2015 IEEE Intelligent Vehicles Symposium (IV)
June 28 - July 1, 2015. COEX, Seoul, Korea

Autonomous Driving at Ulm University: A Modular, Robust, and Sensor-Independent Fusion Approach

Felix Kunz, Dominik Nuss, Jürgen Wiest, Hendrik Deusch, Stephan Reuter, Franz Gritschneder, Alexander Scheel, Manuel Stübler, Martin Bach, Patrick Hatzelmann, Cornelius Wild, and Klaus Dietmayer

Abstract— **The project "Autonomous Driving" at Ulm University aims at advancing highly-automated driving with close-to-market sensors while ensuring easy exchangeability of the particular components. In this contribution, the experimental vehicle that was realized during the project is presented along with its software modules. To achieve the mentioned goals, a sophisticated fusion approach for robust environment perception is essential. Apart from the necessary motion planning algorithms, this paper thus focuses on the sensor-independent fusion scheme. It allows for an efficient sensor replacement and realizes redundancy by using probabilistic and generic interfaces. Redundancy is ensured by utilizing multiple sensors of different types in crucial modules like grid mapping, localization and tracking. Furthermore, the combination of the module outputs to a consistent environment model is achieved by employing their probabilistic representation. The performance of the vehicle is discussed using the experience from numerous autonomous driving tests on public roads.**

I. Introduction

Since the early 1990s, research on autonomous driving has made an enormous progress. Thus, several highly automated driver assistance functions are close to mass production. One of the first projects on autonomous driving was the European PROMETHEUS project [1]–[3] with the autonomous drive from Munich, Germany, to Odense in Denmark in 1995. During the last decade, the Defense Advanced Research Projects Agency (DARPA) organized several challenges for autonomous vehicles. The first challenges in 2004 and 2005 concentrated on off-road vehicles (see e.g. [4]). With the DARPA Urban Challenge, the focus moved to autonomous driving in urban areas. Most successful competitors employed a multi-sensor setup to handle interactions with other road users in the proximity. Laser range finders and radar sensors were commonly used [5], [6] while computer vision played a secondary role. Highway platooning scenarios using vehicle-to-vehicle (V2V) communication have been investigated in the Grand Cooperative Driving Challenge, e.g. [7].

Another widely known project is Google's self driving car which relies on a roof-mounted Velodyne laser range finder and high-precision digital maps. Further impressive demonstrations include the autonomous drive of Broggi's team from Italy to China [8] and their experiment in urban traffic [9]. Both demonstrations heavily rely on computer vision using stereo cameras. In 2013, Daimler presented

All authors are with the Institute of Measurement, Control, and Microtechnology, Ulm University, Ulm, Germany {firstname.lastname}@uni-ulm.de

Fig. 1. The autonomous experimental vehicle.

Bertha [10], an autonomous S-Class equipped with close-to-market sensors. By driving autonomously from Mannheim to Pforzheim on the Bertha Benz memorial route, Daimler researchers impressively demonstrated their expertise. The chosen route comprises urban streets and rural roads, both including traffic lights, roundabouts and pedestrian crossings. The perception system was realized using a stereo camera, two mono cameras, and several short and long-range radar sensors [10]–[12].

Considering aspects of functional safety, future autonomous cars require a multi-sensor setup for environment perception providing the essential redundancy to handle sensor faults. A further advantage of a multi-sensor setup is the possibility to combine the strengths of different sensor types, e.g. the accurate distance and velocity information of a radar sensor and the precise bearing measurement of a camera. For autonomous driving, another key competence of future environment perception systems is the possibility to replace sensors without changing the core of the fusion system. There are two obvious reasons for sensor replacement: On the one hand the integration of a new sensor may enhance the reliability of the system or increase the field of view (FOV). On the other hand the substitution by a cheaper sensor (e.g. with a reduced FOV) facilitates the transfer of autonomous driving functions from luxury cars to smaller cars at reasonable cost.

In contrast to high level track-to-track fusion used, e.g., in [10] and many other autonomous vehicles, the proposed environment perception module is realized using a centralized sensor fusion system. Compared to track-to-track fusion, the required bandwidth for data transmission is significantly increased. However, centralized sensor fusion provides more accurate tracking results since no information is lost [13].

The proposed modular multi-sensor information fusion system for environment perception ensures redundancy and replaceability using sensor-independent probabilistic interfaces. A predecessor of the current system was proposed in [14] where a multi-sensor object tracking system was realized with the Joint Integrated Probabilistic Data Association (JIPDA) filter [15] and the replacement of a sensor simply requires the implementation of a sensor-specific probabilistic measurement model. In this contribution, a Labeled Multi-Bernoulli Filter [16] is used for object tracking instead of the JIPDA algorithm. Furthermore, the principle of sensor-independent probabilistic interfaces is also applied to the localization algorithm as well as the grid map representation of the static environment. Due to modularity and generic interfaces, the proposed concept ensures a reduced but guaranteed performance of the perception system in case of temporary malfunction or degradation of individual sensors. As a first route to test the autonomous vehicle, a 5 km long loop near the university campus was chosen. Although it is quite short, this highly demanding route comprises several traffic lights, roundabouts and pedestrian crossings. Especially the area around the university hospital is challenging due to heavy traffic during visiting hours.

This contribution is organized as follows: First, the system setup of our autonomous car is introduced in Section II. Section III presents the details of the proposed information fusion system. Decision making, motion planning and control of actuators are summarized in Section IV. Finally, performance of the system as well as possible directions for further improvements are discussed.

