ELSEVIER

Robotics and Autonomous Systems 32 (2000) 1–16

Robotics and Autonomous Systems

www.elsevier.com/locate/robot

Overview Paper

Vision-based intelligent vehicles: State of the art and perspectives☆

Massimo Bertozzi[a,*], Alberto Broggi[b], Alessandra Fascioli[a]

[a] *Dipartimento di Ingegneria dell'Informazione, Università di Parma, I-43100 Parma, Italy*

[b] *Dipartimento di Informatica e Sistemistica, Università di Pavia, I-27100 Pavia, Italy*

Received 31 March 1999; received in revised form 1 October 1999

Communicated by F.C.A. Groen

Abstract

Recently, a large emphasis has been devoted to Automatic Vehicle Guidance since the automation of driving tasks carries a large number of benefits, such as the optimization of the use of transport infrastructures, the improvement of mobility, the minimization of risks, travel time, and energy consumption.

This paper surveys the most common approaches to the challenging task of Autonomous Road Following reviewing the most promising experimental solutions and prototypes developed worldwide using AI techniques to perceive the environmental situation by means of artificial vision.

The most interesting results and trends in this field as well as the perspectives on the evolution of intelligent vehicles in the next decades are also sketched out.

Keywords: Intelligent transportation systems (ITS); Automatic vehicle guidance; Intelligent vehicles; Machine vision

1. Introduction

In the last decades in the field of transportation systems a large emphasis has been given to issues such as improving safety conditions, optimizing the exploitation of transport networks, reduce energy consumption and preserving the environment from pollution. The endeavors in solving these problems have triggered the interest towards a new field of research and application, automatic vehicle driving, in which new techniques are investigated for the entire or partial automation of driving tasks. These tasks include following the road and keeping within the correct lane, maintaining a safe distance among vehicles, regulating the vehicle's speed according to traffic conditions and road characteristics, moving across lanes in order to overtake vehicles and avoid obstacles, finding the shortest route to a destination, and moving and parking within urban environments.

The interest in *intelligent transportation systems* (ITS) technologies was born about 20 years ago, when the problem of people and goods mobility began to arise, fostering the search for new alternative solutions. Automatic vehicle driving, intelligent route planning, and other extremely high-level functionalities were selected as main goals. Governmental

☆This research has been partially supported by the Italian National Research Council (CNR) under the frame of the Progetto Finalizzato Trasporti 2 and Progetto Madess 2.

*Corresponding author.

E-mail addresses: bertozzi@ce.unipr.it (M. Bertozzi), broggi@ce.unipr.it (A. Broggi), fascal@ce.unipr.it (A. Fascioli)

0921-8890/00/$ – see front matter
PII: S0921-8890(99)00125-6

institutions activated this initial explorative phase by means of various projects worldwide, involving a large number of research units who worked in a cooperative way producing several prototypes and possible solutions, all based on rather different approaches.

In Europe the PROMETHEUS project (PROgraM for a European Traffic with Highest Efficiency and Unprecedented Safety) started this explorative stage in 1986. The project involved more than 13 vehicle manufacturers and several research units from governments and universities of 19 European countries. Within this framework, a number of different approaches regarding ITS were conceived, implemented, and demonstrated.

In the United States a great deal of initiatives were launched to deal with the mobility problem, involving many universities, research centers, and automobile companies. After this pilot phase, in 1995 the US government established the National Automated Highway System Consortium (NAHSC) [2].

Also in Japan, where the mobility problem is much more intense and evident, some vehicle prototypes were developed within the framework of different projects. Similarly to what happened in the US, in 1996 the Advanced Cruise-Assist Highway System Research Association (AHSRA) was established amongst a large number of automobile industries and research centers [31], which developed different approaches to the problem of Automatic Vehicle Guidance.

The main results of this first stage were a deep analysis of the problem and the development of a feasibility study to understand the requirements and possible effects of the application of ITS technology.

The field of ITS is now entering its second phase characterized by a maturity in its approaches and by new technological possibilities which allow the development of the first experimental products. A number of prototypes of intelligent vehicles have been designed, implemented, and tested on the road. The design of these prototypes was preceded by the analysis of solutions deriving from similar and close fields of research, and exploded with a great flourishing of new ideas, innovative approaches, and novel ad hoc solutions. Robotics, artificial intelligence, computer science, computer architectures, telecommunications, control and automation, signal processing are just some of the principal research areas from which the main ideas and solutions were first derived. Initially, underlying technological devices — such as head-up displays, infrared cameras, radars, sonars — derived from expensive military applications, but, thanks to the increased interest in these applications and to the progress of industrial production, today's technology offers sensors, processing systems, and output devices at very competitive prices. In order to test a wide spectrum of diverse approaches, these prototypes of automatic vehicles are equipped with a large number of different sensors and computing engines.

Section 2 of this paper describes the motivations which underlie the development of vision-based intelligent vehicles, and illustrates their requirements and peculiarities. Section 3 surveys the most common approaches to autonomous Road Following developed worldwide, while Section 4 ends the paper outlining our perspectives in the evolution of intelligent vehicles.

2. Vision-based intelligent vehicles

2.1. Improving vehicles or infrastructures?

Automatic driving functionalities can be achieved acting on infrastructures and vehicles. Depending on the specific application, either choice possesses advantages and drawbacks. Enhancing road infrastructure may yield benefits to those kind of transportation which are based on repetitive and prescheduled routes, such as public transportation and industrial robotics. On the other hand, it requires a complex and extensive organization and maintenance which can become cumbersome and extremely expensive in case of extended road networks for private vehicles use. An ad hoc structuring of the environment can only be considered for a reduced subset of the road network, e.g., for building a fully automated highway on which only automatic vehicles, public or private, can drive.

For this reason, the systems that are expected to be achieved on a short-term basis can only be vehicle-autonomous or at the most, those systems can exploit already existing traffic infrastructures.

In this review only in-vehicle applications are considered, while road infrastructure, inter-vehicle communication, satellite communication, and route planning issues are not covered.

Any on-board system for ITS applications needs to meet some important requirements:

- The final system that will be installed on a commercial vehicle must be robust enough to adapt to different conditions and changes of environment, road, traffic, illumination, and weather. Moreover, the hardware system needs to be resistant to mechanical and thermal stress.
- On-board systems for ITS applications are safety critical systems for which a high degree of reliability is required. Consequently, the project has to be thorough and rigorous during all its phases, from the requirements specification to the design and implementation. An extensive phase of testing and validation is therefore of paramount importance.
- For marketing reasons the design of an ITS system has to be driven by strict cost criteria (it should cost no more than 10% of the vehicle price) thus requiring a specific engineering phase. Operative costs (such as power consumption) need to be kept low as well, since the vehicle's performance should not be affected by the use of ITS apparata.
- The system's hardware and sensors have to be kept compact in size and should not disturb car styling.
- Since ITS systems must be triggered off and controlled by a human driver, they need a user-friendly man–machine interface.

