

Fusion of Occupancy Grid Mapping and Model Based Object Tracking for Driver Assistance Systems using Laser and Radar Sensors

Mohamed Essayed Bouzouraa and Ulrich Hofmann

Abstract—in this paper we present a novel environment perception system based on an occupancy grid mapping and a multi-object tracking. The goal of such a system is to create a harmonic, consistent and complete representation of the vehicle environment as a base for future advanced driver assistance systems. In addition to a mathematical formulation of the problem we present a robust algorithm to detect dynamic obstacles from the occupancy map and show how both, the mapping process and the tracking can benefit from each other. Therefore, the concept of moving objects with associated dynamic cells is introduced. The presented techniques are applicable to both 2D and 3D mapping and can be also extended to correct the ego motion from the occupancy map and the object tracks. Unlike many publications over the last years our work provides real time performance and an accurate detection of obstacles with real laser and radar sensors and can fulfill the requirements of future driver assistance systems.

I. INTRODUCTION

An accurate and complete environment perception is essential to develop advanced driver assistance systems. These systems require the complete detection of static and dynamic obstacles, the shape of the obstacles, free spaces and unexplored areas. The precise estimation of ego motion is also an important requirement in order to correctly interpret the data of the sensors. Many systems only focus on an object based description of the environment making assumptions in form of geometric and dynamic models. Such a description provides abstract data about the environment of the vehicle and can be implemented very efficiently. On the other hand the assumptions made by these approaches can result in an erroneous perception of the environment. They also suffer from the necessity of continually solving the data association problem between measurement data and the detected objects. An other reliable and established method describing the occupied, free and unknown parts of the vehicle environment is the occupancy grid representation known from robotics [1]. This technique is particularly convenient if scanning sensors like laser are used. The occupancy map nevertheless requires high computing resources and has a low abstraction level due to its data

structure which is composed from probability cells. Considering the different advantages of both methods (occupancy map and object tracking), we present in this paper a concept linking these both worlds in order to create a harmonic and complete environment perception as represented in figure 1. The map and the object tracking have a common interface composed from the list of the tracked objects and the corresponding cells in the occupancy grid. This is possible by providing model information about moving obstacles from the tracker and by correcting the object tracking with data from the occupancy map. We think that this is the right way to satisfy the increasing requirements of future driver assistance systems on the perception systems and their sensors.

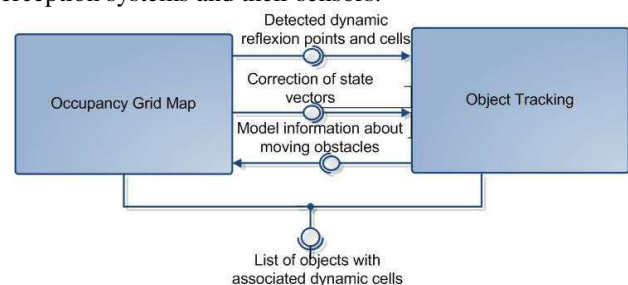


Fig. 1. Communication between the object tracking and the occupancy map

Over the last years, few works tried to simultaneously solve the SLAM problem (Simultaneous Localization And Mapping) and the DATMO (Detection And Tracking Of Moving Objects). In [2] and [3] an approach is presented to tackle the SLAM problem and the DATMO problem at once. The map is used to detect moving objects which are filtered from the map in order to achieve more accurate pose estimation. Unfortunately the mapping quality does not benefit from this approach because the motion of the moving objects is not compensated in the Map. Additionally the two problems are separated so that no communication between the map and the object tracker is possible. In [4] a similar approach is described using additionally short range radar sensors to detect moving objects but not to update the map. A major difference to our work is that these methods classify only the laser scan measurements as static or dynamic. We classify in our work both measurements and cells from the map depending on their dynamic status. In [5] [6] and [7] the authors demonstrate a new filtering technique called Bayesian Occupancy Filter (BOF). This technique can

Mohamed Essayed Bouzouraa is with the Advanced Development Driver Assistance Systems, AUDI AG, Germany, e-mail: essayed.bouzouraa@audi.de.

Ulrich Hofmann is with the Advanced Development Driver Assistance Systems, AUDI AG, Germany, e-mail: ulrich.hofmann@audi.de

estimate the position and the velocity of cells. Unfortunately it requires a velocity measuring sensor and very high computation time. Another drawback of this method is the absence of an abstract object based description. So it is necessary to extract objects from the map which costs further computation time.

