

Obstacle and Lane Detection on ARGO*

Massimo Bertozzi, Alberto Broggi, Gianni Conte, Alessandra Fascioli
Dipartimento di Ingegneria dell'Informazione
Università di Parma, I-43100 Parma, Italy
http://WWW.CE.UniPR.IT/computer_vision
E-Mail: {bertozzi, broggi, conte, fascal}@CE.UniPR.IT

Abstract

This work presents ARGO, the experimental land vehicle developed at the Dipartimento di Ingegneria dell'Informazione of the University of Parma, Italy.

ARGO integrates the GOLD (Generic Obstacle and Lane Detection) system, a stereo vision-based hardware and software architecture that allows to detect both generic obstacles (without constraints on shape, color, or symmetry) on flat roads and the lane position in structured environments (with painted lane markings). In addition, this paper presents a new approach that allows to handle also non-flat roads.

1 Introduction

This paper presents the ARGO Autonomous Vehicle, developed at the Dipartimento di Ingegneria dell'Informazione, University of Parma, Italy. ARGO, a Lancia Thema 2000 passengers car, is an experimental road vehicle integrating the main results [3, 2] of the research conducted in the last few years on MOB-LAB. The MOB-LAB (MOBile LABoratory) vehicle, developed within the PROMETHEUS project and officially presented and demonstrated during the last final PROMETHEUS BMM Meeting (Le Mortefontaine Track, Paris, October 1994), was extensively used by our group to test both algorithms and ad-hoc computer architectures for vision-based navigation. The availability of MOB-LAB allowed to

develop, test, and tune a number of different approaches for autonomous navigation; however, special emphasis was given to Obstacle Detection (OD) and Lane Detection (LD), since they represent basic functionalities for automatic road vehicle driving.

MOB-LAB was manually driven and the activities of the on-board instruments were limited to the displaying of warnings to the driver. Nevertheless, the analysis of the differences between the driver's behavior and the actions suggested by the system allowed to test its reliability under many different road and weather conditions.

After this first research stage, which led to the development of a number of different solutions to the LD problem, the most promising approach [2] was selected to be integrated on board of the new ARGO vehicle, shown in fig. 1. ARGO¹ is now being equipped with actuators for automatic steering. The pair of synchronized low-cost stereo cameras that have been installed on ARGO have a size of 3.2×3.2 cm and are used to acquire pairs of gray-

¹The name ARGO derives from the Italian name of two distinct myths related to *navigation* and *vision*: (i) Argo, the first ship with name in history, used by Jason and a band of 50 mythological heroes to *navigate* all the Mediterranean sea and to fetch the GOLDen fleece. As the Argo was equipped with all those implements and tackling necessary for the management and guiding of the ship in open seas and through the blows of the huge waves, she can represent the archetype of the perfect navigation media; (ii) Argus Παρόπτῆς (the *all-seeing*), the many eyed god with the gift of all-round vision. Since Argus had a hundred eyes and could have some of them sleep while others were awake, he made a fine watchman and thus was appointed by the goddess Hera to watch the cow into which Io (Hera's priestess) had been transformed. Disguised as a Shepherd, Hermes gained Argus confidence and lulled him to sleep. Once asleep Hermes killed Argus. As a memorial, Hera took his eyes and set them into the tail of her favorite bird, the peacock.

*This work was partially supported by the Italian National Research Council (CNR) under the frame of the *Progetto Finalizzato Trasporti 2*.

level images at a resolution of 768×288 pixels (only a single image field instead on a full frame of resolution 768×576).

The *processing engines* that have been evaluated to support the real-time processing of images range from special-purpose dedicated SIMD systems [4] to general-purpose off-the-shelf commercial processors; these alternatives are now being considered and carefully evaluated.

On the other hand, the *software system* currently being ported to ARGO and described in this work is an enhanced version of GOLD [3, 2]. Both functionalities integrated in the GOLD (Generic Obstacle and Lane Detection) system are based on the elimination of the perspective effect [8, 9], from pairs of stereo images the former and from a single image the latter.

Lane Detection, as a stand alone module, relies on the presence of painted road markings, and is based on a low-level feature-extraction process followed by a high-level interpretation, aimed to the recovery of the road geometry and to the reconstruction of occlusions caused by other vehicles on the path. An extension to handle also poorly structured roads (as it happens in RALPH [9], where not only road markings are used, but features running parallel to the road as well) exploiting the high image vertical correlation is now under evaluation.

