Fusion of Object Tracking and Dynamic Occupancy Grid Map

Nils Rexin, Marcel Musch and Klaus Dietmayer

Abstract—Environment modeling in autonomous driving is realized by two fundamental approaches, grid-based and feature-based approach. Both methods interpret the environment differently and show some situation-dependent beneficial realizations. In order to use the advantages of both methods, a combination makes sense. This work presents a fusion, which establishes an association between the representations of environment modeling and then decoupled from this performs a fusion of the information. Thus, there is no need to adapt the environment models. The developed fusion generates new hypotheses, which are closer to reality than a representation alone. This algorithm itself does not use object model assumptions, in effect this fusion can be applied to different object hypotheses. In addition, this combination allows the objects to be tracked over a longer period of time. This is evaluated with a quantitative evaluation on real sequences in real-time.

I. INTRODUCTION

For the planning of movements of an autonomous robot, a modeling of the environment is required in order to determine the corresponding action space for a collision-free execution of a movement. All obstacles in the environment have to be considered. In particular, dynamic objects are of interest because they directly influence the free movement space. There are two fundamentally different approaches for such an environment modeling.

The environment of a vehicle, in the area of autonomous driving, is modeled on the one hand by a feature-based approach with object model assumption, which interprets objects as vectors. On the other hand a grid-based approach is used, which discretizes the environment and all objects in cells and is generally object model free. The environment perception is performed by sensors from which specific information for the representations can be derived with varying degrees of accuracy.

In feature-based approaches, which typically include object tracking, the intention is usually to track solely dynamic objects, although static objects could also be modeled. An example for this approach is the tracking according to Reuter et al. [1], in which the labeled multi-bernoulli (LMB) filter is applied. Based on radar measurements for tracking, an object model is assumed for the estimation of pose and movement.

On the contrary, no object model is assumed in the grid-based approach. For this approach there is a static modeling by an occupancy grid map, see [2], [3], or its extension for estimating dynamic objects with the dynamic occupancy grid map (DOGMa), as in [4], [5]. So, it is possible to estimate the position in static case and the orientation and movement in dynamic case by a particle filter based on laser measurements.

Both approaches model objects and the environment with different focus. Especially with sparse radar measurements, object hypotheses can be generated by tracking through object model assumptions, whereby continuous tracking is difficult due to a situation. Likewise, the dynamic grid map lacks the expansion of the objects and an accurate tracking of the whole objects, but the form of objects are well presented.

A method that combines both forms of representation and enables a smooth transition into each other is desirable. Various fusion methods have been proposed for this purpose [4], [6]–[10].

The key contribution of this paper is a fusion, which does not make any object assumptions, but uses information from the association of an object hypothesis and grid cells to generate new hypotheses. Therefore, any dynamic objects can be tracked during this fusion. Thus, the fusion algorithm is so modular that the environment modeling is decoupled from it and the logic is purely in the fusion. The evaluation shows that the presented fusion allows a more accurate object estimation and a longer temporal tracking of objects.

In order to bring the two representations together, a common basis can be found or these can be related, as with the algorithm presented here. A feature vector can serve as a basis. In the paper by Steyer et al. [9] object hypotheses are directly created based on the DOGMa. Here, however, no fusion is carried out, but only a hypothesis is created on a grid map. The approach of Bouzouara [6] exchanges information between models to improve estimates and create a common list of object hypotheses. For this purpose, model information is also exchanged and only the distance measurement is examined. A more comprehensive fusion with additional context information is performed on [7], [8]. Gies et al. [7] proposes an architecture in which, meta object lists are created as a common basis that are fused on the basis of model-based and constrain-based confidence measures. For this purpose, object hypotheses are created from the DOGMa, which, together with context information and object hypotheses, are fused into a high-level fusion. The approach is also structured in such a way that the environment representations are not adapted.

II. ENVIRONMENT MODELS

This section introduces the representations for the two environment models and points out their differences. This is followed by a concept for the fusion of representations.

For the grid-based approach, the implementation of the dynamic occupancy grid map (DOGMa) according to
Nuss et al. [4] is used. The environment is divided into 15 cm square cells. By this division objects are represented as a set of cell-size point objects, whereby in each cell at most only one such point object exists. Thus, this representation is without object model. As sensor for the DOGMa in this paper an ibeo LUX four layer laser scanner is used with a horizontal angle of view (AOV) of 85° with four layers, further 25° with two layers, a horizontal angular resolution of 0.25°, vertical AOV of 3.2° and a range up to 200 m. Measurements generated by this laser scanner are inserted into a grid map by an inverse sensor model [3]. This measurement grid map serves as an input for the DOGMa, in which the 2D velocity for each cell is estimated by a particle filter in the background. Fig. 1 shows the DOGMa, in which the occupancy probability for each cell is represented by the intensity of the gray scale value. The colored cells illustrate the direction of movement of dynamic point objects, correspondingly coded with the color circle at the bottom right.

