Real-Time Obstacle Detection using Stereo Vision

Massimo Bertozzi, Alberto Broggi, Alessandra Fascioli
Dipartimento di Ingegneria dell'Informazione
Università di Parma, I-43100 Parma, Italy
Tel. +39-521-905707 Fax. +39-521-905723
e-mail: {bertozzi,broggi,fascal}@CE.UniPR.IT

ABSTRACT

This work presents a low-cost stereo vision system aimed to the real-time detection of generic obstacles (without constraints on symmetry or shape) on the path of a mobile road vehicle.

Thanks to a geometrical transform the perspective effect is removed from both left and right stereo images. The difference between the results is used for the detection of free-space in front of the vehicle. The output of the processing is displayed on both an on-board monitor and a control-panel to give a visual feedback to the driver.

The system was tested on MOB-LAB experimental land vehicle, which was driven for more than 3000 km along extra-urban roads and freeways at speeds up to 80 km/h, and demonstrated its robustness with respect to shadows and changing illumination conditions, different road textures, and vehicle movement.

1 INTRODUCTION

The techniques used in the detection of obstacles may vary according to the definition of "obstacles". If "obstacle" means a vehicle, then the detection is based on a search for specific patterns, possibly supported by other features, such as shape [11], symmetry [12], or the use of a bounding box [1]. In this case the processing can be based on the analysis of a single still image.

Conversely, if we intend as obstacle any object that can obstruct the vehicle's driving path or anything raising out significantly from the road surface, obstacle detection is generally reduced to the determination of *free-space* instead of the detection of specific patterns. Thus in order to be fast and robust with respect to camera calibration and vehicle movements, the detection of a *generic obstacle* is reduced to the determination of the *free-space* in front of the vehicle without any 3D world reconstruction; this reduces the traditional problems of *stereo vision techniques*, namely the detection of correspondences between two stereo images (or 3 images in case of trinocular vision [9]).

This work is organized as follows: section 2 presents the basics of the underlying approach used to remove the perspective effect from a monocular image and its application to the processing of stereo images, section 3 describes how obstacle detection is performed, while section 4 ends the paper with an analysis of the results.

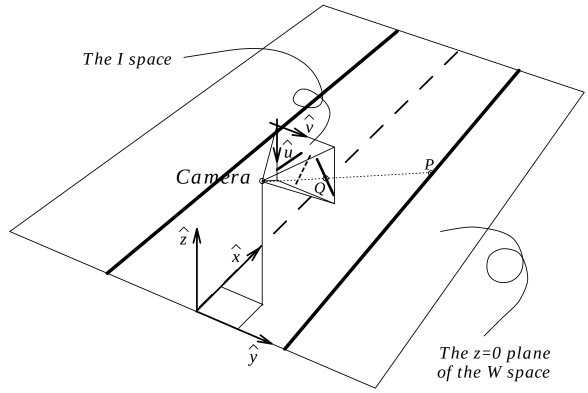


Figure 1: The relationship between the two coordinate systems

2 THE UNDERLYING APPROACH

2.1 The Inverse Perspective Mapping

The *Inverse Perspective Mapping* (IPM) is a well-established technique [2, 8, 10, 3] that allows to remove the perspective effect when the acquisition parameters (camera position, orientation, optics,...) are completely known and when a knowledge about road is given, such as a *flat road hypothesis*. The procedure aimed to remove the perspective effect resamples the incoming image, remapping each pixel toward a different position and producing a new 2-dimensional array of pixels. The resulting image represents a top view of the road region in front of the vehicle, as it were observed from a significant height.

Two Euclidean spaces are defined:

- $\mathcal{W} = \{(x, y, z)\} \in E^3$ representing the $3D$ world space (*world-coordinate system*), where the real world is defined;
- $\mathcal{I} = \{(u, v)\} \in E^2$ representing the $2D$ image space (*screen-coordinate system*), where the 3D scene is

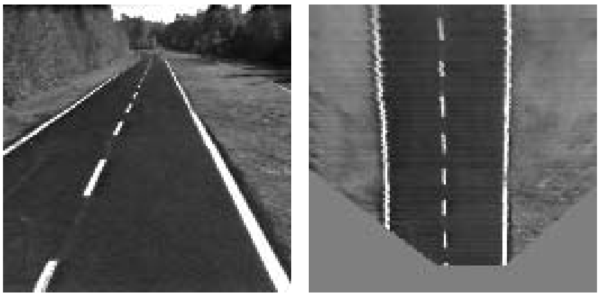


Figure 2: The original and the remapped images

projected.

