This is the author's version of a conference paper presented at the IEEE International Conference on Intelligent Transportation Systems (ITSC).
The final version of record is available at http://dx.doi.org/10.1109/ITSC.2018.8569511

Data Association for Grid-Based Object Tracking Using Particle Labeling

Sascha Steyer, Georg Tanzmeister, Christian Lenk, Vinzenz Dallabetta, and Dirk Wollherr

***Abstract*— Estimating surrounding objects and obstacles by processing sensor data is essential for safe autonomous driving. Grid-based approaches discretize the environment into grid cells, which implicitly solves the data association between measurement data and the filtered state on this grid representation. Recent approaches estimate, in addition to occupancy probabilities, cell velocity distributions using a low-level particle filter. Measured occupancy can thus be classified as static or dynamic, whereby a subsequent tracking of moving objects can be limited to dynamic cells. However, the data association between those cells and multiple predicted objects that are close to each other remains a challenge. In this work, we propose a new association approach in that context. Our main idea is that particles of the underlying low-level particle filter are linked to those high-level objects, i.e., an object label is attached to each particle. Cells are thus associated to objects by evaluating the particle label distribution of that cell. In addition, a subsequent clustering is performed, in which multiple clusters of an object are extracted and finally checked for plausibility to further increase the robustness. Our approach is evaluated with real sensor data in challenging scenarios with occlusions and dense traffic.**

I. Introduction

An accurate model of the environment is essential for autonomous vehicles. Sensor data are required to detect surrounding traffic participants and obstacles of the current local environment. Objects have to be extracted from those sensor data to plan interactive maneuvers and avoid collisions, i.e., to enable safe and intelligent autonomous driving applications. A basic problem of object tracking is the data association between predicted objects of the temporally filtered estimation and new measurements, which also directly applies to the data fusion between different sensors.

A common approach is a high-level fusion, e.g. [1], [2], in which, per sensor, object hypotheses are extracted from the individual measurements by specific features and associated with tracked objects of that sensor. The sensor data fusion is performed afterwards, i.e., based on that abstracted high-level representation, which also requires a high-level association. Various data association concepts exist that can be used for this purpose, e.g., local or global nearest neighbor (NN/GNN) association [3], multiple hypothesis tracking (MHT) [4], joint probabilistic data association (JPDA) [5], or random finite set (RFS) tracking [6], [7].

The sensors used today usually provide a large amount of detections, i.e., an object typically induces multiple measurements, resulting in so called extended object tracking [8]. The abstraction of measurement data to a high-level object representation like a bounding box shape, however, causes information loss and thus also a more error-prone data fusion and association. Hence, it is beneficial to fuse those measurement data in advance, i.e., on a low-level representation.

Occupancy grids [9] enable such a low-level representation, in which the sensor-specific measurement data are modeled in a uniform, discretized grid with occupancy probabilities of each grid cell. Sensor data can thus be fused cell-wise, i.e., the association is implicitly solved by the spatial cell discretization without requiring object assumptions. This concept is also used in the temporal accumulation, resulting in the known grid mapping of static environments.

Dynamic occupancy grids additionally estimate the dynamic state of the grid and thus distinguish static and dynamic occupancy. Hence, moving objects can be extracted and updated by solely considering dynamic occupied cells, whereas static obstacles are estimated in the grid map representation. This simplifies the association as the set of measured occupied cells that can be associated with the predicted objects is reduced to those that are classified as dynamic. Moreover, since static obstacles are precisely modeled in the grid map without requiring to extract them in the same high-level object representation, ambiguities of overlapping areas of predicted objects are significantly reduced.

Occupied cells can be classified as dynamic by determining free/occupied inconsistencies between the accumulated map and new measurements, e.g. [10]–[12]. This concept is used in [12]–[15] to classify parts of the measurement data as dynamic, which are then clustered using a distance- or density-based clustering to extract object hypotheses. These object hypotheses are finally associated with filtered objects using GNN, MHT, or JPDA.

Clustering dynamic cells for extracting new occurring objects is expedient as no previous information is available in such a case. However, especially in urban environments, there are several scenarios where a clustering without additional information fails. For example, multiple objects that temporarily move closely to each other may result in one large cluster, whereas partial occlusions of one object may result in several small clusters. Hence, it is crucial that available information of predicted objects is considered in advance to help solving these ambiguities. In [16], a region of interest of each predicted object is used as possible starting points of a clustering of that object. However, no detailed prediction of expected areas of occupancy or object shape information are considered in those regions of interest.

S. Steyer, G. Tanzmeister, C. Lenk, and V. Dallabetta are with the BMW Group, Munich, Germany. E-mail: sascha.steyer@bmw.de, georg.tanzmeister@bmw.de, christian.cl.lenk@bmw.de, vinzenz.dallabetta@bmw.de

D. Wollherr is with the Institute of Automatic Control Engineering (LSR), Technical University of Munich, Munich, Germany. E-mail: dw@tum.de

This is the author's version of a conference paper presented at the IEEE International Conference on Intelligent Transportation Systems (ITSC).
The final version of record is available at http://dx.doi.org/10.1109/ITSC.2018.8569511

Recent approaches of dynamic occupancy grids use a grid-based particle filter to estimate the dynamic state of the occupied environment, e.g. [17]–[21]. Each particle represents a hypothesis of occupancy at a specific continuous position with a particular velocity. This low-level particle tracking is primarily used to robustly estimate cell velocity distributions and thus predict occupancy probabilities of the grid map. This velocity estimation also helps distinguishing objects that move closely to each other with different velocities.

In our previous work [22], we proposed to associate individual grid cells with existing objects without using a clustering. The predicted object pose and box model shape are used to compute the expected areas of measured occupancy, resulting in cell association probabilities of each object. In addition, the similarity of the estimated cell velocity and the object velocity is considered. The mentioned scenarios where a clustering typically fails can thus be handled this way. However, a main drawback is that the occupancy likelihood approximation of cells extracted from the box representation is not accurate, even though the current visibility of the box edges is also considered. This is critical in overlapping areas of multiple objects, in which conflicts of multiple similar cell association probabilities cannot be robustly solved.

