Machine Vision and Applications (2003) 14: 42–49
Digital Object Identifier (DOI) 10.1007/s00138-002-0093-y

Machine Vision and Applications

Radar and vision data fusion for hybrid adaptive cruise control on highways

U. Hofmann, A. Rieder, E. D. Dickmanns

Universität der Bundeswehr München, Institut für Systemdynamik und Flugmechanik,
Werner-Heisenberg-Weg 39, 85579 Neubiberg, Germany; e-mail: ulrich.hofmann@unibw-muenchen.de

Abstract. A system for hybrid adaptive cruise control (HACC) on high-speed roads designed as a combination of a radar-based ACC and visual perception is presented. The system is conceived to run on different performance levels depending on the actual perception capabilities. The advantages of a combination of the two different types of sensors are discussed in comparison to the shortcomings of each single sensor. A description of the visual lane detection and tracking procedure is given, followed by an overview of the vehicle detection, hypothesis generation, and tracking procedure. Enhanced robustness is achieved by cooperative estimation of egomotion and the dynamics of other vehicles using the lane-coordinate system as a common reference. Afterwards, the assignment of vehicles to lanes and the determination of the relevant vehicle for the longitudinal controller is described.

Keywords: Adaptive cruise control – Dynamic machine vision – Road recognition – Vehicle recognition

1 Introduction

At present, great effort is put into the development of driver assistance and comfort systems, such as lane departure warning, stop-and-go traffic assistance, convoy driving, and adaptive cruise control (ACC). Unfortunately, these efforts mostly result in independent solutions for each task, which do not communicate their knowledge with each other.

The expectation-based multi-focal saccadic, EMS-Vision, system from the University of the Federal Armed Forces Germany in Munich (UBM) bundles the information different experts extract from sensor data and makes it available to all other experts. As a spin-off of the overall system architecture [1], in cooperation with an automotive supplier[1], UBM designed a system for hybrid adaptive cruise control. It is a combination of a radar-based ACC system and visual perception for the detection and tracking of vehicles and lanes.

Several other research groups also use a combination of radar and vision sensors for driver-assistance tasks. In [2] a system is described where a purely radar-based obstacle-detection and tracking is performed. To determine the relevant vehicle for the controller of the own car, the results of an optical lane recognition are used to assign vehicles to lanes. In [3] and [4], the information about the positions of vehicles received from a radar system is fused with a vision-based lane recognition to improve the geometry estimation of the lane. The main difference between the system presented here and these systems is that the radar-based vehicle detection is cross-checked by vision to improve the position estimation of vehicles. Further, the system is conceived to run on different levels of performance that depend on the perception capabilities.

2 System specification

This HACC system is a comfort system designed for motorways and similar roads with white lane markings on both sides of all lanes. This is a well-known domain where the expected obstacles are restricted to road vehicles. The own car (Ego) is driven manually in the lateral direction and is controlled autonomously in the longitudinal direction. A desired speed is set by the human driver. The computer controls the velocity of the car in such a manner that a specified safety distance to other vehicles (OVs) is kept, and the difference between the desired and the actual velocity is as small as possible. It is the driver's responsibility to choose the lane of the road and to decide whether to overtake or not. For example, if there is another vehicle in front of the Ego driving at a velocity slower than the desired speed, the Ego slows down and follows at the speed of the leading car. The accelerations commanded by the velocity controller are restricted to a level such that the passengers feel comfortable. The safety distance to the OV ahead shall not be smaller than, for example, $1.6\,\mathrm{s} \times \mathrm{velocity}$ of the own car. The HACC is not a safety system. The driver always has to be aware of the traffic situation. They are legally responsible for all actions of the car. The driver can overrule the HACC system at any time. The maximal pressure of the braking system that the HACC may command is limited to

Correspondence to: U. Hofmann

[1] We thank our project partner for supplying their profound knowledge on radar-based ACC to the project.

a deceleration of $-2.5\,\mathrm{m/s^2}$, which implies that the HACC is not able to command emergency braking.

3 Scalable performance

The system is designed to operate at different performance levels, as depicted in Fig. 1. The initial system status is given when no cruise control is active, the human driver controls the velocity and the heading direction.

