2010 13th International IEEE
Annual Conference on Intelligent Transportation Systems
Madeira Island, Portugal, September 19-22, 2010

TC2.6

Automatic Information Recognition of Traffic Panels using SIFT descriptors and HMMs

Á. González, L.M. Bergasa, J. Javier Yebes, M.A. Sotelo

Abstract— **This paper presents an algorithm to detect and recognize the information contained in road panels. The aim of this work is to complement the functionality of a traffic signposting inspection system based on computer vision, which is able to collect data related to the maintenance state of traffic signs and panels automatically. In this context, not only a good visibility of the panels is vital for a safe use by road users, but also the suitability of the information contained in the traffic panels. The algorithm presented here, which is based on SIFT descriptors to recognize single characters and also on HMMs to recognize whole words, will be able to make an inventory of the information contained in traffic panels with the aim to check its reliability and brevity automatically. Experimental results and conclusions obtained after analysing a diverse set of real images show the effectiveness of the proposed method.**

I. INTRODUCTION

During the last years, a thorough research has been made in the field of traffic sign detection and recognition. A few examples of the developed works in this area are [1], [2], [3], [4], [5] and [6]. However, automatic visual recognition of the information contained in traffic panels have been hardly tackled. Actually, from our knowledge, there are only two works that deal with this problem [7], [8]. There are several reasons of the absence of works on this topic. First of all, changing lighting conditions are a major problem in outdoor environments. In addition, traffic signs and panels are typically occluded due to the presence of objects between the cameras and the traffic signposts. However, the main reason of the absence of works on automatic information extraction of traffic panels based on computer vision is that there is not a global standardization of the format of the information, because each country has its own traffic signposting regulation. It is even common that different ways of depicting the information on traffic panels coexist in the same country because of the fact that signposting regulation changes but traffic panels are not usually replaced.

Automatic classification of road panels can be very useful for inventory and maintenance purposes, and even further, for driver assistance applications and autonomous vehicles. Road panels provide important information on routes by means of text strings and iconic symbols. The aim of the algorithm presented in this paper is to detect, extract and recognize the visual information contained in road panels. This technique is intended to complement a patented automatic visual inspection system of signs and panels called VISUALISE (VISUAL Inspection of Signs and panEls), which has been developed by the Robesafe Research Group at the University of Alcalá together with a series of recognized and prestiged companies in the road safety and inspection industries, such as Euroconsult[1], 3M-Spain[2] and Safecontrol[3].

VISUALISE is an automatic inspection system, mounted onboard a vehicle, which performs vertical signposting inspection tasks at conventional driving speeds. This system is able to compute the retroreflection values at several distances of traffic signs and panels automatically. The purpose of this system is to analyse such curves in order to decide if they fulfill the regulations related to traffic vertical signposting. Therefore, VISUALISE allows for an improvement in the awareness of road signposting state, supporting planning and decision making on the management and infrastructure operators' side.

The VISUALISE system is based on the light retroreflection principle. It uses an active infrared illuminator with perfectly known features as pattern light source. Part of the infrared light that comes into contact with the traffic signs and panels is reflected. The reflected light is then captured by an stereoscopic system made up of two high-resolution cameras. As a consequence, the luminance level of the traffic signs, which is given in grey-level units by the cameras, is directly proportional to the grade of luminance measured in candels per square meter units (cd/m^2). The relation between the luminance measure and the retroreflection value is defined by considering the distance and angular orientation between the light source, the retroreflective material and the measurement system. This relation (luminance-retroreflection) is set through a prior calibration process.

The inspection process can be divided into two steps. Firstly, a series of video sequences of the roads under review are recorded at night. The reason of recording each sequence at night is because active infrared illumination is used and also the influence of the environmental lighting is lower. Later, each sequence is processed and quality numbers for every sign and traffic panel on the road are computed automatically. Then, these numbers are compared to the values stated by the corresponding regulation, thus indicating if the sign should be replaced or not and its degree of deterioration.