2.2. Which kind of sensors should be employed?

Among the sensors widely used in indoor robotics, tactile sensors and acoustic sensors are of no use in automotive applications because vehicles' speed the former ones, and their reduced detection range the latter ones.

Laser-based sensors and millimeter-wave radars detect the distance of objects by measuring the travel time of a signal emitted by the sensors themselves and reflected by the object, and are therefore classified as *active sensors*. Their common principal drawbacks are the low spatial resolution and slow scanning speed. However, millimeter-wave radars are more robust to rain and fog than laser-based radars, even if more expensive.

Vision-based sensors are defined as *passive sensors* and have an intrinsic advantage over laser and radar sensors: the possibility of acquiring data in a non-invasive way, thus not altering the environment. Image scanning is performed fast enough for ITS applications. Moreover, they can be used for some specific applications for which visual information plays a basic role (such as lane markings localization, traffic signs recognition, obstacle identification) without requiring any modifications to road infrastructures. Unfortunately vision sensors are less robust than millimeter-wave radars in foggy, night, or direct sun-shine conditions.

Active sensors possess some specific peculiarities which, in this specific application, result in advantages over vision: they can measure some quantities, such as movement, in a more direct way than vision and require less performing computing resources as they acquire a considerably lower amount of data.

Notwithstanding, besides the problem of environment pollution, the wide variation in reflection ratios caused by different reasons — such as obstacles' shape or material — and the need for the maximum signal level to comply with some safety rules, the main problem with the use of active sensors is represented by interference among sensors of the same type, which could be critical for a large number of vehicles moving simultaneously in the same environment as, e.g., in the case of autonomous vehicles traveling on intelligent highways.

Hence, foreseeing a massive and widespread use of autonomous sensing agents, the use of passive sensors, such as cameras, obtains key advantages over the use of active ones.

Obviously machine vision does not extend sensing capabilities besides human possibilities (e.g., in foggy conditions or during the night with no specific illumination), but can, however, help the driver in case of failure, e.g., in the lack of concentration or drowsy conditions.

2.3. Machine vision for intelligent vehicles

Some important issues must be carefully considered in the design of a vision system for automotive applications.

In the first place, ITS systems require faster processing than other applications, since the vehicle speed is bounded by the processing rate. The main problem that has to be faced when real-time imaging is concerned and which is intrinsic to the processing of images is the large amount of data, and therefore computation involved. As a result, specific computer architectures and processing techniques must be devised in order to achieve real-time performance. Nevertheless, since the success of ITS apparata is tightly related to their cost, the computing engines cannot be based on expensive processors. Therefore, either off-the-shelf components or ad hoc dedicated low-cost solutions must be considered.

Secondly, in the automotive field no assumptions can be made on key parameters, e.g., the scene illumination or contrast, which are directly measured by the vision sensor. Hence, the subsequent processing must be robust enough to adapt to different environmental conditions (such as sun, rain, fog) and to their dynamic changes (such as transitions between sun and shadow, or the entrance or exit from a tunnel).

Furthermore, other key issues, such as the robustness to vehicle's movements and drifts in the camera's calibration, must be handled as well.

However, recent advances in both computer and sensor technologies promote the use of machine vision also in the intelligent vehicles field. The developments in computational hardware, such as a higher degree of integration and a reduction of the power supply voltage, permit to produce at an affordable price machines that can deliver a high computing power with fast networking facilities. Current technology allows the use of SIMD-like processing paradigms even in general-purpose processors such as the new generation of processors that include multimedia extensions.

In addition to this, current cameras include new important features that permit the solution of some basic problems directly at sensor level. For example, image stabilization can be performed during acquisition, while the extension of camera dynamics allows to avoid the processing required to adapt the acquisition parameters to specific light conditions. The resolution of the sensors has been drastically enhanced, and in order to decrease the acquisition and transfer time, new technological solutions can be found in CMOS sensors, such as the possibility of dealing with pixels independently as in traditional memories. Another key advantage of CMOS-based sensors is that their integration on the processing chip seems to be straightforward.

Many different parameters must be evaluated for the design and choice of an image acquisition device. First of all, some parameters tightly coupled with the algorithms regard the choice of monocular vs binocular (stereo) vision and the sensors' angle of view (some systems adopt a multi-camera approach, by using more than one camera with different viewing angles, e.g., fish eye or zoom). The resolution and the depth (number of bits/pixel) of the images have to be selected as well (this also includes the selection of color vs monochrome images).

Other parameters — intrinsic to the sensor — must be considered. Although the frame rate is generally fixed for CCD-based devices (25 or 30 Hz), the dynamics of the sensor is of basic importance: conventional cameras allow an intensity contrast of 500:1 within the same image frame, while most ITS applications require a 10,000:1 dynamic range for each frame and 100,000:1 for a short image sequence. Different approaches have been studied to meet this requirement, ranging from the use of CMOS-based cameras with a logarithmically compressed dynamic [27,29] to the interpolation and superimposition regarding values of two subsequent images taken from the same camera [23].

In conclusion, although extremely complex and highly demanding, thanks to the great deal of information it can deliver (it has been estimated that humans perceive visually about 90% of the environment information required for driving), computer vision is a powerful means for sensing the environment and has been widely employed to deal with a large number of tasks in the automotive field.

3. Automatic Road Following: An overview of the approaches

Among the complex and challenging tasks that received most attention in Automatic Vehicle Guidance is *Road Following*. It is based on *lane detection* (which includes the localization of the road, the determination of the relative position between vehicle and road, and the analysis of the vehicle's heading direction), and *obstacle detection* (which is mainly based on localizing possible obstacles on the vehicle's path).

In this section, a survey on the most common approaches to Road Following is presented, focusing in particular on vision-based systems.

3.1. Lane detection

In most prototypes of autonomous vehicles developed worldwide, lane following is divided into the following two steps: initially, the relative position of the vehicle with respect to the lane is computed and then actuators are driven to keep the vehicle in the correct position. Some examples of the strategies which may be adopted in this vision-based lateral control problem are discussed in [30].