The first performance step is that the conventional radar-based ACC system is activated. The decision whether an OV might be relevant, for example, if it is driving ahead of the Ego in the own lane at a velocity slower than that of the Ego, is made using the so-called driving tube. This driving tube is fixed parallel to the longitudinal axis of the Ego. Its curvature is estimated from the relative speeds of the 4 wheels using ABS-sensor signals. The system has no knowledge about the relative position to the real lanes.

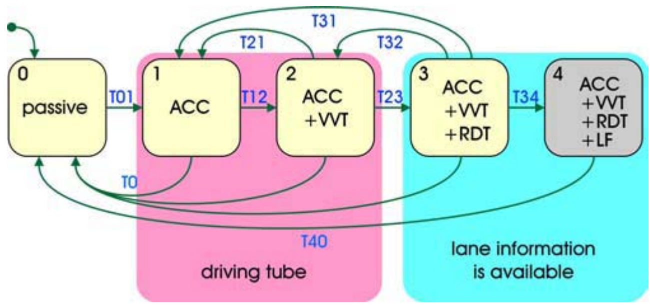


Fig. 1. Scalable performance steps. ACC denotes radar-based adaptive cruise control; VVT denotes visual vehicle validation and tracking; RDT denotes road detection and tracking; LF denotes lane follow

The second performance step is that the OV hypotheses generated by the radar module are validated by vision and the lateral positions relative to the Ego, and the dimensions of the OVs are determined. If the validation is successful, the OVs are inserted into the scene tree as the central knowledge-representation scheme for physical objects in the system (see Sect. 6). All objects are tracked by vision. Also, at the second performance step, no information about the position and shape of the lane is available.

At the third performance step, the detection and tracking of the own lane and the assignment of the validated OVs to a lane (the own and the two adjoining lanes) takes place.

A fourth performance step could be to follow a lane completely autonomously. This step was not in the scope of the project, but is a standard capability in UBM's EMS-Vision system.

The desired performance level is set by the driver via the human-machine interface (HMI). In Fig. 1, the transitions, $\mathrm{T}xy$, between the different performance levels stand for conditions that have to be met for a performance-level change. T01 has to check whether the radar module is ready to perform the ACC task. Transition T12 verifies that vision is present and that the VVT process is already delivering data. Transition T23 checks whether automatic lane detection is completed and that the RDT process is tracking the lane. If the RDT process stops tracking and starts a new automatic detection of the lane while the system is running at performance-level 3, the system changes via transition T32 to level 2 until T23 is satisfied. If the weather or lighting conditions are not suitable for vision, the system changes via T21 or T31 to performance-level 1. If the radar is not active or the driver overrules the computer, the system always changes via T0 to performance-level 0. If the desired performance-level is 4, the system can change via T34 to level 4 and may end the autonomous mode via T40 if the ACC is not active, the driver overrules the system, or lane tracking fails.

4 Sensor properties

The reason for a combination of a radar-based ACC with a vision system is to overcome the shortcomings of a purely radar-based ACC system and of a purely vision-based ACC as well. Shortcomings for the radar system are:

- Reflections on crash barriers can lead to false alarms.
- Two vehicles driving side-by-side at nearly the same speed are hardly distinguishable and may appear as only one obstacle, which is assigned to one lane. That means that a vehicle or motorcycle beside a truck can be invisible to the radar.
- The determination of the lateral position of a vehicle relative to the Ego is considerably less precise than that of the longitudinal distance.
- The own position and the relative positions of the OVs have no reference to the real lanes. This makes the decision whether an obstacle is relevant or not very difficult, especially at larger distances. The risk of false alarms is high.
- The radar-based ACC used suppresses vehicles with a velocity slower than a threshold value and oncoming traffic for vehicle-hypothesis generation. A so-tuned conventional ACC system is not able to handle stop-and-go traffic.

On the other hand, advantages of the radar system are:

- A radar system is independent of weather and lighting conditions.
- The determination of the distances and relative velocities to OVs is very precise.

Advantages of the vision system are the ability to:

- Determine the lateral positions of OVs relative to the Ego with high accuracy.
- Determine the dimensions/shapes of OVs. This enables the system to classify obstacles and make a model-based prediction of its possible behavior.
- Detect and track the own lane.

