Vision Based Intelligent System for Autonomous and Assisted Downtown Driving

Miguel Ángel Sotelo, Miguel Ángel García, and Ramón Flores

Department of Electronics, University of Alcalá,
Alcalá de Henares, Madrid, Spain,
{michael,garrido,flore}s@depeca.uah.es
http://www.depeca.uah.es

Abstract. Autonomous and assisted driving in city urban areas is a challeging topic that needs to be addressed during the following ten to twenty years. In the current work an attempt in this direction is carried out by using vision-based systems not only for autonomous vehicle driving, but in helping the driver recognize vehicles, and traffic signs. Some well consolidated results have been attained on a private test circuit using a commercial Citroen Berlingo as described in this paper.

1 The Global Concept

The work presented in this paper is in the frame of the ISAAC Project, whose general layout is depicted in figure 1. The ISAAC project undertakes the challenge of multivehicle cooperation in an urban-like environment by deploying three operational prototypes. Each vehicle is equipped with a GPS receiver and a colour vision system, so that either autonomous or manual navigation can be carried out. On the other hand, there is also a module onboard each vehicle which is in charge of providing assistance to the main driving system (either autonomous or manual) for safety enhancement during navigation.

The work described in this paper focuses on the development of an onboard vision based system for autonomous and assisted driving on urban city areas. This implies to solve the problem of intelligent unmanned mission execution using vision as the main sensor. According to this, the following remarkable challenges arise: lane tracking on non-structured roads (roads with no lane markers), sharp turn manoeuvres in intersections (very usual in urban areas), vehicle detection, and traffic sign detection and recognition. The system achieves global navigation by switching between Lane Tracking and Intersection Navigation, while detecting vehicles, and traffic signs in its local surrounding. The detailed description of the algorithms developed for lane tracking and intersection navigation is provided in [14]. Nonetheless, a brief summary of both algorithms is presented in this paper for completeness purposes.
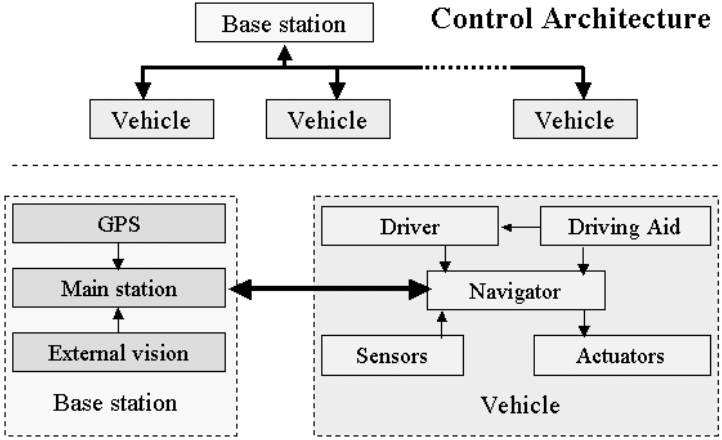
Fig. 1. General layout of the ISAAC Project.

2   Lane Tracking and Intersection Navigation

The mission of this vision based task is to provide correct lane tracking between two consecutive intersections. A polynomial representation is used to model the road edges in the image plane [1] [2]. According to these previous considerations the incoming image is on hardware re-scaled, building a low resolution image of what we call the Area of Interest (AOI), comprising the nearest 20 m ahead of the vehicle. The AOI is segmented basing on colour properties and shape restrictions. Image segmentation must be carried out by exploiting the cylindrical distribution of colour features in the HSI (Hue, Saturation, and Intensity) colour space, bearing in mind that the separation between road and no road colour characteristics is non linear. From the analytical point of view, the difference between two colour vectors in the HSI space can be established by computing the distances both in the chromatic plane, $d_{chromatic}$, and in the luminance axis, $d_{intensity}$, as described in equations 1 and 2.

