Robust Real-Time Lane and Road Detection in Critical Shadow Conditions*

Alberto Broggi
Dipartimento di Ingegneria dell'Informazione
Università di Parma
Parma, ITALY, I-43100

Abstract

This paper presents the vision-based road detection system currently installed onto the MOB-LAB land vehicle. Based on a geometrical transform and on a fast morphological processing, the system is capable to detect road markings even in extremely severe shadow conditions on flat and structured roads. The use of a special-purpose massively architecture (PAPRICA) allows to achieve a processing rate of about 17 Hz.

1 Introduction

Many different vision-based road detection systems have been developed worldwide, each of them relying on different characteristics such as different road models (2D or 3D), acquisition devices (color or monochrome camera), hardware systems (special- or general-purpose, serial or parallel), and computational techniques (template matching, neural networks, mono or stereo vision,...).

The SCARF system [7] (tested on the NAVLAB vehicle at Carnegie Mellon University) uses two color cameras for a color-based image segmentation; the different regions are classified and a Hough-like transform is used to vote for different binary road models. Although the image resolution is reduced from 480×512 to 60×64 pixel, a high performing system (a 10 cells Warp [9]) has been chosen to speed-up the processing. SCARF, capable of detecting even unstructured roads in slow varying illumination conditions, reaches a processing rate of $\frac{1}{3}$ Hz [8].

Also the VITS system (tested on the ALV vehicle, Martin Marietta) relies on two color cameras. It uses a combination of the red and blue color bands to reduce the artifacts caused by shadows. Information on vehicle motion are also used to aid the segmentation process. Tested successfully on straight, single lane roads, it runs faster than SCARF, sacrificing general capability for speed [24].

ALVINN [21] (tested on NAVLAB, CMU) is a neural network based 30×32 video retina designed to detect unstructured roads like SCARF, but it does not have any road model: it learns associations between visual patterns and steering wheel angles, without reasoning about the road location. It has been implemented on the Warp system as well, reaching a processing rate of about 10 Hz [22].

A different neural approach has been developed at CMU and tested on NAVLAB, too: a 256×256 color image is segmented on a 16k processors MasPar MP-2 [19]. A trapezoid road model is used together with the assumption of a constant width road. A reduced version (128×128) runs at a rate of 2.5 Hz [14].

Due to the high amount of data (2 color images) and to the complex operations involved (clustering, Hough transform, non-linear neural functions) these systems have been implemented on extremely powerful hardware engines. Anyway, a lot of different methods have been considered for the speed-up of the processing.

As an example, in VaMoRs (Universität der Bundeswehr, München) monochrome images are processed by custom hardware, focusing on the regions of interest only [11]. These windowing techniques are supported by strong road and vehicle models[1] to predict features in incoming images [10]. The use of a single monochrome camera together with these simple road models allows a fast processing, but unfortunately this approach is not successful in shadow conditions or when road imperfections are found [18].

Also the LANELOK system (General Motors) [15] relies on strong models: it estimates the location of lane boundaries with a curve fitting method [17]. Unfortunately the technique used to correct the shadow artifacts [16] relies on fixed brightness thresholds which is far from being a robust and general approach.

Conversely this paper presents a *low-cost* system capable of reaching *real-time* performances in the detection of *structured* roads (with painted lane markings), and robust enough to tolerate critical *shadow* conditions. The limitation to the analysis of structured environments allows to use simple road models, as well as the processing of monocular monochrome images on special-purpose hardware allows to reach high performances at a low cost. The system has been tested on the MOB-LAB land vehicle [1], taking advantage of the critical analysis of previous approaches [3, 4].

The following section motivates the approach used, while Sect. 3 details its theoretical basis; Sect. 4 and 5

*This work was partially supported by Italian CNR under the framework of the Eureka PROMETHEUS Project.

[1] In this case, the vehicle was driven at high speeds (up to 100 kph) on German highways, which have constant width lanes, and where the road has specific shapes: straight, constant curvature, or clothoid.