Based on the subdivision into cells, the DOGMa does not require an explicit object model assumption, but the whole environment. In general the effort to calculate the particle filter is high and therefore a simple motion model, the constant velocity (CV) model [11], was chosen to keep the state space for the particle filter small. Accordingly, the particle state space comprises two dimensions each for position and velocity. The sum of the weights of the particles of a cell represents the probability of their occupancy. The state vector for a cell consists of the occupancy probability and the two-dimensional position as well as velocity.

In the feature-based approach, objects of an environment are modeled as a vector, i.e. a point, in the state space. A representative for this approach is the LMB tracking according to Reuter et al. [1], which is used in this paper. To identify objects from the data of a measurement, models are used for object assumption. In this work, the tracking is based on sensor data of a long-range radar (LRR) and an ibeo LUX four layer laser scanner, the same as for DOGMa. As object model, a box for a car is assumed. Specifically, this LMB tracking uses a point object tracking, i.e. each object generates at most one measurement at a time step (see [12]). The LRR provides a maximum of one measurement for just one object, which can then be incorporated directly into the LMB tracking. Due to its use as a long-range radar, it has a horizontal AOV of 30°, vertical AOV of 5° and a range up to 250 m. This means that objects in the vicinity can only be detected in a small area. Thus, the LRR is designed for the use case of an adaptive cruise control (ACC) and rather for longitudinal tracking. In contrast, the ibeo laser scanner has an AOV of 85° with four layers and covers a larger area (see specification above at DOGMa). A laser scanner generates more than only one measurement for an object. In order to get only one point for an object for tracking, we preprocess the laser data with a box fitting. This algorithm is not very robust, so that there are problems with the calculation of the box hypotheses and thus also influences the quality of the object hypotheses. It shall be mentioned again that this is a concept and therefore the algorithms are exchangeable. Only any algorithm for the environment modeling is needed. Initially, a hypothesis is created based on the radar measurement and updated by both LRR and ibeo laser box hypotheses. Overall, only dynamic objects are tracked here, since this has a direct influence on the vehicle’s space of movement.

Due to the feature-based approach and the accompanying object representation as a point, there is scarce data in the environment model. For this reason, only a small amount of object hypotheses needs to be calculated, as opposed to the DOGMa approach where each cell that can be an object is calculated by a set of particles. Therefore, tracking uses a more complex motion model, the CTRV model [11].

In summary, the two models in Fig. 2 are illustrated with reference to their arbitrary complexity over the distance, starting from the ego vehicle. Particularly in the immediate vicinity there are objects with versatile and amorphous forms, such as bushes, artworks and pedestrians with prams, scooters or wheelchairs, which can be captured in detail by high-resolution laser scanners. Due to the high level of detail in the vicinity, many measurements are also available. As the distance increases, the sensor coverage area and thus the number of measurements as well as the perceived details decrease. Radars in particular can still detect objects far in the distance, but provide comparatively fewer measurements than laser scanners.

Fig. 2 shows that the complexity of the grid-based approach increases approximately squarish with distance. With the feature-based approach, the complexity increases strongly in the near environment and then flattens out with increasing distance, since with the feature-based approach a large number of object models have to be associated in the detailed measurement. In the distance, only a small amount of data is added, so that only a small amount of more object models need to be matched.

The grid-based approach represents the environment as a square grid map, which also explains the quadratic complexity. This has advantages especially in the vicinity. When creating the grid map, no object models are assumed, just
the measurements are summarized in discrete areas. Since all objects are divided into cells, any form can be mapped as accurately as possible, depending on the cell size. This increases the calculation effort even if only a few measurements are available, because with this approach the complete environment is modeled.

In contrast, for an object model-based approach, such as the feature-based approaches, it is challenging to assign detailed resolved forms and unstructured data to the given object models. Among other things, a lot of preprocessing effort is required to fit, e.g. box models, into large amount of data in the vicinity. On the contrary, in distant areas mainly objects such as vehicles are interesting as object models and the data is sparser, a matching of the data to object models is less complex. Areas far away are covered in particular by LRRs that generate only a small amount of measurement data.

In summary, the dynamic grid approach is advantageous for the near area and the object based approach for the far area. It is desirable to have an approach that combines the two methods in order to use the benefits of both. For the smooth transition of these areas, including the urban area, both environment models should match their information. This is exactly where this work begins and describes a concept for a fusion of the two models.

