2012 15th International IEEE Conference on Intelligent Transportation Systems
Anchorage, Alaska, USA, September 16-19, 2012

Using Grid Maps to Reduce the Number of False Positive Measurements in Advanced Driver Assistance Systems

Dominik Nuss[1], Stephan Reuter[1], Marcus Konrad[2], Michael Munz[3], Klaus Dietmayer[1]

***Abstract*— In Advanced Driver Assistance Systems (ADAS), object tracking is a crucial method to foresee dangerous situations. The Joint Integrated Probabilistic Data Association (JIPDA) offers the advantage, that existence and association uncertainties are considered in multi-target tracking. Recently, real-time capable implementations have been presented. However, the real-time capability is only given, if a certain number of tracked objects is not exceeded. Thus, so called false positive object detections yield a problem. To mitigate this issue, additional information about the vehicle's environment is used to identify measurements that are not relevant. The idea is to focus on moving objects for tracking. As an example, an Occupancy Grid Map is used to distinguish between stationary and non-stationary objects. The approach is evaluated using real-world data of a research vehicle.**

I. Introduction

In recent years, the number of traffic accidents declined due to driver assistance systems, that have become standard in automotive applications. In order to further improve these systems, it will be necessary to predict dangerous situations that arise from interactions with traffic participants in the immediate environment. Hence, object tracking is an essential ability of future advanced driver assistance systems (ADAS). One promising approach for object tracking in ADAS is the Joint Integrated Probabilistic Data Association Filter (JIPDA) [10], which allows an integrated state and existence estimation. In [9], a real-time capable graph-based implementation of JIPDA is described. However, this method is still limited to a certain number of objects that can be tracked. This number exceeds the number of traffic participants, that are relevant in typical traffic scenarios. Still, problems arise due to false positive measurements, that can outnumber the possible number of objects to be tracked. This contribution presents a method to reduce false positive measurements of a laser range finder by using a grid map. Grid maps are a discrete representation of the environment that have their origin in robotic applications [14] but have also established in the ADAS field [18], [4]. Several proposals for environmental representation based on a combination of grid mapping and object tracking have been presented in recent years. [3] combines grid mapping and object tracking in order to provide an extensive environmental model. In [17] and [15], moving objects are detected and filtered out of a stationary occupancy grid map. In [1], the movement state of an object is considered in a dynamic occupancy grid map.

The grid map used in this paper is based on a method proposed in [4]. The method removes moving objects, producing an occupancy grid map holding only stationary objects similar to [17] and [15]. In this contribution, such a grid map is used in a preprocessing step in order to distinguish between measurements of stationary and moving objects when the same route is used a second time. This way, the number of objects that have to be tracked by the JIPDA filter can be reduced.

This paper is organized as follows: In section II and III, grid maps and JIPDA filtering are reviewed. Section IV explains how stationary objects are identified and filtered in JIPDA-tracking, which is evaluated in section V. Finally a conclusion and an outlook is given in section VI.

II. Grid Mapping

A. Occupancy Grids

Occupancy grid maps describe the probability of areas to be occupied.

1) Grid Map Definition: In a grid map ${}^{G}\mathbf{M}$, the environment is divided into n grid cells ${}^{G}c_i$, $1 \leq i \leq n$. A common approach is to create 2-dimensional maps with a constant resolution. Each cell holds one or more values $v({}^{G}c_i)$ of grid cell data. These values represent arbitrary information about the corresponding region of the environment. In occupancy grids, $v({}^{G}c_i)$ stands for the occupancy likelihood of the corresponding cell area, which is a value between 0 (free) and 1 (occupied). In some approaches the grid is defined relatively to the coordinate system of the vehicle [16], which forces the mapping algorithm to rotate the map when the vehicle rotates. Weiss [18] proposes to avoid the rotation of the map by applying only the integral part of the translation to the map and exerting the floating point part and the rotation on the vehicle.

