Title:
Lane Recognition on Poorly Structured Roads - the Bots Dot Problem in California

Authors:

Stefan K. Gehrig	Axel Gern	Stefan Heinrich	Bernd Woltermann
DaimlerChrysler Research RIC/AP, HPC T 728 70546 Stuttgart Germany	DaimlerChrysler Research RIC/AP, HPC T 728 70546 Stuttgart Germany	DaimlerChrysler Research RIC/AP, HPC T 728 70546 Stuttgart Germany	DaimlerChrysler Research RIC/TP, 1510 Page Mill Rd Palo Alto, CA 94304 USA
Phone: ++49-711-17-41484 Fax: ++49-711-17-47054 Stefan.Gehrig@DCX.Com	++49-711-17-47076 ++49-711-17-47054 Axel.Gern@DCX.Com	++49-731-505-2257 ++49-731-505-4115 Stefan.Heinrich@DCX.Com	++1-650-845-2503 ++1-650-845-2555 Bernd.Woltermann@DCX.Com

Technical Categories:

Image analysis, lane recognition, intelligent vehicle

Contact Author:

Stefan K. Gehrig
Research Institute DaimlerChrysler AG
RIC/AP, HPC T 728
70546 Stuttgart
Germany

Phone: ++49-711-17-41484
Fax: ++49-711-17-47054
Stefan.Gehrig@DCX.Com

Lane Recognition on Poorly Structured Roads - the Bots Dot Problem in California

Stefan K. Gehrig, Axel Gern, Stefan Heinrich, Bernd Woltermann

Abstract—
Lane recognition is the basis for many driver assistance systems, including Lane Departure Warning (LDW), the assignment of vehicles to specific lanes, and fully autonomous driving. A major problem of common vision-based lane recognition systems is their susceptibility to weather and poorly structured roads. Especially when driving in adverse weather conditions such as rain or snow, it is difficult to estimate the road course. The contrast between the white lane markings and the pavement is poor, sometimes the colors of the markings are negated. Furthermore the range of sight is reduced enormously causing a bad prediction of the lane parameters, particularly the curvature. We present a solution which relies not only on finding white markings. In addition we are recognizing reflective lane markers and bots dots. These measurements are then integrated in the lane recognition system estimating the position of the vehicle within the lane and the curvature parameters of the road ahead. The system allows us to perform lane departure warning and to drive laterally controlled autonomously even under adverse weather conditions.

Keywords— **Intelligent vehicle, computer vision, stereo vision**

I. Introduction

In the past, many different vision-based lane recognition systems have been presented. Most of them try to find road features such as lane markings or road surface textures. These are matched against a specific geometrical model of the road (e.g. [1], [2], [3]). These features are used to determine the parameters of the chosen model and the position of the car in the lane, for example using a least-squares fitting or a Kalman filter [2]. Many lane recognition systems encounter problems when driving in adverse weather conditions such as rain or snow. Often the contrast between the markings and the pavement is poor (see Fig. 1). The range of sight is reduced enormously, causing a bad prediction of the lane parameters, particularly the curvature. Another problem are non-standard lane markings such as reflective markers or bots dots.

There are many different solutions to this problem. Instead of only searching locally for white lane markings, [4] and [5] analyze the complete image. ALVINN [4] is based on an artificial neural network, which is trained to learn the characteristic features of particular roads under specific conditions. It uses this learned road model to determine how to steer keeping the vehicle in the lane. Problems emerge when driving in unknown areas the neural network isn't trained on. In RALPH [5] the road in front of the car is converted into a bird's-eye view. This low-resolution image is deformed with several road-curvature hypotheses to determine the curvature. The advantage of this approach is that in advance no model of the road is needed. As can be seen in Fig. 1, the overall structure of the road indicates the road course. Disadvantages of this approach are the lack of estimation of the yaw angle of the own vehicle relative to the lane and the accuracy of the curvature estimation due to only analyzing low resolution images. In LOIS [6] the determination of the road course and the position of the vehicle within the lane is reduced to an optimization problem in a multidimensional parameter space. The contribution made by every pixel to a likelihood function is defined by its intensity gradient magnitude and gradient direction. In this approach, there is no need for the typical gradient magnitude thresholding.

S. Gehrig is with DaimlerChrysler AG, HPC T 728, 70546 Stuttgart, Germany. E-mail: Stefan.Gehrig@DaimlerChrysler.Com

Fig. 1. A typical scene when driving on a highway with bots dots. The contrast between the bots dots and the pavement is poor.

