TRAFFIC SIGN DETECTION FOR DRIVER SUPPORT SYSTEMS

A. de la Escalera, J.Mª Armingol, M.A. Salichs
Systems Enginnering and Automation Division, Universidad Carlos III de Madrid, Leganés, Madrid, Spain.
Phone: +916249430, Fax: +916249431, E-mail: escalera@ing.uc3m.es

Abstract: This paper deals with object recognition in outdoor environments. In this type of environments, lighting conditions can not be controlled and predicted, objects can be partially occluded, and their position and orientation are not known *a priori*. The chosen type of objects is traffic and road signs, due to their usefulness for Driver Support Systems and Intelligent Autonomous Vehicles. A genetic algorithm is used for the detection, allowing an invariance localisation to changes in position, scale, rotation, weather conditions, partial occlusion, and the presence of other objects of the same colour.
Keywords: Object Recognition, Genetic Algorithm, Traffic Sign Detection, Driver Support Systems, Intelligent Autonomous Vehicles, Intelligent Transportation Systems, Intelligent Highway Systems.

1.INTRODUCTION.

Traffic sign detection is one of the subjects less studied in the field of Driver Support Systems. Research groups have focused on other aspects, more related with the development of an automatic pilot, as the detection of the road borders or the recognition of obstacles in the vehicle's path such as other vehicles or pedestrians. Nevertheless traffic sign detection have received an increasing interest in the last years. 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.

The extraction or detection of a traffic sign 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. (figure 1-a-b-c).
2. The presence of other objects. Except in the case of highways, the simple case, other objects often surround traffic signs (figure 1-d).
3. Models of all the possibilities of the sign's appearance are not possible to generate off-line, because there are so many degrees of freedom (figure 1-e-f).

The objective of the present paper is that 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. STATE OF THE ART.

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

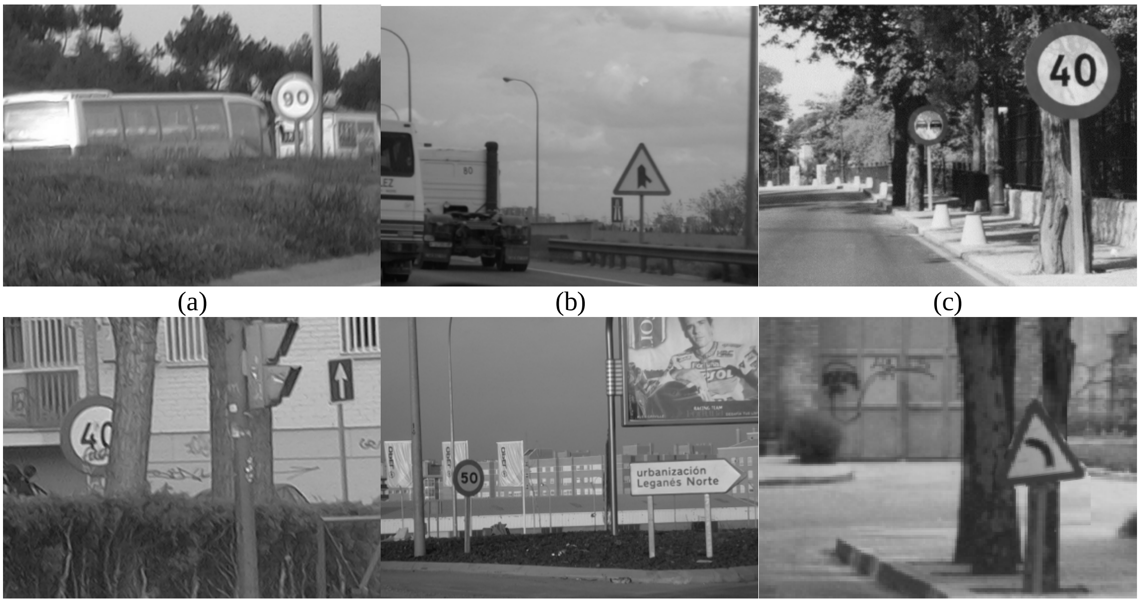
1. Segmentation through colour thresholding, region detection and shape analysis.
2. Segmentation through the border detection in a black and white image and their analysis.

