ELSEVIER

Image and Vision Computing 21 (2003) 247–258

image AND vision COMPUTING

www.elsevier.com/locate/imavis

# Traffic sign recognition and analysis for intelligent vehicles

A. de la Escalera*, J.M[a] Armingol, M. Mata

*Division of Systems Engineering and Automation, Universidad Carlos III de Madrid, C/Butarque 15, 28911 Leganés, Madrid, Spain*

Received 22 October 2000; accepted 24 October 2002

**Abstract**

This paper deals with object recognition in outdoor environments. In this type of environments, lighting conditions cannot be controlled and predicted, objects can be partially occluded, and their position and orientation is not known a priori. The chosen type of objects is traffic or road signs, due to their usefulness for sign maintenance, inventory in highways and cities, Driver Support Systems and Intelligent Autonomous Vehicles. A genetic algorithm is used for the detection step, allowing an invariance localisation to changes in position, scale, rotation, weather conditions, partial occlusion, and the presence of other objects of the same colour. A neural network achieves the classification. The global system not only recognises the traffic sign but also provides information about its condition or state.
© 2002 Elsevier Science B.V. All rights reserved.

*Keywords:* Object recognition; Genetic algorithms; Neural networks; Traffic sign recognition; Driver support systems; Intelligent vehicles; Intelligent transportation systems

## 1. Introduction

### 1.1. Motivations

Traffic sign detection and recognition have received an increasing interest in the last years. This is due to the wide range of applications that a system with this capability provides:

1. *Highway maintenance.* Nowadays, a human operator has to watch a videotape to check the presence and condition of the signs. It is a tedious task because the signs appear from time to time, and because the operator has to pay great attention. The Esprit European project AUTOCAT [4] presents a van developed for the automatic gathering of the traffic sign position.
2. *Sign inventory.* It is basically the same application but in towns and cities. In this case the environment is more difficult than highways. The signs are not always placed perpendicular to the movement of the vehicles, producing a deformed image of the signs; besides, there are occlusions, and other objects with the same colour. There has been little work in this particular environment, as it will be mentioned later.
3. *Driver Support Systems.* Traffic sign detection and classification is one of the less studied subjects in the field of Driver Support Systems. Research groups have been focused on other aspects, more related with the development of an automatic pilot, as the detection of the road borders [12,35] or the recognition of obstacles in the vehicle's path such as other vehicles [7,15] or pedestrians [15,20]. The future Intelligent Vehicles would take some decisions about their speed, trajectory, etc. depending on the signs detected. Although, in the future, it can be part of a fully automated vehicle, now it can be a support to automatically limit the speed of the vehicle, send a warning signal indicating overspeed, warn or limit illegal manoeuvres or indicate earlier the presence of the sign to the driver. The general idea is to support the driver in some tasks, allowing him or her to concentrate in driving. One of the most important and more general works was described by Estable in Ref. [14] about the research by Daimler-Benz on its vehicle VITA-II. As a particular application, a system for the detection of the Stop sign as a blind people support system is presented in Ref. [39].
4. *Intelligent Autonomous Vehicles.* Mobile robots use landmarks as means for their relocalisation [38]. In the case of artificial landmarks, they can be designed as traffic signs [3,32]. The advantage of this idea is that

* Corresponding author. Tel.: +34-916-249480; fax: +34-916-249430.
*E-mail address:* escalera@ing.uc3m.es (A. de la Escalera).

0262-8856/03/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0262-8856(02)00156-7

IPR2025-00943
Tesla EX1029 Page 1

some precise information is added to the sign, for example, indicating the robot the path to follow or some task it has to perform in a particular location.

### 1.2. Difficulties

The extraction or detection of a traffic sign, for its posterior recognition, presents the same difficulties as object recognition in natural environments:

1. Lighting conditions are changeable and not controllable. Lighting is different according to the time of the day, season, cloudiness and other weather conditions, etc. (Fig. 1a–c).[1]
2. The presence of other objects. Except in the case of highways, the simple case, other objects often surround traffic signs. This produces partial occlusions, shadows, etc. (Fig. 1d and e).
3. It is not possible to generate off-line models of all the possibilities of the sign's appearance, because there are so many degrees of freedom. The object size depends on the distance to the camera; the scale for each axis is different if the camera optic axis is not perpendicular to the sign, producing an aspect modification (Fig. 1g); besides, physical condition of a sign changes depending on its age, accidents, etc. An example of a rotated sign is presented in Fig. 1f.

### 1.3. Objectives

The objectives of the present paper are:

1. The system has to be able to detect traffic signs independently of their appearance in the image. Because of that, it has to be invariant to:
   - Perspective distortion.
   - Lighting changes.
   - Partial occlusions.
   - Shadows.
2. In addition, it has to provide information about the presence of possible problems:
   - Lack of visibility.
   - Bad condition.
   - Bad placement.

## 2. State of the art

There are little work that presents complete algorithms for the detection and recognition of traffic signs. Because of that, the most important research for detection will be presented first, followed by the work dedicated to recognition.

[1] All the images of this article can be downloaded from www.uc3m.es/uc3m/dpto/IN/dpin04/xxx/p2/research/its/Its.html.

### 2.1. Sign detection

Traffic sign detection research can be classified in the two following approaches:

1. Segmentation through colour thresholding, region detection and shape analysis.
2. Segmentation through the border detection in a black and white image and their analysis.

There are two groups within the first approach: those who work with standard colour spaces, and those who developed a more exhaustive colour study. Thus, the RGB space is used in Refs. [11,13,27,40]. Because of the lighting changes problem, they use the relations between the colour components or subgroups within the colour space. Other researchers prefer working with spaces more immune to lighting changes. Although the HSI space is the most used [4,21,25,34], the Luv space has been also used [24].

The second subgroup, those who developed a more exhaustive colour study, noted that the HSI space cannot cover all the possible cases. Thus Aoyagi [1] noted that colour sign segmentation cannot be done because the Hue component changes with distance, weather, and sign age. As a conclusion the segmentation stage is not absolutely reliable in perfectly detecting the sign pixels. Several more complex colour classifications have been proposed to solve this problem. Thus a database for the colour pixel classification was developed in Refs. [36,37]. The use of textures has been proposed in Refs. [27,28], while in Ref. [16] a fuzzy classification is used. Neural networks are proposed in Refs. [6,23,26].

