Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems, St. Louis, MO, USA, October 3-7, 2009

Automatic Information Extraction of Traffic Panels based on Computer Vision

Á. González, L.M. Bergasa, M. Gavilán, M.A. Sotelo, F. Herranz, C. Fernández

***Abstract*—Computer vision systems used on road maintenance, either related to signs or to the road itself, are playing a major role in many countries because of the higher investment on public works of this kind. These systems are able to collect a wide range of information automatically and quickly, with the aim of improving road safety. In this context, the suitability of the information contained on the road signs located above the road, typically known as traffic panels, is vital for a correct and safe use by the road user. This paper describes an approach to the first steps of a developing system which will be able to make an inventory and to check the reliability of the information contained on the traffic panels, and whose final aim is to take part on an automatic visual inspection system of signs and panels.**

***Index Terms*—traffic panels; image segmentation; Hough transform; panel reorientation.**

I. Introduction

DESPITE many works have been developed in the field of traffic sign detection and recognition [1], [2], [3], [4], [5], [6], automatic visual classification of the information format contained on traffic panels have not been thoroughly studied yet. Actually, from our knowledge only two works have been found on this topic [7], [8]. The main reason of the absence of works of this kind is that there is not a global standardization of the format of the information contained on panels, because of the fact that each country has its own signaling regulations. However, automatic classification of road panels can be very useful for inventory and maintenance purposes, or even further, for driver assistance applications. Road panels provide drivers important information about the route by means of text strings and symbols. The aim of the algorithm presented on this paper is to detect, extract and classify the visual information contained on road panels. In addition, this technique is intended to complement a patented visual inspection system of signs and panels called VISUALISE (Visual Inspection of Signs and Panels), which has been developed by the Departments of Electronics at the University of Alcalá together with the companies Euroconsult S.A. (http://www.euroconsult.es) and 3M-Spain (http://www.3m.com). Since some aspects of this work are related to the VISUALISE system, a short description of this system is going to be held on section II, while a review of the state of the art on road panels recognition is going to be shown on section III. The algorithm for extracting text on traffic panels is going to be described on section IV. Finally, the experimental results obtained as well as the conclusions are going to be detailed on section V and VI respectively.

Departamento de Electrónica
Escuela Politécnica Superior, Universidad de Alcalá.
Alcalá de Henares, Madrid, Spain.
E-mail: {alvaro.g.arroyo;bergasa;miguel.gavilan;sotelo}@depeca.uah.es

II. Visualise System

VISUALISE system is currently running on its first year of operation in Spain. This system is able to calculate the retrorreflection curves of signs and traffic panels automatically. The purpose of this system is to analyse such curves in order to decide if they fulfill the Spanish regulations related to traffic signaling.

The process can be divided into two steps. First of all, a series of video sequences of the roads under review are recorded at night. The reason of recording each sequence at night is because an active infrared illuminator is used and the influence of the environmental lighting is lower. Later, each sequence is processed and quality numbers for every sign on the road are obtained automatically. The recording system is embedded on a vehicle. It consists on a couple of cameras which are part of a stereo system, an infrared illuminator and a synchronization hardware. The result of the recording is two video sequences, one for the left camera and other for the right camera, which are an intermittent succession of illuminated by the infrared beam and non-illuminated frames. In order to minimize the effect of the environment, the measures of each sign and traffic panel are obtained from the difference between one illuminated frame and the previous non-illuminated frame. In addition, the system has an odometer and a GPS receiver which are used, together with the stereo system, to locate traffic signs and panels. As if that were not enough, the recording process can be easily supervised by an operator thanks to a tactile screen located in front of the co-driver seat. Fig. 1 shows a diagram of the vehicle and the position of the infrared illuminator over the vehicle and the cameras inside it.