II. System Setup

To build a vehicle that reacts autonomously, several sensors and computers need to be integrated. All of them have to work together seamlessly. This section briefly describes the system setup of our autonomous vehicle which is shown in Fig. 1.

A. Hardware Setup

The automated vehicle is a Mercedes-Benz E-Class which has been modified to enable autonomous driving functionalities. A Paravan Space-Drive steering system is mounted to the steering column to obtain the ability to steer by wire. A modification of the ESP (electronic stability program) control unit enables to directly specify brake and acceleration torque via CAN (controller area network) bus. The desired torques are computed by a closed-loop controller on a dSpace real-time system. Additionally, the dSpace system is responsible for low-level operations and sensor triggering which ensures time synchronized measurements. Two computers are installed in the trunk (running Ubuntu 12.04 64 Bit) which are responsible for high-level functions. All three computing systems are connected.

B. Sensor Setup

The experimental vehicle is equipped with several close-to-market sensors. Special care has been taken that the sensors are well integrated in the body and are hardly recognizable. Three IBEO LUX laser scanners are installed in the front bumper covering a total azimuth angle of roughly 210°. Each laser scanner has a horizontal angular resolution of 0.125°, a radial resolution of 0.04 m and a maximum range of up to 200 m. The laser scanners are synchronized and time triggered with a frequency of 12.5 Hz. A forward facing monochrome camera (Baumer TXG14f) with a resolution of 1392 x 1040 px and a horizontal FOV of about 56° is mounted behind the windshield. The camera is synchronized and time triggered with a frequency of 15 Hz. Additionally, the vehicle is equipped with several radar sensors. A Frequency Modulated Continuous Wave (FMCW) based Continental ARS 310 automotive long-range radar is mounted underneath the grille. It features a range from 0.25 m to 200 m in the far field with an angular resolution of 1° (range of 0.25 m to 60 m with an angular resolution of 4° in near field) and runs with a frequency of 15 Hz. The complete sensor coverage is visualized in Fig. 2.

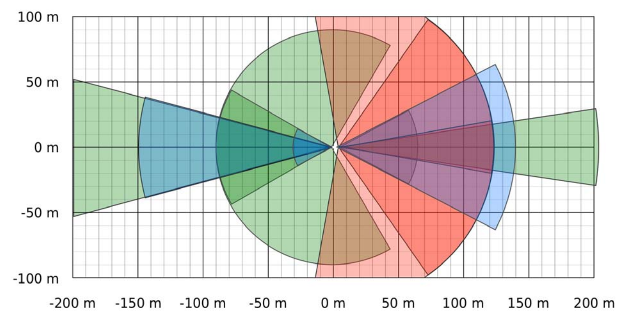


Fig. 2. Sensor coverage of the automated vehicle: cameras (blue), laser scanners (red), radar sensors (green)

To observe the rear area of the vehicle, two Bosch MRR rear radar sensors are mounted behind the rear bumper, rotated ±45° w.r.t. the longitudinal axis. With a very wide aperture angle of up to 150 deg and a range of up to 90 m they cover the whole area behind the vehicle. Additionally, a Bosch LRR 3 FMCW long-range radar is mounted centered behind the rear bumper. It features an aperture angle of 30° and provides objects up to 250 m. Two rearward facing cameras are mounted behind the rear window. A Baumer TXG14f with horizontal FOV of about 20° is used for vehicle tracking. Further, a Baumer TXG06 with a resolution of 776 x 582 px and a horizontal FOV of about 56° is applied for localization. Both cameras are synchronized and time triggered with a frequency of 15 Hz.

A real-time kinematic (RTK) system in combination with a differential GPS is used for mapping (see Section III-B and [17]) and evaluation.

III. Environment Perception

Reliable environment perception is essential for driver assistance systems and autonomous vehicles. The utilization

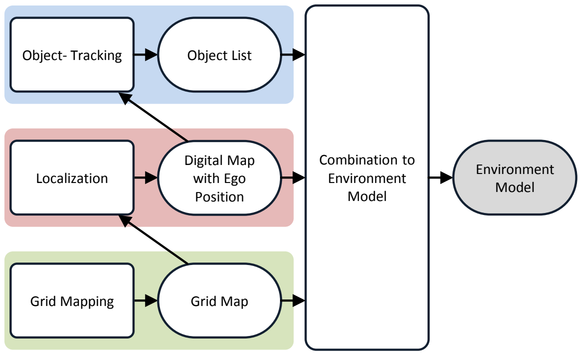


Fig. 3. Perception System Overview

of a multi-sensor system is highly recommended since every sensor has its assets and drawbacks. Consider the following example for illustration: A radar sensor delivers very accurate range and velocity measurements. However, the measurement uncertainty of the yaw angle is relatively high and a reliable classification of the measured objects is often not possible. On the other hand, a mono camera provides an accurate yaw angle information and a reliable classification at the cost of imprecise range and missing velocity information. Thus, a probabilistic fusion of radar and video data improves the estimation accuracy by compensating the weaknesses of one sensor with the strengths of another. Observe that a direct coupling between two or more sensors should not be used since this prohibits, e.g., sensor replacement.