Conversely, some early systems were not based on the preliminary detection of the road's position, but obtained the commands to be issued to the actuators (steering wheel angles) directly from visual patterns detected in the incoming images. As an example, the ALVINN (Autonomous Land Vehicle In a Neural Net) system is based on a neural net approach: it is able to follow the road after a training phase with a large set of images [16].

Nevertheless, since the knowledge of the lane position can be conveniently exploited by other driving assistance functions, the localization of the lane is generally performed.

A few systems have been designed to handle completely unstructured roads, e.g., the SCARF (Supervised Classification Applied to Road Following [9]) and PVR III (POSTECH Road Vehicle [17]) systems are based on the use of color cameras and exploit the assumption of a homogeneously colored road to extract the road region from the images.

More generally, however, lane detection has been reduced to the localization of specific features such as markings painted on the road surface. This restriction eases the detection of the road, nevertheless two basic problems must be faced.

- The presence of *shadows* (projected by trees, buildings, bridges, or other vehicles) produces artifacts onto the road surface, and thus alters the road texture.

  Most research groups face this problem using highly sophisticated image filtering algorithms. For the most part, gray-level images are used, but in some cases color images are used: this is the case of the MOSFET (Michigan Off-road Sensor Fusing Experimental Testbed) autonomous vehicle which uses a color segmentation algorithm that maximizes the contrast between lane markings and road [22].
- *Other vehicles* on the path partly occlude the visibility of the road and therefore also of road markings.

  To cope with this problem, some systems have been designed to investigate only a small portion of the road ahead of the vehicle where the absence of other vehicles can be assumed. As an example, the LAKE and SAVE autonomous vehicles rely on the processing of the image portion corresponding to the nearest 12 m of road ahead of the vehicle, and it has been demonstrated that this approach is able to safely maneuver the vehicle on highways and even on belt ways or ramps with a bending radius down to 50 m [8].

  The RALPH (Rapidly Adapting Lateral Position Handler) system, instead, reduces the portion of the image to be processed according to the result of a radar-based obstacle detection module [24].

  In other systems the area in which lane markings are to be looked for is determined first. The research group of the Laboratoire Régional des Ponts-et-Chaussées de Strasbourg exploits the assumption that there should always be a chromatic contrast between road and off-road (or obstacles), at least in one color component; to separate the components, the concept of *chromatic saturation* is used [7].

Since lane detection is generally based on the localization of *specific patterns* (lane markings), it can be performed with the analysis of a *single* still image. In addition, some assumptions can aid the detection and/or speed-up the processing.

- Due to both physical and continuity constraints, the processing of the whole image can be replaced by the analysis of specific *regions of interest* only (the so-called *focus of attention*), in which the features of interest are more likely to be found. This is a generally followed strategy that can be adopted assuming a priori knowledge on the road environment.

  For example the system developed by the Robert Bosch GmbH research group employs a model both for the road and the vehicle's dynamic to determine the portion of the road where it is most likely to find lane markings [13].
- The assumption of a *fixed or smoothly varying lane width* allows the enhancement of the search criterion, limiting the search to almost parallel lane markings.

As an example, on the PVR III vehicle, lane markings can be detected using both neural networks and simple vision algorithms: two parallel stripes of the acquired image are selected and filtered using Gaussian masks and zero crossing to find vertical edges. The result is matched against a given model (a typical road pattern with parallel lane markings) to compute a steering angle and a *fitness* evaluation indicating the confidence in the result [17].

Analogously, the RALPH system is based on the processing of the image portion corresponding to the road about 20–70 m ahead of the vehicle, depending on the vehicle's speed and obstacles presence. The perspective effect is removed from this portion of the image and the determination of the curvature is carried out according to a number of possible curvature models for a specific road template featuring parallel road markings [24].

- The reconstruction of road geometry can be simplified by assumptions on its *shape*.

The research groups of the Universität der Bundeswehr [20] and Daimler-Benz [12] base their road detection functionality on a specific road model: lane markings are modeled as *clothoids*. In a clothoid the curvature depends linearly on the curvilinear reference. This model has the advantage that the knowledge of only two parameters allows the full localization of lane markings and the computation of other parameters like the lateral offset within the lane, the lateral speed with respect to the lane, and the steering angle.

Other research groups use a polynomial representation for lane markings. In the MOSFET autonomous vehicle, for instance, lane markings are modeled as parabolas [22]. A simplified Hough transform is used to accomplish the fitting procedure.

Similarly, the lane detection system developed at The Ohio State University Center for Intelligent Transportation Research relies on a polynomial curve [25]. It assumes a flat road with either continuous or dashed bright lane markings. The history of previously located lane markings is used to determine the region of interest, thus reducing the portion of the image to be processed. The algorithm extracts the significant bright regions from the image plane and stores them in a vector list. Qualitative parameters such as the convergence of the lines at infinity or the lane width, known or estimated, are used to extract from the list the candidate lane markings. Finally, in order to handle also dashed lines, a low-order polynomial curve is fitted to the computed vectors.

On the contrary, other systems adopt a more generic model for the road. The ROMA vision-based system uses a contour-based method [26]. A dynamic road model permits the processing of small portions of the acquired images therefore enabling real-time performance. Actually, only straight or small curved roads without intersections are included in this model. Images are processed using a gradient-based filter and a programmable threshold. The road model is used to follow contours formed by pixels that feature a significant gradient direction value.

Analogously, a generic triangular road model has been used on the MOB-LAB experimental vehicle by the research groups of the Università di Parma [3] and Istituto Elettrotecnico Nazionale "G. Ferraris", CNR, Italy [10].

- The knowledge of the specific camera calibration together with the assumption of an *a priori knowledge on the road surface/slope* (i.e., a *flat* road without bumps) can be exploited to simplify the mapping between image pixels and their correspondent world coordinates.

The great part of the previously discussed systems exploit the assumption of a flat road in front of the vehicle in the determination of obstacle distance or road curvature, once the specific features of interest have been localized in the acquired image. The GOLD (Generic Obstacle and Lane Detection) system [1] implemented on the ARGO autonomous vehicle and the already mentioned RALPH system, however, exploit this assumption also in the lane determination process. In fact, in both cases the lane markings detection is performed in a different image domain, representing a bird's eye view of the road, which can be obtained thanks to the flat road assumption.

Table 1 summarizes the pros and cons of the assumptions on which the most common approaches to lane detection rely.