$$d_{intensity} = |I_p - I_i| \tag{1}$$

$$d_{chromatic} = \sqrt{(S_p)^2 + (S_i)^2 - 2S_p S_i cos\theta} \tag{2}$$

with

$$\theta = \begin{cases} |H_p - H_i| & if \ |H_p - H_i| < 180° \\ 360° - |H_p - H_i| & if \ |H_p - H_i| > 180° \end{cases} \tag{3}$$

where H_p, H_i, S_p, S_i, I_p, and I_i represent the Hue, Saturation and Intensity values of the pattern (p) and given (i) pixels under comparison, respectively. As can be readily derived from the previous equations, $d_{chromatic}$ measures the distance

between two 2D colour vectors in the chromatic plane while $d_{intensity}$ provides the luminance difference between the pattern pixel and the pixel under consideration. According to this, a cylindrical surface of separation between the road and non-road classes is proposed in an attempt to decouple chromatic changes from luminance changes, as the latter are much greater in outdoor environments despite intensity is not a determinant characteristic in the colour segmentation process. In other words, any given pixel i will be classified as road if the chromatic distance ($d_{chromatic}$) to the colour pattern is bellow some threshold T_{chrom}, and the intensity distance ($d_{intensity}$) is lower than some T_{int}. This constraints the road pixels features in a cylinder around the pattern colour vector. On the other hand, to account for road shape restrictions, thresholds are affected by an exponentially decay factor yielding new threshold values that are updated depending on the distance between the current pixel and the previously estimated road model. Once the segmentation is accomplished, a time-spatial filter removes non-consistent objects in the low resolution image, both in space and time (sporadic noise). After that, the maximum horizontal clearance (absence of non-road sections) is determined for each line in the AOI. The measured points are fed into a Least Squares Filter with Exponential Decay [13] that computes the road edges in the image plane as well as the central trajectory of the road using a second order (parabolic) polynomial.

On the other hand, intersection navigation is completely vision based, and accounts for any angular value between the intersection branches. A simple monocular colour vision system is proposed to navigate on intersections (indeed the same system utilised for road tracking). This fact becomes a major issue, from the technological point of view, as cheap prototype navigation systems will be possible. Two basic manoeuvres can be executed at an intersection: on one hand the vehicle can change its moving direction by turning left or right; on the other hand the vehicle can go ahead and cross the intersection by keeping its current direction. The problem of crossing an intersection is basically the same of tracking the lane, and thus the same algorithmic solution is provided for this kind of manoeuvre. On the contrary, turning right or left at an intersection is quite a different problem that needs to be addressed in a different manner. At an intersection, the vehicle will start a turning manoeuvre until the new road to be tracked is perceived with a sufficiently reliable perspective. Vehicle localisation during the turn is reinforced using a Markov stochastic process, in what will be referred to as Markov Localisation Process [15] hereinafter. For this purpose, the angular trajectory described by the vehicle during the turn is modelled by a random variable denoted by ξ. A probability density function is calculated for all possible positions along the localisation space. Such a function is updated at each iteration time under the typical Markov assumptions, and in doing so it becomes a Markov stochastic process. For further details on this algorithm the reader is referred to [14].

3   Vehicle Detection

Vehicle detection is accomplished by using a monocular colour vision system. Using one single image leads to some limitations on the kind of obstacle that can be detected but provides a simple and fast method compared to optical flow based methods [8]. Other vehicles moving in the same or opposite lane can be reliably detected using the road shape and an estimation of the road width to predict the exact area of the image where the obstacles are expected to appear. Vehicles are then characterised by symmetry and edges features, within the estimated road, as far as usual vehicles have quite a distinguishable artificial shape and size that produces remarkable vertical edges in filtered images.

3.1   Searching Area

The execution time is abruptly reduced by limiting the obstacles detection to some predefined area where the obstacles are more likely to appear. This commonly called Region of Interest (ROI) is a rectangular area covering the central part of the image. In order to robustly detect and track vehicles along the road, two consecutive processing stages are necessary. In the first step vehicles are localised basing on the previously mentioned colour and shape properties, while in the second one the already detected vehicle is tracked using a real time estimator. A detailed description of both processes is given below.