Filtered objects can also be directly extracted by clustering dynamic cells that are estimated by the low-level particle tracking, since the particle population already represents a filtering of occupancy and velocity distributions. As stated in [17], those particles could be extended by a unique ID to reconstruct trajectories of those extracted objects. Similarly, in [18], an object ID of those particles is used to extract object clusters including their mean velocity and gravity center. This ID is propagated by the particle resampling, i.e., duplicated particles inherit the ID, whereas randomly drawn new particles are initialized with a new unique ID. However, the resulting object clusters tend to have similar problems as described before, i.e., a large object can result in several small clusters of estimated objects, whereas objects moving closely to each other converge toward one single cluster.

In this work, we propose a new association concept using those low-level particles. Similar to [17], [18], particles are also extended by an object identifier, which we denote the label of a particle. However, as presented in [22], we retain our concept of using a separate high-level object tracking with an unscented Kalman filter that is updated by a set of associated occupied cells. That way, an accurate object state, including the position, velocity, orientation, acceleration, turn rate, and the size of the box model shape, is robustly filtered. Moreover, merging or deleting objects is decided on this high-level representation. Hence, our main motivation is to use the low-level particle population for improving the association between occupied cells and high-level objects.

The basic idea is that particles can be linked to an object, where the predicted population of all particles linked to an object represents the prediction of its occupancy likelihood. Hence, those labeled particles enable an accurate estimation of the expected areas of occupancy measurements of each object. Cells can thus be associated by analyzing their particle label distribution. Additionally, in a subsequent clustering step, multiple clusters of an object are extracted and checked for plausibility. Finally, only the most certain clusters of an object are used, i.e., uncertain clusters can also be discarded, which further improves the association.

Overall, we propose a new association approach in the context of grid-based object tracking that improves the object estimation and, altogether, the safety of autonomous driving applications. The rest of this paper is structured as follows: Sec. II defines the specific data association problem of this work and gives an overview of our grid-based environment estimation approach. The proposed particle labeling association is presented in Sec. III, which is extended in Sec. IV by the additional clustering step with verification. Results with real sensor data in various scenarios are given in Sec. V.

II. Problem Formulation and Grid-Based Mapping & Tracking Overview

In this section, the addressed data association problem of input measurements and predicted objects is defined regarding the specific representations of this work. In addition, the basic concept of the overall grid-based mapping and tracking approach is explained, especially the low-level particle tracking, as it is directly used to solve the association. Further information of that approach is described in [20]–[22].

A. Problem Formulation

The fused measurement data of various sensors are modeled in a measurement occupancy grid

$$Z_t = \{z_c \,|\, c \in \mathcal{G}\}, \quad z_c = \left[\ \mathrm{bel}(O_{z,c}),\, \mathrm{bel}(F_{z,c})\ \right]^\mathsf{T} \tag{1}$$

that represents measurement data from one time instance t, i.e., not accumulated over time. An occupancy grid representation using the Dempster-Shafer evidence framework [23], [24] is used. Each cell measurement z_c of a cell $c \in \mathcal{G}$ of the grid structure $\mathcal{G}$ contains two measured beliefs $\mathrm{bel}(\cdot) \in [0, 1]$ for the hypotheses occupied O and free F, respectively.

The output of the overall grid-based object tracking approach [22] are temporally filtered objects, called tracks $\tau \in \mathcal{T}_t$, with $\mathcal{T}_t$ defining the set of all estimated tracks at time t. Each track τ is defined by its state s_τ, containing the pose and size of the oriented box model, among others. Hence, cell measurements z_c, or the cell indices c, have to be associated to the corresponding predicted track $\tau \in \mathcal{T}_t$ to perform a measurement update of that track. Relevant cells of the grid structure can be limited to the set

$$\mathcal{G}' = \left\{ c \in \mathcal{G} \mid \mathrm{bel}(O_{z,c}) \geq \Gamma^O_{\min} \right\} \subseteq \mathcal{G}, \tag{2}$$

as only cells above a minimum occupancy measurement threshold $\Gamma^O_{\min} \in (0, 1)$ are considered for this purpose.

Overall, the association problem is described by the surjective function

$$f_a \ : \ \mathcal{G}' \to \mathcal{T}_t \cup \{\xi_\emptyset\} \ : \ c \mapsto \tau \tag{3}$$

that maps occupied cells $c \in \mathcal{G}'$ to tracks $\tau \in \mathcal{T}_t$ given the cell measurements z_c and predicted track states $\hat{s}_\tau$. Cells that are not associated to a track are mapped to an auxiliary

This is the author's version of a conference paper presented at the IEEE International Conference on Intelligent Transportation Systems (ITSC).
The final version of record is available at http://dx.doi.org/10.1109/ITSC.2018.8569511

variable $\xi_\emptyset$, representing static background, clutter, or new objects. As this work handles extended objects rather than point objects, multiple cells can be associated to a track, with

$$\mathcal{C}_\tau = \{c \in \mathcal{G}' \mid f_a(c) = \tau\} \quad \forall \tau \in \mathcal{T}_t \tag{4}$$

defining the set of cells associated to a track τ. However, each cell is at most associated with one track, i.e.,

$$\mathcal{C}_{\tau_1} \cap \mathcal{C}_{\tau_2} = \emptyset \quad \forall \tau_1, \tau_2 \in \mathcal{T}_t,\ \tau_1 \neq \tau_2\,. \tag{5}$$

Fig. 1 exemplarily illustrates this data association problem between occupied cells of the measurement grid Z_t, shown in Fig. 1a, and multiple predicted tracks $\tau \in \mathcal{T}_t$, shown in Fig. 1b. The resulting associated occupied cells $\mathcal{C}_\tau$ of each track τ are visualized in Fig. 1c.

B. Dynamic Grid Mapping and Particle Tracking

An object tracking that directly and solely uses the measurement occupancy grid Z_t as input would require to extract high-level objects for all, arbitrary shaped, occupied areas of the environment. Static obstacles, however, can be modeled and estimated more accurately and efficiently by an occupancy grid map using the same grid representation as the input. Therefore, the object tracking should be limited to dynamic parts of the environment instead, which requires an additional dynamic state estimation of the grid.