2) Grid Map Calculation: This approach was proposed in [6] and is based on [18]. Assuming a laser measurement arrives at every time step t, the following algorithm is processed (see Figure 1):

a) Estimation of the vehicle's position and orientation: In this paper, the position and orientation $\boldsymbol{x}_t$ of the vehicle is estimated by a highly precise GPS System with a positioning error of less than $0.1\,\mathrm{m}$.

b) Transformation rule: In this step, the algorithm creates a temporary grid with cells ${}^{MG}c_i$ holding values $v_t({}^{MG}c_i)$, that result from one measurement sample $\boldsymbol{z}_t$. Such a grid is called measurement grid ${}^{MG}\mathbf{M}_t$. This step can be

[1]D. Nuss, S. Reuter and K. Dietmayer are with the Institute of Measurement, Control, and Microtechnology, Ulm University, Germany, firstname.lastname at uni-ulm.de

[2]M. Konrad is with EvoBus GmbH - Daimler Buses, Neu-Ulm, Germany, marcus.konrad at daimler.com

[3]M. Munz is with the Department Environment Perception, Daimler AG, Ulm, Germany, michael.mi.munz at daimler.com

further divided in two parts. The first part transforms the laser points into the measurement grid. Each point is uncertain in its position caused by an inaccurate measurement of distance and angle on the one hand and errors based on the estimation of the movement of the vehicle on the other hand. These uncertainties are modeled as a two-dimensional Gaussian function. The second part models the likelihoods in free regions using a certain freespace function. The cells of the measurement grid correspond in their position to certain cells ${}^{G}c_i$ of the grid map ${}^{G}\mathbf{M}$. The maximum distance between a measurement and the vehicle is limited by the maximum range parameter, which depends on the used sensor and determines the size of the measurement grid.

c) Grid Map Update: The update algorithm updates the grid map ${}^{G}\mathbf{M}_{t-1}$ with a measurement grid ${}^{MG}\mathbf{M}_t$ by combining the values of the corresponding cells ${}^{MG}c_i$ and ${}^{G}c_i$. This step is based on the Binary Bayes Filter as stated in equation (1). See [5] for more details.

$$v_t({}^{G}c_i) = \frac{S}{1+S} \tag{1}$$

$$\text{with } S = \frac{v_t({}^{MG}c_i)}{1 - v_t({}^{MG}c_i)} \cdot \frac{v_{t-1}({}^{G}c_i)}{1 - v_{t-1}({}^{G}c_i)}$$

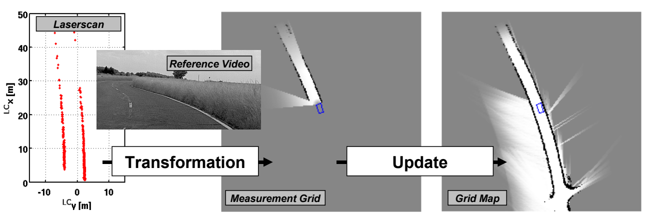

Fig. 1. Overview over the laser based grid mapping using an occupancy grid mapping approach.

B. Avoiding Moving Objects in Occupancy Grid Maps

1) Moving Objects in Grid Maps: In the application that will be presented later in section IV, the term occupied is interpreted as occupied by stationary objects. This means that a cell, that is only temporarily occupied (e.g. by a vehicle) should be marked as free in the occupancy grid. An ordinary grid mapping algorithm as described above usually produces artifacts as shown in Figure 2. Other traffic participants like oncoming vehicles produce laser measurements, which result in occupied cells in the area where the vehicle produced the last laser measurements before it left the laser's field of view. This is a problem, because if an area right in the middle of the street is marked as occupied, no moving objects will be detected there later since the detection algorithm will consider any corresponding measurements as produced by stationary objects. The aim is to produce an occupancy grid ${}^{O}\boldsymbol{M}$, that holds occupied cells only for stationary objects.

2) Grid Based Tracking: The following approach was introduced in [4]. It is based on the idea to track inconsistencies in consecutive occupancy grids. The filtering is realized with a multi-object Kalman filter under assumption of constant velocity. The state consists of the position and velocity of an object.

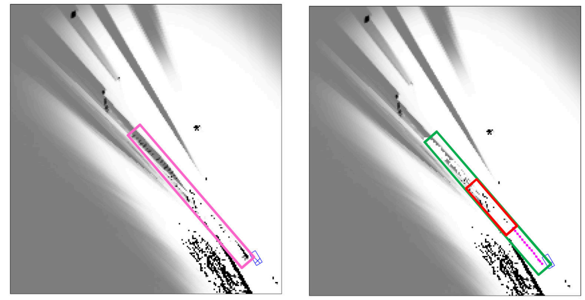

Fig. 2. Regular grid map (left) and grid map without moving objects (right) in a scene with an oncoming vehicle. The affected areas are framed. Only few artifacts remain in the grid map without moving objects (framed red). Magenta dots mark the track of the vehicle.

