Robust Vehicle Tracking Fusing Radar and Vision

Axel Gern, Uwe Franke

DaimlerChrysler Research, T728
D-70546 Stuttgart, Germany

{Axel.Gern, Uwe.Franke}@DaimlerChrysler.com

Paul Levi

Institute of Parallel and
Distributed High-Performance Systems
University of Stuttgart, Germany

Paul.Levi@informatik.uni-stuttgart.de

Abstract

Vehicle detection and tracking is the base for many driver assistance systems including Adaptive Cruise Control (ACC), collision warning and fully autonomous driving. A large detection range is required, especially while driving at higher speeds on highways. A reliable and precise detection is needed even under adverse weather conditions. In this paper we present a fusion approach combining radar and monocular image processing. That enables us to track vehicles up to a distance of 130 m and to assign them reliably to specific lanes.

1 Introduction

In the past, many different vehicle detection systems have been presented. Systems using only one sensor often lack reliability and robustness in specific situations. Monocular vision-based vehicle detection [1] has to deal with problems of localizing vehicles due to the missing depth information in 2d images. Stereo information, as e.g. described in [2], allows reliable depth measurements up to a disparity of 5 pixels. This corresponds in our setup to a distance of 50 to 60 m. A sensor providing data above this range is the radar sensor, which is commercially available in upper class cars for ACC applications. In addition, radar shows a higher robustness to weather compared to vision. Problems occur if the precise lateral localization of the obstacles is needed, e.g. for assigning vehicles to specific lanes. These unprecise lateral measurements stem from the fact that cars don't have defined reflection points.

This paper presents our two-step fusion approach which is based on [3] and was first suggested in [4]. In a first step the radar-based obstacle detection is fused with a monocular localization, which allows us to measure precisely the center of the car. In the second step the vehicle detection is combined with our lane recognition system for improved tracking of vehicles as well as improved lane recognition. The system allows the reliable assignment of vehicles to specific lanes even under adverse weather conditions.

The following sections briefly introduce the Adaptive Cruise Control and our lane recognition system, which is the basis for the lateral obstacle tracking. Section 4 sketches our fusion approach and the tracking of the vehicles using lane information. Section 5 shows results from test drives under adverse weather conditions.

2 Adaptive Cruise Control

In May 1999, DaimlerChrysler introduced the radar-based ACC, called DISTRONIC, in the Mercedes S-class. The system detects and tracks cars up to a distance of 150 m. As already mentioned, in contrast to vision systems, radar is robust with respect to bad weather conditions.

The system measures the following three parameters for every radar obstacle i (see Fig. 1):

1. The distance d_{obj_i}.
2. The relative speed v_{rel,obj_i}, taking advantage of the Doppler effect.
3. The angle φ_{obj_i}.

In contrast to the distance d_{obj_i} and the relative speed v_{rel,obj_i} which are measured precisely, the angle to the vehicle in front is noisy. This stems from the fact that cars don't have defined reflection points such as corner reflectors. Therefore the radar beam is 'sliding' from the left to the right edge and vice versa. Most of the time it detects edges and sometimes it generates phantom obstacles caused by reflections by crash barriers or other cars.

The crucial step for an intelligent cruise control is the assignment of the tracked obstacles to certain lanes. This requires knowledge of the road course. Without an independent lane recognition, provided for example by a vision system, the road shape has to be estimated indirectly. This can be realized using tracked obstacles, the steering wheel angle, the yaw rate and the speed using an appropriate vehicle model.

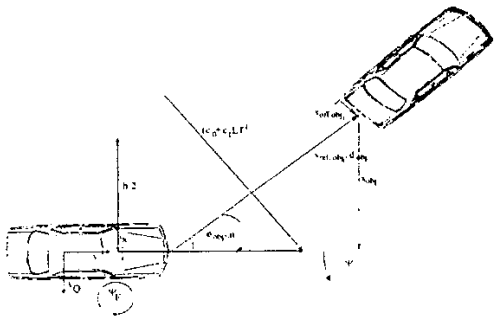

Figure 1: The relevant parameters for the lane recognition system and radar.

Estimating the oncoming lane geometry from these vehicle parameters only is difficult. For example, it is hard to distinguish between the two situations 'the vehicle in front is going into a bend' and 'the vehicle in front is changing lanes'.

3 Vision-Based Lane Recognition

In the past, many different vision-based lane recognition systems have been presented, e.g. [5], [6], [7], [8] and [9]. Similar to the common vision-based lane recognition systems, we try to find road features such as white lane markings which are matched against a specific geometrical model of the road [10]. Hereby, the parameters of the chosen model and the position of the car in the lane are determined using a Kalman filter [5].