0-8186-7190-4/95 $4.00 © 1995 IEEE

present the low- and medium-level processing for road marking detection; Sect. 6 analyzes the results and discusses the current evolution; finally Sect. 7 ends the papers with some concluding remarks.

2 The Underlying Approach

Due to its intrinsic nature, low-level image processing is efficiently performed on SIMD systems by means of a massively parallel computational paradigm. Anyway, this approach is meaningful in the case of generic filterings (such as noise reduction, edge enhancement,...), which consider the image as a mere collection of pixels, independently of their semantic content.

On the other hand, the implementation of more sophisticated filters requires some semantic knowledge. As an example, let us consider the specific problem of road markings detection. Due to the perspective effect induced by the acquisition conditions, the road markings width changes according to their distance from the camera. Thus, the correct detection of road markings should be based on matchings with different sized patterns, according to the specific position within the image. Unfortunately this differentiated low-level processing cannot be efficiently performed on SIMD massively parallel systems, which by definition perform the *same* elaboration on each pixel of the image.

In fact, the perspective effect associates different meanings to the different image pixels, depending on their position in the image. Conversely, after the removal of the perspective effect, each pixel represents the same portion of the road[2], allowing a homogeneous distribution of the information among all the image pixels; now the size and shape of the matching template can be independent of the pixel position.

To remove the perspective effect it is necessary to know the specific acquisition conditions (camera position, orientation, optics,...) and the scene represented in the image (the road, which is now assumed to be *flat*), which constitutes the a-priori knowledge. The processing can be conveniently divided into two steps: the first, exploiting the a-priori knowledge, is a transform (a non-uniform resampling similar to what happens in the human visual system [25, 26]), that generates an image in a new domain where the detection of the features of interest is extremely simplified; the second, exploiting the sensorial data, consists of a mere low-level morphological processing. In this way it is possible:

- to detect the road markings through an extremely simple and fast morphological processing;
- to overcome completely the annoying problems caused by a non uniform illumination (shadows);
- to implement efficiently the detection step on massively parallel SIMD architectures, in order to obtain real-time performances.

3 Removing the Perspective Effect

The procedure aimed to remove the perspective effect reads the incoming image and resamples it, 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 the top.

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 projected.

The image acquired by the camera belongs to the $\mathcal{I}$ space, while the reorganized image is defined as the $z = 0$ plane of the $\mathcal{W}$ space (according to the assumption of a flat road). Fig. 1 shows the relationships between the two spaces $\mathcal{W}$ and $\mathcal{I}$.

The reorganization 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 [20].

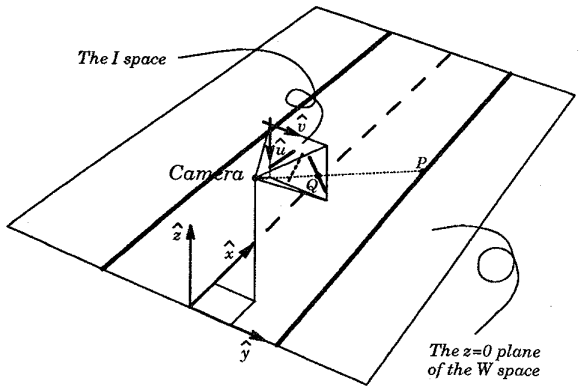

Figure 1: The relationship between the two coordinate systems

3.1 Mapping the $\mathcal{W}$ space to the $\mathcal{I}$ space

In order to generate a 2D view of a 3D scene, the following parameters must be specified [20]:

- *viewpoint*: the camera is placed in $C = (l, d, h) \in \mathcal{W}$;
- *viewing direction*: the optical axis $\hat{o}$ is determined by the following angles:
  - $\overline{\gamma}$: 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 (as shown in fig. 2.a);
  - $\overline{\theta}$: the angle formed by the optical axis $\hat{o}$ and versor $\hat{\eta}$ (as shown in fig. 2.b);
- *aperture*: the camera angolar aperture is 2α;
- *resolution*: the camera resolution is $n \times n$ pixels.

