# A Massively Parallel Approach to Real-Time Vision-Based Road Markings Detection*

Alberto Broggi
Dipartimento di Ingegneria dell'Informazione
Università di Parma, I-43100 Parma, Italy
Phone: (+39) 521 90 5707, Fax: (+39) 521 90 5723, broggi@CE.UniPR.IT

**Abstract**

*This paper presents the vision-based road detection system currently operative onto the MOB-LAB land vehicle. Based on a full-custom low-cost massively parallel system, it achieves real-time performances ($\simeq$ 17 Hz) in the processing of image sequences, thanks to the extremely efficient implementation of the algorithm. Assuming a flat road and the complete set of acquisition parameters to be known (camera position, orientation, optics,...), the system is capable to detect road markings on structured roads even in extremely severe shadow conditions.*

## 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 [8], capable of detecting even unstructured roads in slow varying illumination conditions, uses *two color cameras* [7] for a color-based image segmentation. Although working at low image resolutions, a high performing system (a 10 cells Warp [9]) has been chosen to face the heavy computational load.

The VITS system [23] uses a combination of the red and blue color bands of *two color images* to reduce the artifacts caused by shadows.

ALVINN [20, 21], implemented on the Warp system as well, is a *neural-based system* designed to navigate in unstructured environments like SCARF, but without any road model.

A different *neural* approach [13] tested on NAVLAB as well, is based on the segmentation of a *color* image on a 16k processors MasPar MP-2 [18].

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

As an example, in VaMoRs monochrome images are processed by custom hardware, focusing on the regions of interest only [11]. The use of a single monochrome camera together with strong road models and windowing techniques [10] allow a fast processing, but unfortunately, as discussed in [17], this approach is not successful in severe shadow conditions or when road imperfections are found.

Also the LANELOK system [14] relies on strong models, estimating the location of lane boundaries with a curve fitting method [16]. Unfortunately, as discussed in [15], the technique used to correct the shadow artifacts relies on fixed brightness thresholds which may be easily affected by noise.

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.

Figure 1: The MOB-LAB land vehicle

The system has been tested on the MOB-LAB land vehicle, shown in the photograph of fig. 1, taking advantage of the critical analysis of previous approaches [2, 3].

*This work was partially supported by the Italian Research Council under the framework of the Eureka PROMETHEUS Project – Progetto Finalizzato Trasporti – under contracts n. 93.01813.PF74 and 94.01371.PF74.

The following Section discusses the underlying approach, while Sect. 3 presents its theoretical basis; Sect. 4 discusses the details of the low-level processing for road marking detection; Sect. 5 and 6 analyze the computing architecture and the results obtained; finally Sect. 7 ends the papers with some concluding remarks.

## 2 The Approach

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 by means of traditional *pattern-matching* techniques should be based on matchings with different sized patterns, according to the specific position within the image (see fig. 2).

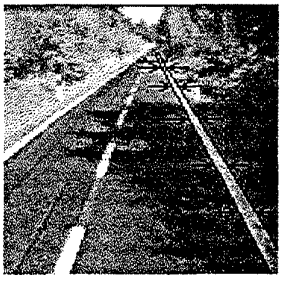


Figure 2: The road markings width changes according to their position within the image

Unfortunately this differentiated low-level processing cannot be efficiently performed on SIMD massively parallel systems, which by definition perform the *same* computation 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 (see fig. 3).

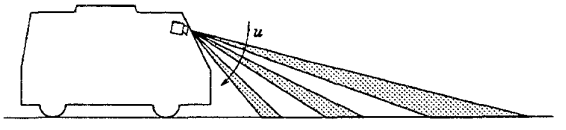



Figure 3: Due to the perspective effect, different pixels represent different portions of the road

Conversely, after the removal of the perspective effect, each pixel represents the same portion of the road, 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*). The processing is thus divided into two steps:

1. the first, exploiting the knowledge on the acquisition process and on the scene represented in the image, is a transform (a non-uniform resampling similar to what happens in the human visual system [24, 25]), that generates an image in a new domain where the detection of the features of interest is extremely simplified;
2. the second, exploiting the sensorial data, consists of a mere low-level morphological processing.

## 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 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 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). 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 [19]. Fig. 4 shows the relationships between the two spaces $\mathcal{W}$ and $\mathcal{I}$.