The main contributions of this work are the robust detection of moving obstacles using an occupancy map and the techniques necessary so that both mapping and object tracking can benefit from each other. We also show how to use the radar measurement to update the map.

The paper is organized as follows: In section II a mathematical problem formulation of the mapping process expanded by the dynamic information of the cells and using the bayes probabilistic is given. We also discuss the different steps of the presented concept. Details about the object tracking and the feedback from the map are shown in section III. Finally experimental results of both object tracking and mapping with real sensor data and a reference system representing the ground truth are demonstrated.

II. OCCUPANCY GRID MAPPING WITH DETECTION OF MOVING OBJECTS

A. Problem formulation

In this section we give a mathematical formulation of the mapping process with detection of moving objects and consideration of cells motion belonging to these objects. This will provide a solid base to understand the problem and make the appropriate approximations to achieve a real time implementation. We use the following random variables to describe the problem.

- 1) M_k : Map at time k .
- 2) $\mathbf{z}_{1..k}$: Vector of sensor measurements until time k
- 3) O_k^x : Occupancy probability of the cell with coordinates \mathbf{X} .
- 4) \mathbf{X}_k : Position of the cell at time k .
- 5) $\mathbf{X}_{k-1..k}^{\text{ego}}$: Change of ego data motion between the steps $k-1$ and k .
- 6) D_k^x : Random variable describing the dynamic state of the cell \mathbf{X} at time k . $D_k^x \in \{'static', 'dynamic'\}$

The goal of the mapping process is to calculate the probability $P(O_k^x, \mathbf{X}_k | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1})$.

Since $P(O_k^x | \mathbf{X}_k) = P(O_k^x)$, we can write the following equation:

$$P(O_k^x, \mathbf{X}_k | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}) = P(O_k^x | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}) \cdot P(\mathbf{X}_k | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}) \quad (1)$$

The first probability can be calculated by applying the bayes rule and making the Markov assumption.

$$P(O_k^x | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}) \propto P(\mathbf{z}_{1..k-1} | O_k^x, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}) \cdot P(O_k^x | \mathbf{z}_{1..k-1}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}) \quad (2)$$

This corresponds to the classical grid map update. The second probability is calculated by introducing the variable D_k^x and using the total probability law.

$$P(\mathbf{X}_k | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}) = \sum_{D_k^x} P(\mathbf{X}_k | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}, D_k^x) \cdot P(D_k^x | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}) \quad (3)$$

$$P(\mathbf{X}_k | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}) = P(\mathbf{X}_k | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}, D_k^x = 'static') \cdot$$

$$P(D_k^x = 'static' | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}) + P(\mathbf{X}_k | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}, D_k^x = 'dyn.') \cdot$$

$$P(D_k^x = 'dyn.' | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1})$$

This equation shows that it is necessary to calculate the probability $P(D_k^x | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1})$ of dynamic state of the cell given the last map and the sensor measurements in order to determine the probability of the new cell position. If the cell is static, the new cell position is only dependent on the change in the ego data motion. Considering a dynamic cell, we can introduce the position \mathbf{X}_{k-1} at the step $k-1$ by using again the total probability law.

$$P(\mathbf{X}_k | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}, D_k^x = 'dyn.') = \int P(\mathbf{X}_k | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1}, D_k^x = 'dyn.', \mathbf{X}_{k-1}) \cdot P(\mathbf{X}_{k-1} | \mathbf{z}_{1..k-1}, \mathbf{X}_{k-2..k-1}^{\text{ego}}, M_{k-2}, D_{k-1}^x = 'dyn.') d\mathbf{X}_{k-1} \quad (4)$$

The last integral corresponds to a prediction step and requires high computing costs to calculate an approximation for it. By making the assumption that two dynamic cells can not reach the same position in the future, we can efficiently calculate the predicted positions of all dynamic cells considering only one previous position for each dynamic cell.

The whole formulation can be extended to additionally correct the ego motion data from the map and sensor measurements. In this case the aim of the mapping process is to calculate the probability $P(O_k^x, \mathbf{X}_k, \mathbf{X}_{k-1..k}^{\text{ego}} | \mathbf{z}_{1..k}, M_{k-1})$.