3.2   Vehicles Detection

The identification of other vehicles is performed according to vertical edge and colour symmetry characteristics, together with temporal constraints for consistence purposes, under the assumption that vehicles generally have artificial rectangular and symmetrical shapes that make their vertical edges easily recognisable from the rest of the environment. This is quite a realistic situation that can be exploited in practice.

Vertical edge and symmetry discriminating analysis. A first discriminating analysis is carried out on the ROI basing on colour and vertical edge features. It permits to obtain candidate edges representing the limits of the vehicles currently circulating on the road. Thus, a symmetry map [3] of the ROI is computed so as to enhance those objects in the scene that present strong colour symmetry characteristics. After that, vertical edges are considered in pairs around those regions of the ROI where a sufficiently high symmetry measure (rejecting uniform areas) has been obtained, in order to account only for couples that represent possible vehicle contours, disregarding those combinations that lead to unrealistic vehicle shapes. Only those regions complying with a realistic edges structure will be validated as candidate vehicles.
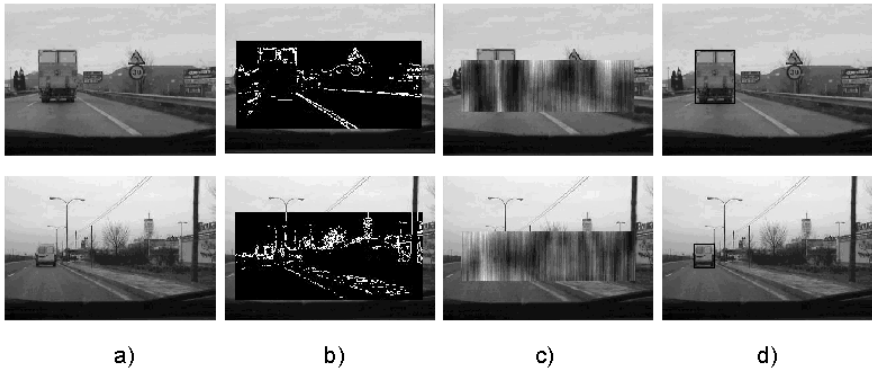
Fig. 2. Vehicle detection examples. a) Original image. b) ROI edge enhancement. c) Symmetry map. d) Position of detected vehicle.

Temporal consistence. Using spatial features as the only criterion for detecting obstacles yields to sporadic incorrect vehicles detection in real situations due to noise. Hence, a temporal validation filter becomes necessary to remove non-consistent objects from the scene [14]. This means that an object validated under the spatial features criterion described in the previous section must be detected several consecutive iterations in the image in order to be considered as a real vehicle. Otherwise it is discarded. A value $t = 0.5s$ has been used in practice to ensure that a vehicle appears in the image in a consistent time sequence. During the time-spatial validation stage a major problem is to identify the appearance of the same vehicle in two consecutive frames. For this purpose, its (x, y) position in correlative frames is used to keep track of temporal consistence of all candidate vehicles. Figure 2 depicts a couple of examples where the original and filtered images are illustrated together with the symmetry map of the ROI and the final position of the detected vehicle.