A different solution is to use other sensors. For example other obstacles, seen by a radar sensor, can be taken into account as described in [7], [8] and [9]. This stabilizes the estimation of the yaw angle of the own vehicle relative to the lane and the curvature parameters enormously. On the other hand, DGPS-based map information systems can be used to stabilize the estimation of the road course (see [10], [11], and [12]).

In order to use the poorly visible (but well audible) bots dots for lane recognition, the typical gradient-based feature extraction delivers only results when keeping the gradient threshold very low. However, a low gradient threshold causes a lot of noise.

In this paper we present a new approach fusing two different types of road features. Firstly, we are using the classic geometrical approach, tracking locally the white markings. Secondly, we use a dedicated bots dots algorithm that recognizes these features without gradient thresholding.

How is this paper organized? In Section II the standard lane recognition algorithm is introduced. Section III describes the additional algorithm that is necessary to deal with "poorly structured" roadways such as Californian roadways marked by bots dots. The fusion of these two feature extraction algorithms is presented in Section IV. Section V briefly describes the sensor system used in our demonstrator. Results are detailed in Section VI. Conclusions and future work comprise the final Section.

II. Vision-Based Lane Recognition

According to the recommendations for highway construction, highways are built under the constraint of slowly changing curvatures. Specifically in the U.S., piece-wise constant curvature segments are recommended. Most lane recognition systems are based on a clothoidal lane model, that is given by the following equation:

$$c(L) = c_0 + c_1 \cdot L. \tag{1}$$

$c(L)$ describes the curvature at the length L of the clothoid, c_0 is the initial curvature and c_1 the curvature-rate, which is called the clothoidal parameter. The curvature is defined as $c = \frac{1}{R}$, where R denotes the radius of the curve.

Besides these curvature parameters, lateral position x_{off} and yaw angle $\Delta\psi$ relative to the lane axis are of interest for modeling the vehicle position within the lane.

Assuming the pinhole-camera model and knowing the camera parameters focal length f, tilt angle α and height-over-ground H, the relation between a point on a marking and its image point $P_i(x_i, y_i)$ can be described by the following equations:

$$x_i = \frac{f}{L}(a \cdot w - x_{off} - \Delta\psi \cdot L + \frac{c_0}{2} \cdot L^2 + \frac{c_1}{6} \cdot L^3) \tag{2}$$

$$L = \frac{H}{\alpha + (y_i/f)} \tag{3}$$

w is the lane width and $a = \pm 0.5$ is used for the left or the right marking. Hence, every measurement is projected onto a virtual measurement directly on the centerline of the lane. In all equations, the trigonometrical functions are approximated by the argument ($\sin x = x, \tan x = x$), because we consider only small angles. These equations allow to determine the relevant run of the curve and vehicle position parameters.

The image processing used to detect the lane marking features is basically edge detection. The edges of the lane markings are detected with gradient operators and a certain edge threshold is applied to reject noise.

Driving at higher speeds, dynamic and kinematic restrictions have to be taken into account. These constraints can be expressed by the following differential equations:

$$\dot{x}_{off} = v \cdot \Delta\psi + v_x \tag{4}$$

$$\dot{\Delta\psi} = \dot{\psi}_{veh} - c_0 \cdot v \tag{5}$$

$$\dot{c}_0 = c_1 \cdot v \tag{6}$$

$$\dot{c}_1 = 0 \tag{7}$$

In these equations, v denotes the longitudinal speed of the vehicle, v_x the lateral speed caused by a possible side slip angle and $\dot{\psi}_{veh}$ the yaw rate. v_x and $\dot{\psi}_{veh}$ are measured by inertial sensors.

Based on the dynamic and kinematic model (Eqn. (4) through Eqn. (7)) the road markings are tracked from frame to frame by using Kalman filter techniques as first proposed by [2]. The geometrical equation (2) is used as the measurement equation updating the filter. The search areas are centered at the predicted position in the image. The size of the regions is determined by calculating the 3σ-area of the expected measurement, assuming a Gaussian noise process.

The above described system is independent of the image source, using a monocular or a stereo camera system. Our first approach as e.g. described in [13] uses a monocular camera and relies on the assumption, that the road is flat. Sometimes problems occur because 'markings' are falsely found on cars cutting in or crash barriers. This causes a wrong state estimation.