B. Dynamic Detection

The aim of the dynamic detection algorithm is to calculate the probability $P(D_k^x | \mathbf{z}_{1..k}, \mathbf{X}_{k-1..k}^{\text{ego}}, M_{k-1})$. For simplicity we only determine the binary information whether the cell is static or not. Our algorithm is based on the comparison between the laser sensor raw data and the occupancy map build from past sensor measurements and ego motion data. The idea of detecting moving objects by observing a difference between the laser reflexion points and the map is a common method used in [4], [2] and [9]. We extend these algorithms to assure a robust and fast dynamic detection by modeling the source of errors. Since we use a laser scanner close to series production, the sensor beams have a field of view up to 2° . As consequence the laser measurement can not be considered as one reflexion point. So we model this fuzziness by a line as represented in figure 2. The free space of the beam is modeled by a triangle.

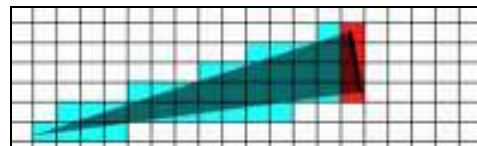


Fig. 2. Modeling the laser sensor beam on the occupancy map

Because the ego motion data, the sensor measurements and the discretization of the map are sources of errors making the comparison unreliable, we extend the region of examination in the map. This operation corresponds to a dilatation of a binary image and is mathematically equivalent to the calculation of the Minkowski sum of a region R and a mask S [12].

$$R \oplus S = \{r + s \mid r \in R, s \in S\} \quad (1)$$

The method can be extended to 3D mapping by using a mask S composed of $3 \times 3 \times 3$ voxels whereas the examined voxel is in the center.

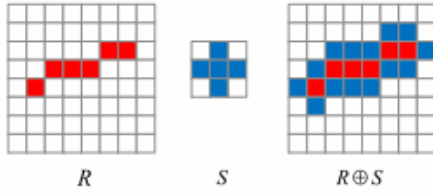


Fig. 3. Dilatation of a rasterized line

After the projection of the new laser data on the map, the detection process is accomplished in two steps and according to certain rules shown in table I.

Firstly if an occupied cell is detected within the dilated area of the measurement, this sensor observation and the corresponding cells are classified as static. If there are other unknown and free cells in the examined area, the sum over the cells of area is calculated and a decision about the dynamic state of measurement is done with the help of a threshold value.

The second step of the dynamic detection is done while updating the map with free space information from the sensor. Map cells which are interleaved by the sensor beams are classified as old dynamic cells. This information will help us later to classify measurements with unknown dynamic state as moving ones if they can be associated to these cells.

TABLE I
DETECTION OF DYNAMIC CELLS AND LASER REFLEXION POINTS

		Occupancy map		
		free	unknown	occupied
Measurement	occupied	Dynamic measurement	Potential dynamic measurement	Static measurement
	free	Free cell	Free cell	Old dynamic cell

Figure 4 shows some results from the first part of the dynamic detection algorithm applied to e.g. 3D laser raw data. The red points are belonging to the ground plane. The green points are classified as dynamic. The blue ones are potential dynamic and must be verified with the free space

information from the sensor.