In [21], we proposed such a dynamic grid mapping approach that estimates evidence masses for the hypotheses static occupancy S, dynamic occupancy D, free space F, and their combined hypotheses. Hence, the unclassified occupancy hypothesis $O = \{S, D\}$ is split into the individual hypotheses S and D, which are separately estimated. The filtered dynamic grid map allows a classification and thus a subdivision of the measured occupancy belief

$$\mathrm{bel}(O_{z,c}) = m(S_{\tilde{z},c}) + m(D_{\tilde{z},c}) + m(O_{\tilde{z},c}) \tag{6}$$

into individual evidence masses $m(\cdot) \in [0, 1]$ of static, dynamic, and the remaining unclassified occupancy evidence, which together form a pseudo-measurement $\tilde{z}$. The resulting occupancy classification of Fig. 1a is shown in Fig. 1d.

Dynamic evidence masses $m(D)$ of the grid map are robustly estimated using a low-level particle tracking. Each particle $\chi \in \mathcal{X}_t$ of the population $\mathcal{X}_t$ represents a hypothesis

$$\left[\, x_\chi^x, x_\chi^y, v_\chi^x, v_\chi^y, o_\chi \,\right]^\mathsf{T} \tag{7}$$

of dynamic occupancy at a specific position $x_\chi \in \mathbb{R}^2$ with a velocity $v_\chi \in \mathbb{R}^2$ and an occupancy value $o_\chi \in [0, 1)$. The particle population $\mathcal{X}_t$ is updated cell-wise by evaluating the set of particles $\mathcal{X}_{t,c}$ associated to a grid cell $c \in \mathcal{G}$ regarding the current positions x_χ. Particles are only drawn in dynamic cells and, to initialize new dynamic evidence, in cells with unclassified occupancy of the filtered map. Static occupancy, in contrast, is directly accumulated by the map without using particles. The main task of the particle tracking is to estimate cell velocity distributions and to initialize and predict dynamic evidence of the dynamic grid map using the corresponding occupancy values o_χ. The filtered particle population of the scenario of Fig. 1 is shown in Fig. 1e.

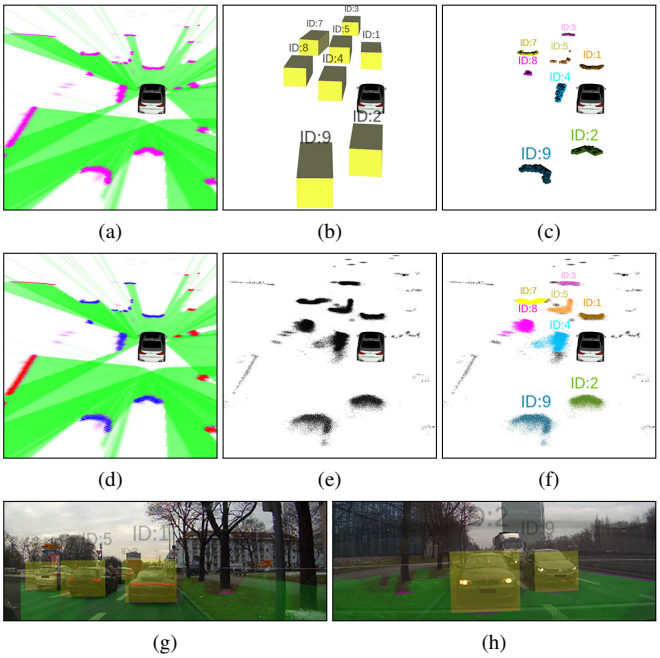


Fig. 1. Overview of the association problem and proposed approach. (a) Measurement occupancy grid Z_t (O: pink, F: green). (b) Predicted tracks τ. (c) Resulting associated occupied cells $\mathcal{C}_\tau$ of each track. (d) Occupancy classification of measurement grid (S: red, D: blue). (e) Particle population. (f) Proposed labeled particle population. (g), (h) Front and rear camera images with an additional projected overlay of (a) and (b).

In addition to the dynamic grid mapping approach of [21], the predicted tracks are used to avoid that occupied cells of slow-moving objects converge from dynamic toward static in this work. Hence, dynamic evidence and thus the corresponding particles are retained in areas of the predicted tracks. Therefore only occupied cells with a dynamic evidence $m(D)$ above a threshold $\Gamma^D_{\min} \in (0, 1)$ have to be considered in the association as defined in (3). Overall, the set of cells that can be associated to a track is limited to

$$\mathcal{G}'' = \{c \in \mathcal{G}' \mid m(D_{\tilde{z},c}) \geq \Gamma^D_{\min}\} \subseteq \mathcal{G}'\,. \tag{8}$$

III. Particle Labeling Association

This section presents the concept of using the low-level particle representation for solving the association between occupied grid cells and object tracks.

A. Particle Label Extension as Object Identifier

Our main idea is to extend the track attributes by linking a set of particles $\mathcal{X}_t^\tau \subseteq \mathcal{X}_t$ to each track $\tau \in \mathcal{T}_t$ that help to solve the association problem as defined in (3). For this purpose, a label index

$$l_\chi \in \mathcal{T}_t \cup \{\xi_\emptyset\} \tag{9}$$

is attached to each particle χ that defines the possible connection to a corresponding track. Fig. 1f illustrates the proposed label extension of the particle population of Fig. 1e. The set of particles linked to a track τ is thus defined by

$$\mathcal{X}_t^\tau = \left\{\chi \in \mathcal{X}_t \mid l_\chi = \tau\right\}, \tag{10}$$

where each particle is at most linked to one track at time t,

$$\mathcal{X}_t^{\tau_1} \cap \mathcal{X}_t^{\tau_2} = \emptyset \quad \forall \tau_1, \tau_2 \in \mathcal{T}_t,\ \tau_1 \neq \tau_2\,. \tag{11}$$

This label is inherited in the particle resampling step to remain the linking to the corresponding track, while completely new drawn particles are initialized with $l_\chi = \xi_\emptyset$. However, the particle label is not evaluated in the particle filter itself, i.e., no object relation is used for the prediction or weighting of the particles to retain a robust low-level tracking without modeling specific extended object assumptions for the point mass particles. Hence, the filtering behavior of the particle tracking remains unchanged. In the following, the population $\mathcal{X}_t$ represents updated particles after resampling, meaning that a particle velocity weighting and update using radar Doppler measurements as presented in [20], [25] are performed in advance. Measured velocities are thus implicitly considered by the updated particle population.

