EMS-Vision: Application to Hybrid Adaptive Cruise Control

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

Institut für Systemdynamik und Flugmechanik,
Universität der Bundeswehr München (UBM),
D-85577 Neubiberg, Germany

ulrich.hofmann@unibw-muenchen.de

Abstract

This article presents a system for hybrid adaptive cruise control (HACC) on high speed roads designed as a combination of a radar-based ACC and visual perception. The system is conceived to run on different performance levels depending on the perception abilities. The advantages of a combination of the two sensor types are discussed in comparison to the shortcomings of each single sensor type. 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. Afterwards, the assignment of vehicles to lanes and the determination of the relevant vehicle for the longitudinal controller is described.

Keywords road recognition, vehicle recognition, dynamic machine vision, adaptive cruise control

1 Introduction

Many automobile firms put great effort into the development of driver assistance and comfort systems like lane departure warning, stop-and-go traffic assistance, convoy driving or adaptive cruise control (ACC). Unfortunatly, these efforts mostly result in independent solutions for each task which do not communicate their knowledge among each other.

The EMS[1]-Vision system of UBM bundles the information different experts extract from sensor data and makes it available for all other experts. As a spin-off of the overall system architecture [1], UBM designed, in cooperation with an automotive supplier[2], a system for hybrid adaptive cruise control. It is a combination of a radar-based ACC system and visual perception for vehicle as well as lane detection and tracking.

[1]Expectation-based Multi-focal Saccadic

[2]We thank our project partner, especially T. Müller, for supplying their profound knowledge on radar-based ACC to the project.

2 System Specification

This HACC system is meant for motorways and similar roads with white lane markings on both sides of all lanes. This is a wellknown domain where the expected obstacles are restricted to road vehicles. The own car (Ego) shall be driven manually in the lateral direction and is controlled autonomously in the longitudinal direction. It is in the driver's responsibility to choose the lane of the road and he has to decide whether to overtake or not. 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 (OV) is kept and the difference between the desired and the actual velocity is as small as possible. E.g. if there is another vehicle in front of the Ego driving with a velocity slower than the desired speed, the Ego slows down and follows with the speed of the leading car. The accelerations commanded by the velocity controller are restricted to a level, that the passengers feel comfortable. The safety distance to the OV ahead shall not be smaller than e.g 1.6sec $\cdot$ velocity of the OV. The driver can overrule the HACC system at any time. The HACC is a comfort system, not a security system. The maximal pressure of the braking system the HACC may command is limited to an acceleration of $-2.5\mathrm{m/s}^2$. That implies that the HACC is not able to command emergency braking. The driver always has to be aware of the traffic situation. He is legally responsible for all actions of the car.

3 Scalable Performance

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

A first performance step is that the conventional radar-based ACC system is activated. The decision whether an

OV might be relevant, which means driving ahead of the Ego in the own lane with a velocity smaller 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.

A 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 as well as 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 Chapter 6). All objects are tracked by vision.

The third performance step is the additional detection and tracking of the own lane and the assignment of the validated OVs to a lane (the own and the two adjoining lanes).

A forth performance step could be to follow a lane completely autonomous. This step was not in the scope of the project, but is a standard ability in UBM's system.

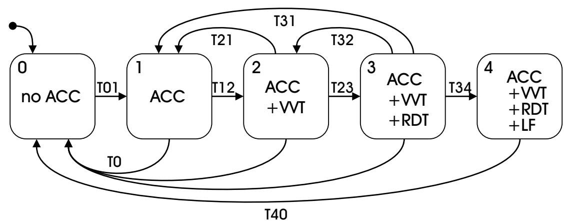

Figure 1: Scalable performance steps
ACC = radar-based Adaptive Cruise Control, VVT = visual Vehicle Validation and Tracking, RDT = Road Detection and Tracking, LF = Lane Follow

The desired performance level is set by the driver via the human machine interface (HMI). In figure 1 the transitions Txx between the different performance levels stand for conditions which 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 the VVT process is already delivering data. Transition T23 checks whether automatic lane detection is completed and the RDT process tracks 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 fulfilled. 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 pure radar-based ACC system and of a pure 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 with nearly the same speed are hardly distinguishable and may appear as only one obstacle, which is assigned to one lane. That means a vehicle or motorcycle beside a truck can be invisible for the radar.
- The determination of the lateral position of a vehicle relative to the Ego is considerably less precise than of the logitudinal 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

- 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 classification of obstacles and to 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 in 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 experimental platform UBM´s Mercedes 500 SEL, dubbed VaMP, is used. See figure 2 and [1].

Figure 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. 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 to 130 meters with an accuracy of ± 1.5m. The radar-based ACC module uses data of the ABS-sensors to calulate 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. From this MarVEye camera configuration [1] only the video data of the high sensitive black-and-white camera and the intensity signal of the 3-chip color camera are evaluated. The platform is not active. This bifocal camera configuration is equipped with an 8mm lens on the $1/2$” chip b/w-camera (wide-angle) and a 25mm lens on the 1/3” 3-chip color camera (tele), which corresponds to a 37.5mm lens on a 1/2” chip camera. For image processing each second field (half image) is taken with a resolution of 768x286 pixels every 40msec.

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 process are running. See figure 3. The Radar process is the interface to the 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 (HCT). HCTs can be used for describing the relative pose (6 degrees of freedom = 6DOF) between physical objects as well as for perspective projection into the image coordinate systems. For details see [1]. In figure 3, the connections between the scene representation, the processes and the computers used are depicted.