According to the recommendations for highway construction, highways are built under the constraint of slowly changing curvatures. Therefore, 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 $c = R^{-1}$, 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 autonomous driving (see Fig. 1).

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:

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

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

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 linearized, because we consider only small angles. These equations allow to determine the relevant road course ahead and the vehicle position parameters.

Our current version 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 50 m to 70 m. The monocular system runs at a cycle time of about 5.5 ms, the binocular one takes about 10 ms time for every cycle. The system allows autonomous driving comfortably at a speed of up to 160 km/h, if the markings are well visible. A spinoff of this approach is marketed as a product in the Lane Departure Warning System (LDW) for heavy trucks since June 2000. A more detailed description of our system can be found in [4].

4 Fusion Approach

The combination of radar and vision is described e.g. in [11], [12] and [13]. [11] suggests to use the information of the radar sensor for segmentation of the image in relevant and non-relevant areas. The fusion of the LOIS lane detection system [9] and radar combined with vision-based obstacle tracking is presented in [12]. It uses the radar for a primary detection of the vehicles defining the search area in the image. A template matching algorithm is used to localize the vehicles in the image. Areas where obstacles are detected are locked for finding lanes and vice versa. Lane and obstacle detection are fused using a combined-likelihood function. [13] uses RALPH [8] to classify the degree of danger that an obstacle poses to the ego vehicle. The obstacle detection is performed by radar. [14] combines RALPH with monocular obstacle tracking to improve the estimation of the road course as well as the position of the vehicle within the lane.

Our fusion approach consists of two steps (see Fig. 2). In the first step radar and vision are fused to detect vehicles (see Sec. 4.1). On the second stage, already detected vehicles are tracked over time using the lane recognition system and different tracking schemes (see Sec. 4.2). Detected obstacles are verified classifying the back-sides of the vehicles using a neural network classifier with receptive fields

as described in [15]. The occlusion module locks areas where an obstacle is already detected. The system allows to precisely localized vehicles up to a distance of 130 m and an improved estimation of the road course as already shown in [3] and [4].

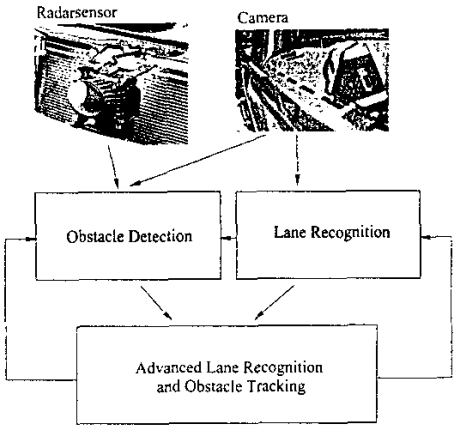

Figure 2: Fusion of radar and vision on two stages.

4.1 Vehicle Detection

The architecture of our vehicle detection module is shown in Fig. 3. A raw lateral position and a search area of the obstacles is given by radar. The search area in the image is defined by the upper left position $P_s(u_s, v_s)$, the width w and the height h of the search box. Different operators ξ_i vote for the center of an obstacle, mapping the two-dimensional search area on a one-dimensional vector:

$$\xi_i : \mathcal{R}^{b \cdot h} \longmapsto \mathcal{R}^b \tag{4}$$

The position of the vehicle is determined by the decision module, combining the different one-dimensional vectors. It can involve a neural network similar to that proposed by [11] for segmentation or other classificators. At the moment our decision module is a simple heuristic, which is satisfying for most situations.

The operators ξ_i involve:

- A symmetry operator (see e.g. [16] or [1]).
- A so-called 'claw'-operator, searching for a pair of vertical edges, tested at different hypothetical car widths.
- An operator searching for horizontal edges.

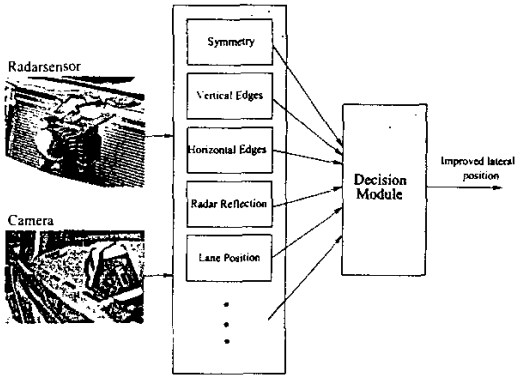

Figure 3: Architecture of the vehicle detection module.