Nevertheless, although these last classification methods are more complex and complete than those that use thresholding with fixed values, they do not take into account the occlusion problem. The presence of an object partially occluding a sign would produce the same effect than a bad segmentation. Except in the case of highways, where due to the sign size and location, occlusions are more difficult, it is a normal case in roads and almost always within cities. Because of this, if the complete algorithm deals with partial occlusions, the colour segmentation stage, although important, is not so decisive as was believed until now.

Among those who work directly with a black and white image are Austerirmeier et al. [5], Besserer et al. [8] and Buker and Mertsching [10] who detect the borders in a pyramidal structure. They do not take occlusions into account. Aoyagi and Asakura [1] present a genetic algorithm (GA) for the sign detection. Their work is explained with more detail because in this article GAs are used too. The goal of their work is to detect speed limit signs. They only work with the bright image because of the Hue variations explained before. After obtaining the laplacian of the original image, there is a thresholding. Those pixels that pass the threshold are analysed later. As

IPR2025-00943
Tesla EX1029 Page 2

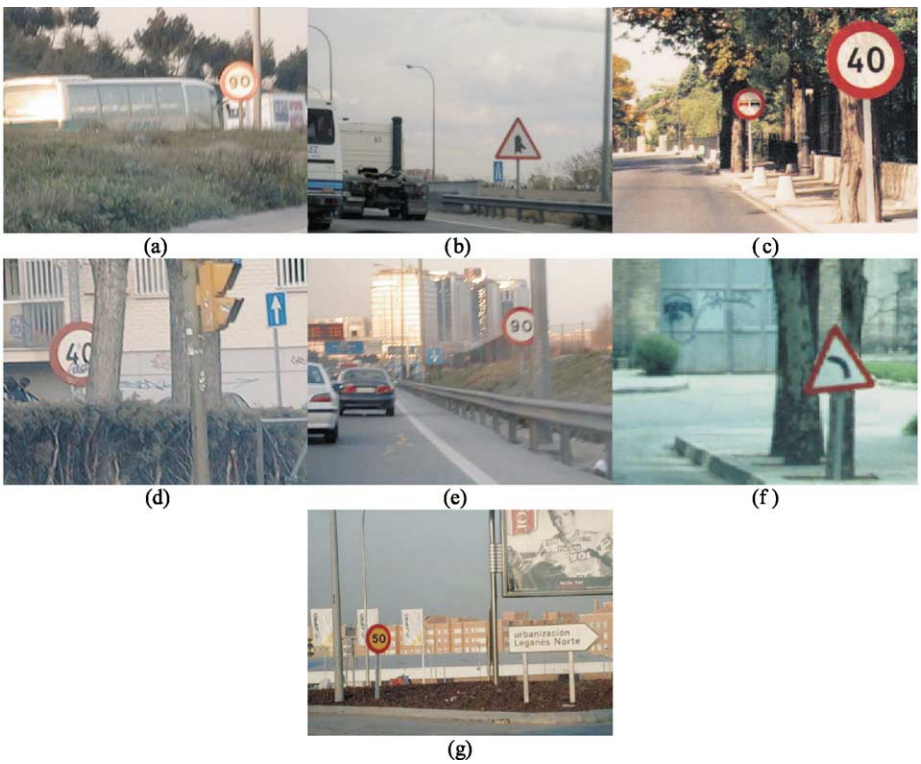


Fig. 1. Traffic sign detection problems. (a) Reflections; (b) not controlled lighting; (c) shadows; (d, e) partial occlusions; (f) sign rotation; (g) shape deformation.

almost all the research groups, they do not take into account different scales for the horizontal and vertical axes; thus they do a matching with a circular pattern. To do this the gene information is the $x$ position (7 bits), the $y$ position (7 bits) and the radius (4 bits). The population is formed of 32 individuals, the selection rate is 30%, 10% the mutation rate and there are 150 iterations. Finally there are multiple cross points. In Ref. [17] several models are generated off-line. They represent the sign borders taking into account scale changes. The object borders presented in the image are obtained and enhanced by a distance transform (chamfer-distance). For the sign detection the models are correlated with the image.

### 2.2. Sign recognition

Once the sign has been detected the recognition is done. In Ref. [21] the regions are obtained after the colour thresholding, and their exterior border is coded; after a complex-log mapping transform (immune to scale and rotation) a FFT is realised. For the final classification a matching with a spectrum database is performed. In Ref. [24] a similar algorithm is adopted but a Back-propagation Neural Network (NN) does the final classification. A Cellular NN localises circular and triangular signs in Ref. [2], however no final classification within each group is performed neither occlusions are taken into account. The input layers of the net are the borders detected in a black and white image and the size of the sign is fixed. The contour is analysed (with no occlusions) in Ref. [27], using the eigenspace of the model bases for the classification. In Ref. [34] the region borders are analysed; after a first classification in triangular and circular signs, a cross correlation makes the final classification. Occlusions and sign deformations are not considered in this work. A Backpropagation NN classifies between speed limit, no-speed-limit and no-sign in Ref. [1]. The input layer is a normalised image of $18 \times 18$ pixels; the hidden layer has 15 neurones whereas the output layer has three neurones. There is no specific mention of occlusions. A Kohonem NN is used in Ref. [29] where the Network is trained considering rotation and occlusion. Another NN, Receptive Field, is proposed in Refs. [30,31] where the inputs are directly the images. In Ref. [13] the borders of the colour-detected regions are analysed through the search of corners with certain angles and analysing the relations between them where a Backpropagation NN performs the final classification. Small rotations and occlusions within the sign are taken into account. In Ref. [9] normalised correlation is used to detect and classify the sign at the same time. The algorithm is immune to lighting changes and occlusions. To detect the difference in the scale and the perspective between the model and the image, the former is changed

on-line by *simulated annealing*. Although the idea is interesting it does not proved to be practical because it has to be done for every sign and therefore the computational cost would be prohibitive.

## 3. Traffic sign detection

Traffic signs are going to be detected trough colour and shape analysis. First the hue and saturation components of the image will be analysed and the regions in the image that fulfilled some colour restrictions will be detected. If the area of one of these regions is big enough a possible sign can be located in the image. The perimeters of the regions are obtained and a global search of possible signs is performed. This search is done with a GA.