[2] A pixel in the lower part of the original images of fig. 5 represents a few cm^2 of the road, while a pixel in the middle of the same images represents a few tens of cm^2, or even more.

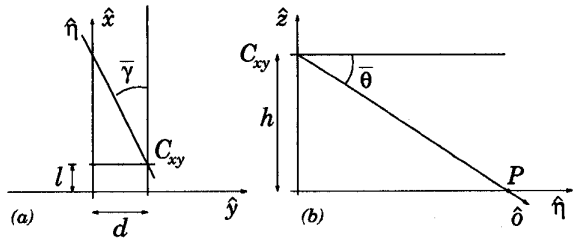

Figure 2: *(a)* the xy plane in the $\mathcal{W}$ space and *(b)* the $z\eta$ plane, assuming the Origin translated onto the projection C_{xy} of C on $z = 0$.

After simple algebraic and trigonometric manipulations [2], the final mapping $f : \mathcal{I} \rightarrow \mathcal{W}$ as a function of u and v is given by:

$$\begin{cases} x(u,v) &= \dfrac{h}{\mathrm{tg}\left[(\overline{\theta}-\alpha)+u\,\frac{2\alpha}{n-1}\right]} \times \\ & \quad \times \cos\left[(\overline{\gamma}-\alpha)+v\,\frac{2\alpha}{n-1}\right]+l \\ y(u,v) &= \dfrac{h}{\mathrm{tg}\left[(\overline{\theta}-\alpha)+u\,\frac{2\alpha}{n-1}\right]} \times \\ & \quad \times \sin\left[(\overline{\gamma}-\alpha)+v\,\frac{2\alpha}{n-1}\right]+d \\ z &= 0 \end{cases} \tag{1}$$

with $u, v = 0, 1, ..., n-1$.

Given the coordinates (u, v) of a generic point Q in the $\mathcal{I}$ space, equations (1) return the coordinates $(x, y, 0)$ of the corresponding point P in the $\mathcal{W}$ space (see fig. 1). As an example, fig. 3.a shows a synthetic computer-generated texture representing the $z = 0$ plane of the $\mathcal{W}$ space. Fig. 3.b presents the result of the application of the *ray-tracing*-like algorithm defined by equations (1), using the set of real parameters [23] of the camera installed on MOB-LAB and $n = 512$. This transform causes an information loss, higher in the central region of fig. 3.b (corresponding to the top of fig. 3.a) and lower in the peripheral region of the same image (corresponding to the bottom of fig. 3.a).

3.1.1 Mapping the $\mathcal{I}$ space to the $\mathcal{W}$ space

The inverse transform $g : \mathcal{W} \rightarrow \mathcal{I}$ (the dual mapping) is given as follows [2]:

$$\begin{cases} u(x,y,0) &= \dfrac{\gamma(x,y,0)-\left(\overline{\gamma}-\alpha\right)}{\frac{2\alpha}{n-1}} \\ v(x,y,0) &= \dfrac{\theta(x,y,0)-\left(\overline{\theta}-\alpha\right)}{\frac{2\alpha}{n-1}} \end{cases} \tag{2}$$

The reorganization process defined by equations (2) removes the perspective effect and recovers the texture of the $z = 0$ plane of the $\mathcal{W}$ space. It consists of scanning the array of pixels of coordinates $(x, y, 0) \in \mathcal{W}$ which form the reorganized image, in order to associate to each of them the corresponding value assumed by the point of coordinates $\Big(u(x,y,0), v(x,y,0)\Big) \in \mathcal{I}$.

Fig. 3.c shows the application of the reorganization process (defined by equations (2)) applied to fig. 3.b. In this case the resolution chosen ($n = 128$) is determined as a good trade-off between information loss and processing time. Note that, as shown in fig. 3.c, the lower portion of the reorganized image is undefined: this is due both to the specific camera position with respect to the z axis, and to the camera angolar aperture.