In this work, we present an algorithm to recognize the information contained in traffic panels in order to complement the functionality of the VISUALISE system. The algorithm is

Authors are with the Department of Electronics, Escuela Politécnica, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, SPAIN. Email: {alvaro.g.arroyo;bergasa;javier.yebes;sotelo}@depeca.uah.es

[1] http://www.euroconsult.es
[2] http://www.3m.com/es
[3] http://www.safecontrol.com.ar/

based on SIFT (Scale-Invariant Feature Transform) descriptors to recognize single characters and symbols, and it is also based on HMMs (Hidden Markov Models) to recognize single words. Therefore, a review of the state of the art on road panels recognition, on SIFT descriptors and on HMMs is going to be shown in section II. A short description of the text segmentation algorithm is going to be presented in section III, while the method to classify and recognize characters and words is going to be described in sections IV and V. Finally, experimental results and conclusions are going to be detailed in sections VI and VII respectively.

II. STATE OF THE ART

Because of the wide diversity of the information contained in traffic panels, as well as the usual problems related to outdoor computer vision systems such as occlusions, shadows and non-controlled lighting conditions, to date there has not been much research on automatic visual classification of the information contained in 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 for blue and white colours. These candidates are classified by correlating the radial signature of their FFT (Fast Fourier Transform) with the pattern corresponding to an ideal rectangular shape. Later, an image transformation is carried out 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 grey-scale by using a SVM classifier. This algorithm is invariant to traslations, rotations, scaling and projective distortion. However, it is severely affected by changing lighting conditions. In addition, the segmentation step needs RGB images, although the classification is done by using grey-scale images. Another point is that 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 in 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 there is no tracking of the candidates, so the information can be inconsistent between two consecutives frames of the same panel.

The most outstanding work on this topic is described 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 (Gaussian Mixture Models) 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 computational capacity. In addition, it provides good results under different 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.

As there are not many works on automatic recognition of the information contained in traffic panels, we have taken OCR (Optical Character Recognition) systems as a starting point. However, given the vast number of papers published on OCR every year, it is impossible to include all the available methods in this section. Therefore, an overview of the main OCR techniques is going to be held.

A character recognition system is generally comprised of a feature extraction stage and a classification step. In feature extraction, the aim is to represent the objects to be classified in terms of some quantifiable measurements that may be easily used in the classification stage. There are many feature extraction methods [9] including template matching, deformable templates, unitary image transforms, graph description, projection histograms, contour profiles, zoning, geometric moment invariants, Zernike moments, spline curve approximations, Fourier descriptors and statistical moments such as the Hu's moment invariants, affine moment invariants and the Tsirikolias-Mertzios moments. On the other hand, the classification techniques include the Euclidean distance measure, cross correlation [10], the minimum discrimination cost classification [11], neural networks, SVM, K-means and genetic algorithms. Recently, a successful feature extraction method, namely SIFT, has been proposed [12]. It has been widely employed in computer vision and object recognition. It has also been applied to handwritten alphabetical character recognition [13] and handwritten chinese character recognition [14].

On the other hand, HMMS are applied in temporal pattern recognition such as speech [15], handwriting [16], gesture recognition [17] and alignment of bio-sequences [18].

III. TEXT EXTRACTION ALGORITHM

The work presented here is based on a previous algorithm developed by the authors [19]. This technique is able to extract the location of the information contained in traffic panels (symbols, number, characters, frame delimiters) automatically, with a hit rate higher than 96% and a false positives rate lower than 2%.

The text extraction algorithm is divided into three main steps. The first one is the segmentation, which is based on a Canny edge detector with a series of improvements: an edge-enhancing filter applied to the image of the road panel, and a dynamic computation of the Canny thresholds in order to achieve a better performance when non-homogeneous illumination of the traffic panel is present, which happens in