The image acquired by the camera belongs to the $\mathcal{I}$ space, while the remapped image is defined as the $z = 0$ plane of the $\mathcal{W}$ space (according to the assumption of a flat road). The remapping process projects the acquired image onto the $z = 0$ plane of the 3D world space $\mathcal{W}$, acting as the dual of a *ray-tracing* algorithm [7]. Fig. 1 shows the relationships between the two spaces $\mathcal{W}$ and $\mathcal{I}$. In order to generate a 2D view of a 3D scene, the following parameters must be known:

- *viewpoint*: the camera position is $C = (l, d, h) \in \mathcal{W}$;
- *viewing direction*: the optical axis $\hat{o}$ is determined by the following angles:
  - $\bar{\gamma}$: the angle formed by the projection (defined by versor $\hat{\eta}$) of the optical axis $\hat{o}$ on the plane $z = 0$ and the x axis;
  - $\bar{\theta}$: the angle formed by the optical axis $\hat{o}$ and versor $\hat{\eta}$;
- *aperture*: the camera angular aperture is 2α;
- *resolution*: the camera resolution is $n \times n$.

After simple algebraic and trigonometric manipulations [2], the inverse transform $g : \mathcal{W} \to \mathcal{I}$ (the dual mapping) is given as follows

$$\left\{ \begin{array}{rcl} u(x,y,0) & = & \dfrac{\theta(x,y,0) - \left(\bar{\theta} - \alpha\right)}{\dfrac{2\alpha}{n-1}} \\ v(x,y,0) & = & \dfrac{\gamma(x,y,0) - \left(\bar{\gamma} - \alpha\right)}{\dfrac{2\alpha}{n-1}} \end{array} \right. \quad (1)$$

where

$$\left\{ \begin{array}{rcl} \gamma(x,y,0) & = & \operatorname{arctg}\left[\dfrac{y-d}{x-l}\right] \\ \theta(x,y,0) & = & \operatorname{arctg}\left[\dfrac{h \sin\gamma(x,y,0)}{y-d}\right] \end{array} \right. . \quad (2)$$

The image acquired by a camera and its remapping are shown in figure 2 The information contained in a monocular *remapped image* has already been successfully used to determine the road markings [3] even in critical shadow conditions; on the other hand, the remapping of *stereo pairs* can deliver information about the presence of elevated obstacles or, in general, of a *non-flat* portion of the road.

2.2 Stereo Inverse Perspective Mapping

Assuming a flat road and no obstacles in front of the vehicle, the IPM algorithm can be used to produce an image representing the road as seen from the top. Using the IPM algorithm with appropriate parameters on stereo images, two different patches of the road surface can be obtained. Moreover the knowledge of the parameters of the whole vision system allows to bring the two road patches to correspondence. This means that, under the flat road hypothesis, pairs of pixels having the same image coordinates in the two remapped images represent the same point in the road plane.

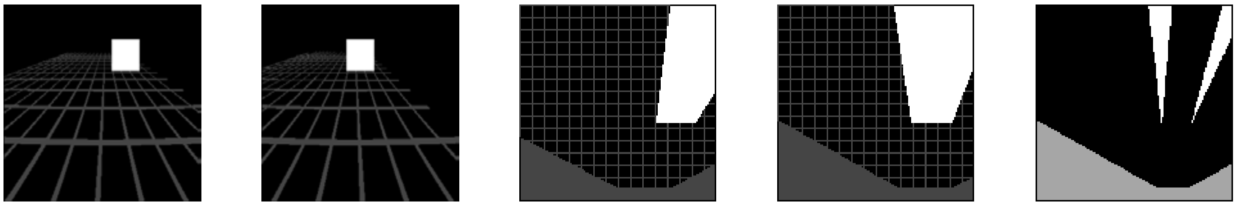
3 OBSTACLE DETECTION

As seen in the previous section any difference in the two reorganized images represents a deviation from the assumption of *flat* road, and thus identifies an obstacle. The flat road hypothesis can be verified computing the difference between the two remapped images: a *generic obstacle* (anything raising out from the road) is detected if the *difference image* presents sufficiently large clusters of non-zero pixels having a specific shape.