a) Generation of a Differential Grid: The algorithm starts with an occupancy grid ${}^{O}\boldsymbol{M}_{t-1}$ that holds no stationary objects and the regular occupancy grid ${}^{G}\boldsymbol{M}_{t-1}$. Both maps are updated with the measurement grid ${}^{MG}\boldsymbol{M}_t$ according to (1). The *differential grid* ${}^{D}\boldsymbol{M}_t$ represents the difference between two consecutive grids:

$$v_t({}^{D}c_i) = v_t({}^{G}c_i) - v_{t-1}({}^{G}c_i). \tag{2}$$

b) Binarisation: A threshold γ separates the differential grid into cells that are either consistent if $v_t({}^{O}c_i) < \gamma$ or inconsistent if $v_t({}^{O}c_i) > \gamma$.

c) Segmentation: For robustness, only clusters of a certain size are considered in the differential grid. Differences that occur only in few cells are deleted. A centroid and a bounding box is calculated for each considered cluster.

d) Data association: The centroids are considered as measurements and are associated to tracks with the nearest neighbor rule. Measurements that can not be associated to any track create new tracks, whereas tracks are deleted if their uncertainty exceeds a certain level due to missing measurements.

e) Elimination of dynamic objects: The grid map ${}^{O}\boldsymbol{M}_t$ is modified by assigning low occupancy values to cells that belong to moving objects. For validation, tracks that only exist for few time steps are handled in a reprocessing step. In this case, the corresponding cells are assigned to their original occupancy values again.

The result of this algorithm is a grid map, that is not affected by moving objects of a certain size, that could be tracked. Vehicles, that do not move while reflecting laser measurements will still result in occupied cells. However, in [4] an evaluation showed that artifacts produced by oncoming vehicles can be reduced significantly, which is important because these artifacts are situated right on the street, where object tracking has to be reliable.

III. JIPDA Filtering

A. Basics of JIPDA

In general, object tracking delivers an estimation about the state of one or more objects which is based on measurements and a priori knowledge about the behavior of the objects.

Statistical filters based on the Bayesian approach are common methods for object tracking. A general problem that occurs when more than one measurement is available, is how to associate the measurements to a track. Multiple measurements can originate from multiple sensors and/or from clutter measurements. A simple solution to this problem is the nearest neighbor association which uses hard decisions based on a spatial comparison between tracks and measurements. More sophisticated methods are based on probabilistic data association (PDA). In this case, not only the most likely measurement of a track is taken into account, but also less likely measurements.

Another problem occurs if not only one but multiple objects have to be tracked. A filter that considers all possible measurement-to-track combinations, is the Joint PDA (JPDA) filter [2]. A filter that additionally estimates the existence probability of tracks, is the Joint Integrated PDA (JIPDA) filter [10]. Especially laser measurements can usually not be uniquely associated to tracks, because they do not contain unique features and clutter measurements usually can not be avoided in an automotive environment. Therefore, JIPDA filters have proven to offer great benefits compared to simple estimation methods because information is not lost due to hard decisions. A drawback are the computational costs that originate from the number of hypotheses, that have to be considered.

A real time capable graph based implementation of a JIPDA filter has been presented in [8]. However, the number of objects, that can be tracked with this method is still limited due to computational costs.

In the JIPDA filter described in [8], each track $\boldsymbol{x}_i$ consists of a Gaussian spatial distribution $p(\boldsymbol{x}_i) = \mathcal{N}(\boldsymbol{x}_i, \widehat{\boldsymbol{x}}_i, \boldsymbol{P}_i)$ with mean $\widehat{\boldsymbol{x}}_i$ and covariance $\boldsymbol{P}_i$ and an existence probability $p(\exists \boldsymbol{x}_i)$. An existent object is defined as relevant and observable. All information about track quality is included in the existence probability. Since no hard decisions take place, measurement-, association-, detection- and existence uncertainties have to be dealt with. Therefore, certain sensor-specific properties have to be modeled.

B. Sensor properties

The sensors are supposed to transmit a spatial measurement vectors $\boldsymbol{z}_s$ with an associated covariance matrices $\boldsymbol{R}$ representing the measurement uncertainty as well as existence measurements $\boldsymbol{p}$.