On the other hand, the processing software is able to detect and track not also each sign and panel on each side of the road, but also each traffic panel located in the way of the road using computer vision techniques. The detection algorithm consists of a shape recognition technique based on the Hough transform, while the tracking is carried out by using Kalman filtering. Thanks to some conversion surfaces previously obtained in a calibration setup, the gray-scale luminance

978-1-4244-5521-8/09/$26.00 ©2009 IEEE

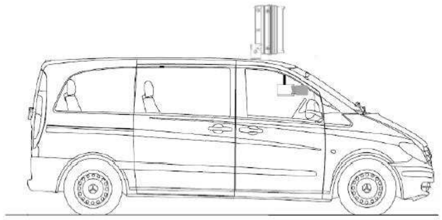


Figure 1: Vehicle used in the recording process

measures of each sign are converted to retrorreflection units, finally obtaining a retrorreflection curve from 15 or 20 meters to over 120 meters. This curve is compared to a decision curve and as a function of this the system indicates if the sign should be replaced or not, according to the regulations. Fig. 2 shows a screenshot of the results got at the processing stage, while a screen capture of the results viewer is shown on Fig. 3.

Figure 2: Processed frame

Figure 3: Results viewer

III. State of the art

Because of the wide diversity of the information contained on traffic panels, as well as the common problems related to outdoor computer vision systems such as: obstructions, shadows and non-controlled environmental lighting, until now there has not been much research on automatic visual classification of the information on road panels. Actually, from our knowledge only two works have been developed. The first one [7] is able to detect candidates to be traffic panels by using an image segmentation in blue and white colours. These candidates are classified by correlating the radial signature of their FFT with the pattern corresponding to an ideal rectangular shape. Later, an image transformation is done by establishing an homography between the original plane and the reoriented one, in order to correct the angular deviation of the panel in the image. Finally, a local adaptative thresholding is applied on the image so that the classification is done for every symbol and character in gray-scale by using a SVM classifier. This algorithm is invariable to traslations, rotations, scaling and projective distortion. However, it is strongly affected by changing lighting conditions. In addition, the segmentation step needs RGB images, although the classification is done by using grayscale images. In addition, this algorithm does not take into account the a priori information that it could be known from the panels, because the kind of information contained on the panel depends on the situation over the panel itself. Thus, a one-against-all classifier is used, but it would be more effective to apply different classifiers depending on the kind of information: alphabetic characters, numbers or symbols. Another problem lies in the fact that it is not done a tracking of the candidates, so the information can be inconsistent between two consecutive frames of the same panel.

The most outstanding work on this topic was carried out by Xilin Chen et al. in [8]. Their algorithm consists of two stages. The first one looks for the traffic panels in the image, while the second one searchs for the text on each panel detected. A priori knowledge of the geometry and other features of the panels is considered to detect them in the image. The text detection is carried out by applying a technique that incorporates edge detection, a segmentation method based on GMM theory and search for lines through a geometrical analysis, so that those characters that belong to the same context are put into groups correctly. The main advantage of this technique is its high compute capacity. In addition, it provides good results under several lighting conditions and it is not affected by rotations and projective distortion. On the other hand, the main drawback of this algorithm lies in the geometrical restrictions used for putting the objects into lines and words, because it does not take into account other features such as size or colour, which can be vital in some contexts. As well as this, the segmentation method based on GMM depends highly on the contrast between foreground and background, which is affected at the same time by lighting conditions.

IV. Text extraction algorithm

An overview of the algorithm developed for extracting text on road panels, which is aimed to be embedded on the VISUALISE system described previously, is going to be held on this section. First of all, it is crutial to make a review of the general features that traffic panels have under the Spanish traffic regulations.

A. Main features of traffic panels.

The design of road panels for the Spanish road network is regulated in [9] and [10]. Their main characteristics can be sum up on the following points:

- All traffic panels, whose aim is to provide useful information to the driver, are rectangular, unlike code signs, which are typically triangular and circular and whose purpose is to point out priorities, prohibitions, obligations and restrictions on the road.
- Road panels have typically blue or white background colors, the foreground objects being white when the background is blue, while they are black in a white background context. The purpose is to provide a high contrast between the background and the foreground. These colors depend on the type of road and the kind of information that is wanted to be supplied. Panels located on highways have a blue background, while panels located on conventional roads have a white background. When a road junction is nearby, the color of the panel that indicates a particular exit to another road depends on the categories of both roads.
- The font used on highway panels is *Highway Gothic*, while on conventional road panels is *Traffic Type Spain*. Fig. 4 shows both fonts.
- Only the first letter of town names are written with a capital letter, except for conventional road panels, which have all their letters capitalised. However, those names which do not correspond to towns are always written with their first letter capitalised, regardless of the type of road. Common nouns are always written with lower case letters.