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 [Buluswar and Draper 1998] [Escalera et al 1997] [Kim and Forsyth 1997] and [Zadeh et al 1998]. 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 [Arnoul et al 1996] [Hibi 1996] [Piccioli et al 1996], the Luv space has been also used [Kang et al 1994].

The second subgroup, those who developed a more exhaustive colour study, noted that the HSI space can not cover all the possible cases. 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 [Priese et al 1995] developed a database for the colour pixel classification. The use of textures has been proposed in [Kim and Forsyth 1997], while in [Jiang and Choi 1998] a fuzzy classification is used. Neural networks are proposed in [Kellmeyer and Zwahlen 1994].

In [Kang et al 1994] and [Hibi 1996] 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. The contour is analysed in [Kim and Forsyth 1997] using the eigenspace of the model bases for the classification. [Piccioli et al 1996] the region borders are analysed; after a first classification in triangular and circular signs. In [Escalera et al 1997] the borders of the colour-detected regions are analysed through the search of corners with certain angles and analysing the relations between them

Nevertheless, although these last classification methods are more complex and complete than those that use thresholding with fixed values, they do not



(a) (b) (c)

(d) (e) (f)

Figure 1. Traffic sign detection problems. (a) Reflections (b) Not controlled lightning (c) Shadows (d) Partial occlusions (e) Shape deformation (f) Sign rotation

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 [Buker and Mertsching 1995] that detect the borders in a pyramidal structure. They do not take occlusions into account. A Cellular Neural Network localises circular and triangular signs in [Adorni et al 1996]. The input layer of the net are the borders detected in a black and white image and the size of the sign is fixed. [Aoyagi and Asakura 1996] presents a genetic algorithm for the sign detection. Their work is explained with more detail because in this article genetic algorithms 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 almost all the research, 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, the y position and the radius. The population is formed by 32 individuals, the selection rate is 30%, 10% the mutation rate and there are 150 iterations. Finally there are multiple crosspoints. In [Gavrila and Philomin 1999] 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. For the sign detection the models are correlated with the image. In [Betke and Makis 1995] 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*.

3. TRAFFIC SIGN DETECTION.

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 for the colour classification since it gives different pieces of information in every component. Nevertheless, instead of choosing some fixed thresholds to form a binary image, the pass from the maximum value to the minimum one follows a ramp (figure 2). This way pixel classification errors through the use of rigid limits as thresholding are meant to be avoided. Thus, for the Hue component the sign's red colour has very low and high values (0 red, 85 green, 170 blue). From two values there is a ramp until the maximum value. As 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 (figure 2). Once both LUTs are

Figure 2. Colour Classification LUTs.

applied, the images are multiplied and normalised to the maximum value of 255. Results can be observed in figure 3. It is important to emphasise that a correct classification will not be necessary because occlusions will be taken into account. Additional problems are shown in figure 3: in (a) and (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.
After this a thresholding is done, and the resulting blobs are located. If the area of one blob is large enough, the blob's borders are found, and a sign is looked for.

3.2. Genetic algorithms.

Once the border of the objects are found, the algorithm has to find the signs presented in the image. It is a search problem. In [Piccioli et al 1996], they used the Hough transform for circles. In the case described here circular signs are ellipses, and therefore five parameter has to be found, been that transform too slow. On the other hand, Genetic algorithms (GAs) [Goldberg 1989] do a parallel search in several directions following an optimisation process that imitate natural selection and evolution.

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 traped in a local maximum or minimum. 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 one specie.
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 to 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:

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

The transform coefficients are:

Figure 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

(a)

(b)

Figure 4. Initialisation. (a) Original images (b) Initial population location

$$a_{00} = E_x \cos\theta \quad a_{01} = E_x sin\theta \quad a_{02} = T_x \qquad (2)$$
$$a_{10} = -E_y sin\theta \quad a_{11} = E_y \cos\theta \quad a_{12} = T_y$$

where:

- Tx is the horizontal displacement.
- Ty is the vertical displacement.
- Ex is the horizontal scale.
- Ey is the vertical scale.
- θ 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.

The initial population is generated randomly in a classic GA, but, this case can help some values to be nearer the final values. To do this, a thresholding of the colour analysis image is performed and the number and position of the blobs are obtained. A fixed number of individuals are assigned to every blob (figure 4). This way, the presence of enough individuals can be guaranteed despite the presence of bigger objects or occlusions. Thus, figure 4, 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.