When the processing is aimed not only at a mere lane detection but also at the tracking of lane markings,

Table 1
Pros and cons of the most typical assumptions in lane detection

	Pros	Cons
Focus of attention	Analysis of a small image portion, fast processing, real-time performance, low cost hardware	Does not fit large features, choice of the region of interest is critical
Fixed lane width	Enhancement of the search criterion (parallel lane markings), robustness to shadows and occlusions	Does not match roads with variable lane width
Road shape	Strong model robust w.r.t. shadows and occlusions, eases road geometry reconstruction, simplifies vehicle control	Requires fitting with complex equations, high computational power needed, fails if the road does not match the model
A priori knowledge on the road surface/slope	Simplifies mapping between image pixels and world coordinates (e.g., determination of obstacle distance, road curvature)	Hypotheses (e.g., flat road) are not always met in real cases, approximations or recalibration needed

the temporal correlation between consecutive frames can be used either to ease the feature determination or to validate the result of the processing. The lane detection module implemented and tested on the ARGO vehicle falls in the first case as it restricts the image portion to be analyzed to the nearest neighborhood of the markings previously detected [5]. In a different manner, the lane detection module developed by the research group of the Istituto Elettrotecnico Nazionale "G. Ferraris", once lane markings are found by means of a triangular model, uses the result of previous computations to validate the current one [10].

3.2. Obstacle detection

The criteria used for the detection of obstacles depend on the definition of what an *obstacle* is (see Fig. 1). In some systems the determination of obstacles is limited to the localization of *vehicles*, which is then based on a search for specific patterns, possibly supported by other features, such as shape, symmetry, or the use of a bounding box.

The researchers of the Istituto Elettrotecnico Nazionale "G. Ferraris" limit the processing to the image portion that is assumed to represent the road,

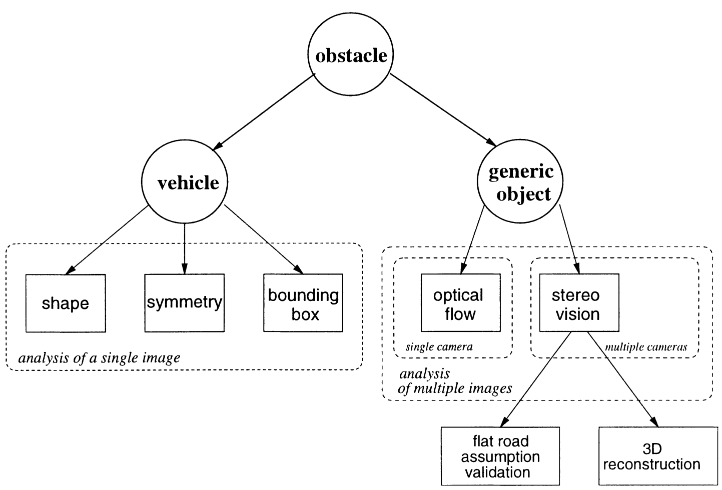


Fig. 1. Depending on the definition of obstacle, different techniques are used.

thus relying on the previously discussed lane detection module. This area of the image is analyzed and borders that could represent a potential vehicle are looked for and examined [10].

Conversely, the obstacle detection algorithm developed at the Universität der Bundeswehr is based on both an edge detection process and an obstacle modelization; the system is able to detect and track up to 12 objects around the vehicle. The obstacle variables continuously updated are: distance, direction, relative speed, relative acceleration, lateral position, lateral speed, and size [20].

When obstacle detection is limited to the localization of specific patterns, as in the previous two examples, the processing can be based on the analysis of a single still image, but the approach is not successful when an obstacle does not match the model.

A more general definition of obstacle, which obviously leads to more complex algorithmic solutions, identifies as an obstacle any object that obstructs the vehicle's driving path or, in other words, anything rising out significantly from the road's surface. In this case, obstacle detection is reduced to identifying the *free space* (the area in which the vehicle can safely move) instead of recognizing specific patterns.

Due to the general applicability of this definition, the problem is dealt with using more complex techniques; the most common ones are based on the processing of two or more images, such as

- the analysis of the *optical flow* field and
- the processing of *non-monocular* images.

In the first case more than one image is acquired by the same sensor in different time instants, whilst in the second one multiple cameras acquire images simultaneously, but from different points of view. Besides their intrinsic higher computational complexity caused by a significant increment in the amount of data to be processed, these techniques must also be robust enough to tolerate noise caused by vehicle movements and drifts in the calibration of the multiple cameras' setup.

The *optical flow*-based technique requires the analysis of a sequence of two or more images: a 2D vector is computed in the image domain, encoding the horizontal and vertical components of the velocity of each pixel. The result can be used to compute ego-motion, which in some systems are directly extracted from odometry; obstacles can be detected by analyzing the difference between the expected and real velocity fields.

As an example, the ROMA system integrates an obstacle detection module that is based on the use of an optical flow technique in conjunction with data coming from an odometer [19].

Similarly, the ASSET-2 (A Scene Segmenter Establishing Tracking v2) is a complete real-time vision system for segmentation and tracking of independently moving objects. Its main feature is that it does not require any camera calibration. It tracks down objects and is capable of correctly handling occlusions amongst obstacles, and automatically tracks down each new object that enters the scene. ASSET-2 initially builds a sparse image flow field and then segmentates it into clusters that feature homogeneous flow variation. Temporal correlation is used to filter the result, therefore improving accuracy [28].

On the other hand, the processing of *non-monocular* image sets requires identifying correspondences between pixels in the different images: two images, in the case of *stereo vision*, and three images, in the case of *trinocular vision*. The advantage of analyzing stereo images instead of a monocular sequence lies in the possibility of directly detecting the presence of obstacles, which, in the case of an optical flow-based approach, is indirectly derived from the analysis of the velocity field. Moreover, in a limit condition where both vehicle and obstacles have small or null speeds, the optical flow-based approach fails while the other can still detect obstacles.

The UTA project (Urban Traffic Assistant) of the Daimler-Benz research group, e.g., aims at an intelligent stop and go for inner-city traffic using stereo vision and obtaining 3D information in real-time. In addition, the UTA demonstrator is able to recognize traffic signs, traffic lights and walking pedestrians as well as the lane, zebra crossings, and stop lines [12].

Also the Massachusetts Institute of Technology group developed a cost-effective stereo vision system. The system is used for three-dimensional lane detection and traffic monitoring as well as for other on-vehicle applications. The system is able to separate partially overlapped vehicles and distinguish them from shadows.

Furthermore, to decrease the intrinsic complexity of stereo vision, some domain-specific constraints are generally adopted.