- Template matching of a simple u-shape in the norm-gradient image, scaled in x and y and translated in the search area as proposed in [17].
- The probability of the center of an obstacle using the radar measurement. The probability is defined by the distribution of center axes relative to the radar measurement. It was calculated by a histogram of measured and labeled angles.
- A function defining the probability of the position of the car within the lane, assuming a Gaussian distribution that the vehicle is driving in the center of the lane.

Assuming a normal car width, the symmetry is a reliable operator if the obstacle has strong vertical edges. But it fails under different lightening conditions, e.g. shadows or sun light from the side. Analysis of symmetry on labeled images has shown similar results testing different expected car widths. We are using the symmetry operator only with a width smaller than most of the cars. This allows us to use the symmetric structure on the backside of cars. The advantage of this approach is that it allows to analyze a complementary feature of vehicles in addition to the 'claw'-operator which only searches for vertical edges. There are lots of symmetric areas in images, e.g. directly on the road, which is an area with little structure and perfect symmetry. Therefore the symmetry operator is normalized to the energy in this area.

4.2 Vehicle Tracking

The vehicle tracking is done by means of a Kalman filter. Longitudinal and lateral tracking are separated.

Getting raw data from the radar sensor, an adequate kine-

matic model for the longitudinal tracking is given by:

$$\dot{d}_{obj_i} = v_{rel,obj_i} \quad (5)$$

$$\dot{v}_{rel_{obj_i}} = 0 \quad (6)$$

The lateral tracking is done fusing the vehicle detection with the lane recognition module. This approach was first suggested in [3]. A similar approach is described in [14]. This approach is motivated by the human strategy when driving in rainy night on a lonely cross-country highway with bad markings. Human drivers will orientate on the tail lights of cars in front, because this enlarges the range of sight and allows to 'see' the road course assuming that the car stays in its lane without significant lateral motion. Every additional vehicle enhances the estimation of the road course.

We use the cars in front in order to improve the estimation of the road course. Tracking the vehicles the relevant parameters can be extracted from the measured distance and angle over time. The basic assumption is that the lateral motion of the leading vehicles relative to the lane is small, which can be expressed by:

$$\dot{x}_{off,obj_i} = 0 \quad (7)$$

Large lateral motion caused by lane changes of the tracked vehicles can be detected by appropriate tests. A central fusion Kalman filter is introduced based on the models of the lane recognition system as described in Sec. 3. Each vehicle adds one state to the differential equation system.

The vehicle measurements are incorporated in the Kalman estimation by an additional measurement equation:

$$x_{off,obj_i} = -x_{off} + x_{obj_i} - \Delta\psi d_{obj_i} + \frac{c_0}{2} d_{obj_i}^2 + \frac{c_1}{6} d_{obj_i}^3 \quad (8)$$

The lateral offset x_{obj_i} of each vehicle i is related to the measured angle φ_{obj_i} via

$$\varphi_{obj_i} \approx sin\varphi_{obj_i} = \frac{x_{obj_i}}{d_{obj_i}}, \varphi_{obj_i} < 5° \quad (9)$$

Every tracked obstacle is used as a measurement of the road course, extending the range of sight enormously. This results in a better estimation of the yaw angle $\Delta\psi$, the curvature c_0 and the clothoidal parameter c_1. Moreover, the lane assignment of the ACC can be improved.

Tracking the vehicles in the image is done using three different components:

1. The vehicle detection system as presented above with minimized search areas.
2. Correlating the rear sides of the vehicles over time.
3. A Hausdorff tracker ([18] and [17]).

According to the predicted position of the vehicle in the next image, the operators are applied and scaled within the 3σ-area independently. The computation time is reduced using a Gaussian pyramid for close vehicles. The aim of combining three different vehicle tracking schemes is to obtain a reliable and redundant system behavior.

The localization of each obstacle is done with a average computation time of about 7 ms on a 700 MHz Pentium III. The classification task is performed in about 2 ms for each obstacle. The tracking takes about 4 ms for every single vehicle.

Again, the estimation of the road course can be improved by fusion the above described system with a DGPS-based Trajectory system as described in [19].

5 Results

The results presented in this section were gathered during a test drive on a winding German highway under adverse weather conditions. A typical scene is shown in Fig. 4. It had been raining and the sun reflected on the wet pavement causing a bad contrast. One obstacle at a distance of about 90 m is tracked. The dark cross shows the radar reflection point on the left side of the car. The brighter lines represent the results of the monocular localization and tracking. The conditions presented in the image are challenging for many lane recognition systems because the contrast of the markings to the pavement is bad. The predicted markings and the estimated centerline of the road using the fusion approach show good results. The range of sight of our classic optical lane recognition system is about 20 m along this test sequence.