Conversely, Obstacle Detection is based on the detection of disparities between pairs of stereo images. As it happens in many systems in which real-time is a basic constraint, the generality of the approach is traded for a faster and simpler processing: in this case, in fact, no full search for correspondences (*homologous points*) is performed [6] but only a check for the consistency with a predefined set of hypotheses, such as a given road model. More precisely, in the original GOLD system (ver. 1.0), as well as in a large number of other similar systems [6], a *planar* road model is assumed, and the target of the processing is aimed to the detection of deviations from this model; such deviations represent objects above the road surface, thus locating potential obstacles. Contrary to similar approaches [7], this first processing step does not return an estimate of the obstacle distance, but only its angular position with respect to the cameras. In other words, no full 3D reconstruction is performed, thus avoiding to deal with complex road



Figure 1: The ARGO experimental vehicle: the figure shows the cameras installed under the upper part of the windshield and the ARGO interior, where the on-board monitor can be seen

and environmental internal models, as in [5]. Instead, a following module determines the obstacle distance in a very simple but reliable qualitative way, so that only the *free space* in front of the vehicle is computed, namely the portion of the road where the vehicle can safely move. This is a further simplification that is exploited to speed-up the processing by eliminating a time-consuming reconstruction of the 3D environment. The output devices currently installed on ARGO are:

- a monitor which, along with different information, shows both lane markings and obstacles with different markers of different colors;
- a set of leds indicating the correct position of the steering wheel. An actuator on the steering wheel will soon replace this output device.

The *original* GOLD system was aimed to the *independent* recovery of the road geometry and the *independent* detection of obstacles. The assumptions that were necessary for the system to work were:

1. the presence of *painted road markings* (for lane detection);
2. a *flat road* in front of the vehicle (for both obstacle and lane detection).

The possibility to eliminate these two assumptions is now being evaluated and is discussed in the final part of this work along with a discussion on the computing engines.

2 The Evolution of



In this section the two functionalities (LD and OD) are briefly summarized (more details can be found in [2]). Both of them share the same underlying approach (*Inverse Perspective Mapping*, IPM), that, once known all the acquisition intrinsic and extrinsic parameters, thanks to a geometrical transform, remaps the acquired images in a new domain, called *road domain*, that represents a bird's eye view of the road surface; obstacle detection is based on the detection of disparities between a stereo pair, while the detection of lane markings is extremely simplified since they can be devised as almost vertical lines with constant width.

Moreover, since both functionalities are based on the processing of images remapped into the *same* domain, the fusion of the result of the two processings is straightforward.

2.1 Obstacle Detection Functionality

The use of the IPM transform on stereo images allows to obtain two patches (*remapped images*) of the road texture which can be brought to correspondence exploiting the knowledge of the vision system setup. Any difference in the two remapped images represents a deviation from the starting hypothesis of flat road and thus identifies a potential obstacle, namely anything rising up from the road surface.

An obstacle is detected when the *difference image*, obtained comparing the two remapped images, presents sufficiently large clusters of non-zero pixels having a specific shape [1]. Thus, the low-level portion of the processing is reduced to the computation of the difference between the remapped images and to the binarization of the resulting image. Conversely, the medium and high level processing steps consist of the analysis of a set of histograms built starting from the difference image [1].

2.2 Lane Detection Functionality

In the remapped image lane markings can be devised as almost *vertical* bright lines of *constant* width, surrounded by a darker region. In this case the pixels belonging to a road marking have a brightness value higher than their horizontal left and right neighbors at a given distance. Thus the first phase of lane detection is based on the search for dark-bright-dark horizontal patterns with a specific size. The brightness value of every pixel is compared to that of its left and right horizontal neighbors and a new grey-level image is computed which encodes the horizontal brightness transitions and thus the presence of lane markings.