Table 2
Comparison of the different approaches to obstacle detection

	Pros	Cons
Analysis of a single image	Simple algorithmic solutions, fast processing, does not suffer from vehicle movements	Loss of info about scene's depth unless specific assumptions are made, not successful when obstacles do not match the model
Optical flow	Detection of generic objects, allows computation of ego-motion and obstacles' relative speed	Computationally complex, sensitive to vehicle movements and drifts in calibration, fails when both vehicle and obstacle have small or null speed
Stereo vision	Detection of generic objects, allows 3D reconstruction	Computationally complex (still domain specific constraints may reduce complexity), sensitive to vehicle movements and drifts in calibration

In the GOLD system, the removal of the perspective effect from stereo images allows to obtain two images that differ only where the initial assumption of a flat road is not valid, thus detecting the free space in front of the vehicle [1].

Analogously, the University of California research unit developed an algorithm that remaps the left image using the point of view of the right image, thus detecting disparities in correspondence with the obstacles. A Kalman filter is then used to track obstacles [18].

Table 2 compares the strong and weak points of the different approaches to the obstacle detection problem.

As mentioned above, a great deal of different techniques have been proposed in the literature and tested on a number of vehicles prototype in order to solve the Road Following problem, but only few of them provide an integrated solution (e.g., lane detection and obstacle detection) which, obviously, leads to both an improved quality of the results and to a faster and more efficient processing.

3.3. Hardware trends

In the early years of ITS applications a great deal of custom solutions were proposed, based on ad hoc, *special-purpose* hardware. This recurrent choice was motivated by the fact that the hardware available on the market at a reasonably low cost was not powerful enough to provide real-time image processing capabilities. As an example, the researchers of the Universität der Bundeswehr developed their own system architecture: several special-purpose boards were included in the Transputer-based architecture of the VITA vehicle [11].

Others developed or acquired ad hoc processing engines based on SIMD computational paradigms to exploit the spatial parallelism of images. Among them, the cases of the 16k Mas-Par MP-2 installed on the experimental vehicle NavLab I [14,15] at the Carnegie Mellon University (CMU) and the massively parallel architecture PAPRICA [6] jointly developed by the Università di Parma and the Politecnico di Torino and tested on the MOB-LAB vehicle.

Besides selecting the proper sensors and developing specific algorithms, a large percentage of this first research stage was therefore dedicated to the design, implementation, and test of new hardware platforms. In fact, when a new computer architecture is built, not only the hardware and architectural aspects — such as instruction set, I/O interconnections, or computational paradigm — need to be considered, but software issues as well. Low-level basic libraries must be developed and tested along with specific tools for code generation, optimization and debugging.

In the last few years, the technological evolution led to a change: almost all research groups are shifting towards the use of off-the-shelf components for their systems. In fact, commercial hardware has nowadays reached a low price/performance ratio. As an example, both the NavLab 5 vehicle from CMU and the ARGO vehicle from the Università di Parma are presently driven by systems based on general-purpose processors. Thanks to the current availability of fast internetworking facilities, even some MIMD solutions are being explored, composed of a rather small number of powerful, independent processors, as in the case of the VaMoRs-P vehicle of the Universität der Bundeswehr on which the Transputer processing system has now been partly replaced by a cluster of three PCs (dual Pentium II) connected via a fast Ethernet-based network [20].

Current trends, however, are moving towards a mixed architecture, in which a powerful general-purpose processor is aided by specific hardware such as boards and chips implementing optical flow computation, pattern-matching, convolution, and morphological filters. Moreover, some SIMD capabilities are now being transferred into the instruction set of the last-generation CPUs, which has been tailored to exploit the parallelism intrinsic to the processing of visual and audio (multimedia) data. The MMX extensions of the Intel Pentium processor, for instance, are exploited by the GOLD system which acts as the automatic driver of the ARGO vehicle to boost up performance.

In conclusion, it is important to emphasize that, although the new generation systems are all based on commercial hardware, the development of custom hardware has not lost significance, but is gaining a renewed interest for the production of embedded systems. Once a hardware and software prototype has been built and extensively tested, its functionalities have to be integrated in a fully optimized and engineered embedded system before marketing. It is in this stage of the project that the development of ad hoc custom hardware still plays a fundamental role and its costs are justified through a large-scale market.

3.4. *Extensive tests of prototype vehicles*

Several groups researching on ITS applications have integrated their most promising approaches and solutions for automatic driving into prototypes of intelligent vehicles. Many of these experimental results were presented and demonstrated during important international events such as the final meeting of the PROMETHEUS project in Paris (1994), the NAHSC demonstration in San Diego, CA (August 1997), and the Automated Vehicle Guidance Demo'98 in Rijnwoude, Netherlands (June 1998), and Demo'99 in East Liberty, OH (July 1999).

Generally, these vehicles prototype underwent several tests on structured environments, such as closed tracks. However, only in a few cases extensive tests were carried out on public roads in real traffic conditions:

Fig. 2. The VaMP prototype vehicle.

- the VaMP prototype was experimented on the route from Munich (Germany) to Odense (Denmark) [21] in 1995, *
- the RALPH system was tested on NavLab 5 through a journey (*No Hands Across America* [24]) from Pittsburgh, PA to San Diego, CA in 1995, † and
- the ARGO experimental vehicle was driven by the GOLD system for nearly 2000 km throughout Italy during the *MilleMiglia in Automatico Tour* [4] in 1998. ‡

These tests allowed to experiment the vehicles in the most different conditions and to bring out their robustness and weakness. The encouraging results obtained by all groups testify the maturity achieved by the research in this field.

3.4.1. *Munich to Odense UBM test*

In 1995 the VaMP autonomous vehicle developed at the Universität der Bundeswehr München (UBM) (see Fig. 2) was experimented on a long-distance test from Munich (Germany) to Odense (Denmark).

The tour was aimed at testing the capability, reliability, and performance of highway automatic driving. In particular, the tasks that were performed automatically were lane keeping, longitudinal control, collision avoidance, and lane change maneuvers. The desired travel speed had to be selected by a human driver, taking into account traffic signs, environmental condi-

* See http://www.unibw-muenchen.de/campus/LRT/LRT13/Fahrzeuge/VaMP_E.html.

† See http://www.cs.cmu.edu/afs/cs/user/pomerlea/www/nhaa.html.

‡ See http://MilleMiglia.CE.UniPR.IT.

tions and his own goals. The safety driver was also in charge of starting the automatic lane change maneuvers and supervising the rear hemisphere, since only a front bifocal vision system was used.