As a consequence it is possible to:

- Determine the shape of the own lane
- Determine the position of the Ego relative to the own lane
- Recognize a lane change depending on the yaw angle and horizontal offset of the Ego's center of gravity (CG) relative to the center of the lane
- Determine the positions of OVs relative to the own lane

The drawbacks are:

- Measurement results depend on the weather and lighting conditions.
- A vision-only ACC has difficulties determining the distances to OVs in longitudinal direction because range information is lost in perspective projection. Consequently, it is rather difficult to get a precise value for the relative velocity.

Radar and vision have complementary properties. A combination of both leads to better overall system performance.

5 Sensors and hardware used

As an experimental platform, UBM's Mercedes 500 SEL, dubbed VaMP, is used. See Fig. 2 and [1].

Fig. 2. Experimental vehicle VaMP

For this project, the vehicle has been equipped with a radar system, which is attached to the center of the front bumper (Fig. 3). It has one radar club with a viewing angle of $\pm 4°$, and it is able to measure the relative velocity and distance to other vehicles in a range from 2–130 m with an accuracy of ± 1.5 m. The radar-based ACC module uses data from the ABS sensors to calculate the curvature of the trajectory of the own vehicle.

The system is able to observe the environment in front of the car with several cameras, which are mounted on a pan camera platform (Fig. 4). From this MarVEye camera configuration [1], only the video data of the highly sensitive black-and-white camera (third camera from left) and the intensity signal of the 3-chip color camera (second camera from left) are evaluated. The platform is not active. This bifocal camera configuration is equipped with a wide-angle lens on the black-and-white camera with a horizontal viewing angle of $\pm 22°$ and a tele lens on the 3-chip color camera with a viewing angle of $\pm 5.5°$. For image processing, each second field (half image) is taken with a resolution of 768×286 pixels every 40 msec.

Only one of the 3 image-processing PCs available in the whole system is used for the vision tasks here. On this computer (comp2), the VVT, the RDT, and the radar processes are running (Fig. 5). The radar process is the interface to the

Fig. 3. Radar

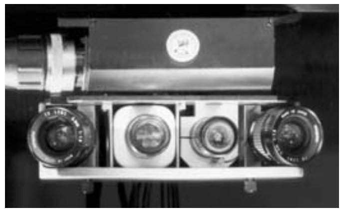


Fig. 4. Camera platform

radar hardware. The actuators get their commands via the controller PC (comp1), where the locomotion expert is running. For details see [1].

6 Scene tree

The scene tree is the internal representation of the outside world. All nodes of the scene tree represent physical objects or virtual coordinate systems. The transformations between scene nodes are described by homogeneous coordinate transformations (HCTs). HCTs can be used for describing the relative pose, 6 degrees of freedom (DOF), between physical objects as well as for perspective projection into the image coordinate systems. For details see [1]. In Fig. 5, the connections between the scene representation, the processes, and the computers used are depicted. For example, the positions and dynamics of other vehicles are estimated relative to the coordinate system of the current own lane. This coordinate system lies in the middle of the lane, tangential to the lane and moves along the lane at a constant distance ahead of the own car. The position and motion of the own car is also described relative to the lane coordinate system. Using the HCTs, for example, the positions of OVs relative to the Ego can be calculated.