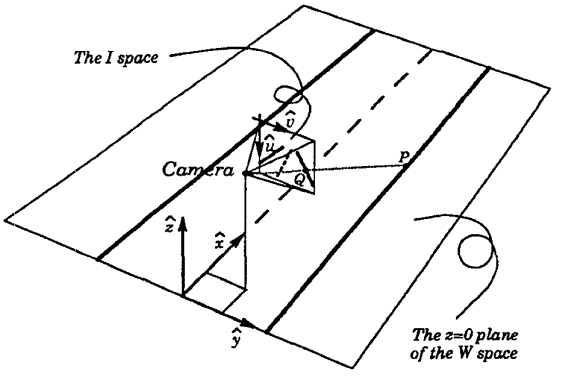



Figure 4: The relationship between the two coordinate systems

### 3.1 $\mathcal{I} \rightarrow \mathcal{W}$ mapping

In order to generate a 2D view of a 3D scene, the following parameters must be specified [19]:

- *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:
  - $\overline{\gamma}$: 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. 5);
  - $\overline{\theta}$: the angle formed by the optical axis $\hat{o}$ and versor $\hat{\eta}$ (as shown in fig. 6);
- *aperture*: the camera angolar aperture is $2\alpha$;
- *resolution*: the camera resolution is $n \times n$.

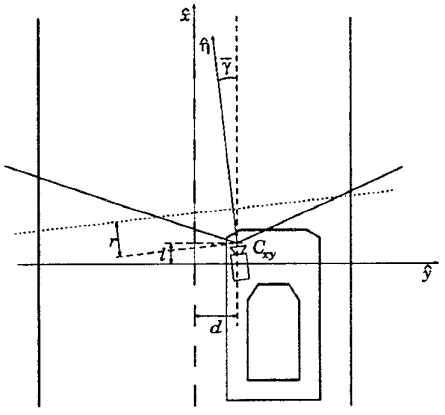



Figure 5: The $xy$ plane in the $\mathcal{W}$ space

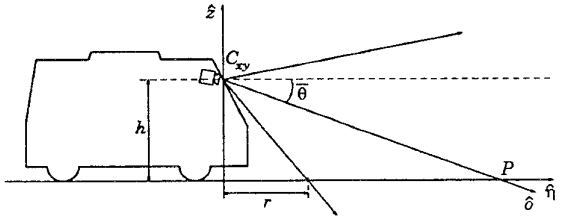



Figure 6: The $z\eta$ plane, assuming the Origin translated onto the projection $C_{xy}$ of $C$ on $z = 0$

After simple algebraic and trigonometric manipulations [1], the final mapping $f : \mathcal{I} \rightarrow \mathcal{W}$ as a function of $u$ and $v$ is given by:

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

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. 4).

### 3.2 $\mathcal{W} \rightarrow \mathcal{I}$ mapping

The inverse transform $g : \mathcal{W} \rightarrow \mathcal{I}$ (the dual mapping) is given as follows [1]:

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

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 $\Big(u(x,y,0), v(x,y,0)\Big) \in \mathcal{I}$.

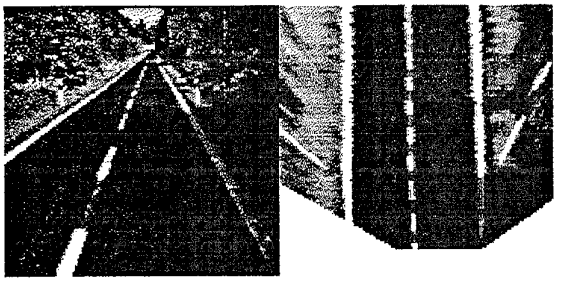


Figure 7: The original and the reorganized images

As an example, fig. 7 shows the original and reorganized images: it is clearly visible that in this case the road markings width is almost invariant within the whole image. The resolution of the reorganized image ($n = 128$) has been chosen as a trade-off between information loss and processing time. Note that 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 shown in fig. 8.

Fig. 9 shows the specific calibration of the camera installed onto MOB-LAB.