The Obstacle Detection and Tracking (ODT) module which was integrated on the VaMP vehicle is based on a recognition process which relies on specific expectations about the appearance of other traffic participants. A combination of two different approaches are used. The single object detector and tracker (SODT) searches for leading vehicles in the same lane by means of a contour following algorithm which extracts the left and right object boundary and then verifies vertical axis symmetry and demands dark wheels and a smooth contour. The second approach, the multiple object detector and tracker (MODT), focuses on the dark area beneath the vehicle. This area is segmented with a contour analysis and then, as a validation, the systems expect two lateral boundaries for vehicles in its own lane and one for those in the neighboring lanes in order to take occlusions into account.

The Road Detection and Tracking (RDT) module estimates a state vector which describes a dynamical model of the vehicle motion and of the road shape. The state vector of the vehicle is composed of the lateral offset, yaw, side slip angle, and pitch angle. The shape of the road is characterized by both a horizontal and vertical clothoid. The estimation is performed by a Kalman filter algorithm which recursively evaluates the measurements of the lane markings position in the image. Left and right lane markings are searched for by applying edge operators of small width in long horizontal search windows. A straight line is then fitted to the detected edge elements.

During this test the system covered more than 1600 km and a percentage of 95% was driven automatically. More than 400 lane change maneuvers were performed automatically. Driving parameters such as the estimated horizontal curvature and estimated lane width have been analyzed subsequently.

Problems occurred at construction zones where the lane markings were ambiguous and could not be correctly interpreted by the RDT module. A major weak point of the system turned out to be the limited operating range of the video cameras which happened to be dazzled by the glare of the sun. Other problems were due to the simple object models on which the ODT module relies. In fact, these hypotheses are not only met by vehicles but sometimes by other objects like certain patterns on the road or rarely bridges' shadows, and this may therefore generate false alarms.

3.4.2. No Hands Across America

In July 1995 researchers from the Robotics Institute of Carnegie Mellon University performed a trans-continental journey from Pittsburgh, PA, to San Diego, CA, in a Pontiac Trans Sport which drove itself most of the way. This trip was the culmination of over a decade of research funded by the US Department of Defense and the US Department of Transportation.

The vehicle, called the NavLab 5 (see Fig. 3), was outfitted with a portable computer, a windshield mounted camera, a GPS receiver, a radar system for obstacle detection as well as some other supplementary equipment. The determination of road curvature was based on the analysis of a picture of the scene ahead and radar data. Once it had this information, it was able to produce steering commands which kept the vehicle in its lane (throttle and brakes were human operated).

The brain of NavLab 5 is a vision-based software system dubbed RALPH. As the vehicle moves along, a video camera mounted just below the rearview mirror reads the roadway, imaging information including lane markings, oil spots, curbs and even ruts made in snow by car wheels. Lane detection is based on the determination of features running parallel to the road, aided by the flat road assumption. Obstacle detection is mainly based on radar information. The driving system runs on the PANS (Portable Advanced Navigation Support) hardware platform. The platform provides a computing base and I/O functions for the system, as well as position estimation, steering wheel control and safety monitoring.

Thanks to this trip the researchers were able to prove the roadworthiness of on-road autonomous lane keeping and lateral roadway departure warning and support systems. The NavLab 5 was able to drive itself 98.2% of the way (2797/2849 miles). Some other statistical data about the tour are reported in Table 3.

The system proved to be robust with respect to real road and traffic conditions. The major problems encountered were due to rain, low sun reflections, shadows of overpasses, construction zones, road and road markings deterioration. A strong point of the system,

Fig. 3. The NavLab 5 Pontiac Transport and the No Hands Across America logo.

however, is the ability to use any feature parallel to the road (such as road boundary or tyre wear marks) in the determination of road curvature: thanks to this characteristic RALPH showed enforced Road Following ability even in case of poor lane markings.

3.4.3. MilleMiglia in Automatico

The ARGO experimental autonomous vehicle (shown in Fig. 4) developed at the Dipartimento di Ingegneria dell'Informazione of the Università di Parma, Italy, is a Lancia Thema passenger car equipped with a vision-based system that allows to extract road and environmental information for the automatic driving of the vehicle, and with different output devices used to test the automatic features.

Only low cost passive sensors (cameras) are used on ARGO to sense the surrounding environment: by means of stereo vision, obstacles on the road are detected and localized, while the processing of a single monocular image allows to extract the road geometry in front of the vehicle. Furthermore, an I/O board is used to acquire information about the velocity of the vehicle, and other user data.

Table 3
Statistical data of the No Hands Across America tour

Total autonomous statistics	
Autonomous driving percentage (%)	98.2 (4503/4587 km)
Longest autonomous segment (km)	111
Average speed ($km\,h^{-1}$)	102.72

Fig. 4. The ARGO experimental vehicle and the MilleMiglia in Automatico logo.

The approach followed for lane detection is based on the extraction of features from a monocular gray-tone image in which the perspective effect has been removed. The features are selected and matched against either a straight road model (during detection) or the previous result (during tracking). Obstacle detection is based on the verification of hypothesis about road shape such as planarity.

The ARGO vehicle has autonomous steering capabilities: the result of the processing (position of obstacles and geometry of the road) is used to drive an actuator on the steering wheel in order to follow the road; moreover human-triggered lane change maneuvers can be automatically performed. For debug purposes, the result of the processing is also fed to the driver through a set of output devices installed on-board of the vehicle: an acoustical device warns the driver in case dangerous conditions are detected, while a visual feedback is supplied to the driver by displaying the results both on an on-board monitor and on a led-based control panel.

In order to extensively test the vehicle under different traffic situations, road environments, and weather conditions, a 2000 km journey was carried out in June 1998. During this test, ARGO drove itself for about 94% of the total trip along the Italian highway network, passing through flat areas and hilly regions including viaducts and tunnels. The Italian road network is particularly suited for such an extensive test since it is characterized by quickly varying road scenarios which include changing weather conditions and a generally high amount of traffic.