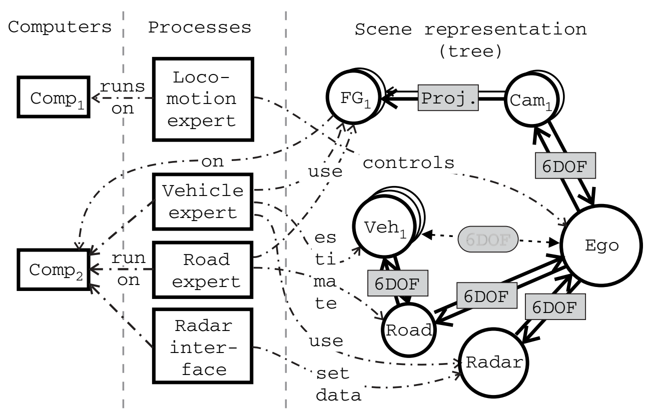


Fig. 5. EMS-Vision databases: hybrid adaptive cruise control

7 Overview of lane detection and tracking

Lane detection and tracking uses horizontal search windows to extract lane markings from grayscale images. Lane markings are characterized by dark-bright-dark grayscale transitions that match an expected width and orientation at expected positions. The expected positions and orientations of the lane markings are calculated from timestep to timestep using a 3-D model of the lane, which is projected onto the image plane. The differences between the measured and the expected positions are taken to innovate the 3-D model. In order to initialize the tracking process, a setup phase is necessary first. Lane markings are detected by extracting their edges. Therefore, the correlations between ternary masks of the kind [-1,-1,-1,0,+1,+1,+1] and the grayscale values along the search paths are calculated. A grayscale transition from dark to bright results in a maximum, and from bright to dark in a minimum correlation value. For the decision on a max or min, a suitable grayscale threshold has to be found first. Therefore, the value for a threshold is successively decreased until sufficiently many maximums and minimums are found and the threshold is larger than an allowed minimum value. Hence, two regression lines through the left and right lane markings are calculated. This is done using the image of the wide-angle camera only because, on motorways and similar roads, the influence of lane curvature is negligible at near distances (6–30 m). Under the assumptions that the position of the Ego relative to a lane has a horizontal offset smaller than half the lane width and a yaw angle smaller than, e.g., $5°$, the following items are calculated from the regression lines:

- horizontal offset of the Ego's CG relative to the skeleton line (center of the lane)
- yaw angle between the longitudinal axis of the Ego and the tangent at the skeleton line measured at the CG
- lane width

These first approximations are taken as starting values for the extended Kalman filter (EKF) [5]. During this automatic lane detection, the driving tube is used for the decision on the relevance of OVs. The lane geometry is described by a moving average clothoid model. For details see [6] or [7]. The state variables which are estimated by the EKF can be differentiated into two kinds:

- shape parameters of the model, which are the horizontal and vertical curvature, the lane width, and their changes
- position parameters, which are the horizontal offset of the CG of the Ego to the skeleton line, the yaw angle, and the pitch angle of the vehicle body relative to the lane

By successively increasing the lookahead distance from near to far distances, the model reliably approaches the real lane markings by determining their curvatures. In the wide-angle image, the search windows are set such that the lookahead distance in 3-D-space is 6–40 m, and 30–100 m for the tele-image. If the number of extracted features in the tele-image is less than a certain minimal number for several cycles, the lookahead distance is shortened and afterwards successively extended from near to far. This increases the robustness of lane tracking.

Before feature extraction is started, all search windows are checked to determine whether a vehicle obscures the expected lane markings. To do this, the bounding box for each OV is tested to determine whether it intersects with any search window. If an intersection exists, the search window is clipped (Fig. 6). If the resulting search path is too short, this measurement is disabled.

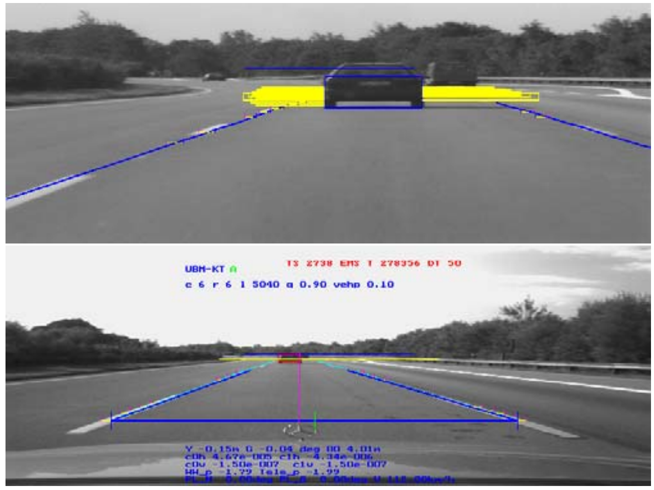


Fig. 6. Clipping search windows using bounding boxes of OVs

8 Vehicle detection, hypothesis generation, and tracking

In order to control the velocity of the Ego correctly, the HACC system has to detect all vehicles that are potential obstacles. New vehicle hypotheses are generated by evaluating the radar measurements. Within the radar module, a preprocessing of the radar measurements takes place, wherein reflections with a similar distance, relative velocity, and amplitude are grouped together. The radar system creates a list of potential vehicles every 60 msec. These measurements first have to be assigned to the existing OV hypotheses of the scene tree. This is accomplished by defining a capture area around each OV hypothesis and assigning to it all radar measurements that lie within it (Fig. 7). For details see [8].

