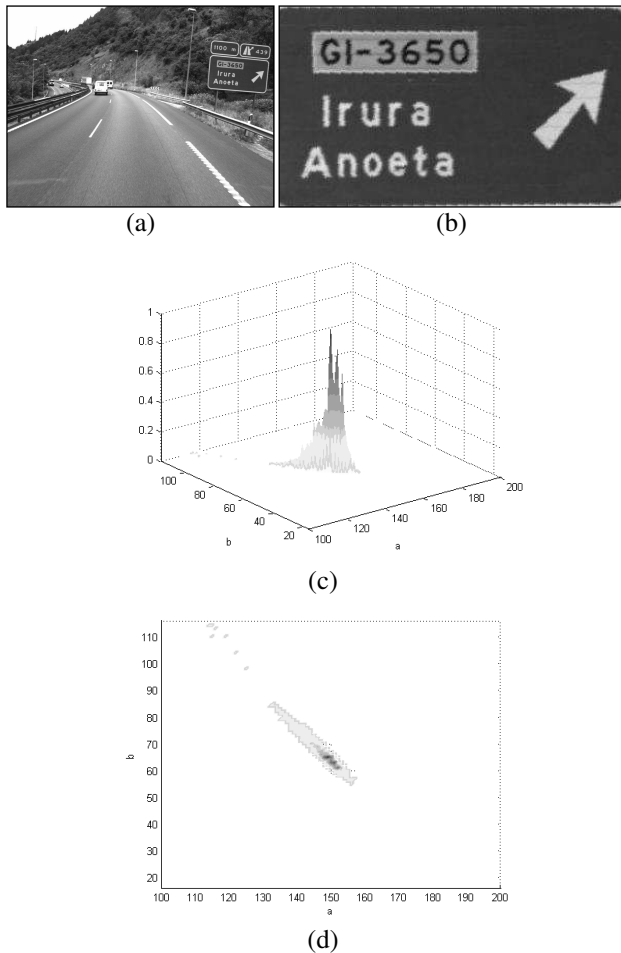


Fig.6. 3D color histogram distribution of a road panel. (a) Original image. (b) Extracted and reoriented panel. (c) AB 3D plot (d) AB 3D plot top view.

A closer look of the panel histogram shows that there is always a maximum peak which represents the most common background color. Hence, the number of pixels with this color is always greater than anyone with any other chrominance combination. Following this observation, it's easy to know where the panel background lies, and thereby, how to apply a suitable segmentation that could allow us to distinguish the objects on the plane. The main consequence of this is that we don't need to know which color the panels are, we just examine the color histogram and then we apply a suitable segmentation. Several techniques have been studied in order to find the best geometrical approach that could define the

region which best identifies the panel background color and separates it from the rest. The first solution adopted was to obtain a contour curve at a specified level in the 3D histogram. At the beginning we thought it was going to offer a good resolution to the problem but in some cases foreground colors are also shared by a substantial amount of pixels and by that they distort the geometrical boundary we are looking for, so we rejected the idea.

It is important to understand how LAB works in order to find a well suited solution to the problem. If we plot a 2D graphical representation of every possible combination of chrominance components A and B, it can be showed that rotating a radial division with one vertex fixed in the center of the AB plane let us to select each of the gamut of colors of interest at a time (Fig. 7.a), except for the gamut of grey levels, where a polygonal approximation of a circle can be used (Fig. 7.b).

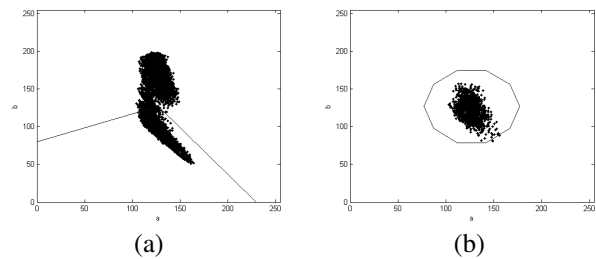


Fig.7. AB chrominance color distribution and segmentation (a) Example of blue thresholding trough a radial division. (b) Example of grey level thresholding trough a polygonal approximation of a circle.