### 3.1. Colour classification

The use of colour analysis is basic because traffic signs are designed thinking using colours to reflect the message of the sign. This way, the chosen colours stand out from the environment. HSI is the chosen colour space, in this article, for the colour analysis since it gives different pieces of information in every component. This step has two goals: to trigger the whole algorithm and as an intermediate step to detect the border of the sign.

Only the Hue and Saturation components are taking into account in order to avoid lighting condition. Two look-up-tables (LUTs) are constructed (Fig. 2), one for the Hue component and another for the Saturation one. The general idea is that if one component has a wrong value, it could be helped by the other component. Thus, for the Hue component the sign's red colour has very low, 0°, and high, 360°, values. The correspondence for an image with 255 different values is 0 red, 85 green, 170 blue, 255 red. From two values there is a ramp until the maximum value

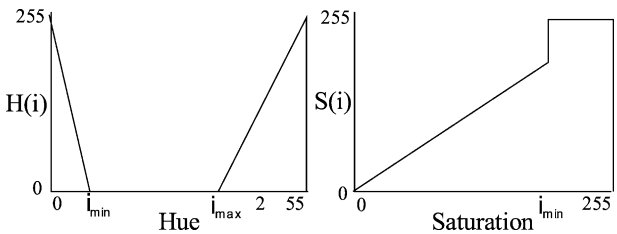


Fig. 2. Colour classification LUTs.

following the formula:

$$H(i) = \begin{cases} 255\dfrac{i_{\min} - i}{i_{\min}} & 0 \le i \le i_{\min} \\ 0 & i_{\min} < i < i_{\max} \\ 255\dfrac{i - i_{\max}}{i_{\max}} & i_{\max} \le i \le 255 \end{cases} \tag{1}$$

where $i$ is the former Hue component value and $H(i)$ the new one.

For the saturation component, its value will be higher as much redness as the sign contains. The LUT will follow a ramp until a saturation value and from that point will have the maximum value, following the formula:

$$S(i) = \begin{cases} i & 0 \le i \le i_{\min} \\ 255 & i_{\min} \le i \le 255 \end{cases} \tag{2}$$

where $i$ is the former Saturation component value and $S(i)$ the new one.

Once both LUTs are applied, the images are multiplied and normalised to the maximum value of 255. The reason to do this instead of choosing some fixed thresholds for both component and through an AND operator form a binary image, is that pixel classification errors are meant to be avoided. Both components help each other and if one has a low value the other can correct it and pass the combined threshold. Results and additional problems can be observed in Fig. 3. It is important to emphasise that a correct classification, although positive, will not be necessary because occlusions will be taken into account. In (a) and

Fig. 3. Colour detection. (a, d) Other objects with the same colour presence; (b) union of two signs to form only one object; (c) some points of the sign are badly classified.

(d) all the pixels belonging to the sign have high values but there are other objects with the same colour; in (b) all the pixels are well classified but there are other objects connected to the sign; in (c) part of the sign has some values much lower than the rest, and the sign will be split in three different objects.

### 3.2. Genetic algorithms

GAs [18] do a parallel search in several directions following an optimisation process that imitate natural selection and evolution. To do this, there is a set of possible solutions (the individuals) that exchange information depending on the fitness of the result in the search of the global maximum. GAs robustness resides in their ability to reach a global maximum surrounded by local ones. This is precisely the traffic sign case as can be seen in Fig. 4a, where the detection of the triangular sign in the image is desired. All the possible triangular signs, only taking into account different scales, have been tested for every pixel (Fig. 4b). The formula to measure similarity between the sign and the image will be explained later. Several local maximum can be observed around the global one. They correspond to the corner of the sign where other triangular signs with smaller size than the real one fit.

Search algorithms have to find a balance between two opposite task: exploration of the whole search space and the exploitation of certain zones. With exploration the search space is covered looking for promising areas in which a more detailed search has to be done; that is the exploitation task, where the best solution is looked for in a zone known as good. The risk is being trapped in a local maximum or minimum. Hashing methods are the extreme case of exploration where gradient-based methods (hillclimbing) are the extreme for exploitation. Although GAs tend to find a balance between those two goals, they are better at exploitation than exploration. The case presented in this article is not an exception and therefore, the danger of a premature convergence has to be avoided, that is equivalent to the loss of genetic richness in a specie.

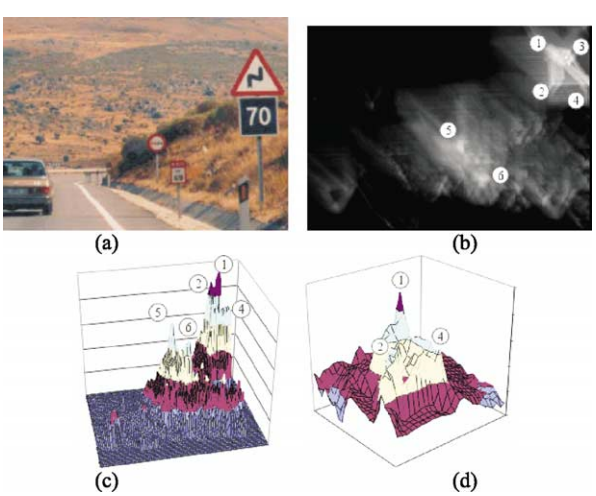


Fig. 4. Local and global maximum. (a) Original image; (b) best fitness for a triangular sign; (c) elevation map of the previous; (d) map detailed of the maximum peaks.

GA steps are:

1. Population initialisation.
2. Fitness evaluation for every individual.
3. Good solution selection to produce a new population.
4. Producing the new population.
5. Evaluation of the new results.
6. Exchanging the old population with the new one.

Steps 3–6 are done a fixed number of times (generations) or until the solution provided by the best individual reach a certain value.

### 3.3. Gene codification

Gene codification starts from a sign model representing a sign at a fix distance and perpendicular to the optical axes. The considered modifications are a change in the position and in the scales, due to the sign being farther or nearer than the model, or because the optical axis is not perpendicular to the sign producing a deformation in it, which is due to the magnification difference for every axis.

All these factors can be expressed if there is an affine transform between the ideal model without deformations and the model that is being looked for in the image:

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Xm \\ Ym \\ 1 \end{bmatrix} \tag{3}$$

The transform coefficients are:

$$\begin{aligned} &a_{00} = E_x \cos\theta \qquad a_{01} = E_x \sin\theta \qquad a_{02} = T_x \\ &a_{10} = -E_y \sin\theta \qquad a_{11} = E_y \cos\theta \qquad a_{12} = T_y \end{aligned} \tag{4}$$

where $T_x$ is the horizontal displacement, $T_y$ the vertical displacement, $E_x$ the horizontal scale, $E_y$ the vertical scale, and $\theta$ is the horizontal rotation.