(a) Highways (b) Conventional roads

Figure 4: Fonts used on traffic panels

A hierarchy can be established to represent the way the information is depicted on a traffic panel. Typically, each panel can be broken into information frames, while each frame consists of individual objects, such as characters, symbols and road identifiers. Furthermore, characters can be put together into words. The aim of the segmentation algorithm is to extract all the elements of the hierarchy by preserving their essential characteristics, in order to get an optimum classification later. Some dynamic structures have been developed to make this hierarchichal division possible. Depending on the connections between each member of the hierarchy, different links between them are implemented. Finally, the result is a hierarchichal tree structure like the one shown in Fig. 5.

B. Image segmentation.

The segmentation is the key step of the algorithm, because the better the segmentation results are, the better the recognition is. Thus, several segmentation techniques have been studied. The method described on this paper turned out to be the best one by far. It is based on a Canny edge detector [11], but a series of improvements have been developed. Firstly, an edge-enhancing filter is applied to the original image of the traffic panel. The convolution mask of this filter is the next one:

$$M = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (1)$$

Secondly, owing to the fact that the global illumination of each panel can be different, thresholds used by the Canny edge detector are calculated dinamically. Computation of these values can be divided into 5 steps:

1) The histogram of the normalized module of the gradient of the image is computed with only 64 bins: $H = [h_i] = [h_1, h_2, ..., h_{64}]$.
2) The cumulative sum of the previous histogram is calculated in the following way: $S = [s_i] = [h_1, h_1 + h_2, ..., h_1 + h_2 + ... + h_{64}]$.
3) An intermediate threshold is computed by using the next expression: $C = N_{rows} \cdot N_{columns} \cdot 0.7$. It only depends on the image size.
4) We search for the first value of S which is higher than C. This value belongs to an specific bin i. Thus, the high threshold used on the Canny detector turns out to be: $T_H = \frac{i}{64}$.
5) The low threshold is then computed via this expression: $T_L = 0.4 \cdot T_H$.

Fig. 7(b) shows an example of this segmentation technique.

C. Panel reorientation.

Once the image has been segmented, a transformation is applied to the resulting image so that inclination of the panel in the image is corrected. We look for the horizontal borders of the panel in the image, in order to calculate their deviation respect to the horizontal axis. This process can be divided into 4 stages:

1) The Hough transform is calculated over the top third-part of the segmented image but only those lines whose angles respect to the horizontal axis are minimum, are taking into account.
2) The same process is done over the bottom third-part of the segmented image.
3) The average angle α of the lines under study is supposed to be the inclination of the panel in the image.
4) A rotation is applied to the segmented image. The transformation matrix is:

$$T = \begin{bmatrix} \cos\alpha & \sin\alpha & (1-\cos\alpha)x_0 - \sin\alpha \cdot y_0 \\ -\sin\alpha & \cos\alpha & \sin\alpha \cdot x_0 + (1-\cos\alpha)y_0 \end{bmatrix} \quad (2)$$