B. Cell Association Using Particle Label Evaluation

Dynamic occupied grid cells $c \in \mathcal{G}''$ as defined in (8) are associated with tracks $\tau \in \mathcal{T}_t$ by evaluating the particle label distribution in a cell. The ratio

$$r_c^\tau = \frac{|\mathcal{X}_{t,c}^\tau|}{|\mathcal{X}_{t,c}|} \in [0,1] \tag{12}$$

of the number of particles $|\mathcal{X}_{t,c}^\tau|$ linked to a track τ compared to the total number of particles $|\mathcal{X}_{t,c}|$ in that cell c indicates the origin of an occupancy measurement. The track

$$\tau_c^* = \underset{\tau \in \mathcal{T}_t}{\arg\max}\ r_c^\tau\,, \tag{13}$$

with the most corresponding particles in that cell results in the best fitting association. As stated in (5), each cell is associated at most with one track. Hence, a cell is associated

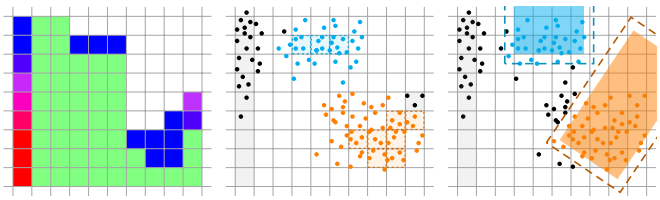
$$f_a(c) = \begin{cases} \tau_c^*\,, \text{if } r_c^{\tau_c^*} \geq \Gamma_{\min}^r \\ \xi_\emptyset\,, \text{else} \end{cases} \quad \forall c \in \mathcal{G}'' \tag{14}$$

to the best fitting track τ_c^* if the ratio of particles is above a threshold $\Gamma_{\min}^r$, otherwise no track is associated to that cell.

Within the remaining set of unassociated dynamic occupied cells $c \in \mathcal{G}''$ with $f_a(c) = \xi_\emptyset$, potential new object tracks are analyzed using a combination of a density-based clustering and a region growing as proposed in [22]. The resulting set of newly detected tracks is denoted by $\mathcal{T}_t^{\text{new}}$.

C. Particle Label Update / Reselection of Linked Particles

After the measurement update of the tracks, the linking between the particles and tracks has to be updated. On the one hand, particles that are too far away from the bounding box of the corresponding track should not retain linked to that track. On the other hand, particles without a corresponding track, i.e., $l_\chi = \xi_\emptyset$, but that are inside the bounding box of a track and have a similar velocity, may be linked to that track. This concept basically corresponds to a gating, i.e., defining which area of a track is generally valid and which not.



(a) (b) (c)

Fig. 2. Schematic illustration of particle labeling association concept. The scenario shows a guardrail on the left side, partly wrongly classified as dynamic, and two existing tracks with one turning right. (a) Classified measurement grid. (b) Labeled particle population and associated cells ($\xi_\emptyset$: black, τ_1: cyan, τ_2: orange). (c) Update of particle labels (filled rectangles: bounding boxes of updated tracks; dashed lines: valid gating areas of tracks).

The first case, i.e., unlinking all particles that do not match with the linked track anymore, is described by

$$\begin{aligned} &l_\chi \leftarrow \xi_\emptyset \quad \forall\, \chi \in \mathcal{A}_1^\tau\,,\ \tau \in \mathcal{T}_t\,, \\ &\mathcal{A}_1^\tau = \left\{\chi \in \mathcal{X}_t^\tau \mid x_\chi \notin b(s_\tau, \Delta_b)\right\} \end{aligned} \tag{15}$$

The position x_χ of a particle χ is compared with the bounding box $b(s_\tau, \Delta_b)$ of the track τ which depends on the track pose and length/width that are part of the state s_τ, where the valid area is enlarged by a distance parameter Δ_b. This unlinking strategy implicitly includes resetting labels of deleted tracks.

The second case, i.e., linking new particles to a track, including newly extracted tracks $\mathcal{T}_t^{\text{new}}$, is described by

$$\begin{aligned} &l_\chi \leftarrow \tau \quad \forall\, \chi \in \mathcal{A}_2^\tau\,,\ \tau \in \mathcal{T}_t \cup \mathcal{T}_t^{\text{new}}\,, \\ &\mathcal{A}_2^\tau = \big\{\chi \in \mathcal{X}_t^{\xi_\emptyset} \mid x_\chi \in b(s_\tau)\ \wedge\ v_\chi \in v_g(s_\tau)\ \wedge \\ &\qquad \left(\nexists \tilde{\tau} \in \{\mathcal{T}_t \cup \mathcal{T}_t^{\text{new}}\} \backslash \{\tau\} : x_\chi \in b(s_{\tilde{\tau}})\right)\big\}\,. \end{aligned} \tag{16}$$

A particle is only newly linked to a track, if it is directly inside the not enlarged bounding box $b(s_\tau)$ of exactly one track. In addition, the particle velocity v_χ is compared to a defined valid velocity gating interval $v_g(s_\tau)$ of the track. In overlapping areas of multiple tracks, no new particles are linked to any track. In such a case, particles that remained their label, without being unlinked by (15), are used to solve these ambiguities.

The different steps of the particle labeling concept are illustratively summarized in Fig. 2. Dynamic occupied cells of the classified measurement grid, shown in Fig. 2a, are associated to the two existing tracks by evaluating the particle label distributions, cf. Fig. 2b. After the measurement update of the tracks, the particle labels are updated, while particles outside the gating area are unlinked from that track and particles without a label that are inside the bounding box of a track are added to that track, see Fig. 2c.

IV. Additional Clustering with Verification

In this section, the particle labeling association of the previous section is extended by a subsequent clustering step. Multiple clusters of a track are extracted and eventually checked for plausibility. The motivation of this additional step is twofold. First, dynamic cells that have not been associated to any track by the particle population but are directly adjacent to cells associated with a track may be added to

that cluster. This is especially crucial for cells that are not directly covered by the estimated box model, i.e., outside the bounding box $b(s_\tau)$. For example, a preceding vehicle that turns quickly where a wrongly estimated orientation does not cover the long side of the detected L-shape. Secondly, the particle labeling association may associate individual cells close to a track, which, however, are not part of that track. For example, this may occur at road boundaries or curbs with some parts wrongly classified as dynamic, or new emerging objects that have not been extracted as a separate track in areas next to an existing track.