As an example, fig. 3.e shows the effect of the reorganization procedure applied on the original frame shown in fig. 3.d: it is clearly visible that in this case the road markings width is almost invariant within the whole image.

4 The Low-Level Processing

4.1 Identifying road markings

The assumptions used in the definition of a *"Road Marking"* are the following: a road marking in the $z = 0$ plane of the $\mathcal{W}$ space (i.e. in the reorganized image) is represented by a quasi-vertical bright line of constant width surrounded by a dark region (the road). Thus the pixels belonging to a road marking have a brightness value higher than their left and right neighbors. The detection is thus reduced to the determination of horizontal black-white-black transitions.

For each image line $x = 0, 1, ..., n-1$, every pixel $P = (x, y, 0)$ compares its brightness value $b(x, y, 0)$ with its left and right ones at distance m: $b(x, y-m, 0)$ and $b(x, y+m, 0)$, with $m \geq 1$. A new pixel value $r(x, y, 0)$ is computed according to the following rule:

$$r(x,y,0) = \begin{cases} \Big(b(x,y,0)-b(x,y-m,0)\Big)+ \\ \quad \Big(b(x,y,0)-b(x,y+m,0)\Big), \\ \quad \text{if}\Big(b(x,y,0) \geq b(x,y-m,0)\Big)\,\& \\ \quad \&\Big(b(x,y,0) \geq b(x,y+m,0)\Big) \\ 0 \quad \text{otherwise} \end{cases} \tag{3}$$

According to equation (3),

$$r(x,y,0) \neq 0 \implies \begin{cases} r(x,y-m,0) = 0 \\ r(x,y+m,0) = 0 \end{cases} . \tag{4}$$

The choice of m depends on the road markings width, which is assumed to be in a known range.

Considering $m = 2$, the resulting image[3] is shown in fig. 4.a.

[3]The rule described by equation (3) can be expressed in a more compact and simple way thanks to grey-tone mathematical morphology notations [12]: the new pixel value $r(x, y, 0)$ is given by:

$$r(x,y,0) = \begin{cases} d(x,y,0)-b(x,y,0) & \text{, if } d(x,y,0) \geq b(x,y,0) \\ 0 & \text{, otherwise} \end{cases}$$

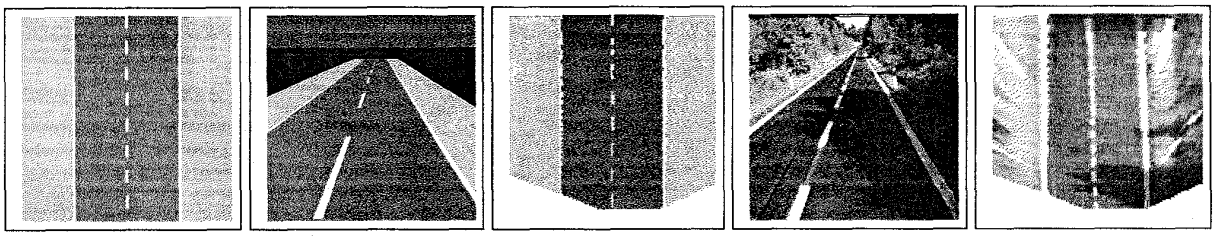

Figure 3: *a*) The model image; *b*) the projection of the previous model on the $\mathcal{I}$ space; *c*) the result of the reorganization procedure applied on the previous image; *d*) the original image; *e*) the reorganized image.