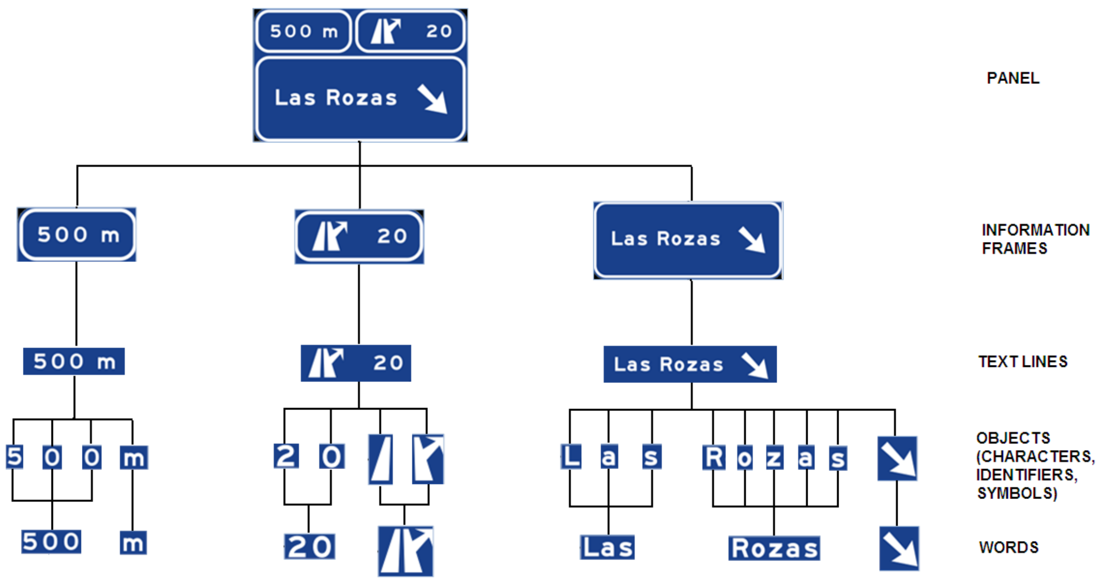


Figure 5: Hierarchical division of the information contained on the traffic panel

being (x_0, y_0) the center of the image. The relation between pre-rotation coordinates $(x_i.y_i)$ and post-rotation coordinates (x_f, y_f) is:

$$\begin{bmatrix} x_f \\ y_f \end{bmatrix} = T \cdot \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (3)$$

Fig. 6 shows an example of a panel reorientation.

(a) Segmented image (b) Lines used for the angle calculation (c) Reoriented image

Figure 6: Correction of the panel inclination

D. Extraction of every part of the panel.

Because of the fact that traffic panels are made of horizontal bands of a reflective material, some undesirable horizontal bordes can appear in the segmented image occasionally. In order to avoid a wrong extraction of the foreground objects, horizontal edges and vertical edges are computed separatedly, as shown in Fig. 7(c) and Fig. 7(d). Then, we look for horizontal lines in the horizontal edges image, as shown in Fig. 7(e), taking them into account as one in case they are too close. The same process is carried out in order to find the vertical borders of each information frame of the panel under study, but using the vertical edges image. The result is shown in Fig. 7(f). Each information frame of the panel is then separated. As some traffic panels have two upper information frames, we look for a clear vertical separation on the second third-part of the upper information frame. Later, text lines are searched on every information frame separatedly. What we do is to compute the projection on the vertical axis of every vertical edge point (Fig. 7(g)). Typically, zero-crosses mean a different line. However, the average height of all the text lines is computed, in case it is necessary to split too wide text lines into two or more lines. Fig. 7(h) shows the limits of each text line: the upper limit with green color and the lower limit with red color. Once we have found every text line, we look for objects on each one by using the projection on the horizontal axis of every horizontal edges point, as shown in Fig. 7(i). We achieve the separation of each object by looking for zero-crosses. When a found character is too wide respect to the line height, we break it into small objects by searching for minimums on the projection. Therefore, we get the position of each object, as shown in Fig. 7(j).

V. Experimental results

A series of experiments have been carried out with different images of 105 traffic panels located above the road (77 white background panels and 28 blue background panels), obtained from a distance of 15 or 20 meters, when the panels are detected for the first time, up to 50 meters. This distance range has turned out to be optimum, because the panels are neither completely detected nor correctly illuminated by the infrared illuminator below 15 meters, while they are too far away to distinguish the characters over 50 or 60 meters. Altogether, 559 images have been used, which 412 correspond to white background panels and 147 correspond to blue background panels. Figure 8 shows some examples of the source images used and the results obtained, while table I shows a series of statistics obtained from this set of source images. It is important to notice that, because of the hierarchichal division implemented, a wrong extraction of an upper member of the hierarchy can affect to the detection of the lower members of the hierarchy. Thus, detection rate decreases while we go down on the above-mentioned hierarchy, especially from text lines downwards.