These problems can be solved using stereo information. Every point on the markings found in one image is correlated against a small region in the second image. This delivers three-dimensional information allowing a vertical modeling.

In our approach, we divide the vertical modeling into two components:
1. A linear part, described by the tilt angle α.
2. A non-linear part, described by the vertical curvature, approximated by a clothoid.

German Highways are designed according to a parabolic vertical curvature c_v. The vertical and horizontal curvature models are separated. The parabola curvature is approximated using a clothoid as described in [14]:

$$c_v(L) = c_{v,0} + c_{v,1} \cdot L \tag{8}$$

Tilt angle and vertical curvature are estimated in one joint Kalman filter using the following measurement equation for the height over ground y:

$$y = H - \alpha \cdot L + \frac{c_{v,0}}{2} \cdot L^2 + \frac{c_{v,1}}{6} \cdot L^3 \tag{9}$$

The change $\Delta\alpha$ of tilt angle α between two cycles caused by ground bumps or acceleration/deceleration is estimated using a model of an oscillating damped system:

$$\ddot{\Delta\alpha} + 2 \cdot D \cdot \omega_0 \cdot \dot{\Delta\alpha} + \omega_0^2 \cdot \Delta\alpha = k_{max} \cdot \omega_0^2 \cdot a \tag{10}$$

The parameters are defined as follows:
- $\Delta\alpha$: change of tilt angle caused by ground bumps or acceleration/deceleration.
- D: spring constant of the oscillating system.
- $\omega_0 = 2 \cdot \pi \cdot f$: angular frequency and f the resonant frequency.
- a: longitudinal acceleration of the vehicle, measured by an inertial sensor.
- k_{max}: gain.

Summarizing, the stereo lane recognition system as described above consists of two independent Kalman filters – one estimating the horizontal curvature, the other estimating the tilt angle and the vertical curvature. They are the central components and basis for the fusion approach.

The calibration of the camera can be performed iteratively. After crudely estimating the initial yaw and pitch angle, these values are observed, and the correct calibration values are obtained by low-pass filtering the estimates over hours of driving time.

One current version used for autonomous driving on German highways runs on a single Pentium II at 400MHz and tracks two markings reliably. Under good weather conditions the system analyzes up to 150 search windows at a range of sight of 50m to 70m. The monocular system runs at a cycle time of about 5.5ms, the binocular takes about 10ms time for every cycle. The system allows to drive comfortable autonomously with a speed up to 160km/h if the markings are well visible.

III. The Bots Dot Recognition Algorithm

The lane recognition system described above works well on reasonably well structured highways. However, along the west coast of the United States, a special kind of lane marking is often used: The bots dots (also commonly referred to as bot dots) - occasionally reflective dot- or square-shaped points that protrude slightly out of the ground to make an audible noise when crossing the lane. In cases, when bots dots are set densely (more than 5 bots dots per meter), the regular lane marking algorithms still receive sufficient edge information to fit a lane. The more common case of sparsely placed bots dots calls for an extra feature extraction method. To avoid noisy measurements, we consciously renounce on gradient-based methods.

Based on a priori knowledge, the proposed method is using the morphological properties of small spot-like road-markings. Starting from an input image, we are searching for bots dot lane markings within regions of interest (ROI's). These ROI's are defined by the prediction of the road model. Together with the minimum and maximum measurement distance and with the variance of the road model parameters, the ROI's cover the lane markings as rectangular regions within the image (see e.g. Figure 7). In order to enhance potential lane-marking pixels, a matched filter is applied within the ROI's. The gray values for an image row within the ROI is shown in Figure 2. To determine whether an image-pixel (41) at position (x0) belongs to a lane marking, its intensity as well as the statistical information of the surrounding area (S1,S2) is taken into account. The width (d) of the matched filter is taken from the lane-marking width of the road-model. Within (S1,S2), the average value (A1,A2) and the standard deviation (σ_1, σ_2) is used to compute a threshold (T1,T2). Comparison of the gray value (40) with the threshold (42) generated from (T1,T2) is leading to a binarized image, with enhanced lane marking pixels as shown in Figure 3. An additional stage now checks for morphological properties such as shape, closeness, holes, and compactness. Only

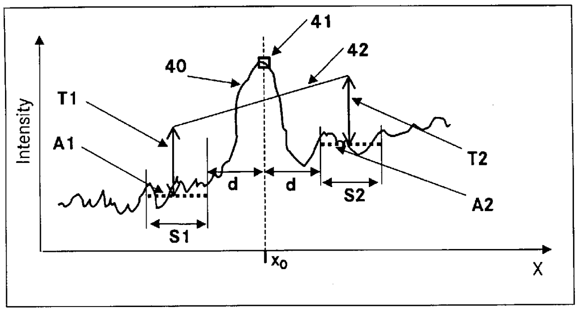

Fig. 2. Grayvalue intensity of an image row.