the majority of cases. Once the image has been segmented, a geometrical transformation is applied to the resulting image so that the inclination of the panel, due to perspective distortion, is corrected. Finally, the location of the different elements of the panel (delimiters and objects) is carried out. This technique works with horizontal edges and vertical edges of the image separately so as to avoid a wrong extraction of the foreground objects. Firstly, it looks for text lines by projecting vertical edge points on the vertical axis, as zero-crosses mean a different line, and then it looks for objects (characters, numbers and symbols) in each text line by projecting horizontal edge points on the horizontal axis. A couple of examples that show the performance of this text extraction algorithm are shown in Fig. 1.
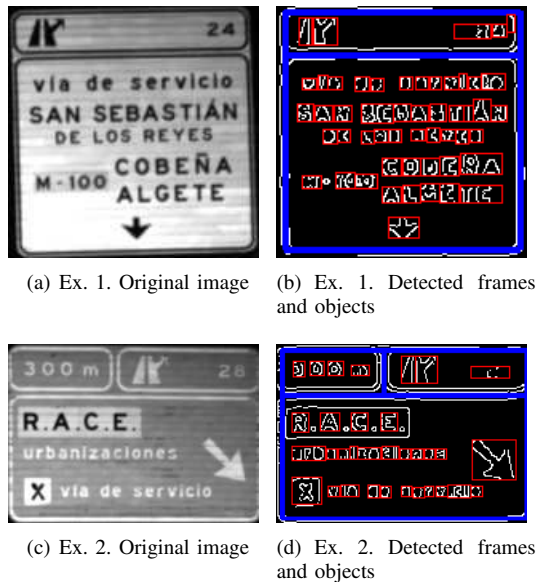

(a) Ex. 1. Original image

(b) Ex. 1. Detected frames and objects

(c) Ex. 2. Original image

(d) Ex. 2. Detected frames and objects

Fig. 1.   Examples of the text extraction algorithm's performance

IV. OPTICAL CHARACTER RECOGNITION

Once we know where the objects are placed, a character recognition technique is applied. Firstly, we have used SIFT keypoints to describe alphabetic characters, numbers and symbols. A SIFT keypoint is a circular image region with an orientation (see Fig. 2). It is described by a geometric frame of four parameters: the keypoint center coordinates x and y, its scale (the radius of the region) and its orientation angle. However, as many characters are symmetrical, like 6 and 9 or b and d, we have left out the fourth parameter, the orientation angle, in order to avoid wrong recognition of symmetrical characters.

Then, a descriptor vector for each keypoint is computed. The gradient at each pixel of the keypoint is regarded as a sample of a three-dimensional elementary feature vector, formed by the pixel location and the gradient orientation. Samples are weighed by the gradient norm and accumulated in a 3-D histogram. An additional Gaussian weighting function is applied to give less importance to gradient further away from the keypoint center. Gradient orientations are
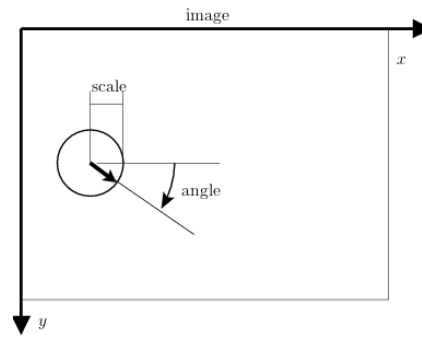

Fig. 2.   SIFT keypoints are circular image regions with an orientation

quantized into 8 bins and the spatial coordinates into 4 each one. The result is a 4×4 array of spatial histograms of the image gradients with 8 bins each histogram. This array forms the SIFT descriptor of the region. Since there are $4 \times 4 = 16$ histograms each one with 8 bins, the descriptor vector of each keypoint has 128 elements.

SIFT keypoints of objects are first extracted from a set of reference images and the corresponding descriptor vectors are stored in a database. Fig. 3 shows the SIFT keypoints extracted from some of the reference images. At the moment, the set of reference images is formed by 202 images of 32×32 pixels. These images correspond to all the uppercase and lowercase letters, all the numbers since 0 to 9 and also 13 common symbols that are usually depicted on a traffic panel. In addition, each character, number and symbol have samples for white-background panels and blue-background panels. This means that we are using only one or at maximum two reference images for each object. This is one of the main advantages of using SIFT as local feature descriptor.