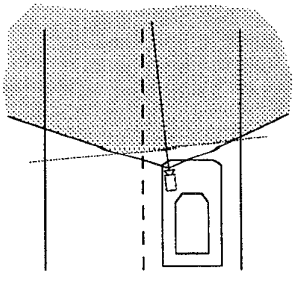


Figure 8: The visible portion of the road

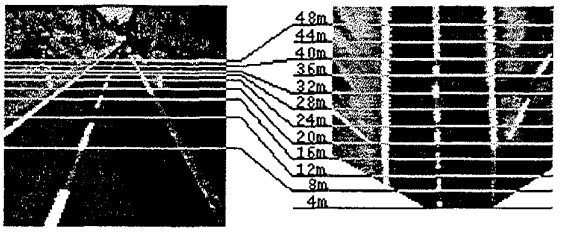



Figure 9: Calibration of the reorganized image

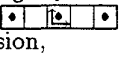
# 4 Morphological Processing

## 4.1 Road markings identification

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 darker 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 neighbors at distance $m$: $b(x, y-m, 0)$ and $b(x, y+m, 0)$, with $m \geq 1$. According to grey-tone mathematical morphology notations [12], the new pixel value $r(x, y, 0)$ is given by:

$$r(x,y,0) = \begin{cases} d(x,y,0) - b(x,y,0) \\ \quad \text{if } d(x,y,0) \geq b(x,y,0) \\ 0 \quad \text{otherwise} \end{cases} \qquad (3)$$

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

According to the previous expression,

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

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

Considering $m = 2$, the filtered image is shown in fig. 10.a.

## 4.2 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; 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 enhancement of the filtered image is performed through a few iterations (say $h$) of a *geodesic dilation* [22] with the following binary structuring element  , where

$$c(x,y,0) = \begin{cases} 1 \ , & \text{if } r(x,y,0) \neq 0 \\ 0 \ , & \text{otherwise} \end{cases} \qquad (5)$$

is the *control image* [22]. The enhanced image (with $h = 8$) is shown in fig. 10.b.

The binarization is performed by means of an adaptive threshold:

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

where $e(x, y, 0)$ represents the enhanced image, $m(x, y, 0)$ the maximum value computed in a given $c \times c$ neighborhood, and $k$ is a constant. The result of the binarization of fig. 10.b, considering $k = 2$ and $c = 11$, is presented in fig. 10.c.

Fig. 10.d shows the representation in the $\mathcal{I}$ space of the binarized result of fig. 10.c, while fig. 10.e presents its final superimposition onto the original image.

# 5 The Computing Architecture

The discussed approach has been implemented on PAPRICA [4, 5] massively parallel architecture. PAPRICA is a *low-cost special purpose* coprocessor containing 256 single-bit Processing Elements (PEs) disposed on the nodes of a 2D grid, with full 8-neighbors connectivity. The PAPRICA prototype currently installed onto MOB-LAB is composed of an array of $4 \times 4$ full custom ICs (1.5 $\mu$m CMOS, 45 mm$^2$, $\simeq$ 35000 transistors), each of them containing a sub-array of $4 \times 4$ PEs. Each PE has an internal memory composed of 64 bits; data flow between the image memory and the PA through a 16-bit data bus. In the current implementation, the PAPRICA coprocessor is integrated on a single VME board (6U) connected to a SPARC-based workstation. It comprises 5 major functional parts:

1. the Program Memory, storing up to 256k instructions;
2. the Image Memory (8 MBytes of static RAM), whose 16-bit memory fetch time is 150 ns;
3. the Processor Array, with a 250 ns cycle time;
4. the Frame Grabber device, able to grab and store into PAPRICA Image Memory $512 \times 512$ 8 bit/pixel grey-tone images at 25 Hz;
5. the Control Unit, managing the activities of the whole system.

Table 1 presents the performance of the current implementation of the algorithm on PAPRICA massively parallel architecture: the complete acquisition and processing of a single frame takes less than 55 ms, thus allowing the processing of about 17 frames per second.

Figure 10: *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.

| Operation | Time |
|---|---|
| Image acquisition ($512 \times 256$) | 20 ms |
| Perspective effect removal ($128^2$) | 2.8 ms |
| Low-level processing ($128^2$) | 30 ms |
| Medium-level processing ($128^2$) | in pipeline |
| Warnings to the driver | negligible |

Table 1: Timings on PAPRICA system

PAPRICA system includes a hardware extension explicitly designed for an efficient removal of the perspective effect through a look-up table [6], while it is currently under evaluation the possibility to reorganize the image optically, with an ad-hoc shaped lens.

The medium-level processing, not discussed here, reconstructs the road markings, and compares them to standard positions, with the aim of warning 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, 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 narrow rural roads. Fig. 11 presents a few results of the processing of images acquired under different conditions[1].