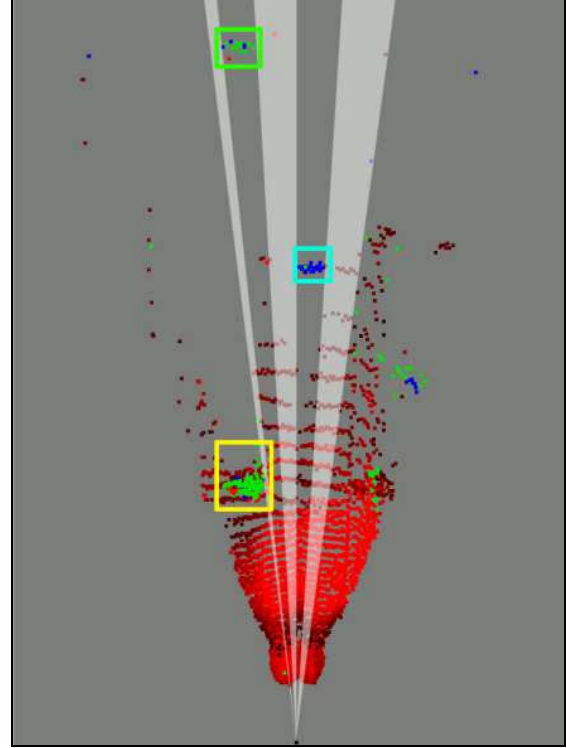


Fig. 4. Results from the dynamic detection algorithm

C. Map update from radar sensor

In addition to the 2D or 3D laser range sensors, we use a long range radar providing distance and velocity measurements, but no information about the orientation, length and width of objects. The main advantage of the radar sensor is the availability of a velocity measurement thanks to Doppler principle. In our concept we can associate the velocity measurement provided by the radar with the laser scanner measurements by using for example the nearest neighbor method. If a radar measurement can not be associated with the laser data, or if the radar is used as stand alone sensor, the map is updated by an appropriate radar sensor model. In order to represent vehicles in the map with the radar measurements, we did several reference measurements. By knowing the exact position and shape of a target vehicle and comparing it with the distance and velocity values of the radar, we can statically determine the uncertainties of the radar measurement in the x- and y-directions as represented in figure 5.

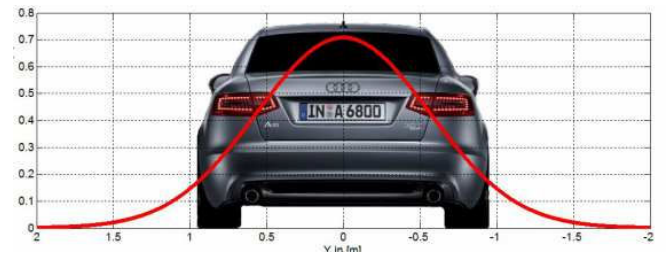


Fig. 5. The probability density function of the lateral radar measurement

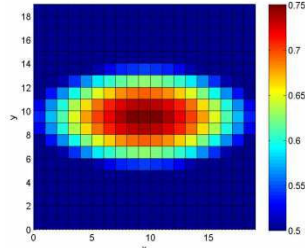


Fig. 6. The two dimensional probability generated by the radar sensor model

The radar reflexion center can be described by a two dimensional Gaussian distribution as mentioned in equation 5 and represented in figure 6. x_{radar} and y_{radar} describe the radar measurements. σ_x^2 and σ_y^2 are the variances of the measurements. The orientation of the ellipse corresponds to the motion direction of the object, since the radar is not able to measure the object heading. In order to save computation time the probability distribution is discretized and rotated offline in a look-up-table for different angles.

$$f(x, y) = 1/2\pi \cdot 1/\sigma_x \cdot 1/\sigma_y \cdot e^{-\left(\frac{(x-x_{radar})^2}{2\sigma_x^2} + \frac{(y-y_{radar})^2}{2\sigma_y^2}\right)} \quad (5)$$

D. Clustering and Object Extraction

The aim of the clustering and the object extraction is to group the 2D or the 3D laser data and cells classified as dynamic in a first step. The grouping algorithm uses a distance criterion to merge the points. Figure 7 shows results from this grouping algorithm. The grouped points are represented in green.

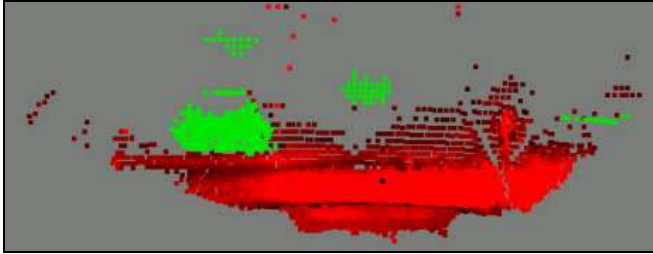


Fig. 7. Results from the grouping algorithm

In a second step a state vector and the geometric form of the object are extracted. We use the algorithm presented in [13] and we extend it to segment 3D data. The approach transforms laser scan measurements into a list of linear segments as represented in figure 8. An L-shape detection additionally allows to determine the orientation of the object.