3.3   Vehicle Tracking

The position of the vehicle detected in the previous stage is tracked in two steps: position measurement and position estimation. The image contained inside the bounding box of the vehicle detected in the previous iteration is used as template in order to detect the updated position of the vehicle in the current image basing on a best fit correlation. After that, data association for position validation is carried out using the (x, y) location of the newly detected vehicle. Basically it must be determined whether some of the objects in the current frame corresponds to the vehicle under tracking. For this purpose, a limited searching area is specified around the vehicle position yielding to efficient and fast detection. Likewise, a minimum correlation value and template size are established so as to determine the end of the tracking process, whenever poor
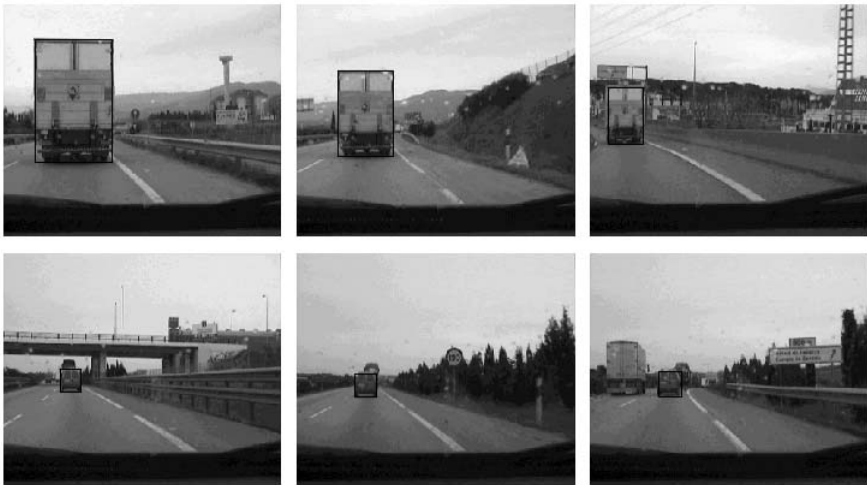
Fig. 3. Examples of vehicle tracking in real traffic situations.

correlations are attained or in case the vehicle gets too far away or out of the scene. The vehicle position measurements are then filtered using a recursive least squares estimator with exponential decay [13]. To avoid the problem of partial occlusions, the previously estimated vehicle position is kept during 5 consecutive iterations without obtaining any validated position, before considering that the vehicle track has been lost. If this happens, vehicle tracking is stopped and the vehicle detection stage is started again. To illustrate the vehicle tracking algorithm, figure 3 shows some real traffic situations in which the preceding vehicle position is tracked in a sequence of images.

3.4   Adaptive Navigation

Upon detecting the position of the preceding vehicle proper actions must be taken in order to ensure safe navigation, in an adaptive cruise control manner. Thus, if the time separation between the preceding vehicle and the ego-vehicle is below some predefined safety interval ($2s$ in this work) the ego-vehicle velocity is accordingly modified. Further navigation actions can also be undertaken in case the preceding vehicle performs a sudden braking manoeuvre in the limits of the safety interval. This could lead to an inminent crash unless the ego-vehicle rapidly detects the braking situation and reacts accordingly producing also a fast braking. In order to achieve this goal, the activation of the preceding vehicle braking lights must be detected as it will clearly indicate a braking down manoeuvre. The position of braking lights changes depending on each vehicle model and company, and thus, a detailed search should be carried out so as to accurately locate them inside the vehicle bounding box. Nonetheless, there is

Fig. 4. Detection of a sudden braking of the preceding vehicle by continuously monitoring the activation of the braking lights.

some a priori information that can obviously be used to enhance and ease the searching process, as long as braking indicators are usually two red lights simmetrically located near the left and right sides of the rear part of the vehicle. This constraint will help locate the braking lights for different types of vehicles. The next step implies to detect a sudden activation of the braking lights, as previously mentioned, by continuously monitoring their luminance component. In case this occurred, an alarm would be raised to the vehicle navigator so as to produce an emergency braking down manoeuvre. Figure 4 depicts an example where a sudden braking of the preceding vehicle is detected following the previously described scheme.

4   Traffic Sign Detection

The vision based traffic sign detection module developed in this work manages 289x352 colour images in RGB (Red, Green, Blue) format. As in the previous sections, a ROI is defined intended to reduce the processing time by focusing on the area of the image where traffic sings are more likely to appear. For this purpose, the 10% upper part and the 30% lower part of the image are disregarded for visual computations. On the other hand, appropriate choice of the colour features to use in the process is of crucial importance in order to attain proper and fast detection. Accordingly, only the Red component is considered as it provides a high capacity for colour discriminating in the visual analysis of traffic signs, and no further preprocessing is needed after digitalisation. In an attempt to carry out a preattentive strategy, a coarse analysis of vertical edges is performed in a first stage basing on differential characteristics computed on the Red component of the image using a gradient filter, as depicted in figure 5, where several images containing traffic signs are illustrated together with their associated gradient images after applying a vertical edge operator.