The analysis of the data collected during the tour allowed the computation of a number of statistics regarding system performance (see Fig. 5). In particular, for each stage of the tour the average and the maximum speed of the vehicle during automatic driving were computed. The average speed was strongly influenced by the heavy traffic conditions (especially on Torino, Milano and Roma's by-passes) and by the presence of toll stations, junctions, and road works. The automatic driving percentage and the maximum distance automatically driven show high values despite the presence of many tunnels (particularly during the Appennines routes Ancona–Roma and Firenze–Bologna) and of several stretches of road with absent or worn lane markings (Ferrara–Ancona and Ancona–Roma) or even no lane markings at all (Firenze–Bologna). It is

Leg	Date	Departure	Arrival	kilometers	average speed [km/h]	maximum speed [km/h]	percentage of automatic driving	maximum distance in automatic [km]
1	1st June 1998	Parma	Torino	245	86.6	109	93.3	23.4
2	2nd June 1998	Torino	Pavia	175	80.2	95	85.1	42.2
3	2nd June 1998	Pavia	Ferrara	340	89.8	115	86.4	54.3
4	3th June 1998	Ferrara	Ancona	260	89.8	111	91.1	15.1
5	4th June 1998	Ancona	Roma	365	88.4	108	91.1	20.8
6	5th June 1998	Roma	Firenze	280	87.5	110	95.4	30.6
7	6th June 1998	Firenze	Parma	195	89.0	123	95.1	25.9

Fig. 5. Statistical data regarding system performance during the MilleMiglia in Automatico tour.

fundamentally important to also note that some stages included passing through toll stations and transiting in by-passes with heavy traffic and frequent queues during which the system had to be switched off.

Apart from some difficulties due to the sometimes poor road network infrastructures (absent or worn lane markings), the major problems encountered were the sun light's reflection on the windscreen causing the acquisition of oversaturated images, and the cameras' inadequacy to quick changes in the illumination (such as at the entrance or exit from a tunnel) provoking a degradation in the image quality.

4. Conclusions and perspectives on intelligent vehicles

Some common considerations can be drawn from the three experiments described: all of them succeeded in reaching an extremely high percentage of automatic driving, but although in the first experiments some special-purpose hardware was needed, more recent experiments benefited from the latest technological advances and used only commercial hardware.

Anyway, though for this kind of application computing power do not seem to be a problem any more, still some problems remain regarding data acquisition. It has been shown that the main difficulties encountered during the demos were due to light reflections and non-perfect conditions for image acquisition (wet road, direct sunshine on the cameras, tunnels and bridges' shadows). As a common framework for the next years of research a great deal of work will be addressed towards the enhancement of sensor's capabilities and performance, including the improvement of gain control and sensitivity in extreme illumination conditions.

The promising results obtained in the first stages of the research on intelligent vehicles demonstrate that a full automation of traffic (at least on motorways or sufficiently structured roads) is technically feasible. Nevertheless, besides the technical problems there are some issues that must be carefully considered in the design of these systems such as the legal aspects related to the responsibility in case of faults and incorrect behavior of the system, and the impact of automatic driving on human passengers.

Therefore, a long period of exhaustive tests and refinement must precede the availability of these systems on the general market, and a fully automated highway system with intelligent vehicles driving and exchanging information is not expected for a couple of decades.

For the time being, complete automation will be restricted to special infrastructures such as industrial applications or public transportation. Then, automatic vehicular technology will be gradually extended to other key transportation areas such as the shipping of goods, e.g., on expensive trucks, where the cost of an autopilot is negligible with respect to the cost of the vehicle itself and the service it provides. Finally, once technology has been stabilized and the most promising solution and best algorithms freezed, a massive integration and widespread use of such systems will also take place with private vehicles, but this will not happen for another two or more decades.

References

[1] 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.

[2] D. Bishop, Vehicle-highway automation activities in the United States, in: Proceedings of the International AHS Workshop, US Department of Transportation, 1997.

[3] A. Broggi, S. Bertè, Vision-based road detection in automotive systems: A real-time expectation-driven approach, Journal of Artificial Intelligence Research 3 (1995) 325–348.

[4] A. Broggi, M. Bertozzi, A. Fascioli, The 2000 km test of the ARGO vision-based autonomous vehicle, IEEE Intelligent Systems (1999) 55–64.

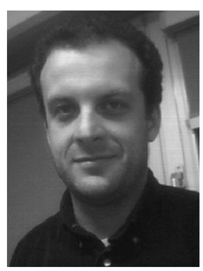
[5] A. Broggi, M. Bertozzi, A. Fascioli, G. Conte, Automatic Vehicle Guidance: The Experience of the ARGO Vehicle, World Scientific, Singapore, 1999.

[6] A. Broggi, G. Conte, F. Gregoretti, C. Sansoè, L.M. Reyneri, The evolution of the PAPRICA system, Integrated Computer-Aided Engineering Journal (special issue on Massively Parallel Computing) 4 (2) (1997) 114–136.

[7] P. Charbonnier, P. Nicolle, Y. Guillard, J. Charrier, Road boundaries detection using color saturation, in: Proceedings of the Ninth European Signal Processing Conference '98, September 1998.

[8] A. Coda, P.C. Antonello, B. Peters, Technical and human factor aspects of automatic vehicle control in emergency situations, in: Proceedings of the IEEE International Conference on Intelligent Transportation Systems '97, 1997.