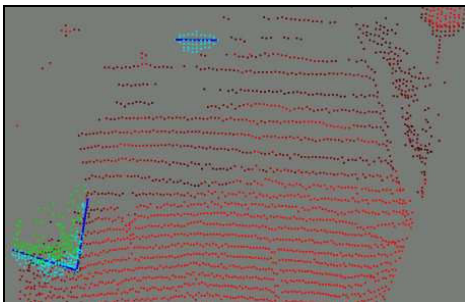


Fig. 8. Results from the segmentation algorithm

III. MULTI-OBJECT TRACKING

A. Object state vectors with dynamic cells

In order to demonstrate the successful feasibility of our concept we use a simple Multi Target Tracking (MTT) [13] with Global Nearest Neighbor (GNN) [10] to associate the measurements with the objects tracked. These methods can be simply upgraded to enhance the performance of the object tracker by introducing for example the Interacting Multiple Model [11] and the Multi Hypothesis Tracking (MHT) approaches. The state vector $x^T = \left(x, \dot{x}, \ddot{x}, y, \dot{y}, \ddot{y}\right)^T$ is

estimated by a Kalman filter based on a constant acceleration model. Equations 6 and 7 show the system and measurement matrices. If the radar is directly used to update the tracking, its velocity measurement is additionally integrated.

$$A = \begin{pmatrix} 1 & \Delta T & 1/2 \Delta T^2 & 0 & 0 & 0 \\ 0 & 1 & \Delta T & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \Delta T & 1/2 \Delta T^2 \\ 0 & 0 & 0 & 0 & 1 & \Delta T \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (6)$$

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \quad (7)$$

The width and length of the object are estimated by a low pass filter. In order to reduce errors from the segmentation algorithm the orientation is updated from a measurement only if the difference to the motion direction is lower as a certain threshold.

Our new idea is to extend the notion of the object states classically composed from its motion data (position, velocity) and its geometric data (e.g. length and width) by introducing the list of associated and accumulated cells from the occupancy map (see figure 9). This new approach will open up new possibilities to improve the objects tracking. The cells associated with the state vector of an object contain their occupancy probability and allow a more precise and complete description of the object geometric shape accumulated over the past as represented in figure 13 for the pick-up truck. This kind of information is not necessarily available in the state vectors and can be lost due to assumptions made in the models (e.g. a bounding box model). Furthermore the probabilities of the cells belonging to one object offer a control mechanism to the models used in the estimation of the object state. If the majority of object cells have a low probability, the motion model and the geometric model used to track this object may be improper to it. Furthermore a cell management is responsible to maintain the cells with high occupancy probability values within the list of associated cells and delete those which are not confirmed from the cell list.

All these ideas will help us to achieve a very robust tracking in highly dynamic areas with occlusion and with unknown object types, even if errors from environment sensors and ego motion data exist.

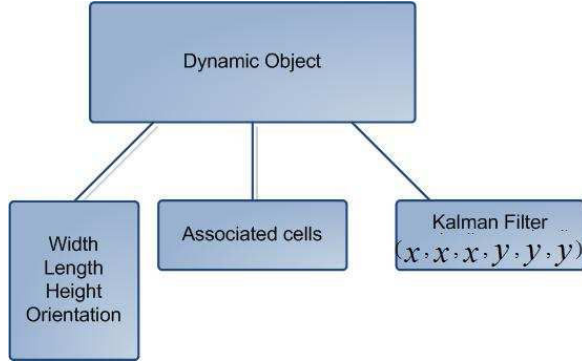


Fig. 9. Representation of a dynamic object

B. Prediction of dynamic cell with subcell calculation

A decisive step in the mapping process is the prediction of the dynamic cells on the map according to the state vectors of the objects associated with these cells. This would enhance the map quality and prevent the trail caused by dynamic objects (see the left part of figure 11). A simple motion model considering the rotation of the object is used to calculate the predicted position of each cell (figure 10). We maintain in parallel a list of these dynamic cells belonging to tracked objects with a reference to the occupancy map. So we can implement the prediction very efficiently and without overwriting other dynamic cells. After the prediction step, the map is updated with the new laser measurement using the bayes rule.

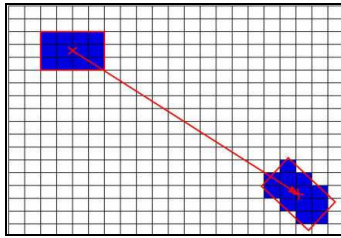


Fig. 10. Prediction of the dynamic cells associated with a moving object
To assure a better accumulation of the dynamic cells we introduce uncertainties in the prediction process by setting the neighborhood of a dynamic cell with a low occupancy probability. In order to achieve a better accuracy than the resolution of the occupancy grid allows, we calculate the exact position of the obstacle in the cell. By storing this exact position we can detect very small motion of the objects. We call this technique subcell calculation.