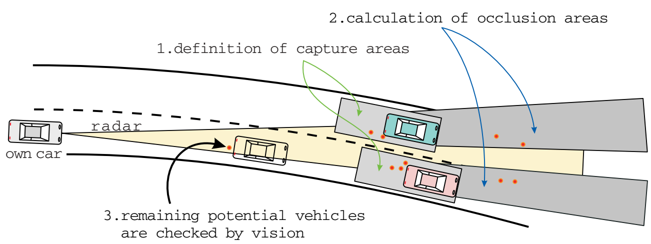


Fig. 7. Vehicle hypothesis generation

The existing vehicle hypotheses are sorted with respect to the distance from Ego. Then, the angular range covered by each vehicle hypothesis is calculated, and the radar measurements left over are checked to determine whether they lie in such an area with a larger distance than the corresponding vehicle hypothesis. If this is true, the measurement is rejected. Remaining radar measurements, which could not be assigned to an existing vehicle hypothesis or occlusion area, are candidates for new vehicle hypotheses. These are checked by vision. If the validation is successful, a new vehicle hypothesis is added to the scene tree. If a hypothesis is not updated either by radar or by vision for several cycles, it is removed from the scene tree. All vehicle hypotheses in the scene tree are tracked. At the position of a candidate for a new vehicle hypothesis, a box model is initialized to fit the shape of the potential vehicle. The orientation of the box in 3-D is assumed to be parallel to the lane at this distance. Depending on the yaw angle relative to the Ego, the length or the width of the box is estimated. Furthermore, the lateral position and lateral velocity, the longitudinal position, and the speed and acceleration are estimated via EKF for each OV. For details see [8]. As already mentioned in Sect. 6 the positions and dynamics of OVs and the egomotion are estimated relative to the lane coordinate system. This means that, implicitly, a separation between the dynamics of the Ego and of OVs is performed, which results in an enhanced robustness during the tracking of OVs.

9 Vehicle of relevance

In order to decide which object is relevant for the longitudinal controller, it is necessary to determine the positions and the future behavior of other vehicles relative to the Ego. The relevance decision is made by assigning them to the lanes of the road with the implicit assumption that OVs keep their lane most of the time.

The pure radar-based ACC system (performance level 1) can only use the driving tube for the relevance decision. The driving tube is fixed with the longitudinal axis of the own vehicle. Its only parameter is the curvature, which is calculated from the speeds of the 4 wheels measured with the ABS sensors. It has no reference to the real lane geometry (Fig. 8).

Fig. 8. Driving tube (bright overlay) and lane model (dark) while the Ego is driving near the right border of the lane but still inside the lane

Movements inside the lane result in an alternating curvature of the driving tube. To reduce this behavior, the curvature

Fig. 9. Driving tube and lane model while performing a lane change

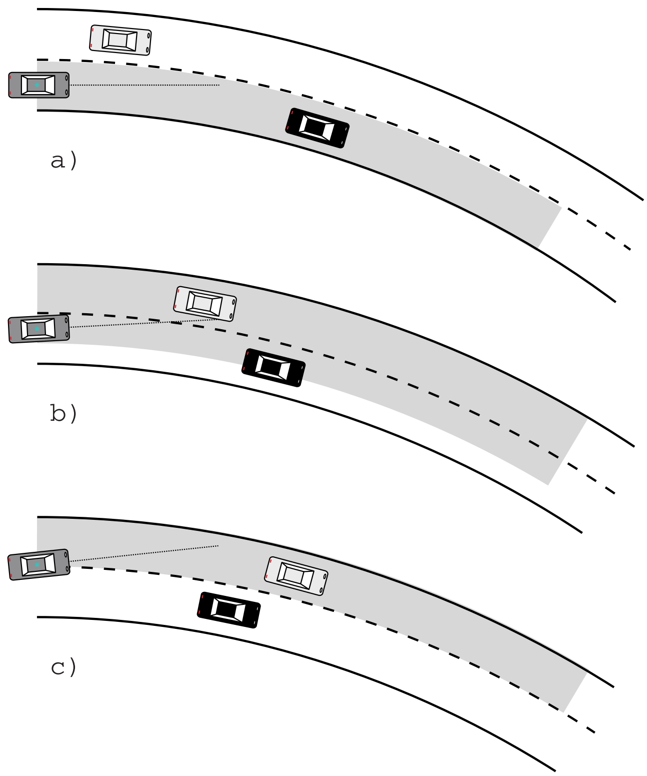


Fig. 10. Relevance decision area as function of the horizontal offset

of the driving tube is calculated by lowpass filtering. As a consequence, the driving tube lags behind or overshoots the value of the real lane curvature if the steering angle changes strongly. The decision for relevance using the driving tube easily leads to false alarms, especially at far distances. For example, if the Ego passes a vehicle in a left curved lane, it could become the relevant vehicle if the driving tube changes its curvature because of steering angle perturbations of the Ego within the lane.