The presented concept is a hierarchical modular environment perception system (HMEP) [18]. It is divided into three perception layers: grid mapping, localization and object tracking (see Fig. 3). Each layer produces an element of the environment model fusing data of all available sensors. Thus, sensor fusion is applied in each layer individually with the objective to estimate the corresponding element of the environment. Generally, perception layers are allowed to use the results of other layers in addition to sensor data. To avoid loops, layers are ordered hierarchically and only results of lower layers may be used as additional input data in each perception module. The order increases with the abstraction level of the environment model element. Results of the different layers can be redundant or even contradictory. Therefore, a combination module combines all elements to a consistent environmental model.

To facilitate a mathematically well-founded combination of the modules, the output of the three layers has to be probabilistic, i.e., the output has to incorporate uncertainty information. Further, the individual layers are required to provide a generic sensor interface which enables the incorporation of an additional sensor or the replacement of an existing one without changing the fusion core of the perception layer. Finally, the functionality of a perception layer may not rely on a specific sensor type or even a specific sensor.

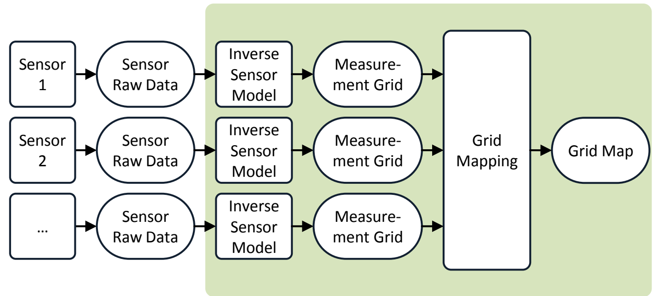


Fig. 4. Grid Mapping System

A. Grid Mapping

The lowest layer is the grid mapping layer. It divides the environment into single grid cells and estimates the occupancy state of each cell according to the classical occupancy grid mapping approach [19], [20]. In the implemented system, data of the three laser scanners is used as input. Classical occupancy grids are not eligible for estimation of dynamic environments, since they do not incorporate a process model. Therefore, in the implemented system the grid map has a fading memory and quickly adapts to the current situation. Further, the combination module uses the grid map information mainly for estimating stationary boundaries. New sensors can be integrated into the grid mapping layer by creating an inverse sensor model for the sensor as pre-processing filter (see Fig. 4) [21]. The inverse sensor model estimates the occupancy probability of each grid cell based on a sensor measurement. This is called a measurement grid. The grid mapping algorithm updates the grid map with the measurement grid using the binary Bayes-filter [20]. The result is the estimated occupancy probability of each grid cell in the vehicle environment.

B. Localization

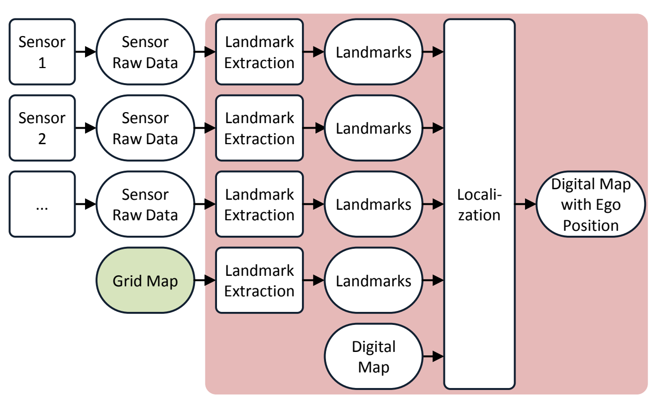
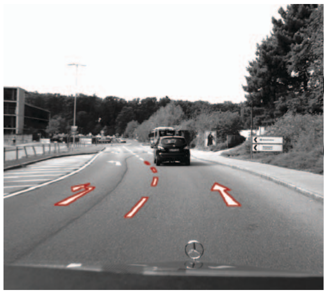
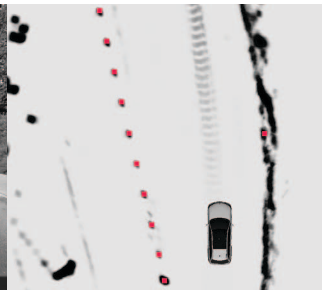


Fig. 5. Localization System

Figure 5 shows a schematic overview of the generic localization module used for the automated test vehicle. The vehicle's pose is estimated by making use of the concept of Monte-Carlo Localization (MCL) on a feature-based map representation [22], [23]. Features are extracted from the grid map built with three front-facing laser scanners [24]



(a) Image



(b) Grid map

Fig. 6. Detected features in the image (lane markings) and grid map (wooden posts) using the MSER algorithm