As it has already been commented, most common panel background colors are blue and white, and after various experimental tests we have been able to obtain which vertexes are better suited for selecting them. Each pixel which lies outside the region is classified as an object, the ones which lays inside the region is classified as background.

4.2 Luminance analysis and segmentation

Besides the color, the illumination of the road panel also plays an important role on the segmentation. The same panel can be found in different intensity conditions in different scenes and they can be wrong classified if luminance is not taken into account. The luminance distribution along a real road panel like chrominance is not homogeneous and the different pixels within the panel background have not the same luminance. This implies we need to know where we should establish again a threshold to discriminate the background from the rest. After studying and analyzing our panel images test-bench we have observed that every luminance

histogram presents a clear maximum peak around a point which corresponds to the most repeated luminance value of the panel.

Taking a threshold above and another one below the luminance histogram maximum, we can decide what background is and what background is not. This can be observed in Fig. 8.

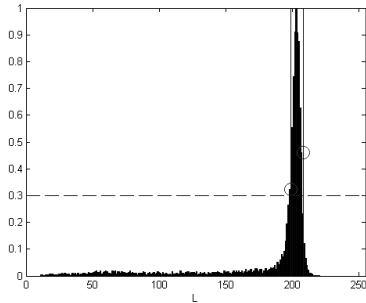


Fig.8. Luminance thresholding example where circles represent last value above the threshold chosen in this case at 0.3 (normalized value).

5 Characters arrangement

Once we have segmented the image we need to detect every potential character on the panel. We compute connected components labeling so that we can tag every object, and then write down its physical and geometrical properties. We apply then a filter to reject those with uncommon aspect relation, size and position, such as arrows, signs, line divisions, or noisy fragments. The objective is to gather to the extent possible only the characters on the panel.

Those objects which successfully pass through the filter are likely to be characters, but maybe they are not. In the worst case OCR systems would help us to know which are and which are not. In any case, every potential character is then arranged in horizontal rows by means of a partitional clustering. What we want is to guess how many text lines are actually on the panel and how many text lines are actually on the panel and assign each object to one of them. To do that, we first choose a random object and create a row with top and bottom boundary equals to those from the rectangle in which the object is inscribed. We continue picking random objects and computing their center to see if there is any row to which it could belong. If it doesn't, we create a new row.

Once every potential text line has been discovered, and each potential character has been arranged into one of them, we put characters and text lines in order horizontally and vertically respectively. This is done by means of quickSort, one of the fastest sort algorithms on average [12], so excellent performance rates are achieved. Next we break lines into words. To do that we just compute an array with the distances

between characters of each row and pay special attention to abrupt hops. Choosing a convenient threshold allow us to determine when these changes take place and then place each potential character into its corresponding word.

6 Experimental results

There were various parameters we could adjust in order to control the overall performance. Some of the most important include the vertexes of the radial division in the chrominance segmentation and the choosing of the thresholds in the luminance segmentation. We had also to choose some thresholds for the area, size, and relation aspect filter. Besides of that, some other parameters were adjusted for the character arrangement into words.

After testing with various road sign panels we have finally chosen the configuration which better results offered. An example can be visualized on Fig. 9.