(a)   (b)   (c)   (d)

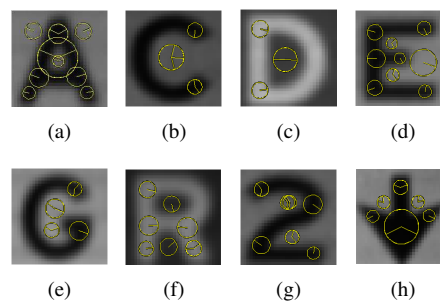(e)   (f)   (g)   (h)

Fig. 3.   SIFT keypoints of some of the reference images

Then, a character or symbol is recognized in a new image by individually comparing each feature vector from the new image to the database and finding candidate matching features based on Euclidean distance of their feature vectors. Therefore, each feature vector from the new image votes for a certain object from the database. We sort each candidate in ascending order by taking into account the Euclidean distance and we compute the ratio between the distance to the first candidate and the distance to the second candidate. Here, second candidate alludes to the following candidate that is different to the first one in this ascending classification. This
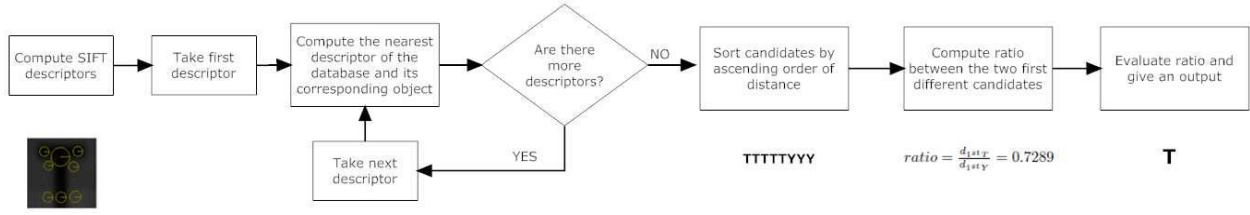
Fig. 4. Block diagram of the process of classifying an object

process is shown in Fig. 4. We have seen that if the ratio is higher than 0.4, the first candidate is the right solution. However, if the ratio is between 0.1 and 0.4, either the first candidate or the second one can be the solution to the OCR problem in an equally probable way, while the first candidate is not the solution at all if the ratio is lower than 0.1. In this case, we give an unknown character as solution ("?").

In order to increase the effectiveness of the recognizer, we use the a priori information that we know about traffic panels, which can be found in the corresponding regulation in law [20], [21], because certain symbols and characters are only located in determined parts of particular types of panels. For instance, Fig. 1(c) shows a panel that can be divided into three subpanels: two on the upper side and one on the lower side. The current regulation in law states that only certain kind of information must be depicted on the upper-side subpanels. Therefore, we apply restrictive recognizers to these subpanels, as only numbers, certain symbols and certain characters can be found there. On the other hand, sometimes the previous step of dividing the panel into subpanels is inaccurate, because the lines that delimit these subpanels are not correctly detected. In this context, we get feedback from the recognition step in order to delineate the subpanels limits correctly.

V. WORD RECOGNITION

The output characters given by the OCR are just an estimation. Therefore, once the character recognizer has been applied, the next step is to identify the words. For this purpose, a word recognizer based on HMMs is used. A dictionary, which includes all the words that the system is able to recognize, has been created. Each word from the dictionary is represented by a different model $\lambda(A, B, \pi)$, which is defined by the state transition probabilities $A = \{a_{ij}\}$, the emission probabilities $B = \{b_j(k)\}$ (the probability that symbol v_k is emitted from state j^{th}), and the initial state distribution $\pi = \{\pi_i\}$. Each model has $2 \cdot n$ states, where n is the number of letters of the word, and it can be represented as in Fig. 5. We are using $2 \cdot n$ states because words can be written either in capital letters or in small letters or the first letter in uppercase and the following ones in lowercase. Therefore, there are two states for each letter of the word, one for the capital letter and another one for the lowercase letter. The transition probabilities from the initial state to other states correspond to the initial state distribution.