The following probabilities are needed to describe all sensor-specific properties [9]:

- **true positive** probability (a relevant target exists under condition of the measurement $\boldsymbol{z}_j$):
  $p_{TP} = p(\exists x|\boldsymbol{z}_j)$,
- **false positive** probability (no relevant target exists under condition of the measurement $\boldsymbol{z}_j$):
  $p_{FP} = p(\nexists x|\boldsymbol{z}_j) = 1 - p_{TP}$,
- **true negative** (or detection) probability (the sensor will report a measurement under condition of an existing target $\boldsymbol{x}_i$):
  $p_{TN} = p(\exists z|\boldsymbol{x}_i)$ and
- **false negative** probability (the sensor will not report a measurement under condition of an existing target $\boldsymbol{x}_i$):
  $p_{FN} = p(\nexists z|\boldsymbol{x}_i) = 1 - p_{TN}$.

In summary, p_{TP} and p_{TN} have to be modeled. Various modeling methods have been proposed, see [9] for example. It is a common approach to model p_{TN} as a constant and spatial distributed function that originates from the sensor's characteristics, whereas p_{TP} is derived separately in each case from existence measurements $\boldsymbol{p}$. In general, the tracking process benefits from measurements even if p_{TP} is low, as long as it is modeled correctly.

C. Sensor Setup for JIPDA Tracking

This paper focuses on purely laser scanner based object detection. The detection of traffic participants is narrowed down to searching for regular passenger cars. In terms of laser measurements, these vehicles usually form L- or I-shapes in the birds-eye view of reflected laser points. In a first step, the laser points are segmented. After that, the iterative end point fit (IEPF) algorithm [13] is used for detecting edges in the contours, see Figure 4. This way, L- and I-shapes are detected. The algorithm estimates the center of corresponding vehicles by fitting rectangular boxes in L- and I-shaped clusters. In summary, the spatial part of a measurement $\boldsymbol{z}_s$ includes the position and orientation of a box and its uncertainties.

The quality (and with it the existence part $\boldsymbol{p}$) of such a measurement can be estimated by evaluating the remaining difference δ between the laser measurement points and the resulting box. Here, δ is defined as the maximum distance between a laser point of one segment and the corresponding box margins. This way the quality of a measurement is defined over its source being formed rectangular.

IV. Avoiding Stationary Objects as Measurements

The laser based object detection described above is a very appropriate approach for this paper, because it helps to illustrate the general idea of using the grid map as environmental representation in order to overcome tracking problems.

Figure 3 shows the laser measurement points of one frame in a typical rural scene. The number of detected boxes can be controlled by comparing the difference δ between laser measurement points and resulting boxes to a certain tolerance threshold δ_{tr} and omitting measurements $\boldsymbol{z}_j$ if $\delta_{\boldsymbol{z}_j} > \delta_{tr}$. The higher the threshold, the greater will be the number of measurements.

Theoretically, the JIPDA filter benefits from all measurements, as long as the corresponding existence part $\boldsymbol{p}$ and