Figure 9 shows the driving tube and the lane model during a lane change. It can be seen, that at near distances (6–30 m) on high speed roads the driving tube can be approximated by a straight lane, because the curvature has nearly no influence.

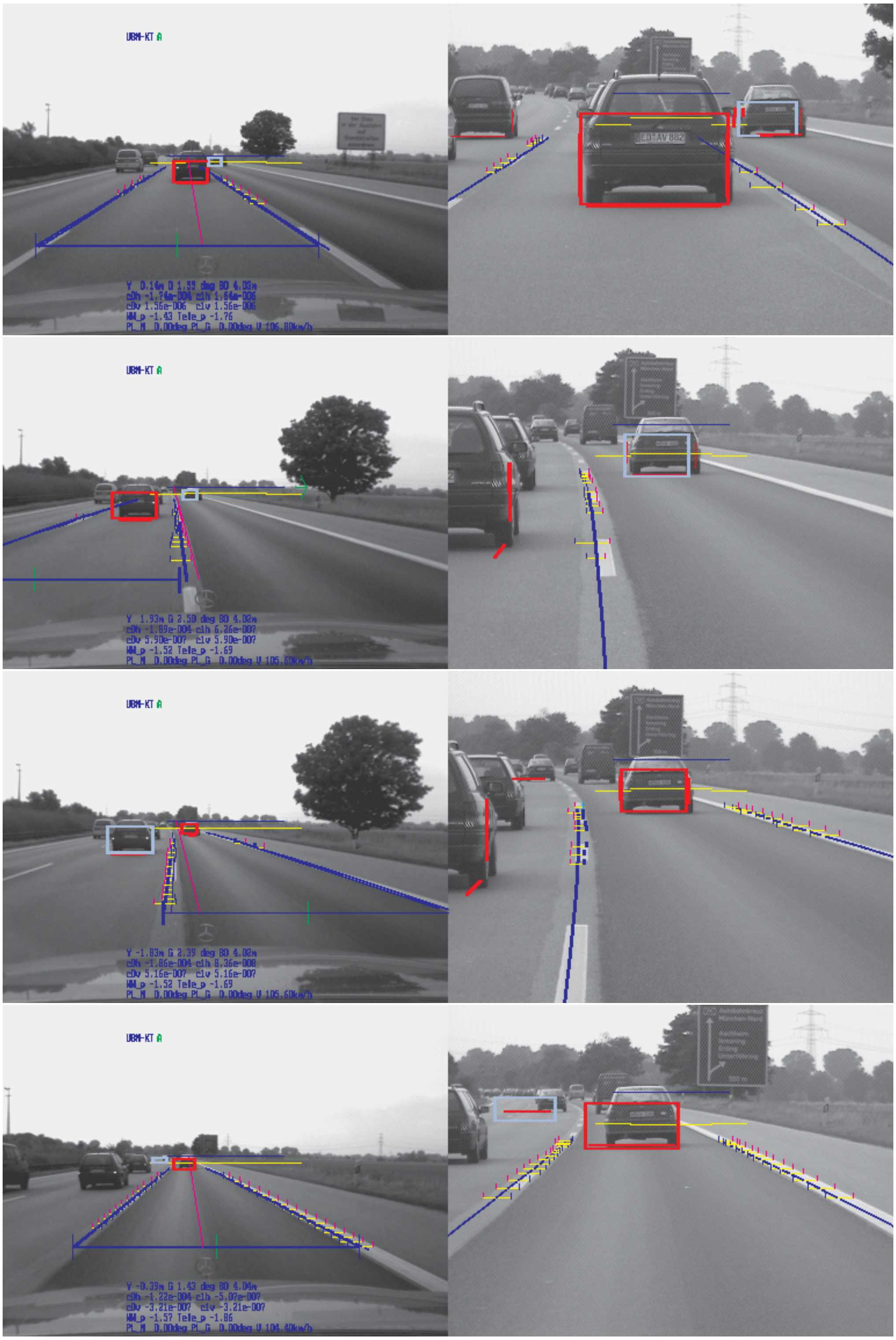


Fig. 11. Lane change to the right

At far distances a relevance decision based on the driving tube would definitely be wrong.