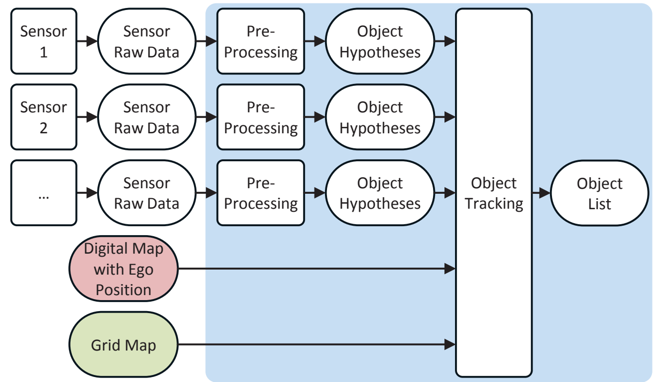


Fig. 7. Tracking System

using Maximally Stable Extremal Regions (MSERs) [25]. MSERs are robust features that are additionally extracted from grayscale images of a front-facing camera within a pre-defined region of interest (ROI) [17]. Features found in the grid map consist of tree trunks, road signs, posts, etc. and deliver well-suited point landmarks for the localization process. Features from the camera mainly consist of lane markings. Figure 6 shows an exemplary scene with MSERs found in the grid map and the corresponding video image. The point landmarks that are found in the grid map (red dots in Figure 6b) correspond to the wooden posts on the left roadside (cf. Figure 6a). Landmarks are represented by a UTM easting and northing coordinate for the MSER centroids as well as an area estimate. In case of video detections, the landmark positions are obtained by applying the flat world assumption. The fact that only a quite small feature vector is stored in the database instead of a whole grid map (cf. for example [**?**]) reduces the amount of data drastically. The database holds about 400 landmarks or approximately 22 kB of memory per kilometer. During an unsupervised mapping process the observed landmarks are tracked (in order to reduce clutter) and stored in a geospatial database. Therefore, a high-precision RTK system is utilized that provides a global accuracy of up to 2 cm. For the task of self-localization, the pose $\hat{\boldsymbol{p}} = [\,E\;N\;\psi\,]^{\mathrm{T}}$ with UTM easting, northing and yaw angle is estimated using a particle filter which incorporates online measurements and the pre-built map. Particles are predicted with the dead reckoning approach based on the on-board odometry sensors that deliver an estimation for the current velocity and yaw rate of the ego vehicle. In the innovation step, the weight of the particles is updated using a likelihood function [17] which evaluates the correspondence of the current sensor measurements with the expected measurements. Therefore, longitudinal and lateral position measurement uncertainties are specified for each sensor. The expected measurements are given by the relative position of the landmarks in the digital map with respect to the particle's state. Finally, the estimated vehicle pose is given by the weighted mean of the particles. Further details and evaluation results can be found in [17], especially regarding the calculation of the likelihood function, which is the core element of the localization algorithm.

Due to the generic approach, additional sensors (e.g. radar) can be added to the system quite easily by deriving appropriate features. Likewise, the absence of one of the sensors can be handled: A missing laser scanner decreases the longitudinal precision, while a loss of camera measurements results in a larger lateral uncertainty.

C. Multi-Object Tracking

The perception of dynamic objects in the environment is realized using a generic centralized multi-sensor information fusion module. Figure 7 illustrates the realization of the vehicle tracking module using radar, laser and video sensors. To reduce computational complexity, pre-processing algorithms are applied to obtain object hypotheses from the sensor's raw data. The radar simply delivers a list of (possible) target measurements, the raw data of the laser range finder is clustered using a box fitting algorithm [26] and the cascaded classifier presented in [27] is used to detect vehicles in the video image. In addition to the measurement vector and the measurement noise, some of the sensors provide confidence values (or the true positive probability). Apart from the sensor measurements, the fusion module may utilize information from the grid map and/or the digital map to obtain a spatial birth probability or prior information like, e.g., maximum velocity or possible orientations.

The fusion module is realized using the Labeled Multi-Bernoulli (LMB) filter [16], [28] which is an efficient approximation of the multi-target Bayes (MTB) filter [29]. The LMB filter outputs an object list containing the spatial distribution and the existence probability of each track. Since the LMB filter incorporates the estimation of the track labels in the filtering algorithm, it does not require the heuristic post-processing algorithms used in other approximations of the MTB filter (e.g. the Probability Hypothesis Density filter [30], [31]) to obtain the tracking result. Making use of the accurate update equations of the Generalized Labeled Multi-Bernoulli (GLMB) filter [32], the LMB filter obtains excellent results in challenging scenarios with closely-spaced objects where the track-to-measurement association is often ambiguous. In contrast, the performance of multi-instance Kalman filters degrades significantly in such scenarios due to heuristic track-to-measurement association.

In the proposed tracking system, each sensor is connected to the LMB filter using a sensor-specific measurement model providing measurement functions, field of views, detection probabilities and false positive probabilities. The integration of additional sensors does not require any changes to the LMB filter itself, it only requires the implementation of sensor-specific measurement models. Additionally, the sensor's perception capabilities are modeled using the Dempster-Shafer approach presented in [33] which, e.g., facilitates an integrated handling of the increased number of missed detections in video images during bad weather conditions.

Tracking of other vulnerable road users like bikes or pedestrians is implemented by a second LMB filter instance using laser and video measurements. Obviously, different pre-processing algorithms have to be applied here which require adapted measurement models. The different application and the adapted sensor setup do not involve any changes to the LMB fusion module. Thus, this example already indicates the benefits of a generic multi-sensor information fusion module.