Fig. 3. Filter result of the bots dot detection run over the whole image. For our algorithm, we only search within our lane marking windows

blobs, which qualify as "reflective marker" will be used for further computation. The image coordinates from the remaining bots dot pixels will be clustered and fed into the tracking Kalman filter if they fit the road model within a 3-σ environment. The proposed bots dot algorithm is not restricted to reflective marker recognition. It is also possible to detect lane marking lines even if they are degraded from wearout and aging.

As explained above, one can see that the non-linear thresholding step in this algorithm is different from a simple gradient thresholding. After searching within a lane marking window, the candidate bots dot pixels are clustered and used for measurement in the following way:

- Begin looking for bots dot candidates at the predicted lane marking position, which is obtained via Kalman filter.
- When more than a minimum number of adjacent bots dot pixels are found, mark their innermost position as a potential bots dot measurement.

The above described algorithm is only run, when the classic lane recognition algorithm fails to find enough lane markings. This way, we can prevent to pick up noise on well-painted roads and are still able to obtain an acceptable lane estimate on poorly structured bots dot roads. This feature extraction method runs on a P-III PC operating on

Fig. 4. Installed camera system in our research vehicle

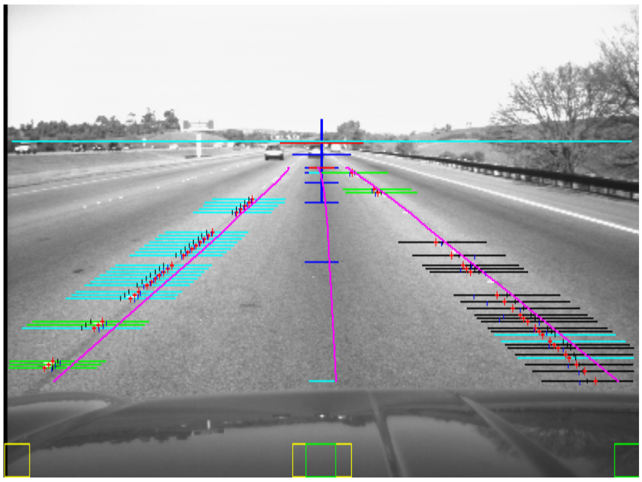

Fig. 5. California freeway scene on Interstate I 280. Bots dot algorithm active. The black windows constitute bots dot measurements.

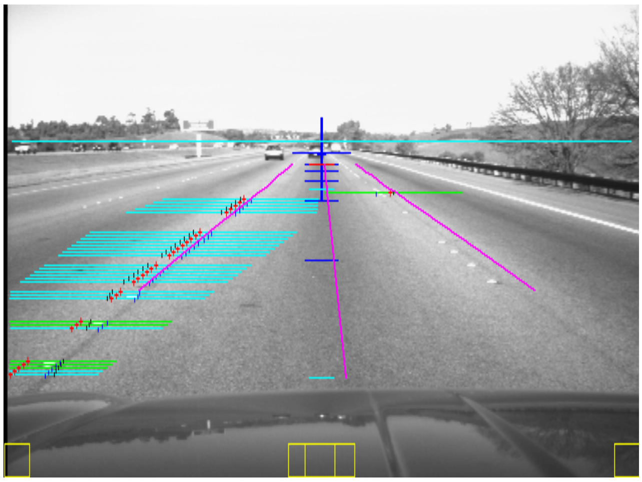

Fig. 6. California freeway scene on Interstate I 280. Standard lane recognition only.

a monocular image in well under 20ms computation time.