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 [Huttenlocher et al 1993]. 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:

$$H(A,B) = max(h(A,B), h(B,A)) \qquad (3)$$

where:

$$h(A,B) = \max_{a\in A}(\min_{b\in B}\|a-b\|) \qquad (4)$$

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 distances are

(a) (b) (c)

Figure 5. Fitness function (a) Best initial individual (b) best intermediate results (d) Best final individual

taken into account.

$$h_k(AB) = K^{th}_{a\in A} \min_{b\in B} \|a-b\| \quad (5)$$

With this last distance the measurement can be immune to occlusion and noise.
In this case, the two sets are the borders of the blobs detected in the previous step and the affine transforms coded in the chromosomes of every individual. The distance transform is calculated and the number of points whose value is less than a certain threshold is obtained. The fitness is the relation between this number and the total number of the model points. Another advantage of this fitness function is that it allows stopping the generations if the percentage is high enough.
Thus, in figure 5, three generations are shown: a) the best individual for the first generation, (b) a partial generation and (c) the last. Points in grey belong only to the model, and in white to the image.

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. By this process genes of good solutions are extended 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.

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 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).

For the following given parameters:

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

the result of the algorithm is illustrated in the traffic sign detection examples shown in figure 6.

4. CONCLUSIONS.

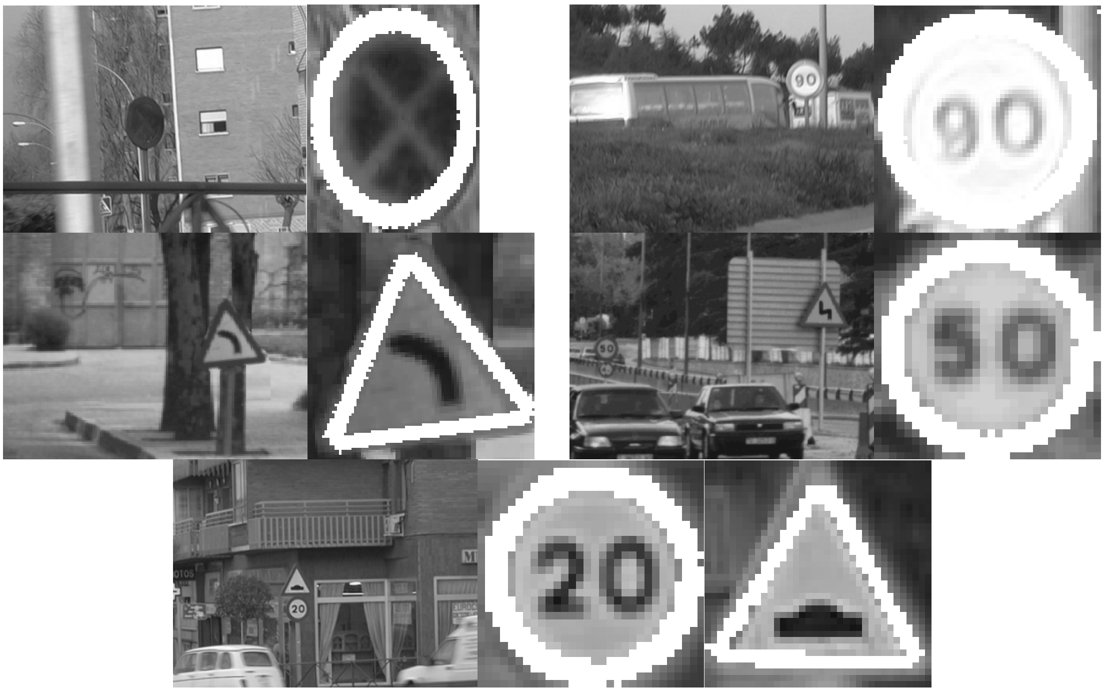


Figure 6. Traffic sign detection examples.

In the paper, an algorithm for the detection 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. It is believed that the system is useful for other applications such as maintenance and inventories of traffic sign in highways and or cities.

The algorithm has been programmed in MMX. This allowed the algorithm to work in real time. The average time for one generation is 17ms for a Petium III 359Mhz processor.

5. ACKNOWLEGMENT.

The authors gratefully acknowledge the support by the Spanish Government under CICYT Project TAP1999-0214.