## 6 Critical Analysis

The assumptions on which the discussed approach is based are (i) flat road and (ii) visible road markings.

In the first case, if the road has irregular bumps, the road markings in the reorganized image present a convergent or divergent behavior (see figure 12). In this case the shape of the road can be recovered only if the assumption of a fixed-width lane is assumed. Moreover, from extensive experimental tests, it has been noted 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 50 meters).

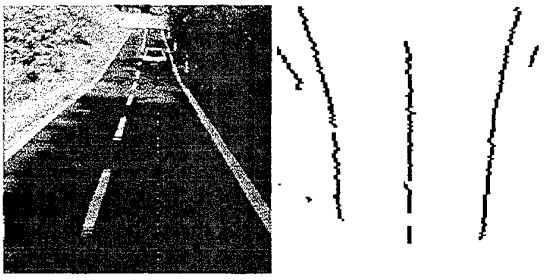


Figure 12: Road markings detection in case of a non-flat road

Conversely, to overcome the problem of invisible road markings, namely in presence of obstacles (other vehicles) on the path, it is currently under evaluation and test the application of the reorganization process to a couple of stereo images. Moreover, since on a flat scene 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.

## 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.

The discussed approach allows:

- 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.

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.

[1] A couple of sequences in MPEG format are available in `http://WWW.CE.UniPR.IT/ computer_vision/ applications.html`, showing lane detection in particularly challenging conditions.

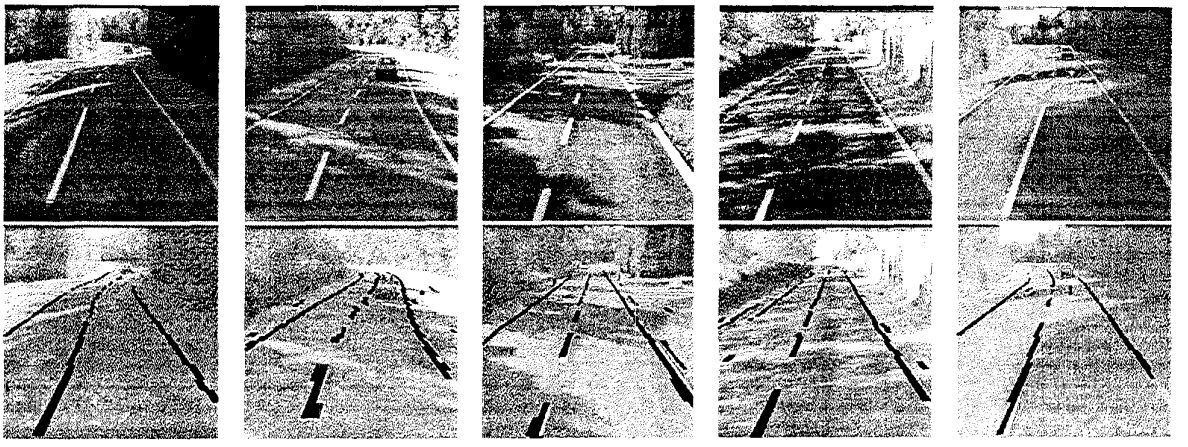


Figure 11: Top: original images acquired from MOB-LAB vehicle; Bottom: the superimposition of the results of road markings detection onto brighter versions of the original images (images taken during the demonstration of the MOB-LAB vehicle at the final meeting of the PROMETHEUS project, Mortefontaine track, Paris, October 1994).

## References


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

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

[3] 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, Sarasota, FL, December, 5-7 1994.

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

[5] A. Broggi, G. Conte, F. Gregoretti, C. Sansoè, and L. M. Reyneri. The Evolution of the PAPRICA System. *Integrated Computer-Aided Engineering Journal*, 1995. Accepted for publication.

[6] A. Broggi, F. Gregoretti, and C. Sansoè. A Smart Sensor for the Real-Time Detection of Road Markings through the Removal of the Perspective Effect. In *Proc. Intl Symposium on Automotive Technology and Automation*, Stuttgart, Sept. 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, 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, 1991.

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

[13] 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.

[14] S. K. Kenue. LANELOK: An Algorithm for Extending the Lane Sensing Operating Range to 100 Feet. In *Proc. SPIE* , v. 1388, pp. 222–233, 1991.

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

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

[17] 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.

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

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

[20] 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.

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

[22] J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, London, 1982.

[23] 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), May 1988.

[24] 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.

[25] 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.