IV. Fusion

The central component of our approach is a combination of the different feature extraction methods that provide measurements to the Kalman filter, based on the dynamical model described by Eqn. (4) through Eqn. (7). Each sensor, such as the lane recognition system, the bots dot detection system, or a GPS-based trajectory system [12] provides additional measurements for a robust and reliable estimation of the road course and/or the position of the own vehicle within the lane. The basic idea of this multi-sensor fusion approach is to overcome the sensitivity of all single sensor systems with respect to unpredictable measurement disturbances and sensor failures, e.g. caused by weather conditions. Note that our fusion is not limited to different sensors but also includes different feature extraction methods using the same sensor.

V. The Sensor System

We equipped a Jeep Liberty with our lane recognition system (see Figure 4 for the camera setup).

Specifically, the cameras we use deliver grayscale half-images with a resolution of 768 by 284 pixels at a $25Hz$ rate. These images are processed using an off-the-shelf Pentium III PC with 1000MHz. Using such a system, it is straightforward to process stereo images for lane recognition in real-time.

VI. Results

A. Algorithmic Results

To demonstrate the necessity of a bots dot recognition algorithm, we compared the performance of the lane recognition system with and without bots dot algorithm. For well-structured roads, no difference was perceived. However, when lanes were only marked with bots dots, the availability of the system dropped significantly.

To get an impression of the system performance, the search windows with their results and the estimated lane is shown in Figures 5 and 6. In Figure 5, the bots dot algorithm was active which yielded a good estimate of the lane geometry. Without the bots dot algorithm, not enough lane markings were found to even initialize the system (see Figure 6).

What makes the bots dot recognition so powerful compared to the edge-based approach? Given the non-linear nature of the feature extraction, lower thresholds to find features can be applied. The particular strong feature of the bots dot algorithm is the recognition rate of bots dots close to the car (up to 20m). Hence the estimation of the offset w.r.t. the lane center can be robustly estimated. However, projecting the Kalman filter result further away from the car (50m in our example), a deviation from the estimated road geometry to the real geometry becomes visible (see Figure 7).

On the other hand, running the same scene without bot dot recognition yields no usable result due to the lack of measurements. The algorithm takes dozens of frames to

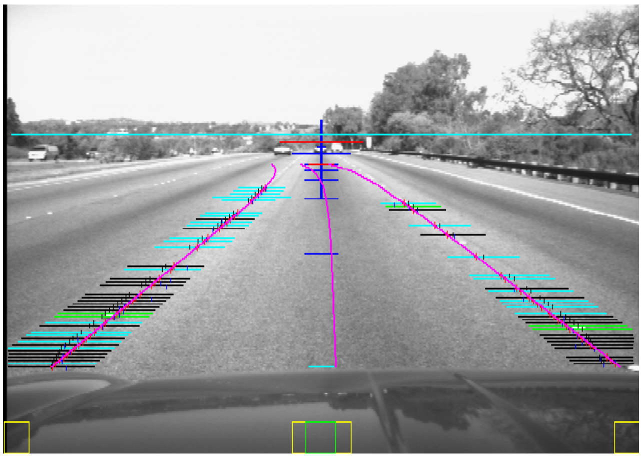

Fig. 7. California freeway scene on Interstate I 280. Bots dot algorithm active. The offset estimate is accurate.

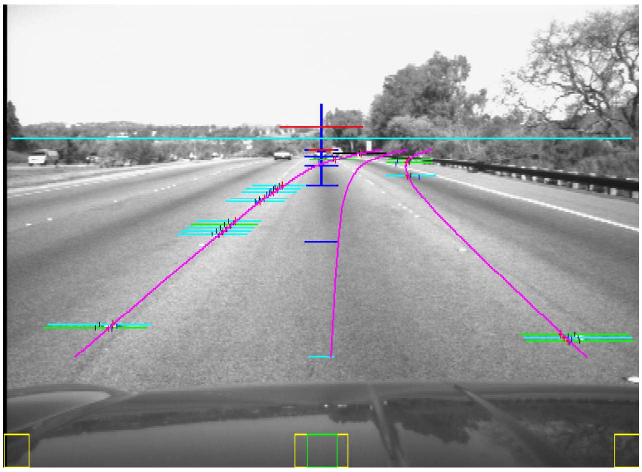

Fig. 8. California freeway scene on Interstate I 280. Standard lane recognition only. Hardly any measurements are found with the edge-based feature extraction

initialize rather arbitrary on random edges and performs re-initialization on a regular basis (see Figure 8).