D. Environment Model

The combination module combines the grid map, the localization result, the digital map and the object list to a consistent environment model (see [18] for details) as depicted in Fig. 3. Summarized, the combination module replaces the grid map with boundary lines, which limit the drivable space of the vehicle (see Fig. 8). The boundary lines represent stationary obstacles. Although dynamic objects may also lead to occupied cells in the grid map, they should not be considered when creating the boundary lines. This is achieved with a probabilistic approach. The combination module calculates for each grid cell the probability for three exclusive events:

- The grid cell is free
- The grid cell is occupied and corresponds to an object included in the object list
- The grid cell is occupied by an obstacle not included in the object list

The calculation takes into account the occupancy probability of the grid cell as well as the state estimate (with according uncertainty) and the existence probability of object tracks. Based on a certain threshold, the boundary lines exclude all obstacles from the drivable corridor which are not included in the tracking list. Note that this is only possible because grid map and object tracking module represent the environment in a probabilistic manner. The idea of the combination module is to achieve a conservative fallback behavior. Consider the following example: If a leading car is precisely tracked and sharply represented in the grid map, there will be no grid cell with a high probability for being occupied by a stationary or unknown object. As a result, the boundary lines are open and the leading car is solely represented in the object list. The ego vehicle can follow the leading vehicle with an appropriate distance. If the tracking module misses the leading vehicle or the estimated existence probability of the leading vehicle is very low, the boundary lines will close. Consequently, the ego vehicle will keep a greater distance to the leading vehicle in order to be able to come to a full stop at the leading vehicle's position. As soon as the tracking algorithm catches the leading vehicle, the boundary lines open again. Thus, the test vehicle will not collide with objects missed by the tracking algorithm while ensuring an appropriate following behavior, if the leading vehicle is precisely tracked. Obviously, this resembles human behavior by driving faster in case of high knowledge about the environment and by reducing the speed with increasing uncertainty.

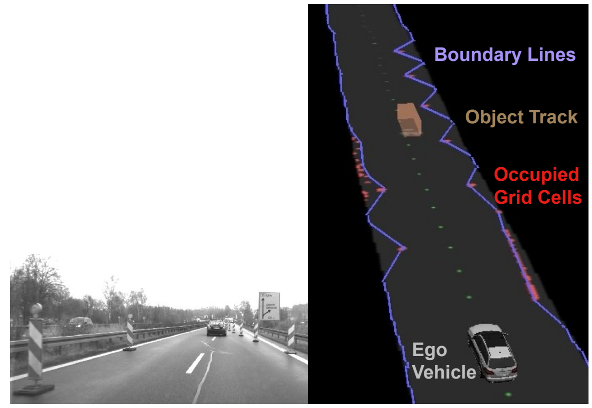


Fig. 8. Visualization of the environment model. Boundary lines (purple) exclude occupied grid cells (red) from the drivable space only if they do not originate from tracked objects (orange).

IV. Motion Planning

A. Overview

The objective of the motion planning module is the generation of a safe vehicle trajectory which offers a high level of driving comfort. Figure 9 depicts the integration of the motion planning module into the system architecture. The fact that the system is evaluated on a static course renders the need for a separate navigation-layer superfluous. Instead, the route is defined by a routing-file containing a vector of lane IDs. However, the interface allows for a simple integration of a dynamic routing algorithm. The situation analysis module interprets the environment model and creates long-term predictions for other vehicles. This is accomplished by the calculation of track-to-lane relations and the prediction of all tracked vehicles along the lanes they are associated with. In case that a track is associated with more than one lane, a prediction is generated for each of them. The predicted environment model contains all information of the environment model as well as the long-term predictions. The decision making module analyzes this model to infer permitted maneuvers at certain passages of the route which require an additional evaluation of traffic rules. Moreover, the digital map contains virtual stopping points preventing the vehicle from entering roundabouts, intersections or crosswalks. These stopping points may be cleared solely by the decider through corresponding maneuver permissions. The trajectory generated by the motion planning module is passed to the vehicle control module which consists of a separate longitudinal and lateral controller. Longitudinal control is implemented as a state space controller whereas lateral control follows the approach described in [34].

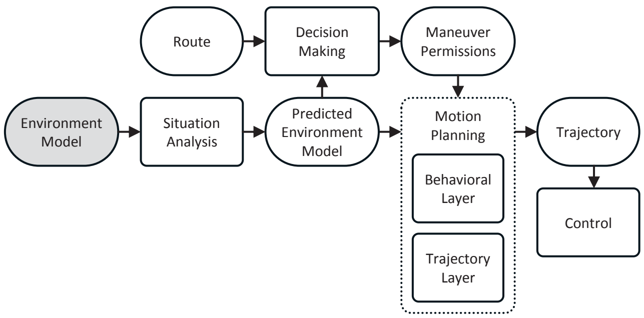


Fig. 9. Motion Planning System

B. Behavioral Layer

In order to mimic human driving the ego vehicle's motion is expressed as a composition of longitudinal and lateral motion w.r.t. the course of the road as described in [35]. In this context, longitudinal motion corresponds to the movement along the center line of the ego vehicle's lane. The composition can be stated using a Frenét frame in which the longitudinal motion corresponds to the movement along the abscissa and lateral motion to the movement along the ordinate. An advantage of this approach is that the trajectory planner's objectives for lateral and longitudinal motion w.r.t. the center line may be set independently. The behavioral layer utilizes this convenient interface to generate the desired driving behavior: Depending on the current situation it chooses the appropriate strategy to keep a target velocity, follow a leading vehicle or reach a specific state e.g. to stop at a certain point. Moreover, the behavioral layer may release constraints on the lateral targets to allow for entering the opposite lane to circumnavigate an obstacle. As the trajectory planning algorithm has a limited planning horizon of 3 seconds, the behavioral layer is responsible for assuring driving safety over a longer period of time e.g. by restricting the maximum velocity.