A. Adaptive Dynamic Evidence Threshold Selection

First, the dynamic evidence masses $m(D_{\tilde{z},c})$ of all cells $c \in \mathcal{C}_\tau$ that are associated to a track τ by the particle labels as defined in (14) are analyzed. An adaptive threshold

$$\gamma(\tau) = \max\left(\Gamma^D_{\min},\, \Gamma^{D,a}_{\min} \frac{1}{|\mathcal{C}_\tau|} \sum_{c\in\mathcal{C}_\tau} m(D_{\tilde{z},c})\right) \tag{17}$$

is calculated that depends on the mean dynamic evidence mass of the set $\mathcal{C}_\tau$ and a factor $\Gamma^{D,a}_{\min} \in [0,1]$. This dynamic evidence threshold $\gamma(\tau)$ is at least $\Gamma^D_{\min}$ as defined before in (8). The set $\mathcal{C}_\tau$ is thus reduced to the subset

$$\mathcal{C}'_\tau = \left\{c \in \mathcal{C}_\tau \,|\, m(D_{\tilde{z},c}) \geq \gamma(\tau)\right\}, \tag{18}$$

i.e., insignificant cells regarding the dynamic evidence mass are discarded. Such lower dynamic evidence masses may occur, e.g., for cells with an uncertain static/dynamic classification or noisy radar detections that are modeled with a low occupancy belief. Overall, this concept reduces false positives of the association and enables a more sensitive clustering of neighboring dynamic cells as described in the following.

B. Clustering of Associated Cells and Neighbors

Secondly, cells of the set $\mathcal{C}'_\tau$ are evaluated in terms of a clustering. In addition, neighboring unassociated dynamic occupied cells of the set

$$\begin{aligned}\mathcal{C}'_{\xi_\emptyset} &= \left\{c \in \mathcal{G}'' \mid f_a(c) = \xi_\emptyset \;\wedge\; r^{\xi_\emptyset}_c \geq \Gamma^r_{\min}\right\} \\ &\subseteq \mathcal{G}'' \setminus \bigcup_{\tau\in\mathcal{T}_t} \mathcal{C}_\tau\,, \quad r^{\xi_\emptyset}_c = 1 - \sum_{\tau\in\mathcal{T}_t} r^\tau_c\end{aligned} \tag{19}$$

are added to these clusters. However, unassociated cells with an ambiguous particle label distribution are excluded in $\mathcal{C}'_{\xi_\emptyset}$, i.e., only cells with no significant track origin are added in this clustering step. Hence, the ratio $r^{\xi_\emptyset}_c$ of the number of particles without an object label compared to all particles in that cell has to be above the threshold $\Gamma^r_{\min}$, cf. (12), (14). A single-linkage hierarchical clustering with a distance threshold $\Gamma^d_{\max}$ between the cell centers x_c is used, i.e., two cells c_1 and c_2 are associated to the same cluster if

$$\begin{aligned}&\|x_{c_1} - x_{c_2}\| \leq \Gamma^d_{\max}\,, \\ &c_1 \in \mathcal{C}'_\tau\,,\; c_2 \in \mathcal{C}'_\tau \cup \left\{c \in \mathcal{C}'_{\xi_\emptyset} \,|\, m(D_{\tilde{z},c}) \geq \gamma(\tau)\right\},\end{aligned} \tag{20}$$

which also requires that $m(D_{\tilde{z},c})$ of an unassociated cell is above the adaptive threshold $\gamma(\tau)$ of that track. Overall, this results in $|\mathcal{K}_\tau|$ clusters $\mathcal{C}^k_\tau$, $k \in \mathcal{K}_\tau$, of a track τ with

$$\mathcal{C}^k_\tau \subseteq \mathcal{C}'_\tau \cup \mathcal{C}'_{\xi_\emptyset},\; \mathcal{C}^k_\tau \cap \mathcal{C}'_\tau \neq \emptyset,\; \mathcal{C}'_\tau \subseteq \bigcup \mathcal{C}^k_\tau \;\; \forall k \in \mathcal{K}_\tau\,. \tag{21}$$

Each cluster $\mathcal{C}^k_\tau$ contains at least one cell of the set $\mathcal{C}'_\tau$, whereas the union of all clusters of a track includes all cells of the set $\mathcal{C}'_\tau$, and the intersection of two clusters results in the empty set, i.e., a cell is at most associated to one cluster.

C. Cluster Score Verification

Thirdly, the clusters are checked for plausibility before they are finally associated. Therefore, the clusters $\mathcal{C}^k_\tau$ of each track τ are compared to its predicted state $\hat{s}_\tau$. The ratio of the number of cells that are inside the predicted bounding box is calculated as

$$\lambda^x_{\tau,k} = \frac{\left|\left\{c \in \mathcal{C}^k_\tau \mid x_c \in b(\hat{s}_\tau)\right\}\right|}{|\mathcal{C}^k_\tau|}\,. \tag{22}$$

This ratio indicates the conformity in terms of the position. In addition, the length $b_l(\mathcal{C}^k_\tau)$ and width $b_w(\mathcal{C}^k_\tau)$ of each cluster are calculated, i.e., an oriented minimum bounding box is extracted using the predicted track orientation. This extracted bounding box geometry is compared

$$\Delta b^l_{\tau,k} = \frac{|b_l(\mathcal{C}^k_\tau) - b_l(\hat{s}_\tau)|}{b_l(\hat{s}_\tau)}, \quad \Delta b^w_{\tau,k} = \frac{|b_w(\mathcal{C}^k_\tau) - b_w(\hat{s}_\tau)|}{b_w(\hat{s}_\tau)} \tag{23}$$

with the length $b_l(\hat{s}_\tau)$ and width $b_w(\hat{s}_\tau)$ of the bounding box $b(\hat{s}_\tau)$ of the predicted track. This relative geometry conformity is mapped to a factor

$$\lambda^b_{\tau,k} = \exp\left(-\eta_b \min(\Delta b^l_{\tau,k}, \Delta b^w_{\tau,k})\right) \in [0,1]\,, \tag{24}$$

with a scaling parameter parameter $\eta_b \in \mathbb{R}^+$. Only the more fitting side of the box is considered in $\lambda^b_{\tau,k}$, i.e., it is sufficient if only the length or the width of the track is represented by a cluster. Overall, a cluster score

$$\lambda_{\tau,k} = \lambda^x_{\tau,k}\, \lambda^b_{\tau,k} \in [0,1] \tag{25}$$

is calculated that combines the conformity of the position and the box geometry compared to the predicted track. This concept can be extended by analyzing additional aspects like the cell velocities regarding the particle velocities v_χ, adjacent static occupied evidence $m(S_{\tilde{z},c})$, or the included freespace evidence $\mathrm{bel}(F_z)$.