C. Feedback from the occupancy map

In order to maintain the consistency between the occupancy grid map and the object tracking, we introduce a feedback step from the map. This corresponds to a measurement update in the object tracking after the map is updated. In this case the occupancy map is considered as a virtual object sensor for the object tracking. After the prediction of the

dynamic cells with model information from the tracking, a raw data laser update is accomplished on the map. The difference between the content of the map and the object tracking resulting from errors in the dynamic models of the objects is used to correct the object tracking.

IV. RESULTS

In this section results from the occupancy map and the object tracking are presented. The occupancy grid map has a width of 140m and a length of 140m with a resolution of 20cm x 20cm. In order to study the accuracy of our algorithm we use a reference system composed from a high accurate IMU (Inertial Measurement Unit) with DGPS (Differential Global Positioning System). Both ego and target vehicles are equipped with this system. Additionally a Wireless local Area Network communication allows to synchronize and to record the data. The generated ground truth is then used to test the algorithms by projecting the exact shape of the target vehicle on the map (see figure 11) and by directly comparing the reference data with the state vector data from the tracking. Figure 11 shows the improvement of the map by tracking the target object and predicting the dynamic cells on the map. Here the shape of the target vehicle is correctly represented and the trail caused by the missing compensation of the objects motion disappears from the map. Figure 12 demonstrates the performance of the algorithms at high speeds and with oncoming vehicle. Thanks to radar sensor the vehicle in front of the target vehicle is also detected.

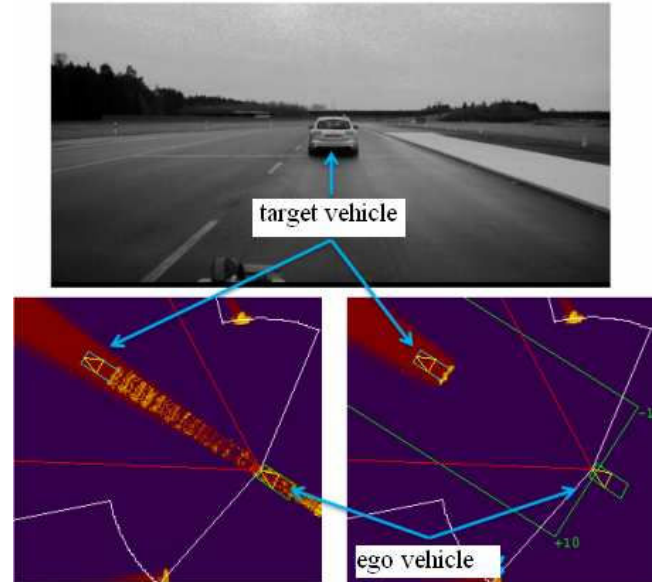


Fig. 11. Improvement of the 2D occupancy map by predicting and updating the dynamic cells. (Left without compensation of dynamic cells, right with compensation of dynamic cells)

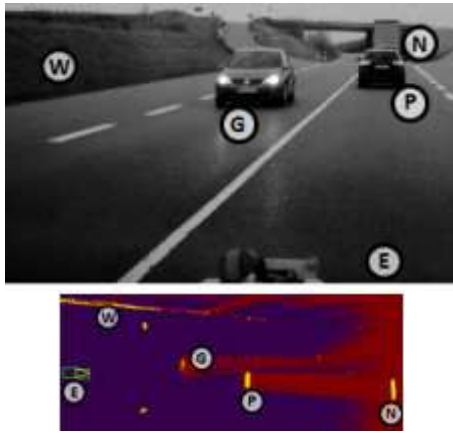


Fig. 12. 2D occupancy grid map in a scene with high speeds and oncoming traffic

The 3D map and the tracking provide very good results in different scenes although here the radar sensor is not used and the velocity can not be measured directly. In the different figures the static part of the environment is represented by the red voxels and the dynamic part is represented by the blue voxels. Besides the occupancy probability is coded by the transparency of the voxels and the height by the darkness. In Figures 17 and 18 the position and the velocity signals from both tracking and ground truth are represented for a high dynamic scene.