C. Trajectory Planning

The trajectory planning concept follows the optimal control approach described in [35]. An important requirement is a comfortable driving behavior which is strongly related to the jerk. As a consequence, the time integral of the squared jerk is chosen as the Lagrange term of the objective function. A key idea is to assure temporal consistency in the optimal solution. This means a subsequent planning step should yield a trajectory which is consistent with the remainder of the previously calculated one. This requirement is met by creating a set of optimal solutions to the unrestricted optimization problem for different target states and times, disregarding all obstacles and constraints on the kinematics. It can be shown that the optimal solutions belong to the function class of quintic polynomials [36]. Constraints are regarded in a second step, where all solutions which fail any of the restrictions are discarded and the remaining solution with the lowest cost is chosen.

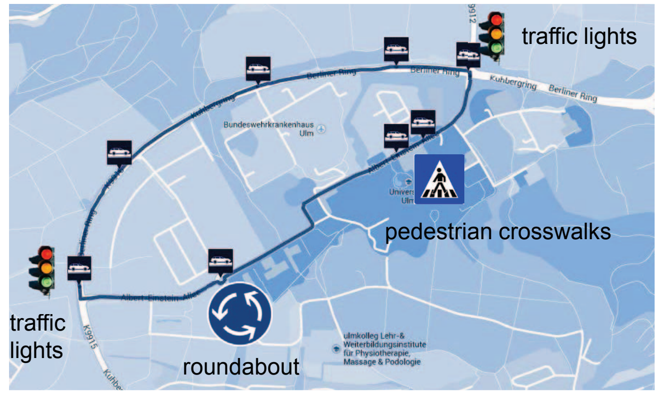


Fig. 10. Route around the campus of Ulm University.

The complete trajectory planning algorithm comprises the following steps: First of all, the ego vehicle's state is transformed to Frenét coordinates. The planner then creates separate trajectory sets in Frenét coordinates for both lateral and longitudinal target states defined by the behavioral layer. These trajectories are expressed in the form of quintic polynomials. Longitudinal and lateral trajectory sets are combined and the resulting trajectories are transformed back to UTM coordinates. During this transformation, all trajectories are checked against the violation of constraints such as maximum velocity, acceleration, curvature and curvature rate ensuring their feasibility. Moreover, collision checks are performed against static and dynamic obstacles as described in [37]. The remaining trajectories are evaluated w.r.t. their accumulated cost and the best trajectory with the lowest cost is passed to the vehicle control.

V. Results

The autonomous car of Ulm University with the proposed probabilistic information fusion system has been presented in a press event in July 2014. Figure 10 shows the route including the campus of Ulm University. Although the route is only 5 km long, it includes a variety of challenges for autonomous vehicles: traffic lights, crosswalks, and roundabouts. Additionally, half of the track does not contain any lane markings. The huge amount of pedestrians and other vulnerable road users in the area around the university hospital, as well as vehicles suddenly turning into the ego vehicle's driving path (especially during rush hour), pose a challenge for the environment perception. The speed limit on the course varies between 50 and 70 km/h. A video providing some impressions can be found at [38]. Additional informations about the route, vehicle, and team as well as press reports are available at [39].

A. Performance evaluation

The proposed information fusion system of our autonomous vehicle provides a robust and reliable environment perception. The combination of grid map and object tracking successfully detects all road users and other obstacles in the vehicle's environment where the presented fallback concept of the environment model ensures a correct behavior of the vehicle. Using a constant turn rate and acceleration model for vehicle tracking leads to smooth object trajectories and facilitates a smooth adaptive cruise control mode. The mean computation time for an update of the LMB filter with the measurements of one sensor is slightly above 2 ms for challenging urban scenarios [40] using a single core of an Intel i7-2600 CPU. Even on winding rural roads, the mean state estimation error for the leading vehicle is below 30 cm at distances between 30 to 100 m.

The localization module delivers a highly accurate estimation of the ego position. Relying on features of video and laser sensors, standard GPS measurements are only required for initialization. Due to the probabilistic multi-sensor approach, the localization module is capable of getting along with missing lane markings in some parts of the road. The localization accuracy of the sensor setup is evaluated in [17], the mean values of longitudinal and lateral error are below 0.09 m with a standard deviation of less than 0.17 m. The mean orientation error is $0.01°$ with a standard deviation of $0.228°$. The superior performance of the fusion approach compared to video only or laser only localization is also illustrated in [17]: Using only video sensors, the longitudinal error increases due to utilization of the flat world assumption and pitching of the ego vehicle. In contrast, the lateral error slightly increases in the laser only system. However, the accuracy of single sensor localization is still sufficient if enough landmarks for the particular sensor setup are present.

The motion planning algorithm was triggered at a cycle time of 150 ms which has shown to be sufficient to safely react to dynamic changes in the environment. Trajectory planning in combination with vehicle control has shown to generate a smooth, comfortable driving behavior. In order to increase driving safety, lateral acceleration has been restricted to a maximum of 2 m/s^2. The limitation of the longitudinal acceleration depends on the current speed, allowing higher acceleration values at low velocities. The generated trajectories have a length of 3 s and are sampled at 20 Hz. The behavioral layer has proven to generate safe long-term objectives for the trajectory planner leading to a defensive driving style which avoids critical situations. Moreover, the anticipatory nature of the decision making module ensures correctness with regard to traffic rules such as the right-of-way at the roundabout and the consideration of traffic lights and crosswalks.