All cluster scores of a track are normalized with respect to the best cluster score of that track, i.e.,

$$\overline{\lambda}_{\tau,k} = \frac{\lambda_{\tau,k}}{\max\limits_{k'\in\mathcal{K}_\tau} \lambda_{\tau,k'}}\,. \tag{26}$$

Finally, cells of a cluster are associated to the corresponding track

$$f'_a(c) = \begin{cases}\tau\,, \text{if } c \in \mathcal{C}^k_\tau \;\wedge\; \overline{\lambda}_{\tau,k} \geq \Gamma^\lambda_{\min} \\ \xi_\emptyset\,, \text{else}\end{cases} \quad \forall c \in \mathcal{G}''\,, \tag{27}$$

if the normalized score of the cluster is above a threshold $\Gamma^\lambda_{\min} \in [0,1]$. Due to the normalization, the best fitting cluster of a track is always associated to that track, whereas additional clusters are only added if they have a similar score.

This is the author's version of a conference paper presented at the IEEE International Conference on Intelligent Transportation Systems (ITSC).
The final version of record is available at http://dx.doi.org/10.1109/ITSC.2018.8569511

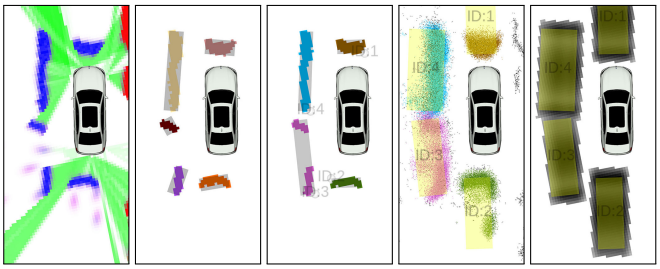
V. Experimental Results

The proposed approach has been tested in various real traffic scenarios. The test vehicle is equipped with four laser scanners and four short-range radar sensors. The implementation is based on fast parallel GPU computing that enables a real-time application in the autonomous driving test vehicles. A grid cell size of $0.15\,\mathrm{m} \times 0.15\,\mathrm{m}$ and a maximum of 100 particles per cell is used. In the following, the particle labeling association is compared with a clustering that does not use any information from the predicted tracks, with our previous association approach [22], and with the proposed additional clustering with verification.

Fig. 3 shows an urban scenario where the ego vehicle is surrounded by multiple objects. The vehicle on the south-west of the ego vehicle is, due to the restricted sensor field of view, split into two separate areas of occupancy measurements, cf. Fig. 3a. Those ambiguities cannot be directly solved by a clustering without prior knowledge. It either results in multiple object hypotheses of the bottom left vehicle, demonstrated in Fig. 3b, or, with a larger distance threshold, in one big cluster that actually belongs to multiple objects. Hence, a correct association would require a subsequent split or merge of the object hypotheses formed by the individual clusters. The predicted state of the tracks is required to solve these ambiguities. Fig. 3c shows the correct association using the predicted labeled particle population, which is visualized in Fig. 3d. Our previous approach [22] is demonstrated in Fig. 3e, in which cell occupancy association probabilities are calculated from the bounding box model of the predicted tracks. That approach also solves those ambiguities and results in the same correct association in that case.

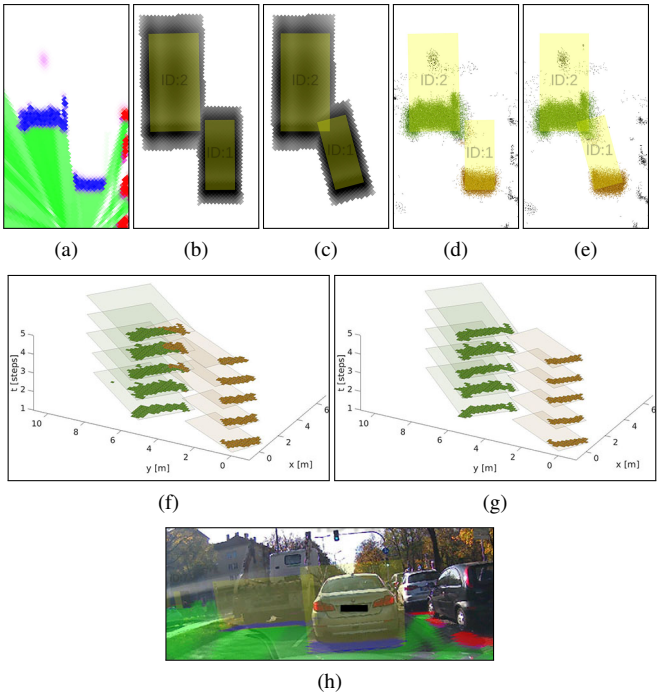
In scenarios with multiple tracks moving closely to each other, however, this box model occupancy likelihood approximation is error-prone to inaccuracies of the predicted track state. Fig. 4 shows such a scenario with two vehicles closely moving in parallel. Their correct estimation is illustrated in Fig. 4b. To demonstrate the error-proneness in case of an incorrect state estimation, an orientation offset of $15°$ is added to the predicted track on the right side (ID 1). Hence, the front left corner of that track overlaps with the rear right corner of the other vehicle (ID 2), as shown in Fig. 4c. Ambiguous cells in the overlapping area can thus be associated to the wrong track, leading to wrong measurement updates, which then further reinforce the incorrect state estimation and thus the association of the next time instance, cf. Fig. 4f. In contrast, the particle population linked to a track represents a detailed occupancy likelihood approximation of that track, see Fig. 4d. This approach is robust against overlapping ares of the box models, cf. Fig. 4e and Fig. 4g, since the particles remain their label from the previous time instance in such a case, i.e., the label is only updated when no ambiguities occur.