The accumulated projection of the vertical edges computed upon gradient images provides quite a valuable and useful information for discriminating the positions of the image where traffic signs are located, as can be derived from observation of figure 6. Traffic signs can be clearly characterized by the presence
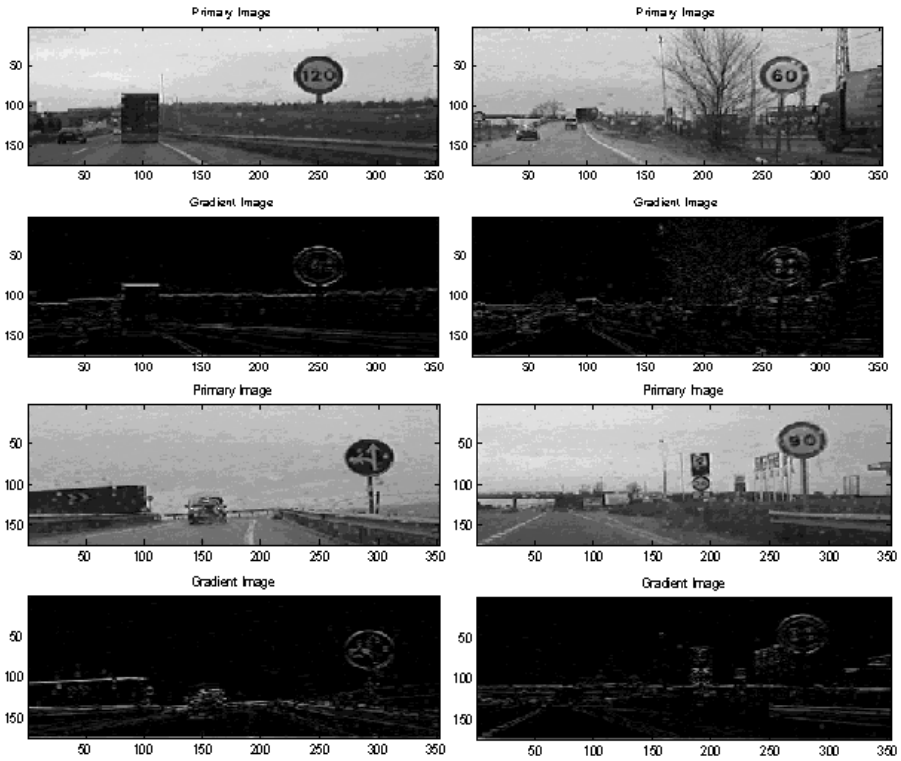
Fig. 5. Original and filtered images using a vertical gradient operator.

of two consecutive peaks of high value very close to each other in the projection profile. This fact constitutes the basis for the traffic sign detection algorithm proposed in this work. As a first step, an adaptive thresholding is performed aiming at removing the common offset component in the projection profile. For this purpose, a threshold u is computed as expressed in equation 4.