B. Challenging environment conditions

The consequent application of several sensors with different characteristics enables an outstanding performance of the system in challenging environment conditions like sunset, darkness, or rain. In the current system setup, the conditions mainly influence the performance of the high resolution video camera which is optimized for industrial applications. Due to the probabilistic fusion approach, the decreased performance of the camera only reduces the accuracy of the environment model while ensuring the reliability of the system. Since object classification in the tracking system is mainly based on video information and object dynamics, the classification accuracy diminishes. Furthermore, the uncertainty of the objects' lateral positions slightly increases due to the absence of the precise bearing measurements. Due to missed detections of the video landmarks, the localization module automatically relies on the laser measurements. As shown in [17], the mean value of the lateral localization error using only laser measurements is only 0.12 m which is still sufficient for providing autonomous driving functions.

VI. Conclusion

This contribution presented system setup, probabilistic sensor-independent fusion approach, and motion planning module of the autonomous vehicle at Ulm University. Test drives in public traffic illustrated the performance and robustness of the proposed system. The modular architecture facilitates the exchangeability of the sensors. Due to the probabilistic sensor-independent interfaces used in grid mapping, localization, and tracking, a modification of the sensor setup does not require any adaption to the fusion algorithms. The sensor setup covers the most significant areas of the vehicle's environment with at least two sensors. Fusing the obtained sensor information in a probabilistic way, the environment perception system exploits the redundancy and complementarity of the sensor setup. Motion Planning has shown to cope well with the challenges of the route, generating a safe and comfortable driving behavior. The modular structure and plain interfaces allow for a convenient adaption to future driving scenarios.

In order to demonstrate the sensor-independence of the proposed fusion approach, a novel autonomous vehicle equipped with another sensor setup is in preparation. Besides the integration of additional sensors, an extension of the grid map to estimate dynamic environments is desirable. The main reason is that a process model enables the grid map to predict its state into the future improving the results of the Bayes filter. Further improvements of the object tracking module are possible by using extended object tracking algorithms [41]–[43] which incorporate the preprocessing or clustering of the sensor raw data into the tracking algorithm. However, these algorithms require further research to reduce computational cost and to facilitate the incorporation of different sensor types. A drawback of the presented localization approach (but also of the approaches used in other popular autonomous vehicles) is the need for a highly-precise digital map that is built in advance. A more sophisticated approach is learning the map online in order to deal with a dynamic and ever changing environment, e.g. using Simultaneous Localization and Mapping (SLAM) [20], [44].

REFERENCES

[1] E. Dickmanns, B. Mysliwetz, and T. Christians, "An integrated spatio-temporal approach to automatic visual guidance of autonomous vehicles," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 20, no. 6, pp. 1273–1284, 11 1990.

[2] U. Franke, S. Mehring, A. Suissa, and S. Hahn, "The daimler-benz steering assistant: a spin-off from autonomous driving," in *Proceedings of the Intelligent Vehicles Symposium*, 11 1994, pp. 120–124.

[3] E. Dickmanns, R. Behringer, D. Dickmanns, T. Hildebrandt, M. Maurer, F. Thomanek, and J. Schiehlen, "The seeing passenger car 'vamors-p'," in *Proceedings of the Intelligent Vehicles Symposium*, 11 1994, pp. 68–73.

[4] S. Thrun *et al.*, "Stanley: The robot that won the darpa grand challenge," *Journal of Field Robotics*, vol. 23, no. 9, pp. 661–692, 2006. [Online]. Available: http://dx.doi.org/10.1002/rob.20147

[5] C. Urmson *et al.*, "Autonomous driving in urban environments: Boss and the urban challenge," *Journal of Field Robotics Special Issue on the 2007 DARPA Urban Challenge, Part I*, vol. 25, no. 8, pp. 425–466, 6 2008.

[6] S. Kammel, J. Ziegler, B. Pitzer, M. Werling, T. Gindele, D. Jagzent, J. Schrder, M. Thuy, M. Goebl, F. v. Hundelshausen, O. Pink, C. Frese, and C. Stiller, "Team annieway's autonomous system for the 2007 darpa urban challenge," *Journal of Field Robotics*, vol. 25, no. 9, pp. 615–639, 2008. [Online]. Available: http://dx.doi.org/10.1002/rob.20252

[7] A. Geiger, M. Lauer, F. Moosmann, B. Ranft, H. Rapp, C. Stiller, and J. Ziegler, "Team annieway's entry to the 2011 grand cooperative driving challenge," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 13, no. 3, pp. 1008–1017, 9 2012.

[8] A. Broggi, M. Buzzoni, S. Debattisti, P. Grisleri, M. Laghi, P. Medici, and P. Versari, "Extensive tests of autonomous driving technologies," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 3, pp. 1403–1415, 9 2013.

[9] A. Broggi, P. Cerri, S. Debattisti, M. Laghi, P. Medici, M. Panciroli, and A. Prioletti, "Proud-public road urban driverless test: Architecture and results," in *IEEE Intelligent Vehicles Symposium Proceedings*, 6 2014, pp. 648–654.

[10] J. Ziegler *et al.*, "Making bertha drive - an autonomous journey on a historic route," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 2, pp. 8–20, 2014.

[11] U. Franke, D. Pfeiffer, C. Rabe, C. Knoeppel, M. Enzweiler, F. Stein, and R. Herrtwich, "Making bertha see," in *IEEE International Conference on Computer Vision Workshops (ICCVW)*, 12 2013, pp. 214–221.

[12] J. Dickmann, N. Appenrodt, H.-L. Bloecher, C. Brenk, T. Hackbarth, M. Hahn, J. Klappstein, M. Muntzinger, and A. Sailer, "Radar contribution to highly automated driving," in *11th European Radar Conference*, 10 2014, pp. 412–415.