Fig. 5 demonstrates a scenario where the particle labeling association, shown in Fig. 5c, fails. Cells of a nearby curb, which are wrongly classified as dynamic, but are not extracted as a separate track, are associated to the closely



(a) (b) (c) (d) (e)

Fig. 3. Scenario with a vehicle (bottom left) with two separate areas of measured occupancy due to a limited sensor field of view. (a) Classified measurement grid. (b) Clustering without prior knowledge. (c) Proposed particle labeling association. (d) Predicted tracks and labeled particle population. (e) Predicted tracks and occupancy association probabilities.

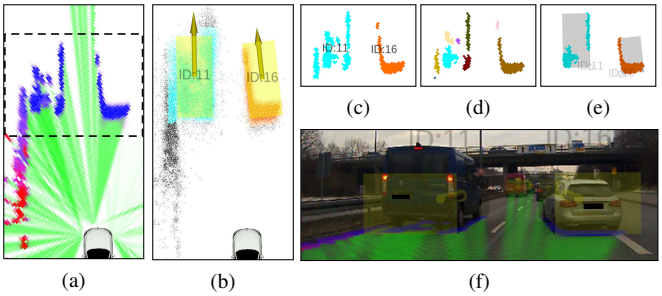


(a) (b) (c) (d) (e)

(f) (g)

(h)

Fig. 4. Scenario with two closely moving vehicles with a modified orientation estimation. Orientation of track 1 is modified in (c), (e), (f), (g) by an offset of $15°$. (a) Classified measurement grid. (b)+(c) Occupancy association probabilities. (d)+(e) Labeled particle population. (f)+(g) Associated cells over time (z-axis) and predicted track box model with $15°$ orientation offset of track 1, (f) corresponds to (c), (g) to (e). (i) Camera image.



(a) (b) (c) (d) (e)

(f)

Fig. 5. Scenario with a vehicle moving closely to a road boundary with a curb that is wrongly classified as dynamic. (a) Classified measurement grid. (b) Labeled particle population and predicted tracks. (c) Particle labeling association. (d) Additional clustering process (color of a cluster is not related to the corresponding track here). (e) Finally associated clusters and resulting measurement minimum bounding boxes. (f) Camera image.

moving track. The additional clustering process, however, distinguishes those cells of the curb as a separate cluster, cf. Fig. 5d. Hence, by evaluating the individual scores of each cluster, that outlying cluster results in a low cluster score and is thus excluded. The selection of valid clusters is chosen conservatively, i.e., a high threshold $\Gamma^{\lambda}_{\text{min}}$ is used such that the score of an cluster has to be similar to the best cluster of that track, cf. (27). A track is updated by an extracted minimum bounding box of the associated cells as described in [22]. Hence, false negatives, i.e., occupied cells of that track which have been omitted in the association, only lead to a smaller, more conservative minimum bounding box that still correctly updates the track. False positives, in contrast, result in an overestimation of the bounding box of that measurement and thus also of that filtered track, which is more critical. This cluster score verification concept can also be extended by evaluating multiple hypotheses, e.g., using MHT or RFS, with different measurement updates of arbitrary combinations of the clusters over time, which, however, also increases the computational effort.

Another challenging urban scenario with multiple surrounding vehicles and an approaching scooter is shown in Fig. 6. Multiple time instances are visualized that illustrate the movement of the scooter, which is very close to the other vehicles and the ego vehicle and thus also causes occlusions. As demonstrated by the different processing steps and the successive time steps, the proposed particle labeling association with the subsequent clustering and verification correctly associated the measured occupied cells to the corresponding tracks and thus successfully distinguishes those different objects. Further results of the proposed association concept and the overall grid-based environment estimation approach are demonstrated in the attached video.

VI. Conclusions

This paper has presented a new data association approach in the context of grid-based object tracking. Particles of the underlying low-level particle tracking, primarily used to estimate the dynamic state of the grid, are linked to filtered object tracks by an attached label. The predicted particle population linked to a track represents its expected areas of measured occupancy, which is used for the association between those measurements and tracks. As shown in various real traffic scenarios, this approach is especially useful to solve ambiguities, e.g., caused by occlusions or dense traffic, where a clustering without using information of the predicted tracks often fails. In contrast to our previous approach, those low-level particles enable a detailed approximation of the occupancy likelihood of a track. Furthermore, based on those individually associated cells, multiple clusters of a track are extracted and analyzed. Wrongly associated cells of neighboring dynamic areas, which are not part of that track, can thus be detected and rejected. Overall, the proposed association approach is basically robust against a wrong dynamic classification, e.g., guardrails close to a track, closely moving objects with an inaccurate prediction, clutter measurements, or occlusions with gaps between clusters.