$$u = \mu + \mu^+ \tag{4}$$

where μ stands for the average value of the projection profile, while μ^+ represents the average of all points in the projection profile whose value is greater than μ. The resulting threshold is depicted in figure 6 for the four edge images shown in figure 5. As can be appreciated, the peaks corresponding to traffic signs positions remain after thresholding. Finally, the coarse detection phase ends by removing narrow peaks from the projection profile. This yields a set of candidate image regions that highly reduces and constraints the portions of the image where traffic signs are likely to appear, as depicted in figure 7. A further analysis should be accomplished so as to validate the existence of circular or triangular shapes within the candidate regions, prior to the traffic sign recognition stage.
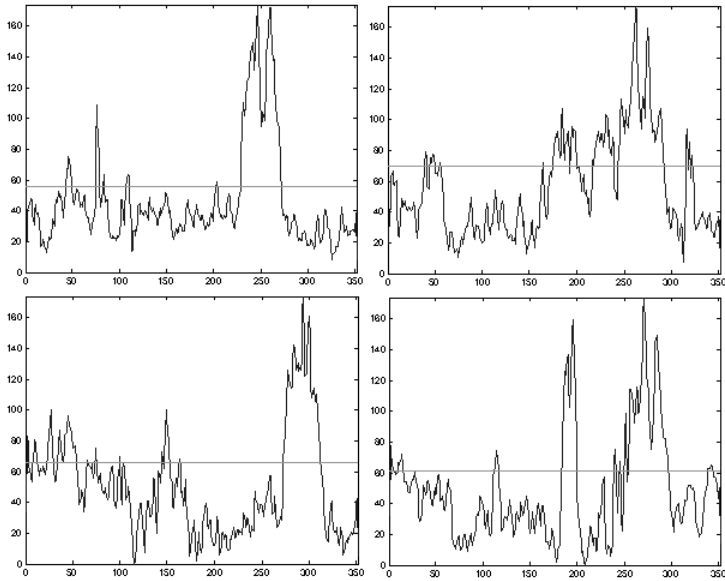
Fig. 6. Adaptive threshold u computed over the projection profile.

The algorithm described in this section has been implemented on a PentiumIV running at 2.5GHz, yielding an average execution time of $106ms$.

5   Discussion

The complete vision and DGPS based system for autonomous and/or assisted navigation described in this paper has been implemented on the so-called Babieca prototype vehicle, an electric commercial Citroen Berlingo, as a first step towards its long-term deployment on urban scenarios. Practical experiments were conducted on a private circuit located at the *Industrial Automation Institute* in Arganda del Rey (Madrid) to test the autonomous navigation system. The circuit is composed of several stop stations, streets, intersections, and roundabout points, trying to emulate an urban quarter. Babieca ran over hundreds of kilometres in lots of successful autonomous missions carried out along the test circuit. Besides, a live demonstration exhibiting the system capacities on autonomous navigation was carried out during the IEEE Conference on Intelligent Vehicles 2002, in a private circuit located at Satory (Versailles), France. A complete set of video files demonstrating the operational performance of the system in real tests can be retrieved from ftp://www.depeca.uah.es/pub/vision. Practical trials have also been carried out on real urban and highway scenarios so as to test the validity of the vision based driving assistance systems developed in this work. Either the vehicle detection or the traffic sign module have proven to be useful for the
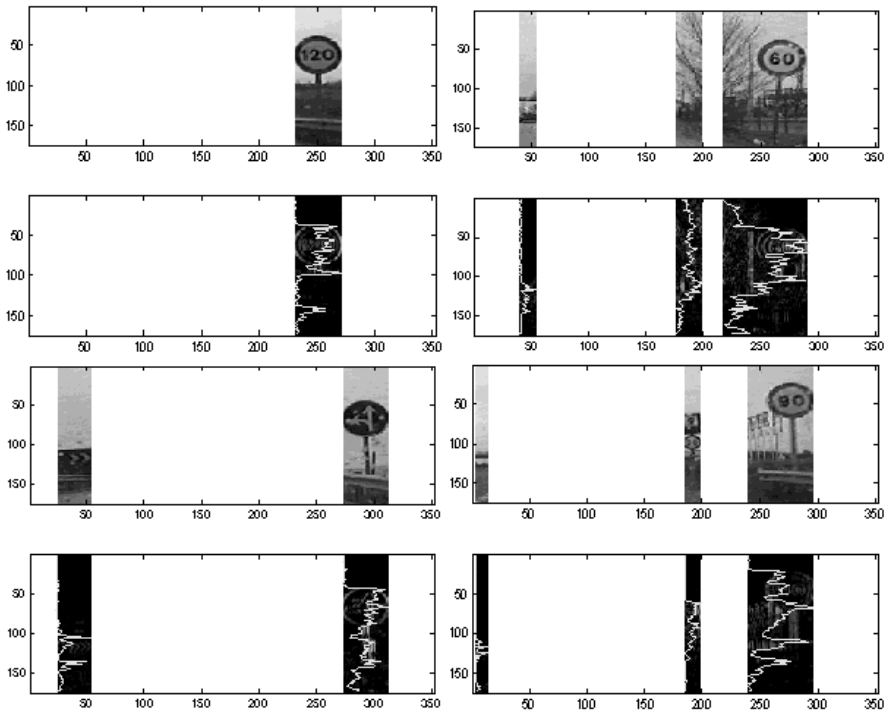
Fig. 7. Candidate image regions where traffic signs are likely to be located.