The road of the test drive, including Fig. 4, consists of a straight segment, going into a wide curve and than into a sharper curve with a radius of about 700 m. The ego-vehicle approaches the vehicle in front, which changes lane during the sequence. At first, the distance to the car in front is about 95 m, in the end it is about 26 m.

Figure 5 depicts the angle to the vehicle in front during the test sequence. The diagram show the angle φ_{obj_i} obtained by radar and by the monocular tracking system. The raw as well as the filtered angles are presented. As can be seen, the radar is 'sliding' on the vehicle's rear side. The raw radar measurements are very noisy. The radar measurement failed twice, which is shown by values larger than 0.1 rad. The localization system recognized the vehicle at a distance of 95 m and classified it as car. The monocular tracking system tracked reliably the center of the car till the end of the sequence with a vehicle distance of 26 m. The

Figure 4: A typical scene during our test drive. It was raining and the sun reflected on the wet pavement causing a bad contrast. One obstacle at a distance of about 90 m is tracked. The dark cross shows the radar reflection point on the left side of the car. The brighter lines represent the results of the monocular localization and tracking. The predicted markings and the estimated centerline of the road using the presented fusion approach is shown.

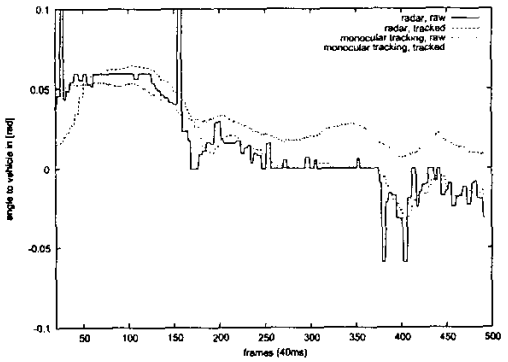

Figure 5: Results from the monocular tracking. The diagram shows the angle φ_{obj_i} estimated by radar and the results obtained by monocular tracking. The raw as well as the filtered angles are presented. As can be seen, the radar is 'sliding' on the vehicle's rear side. The raw radar measurements are very noisy, it fails twice, which is shown by values larger than 0.1 rad. In contrast to that, the monocular tracking shows very good results.

computation time was controlled using a Gaussian pyramid.

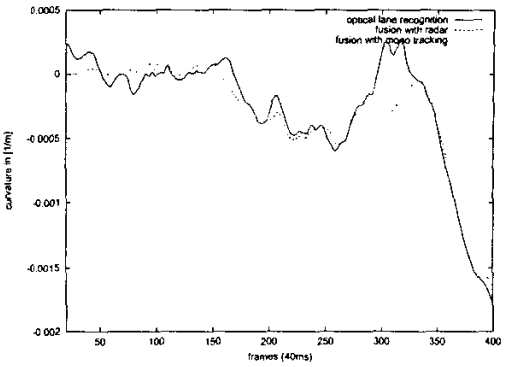

Figure 6: Results from the fusion approach, showing the curvature estimation of the optical lane recognition system and the fusion approach with the raw radar measurements and the monocular tracking system. The estimation of the optical lane recognition system is noisy. As can be seen, the curvature estimation can be improved by taking other vehicles into account.

Figure 6 shows the estimated curvature c_0 of this test drive. Depicted is the optical lane recognition and the fusion approach with the raw radar measurements and with the monocular tracking. The range of sight for the marking detection is reduced to about 20 m. Therefore, the curvature estimation of the optical lane recognition system is noisy. This noise is even larger in the estimation of the clothoid parameter c_1. If other vehicles are taken into account, the noise in the curvature estimation can be reduced. Particularly at larger distances, the curvature estimation is improved (frame 0-150).

In Fig. 7 the lateral offset of the vehicle in front is shown: Firstly, the geometrical lane assignment of the raw radar data and secondly, the estimated offset using the fusion approach. As can be seen, the vehicle in front is changing lane. The estimation of the geometrical offset of the raw radar measurements is very noisy. The measurement failures described above are presented with a lateral offset larger than 10 m. The monocular tracking as well as the fusion of the lane recognition system with the obstacle detection improves the estimation of the lateral offset. The precise estimation of the center of the vehicle as well as the vehicle width allows an early decision about the leading vehicle.