There are two reasons for coding these last parameters instead of the six transform coefficients: there is a parameter less and, above all, solutions without a physical meaning can be discharged easily.

### 3.4. Initialisation

In a classic GA, the initial population is generated randomly, but, in this case, as some information is known from the colour analysis, some values can be obtained that will be nearer to the final one than a random start. To do this, a thresholding of the colour analysis image is performed and the number and position of the blobs are obtained. If no blob is bigger than a certain value there is no sign and no posterior

analysis is performed. Otherwise, a fixed number of individuals are assigned to every blob (Fig. 5). This way, the presence of enough individuals can be guaranteed despite the presence of bigger objects or occlusions. Thus, Fig. 5, left column illustrates how the speed limit sign has been divided in three parts due to occlusion and colour variations. On the central column there is an object larger than the sign and on the right column there are also multiple objects.

Intuitively, one piece of information of the gene should be blob gravity centre. Two problems appeared when it was done:

1. If there was only one sign, all the individuals had the same value, loosing genetic richness (it could not be changed but by mutation).
2. In the presence of occlusions, it causes the non detection of the sign because the same effect is produced.

As a conclusion, the left upper corner is coded.

### 3.5. Fitness evaluation for every individual

The rule to know the fitness, or how near to the best solution the individual is, is based on the Hausdorff distance [22]. This distance indicates how two shapes differ between them. Thus, if there are two point sets $A = \{a_1, \ldots, a_m\}$ and $B = \{b_1, \ldots b_n\}$ the Hausdorff distance is defined as:

$$H(A, B) = \max(h(A, B), h(B, A)) \tag{5}$$

where:

$$h(A, B) = \max_{a \in A} (\min_{b \in B} \|a - b\|) \tag{6}$$

The $h(A, B)$ function is named the direct Hausdorff distance. It identifies the point belonging to the set $A$ that is farthest under the chosen norm from any point of the set $B$.

A variation is the partial Hausdorff distance where only the $k$ shortest distances are taken into account.

$$h_k(AB) = K^{th}_{a \in A} \min_{b \in B} \|a - b\| \tag{7}$$

With this last distance the measurement can be immune to occlusion and noise. In this case, the two sets are:

1. The borders of the blobs detected in the previous step.
2. The affine transforms coded in the chromosomes of every individual.

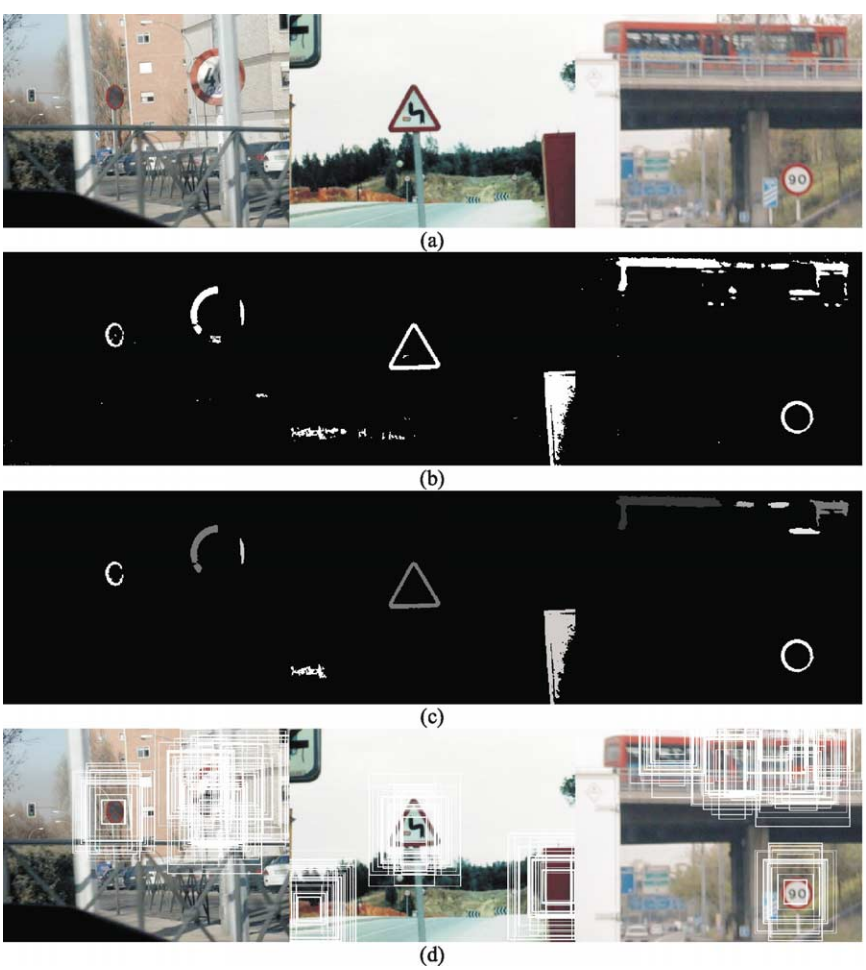


Fig. 5. Initialisation. (a) Original image; (b) thresholded image; (c) blob detection; (d) initial population location.

The Hausdorff distance transform is calculated and the number of points whose value is less than a certain threshold is obtained:

$$c(a,B) = \begin{cases} 1 & h(a,B) < D_{\max} \\ 0 & h(a,B) \geq D_{\max} \end{cases} \tag{8}$$

$$C(A,B) = \sum_{a \in A} c(a,B) \tag{9}$$

The fitness is the relation between this number and the total number of the model points, card($A$):

$$F(A,B) = \frac{C(A,B)}{\mathrm{card}(A)} \tag{10}$$

This fitness function has two main advantages

- It can be immune to occlusion and noise.
- It allows stopping the generations if the percentage is high enough.

Thus, in Fig. 6, four generations are shown. In (a) the best individual for the first generation is shown, in (b) and (c) the best individual for two intermediate generations are shown and in (d) the best individual for the last generation is shown. Points in blue belong only to the model, in yellow only to the image and in white to both.