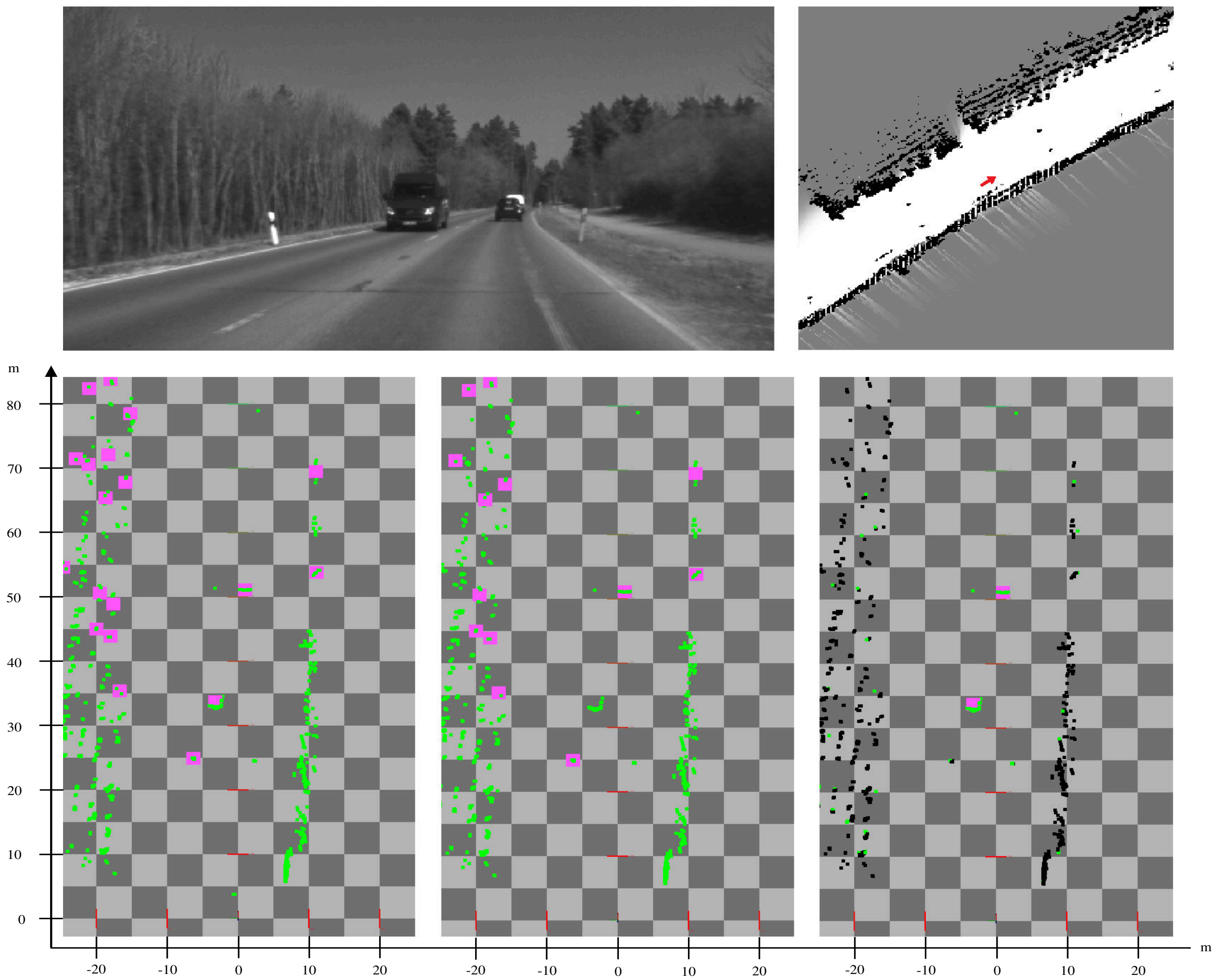

Fig. 3. Laser measurements (bottom), grid map (top right), and reference video (top left). Detected L- and I-shapes are marked with magenta squares. Bottom left: tolerance threshold δ_{tr} = 0.5 m, many objects are detected. Bottom middle: δ_{tr} = 0.2 m, fewer objects are detected, but one vehicle is missed. Bottom right: δ_{tr} = 0.5 m, laser measurement segments that lie in an occupied area of the grid map are omitted (black).

spatial uncertainties $\boldsymbol{R}$ are correct. In practice, the number of measurements has to be limited due to computational reasons. Hence, a low threshold δ_{tr} has to be chosen for the sake of real-time capability, with the result that some vehicles will be missed (p_{FN} rises).

One approach to overcome this drawback is to define only moving objects as relevant. If it is possible to distinguish between stationary and moving objects, the overall number of detected objects can be reduced by omitting stationary objects. This way, the threshold δ_{tr} can be set on a high level again without producing too many measurements. Hence, the algorithm is still real-time capable and at the same time p_{FN} is kept low.

Using an occupancy grid map for this purpose yields the advantage that the object's property of being stationary can be evaluated with only one single measurement despite the fact that the laser points do not provide any information about an object's velocity. Therefore, a measurement $\boldsymbol{z}_s$ is evaluated by analyzing its corresponding position in the occupancy grid map. If it lies in an occupied area of the grid map, it is omitted. As a result, δ_{tr} can be chosen in a way that the number of false positive measurements is reduced and at the same time the number of missed detections is increased in comparison the approach without the grid map.

So far, the benefits of this approach for JIPDA tracking have been outlined. However, a number of other state estimation algorithms are also limited in the number of measurements they can process, especially other PDA filters (e.g. the random finite set filter [7]). These approaches can benefit from the occupancy grid map as well.