In contrast, the visual lane detection and tracking process is able to calculate the position of the Ego relative to the lane with 6DOF and is able to determine the shape parameters of the lane. For the assignment of OVs to the lanes and the decision of their relevance, three cases can be differentiated:

1. Ego is driving inside the own lane and no lane change is indicated or assumed.
2. Ego is driving inside the own lane and a lane change to the left is assumed or notified by the left indicator.
3. Ego is driving inside the own lane and a lane change to the right is assumed or notified by the right indicator.

The detection of a lane change can be done by observing the yaw-angle and the horizontal offset of the Ego. If the predicted horizontal offset of the Ego at a lookahead distance of 10 m is larger than 60% of the lane width, a lane change is assumed.

In case 1, the relevance decision area (RDA) is identical to the current own lane (Fig. 10a). A vehicle will be assigned to the own lane if the horizontal offset is smaller than half the width of the lane. If a vehicle is already assigned to the own lane, it is associated with it as long as the horizontal offset is smaller than half the width of the lane plus half the width of the vehicle.

During a lane change, it is reasonable to hang on to the current own lane for lane tracking until the CG has a horizontal offset larger than half the lane width and then to change the tracked lane. But performing the relevance decision with respect to the current lane will not lead to satisfactory behavior because the Ego will leave the lane within a short time. During and after the lane change, its velocity has to be controlled in consideration of the vehicles in the desired lane.

Normally, an overtake maneuver (case 2) is performed for driving faster than in the current lane. In order to overtake, a strong acceleration is needed at the beginning of the maneuver. Therefore, human drivers mostly accept a safety distance shorter than otherwise chosen. This means that increasing the velocity is performed by decreasing the safety distance in the current lane. The switch for the left indicator could be used for starting acceleration by shortening the allowed safety distance to the leading vehicle.

Simultaneously, the RDA should be extended to the desired lane. Its width in the current lane should be successively decreased as a function of the horizontal offset (Fig. 10b). As long as the CG of the Ego is inside the current own lane, the width of the RDA ranges from that part of the own lane, which is still covered by the vehicle shape, over the complete width of the desired lane. The extended RDA is parallel to the skeleton line of the current lane. If the horizontal offset of the Ego is larger than half the lane width, the new lane becomes the own lane and the RDA becomes identical to the new own lane (Fig. 10c). Case 3 is nearly the same as case 2, but no acceleration by shortening the safety distance in the current lane is allowed.

Afterwards, all OVs are sorted according to their distance to the Ego and only the nearest OV within the RDA is set to be relevant. The velocity and distance to the relevant OV is communicated to the vehicle controller for the longitudinal motion for adjusting the speed of the Ego in a way that the convoy distance is larger than a desired value, e.g., $1.6\,\mathrm{s} \times$ velocity of the own car.

10 Experimental results

In Fig. 11, a lane change from the middle lane to the right lane is manually performed. You can see on the left the image of the wide-angle camera and on the right the image of the tele camera. The top image shows the own car following the OV with the velocity of the vehicle ahead and the desired safety distance. The relevant vehicle is marked with a red box (dark), while other tracked vehicle hypotheses are characterized by gray boxes (bright). The second image shows the own car just one timestep before the center of gravity of the own car crosses the lane markings. The relevant vehicle is still the vehicle ahead in the middle lane. The third image is just after the lane change is performed. The relevant vehicle has changed from one in the middle lane to the vehicle ahead in the right lane (dark box). The bottom image illustrates the approach to the vehicle ahead.