### 3.6. Good individuals selection to produce a new population

Parents are chosen to produce new children in every generation. The selection process is determined by the fitness of the solution. This process extends genes of good solutions through the population. This selection is usually done by the roulette method: the probability of being selected is proportional to the normalised fitness of all the individuals. This way, the best individuals have more offspring (although as it is a stochastic process sometimes it is the opposite). Sometimes, the roulette method has problems with triangular signs. As it was mentioned earlier, GAs are better for exploiting than exploring, because of that they can be trapped in local maximums since most of the individuals are concentrated in that zone. This is like loosing genetic richness in one specie and triangular signs tend to cause this. The reason is that several individuals can overlap two sides of their triangle on the model's one. It is convenient to delay convergence despite the risk of slowing it down to assure the global maximum. To do this the ranking method is used. Individuals are sorted by their fitness and they receive the selection probability following a straight line:

$$F_{\mathrm{new}}(A,B) = \frac{\max F_{\mathrm{new}} - \min F_{\mathrm{new}}}{N_{\mathrm{Pop}}} \mathrm{ranking}(F(A,B)) + \min F_{\mathrm{new}} \tag{11}$$

where $N_{\mathrm{Pop}}$ is the number of individuals of the population, and $\max F_{\mathrm{new}}$ and $\min F_{\mathrm{new}}$ the maximum and minimum fitness.

### 3.7. The new population

The crossover step looks for a good combination of genes, not individuals. For every couple of individuals several cross points are randomly chosen. Later, there are

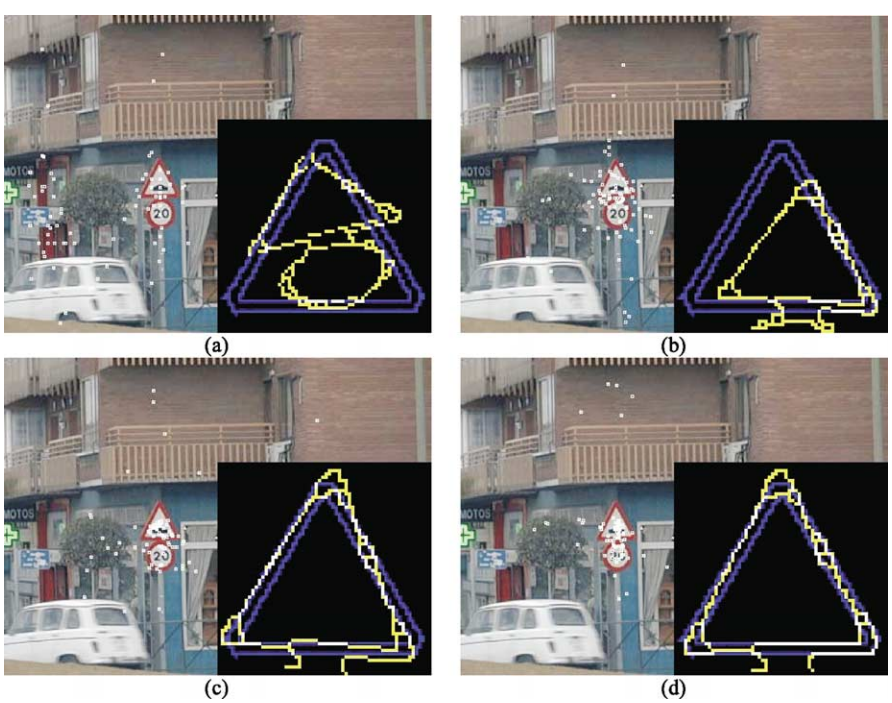


Fig. 6. Fitness function. (a) Best initial individual; (b, c) best intermediate results; (d) best final individual.

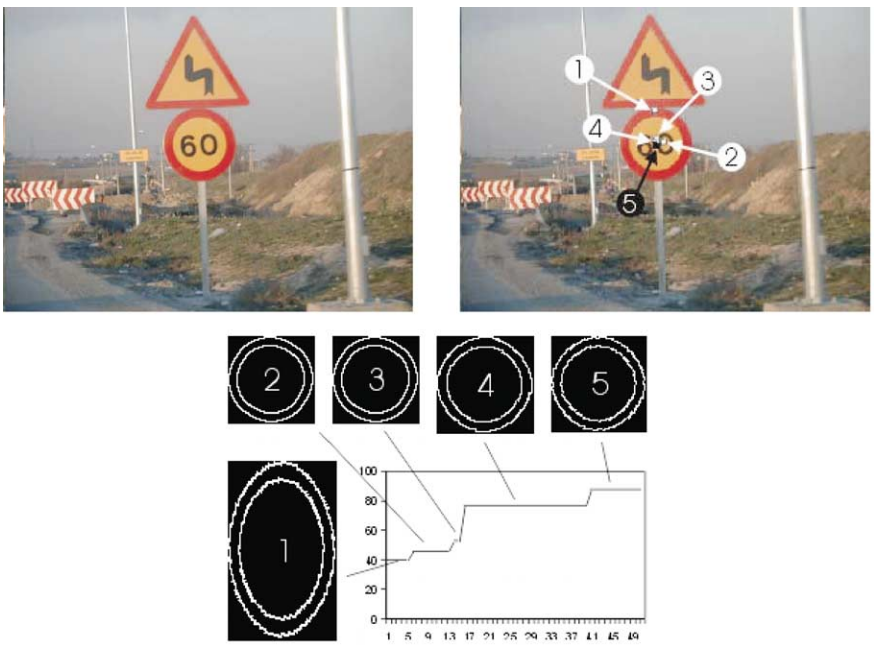


Fig. 7. Maximum evolution.

small changes in the children (mutation). If they produce bad individuals they will be eliminated in the next selection, whereas if they are considered good, their information would propagate through the population. Finally, the best individual of the previous generation is kept (elitism); this way the new result can guarantee that, at least it is as good as the old one. One process can be observed in Fig. 7, where the evolution of the best individual is shown.

For the following given parameters:

- Horizontal displacement: 9 bits, range 0/384 pixels.
- Vertical displacement: 9 bits, range 0/286 pixels.
- Horizontal scale: 7 bits, range 0.25/1.3 (30/157 pixels).
- Vertical scale: 7 bits, range 0.25/1.3 (30/157 pixels)
- Horizontal rotation: 5 bits, range—15°/15°
- Population: 101 individuals
- Crossover probability: 60%
- Mutation probability: 3%
- Maximum number of generations: 51
- Escape value: 80%.