Different illumination conditions, such as shadows or sunny blobs, cause road markings to have different brightness values; anyway the pixels representing road markings maintain a brightness value higher than their horizontal neighbors. Thus the image is enhanced taking advantage of its vertical correlation, and then binarized using an adaptive threshold

The goal of the following medium-level processing is the determination of the almost vertical lines in the binarized image. This step differs from the one described in [2] (GOLD ver. 1.0). The binarized image is scanned line by line horizontally starting from the bottom in order to build *chains* of non-zero pixels. When a non-zero pixel is found, the following actions are taken: if the distance between the pixel and the nearest extremum of a chain is less than a given threshold, the pixel is assigned to the chain, else a new chain formed by only this pixel is started. These chains are then segmented.

Since road markings are often not entirely detectable in the binarized image, the system tries to join the polylines into longer ones, thus filling the gaps produced by either obstacles or an ineffective low-level processing; dashed lane markings are also converted into continuous lines.

The decision whether to join the polylines is based on the evaluation of many parameters, such as the horizontal and vertical distance between the nearest extrema, the angles formed by the final segments of each polyline and the segment that joins the nearest extrema of the polylines. Since the parameters taken

into account depend only on the ending segments of each pair of polylines, the joined polylines need to be evaluated also globally: those featuring a too high or too variable curvature and those that are shorter than a given threshold are discarded.

Then, starting from the assumption of a two-lane road with smoothly varying lane width, the system searches for parallel polylines; for sake of simplicity a loose concept of parallelism is used: only the *horizontal* distance between two polylines is used to determine their parallelism. For this purpose, the intersections among the polylines and a set of equidistant horizontal lines are considered in order to compute the horizontal distances between each pair of polylines.

Thanks to the knowledge of the vision system setup and to the flat road hypothesis, the spatial relationship between pixels of the acquired image and the road is known, thus allowing to estimate both the road geometry and the vehicle position with respect to the lane.

3 Discussion

The previous sections described the new version (ver. 1.1) of GOLD as it was tested both in laboratory, thanks to pre-recorded video streams and digitalized stereo pairs, and on board of the MOB-LAB experimental land vehicle for more than 3000 km.

This section presents a new cooperative approach that is currently under final evaluation that allows to perform a fusion of the results of the two modules (lane detection and obstacle detection) and takes into account also the high temporal correlation of image sequences; this enhancement is now being ported to ARGO.

A final discussion of the computing system is also presented.

3.1 A Cooperative Approach

Once obstacles have been detected and localized, the knowledge of their position can be of great help to the lane detection module, since obstacles generally obstruct the visibility of road markings: the regions occluded by obstacles will not be considered in the

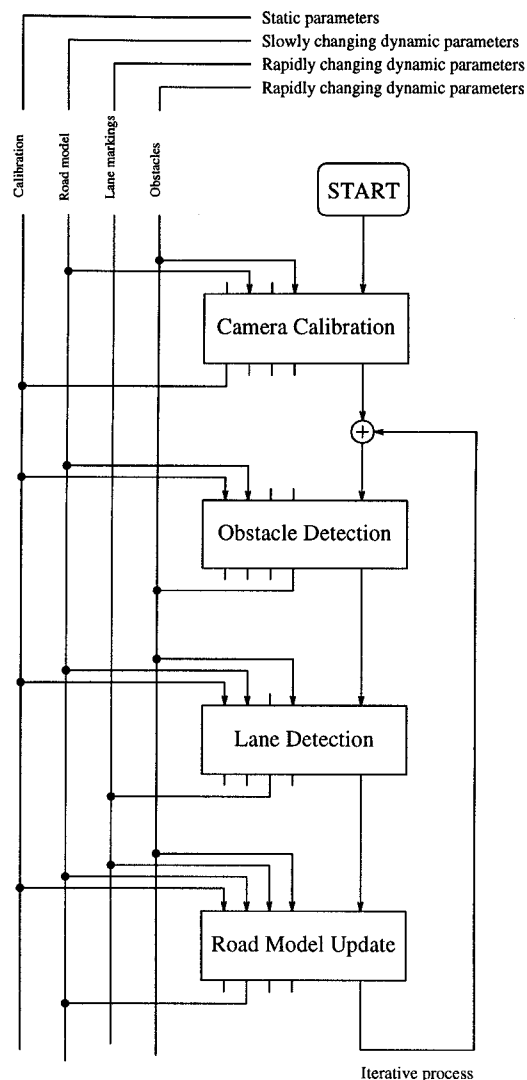


Figure 2: Block diagram of the cooperative approach: the left part depicts the data stream, while the right part shows the control flow of the algorithm

search for road markings, thus decreasing noisy features that may disturb the retrieval of the road geometry.