In fact, in the difference image obstacles are generally represented by two triangles, generated by the two different angles of view of the stereo system [5], as shown in figure 3. The obstacle detection process is based on the localization of pairs of these triangles. Unfortunately due to the texture, the irregular shape, and the non-homogeneous brightness of real obstacles, the detection and localization of the triangles become more difficult. Nevertheless in the difference image some clusters of pixels with a quasi-triangular shape are anyway recognizable, even if they are not clearly disjointed. Moreover, in case two or more obstacles are present in the scene at the same time, more than two triangles appear in the difference image. A further problem is due to partially visible obstacles which produce a single triangle.

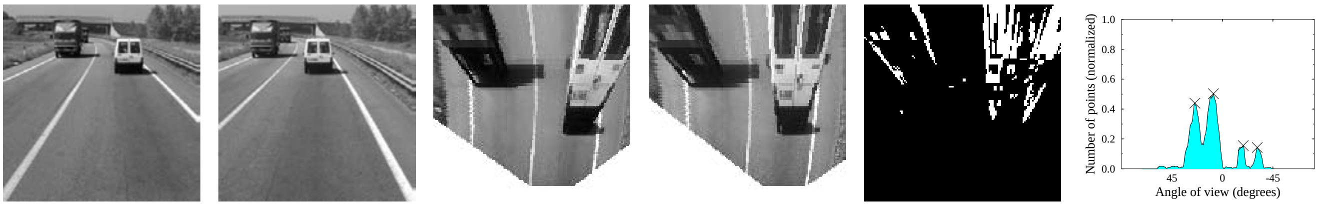
A *polar histogram* is used for the detection of triangles: given a point F in the $z = 0$ plane of the $\mathcal{W}$ space (called *focus*) the polar histogram is computed scanning the difference image and counting the number of overthreshold pixels for every straight line originating from the focus F. Calling $C^{(L)}_{xy}$ and $C^{(R)}_{xy}$ the projections of the two cameras onto the road plane, when F is placed in the middle of $\overline{C^{(L)}_{xy} C^{(R)}_{xy}}$, the polar histogram presents an appreciable peak corresponding to each triangle. Since the presence of an obstacle produces two disjoint triangles in the difference image, obstacle detection is reduced to the search for pairs of adjacent peaks. The use of shape, amplitude, and closeness of peaks allows to join the ones belonging to the same obstacle, thus obtaining the angle of view for each obstacle. Moreover the use of both the difference image and the positions of peaks allows to estimate the obstacles distance.

The whole process is shown in figure 4



(a) (b) (c) (d) (e)

Figure 3: An homogeneous ideal object: a) left view; b) right view; c) remapped left view; d) remapped right view; e) difference between (c) and (d) showing in light grey the area not seen by both cameras



(a) (b) (c) (d) (e) (f)

Figure 4: Obstacle detection in case of two vehicles with different shape and color in front of the stereo system: (a) left image; (b) right image; (c) left remapped image; (d) right remapped image; (e) difference image; (f) polar histogram showing four maxima corresponding to two objects.

4 IMPLEMENTATION

Due to the specific field of application, the response time of the system is a major critical point, since it affects directly the maximum speed allowed for the vehicle. For this reason the PAPRICA system (PArallel PRocessor for Image Checking and Analysis) [4] has been used for low-level processing. PAPRICA is a low-cost special-purpose massively parallel architecture composed of 256 PEs working in SIMD fashion, developed in cooperation with Politecnico di Torino, Italy. The scheme of the whole system is represented in figure 5.

The image remapping is performed in hardware by the PAPRICA system (a 128×128 pixel image is reorganized in less than 3 ms), which is also capable to acquire pairs of images coming from two different cameras at video rate (25 full frames/s or 50 image fields/s as in the current implementation).

The successive processing, aimed to determine the disparities between the two remapped images, is still performed in parallel by the PAPRICA system: it is based on a pixelwise difference followed by a threshold and a simple morphological opening [6] aimed to the removal of small-sized details in the thresholded image. The current implementation on PAPRICA system runs in about 25 ms.