Some traffic sign detection examples are shown in Fig. 8. They represent some real situations: other object with the same colour (a, f and g), perspective distortion (b) rotations (e), uncontrolled lighting as a sign against the light (d), and reflected light (c).

### 3.8. First sign state analysis

Once the sign has been detected, a first analysis can be performed to know about its state. It is the case of the signs presented in Fig. 9 where (a) was detected in Fig. 7, (b) in Fig. 8a, (c) and (d) in Fig. 1c, (e) in Fig. 1e and (f) in Fig. 3b. The Hue and Red components of the sign border are taken and the mean and variance are calculated. Note that the Hue component is shown in the interval (−128, 127) for a better view. Analysing the results, shown in Table 1, it can be observed how occlusion affects the Hue values, producing a high variance. The same can be said about the red component where the presence of shadows increases its variance. This way the algorithm find out if a sign is occluded or the visibility can be affected by shadows. It would be a help for the classification step.

## 4. Traffic sign classification

As it has been mentioned earlier, many researchers propose NNs because they have the ability to generalise from the training patterns and some can give the right answer despite the presence of occlusions. One requirement is that the NN has not to be retrained again when new signs are included, but it should have the ability to increase the information volume without losing what is had already learn.

The process is executed in two steps. First, the extraction of the information contained in the sign and second, its recognition by the NN.

### 4.1. Information extraction

This segmentation step has to be performed using RGB space to obtain the sign contents because one of the well-known problems of the HSI space is its instability to obtain

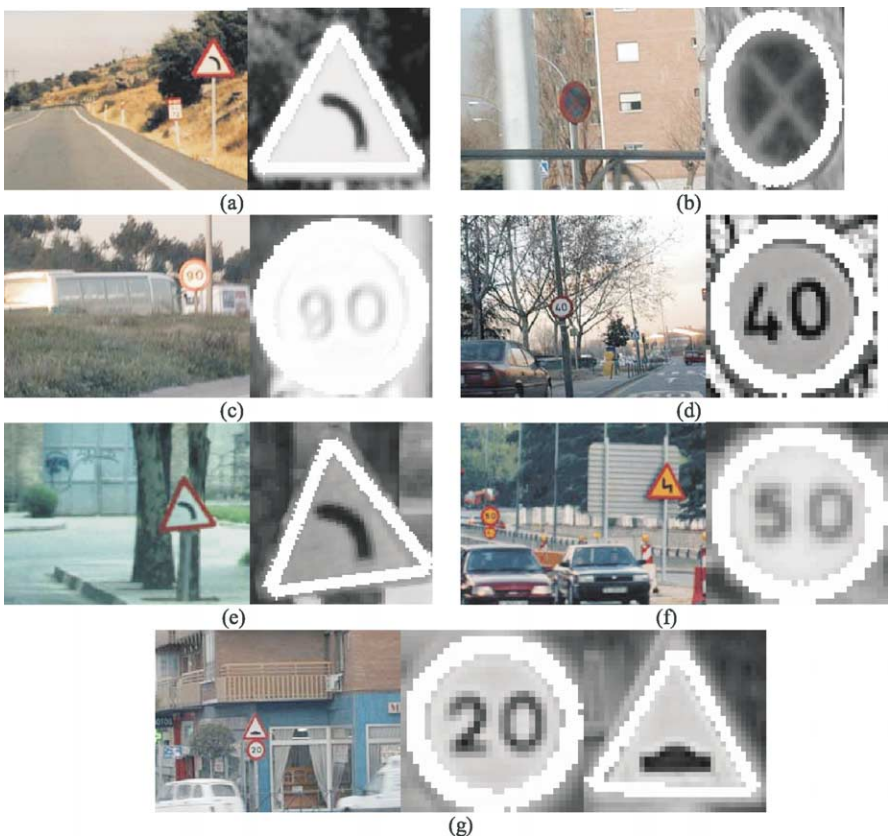


Fig. 8. Traffic sign detection examples.

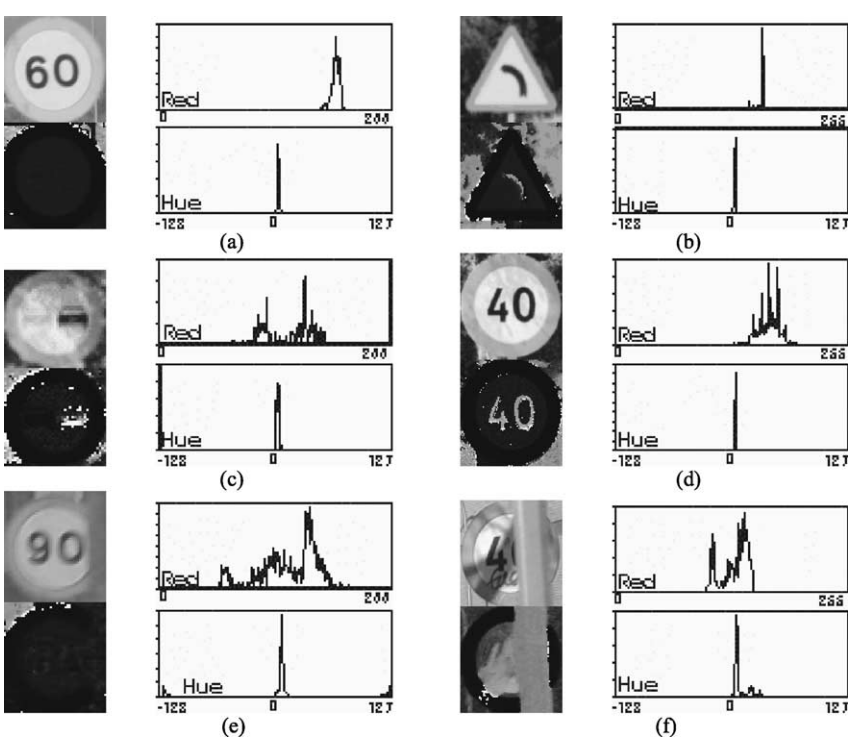


Fig. 9. First sign state analysis. For every sign the superior figure is the red component border histogram and the inferior the Hue.