6. REFERENCES.

[Aoyagi and Asakura 1996] Y. Aoyagi, T. Asakura. A study on traffic sign recognition in scene image using genetic algorithms and neural networks. Proceedings of the 22nd Int. Conf. on Industrial Electronics, Control, and Instrumentation, pages. 1838-43, Taipeis, August 1996.

[Adorni et al 1996] G. Adorni, V. Dándrea, G. Destri, M. Mordoni. Shape searching in real world images: a CNN-based approach. Proceedings of the 4th IEEE Workshop on Cellular Neural Networks and their Applications, pages 213-218, Seville, June, 1996.

[Arnoul et al 1996] 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. Proceedings of the IEEE Conference on Intelligent Vehicles, pages: 141-6, Tokyo, September 1996.

[Betke and Makis 1995] M. Betke, N. Makis. Fast object recognition in noisy images using simulated annealing. Proceedings of IEEE International Conference on Computer Vision, pages 523-530,. ; Cambridge, MA, June 1995.

[Buker and Mertsching 1995] U. Buker, B. Mertsching. Parallel evaluation of hierarchical image databases. Journal of Parallel and Distributed Computing, 31 (2): 141-52, December 1995.

[Buluswar and Draper 1998] S. Buluswar, B. Draper. Color recognition in outdoorr images. Proceedings of IEEE 6th International Conference on Computer Vision, pages 171-177, Bombay, January 1998.

[Escalera et al 1997] A. de la Escalera, L. Moreno, M.A. Salichs, J. Mª. Armingol. Road Traffic Sign Detection and Classification. IEEE Transaction on Industrial Electronics 44 (6): 848- 859, December 1997

[Gavrila and Philomin 1999] D.M. Gavrila, V. Philomin. Realtime object detection using distance transforms. Proceedings of the Seventh IEEE International Conference on Computer Vision, pages 87-93, Kerkyra, Septembre, 1999.

[Goldberg 1989] D.E. Goldberg. Genetic Algorithms in Search, Optimization and Machine Learnig. Addison-Wesley Publishing Company Inc, 1989

[Hibi 1996] T. Hibi. Vision based extraction and recognition of road sign region from natural colour image, by using HSL and coordinates transformation. Proceedings of the 29th International Symposium on Automotive Technology and Automation. Robotics, Motion and Machine Vision in the Automotive Industries, pages. 201-206, Florence, June 1996.

[Huttenlocher et al 1993] D.P. Huttenlocher, G.A. Klanderman, W.J. Ruklidge. Comparing images using the Haussdorff distance. IEEE PAMI 15 (9):850-863, September,1993.

[Jiang and Choi 1998] G.Y. Jiang, T.Y. Choi. Robust detection of landmarks in color image based on fuzzy set theory. Proceedings of the Fourth International Conference on Signal Processing, pages 968-71, Beijing, October 1998

[Kang et al 1994] D.S. Kang, N.C. Griswold, N. Kehtarnavaz. An invariant traffic sign recognition system based on sequential color processing and geometrical transformation. Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation, pages 88-93, Dallas, TX, April 1994.

[Kellmeyer and Zwahlen 1994] D.L. Kellmeyer, H.T. Zwahlen. Detection of highway warning signs in natural video images using color image processing and neural networks. Proceedings of the IEEE International Conference on Neural Networks, pages 4226-31, Orlando, July 1994.

[Kim and Forsyth 1997] S.K. Kim, D.A. Forsyth. A new approach for road sign detection and recognition algorithm. Proceedings of the 30th International Symposium on Automotive Technology and Automation. Robotics, Motion and Machine Vision in the Automotive Industries, pages. 171-8, Florence, June 1997.

[Piccioli et al 1996] G. Piccioli, E. de Micheli, P. Parodia, M. Campani. Robust method for road sign detection and recognition. Image and Vision Computing 14 (3): 209-23, April, 1996.

[Priese et al 1995] L. Priese, R. Lakmann, V. Rehrmann. Ideogram identification in a realtime traffic sign recognition system. Proceedings of Symposium on Intelligent Vehicles, pages 310-14, Detroit, MI, September 1995.

[Zadeh et al 1998] M.M. Zadeh, T. Kasvand, C.Y. Suen. Localization and recognition of traffic signs for automated vehicle control systems. Proceedings of the SPIE vol.3207; pages 272-82, 1998