4.2 Image enhancement and binarization

Due to different illumination conditions (e.g. in presence of shadows), the road markings may have different brightness, yet maintaining their superiority relationship with their horizontal neighbors. Thus a simple threshold seldom gives a satisfactory binarization and consequently an image enhancement is required, as well as an adaptive binarization. Exploiting the property expressed by equation (4), the left and right neighbors of a road marking line assume a zero value in the filtered image. Thus the execution of a few iterations (say h) of the following rule[4]

$$e^{(i+1)}(x,y,0) = \begin{cases} \max\Big(e^{(i)}(x+1,y,0), e^{(i)}(x-1,y,0), \\ \quad e^{(i)}(x,y+1,0), e^{(i)}(x,y-1,0)\Big), \\ \quad \text{if } e^{(i)}(x,y,0) \neq 0 \\ 0 \quad \text{otherwise} \end{cases} \tag{5}$$

starting from $e^{(0)}(x,y,0) = r(x,y,0)$ generates an enhanced image, as shown in fig. 4.b (with $h = 8$).

The binarization is performed by means of an adaptive threshold:

$$t(x,y,0) = \begin{cases} 1 & \text{if } e^{(h)}(x,y,0) \geq \dfrac{m(x,y,0)}{k} \\ 0 & \text{otherwise} \end{cases}, \tag{6}$$

where $m(x,y,0)$ is the maximum value computed in a given $c \times c$ neighborhood, and k is a constant. The result of the binarization of fig. 4.b, considering $k = 2$ and $c = 11$, is presented in fig. 4.c.

Fig. 4.d shows the representation in the $\mathcal{I}$ space of the binarized result of fig. 4.c, while fig. 4.e presents its final superimposition onto the original image. The high quantization visible in the lower region of these images is the effect of the choice of a medium resolution (128×128) for the reorganized image.

where $d(x,y,0)$ is defined as the grey-tone dilation [12] of the original brightness image by the following binary structuring element:  .

[4]Note that equation (5) can be efficiently expressed as a dilation with the following binary structuring element  contextualized (see [13]) to the state of the central pixel.

5 The Medium-Level Processing

The medium-level processing is reduced to the determination of the best concatenations of the pixels representing road markings. This is done by means of a serial scanning of the image: first an histogram is computed in the lower region of the binary image, and a threshold is applied; then, starting from the over-threshold positions, a neighborhood search is performed, keeping track of the search direction. Where no black neighbors are found (namely a gap in the road boundary is met) the search is continued in the previously determined direction, according to the *best continuation* Gestalt principle. The result is then compared to the position of road markings in normal conditions (vehicle in the center of the right lane) and the difference is then used to warn the driver in dangerous situations.

Due to the high effectiveness of the low-level processing and to the high correlation between two subsequent frames in a sequence (thanks to the fast processing), the medium-level step is extremely fast; it is performed in pipeline by a sequential architecture during the low-level processing of the following frame. This system is currently integrated on the MOB-LAB land vehicle, and has been proven to be effective in a number of different road conditions running at about 50 kph on very narrow rural roads.

6 Performance Analysis and Current Evolution

The discussed algorithm has been integrated on PAPRICA [5, 6] massively parallel architecture featuring an hardware extension for the efficient removal of the perspective effect (based on a look-up table). Table 1 presents the performance of the current implementation: the complete acquisition and processing of a single frame takes less than 55 ms, thus allowing the processing of about 17 frames per second.

Beside being easily implementable on any SIMD massively parallel processor based on a morphological computational paradigm, the presented approach is extremely robust with respect to the noise caused by sunny blobs on shaded roads. Fig. 5 presents a few results of the processing of images acquired under different conditions[5].

[5]A couple of sequences in MPEG format (200 and 1000 frames respectively) are available in `http://WWW.CE.UniPR.IT/computer_vision/applications.html`, showing lane detection