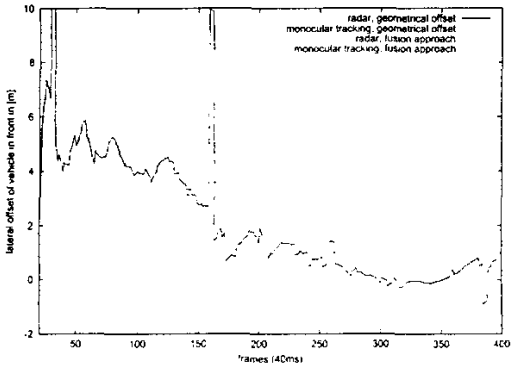

Figure 7: Results from the fusion approach, showing the measured lateral offset of the car in front. The estimation of the geometrical offset using the the raw radar measurements is very noisy. The measurement failures described above are presented with a lateral offset larger than 10 m. The monocular tracking as well as the fusion of the lane recognition system with the obstacle detection improves the estimation of the lateral offset.

6 Summary and Future Work

This paper presents the two-step fusion approach combining vehicle detection by radar and vision and our optical lane recognition system. It is shown that the lane recognition and the ACC benefit mutually from each other with the proposed Kalman filter approach.

The monocular tracking of the radar obstacles significantly improves the estimation of the center of the obstacles and therefore improves the lane estimation as well as the lane assignment of obstacles.

Further investigations will be conducted on the following topics:

- Multi-hypothesis tracking.
- Improved estimation of the vertical curvature and the tilt angle from the tracked vehicles as described in [14].

References

[1] Uwe Regensburger, "Zur Erkennung von Hindernissen in der Bahn eines Strassenfahrzeugs durch maschinelles Echtzeitsehen," Dissertation, Universitaet der Bundeswehr Muenchen, Fakultaet fuer Luft- und Raumfahrttechnik, 1993.

[2] U. Franke and I. Kutzbach, "Fast Stereo based Object Detection for Stop&Go Traffic," in *Proc. IEEE Conference on Intelligent Vehicles*, 1996.

[3] 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.

[4] Axel Gern, Uwe Franke, and Paul Levi, "Advanced Lane Recognition – Fusing Vision and Radar," in *Proc. IEEE Conference on Intelligent Vehicles*, 2000, p. 45.

[5] 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, p. 161.

[6] K. Kluge and C. Thorpe, "Representation and recovery of road geometry in YARF," in *Proc. IEEE Conference on Intelligent Vehicles*, 1992, p. 114.

[7] Jinyou Zhang and Hans-Hellmut Nagel, "Texture-Based Segmentation of Road Images," in *Proc. IEEE Conference on Intelligent Vehicles*, 1994, p. 260.

[8] Dean Pomerleau, "RALPH: Rapidly Adapting Lateral Position Handler," in *Proc. IEEE Conference on Intelligent Vehicles*, 1995, p. 506.

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

[10] U.Franke, "Real time 3D-road modeling for autonomous vehicle guidance," in *Selected Papers of the 7th Scandinavian Conference on Image Analysis*, 1992, p. 277, World Scientific Publishing Company.

[11] Uwe Handmann, Gesa Lorenz, Thomas Schnitger, and Werner von Seelen, "Fusion of Different Sensors and Algorithms for Segmentation," in *Proc. IEEE Conference on Intelligent Vehicles*, 1998.

[12] M. Beauvais and S.Lakshmanan, "CLARK: a heterogeneous sensor fusion method for finding lanes and obstacles," in *Image and Vision Computing, Vol. 18*, 2000, p. 397.

[13] Dirk Langer and Todd Jochem, "Fusing Radar and Vision for Detecting, Classifying and Avoiding Roadway Obstacles," in *IEEE Conference on Intelligent Vehicles*, 1996.

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

[15] Christian Woehler, "Neuronale Zeitverzoegerungsnetzwerke fr die Bildsequenzanalyse und ihre Anwendung in fahrzeuggebundenen Bildverarbeitungssystemen," Dissertation, Rheinische Friedrich-Wilhelms-Universitaet Bonn, Bonn, 2000.

[16] A. Kuehnle, "Symmetry-based recognition of vehicle rears," in *Pattern Recognition Letter*. 1991, p. 249, Vol. 12.

[17] Thomas Kalinke, Christos Tzomakas, and Werner von Seelen, "A Texture-based Object Detection and an adaptive Model-based Classification," in *Proc. IEEE Conference on Intelligent Vehicles*, 1998, p. 143.

[18] Daniel P. Huttenlocher, Gregorya A. Klanderman, and William J. Ruchlidge, "Comparing Images Using the Hausdorff Distance," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1993, p. 850, Vol. 15, No. 9.

[19] 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.