Running both feature extraction algorithms on our P-III computer yielded a computation of well under $40ms$ for every cycle consisting of image acquisition, feature extraction, tracking, and visualization.

B. System Performance Results

In order to determine the precision of lane recognition results, a comparison to ground truth is necessary. Visual inspection of hundreds of scenes off-line and driving dozens of hours checking the system online yielded the following qualitative result: Robust and precise estimation of the offset and yaw angle even with very poorly structured roads has been achieved. However, road curvature is harder to estimate due to the lack of measurements further away.

Another very important quality criterion of a lane recognition system is the availability of the lane information. With the bots dot recognition algorithm, about 90% of the driving time, lane information is available. Without this algorithm, availability drops down to about half the time. Basis for these statistics were dozens of hours driving time along the west coast on freeways marked with bots dots and poorly marked rural roads.

VII. Conclusions and Future Work

In this paper a lane recognition system for American roadways was introduced. The standard edge-based lane marking approach fails for roads that are only marked by square or circular reflective lane markings, so-called bots dots. With an augmentation of the standard approach this performance gap could be filled and a robust recognition performance on poorly structured roadways could be achieved. As the next step, information about the vehicle state, velocity, and heading direction, will be used to better predict the estimated lane geometry. This information will be obtained from the standard vehicle sensors. Due to the lack of a yaw rate sensor in our demonstrator, we will use the heading information instead.

References

[1] K. Kluge and C. Thorpe, “Representation and recovery of road geometry in YARF,” in *Proc. IEEE Conference on Intelligent Vehicles*, 1992, pp. 114–119.

[2] E.D. Dickmanns and A. Zapp, “A curvature-based scheme for improved road vehicle guidance by computer vision,” in *Proc. SPIE Conference on Mobile Robots*, 1986, pp. 161–168.

[3] Jinyou Zhang and Hans-Hellmut Nagel, “Texture-Based Segmentation of Road Images,” in *Proc. IEEE Conference on Intelligent Vehicles*, 1994, pp. 260–265.

[4] Dean Pomerleau, “Progress in Neural Network-based Vision for Autonomous Robot Driving,” in *Proc. IEEE Conference on Intelligent Vehicles*, 1992, pp. 391–396.

[5] Dean Pomerleau, “RALPH: Rapidly Adapting Lateral Position Handler,” in *Proc. IEEE Conference on Intelligent Vehicles*, 1995, pp. 506–511.

[6] Karl Kluge and Sridhar Lakshmanan, “A Deformable-Template Approach to Lane Detection,” in *Proc. IEEE Conference on Intelligent Vehicles*, 1995.

[7] Z.Zomotor and U.Franke, “Sensor fusion for improved vision based lane recognition and object tracking with range-finders,” in *Proc. IEEE Conference on Intelligent Transportation Systems*, 1997.

[8] A. Gern, U. Franke, and P. Levi, “Advanced lane recognition - fusing vision and radar,” in *Proceedings of the Intelligent Vehicles 2000 Symposium*, 2000, pp. 55–60.

[9] Frank Dellaert, Dean Pomerleau, and Chuck Thorpe, “Model-Based Car Tracking Integrated with a Road-Follower,” in *International Conference on Robotics and Automation*, 1998.

[10] Dirk Langer and Bala Kumar, “Integrating Radar and Carrier Phase GPS for Classifying Roadway Obstacles,” in *IEEE Conference on Intelligent Transportation Systems*, 1997.

[11] C. K. H. Wilson, S. Rogers, and S. Weisenburger, “The Potential of Precision Maps in Intelligent Vehicles,” in *Proc. IEEE Conference on Intelligent Vehicles*, 1998.

[12] Axel Gern, Thomas Gern, Uwe Franke, and Gabi Breuel, “Robust Lane Recognition Using Vision and DGPS Road Course Information,” in *Proc. IEEE Conference on Intelligent Vehicles*, 2001.

[13] U. Franke, “Real time 3D-road modeling for autonomous vehicle guidance,” in *Selected Papers of the 7th Scandinavian Conference on Image Anaysis*. 1992, pp. 277–284, World Scientific Publishing Company.

[14] B. Mysliwetz, *Parallelrechner-basierte Bildfolgen-Interpretation zur autonomen Fahrzeugsteuerung*, Ph.D. thesis, Universitaet der Bundeswehr Muenchen, Fakultaet fuer Luft- und Raumfahrttechnik, 1990.