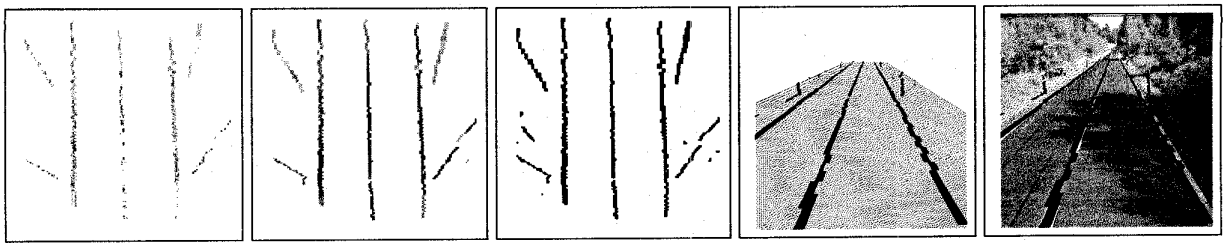

Figure 4: *a*) The filtered image; *b*) the enhanced image; *c*) the binarized image; *d*) the projection of the previous result on the $\mathcal{I}$ space; *e*) the superimposition of the previous image onto the original one.

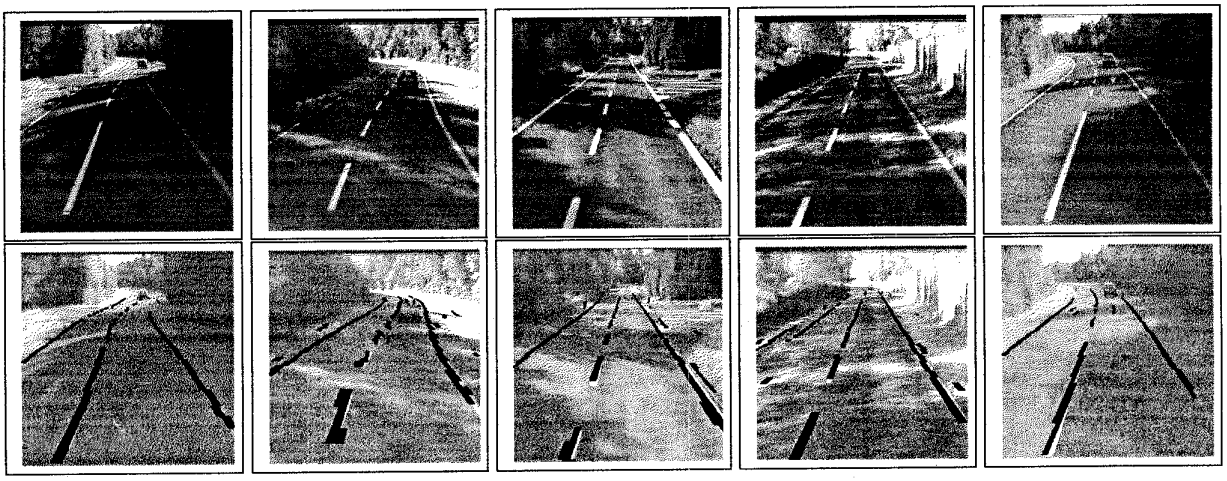

Figure 5: Top: original images acquired from MOB-LAB vehicle; Bottom: the superimposition of the results of road markings detection onto brighter versions of the original images (images taken during the demonstration of the MOB-LAB vehicle at the final meeting of the PROMETHEUS project, Mortefontaine track, Paris, Oct. 94).

Operation	Time
Image acquisition (512 × 256)	20 ms
Perspective effect removal (128 × 128)	2.8 ms
Low-level processing (128 × 128)	30 ms
Medium-level processing (128 × 128)	in pipeline
Warnings to the driver	negligible

Table 1: Timings on PAPRICA system

Unfortunately this approach relies on the implicit assumption of *visible* road markings, namely when no obstacles are on the path. The worst case is represented by a white vehicle on the right lane, since it induces black-white-black transitions in the reorganized image that are erroneously identified as road markings. This problem can be solved by the application of the mentioned approach to a couple of stereo images (see fig. 6). The independent detection of the road markings on both images would be disturbed by the presence of vehicles. Conversely, the fusion of the two results (obtained with a simple logical intersection, thanks to the calibration of the two image reorganization processes) improves and makes more reliable the detection (see fig. 6.g.) In this case the problems of stereo vision are then easily solved by the low-level portion of the processing, thus reaching a high computational efficiency.