Table 1
First sign state analysis results

| | | Fig. 9a | Fig. 9b | Fig. 9c | Fig. 9d | Fig. 9e | Fig. 9f |
|---|---|---|---|---|---|---|---|
| Hue | Mean | 2.53 | 2.44 | 2.26 | 2.76 | 7.57 | 13.84 |
| | Variance | 1.07 | 0.77 | 1.32 | 0.56 | 7.30 | 61.24 |
| Red | Mean | 193.69 | 158.51 | 140.03 | 168.16 | 131.05 | 142.73 |
| | Variance | 4.29 | 5.29 | 25.20 | 11.05 | 14.46 | 32.09 |

the Hue value of the grey colours. Although the I component could have been used, the colour information would have been lost because a dark red pixel belonging to the sign border would have the same value as a dark grey. The same occurs with other colours belonging to objects occluding the sign.

Because of that the following transform is done in the RGB space:

$$g(x,y)=\begin{cases}\dfrac{f_R(x,y)+f_G(x,y)+f_B(x,y)}{3} & \text{if } T_{\mathrm{GMIN}}<\dfrac{f_G(x,y)}{f_R(x,y)}<T_{\mathrm{GMAX}} \\ & T_{\mathrm{BMIN}}<\dfrac{f_B(x,y)}{f_R(x,y)}<T_{\mathrm{BMAX}} \\ 255 & \text{otherwise}\end{cases} \tag{12}$$

obtaining a black and white image with 256 grey levels. Its histogram is obtained and usually presents three maximums, the first one is the signed contained information, the second the sign background and the last one the white pixels. The threshold is the minimum value between the first two maximums.

### 4.2. Recognition

The NN used for the classification follows the Adaptive Resonance Theory paradigm [19] for binary inputs, ART1. It is capable of developing stable clustering of arbitrary sequences of input patterns by self-organisation. It has two layers with bottom–up and top–down pathways between them. The output layer is designed as a competitive network, only the node which receives the largest total input is activated. This node generates a top–down pattern that is compare with the input. If the difference is big enough the node is inhibited and the cycle begins again. When the similarity is high enough there is a resonant state of the network: a category prototype vector matches the current input vector close enough. If not, a signal is generated meaning a different pattern is presented and new neurones are created. Then, it is not necessary to retrain the network if new patterns have to be added. In addition, the network refines and enriches its knowledge with the use. Then ART1 NNs imitate the human ability of combining learning and memory [33]. During a standard use of the ART1 operation, the network groups the patterns in clusters where every output neurone is the centre of one cluster. In the case proposed in this article, the sign type is known in advance, and therefore, the output neurones are labelled with the class of every training sign pattern.

Table 2
Output information

| Figure | Class | Recognition value | Occlusions | Shadows | Rotation (deg) |
|---|---|---|---|---|---|
| Fig. 1b | Merging traffic from right | 0.87 | No | No | 2 |
| Fig. 1e | Maximum speed limit 90 | 0.79 | Partial | No | – |
| Fig. 1f | Left bend | 0.95 | No | No | 8 |
| Fig. 1g | Maximum speed limit 50 | 0.90 | No | No | – |
| Fig. 3b | Double bend | 0.72 | No | No | 3 |
| Fig. 3c | Merging traffic from right | 0.69 | No | No | 1 |
| Fig. 5-left | Maximum speed limit 40 | 0.60 | Yes | Yes | – |
| Fig. 6 | Maximum speed limit 20 | 0.76 | No | No | – |
| Fig. 6 | Uneven road | 0.89 | No | No | 2 |
| Fig. 7 | Maximum speed limit 60 | 0.90 | No | No | – |
| Fig. 7 | Double bend | 0.88 | No | No | 3 |
| Fig. 8d | Maximum speed limit 40 | 0.77 | No | No | – |
| Fig. 8f | Maximum speed limit 50 | 0.75 | No | No | – |
| Fig. 9d | Maximum speed limit 40 | 0.65 | No | Yes | – |

### 4.3. Second sign state analysis

Information about occlusions and shadows is obtained with the first sign analysis and it is important for the classification. If during the classification step the network discovers that there is no learnt pattern similar enough to the sign presented, the system verify if there is occlusion. In the case of occlusion, the maximum value is taken although it was low, if there is no occlusion the algorithm warns the user and ask for new sign type or a new node for a known class. On the other hand, if the network response is high but there is an occlusion, it means that the occlusion is small and affects only the sign border. The results are shown in Table 2 where the signs presented in the images are classified by the NN, the recognition value provided by Eq. 10 (0 no recognition, 1 full recognition) and some information about lack of visibility (occlusion) or bad condition (rotation) is provided. The algorithm takes 8.8 ms per generation for an AMD Duron at 1 GHz.

## 5. Conclusions

In the article, an algorithm for the detection and recognition of traffic sign have been proposed, although

the algorithm has been used for traffic signs it can be generalised to deal with other kinds of objects. The known difficulties that exist for object recognition in outdoor environments have been deal with. This way the system is immune to lighting changes, occlusions and object deformation being useful for Driver Support Systems. There are few previous works that deal with these last two problems. In addition to recognising the sign and as a novelty, the system provides information about its state. This way the algorithm, not only knows the detected sign, but also the confidence in its detection. It is believed that this can help the recognition algorithm in a continuous mode when the same sign will appear in several images. Due to this knowledge of the sign status, it is believed that the system is useful for other applications such as maintenance and inventories of traffic sign in highways and or cities.

## References


[1] Y. Aoyagi, T. Asakura, A study on traffic sign recognition in scene image using genetic algorithms and neural networks, 22nd International Conference on Industrial Electronics, Control, and Instrumentation, IEEE August (1996).

[2] G. Adorni, V. Dándrea, G. Destri, M. Mordoni, Shape searching in real world images: a CNN-based approach, Fourth Workshop on Cellular Neural Networks and their Applications, IEEE June (1996).

[3] G. Adorni, M. Mordonini, A. Poggi, Autonomous agents coordination through traffic signals and rules, Conference on Intelligent Transportation Systems, IEEE November (1997).

[4] P. Arnoul, M. Viala, J.P. Guerin, M. Mergy, Traffic signs localisation for highways inventory from a video camera on board a moving collection van, Intelligent Vehicles Symposium, IEEE September (1996).

[5] H. Austerirmeier, U. Büker, B. Merstching, S. Zimmermann, Analysis of traffic scenes using the hierarchical structure code, International Workshop on Structural and Syntactic Pattern Recognition August (1992).