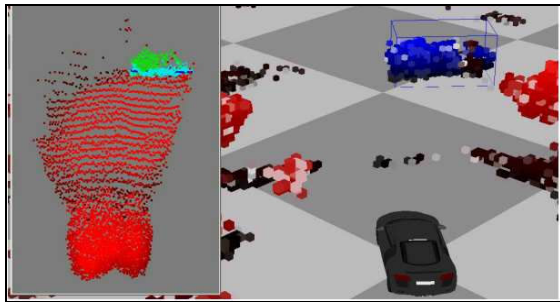


Fig. 13. Detection and tracking of a moving crossing vehicle

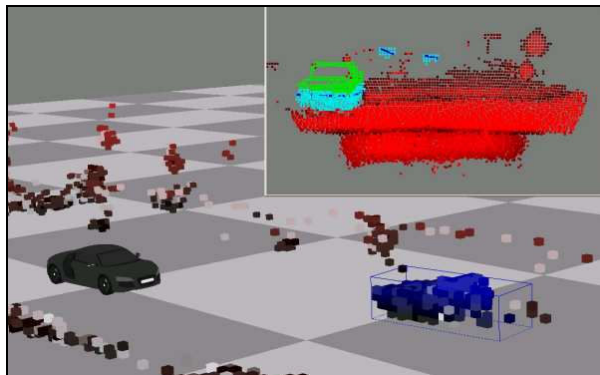


Fig. 14. Tracking of a pick-up vehicle with the bounding box model and the associated dynamic cells

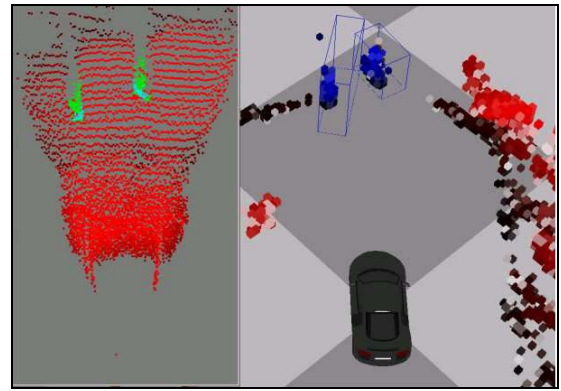


Fig. 15. Detection and tracking of pedestrians

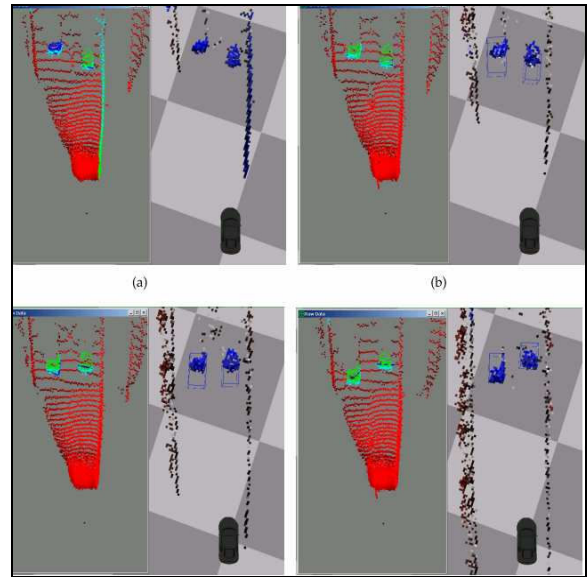


Fig. 16. Convergence of the state vectors of tracked objects

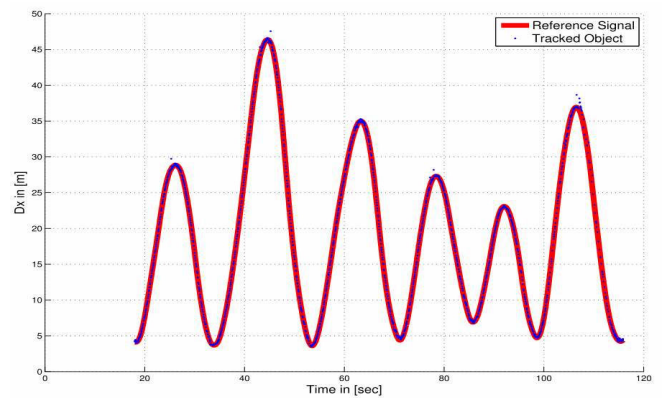


Fig. 17. Comparison of the distance signal from a tracked object and the corresponding reference signal from ground truth