Conversely the polar histogram and the detection of its maxima are computed by a traditional serial system and the results are displayed on an on-board monitor and, in a more condensed format, also on a led-based control panel to generate a visual feedback to the driver. Since the last computation is data-dependent only an estimate of the average processing time required by the host computer can be made, ranging from 20 to 30 ms.

5 CONCLUSION

In this work a technique for obstacle detection has been presented. The main innovative contribution of this work is the use of the IPM technique to simplify both low and medium level processing steps. The whole system was tested with images acquired by the stereo vision system installed on-board of the experimental vehicle MOB-LAB (the experimental MOBile LABoratory developed by the Italian research units within the PROMETHEUS project). It was driven for more than 3000 km, along extra-urban roads and freeways, under different traffic and illumination conditions, at speeds up to 80 km/h. During the tests the system demonstrated to be reliable and robust with respect to noise caused by shadows, different road textures, and vehicle movements. Figure 6 shows some results obtained in a number of different situations (with none, one, or two obstacles with different shapes and colors); they are displayed with black marker superimposed on a brighter version of the left image; the marker's position encodes both the distance and width of obstacles.

References

[1] M. Bertozzi, A. Broggi, and S. Castelluccio. A Real-Time Oriented System for Vehicle Detection. *Journal of Systems Architecture*, 1996. In press.

[2] A. Broggi. A Massively Parallel Approach to Real-Time Vision-Based Road Markings Detection. In I. Masaky, editor, *Procs. IEEE Intelligent Vehicles'95*, pages 84–89, Detroit, 1995.