V. EVALUATION

A research vehicle with a four layer laser sensor provides laser measurements from a test drive on a heavily frequented rural road. First, a grid map without moving objects is generated. In a second drive, laser based vehicle detection

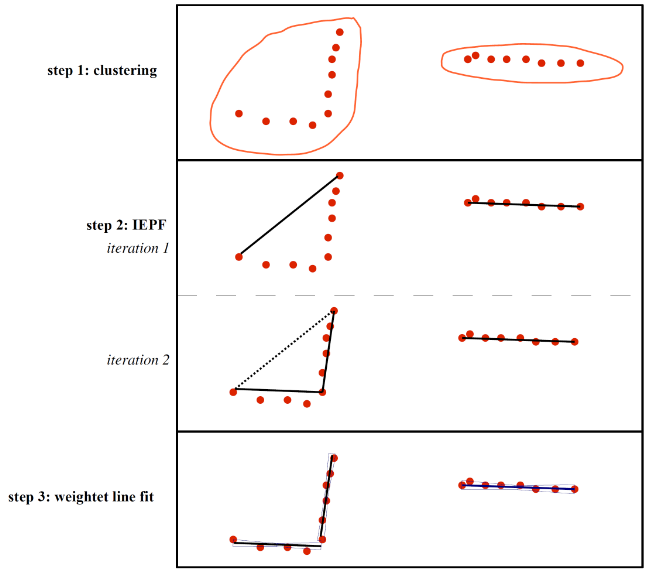

Fig. 4. Iterative End Point Fit (IEPF).

and JIPDA-tracking is evaluated using the grid map taken from the first drive. The region of interest was limited to 60 m. 830 Frames were manually labeled, which included a total number of 551 relevant vehicles.

A. Results of the Object Detection

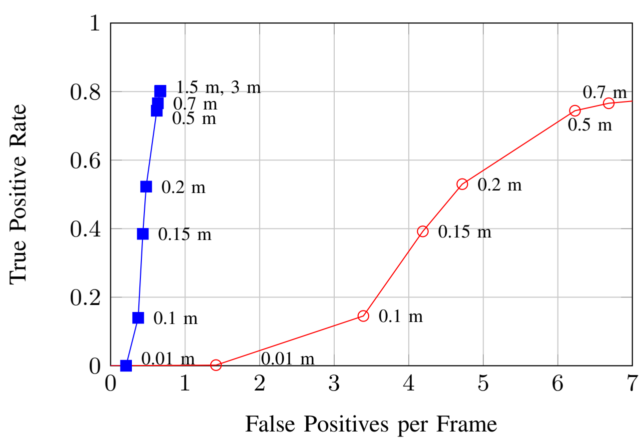

Fig. 5. ROC graphs for the laser based object detection with grid map (blue squares) and without grid map (red circles) for several tolerance thresholds δ_{tr} (distance values).

Figure 5 shows the ROC graph of the laser scanner based object detection with and without grid map. The graphs show, that the false positive rate (FPR) is significantly reduced for a constant tolerance threshold δ_{tr}, whereas the true positive rate (TPR) is only slightly reduced. The reduction of the TPR originates from a small number of vehicles, that have been omitted because they were incorrectly considered as stationary objects due to remaining artifacts or inaccurate localization via GPS. With increasing tolerance threshold δ_{tr} the TPR with grid map converges, because some vehicles do not produce L- or I-shapes. Also the FPR converges, since the total number of moving objects in the sequence is limited.

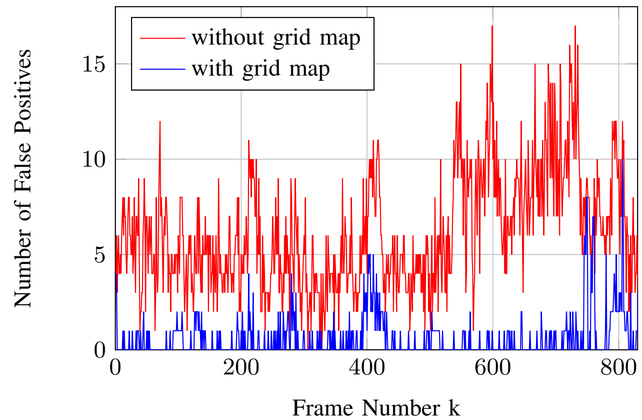

Fig. 6. Number of false positive measurements in the ROI with and without grid map.