Similarly, supposing to know the road geometry (namely the position of road markings within the image), and the precise position of obstacles, a pair of stereo images can be analyzed to validate the assumption of planarity of the road surface: the stereo pair can be used to calibrate the stereo system and adapt it to new road surfaces, such as hills, bridges, or non-planar highway ramps; the new model will replace the old one and will be used by both OD and

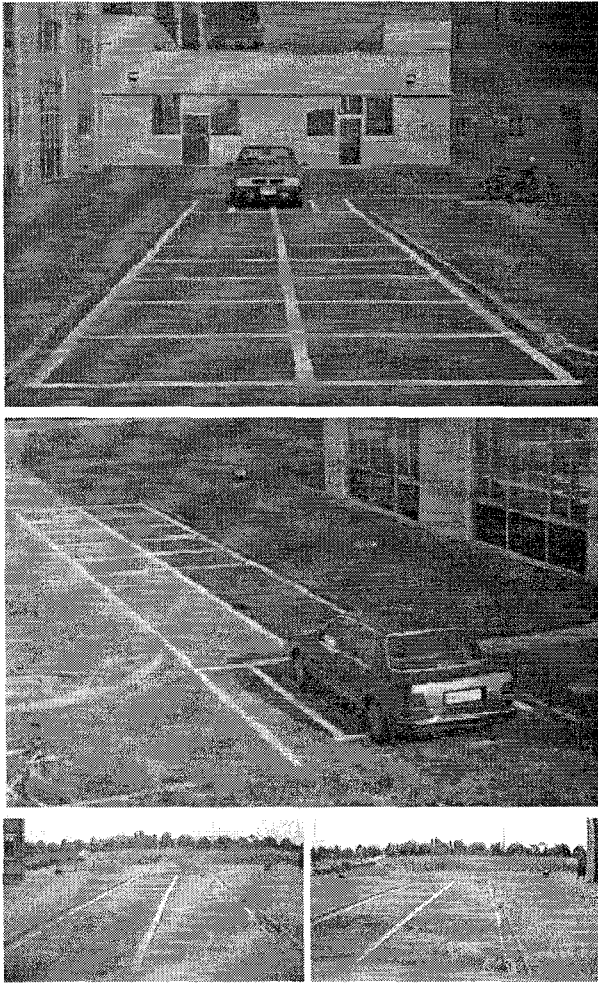


Figure 3: The road segment with a painted grid of known size and the images acquired by the stereo vision system

LD modules.

These considerations led to the development of an integrated cooperative approach in which the OD and LD modules are triggered alternatively; each module exploits the results produced by the previous module and feeds the following one with its own results, thus keeping an updated data structure describing three *dynamic* parameters: (1) obstacles position, (2) road geometry, and (3) road surface. This iterative process is initialized by hand when no obstacles are visible and the road in front of the vehicle is flat; the initialization phase is used to determine the calibration of the acquisition system, i.e. the *static* parameters.

More precisely, this integrated approach, depicted in figure 2, is described as follows.

3.1.1 System Calibration

The system needs to be initialized in known conditions. Since the calibration of the acquisition system consists in the determination of the angles the two cameras' optical axes form with vertical and horizontal directions, the geometry of the road surface in front of the vehicle must be known. Hence the calibration is triggered by hand when the portion of the road visible by both cameras is flat and no obstacle is present. A road segment in front of our building (see fig. 3) has been painted with a grid of known size to ease the calibration phase. This phase takes 30 ÷ 40 seconds, since an exhaustive search in the parameters' space is performed: each parameter (i.e. angles determining the cameras' orientations) scans a given interval and for each parameters configuration the two *remapped* stereo images are computed according to a flat road model. The set of parameters for which the two remapped stereo images present the minimum pixel-wise difference is chosen to be the correct cameras calibration, which remain fixed for the whole run.

3.1.2 Obstacle Detection

Using the given road model and the knowledge of the acquisition system parameters, obstacles are detected as discussed in section 2.1

3.1.3 Lane Detection

The left image is processed as discussed in section 2.2 and the geometry of the road extracted.