[6] N. Bartneck, W. Ritter, Colour segmentation with polynomial classification, 11th International Conference on Pattern Recognition, IAPR August (1992).

[7] M. Bertozzi, A. Broggi, GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection, IEEE Transactions on Image Processing 7 (1) (1998) 62–81.

[8] B. Besserer, S. Estable, B. Ulmer, D. Reichqardt, Shape classification for traffic sign recognition, First International Workshop on Intelligent Autonomous Vehicles, IFAC April (1993).

[9] M. Betke, N. Makis, Fast object recognition in noisy images using simulated annealing, International Conference on Computer Vision, IEEE June (1995).

[10] U. Buker, B. Mertsching, Parallel evaluation of hierarchical image databases, Journal of Parallel and Distributed Computing 31 (2) (1995) 141–152.

[11] S. Buluswar, B. Draper, Color recognition in outdoor images, Sixth International Conference on Computer Vision, IEEE January (1998).

[12] E.D. Dickmanns, A. Zapp, A curvature-based scheme for improving road vehicle guidance by computer vision, Mobile Robots, SPIE October (1986).

[13] A. de la Escalera, L. Moreno, M.A. Salichs, J.Ma. Armingol, Road traffic sign detection and classification, IEEE Transactions on Industrial Electronics 44 (6) (1997) 848–859.

[14] S. Estable, J. Schick, F. Stein, R. Janssen, R. Ott, W. Ritter, Y.J. Zheng, A real-time traffic sign recognition system, Intelligent Vehicles Symposium August (1994).

[15] U. Franke, D. Gavrila, S. Gorzig, F. Lindner, F. Paetzold, C. Wohler, Autonomous driving approaches downtown, IEEE Intelligent Systems 13 (6) (1998) 40–48.

[16] G.Y. Jiang, T.Y. Choi, Robust detection of landmarks in color image based on fuzzy set theory, Fourth International Conference on Signal Processing, IEEE October (1998).

[17] D.M. Gavrila, V. Philomin, Real-time object detection using distance transforms, International Conference on Computer Vision, IEEE September (1999).

[18] D.E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, Reading, MA, 1989.

[19] G.A. Carpenter, S. Grossberg, A massively parallel architecture for a self-organizing neural pattern recognition machine, Computer Vision, Graphics, Image Processing 37 (1987) 54–115.

[20] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner, A image processing system for driver assistance, Image and Vision Computing 18 (5) (2000) 367–376.

[21] T. Hibi, Vision based extraction and recognition of road sign region from natural colour image, by using HSL and coordinates transformation, 29th International Symposium on Automotive Technology and Automation, Robotics, Motion and Machine Vision in the Automotive Industries, ISATA June (1996).

[22] D.P. Huttenlocher, G.A. Klanderman, W.J. Ruklidge, Comparing images using the Haussdorff distance, IEEE Transactions on PAMI 15 (1993) 850–863.

[23] R. Janssen, W. Ritter, F. Stein, S. Ott, Hybrid approach for traffic sign recognition, Intelligent Vehicles Symposium, IEEE July (1993).

[24] D.S. Kang, N.C. Griswold, N. Kehtarnavaz, An invariant traffic sign recognition system based on sequential color processing and geometrical transformation, Southwest Symposium on Image Analysis and Interpretation, IEEE April (1994).

[25] N. Kehtarnavaz, N.C. Griswold, D.S. Kang, Stop-sign recognition based on color/shape processing, Machine Vision and Applications 6 (4) (1993) 206–208.

[26] D.L. Kellmeyer, H.T. Zwahlen, Detection of highway warning signs in natural video images using color image processing and neural networks, International Conference on Neural Networks, IEEE July (1994).

[27] S.K. Kim, D.A. Forsyth, A new approach for road sign detection and recognition algorithm, 30th International Symposium on Automotive Technology and Automation, Robotics, Motion and Machine Vision in the Automotive Industries, ISATA June (1997).

[28] R.C. Luo, H. Potlapalli, D. Hislop, Natural scene segmentation using fractal based autocorrelation, International Conference on Industrial Electronics, Control, Instrumentation, and Automation, Power Electronics and Motion Control, IEEE November (1992).

[29] R.C. Luo, H. Potlapalli, D. Hislop, Neural network based landmark recognition for robot navigation, International Conference on Industrial Electronics, Control, Instrumentation, and Automation, Power Electronics and Motion Control, IEEE November (1992).

[30] R.C. Luo, H. Potlapalli, D. Hislop, Translation and scale invariant landmark recognition using receptive field neural networks, IEEE/RSJ International Conference on Intelligent Robots and Systems (1992) 527–533.

[31] R.C. Luo, H. Potlapalli, D. Hislop, Traffic sign recognition in outdoor environments using reconfigurable neural networks, International Joint Conference on Neural Networks, IEEE October (1993).

[32] R.C. Luo, H. Potlapalli, Landmark recognition using projection learning for mobile robot navigation, International Conference on Neural Networks, IEEE June (1994).

[33] A.S. Pandya, R.B. Macy, Pattern Recognition with Neural Networks in C++, CRC Press, Boca Raton, 1996.

[34] G. Piccioli, E. de Micheli, P. Parodia, M. Campani, Robust method for road sign detection and recognition, Image and Vision Computing 14 (3) (1996) 209–223.
[35] D. Pommerleau, T. Jochem, Rapidly adapting machine vision for automated vehicle steering, IEEE Expert 11 (2) (1996) 19–27.
[36] L. Priese, J. Klieber, R. Lakmann, V. Rehrmann, R. Schian, New results on traffic sign recognition, Intelligent Vehicles Symposium, IEEE October (1994).
[37] L. Priese, R. Lakmann, V. Rehrmann, Ideogram identification in a real-time traffic sign recognition system, Intelligent Vehicles Symposium, IEEE September (1995).
[38] M.A. Salichs, J.M[a] Armingol, L. Moreno, A. de la Escalera, Landmark perception planning for mobile robot localization, Integrated Computer-Aided Engineering 6 (4) (1999) 303–318.
[39] A.L. Yuille, D. Snow, M. Nitzberg, Signfinder: using color to detect, localize and identify informational signs, Sixth International Conference on Computer Vision, IEEE January (1998).
[40] M.M. Zadeh, T. Kasvand, C.Y. Suen, Localization and recognition of traffic signs for automated vehicle control systems, Intelligent Transportation Systems, SPIE (1998).