The size of the state transition matrix A is $(2 \cdot n) \times (2 \cdot n)$. The transition probabilites of the model are defined as
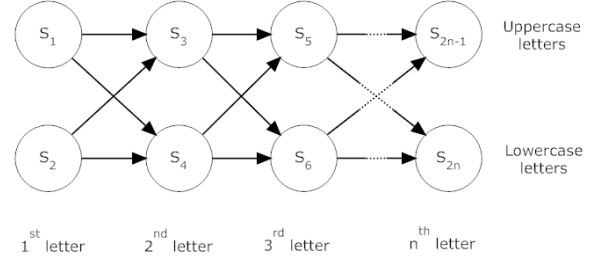

Fig. 5. Hidden Markov Model

follows:

- The model considers that it is more likely to have a capital letter as first letter than a small letter. Thus, $\pi_1 = 0.6$ and $\pi_2 = 0.4$.
- Transition from the first uppercase letter to the second lowercase letter (a_{14}) and from the first uppercase letter to the second uppercase letter (a_{13}). We consider that, if the first letter has been uppercase, it is more probable that the second is lowercase. Thus, $a_{13} = 0.3$ and $a_{14} = 0.7$.
- Transition between letters of the same type from the second letter ($a_{[2i+1][2i+3]}$ and $a_{[2i+2][2i+4]}$, $1 \leq i \leq (n-2)$). We consider that it is unlikely to change from lowercase to uppercase and vice versa. Thus, $a_{[2i+1][2i+3]} = 1$, $a_{[2i+2][2i+4]} = 1$, $a_{[2i+1][2i+4]} = 0$ and $a_{[2i+2][2i+3]} = 0$.

The state transition matrix of a three-letter word is shown in (1):

$$A = \begin{pmatrix} 0 & 0 & 0.3 & 0.7 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (1)$$

On the other hand, the emission matrix B is computed from the confusion matrix C of the OCR. This matrix represents the probabilities $c_{v_i v_j}$ of giving the symbol v_j as output when the input is v_i. There are 55 characters altogether: 27 uppercase letters, 27 lowercase letters and the multi-purpose character "?" which represents any character with equal probability. Therefore, the confusion matrix is a square matrix of 55×55 elements, as in (2):

$$C = \begin{pmatrix} c_{A,A} & c_{A,B} & \cdots & c_{A,Z} & c_{A,a} & \cdots & c_{A,z} & c_{A,?} \\ c_{B,A} & c_{B,B} & \cdots & c_{B,Z} & c_{B,a} & \cdots & c_{B,z} & c_{B,?} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{?,A} & c_{?,B} & \cdots & c_{?,Z} & c_{?,a} & \cdots & c_{?,z} & c_{?,?} \end{pmatrix} \tag{2}$$

Therefore, the ith row of the matrix C is the set of emission probabilities of any state that represents the symbol v_i. For instance, the emission matrix of the model that defines the word "DEL" is shown in (3):

$$B = \begin{pmatrix} c_{D,A} & \cdots & c_{D,Z} & c_{D,a} & \cdots & c_{D,z} & c_{D,?} \\ c_{d,A} & \cdots & c_{d,Z} & c_{d,a} & \cdots & c_{d,z} & c_{d,?} \\ c_{E,A} & \cdots & c_{E,Z} & c_{E,a} & \cdots & c_{E,z} & c_{E,?} \\ c_{e,A} & \cdots & c_{e,Z} & c_{e,a} & \cdots & c_{e,z} & c_{e,?} \\ c_{L,A} & \cdots & c_{L,Z} & c_{L,a} & \cdots & c_{L,z} & c_{L,?} \\ c_{l,A} & \cdots & c_{l,Z} & c_{l,a} & \cdots & c_{l,z} & c_{l,?} \end{pmatrix} \tag{3}$$

The word recognizer works as follows. The output of the OCR $O = O_1 O_2 \ldots O_T$ is the input to the word recognizer. In other words, O is the set of emitted characters. We compute $P(O|\lambda)$ for each model, that is, for each word of the dictionary, by using the forward-backward algorithm. $P(O|\lambda)$ is the probability that model λ has generated the set of observations O. The solution is the word that maximizes $P(O|\lambda)$. Table I shows a series of examples of how the word recognizer based on HMMs works.