Moreover, since on a flat road the two reorganized images are identical, as soon as an obstacle is approached the two reorganized images differ. The difference image can be used for obstacle detection (see fig. 6.i).

7 Conclusion

This paper presented an approach to real-time road detection, working on flat roads with painted road markings. It has been demonstrated to be robust with respect to extremely critical shadow conditions and global illumination changes; an extension to obstacle detection is now under evaluation.

It has been implemented on the special-purpose and low-cost massively parallel system PAPRICA and integrated onto the MOB-LAB land vehicle, reaching a processing rate of about 17 Hz.

in particularly challenging conditions. It is possible to note that the vehicle's pitch does not disturb the processing even if the reorganized image represents a very large road area in front of the vehicle (up to 40 meters).

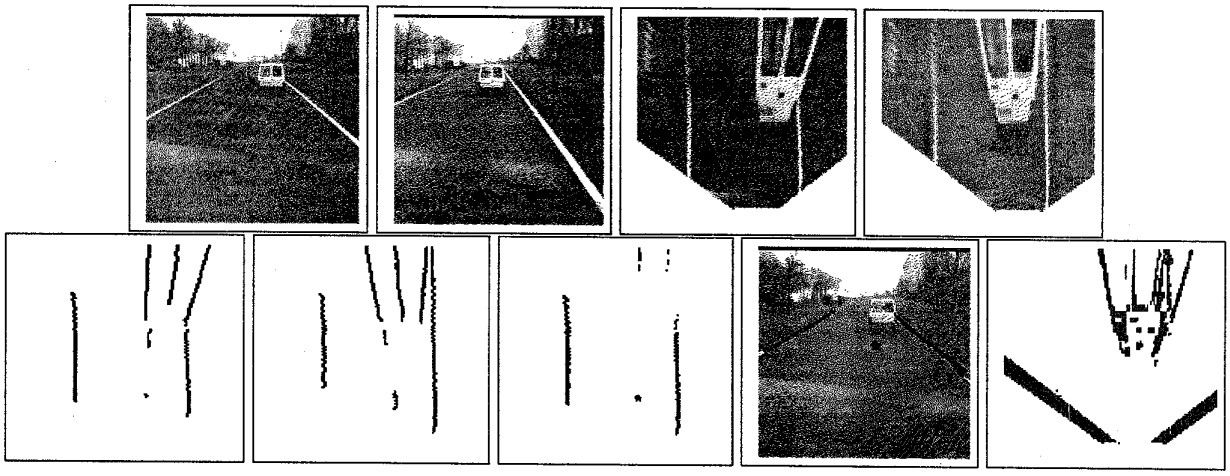

Figure 6: *a*) left image; *b*) right image; *c*) left reorganized image; *d*) right reorganized image; *e*) left binarized image; *f*) right binarized image; *g*) logical intersection between *e* and *f*; *h*) superimposition of the result onto the original image; *i*) difference between images *c* and *d*, used for obstacle detection.

References

[1] G. Adorni, A. Broggi, V. D'Andrea, and G. Conte. Real-Time Image Processing for Automotive Applications. In P. A. Laplante and A. D. Stoyenko, eds., *Real-Time Image Processing: Theory, Techniques and Applications*. IEEE Press, 1996. In press.

[2] A. Broggi. An Image Reorganization Procedure for Automotive Road Following Systems. In *Proc. 2nd IEEE Intl. Conf. on Image Processing*, 1995. In press.

[3] A. Broggi. Parallel and Local Feature Extraction: a Real-Time Approach to Road Boundary Detection. *IEEE Trans. on Image Processing*, 4(2):217–223, February 1995.

[4] A. Broggi and S. Bertè. A Morphological Model-Driven Approach to Real-Time Road Boundary Detection for Vision-Based Automotive Systems. In *Proc. 2nd IEEE Workshop on Applications of Computer Vision*, pages 73–90, 1994.