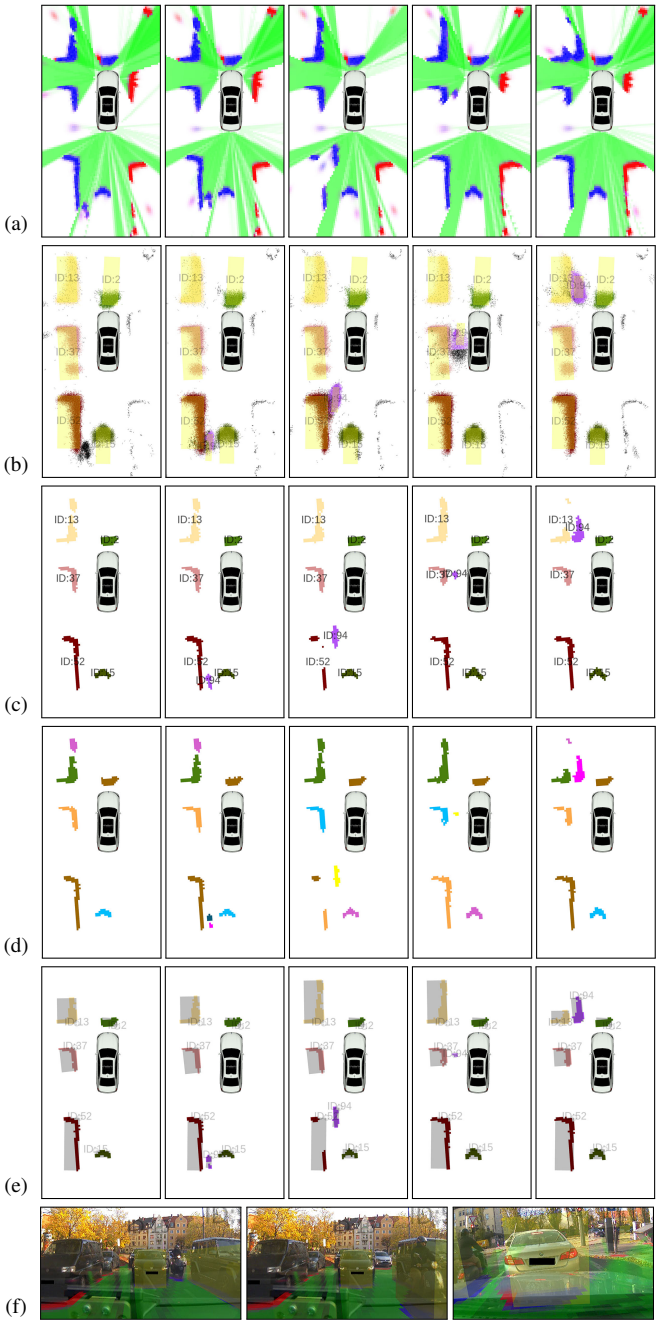


Fig. 6. Scenario with multiple surrounding stopped vehicles and an approaching scooter. Columns show evaluation over time. (a) Classified measurement grid. (b) Predicted tracks and labeled particle population. (c) Particle labeling association. (d) Additional clustering step (color of a cluster is not related to the corresponding track here). (e) Finally associated clusters and measurement minimum bounding boxes. (f) Camera images.

References

[1] K. C. Chang, R. K. Saha, and Y. Bar-Shalom, "On optimal track-to-track fusion," *IEEE Trans. Aerosp. Electron. Syst*, vol. 33, no. 4, pp. 1271–1276, Oct. 1997.

[2] S. Matzka and R. Altendorfer, "A comparison of track-to-track fusion algorithms for automotive sensor fusion," in *Multisensor Fusion and Integration for Intelligent Systems*. Springer, 2009, pp. 69–81.

[3] P. Konstantinova, A. Udvarev, and T. Semerdjiev, "A study of a target tracking algorithm using global nearest neighbor approach," in *Proc. Int. Conf. on Comput. Syst. and Technol.*, 2003, pp. 290–295.

[4] S. S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 19, no. 1, pp. 5–18, 2004.
[5] T. E. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Multi-target tracking using joint probabilistic data association," in *Proc. IEEE Conf. on Decision and Control including the Symp. on Adaptive Processes*, 1980, pp. 807–812.
[6] B. N. Vo, S. Singh, and A. Doucet, "Sequential monte carlo methods for multitarget filtering with random finite sets," *IEEE Trans. Aerosp. Electron. Syst*, vol. 41, no. 4, pp. 1224–1245, Oct 2005.
[7] S. Reuter, "Multi-object tracking using random finite sets," Ph.D. dissertation, Universität Ulm, 2014.
[8] K. Granström and M. Baum, "Extended object tracking: Introduction, overview and applications," *CoRR*, vol. abs/1604.00970, 2016.
[9] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *Computer*, vol. 22, no. 6, pp. 46–57, Jun. 1989.
[10] C.-C. Wang and C. Thorpe, "Simultaneous localization and mapping with detection and tracking of moving objects," in *Proc. IEEE Int. Conf. Robot. Autom.*, vol. 3, 2002, pp. 2918–2924.
[11] J. Moras, V. Cherfaoui, and P. Bonnifait, "Credibilist occupancy grids for vehicle perception in dynamic environments," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 84–89.
[12] T.-D. Vu, O. Aycard, and N. Appenrodt, "Online localization and mapping with moving object tracking in dynamic outdoor environments," in *Proc. IEEE Intell. Veh. Symp.*, 2007, pp. 190–195.
[13] T.-D. Vu, J. Burlet, and O. Aycard, "Grid-based localization and online mapping with moving objects detection and tracking: new results," in *Proc. IEEE Intell. Veh. Symp.*, 2008, pp. 684–689.
[14] M. Bouzouraa and U. Hofmann, "Fusion of occupancy grid mapping and model based object tracking for driver assistance systems using laser and radar sensors," in *Proc. IEEE Intell. Veh. Symp.*, 2010, pp. 294–300.
[15] R. Jungnickel and F. Korf, "Object tracking and dynamic estimation on evidential grids," in *Proc. IEEE Intell. Transp. Syst. Conf.*, 2014, pp. 2310–2316.
[16] K. Mekhnacha, Y. Mao, D. Raulo, and C. Laugier, "The 'fast clustering-tracking' algorithm in the bayesian occupancy filter framework," in *Proc. IEEE Int. Conf. Multisensor Fusion and Integration for Intell. Syst.*, 2008, pp. 238–245.
[17] R. Danescu, F. Oniga, and S. Nedevschi, "Modeling and tracking the driving environment with a particle-based occupancy grid," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1331–1342, Dec. 2011.
[18] L. Rummelhard, A. Nègre, and C. Laugier, "Conditional monte carlo dense occupancy tracker," in *Proc. IEEE Intell. Transp. Syst. Conf.*, 2015, pp. 2485–2490.
[19] D. Nuss, S. Reuter, M. Thom, T. Yuan, G. Krehl, M. Maile, A. Gern, and K. Dietmayer, "A random finite set approach for dynamic occupancy grid maps with real-time application," *arXiv preprint arXiv:1605.02406*, 2016.
[20] G. Tanzmeister and D. Wollherr, "Evidential grid-based tracking and mapping," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 6, pp. 1454–1467, June 2017.
[21] S. Steyer, G. Tanzmeister, and D. Wollherr, "Grid-based environment estimation using evidential mapping and particle tracking," *IEEE Trans. Intell. Veh.*, 2018, submitted.
[22] S. Steyer, G. Tanzmeister, and D. Wollherr, "Object tracking based on evidential dynamic occupancy grids in urban environments," in *Proc. IEEE Intell. Veh. Symp.*, 2017, pp. 1064–1070.
[23] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *Ann. Math. Statist.*, vol. 38, no. 2, pp. 325–339, April 1967.
[24] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
[25] D. Nuss, T. Yuan, G. Krehl, M. Stuebler, S. Reuter, and K. Dietmayer, "Fusion of laser and radar sensor data with a sequential monte carlo bayesian occupancy filter," in *Proc. IEEE Intell. Veh. Symp.*, 2015, pp. 1074–1081.