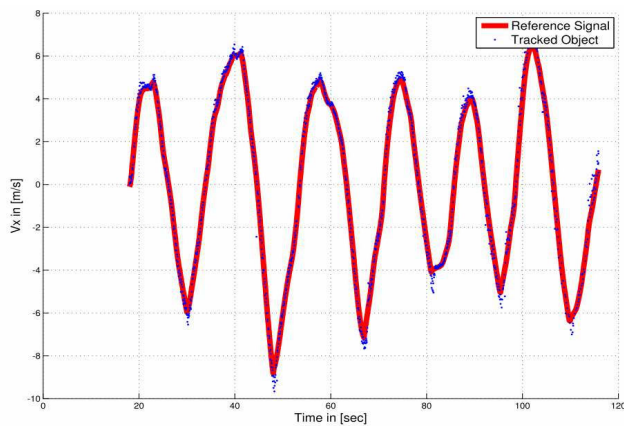


Fig. 18. Comparison of the velocity signal from a tracked object and the corresponding reference signal

V. CONCLUSION AND FURTHER WORK

We presented in this paper a novel approach linking the occupancy map representation and the model based object tracking which is applicable for 2D and 3D environment perception for driver assistance systems. The approach uses laser range and radar sensor and can work with one or both sensors. A robust algorithm to detect moving obstacles from laser measurements and the occupancy grid map is shown. Both the map and the tracking benefit from these methods since the classical object state representation is extended by associated dynamic cells. These cells are predicted and updated on the map. Results with real sensor data and a reference system show that our algorithms are robust and very accurate. The 2D and 3D versions work both in real time on a laptop with 2.1GHz processor. In future work we will focus on improving the object detection and tracking by using better segmentation algorithms and by integrating the orientation in the dynamic model of the Kalman filter. We will also extend the concept by separately describing the dynamic cells which can not be associated to objects and by correcting the whole object state from the map feedback.

REFERENCES

- [1] Elfes A., Using occupancy grids for mobile robot perception and navigation. *IEEE Computer* 22(6), 46–57
- [2] Chieh-Chih Wang and Chuck Thorpe, Simultaneous Localization and Mapping with Detection and Tracking of Moving Objects, *IEEE International Conference on Robotics and Automation*, 2002.
- [3] Chieh-Chih Wang, Charles Thorpe and Arne Suppe, Ladar-Based Detection and Tracking of Moving Objects from a Ground Vehicle at High Speeds, *IEEE Intelligent Vehicles Symposium*, 2003.
- [4] Trung-Dung Vu, Olivier Aycard and Nils Appenrodt, Online Localization and Mapping with Moving Object Tracking in Dynamic Outdoor Environments, *IEEE Intelligent Vehicles Symposium*, Istanbul, Turkey, 2007.
- [5] C. Coué, C. Pradalier and C. Laugier, Bayesian Programming Multi-Target Tracking : an Automotive Application, *International conference on Field and Service Robotics*, 2003.
- [6] Christophe Coué, Cédric Pradalier, Christian Laugier, Thierry Fraichard and Pierre Bessière, Bayesian Occupancy Filtering for Multitarget Tracking: An Automotive Application, *The International Journal of Robotics Research*, Vol. 25, No. 1, pp. 19-30, January 2006

- [7] Kamel Mekhnacha1, Yong Mao, David Raulo1 and Christian Laugier, Bayesian occupancy filter based “Fast Clustering-Tracking” algorithm, *International Conference on Intelligent Robots and Systems*, 2008
- [8] Hess, J. M, Characterizing Dynamic Objects in 3D Laser Range Data, Albert-Ludwigs-Universität Freiburg, Faculty of applied Sciences, Department of Computer Science, Autonomus Intelligent sytem Laboratory, Master thesis, 2008.
- [9] Petrovskaya A. , Thrun S. , Model Based Vehicle Tracking in urban environments, *proceedings of the 2009 IEEE International Conference on robotics and automation*, 2009.
- [10] Pavlina Konstatinova, Alexander Udvarov and Tzvetan Semerdjiev, A Study of a target Tracking Algorithm Using Global Nearest Neighbor Approach, *International Conference on Computer Systems and Technologies*, 2003.
- [11] Blom, “The interacting multiple model algorithm for systems with markovian switching coefficients,” *IEEE Transactions on Automatic Control*, vol. 33, pp. 780–783, 1988.
- [12] Steger, Carsten, Markus Ulrich and Christian Wiedemann, Machine Vision Algorithms and Applications, *Wiley-VCH*, Weinheim, 2007
- [13] Mendes, A., L.C. Bento and U. Nunes., Multi-target detection and tracking with a laser scanner, *Proceedings of the 2004 IEEE Intelligent Vehicle*, 796– 801, June 2004