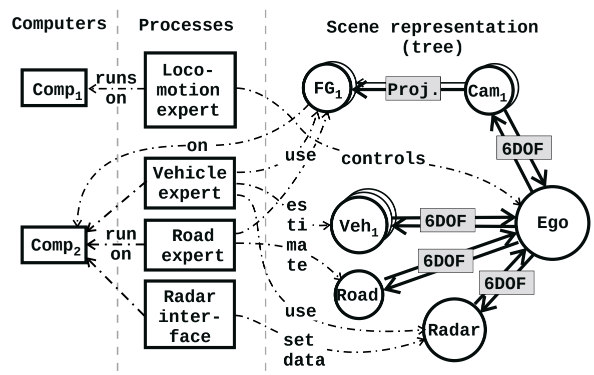

Figure 3: EMS-Vision data bases: hybrid adaptive cruise control

7 Overview of Lane Detection and Tracking

Lane detection and tracking uses horizontal search windows to extract lane markings from grey-level images. Lane markings are characterized by dark-bright-dark grey-level transitions which 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 3D-model of the lane which is projected into the image plane. The differences between the measured and the expected positions are taken to innovate the 3D-model. In order to initalize 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 grey-level values along the search paths are calculated. A grey-level 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 grey-level threshold has to be found first. Therefore, the value for a threshold is successively decreased until enough maxima and minima are found and the threshold is larger than an allowed minimal 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 (6-30m). Under the assumptions that the postion of the Ego relative to a lane has a horizontal offset smaller than half the lane width and a yaw angle smaller e.g. 5 degree, the following items are calulated from the regression lines:

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

These first approximations are taken as starting values for the Extended Kalman Filter (EKF). 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 [2] or [3]. The state variables which are estimated by the EKF can be differentiated in two kinds:

- The shape parameters of the model, which are the horizontal and vertical curvature, the changes in horizontal and vertical curvature, the lane width and change in lane width along the lane.
- The 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 in a manner, that the lookahead distance in 3D-space ranges from 6 to 40m, and for the tele-image from 30 to 100m. 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 whether a vehicle obscures the expected lane markings. To do this, the bounding box for each OV is tested whether it intersects with any search window. If an intersection exists, the search window is clipped (figure 4). If the resulting search path is too short this measurement is disabled.

Figure 4: Clipping of 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, which 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 where reflections with a similar distance, relative velocity and amplitude are grouped together. The radar system creates a list of potential vehicles every 60msec. 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 all radar measurements to it which lie within it (details see [4]). 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 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 neither by radar nor 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 3D 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, speed and acceleration is estimated via EKF for each OV (for details see [4]).

9 Vehicle of Relevance

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

The pure radar-based ACC system (performance level one) 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. See figure 5.

Figure 5: 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 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. E.g. 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 6 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. At far distances a relevance decision based on the driving tube would definitely be wrong.

Figure 6: Driving tube and lane model while performing a lane change

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 10m 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. See figure 7a. 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.

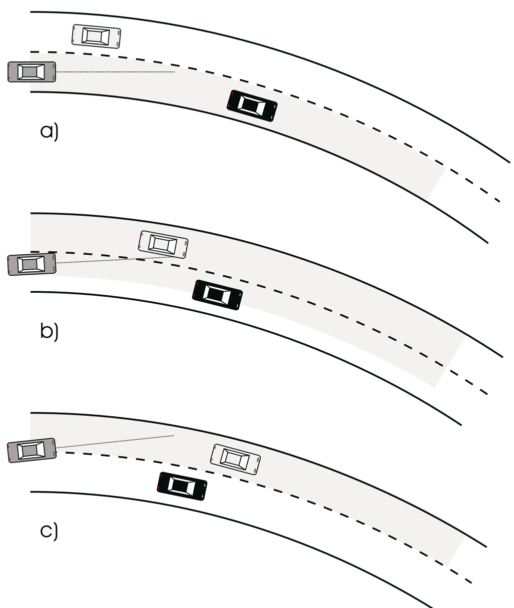

Figure 7: Relevance decision area as function of the horizontal offset

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. That means, 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 function of the horizontal offset. See figure 7b. 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 RDA´s trajectory 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. See figure 7c. 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 other vehicle is communicated to the longitudinal vehicle controller adjusting the speed of the Ego in a way that the convoy distance is larger than a desired value, e.g. $1.6\text{sec} \cdot v_{\text{relevant\_OV}}$.

10 Conclusions and Outlook

It has been shown that 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 performance mode 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 easily be performed using the knowledge about the position of the Ego relative to the own lane. Convoy driving using activated lateral control is possible with this system if the speed of the leading car is sufficently large such that it is not suppressed by radar measurement preprocessing. For speeds slower than that, first experiments are being made using trinocular stereovision for handling stop-and-go traffic (for details see [4]).

References

[1] R. Gregor, M. Lützeler, M. Pellkofer, K. H. Siedersberger, and E. D. Dickmanns. EMS-Vision: A perceptual system for autonomous vehicles. In this volume.

[2] E. D. Dickmanns. Dynamic computer vision for mobile robot control. In *Proc. 19th Int. Symp. and Expos. on Robots*, Sydney (Australia), November 1988.

[3] Reinhold Behringer. Visuelle Erkennung und Interpretation des Fahrspurverlaufs durch Rechnersehen für ein autonomes Straßenfahrzeug. In *Fortschrittsberichte*, volume 310. VDI Verlag, Düsseldorf, Germany, 1996.

[4] A. Rieder. *Fahrzeuge Sehen*. PhD thesis, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, (to appear).