TABLE I

WORD'S RECOGNIZER BASED ON HMMS

Original word	Input	Output
Badajoz	??d?joz	Badajoz
Esparragalejo	?spefregafeje	Esparragalejo
Duque	?u?U?	Duque
MERIDA	?Ee?DA	MERIDA

VI. EXPERIMENTAL RESULTS

A series of experiments have been carried out with images of 175 different panels (with different number of text lines and different words for both blue and white background) located above the road, obtained from a distance of 20 meters, when the panels are totally detected on their whole for the first time, up to 50 meters. We work on this distance range because the panels are neither completely detected nor correctly illuminated by the infrared illuminator below 15-20 meters, while they are too far away to distinguish the characters over 50-60 meters. Table II shows the recognition rate for most of the characters, numbers and symbols which we have been working with. These results have been obtained from a set of 4365 observations, giving a global recognition rate of 88.04%. All these results have been obtained without taking into account the a priori information of the panels as it was explained in Section IV. In other words, an one-against-all classifier has been used in order to know the gross discrimination power of the classifier under study. As it can

be seen in Table II, the recognition rate of every character is higher than 80% in most cases. However, there are some characters that have a low recognition rate, such as "1", "i", "I" and "l", because they cannot be distinguished between them as they have similar SIFT descriptors. The only way to differentiate them, especially the number "1" from the other characters, is from context, as the symbol "1" is likely to be with other numbers. However, it is not completely necessary to reach a 100% recognition rate for characters, as there is a subsequent step that consists of recognising a word from the set of single observations by using HMMs, as it was explained in Section V. Table III shows the recognition rate of the implemented word recognizer for both white-background and blue-background panels. A higher words recognition rate is obtained for white-background panels because of the fact that the OCR achieves better results for this kind of panels, since contrast between background and foreground elements is higher in white-background than in blue-background panels (as it can be seen in Fig. 1, contrast between white and black colours is higher than contrast between blue and white colours, as we are working with grey-scale images). These results have been obtained from 470 different words from a dictionary of more than 3200 words.

TABLE II

OCR'S RECOGNITION RATE

A	98.8%	Y	100.0%	z	87.5%
B	85.0%	Z	100.0%	0	84.0%
C	95.1%	a	83.7%	1	69.7%
D	94.0%	b	77.8%	2	87.5%
E	92.0%	c	94.9%	3	97.2%
F	57.1%	d	94.1%	4	100.0%
G	100.0%	e	87.6%	5	90.2%
H	93.3%	f	80.0%	6	77.1%
I	55.0%	g	50.0%	7	75.8%
J	92.9%	h	55.6%	8	68.6%
L	100.0%	i	59.3%	9	90.0%
M	91.7%	j	89.5%	⬈	92.9%
N	84.3%	l	19.7%	⬈	100.0%
Ñ	100.0%	m	84.6%	⬊	98.3%
O	97.0%	n	58.7%	⬇	94.0%
P	97.0%	o	83.6%	⊙	66.7%
Q	88.9%	p	100.0%	⛽	100.0%
R	84.7%	q	100.0%	⌂	100.0%
S	94.7%	r	82.7%	✗	100.0%
T	97.5%	s	73.9%	⊟	100.0%
U	94.3%	t	96.8%	1⃞	100.0%
V	97.4%	u	92.0%	-	-
X	94.4%	v	66.7%	-	-

VII. CONCLUSIONS AND FUTURE WORK

The method described in this paper has turned out to be really effective to recognise the information contained in road