	White background panels	Blue background panels	Global results
% Correct detection of information frames	95.21 %	85.78 %	93.14 %
% Correct detection of text lines	98.39 %	99.07 %	98.56 %
% False positives in text lines extraction	0.60 %	0.47 %	0.56 %
% Correct detection of objects	97.08 %	95.35 %	96.64 %
% False positives in objects extraction	1.66 %	1.89 %	1.72 %
% Correct detection of words	86 %	79.83 %	84.35 %
% False positives of words extraction	3.18 %	3.06 %	3.15 %

Table I: Results obtained

VI. Conclusions and future work

The algorithm described on this paper has turned out to be really effective to extract the information contained on road panels in a wide variety of lighting conditions. It is not only able to locate the characters and symbols of a panel with an impressive accuracy when the contrast between background and foreground is good, but also in those images which have a really low contrast, since a noise tolerant adaptive segmentation method has been developed. Unlike other techniques, this one does not need to work with color images. In addition, it takes into account the a priori knowledge of the traffic panels, especially when separating the information frames, in order to get an optimum result. While the existing methods are able to locate the 88.9% of the information correctly with a false hit rate up to 9%, the proposed technique has a hit rate of 96% and a false hit rate less than 2%. In addition, this approach uses simple computer vision techniques such as edge detectors together with basic geometric features, while the existing methods make use of more complicated algorithms with a higher computational cost, such as GMM or FFT. Furthermore, it is unaffected by changing conditions of the illumination, although the information extraction process can be influenced by light reflections. However, the purpose is to do a multiframe analysis of the results in the future, in order to make the algorithm stronger against this kind of situations. As well as this, the next step of the algorithm would consist in recognising each extracted object. We propose to use different classifiers depending on the kind of object, thus, a classifier for characters, a classifier for numbers, another one for symbols and so on. In this case, it would be necessary to use the a priori knowledge of the road panels theory, because there are different type of objects depending on their situation over the panel itself. To conclude, the proposed technique can be easily adapted to other traffic signaling regulations, because it does not apply a segmentation method based on color, and it only works with geometrical features.

Acknowledgment

This work has been financed with funds from the Ministerio de Educación y Ciencia through the project DRIVER-ALERT (TRA2008 - 03600), as well as from the Comunidad de Madrid through the project Robocity2030 (CAM - S - 0505 / DPI / 000176). The authors would also like to thank all the people that have worked on the development of the VISUALISE system for their help and support in making this work possible.

References

[1] A. de la Escalera, J.M. Armingol and M. Mata, "Traffic sign recognition and analysis for intelligent vehicles," Image and vision computing, vol. 21, pp. 247–258, 2003.

[2] A. de la Escalera, J.M. Armingol, J.M. Pastor and F.J. Rodriguez, "Visual sign information extraction and identification by deformable models for intelligent vehicles," IEEE Trans. on Intelligent Transportation Systems, vol. 5, pp. 57–68, 2004.

[3] C. Fang, S. Chen and C. Fuh, "Road sign detection and tracking," IEEE Trans. on Vehicular Technology, vol. 52, pp. 1329–1341, 2003.

[4] C.Y. Fang, C.S. Fuh, P.S. Yen, S. Cherng and S.W. Chen, "An automatic road sign recognition system based on a computational model of human recognition processing," Computer Vision and Image Understanding, vol. 96, pp. 237–268, 2004.

[5] M.A. García Garrido, M.A. Sotelo and E. Martín-Gorostiza, "Fast traffic sign detection and recognition under changing lighting conditions," IEEE Intelligent Transportation Systems Conf., 2006.

[6] A. Farag and A.E. Abdel-Hakim, "Detection, categorization and recognition of road signs for autonomous navigation," Procs. of ACIVS, pp. 125–130, 2004.