The number of false positive object detections in each frame with a tolerance threshold $\delta_{tr} = 0.5$ m is plotted in Figure 6. The high number of false positive objects without grid map around $k = 700$ is caused by a forest, which produces a great amount of laser segments, as illustrated in Figure 3. At $k = 800$, the localization accuracy in the grid map was reduced due to a weak GPS signal.

B. JIPDA Tracking Results

The tracking results of the JIPDA tracker using a detection tolerance threshold $\delta_{tr} = 0.5$ m are evaluated with the optimal subpattern assignment metric (OSPA) [12], [11]. The OSPA metric calculates a scalar value for each time step which incorporates the position error as well as the cardinality error. In the evaluation a cut-off parameter $c = 10$ and the order $p = 1$ are used. Due to the cut-off parameter, the maximum value of the OSPA distance is 10 and is e.g. obtained if the labeled set is empty and the JIPDA tracker has one or more tracks or vice versa.

In Figure 7, the number of labeled ground truth objects in the scene as well as the estimated number of tracks of the JIPDA tracker are shown. Except for a small number of time steps, the number of labeled objects is equal to the number of tracked objects.

Figure 8 shows the OSPA distance for the investigated sequence. Around $k = 200$, the OSPA distance increases several times within on timestep to a value around five, which is due to short-time cardinality errors during object appearance or disappearance. For $300 < k < 500$, the OSPA distance increases to 10 every time an object appears or disappears, since one of the sets is empty when the cardinality errors occur. For $k > 500$, the OSPA distance is very small most of the time. Thus, the number of ground truth objects is equal to the number of tracks and additionally the tracks are located very close to the labeled position.

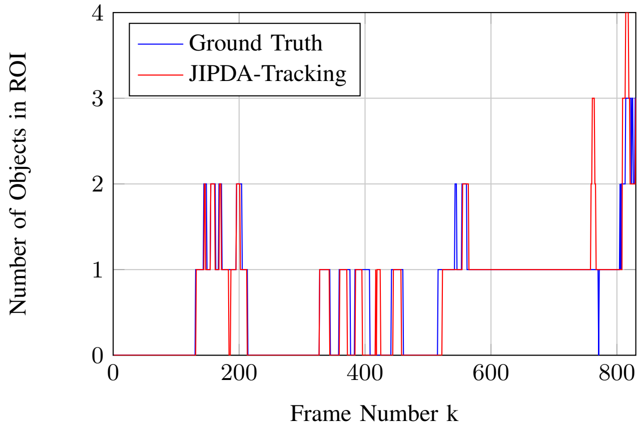

Fig. 7. Number of labeled objects in the scene vs. the number of tracks in the JIPDA tracker.

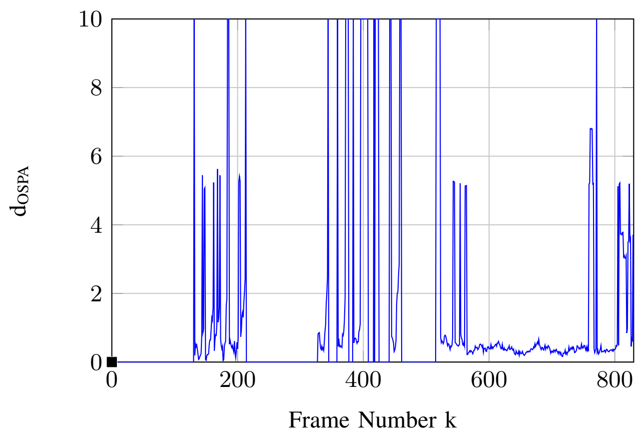

Fig. 8. OSPA of JIPDA tracking using only measurements of moving objects determined by the grid map.

VI. Conclusion and Future Works

In this contribution, an approach to use an in advance created grid map to reduce the number of false positive measurements is presented. Using the grid map, a solid distinction between moving and stationary objects is possible. When generating the grid map, moving objects were filtered, to make sure that only stationary objects result in occupied areas of the grid map. If the same route is driven again, a detection algorithm can be focused on moving objects, which yields a lower false positive rate for the received measurements. At the same time, the tolerance threshold of objects, that are classified as relevant can be increased. This allows a JIPDA filter to run in real time and provide a reliable tracking of vehicles.