3.1.4 Road Model Update

The results of the previous step indicates where lane markings are located or, in case of occlusions caused by obstacles, where they are expected to be.

Assuming to know the position of obstacles within the cameras' field of view, the portion of lane markings that has been generated to overcome occlusions can be discarded. Thus the remaining part identifies features that belong to the road surface; these features are particularly important because they represent horizontal brightness transitions (lane markings) in the left remapped image and the search for

correspondent patterns in the right remapped image (on the same epipolar line [6]) allows to determine the “height” of the road markings with respect to the original $z = 0$ plane used for calibration.

Since lane markings belong to the road surface, this step is used to update the internal model of the road surface. The computation of the “height” of a rather large number of points that represent road markings, followed by a temporal average, allows to filter and compensate for noise caused by sudden vehicle movements. On the other hand, the use of a focus of attention limited only to points that represent road markings, simplifies and speeds-up the processing.

3.2 The Computing Engine

In the original GOLD system presented in [2], a high processing rate (10 Hz) was obtained thanks to the use of the PAPRICA special-purpose hardware system: PAPRICA is a massively parallel SIMD architecture composed of 256 processing elements (PEs), enhanced with a set of features explicitly included to speed-up the execution of the GOLD functionalities, such as the possibility to acquire and store in the image memory pairs of grey-level stereo images at video-rate and to apply geometrical transforms (such as the IPM) thanks to a dedicated hardware module.

Thanks to the analysis of the main system bottlenecks, a new improved version of the PAPRICA architecture has already been developed [4] and is currently under fabrication.

In addition, the evolution toward a general purpose architecture is now being evaluated: currently the GOLD system is implemented on a 200 MHz Pentium processor with MMX (Multimedia Extensions) technology. The PCI frame grabber allows to acquire both stereo images directly in the system memory using DMA. The use of a MMX technology allows to speed-up the computation thanks to the possibility to exploit the specific features of such processors, such as SIMD computations or saturation arithmetics.

In particular, since all the computations involve the use of gray-level images, pixels are represented by one byte each; this allows to use the specific instructions offered by MMX, such as the allocation of 8 pixels on the same processing word (64 bits) and

the processing with a parallelism of 8.

Unfortunately, since up to now no support is given by public domain compilers (GNU GCC ver. 2.7.2) to the use of MMX-based instructions, we had to develop our routines directly in assembly language.

The first preliminary results that we had showed that the exploitation of these specific features at assembly level, even if the code was not fully optimized, allowed to achieve a speed-up ranging from one order of magnitude in the low-level image processing steps to a factor 3-4 in the remaining part.

References

- [1] M. Bertozzi and A. Broggi. Real-Time Lane and Obstacle Detection on the GOLD System. In I. Masaky, editor, *Proceedings IEEE Intelligent Vehicles'96*, pages 213–218, Tokyo, Japan, September 18-20 1996. IEEE Computer Society.
- [2] M. Bertozzi and A. Broggi. GOLD: a Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection. *IEEE Transactions on Image Processing*, 1997. In press.
- [3] M. Bertozzi and A. Broggi. Vision-based Vehicle Guidance. *IEEE Computer*, 30(7):49–55, July 1997.
- [4] A. Broggi, G. Conte, F. Gregoretti, C. Sansoè, and L. M. Reyneri. The Evolution of the PAPRICA System. *Integrated Computer-Aided Engineering Journal - Special Issue on Massively Parallel Computing*, 4(2):114–136, 1997.
- [5] E. D. Dickmans and B. D. Mysliwetz. Recursive 3-D Road and Relative Ego-State Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:199–213, May 1992.
- [6] O. Faugeras. *Three-dimensional Computer Vision, a Geometric Viewpoint*. MIT Press, Cambridge, 1993.
- [7] D. Koller, J. Malik, Q.-T. Luong, and J. Weber. An integrated stereo-based approach to automatic vehicle guidance. In *Proceedings of the Fifth ICCV*, pages 12–20, Boston, 1995.
- [8] L. Matthies. Stereo vision for planetary rovers: stochastic modeling to near real-time implementation. *International Journal of Computer Vision*, 8:71–91, 1992.
- [9] D. Pomerlau and T. Jochem. Rapidly Adapting Machine Vision for Automated Vehicle Steering. *IEEE Expert*, 11(2), April 1996.