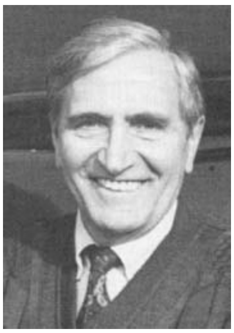
11 Conclusions and outlook

The combination of radar and vision leads to a system with enhanced performance, capable of handling several tasks jointly using a common knowledge base. The system can select an appropriate level of performance depending on the hardware status or the performance of the experts. Monitoring the performance of the vision experts takes the weather and lighting conditions implicitly into account. Lane-departure warning can be performed easily using the knowledge about the position of the Ego relative to the own lane. The implicit separation between egomotion and the dynamics of OVs using the lane coordinate system as a common reference clearly improves the robustness of vehicle tracking. Convoy driving using activated lateral control is possible if the speed of the leading car is sufficiently large, such that it is not suppressed by radar measurement preprocessing. For speeds slower than that, first experiments are under way using trinocular stereovision for handling stop-and-go traffic. For details see [8].

References

1. Gregor R, Lützeler M, Pellkofer M, Siedersberger KH, Dickmanns ED (2001) A vision system for autonomous ground vehicles with a wide range of maneuvering capabilities. In: Schiele B, Sagerer G (eds) Proceedings of 2nd International Workshop on Computer Vision Systems, Vancouver, Canada, July. Berlin Heidelberg New York, Springer, pp 1–20
2. Jochem T, Langer D (1996) Fusing radar and vision for detecting, classifying and avoiding roadway obstacles. In: Proceedings of the IEEE Symposium on Intelligent Vehicles, Tokyo, September, pp 333–338
3. Franke U, Zomotor Z (1997) Sensor fusion for improved vision based lane recognition and object tracking with range-finders. In: Proceedings of the IEEE Conference on Intelligent Transportation Systems Conference, Boston, November
4. Franke U, Gern A (2000) Advanced lane recognition - fusing vision and radar. In: Proceedings of the IEEE Symposium on Intelligent Vehicles, Dearborn, Mich., October, pp 45–51
5. Bierman GJ (1997) Factorization methods for discrete sequential estimation. In: Mathematics in science and engineering, vol 128. Academic Press, New York
6. Dickmanns ED (1988) Dynamic computer vision for mobile robot control. In: Jarvis RA (ed) Proceedings of the 19th International Symposium and Exposition on Robots, Sydney, November, pp 314–327
7. Behringer R (1996) Visuelle Erkennung und Interpretation des Fahrspurverlaufs durch Rechnersehen für ein autonomes Straßenfahrzeug. In: Fortschrittsberichte, vol 310. VDI, Düsseldorf
8. Rieder A (2000) Fahrzeuge Sehen. PhD thesis, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik

André Rieder received his Ph.D. in 2000 from the University of Federal Armed Forces Munich, Germany (UBM). He is currently working as a member of technical staff at Sarnoff Corporation in Princeton, NJ. He specializes in the design and integration of vision systems for autonomous and semi-autonomous vehicles. Key accomplishments include the development of active gaze control for the German DoD project PRIMUS, the design and implementation of a hybrid automatic cruise control system at UBM, and the integration of stereo vision systems for the DARPA program PerceptOR and the US/German DOD program AutoNav. As a technical leader within Sarnoff's Automotive Initiative, he is currently working towards market ready vision systems for standard automobiles.



Ernst Dickmanns studied Aerospace Engineering at the RWTH Aachen, Germany. Till 1975 he has been with the German Research Establishment for Aero- and Astronautics, working in flight dynamics and control engineering; graduate studies in this field at Princeton University. He received his Dr.-Ing. from RWTH Aachen and performed a Post-doctorate Research Associateship at NASA Huntsville. Since 75 he is professor for control engineering at the Aero-Space Department of the UniBw Munich. In his research field 'Dynamic machine vision' he has pioneered applications in satellite docking, ground vehicle guidance, in on-board-autonomous landing approaches of aircraft, autonomous landmark navigation for helicopters, and grasping of free-floating objects.

Ulrich Hofmann (1971) received his diploma in electrical engineering (1997) at the University of Erlangen-Nuernberg, Germany. At present he is a Ph.D. student of Prof. Dickmanns at the University of the Federal Armed Forces Munich, Germany (UBM). Within the UBM-project for a hybrid adaptive cruise control system he was responsible for the robust lane recognition and together with André Rieder responsible for the design of the system concerning the relations between lane and vehicle recognition. In the US/German DOD program AutoNav he developed vision algorithms for the detection and tracking of a negative obstacle in order to perform an autonomous jink. Currently his main interests lie in the development of driver assistance applications.