Further improvements of this approach can be expected. This contribution focused on laser measurements, but grid maps can be produced from other sensors as well [6]. Hence a combination of sensors, that each produce their own grid map is promising.

The evaluation showed that a precise self-localization is indispensable for this approach. A GPS-based localization is not reliable due to multipath propagation and satellite occlusion problems. Hence, alternatives for self-localization like simultaneous localization and mapping (SLAM) or map matching approaches will have to be included in this approach. Finally, it is desirable to find a memory-efficient and generic way to store grid maps and make them available and updateable like open street maps.

References

[1] M.E. Bouzouraa and U. Hofmann. Fusion of occupancy grid mapping and model based object tracking for driver assistance systems using laser and radar sensors. In *Intelligent Vehicles Symposium (IV), 2010 IEEE*, pages 294 –300, 6 2010.

[2] J. Dezert and Y. Bar-Shalom. Joint probabilistic data association for autonomous navigation. *IEEE Transactions on Aerospace and Electronic Systems*, 29(4):1275 –1286, oct 1993.

[3] Jan Effertz. *Autonome Fahrzeugführung in urbaner Umgebung durch Kombination objekt- und kartenbasierter Umfeldmodelle*. PhD thesis, Technischen Universität Carolo-Wilhelmina zu Braunschweig, 2009.

[4] Marcus Konrad, Manuel Fuchs, Otto Löhlein, and Klaus Dietmayer. Detektion und tracking dynamischer objekte in occupancy grids. In *7. Workshop Fahrerassistenzsysteme FAS 2011*, pages 105–114, 2011.

[5] Marcus Konrad, Magdalena Szczot, and Klaus Dietmayer. Road course estimation in occupancy grids. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, pages 412–417, 2010.

[6] Marcus Konrad, Magdalena Szczot, Florian Schüle, and Klaus Dietmayer. Generic grid mapping for road course estimation. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, pages 851–856, 2011.

[7] Ronald P.S. Mahler. *Statistical Multisource-Multitarget Information Fusion*. Artech House Inc., Norwood, 2007.

[8] Mirko Mählisch. *Filtersynthese zur simultanen Minimierung von Existenz-, Assoziations- und Zustandsunsicherheiten in der Fahrzeugumfelderfassung mit heterogenen Sensordaten*. PhD thesis, Ulm University, 2009.

[9] Michael Munz, Klaus Dietmayer, and Mirko 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.

[10] D. Musicki and R. Evans. Joint integrated probabilistic data association - jipda. In *Proceedings of the Fifth International Conference on Information Fusion 2002*, volume 2, pages 1120 – 1125, 2002.

[11] B. Ristic, Ba-Ngu Vo, D. Clark, and Ba-Tuong Vo. A metric for performance evaluation of multi-target tracking algorithms. *IEEE Transactions on Signal Processing*, 59(7):3452 – 3457, July 2011.

[12] D. Schuhmacher, B.-T. Vo, and B.-N. Vo. A consistent metric for performance evaluation of multi-object filters. *IEEE Transactions on Signal Processing*, 56(8):3447 –3457, 8 2008.

[13] Ali Siadat, Axel Kaske, Siegfried Klausmann, Michel Dufaut, and René Husson. An optimized segmentation method for a 2d laser-scanner applied to mobile robot navigation. In *IFAC Symposium on Intelligent Components and Instruments for Control Applications*, 1997.

[14] Sebastian Thrun, Wolfram Burgard, and Dieter Fox. *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, 2005.

[15] Christopher 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*, 25(8):425–466, June 2008.

[16] Trung-Dung Vu, Olivier Aycard, and Nils Appenrodt. Online localization and mapping with moving object tracking in dynamic outdoor environments. In *Proceedings of the IEEE Intelligent Vehicles Symposium*, pages 190–195, 2007.

[17] Chieh-Chih Wang, Charles Thorpe, Sebastian Thrun, Martial Hebert, and Hugh Durrant-Whyte. Simultaneous localization, mapping and moving object tracking. *Int. J. Rob. Res.*, 26(9):889–916, September 2007.

[18] Thorsten Weiss, Bruno Schiele, and Klaus Dietmayer. Robust driving path detection in urban and highway scenarios using a laser scanner and online occupancy grids. In *Proceedings of IEEE Intelligent Vehicles Symposium*, pages 184 –189, 6 2007.