driver in keeping a safety distance and fulfilling the traffic rules, respectively, in real conditions. On the other hand, the integration of the vehicle detection task into the global navigation system has been succesfully achieved, as demonstrated in the previously mentioned video files. The fact that no extremely high precision is needed in the DGPS signal, together with the use of a single colour camera, results in a low cost final system suitable for midterm commercial development. Our current work focuses on the detection of pedestrians using a combination of vision and laser sensors, as it would be a major issue in increasing traffic safety.

References

1. M. Bertozzi and 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.
2. M. Bertozzi, A. Broggi and A. Fascioli.: Vision-based intelligent vehicles: State of the art and perspectives. Robotics and Automation Systems **32** 2000. 1–16.
3. A. Broggi, S. Nichele, M. Bertozzi, and A. Fascioli.: Stereo Vision-based Vehicle Detection. In Proceedings of the IEEE Intelligent Vehicles Symposium. Detroit, USA. October (2000)
4. P. Charbonnier, P. Nicolle, Y. Guillard and J. Charrier.: Road boundaries detection using color saturation. Proceedings of the Ninth European Signal Processing Conference '98. September (1998)
5. T. De Pedro, R. Garcia, C. Gonzalez, J. E. Naranjo, J. Reviejo and M. A. Sotelo.: Vehicle Automatic Driving System Based on GNSS. Proceedings of the International Conference on Intelligent Vehicles. Seville, (2001)
6. E. D. Dickmanns, R. Behringer, D. Dickmanns, T. Hildebrant, M. Mauer, F. Thomanek, and J. Shielhlen.: The seeing passenger car 'VaMoRs-P'. In Proc. of Int. Symp. on Intelligent Vehicles, Paris. (1994)
7. U. Franke, D. Gavrila, S. Gorzig, F. Lindner, F. Paitzold and C. Wohler.: Autonomous driving goes downtown. Proceedings of the IEEE Intelligent Vehicles Symposium. Stuttgart, Germany. (1998)
8. A. Giachetti, M. Campani and V. Torre.: The use of Optical Flow for Road Navigation. IEEE Transactions on Robotics and Automation. **14** No 1, (1998)
9. R. C. Gonzales and R. E. Wood.: Digital Image Processing. Addison-Wesley, Reading, MA, (1992)
10. N. Ikonomakis, K. N. Plataniotis and Venetsanopoulos.: Color Image Segmentation for Multimedia Applications. Journal of Intelligent and Robotic Systems. (2000)
11. W. Li, X. Jiang and Y. Wang.: Road recognition for vision navigation of an autonomous vehicle by fuzzy reasoning. Fuzzy Sets and Systems. **93** Elsevier. (1998)
12. R. Gregor, M. Lutzeler, M. Pellkofer, K. H. Siedersberger, and E. D. Dickmanns.: EMS-Vision: A Perceptual System for Autonomous System. IEEE Transactions on Intelligent Transportation Sustems. **3**, No. 1, (2002)
13. H. Schneiderman and M. Nashman.: A Discriminating Feature Tracker for Vision-Based Autonomous Driving. IEEE Transactions on Robotics and Automation. **10** NO 6. (1994)
14. M. A. Sotelo.: Sistema de Navegación Global Aplicado al Guiado de un Vehículo Autónomo Terrestre en Entornos Exteriores Parcialmente Conocidos. Phd Thesis Dissertation. University of Alcalá. (2001).
15. D. F. Wolfram Burgard and S. Thrun.: Active Markov Localization for Mobile Robots. Elsevier Preprint. (1998)