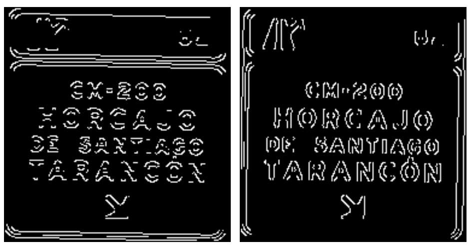
[7] A.V. Reina, R.L. Sastre, S.L. Arroyo and P.G. Jiménez, "Adaptive traffic road sign panels text extraction," Proc. of the 5th WSEAS Int. Conf. on Signal Processing, Robotics and Automation, pp. 295–300, 2006.

[8] X. Chen, W. Wu and J. Yang, "Detection of text on road signs from video," IEEE Trans. on Intelligent Transportation Systems, vol. 6, pp. 378–390, 2005.

[9] Ministerio de Fomento, "Norma 8.1-IC, señalización vertical, de la instrucción de carreteras," Boletín Oficial del Estado, no. 25, pp. 4049–4106, 2000.

[10] Ministerio de la Presidencia, "Real decreto 1428/2003, reglamento general de circulación," Boletín Oficial del Estado, no. 25, pp. 45684–45772, 2003.

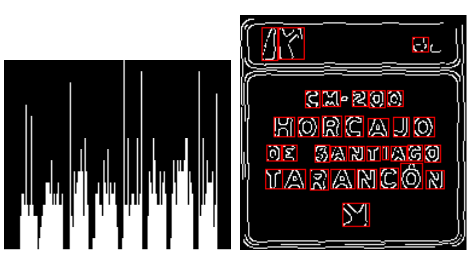
(a) Original image (b) Segmented image



(c) Horizontal edges (d) Vertical edges

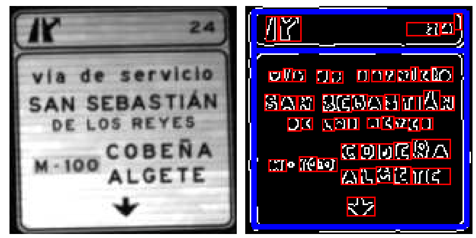
(e) Horizontal lines (f) Information frames

(g) Lower frame vertical projection (h) Text lines

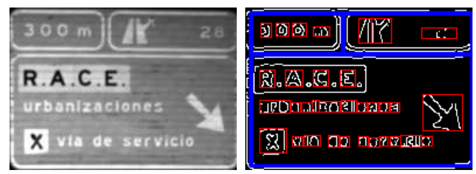


(i) Horizontal projection of the 5th text line (j) Objects detected

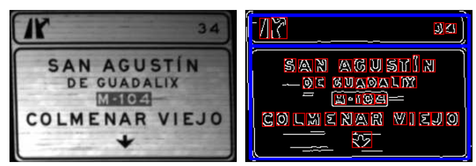
Figure 7: Image segmentation



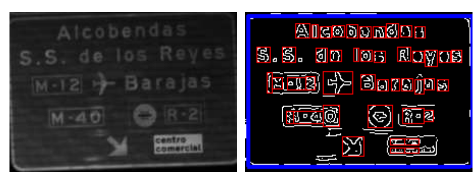
(a) Ex.1 .Original image (b) Ex. 1. Frames and objects detected



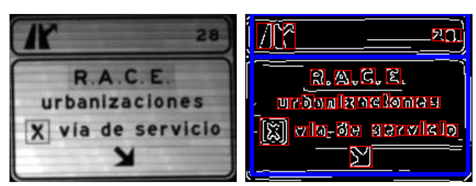
(c) Ex. 2. Original image (d) Ex. 2. Frames and objects detected



(e) Ex. 3. Original image (f) Ex. 3. Frames and objects detected



(g) Ex. 4. Original image (h) Ex. 4. Frames and objects detected



(i) Ex. 5. Original image (j) Ex. 5. Frames and objects detected

Figure 8: Some examples of the results obtained