[5] A. Broggi, G. Conte, F. Gregoretti, C. Sansoè, L. M. Reyneri. The PAPRICA Massively Parallel Processor. In *Proc. IEEE Intl. Conf. on Massively Parallel Computing Systems*, pages 16–30, 1994.

[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*, 1995. In press.

[7] J. Crisman and C. Thorpe. Color Vision for Road Following. In C. E. Thorpe, editor, *Vision and Navigation. The Carnegie Mellon Navlab*, pages 9–24. Kluwer Academic Publishers, 1990.

[8] J. Crisman and C. Thorpe. SCARF: A Color Vision System that Tracks Roads and Intersections. *IEEE Trans. on Robotics and Automation*, 9(1):49–58, February 1993.

[9] J. D. Crisman and J. A. Webb. The Warp Machine on Navlab. *IEEE Trans. on PAMI*, 13(5):451–465, May 1991.

[10] E. D. Dickmans and B. D. Mysliwetz. Recursive 3-D Road and Relative Ego-State Recognition. *IEEE Trans. on PAMI*, 14:199–213, May 1992.

[11] V. Graefe and K.-D. Kuhnert. Vision-based Autonomous Road Vehicles. In I. Masaki, editor, *Vision-based Vehicle Guidance*, pages 1–29. Springer Verlag, 1991.

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

[13] W. D. Hillis. *The Connection Machine*. MIT Press, Cambridge, Ma., 1985.

[14] T. M. Jochem and S. Baluja. A Massively Parallel Road Follower. In M. A. Bayoumi, L. S. Davis, and K. P. Valavanis, editors, *Proc. Computer Architectures for Machine Perception*, pages 2–12, 1993.

[15] S. K. Kenue. LANELOK: An Algorithm for Extending the Lane Sensing Operating Range to 100 Feet. In *Proc. of SPIE - Mobile Robots V*, vol. 1388, pp. 222–233, 1991.

[16] S. K. Kenue. Correction of Shadow Artifacts for Vision-based Vehicle Guidance. In *Proc. of SPIE - Mobile Robots VIII*, vol. 2058, pp. 12–26, 1994.

[17] S. K. Kenue and S. Bajpayee. LANELOK: Robust Line and Curvature Fitting of Lane Boundaries. In *Proc. of SPIE - Mobile Robots VII*, vol. 1831, pp. 491–503, 1993.

[18] K. Kluge and C. E. Thorpe. Explicit Models for Robot Road Following. In C. E. Thorpe, editor, *Vision and Navigation. The Carnegie Mellon Navlab*, pages 25–38. Kluwer Academic Publishers, 1990.

[19] MasPar Computer Corporation, Sunnyvale, California. *MP-1 Family Data-Parallel Computers*, 1990.

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

[21] D. A. Pomerleau. Neural Network Based Autonomous Navigation. In C. E. Thorpe, editor, *Vision and Navigation. The Carnegie Mellon Navlab*, pages 83–93. Kluwer Academic Publishers, 1990.

[22] D. A. Pomerleau. *Neural Network Perception for Mobile Robot Guidance*. Kluwer Academic Publishers, 1993.

[23] R. Tsai. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. In *Proc. IEEE Intl. Conf. on CVPR*, pages 364–374, Miami Beach, FL, 1986.

[24] M. A. Turk, D. G. Morgenthaler, K. D. Gremban, and M. Marra. VITS - A Vision System for Autonomous Land Vehicle Navigation. *IEEE Trans. on PAMI*, 10(3), 1988.

[25] J. M. Wolfe and K. R. Cave. Deploying visual attention: the guided model. In *AI and the eye*, pages 79–103. A.Blake and T.Troscianko, 1990.

[26] B. Zavidovique and P. Fiorini. A Control View to Vision Architectures. In V. Cantoni, editor, *Human and Machine Vision: Analogies and Divergencies*, pages 13–56. Plenum Press, 1994.