panels in a wide variety of lighting conditions, as it is not only able to recognise most of the words on a traffic panel when the contrast between background and foreground is good, but also in those cases with really low contrast. The reason is that a character recognition method based on SIFT descriptors together with a subsequent word recognizer based on HMMs have been implemented. The SIFT descriptors are invariant to scale, affine distortion and partially invariant to illumination changes. As a result, the character recognizer can robustly identify objects even under partial occlusion. The character recognizer achieves a detection rate higher than 80% for most characters, numbers and symbols, while the word recognizer obtains a recognition rate of 96%. These results have been obtained by using a really small training database with one or two samples for each object in most of the cases. Nowadays, we are working on improving these results by increasing the size and enhancing the quality of the training data. Another way of improving these numbers consists of using the a priori information that we know about the traffic panels, as certain objects are located only at certain parts of the panels, and also by getting feedback from the extracted information at each step of the recognition process. As it has been said before, this method has turned out to be strong when heterogeneous illumination is present, although the information extraction process can be influenced by light reflections. The purpose is to do a multiframe analysis of the results in order to make the algorithm stronger against this kind of temporary effects. Another advantage of the proposed technique is its easy adaptation to other traffic signposting regulations, because it only works with geometrical features and does not use a method based on colour. Finally, as we have geographical measures obtained by a GPS receiver, we are working on using some OGC (Open Geospatial Consortium) services, such as WMS (Web Map Service) or WCS (Web Coverage Service), which can provide geospatial information, like names of municipalities, orographic elements or names of roads, from GPS or UTM coordinates, for increasing the effectiveness of the recognition algorithm. Therefore, a word dictionary limited to the geographical area where the traffic panel is located could be used, instead of an only dictionary for the whole country.

ACKNOWLEDGMENT

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 Transactions on Intelligent Transportation Systems*, vol. 5, pp. 57-68, 2004.

[3] C. Fang, S. Chen and C. Fuh, Road sign detection and tracking, *IEEE Transactions 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 Conference*, 2006.

[6] A. Farag and A.E. Abdel-Hakim, Detection, categorization and recognition of road signs for autonomous navigation, *Proceedings 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, *Proceedings of the 5th WSEAS International Conference 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 Transactions on Intelligent Transportation Systems*, vol. 6, pp. 378-390, 2005.

[9] O.D. Trier, A.K. Jain and T. Taxt, Feature extraction methods for character recognition - a survey, *Pattern Recognition*, vol. 29, pp. 641-662, 1996.

[10] G.L. Cash and M. Hatamian, Optical character recognition by the method of moments, *Computer Vision, Graphics and Image Processing*, vol. 39, pp. 291-310, 1987.

[11] K. Tsirikolias and B.G. Mertzios, Statistical pattern recognition using efficient two-dimensional moments with applications to character recognition, *Pattern Recognition*, vol. 26, pp. 877-882, 1993.

[12] D.G. Lowe, Distinctive Image Features from Scale Invariant Keypoints, *International Journal of Computer Vision*, vol. 60, pp. 91-110, 2004.

[13] J. Rodriguez and F. Perronnin, Local Gradient Histogram Features for Word Spotting in Unconstrained Handwritten Documents, *International Conference on Frontiers in Handwritten Recognition*, pp. 7-12, Montreal, Canada, 2008.

[14] Z. Zhang, L. Jin, K. Ding and X. Gao, Character-SIFT: a novel feature for offline handwritten Chinese character recognition, *10th International Conference on Document Analysis and Recognition*, pp. 763-767, 2009.

[15] B.H. Juang and L.R. Rabiner, Hidden Markov Models for speech recognition, *Technometrics*, vol. 33, no. 3, pp. 251-272, 1991.

[16] J. Hu, M.K. Brown and W. Turin, HMM Based Online Handwriting Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 10, pp. 1039-1045, 1996.

[17] T. Starner and A. Pentland, Real-time american sign language recognition from video using Hidden Markov Models, *Proceedings of International Symposium on Computer Vision*, 1995.

[18] R. Durbin, S.R. Eddy, A. Krogh and G. Mitchison, Biological sequence analysis: Probabilistic models of proteins and nucleic acids, *Cambridge University Press*, 1999.

[19] Á. González, L.M. Bergasa, M. Gavilán, M.A. Sotelo, F. Herranz and C. Fernández, Automatic Information Extraction of Traffic Panels based on Computer Vision, *12th International IEEE Conference on Intelligent Transportation Systems*, 2009.

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

[21] Ministerio de la Presidencia, Real Decreto 1428/2003, reglamento general de circulación, *Boletín Oficial del Estado*, no. 25, pp. 45684-45772, 2003.