[13] Y. Bar-Shalom, P. K. Willett, and X. Tian, *Tracking and Data Fusion*. Storrs, CT, USA: YBS Publishing, 2011.

[14] M. Munz, M. Mählisch, and K. Dietmayer, "Generic centralized multi sensor data fusion based on probabilistic sensor and environment models for driver assistance systems," *IEEE Intelligent Transportation Systems Magazine*, vol. 2, no. 1, pp. 6–17, 2010.

[15] D. Musicki and R. Evans, "Joint integrated probabilistic data association: JIPDA," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, no. 3, pp. 1093–1099, 7 2004.

[16] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietmayer, "The labeled multi-Bernoulli filter," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3246 – 3260, 2014.

[17] H. Deusch, J. Wiest, S. Reuter, D. Nuss, M. Fritzsche, and K. Dietmayer, "Multi-sensor self-localization based on maximally stable extremal regions," in *IEEE Intelligent Vehicles Symposium*, 6 2014, pp. 555–560.

[18] D. Nuss, M. Stuebler, and K. Dietmayer, "Consistent environmental modeling by use of occupancy grid maps, digital road maps, and multi-object tracking," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, June 2014, pp. 1371–1377.

[19] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *Computer*, vol. 22, no. 6, pp. 46–57, 6 1989.

[20] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, 2005.

[21] D. Nuss, M. Thom, A. Danzer, and K. Dietmayer, "Fusion of laser and monocular camera data in object grid maps for vehicle environment perception," in *International Conference on Information Fusion (FUSION)*, July 2014, pp. 1–8.

[22] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Monte carlo localization for mobile robots," in *IEEE International Conference on Robotics and Automation (ICRA)*, vol. 2, 1999, pp. 1322–1328.

[23] P. Jensfelt, D. Austin, O. Wijk, and M. Andersson, "Feature based condensation for mobile robot localization," in *IEEE International Conference on Robotics and Automation (ICRA)*, vol. 3, 2000, pp. 2531–2537.

[24] J. Wiest, H. Deusch, D. Nuss, S. Reuter, M. Fritzsche, and K. Dietmayer, "Localization based on region descriptors in grid maps," in *Intelligent Vehicles Symposium (IV), IEEE*, 2014, accepted for publication.

[25] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," in *In British Machine Vision Conference*, 2002, pp. 384–393.

[26] M. Munz, K. Dietmayer, and M. Mählisch, "A sensor independent probabilistic fusion system for driver assistance systems," in *Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems, St. Louis, MO, USA*, 2009.

[27] M. Gabb, O. Löhlein, R. Wagner, A. Westenberger, M. Fritzsche, and K. Dietmayer, "High-performance on-road vehicle detection in monocular images," in *Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems*, 2013, pp. 336–341.

[28] S. Reuter, "Multi-object tracking using random finite sets," Ph.D. dissertation, Ulm University, 2014.

[29] R. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Artech House Inc., Norwood, 2007.

[30] ——, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, 10 2003.

[31] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, 11 2006.

[32] B.-T. Vo and B.-N. Vo, "Labeled random finite sets and multi-object conjugate priors," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3460–3475, 2013.

[33] D. Meissner, S. Reuter, E. Strigel, and K. Dietmayer, "Intersection-based road user tracking using a classifying multiple-model PHD filter," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 2, pp. 21–33, 2014.

[34] R. Mayr, *Regelungsstrategien für die automatische Fahrzeugführung*. Heidelberg: Springer, 2001.

[35] M. Werling, J. Ziegler, S. Kammel, and S. Thrun, "Optimal trajectory generation for dynamic street scenarios in a frenét frame," in *IEEE International Conference on Robotics and Automation*, 5 2010, pp. 987–993.

[36] M. Werling, "Ein neues Konzept für die Trajektoriengenerierung und -stabilisierung in zeitkritischen Verkehrsszenarien," Ph.D. dissertation, Karlsruher Institut für Technologie, 2011.

[37] J. Ziegler and C. Stiller, "Fast collision checking for intelligent vehicle motion planning," in *IEEE Intelligent Vehicles Symposium (IV)*, 6 2010, pp. 518–522.

[38] Ulm University. (2014) Driving autonomously in ulm (video). [Online]. Available: http://youtu.be/R44p69-UHDg

[39] ——. (2014) Project homepage autonomous driving. [Online]. Available: https://www.uni-ulm.de/in/automatisiertes-fahren

[40] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietmayer, "Multi-object tracking using labeled multi-Bernoulli random finite sets," in *Proceedings of the 17th International Conference on Information Fusion*, 2014, pp. 1–8.

[41] J. W. Koch, "Bayesian approach to extended object and cluster tracking using random matrices," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, Issue 3, pp. 1042–1059, July 2008.

[42] K. Granström, S. Reuter, D. Meissner, and A. Scheel, "A multiple model PHD approach to tracking of cars under an assumed rectangular shape," in *Proceedings of the 17th International Conference on Information Fusion*, 2014.

[43] C. Lundquist, K. Granström, and U. Orguner, "An extended target CPHD filter and a Gamma Gaussian Inverse Wishart implementation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 472–483, 2013.

[44] M. Adams, J. Mullane, and B.-N. Vo, "Circumventing the feature association problem in SLAM," *Intelligent Transportation Systems Magazine, IEEE*, vol. 5, no. 3, pp. 40–58, 2013. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6565468