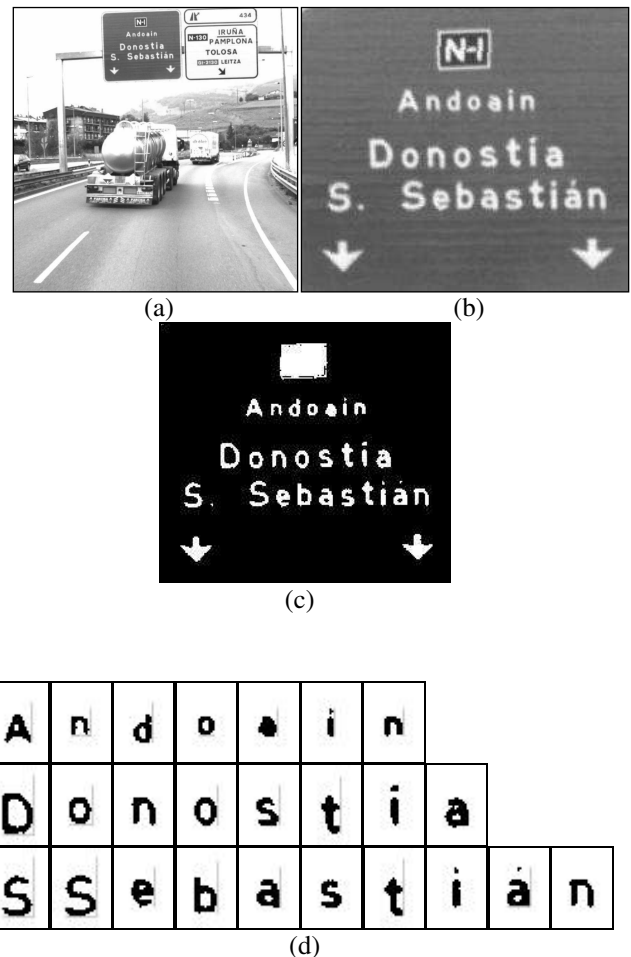


Fig.9. Complete process example. (a) Original image. (b) Blue panel extracted and reoriented. (c) Blue panel segmented (d) Characters extraction and arrangement.

7 Conclusions

This paper describes a method for the detection and extraction of text characters on road traffic panels. OCR systems can thus recognize different characters one by one easily. Panel text strings provide useful information about the route and their recognition sets the first step to deal with the complete interpretation of traffic route panels. As we have seen before, the system is invariant to rotations, changes of scale and different positions. Adaptive chrominance and luminance segmentation let us to deal with so many different kinds of outdoors environments, and in fact it makes the system independent to the text and panel background color and intensity values. Future work lines could be developed around the classification and interpretation of different road guidance panels and its combination with the recognition of basic traffic signs.

Acknowledgements:

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

References:

- [1] A. de la Escalera, J.M. Armingol, and Mata M. Traffic sign recognition and analysis for intelligent vehicles, *Image and vision computing*, 21:247-258, 2003.
- [2] A. de la Escalera, J.M. Armingol, J.M Pastor, and F.J Rodríguez. Visual sign information extraction and identification by deformable models for intelligent vehicles, *IEEE trans. on intelligent transportation systems*, 5(2):57-68, 2004.
- [3] C. Fang, S. Chen, and Fuh C. Road sign detection and tracking, *IEEE trans. on vehicular technology*, 52(5):1329-1341, September 2003.
- [4] C.Y. Fang, C. S. Fuh, P.S. Yen, S. Cherng, and Chen S.W. An automatic road sign recognition system based on a computational model of human recognition processing, *Computer Vision and Image Understanding*, 96:237-268, August 2004.
- [5] A. Farag and A. E. Abdel-Hakim. Detection, categorization and recognition of road signs for autonomous navigation, *Proc. of ACIVS*, pages 125-130, September 2004.
- [6] S. Lafuente-Arroyo, P. Gil-Jiménez, R. Maldonado-Bascón, F. López-Ferreras and S. Maldonado Bascón, Traffic sign shape classification evaluation I: SVM using distance to borders. *Proc. IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005.
- [7] J. Gao and J. Yang, An adaptive algorithm for text detection from natural scenes, *Proceedings of CVPR 2001*, Vol.2, pp. 84-89.
- [8] H. Li, D. Doermann, O. Kia, Automatic text detection and tracking in digital video, *IEEE Trans. on IP*, 9(1), pp 147-156, 2000.
- [9] H. Liu, D. Liu, and J. Xin, Real time recognition of road traffic sign in motion image based on genetic algorithm, *Proc. Of the 1st Int. Conference on Machine Learning and Cybernetics*, pp. 83-86, November 2002.
- [10] P. Gil-Jiménez, S. Lafuente-Arroyo, H. Gómez-Moreno, F. López-Ferreras, and S. Maldonado-Bascón, Traffic sign shape classification evaluation II: FFT applied to the signature of blobs, *Proc. IEEE Intelligent Vehicles Symposium*, Las Vegas, USA, June 2005.
- [11] Hartley, R., and Zisserman, A. *Multiple view geometry in computer vision*, Cambridge University Press, 2004.
- [12] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein, *Introduction to Algorithms*, Second Edition, MIT Press and McGraw-Hill, 2001, Chapter 7: Quicksort, pp.145-164.