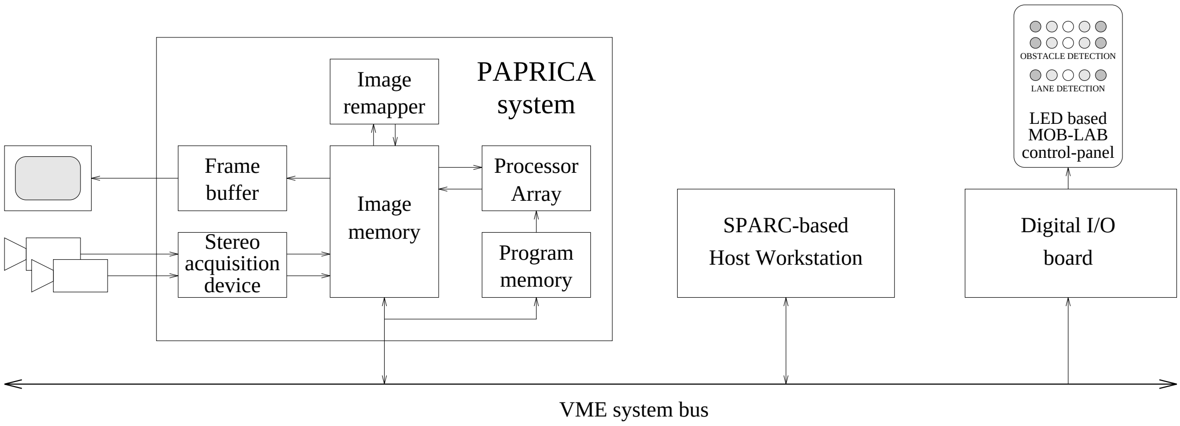


Figure 5: The logical organization of the system

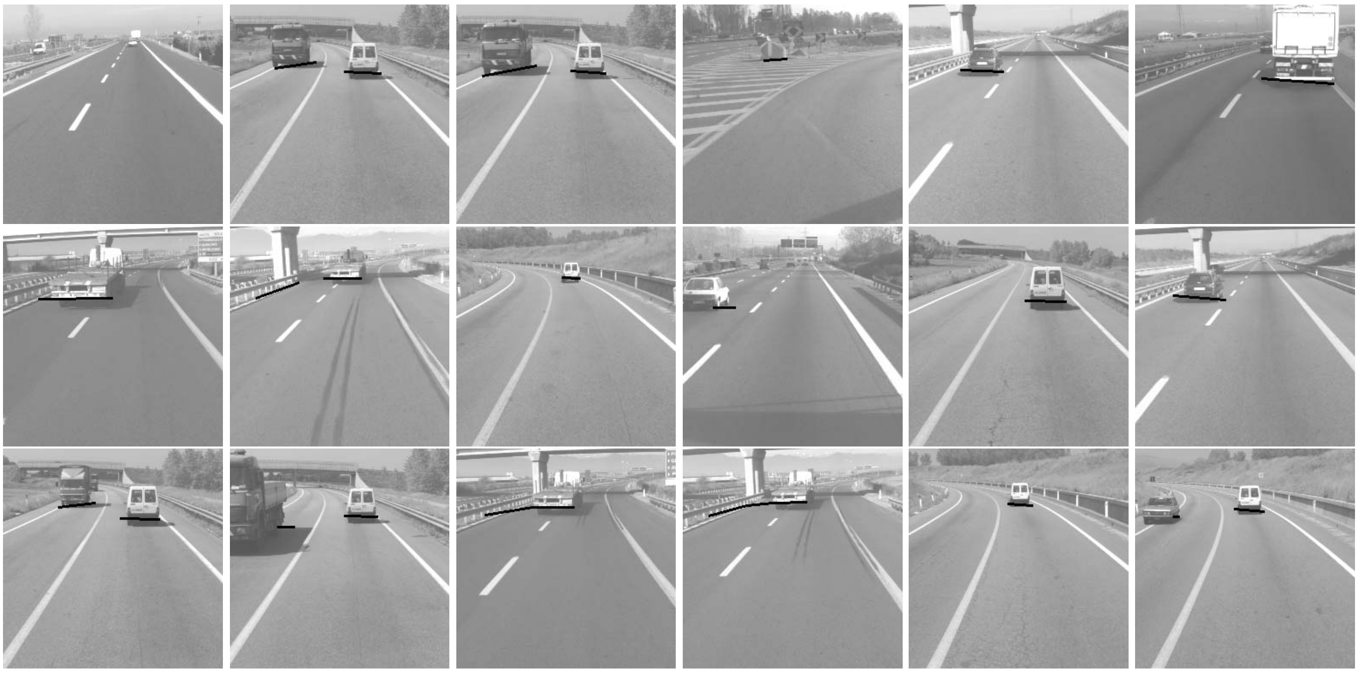


Figure 6: Results of obstacle detection in different road conditions

[3] A. Broggi. Robust Real-Time Lane and Road Detection in Critical Shadow Conditions. In *Procs. IEEE Intl. Symposium on Computer Vision*, Coral Gables, Florida, 1995.

[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*, 1995. Accepted for publication.

[5] A. Fascioli. Localizzazione di ostacoli mediante elaborazione di immagini stereoscopiche. Master's thesis, Università degli Studi di Parma, 1995.

[6] R. M. Haralick, S. R. Sternberg, and X. Zhuang. Image Analysis Using Mathematical Morphology. *IEEE Trans. on PAMI*, 9(4):532–550, 1987.

[7] W. M. Newman and R. F. Sproull. *Principles of Interactive Computer Graphics*. McGraw-Hill,1981.

[8] D. Pomerleau. RALPH: Rapidly Adapting Lateral Position Handler. In I. Masaky, editor, *Procs. IEEE Intelligent Vehicles'95*, pages 506–511, Detroit, 1995.

[9] B. Ross. A practical stereo vision system. In *Proceedings Intl. Conf. on Computer Vision and Pattern Recognition*, pages 148–153, Seattle, 1993.

[10] Y. Ruichek and J.-G. Postaire. Real-Time Neural Vision for Obstacle Detection Using Linear Cameras. In I. Masaky, editor, *Procs. IEEE Intelligent Vehicles'95*, pages 524–529, Detroit, 1995.

[11] F. Thomanek, E. D. Dickmanns, and D. Dickmanns. Multiple object recognition and scene interpretation for autonomous road vehicle guidance. In *Procs. IEEE Intelligent Vehicles'94*, pages 231–236, Paris, 1994.

[12] T. Zielke, M. Brauckmann, and W. von Seelen. Intensity and Edge-based Symmetry Detection with an Application to Car-Following. *CVGIP: Image Understanding*, 58:177–190, 1993.