[9] J.D. Crisman, C.E. Thorpe, UNSCARF, a color vision system for the detection of unstructured roads, in: Proceedings of the IEEE International Conference on Robotics and Automation, Sacramento, CA, April 1991, pp. 2496–2501.
[10] S. Denasi, C. Lanzone, P. Martinese, G. Pettiti, G. Quaglia, L. Viglione, Real-time system for road following and obstacle detection, in: Proceedings of the SPIE on Machine Vision Applications, Architectures, and Systems Integration III, October 1994, pp. 70–79.
[11] E.D. Dickmanns, Expectation-based multi-focal vision for vehicle guidance, in: Proceedings of the Eighth European Signal Processing Conference '95, Trieste, Italy, September 1995, pp. 1023–1026.
[12] U. Franke, D. Gavrila, S. Görzig, F. Lindner, F. Paetzold, C. Wöhler, Autonomous driving goes downtown, in: Proceedings of the IEEE Intelligent Vehicles Symposium '98, Stuttgart, Germany, October 1998, pp. 40–48.
[13] J. Goldbeck, D. Graeder, B. Huertgen, S. Ernst, F. Wilms, Lane following combining vision and DGPS, in: Proceedings of the IEEE Intelligent Vehicles Symposium '98, Stuttgart, Germany, October 1998, pp. 445–450.
[14] T.M. Jochem, S. Baluja, Massively parallel, adaptive, color image processing for autonomous road following, in: H. Kitano (Ed.), Massively Parallel Artificial Intelligence, AAAI Press/MIT Press, Cambridge, MA, 1993.
[15] T.M. Jochem, S. Baluja, A massively parallel road follower, in: M.A. Bayoumi, L.S. Davis, K.P. Valavanis (Eds.), Proceedings of the IEEE Computer Architectures for Machine Perception, New Orleans, LA, December 1998, pp. 2–12.
[16] T.M. Jochem, D.A. Pomerleau, C.E. Thorpe, MANIAC: A next generation neurally based autonomous road follower, in: Proceedings of the Third International Conference on Intelligent Autonomous Systems, Pittsburgh, PA, February 1993.
[17] K.I. Kim, S.Y. Oh, S.W. Kim, H. Jeong, C.N. Lee, B.S. Kim, C.S. Kim, An autonomous land vehicle PRV II: Progresses and performance enhancement, in: Proceedings of the IEEE Intelligent Vehicles Symposium '95, Detroit, MI, September 1995, pp. 264–269.
[18] D. Koller, J. Malik, Q.-T. Luong, J. Weber, An integrated stereo-based approach to automatic vehicle guidance, in: Proceedings of the Fifth International Conference on Computer Vision, Boston, MA, 1995, pp. 12–20.
[19] W. Kruger, W. Enkelmann, S. Rossle, Real-time estimation and tracking of optical flow vectors for obstacle detection, in: Proceedings of the IEEE Intelligent Vehicles Symposium '95, Detroit, MI, September 1995, pp. 304–309.
[20] M. Lützeler, E.D. Dickmanns, Road recognition with MarVEye, in: Proceedings of the IEEE Intelligent Vehicles Symposium '98, Stuttgart, Germany, October 1998, pp. 341–346.
[21] M. Maurer, R. Behringer, F. Thomanek, E.D. Dickmanns, A compact vision system for road vehicle guidance, in: Proceedings of the 13th International Conference on Pattern Recognition, Vienna, Austria, 1996.
[22] S.L. Michael Beuvais, C. Kreucher, Building world model for mobile platforms using heterogeneous sensors fusion and temporal analysis, in: Proceedings of the IEEE International Conference on Intelligent Transportation Systems '97, Boston, MA, November 1997, p. 101.
[23] M. Mizuno, K. Yamada, T. Nakano, S. Yamamoto, Robustness of lane mark detection with wide dynamic range vision sensor, in: Proceedings of the IEEE Intelligent Vehicles Symposium '96, Tokyo, 1996, pp. 171–176.
[24] D.A. Pomerleau, T.M. Jochem, Rapidly adapting machine vision for automated vehicle steering, IEEE Expert 11 (2) (1996).
[25] K.A. Redmill, A simple vision system for lane keeping, in: Proceedings of the IEEE International Conference on Intelligent Transportation Systems '97, Boston, MA, November 1997, p. 93.
[26] R. Risack, P. Klausmann, W. Krüger, W. Enkelmann, Robust lane recognition embedded in a real-time driver assistance system, in: Proceedings of the IEEE Intelligent Vehicles Symposium '98, Stuttgart, Germany, October 1998, pp. 35–40.
[27] U. Seger, H.-G. Graf, M.E. Landgraf, Vision assistance in scenes with extreme contrast, IEEE Micro (1993) 50–56.
[28] S.M. Smith, J.M. Brady, ASSET-2: Real-time motion segmentation and shape tracking, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (8) (1995) 814–829.
[29] C.G. Sodini, S.J. Decker, A 256 × 256 CMOS brightness adaptive imaging array with column-parallel digital output, in: Proceedings of the IEEE Intelligent Vehicles Symposium '98, Stuttgart, Germany, October 1998, pp. 347–352.
[30] C.J. Taylor, J. Košecká, R. Blasi, J. Malik, A comparative study of vision-based lateral control strategies for autonomous highway driving, International Journal of Robotics Research 18 (5) (1999) 442–453.
[31] H. Tokuyama, Asia–Pacific projects status and plans, in: Proceedings of the International AHS Workshop, US Department of Transportation, 1997.



Massimo Bertozzi received the Dr. Eng. (Master) degree in Electronic Engineering (1994) from the Università di Parma, Italy, discussing a Master's Thesis about the implementation of Simulation of Petri Nets on the CM-2 Massive Parallel Architecture. From 1994 to 1997 he was a Ph.D. student in Information Technology at the Dipartimento di Ingegneria dell'Informazione, Università di Parma, where he chaired the local IEEE student branch. During this period his research interests focused mainly on the application of image processing to real-time systems and to vehicle guidance, the optimization of machine code at assembly level, and parallel and distributed computing. After obtaining the Ph.D. degree, in 1997, he has been holding a permanent position at the same department. He is a member of IEEE and IAPR.

Alberto Broggi received both the Dr. Eng. (Master) degree in Electronic Engineering (1990) and Ph.D. degree in Information Technology (1994) from the Università di Parma, Italy. From 1994 to 1998 he was a Full Researcher at the Dipartimento di Ingegneria dell'Informazione, Università di Parma, Italy. Since 1998 he is Associate Professor of Artificial Intelligence at the Dipartimento di Informatica e Sistemistica, Università di Pavia, Italy. His research interests include real-time computer vision approaches for the navigation of unmanned vehicles, and the development of low-cost computer systems to be used on autonomous agents. He is the coordinator of the ARGO project, with the aim of designing, developing and testing the ARGO autonomous prototype vehicle, equipped with special active safety features and enhanced driving capabilities. Professor Broggi is the author of more than 100 refereed publications in international journals, book chapters, and conference proceedings. Actively involved in the organization of scientific events, Professor Broggi is on the Editorial Board and Program Committee of many international journals and conferences and has been invited to act as Guest-Editor of journals and magazines theme issues on topics related to Intelligent Vehicles, Computer Vision Application, and Computer Architectures for Real-Time Image Processing. Professor Broggi is the Newsletter Editor and member of the Conference and Publication Committees of the IEEE Intelligent Transportation Systems Council and will be the Program Chair of the next IEEE Intelligent Vehicles Symposium, Detroit, MI, 2000.

Alessandra Fascioli received the Dr. Eng. (Master) degree in Electronic Engineering (1996) from the Università di Parma, Italy, discussing a Master Thesis on stereo vision-based obstacle localization in automotive environments. In January 2000 she received the Ph.D. degree in Information Technology at the Dipartimento di Ingegneria dell'Informazione, Università di Parma. Her research interests focus on real-time computer vision for automatic vehicle guidance and image processing techniques based on the Mathematical Morphology computational model. She is a member of IEEE and IAPR.