III. ALGORITHM OF FUSION

A. Concept of Fusion

The algorithm presented in this paper proposes how the information from the DOGMa and LMB tracking techniques can be combined. At the beginning, an overview of the developed algorithm is given before the implementation is discussed in more detail. For clarification it should be mentioned once again that two existing methods are used here and the focus of the presented algorithm is on the possibility of fusion of information from both techniques. The representations of these two methods therefore serve as input for the developed concept. Fig. 3 schematically illustrates this concept for the fusion algorithm. The scheme shows that the modules of the fusion have no effect on the representations. This fusion is described as an object model independent fusion, since in fusion and especially in information adaptation the individual cells of the DOGMa are used. Thus, the fusion can be applied to different object types.

Information about cells is known from the DOGMa and from tracking there are point object hypotheses, which also describe an extension of the objects. Thus, for a fusion an association between grid cells and object hypotheses is necessary. It should be noted that different information about an object can exist through the environment representations, leading to several hypotheses. For the final result the best fitting hypothesis is selected from the hypothesis of object tracking, fused hypothesis and the predicted hypothesis. The individual modules are explained in the following.

B. Association Between Grid Cells and Object Hypotheses

At the beginning of the algorithm, the grid cells must be assigned to the tracking object hypothesis. For this purpose, it is possible to generate object hypotheses from the DOGMa, which are then compared with the tracking result, as is done by clustering at [7], [9]. However, this produces incorrect measurements (e.g. false-positives) which have to be corrected in further steps. Therefore, the corresponding cells of the grid map are identified here on the basis of the object hypotheses for simplification. This is due to the fact that information is available from both techniques for a fusion and the focus is on this. In this way, the cells are examined under the object hypothesis and in the immediate vicinity for dynamic cells, because only the dynamic cells are relevant. In addition to the found cells, all those dynamic cells are added, which are connected to these cells by a 8-connected neighborhood. Dynamic cells are defined as those which have a occupancy probability greater than 0.7 and whose velocity is clearly in one direction, with a design parameter. For the expression of the velocity the distance by the mahalanobis distance between the velocity vector of a cell and the static, i.e. the zero velocity, is used [4]. Two examples of the association is shown in Fig. 4. Thereby parts of the DOGMa are pictured and the object hypothesis of tracking as a red box. Figure 4a illustrates a hypothesis that was initially incorrectly oriented. In the association, the green cells below are linked to this hypothesis by the neighborhood relationship.

If no cell is found under the object hypothesis, the area larger by a factor of 1.5 of object length is inspected. This is for the case that the rotation or position of the hypothesis
does not match the representation of the DOGMa, then an assignment of dynamic cells to the object hypothesis of the tracking can take place nevertheless. This case is shown in the Fig. 4b. The object hypothesis (red) lies next to the dynamic cells. A green box for the fusion hypothesis is shown. This result is based on the adaptation of the associated cells. The orientation of the hypothesis is fitted to the form in the DOGMa.

C. Realization of Fusion

Due to the representations, the estimated states of the object hypotheses of the tracking may differ from those of the associated cell group. Strictly speaking, the average of velocity and orientation over this cell group is calculated and these may have different values. Especially when initializing object hypotheses of tracking, for example, cross-traffic is set up misaligned due to the point object tracking, the assumption that the car hypotheses are aligned longitudinally and the low coverage of the LRR in the intersection area (see Fig. 4a). Whereas based on the object model free approach of DOGMa, the measurements are registered according to the cells and thus reflect the occupancy of the environment independent of object assumptions, whereby the form is discretized in cells. Therefore, the form of the objects is more realistic than represented by the tracking approach. This is presented in Fig. 5, with a section of the DOGMa close to the ego vehicle with hypotheses and the corresponding video frame shown.

In both images of Fig. 5 are illustrated the object hypothesis of tracking in red and those of the fusion in green. This is a sample of sequence #2 and the object tracking is only using LRR data. In the video frame (Fig. 5b) the hypotheses for illustration are drawn as 3D boxes in the image. This shows that the hypothesis of the fusion is more realistic. In order to see this in the DOGMa, Fig. 5a shows the section at the same time. Additionally there is a red dot with a line in the image. This is an LRR measurement which is used by object tracking. The estimation of the cells goes to the left according to the color circle. Since the vehicle first entered the sensor area of the LRR, a longitudinal hypothesis was first set up by the tracking. For this reason, the information of the associated cell group can be used to adjust the position and orientation of object hypothesis.

The orientation is determined according to the averaged velocity vector of the cell group and the object hypothesis of tracking is rotated with these. For the position, it is determined which corner point is visible from the vehicle. Starting from this corner point, the hypothesis is shifted from the tracking.

The estimation of the velocity and thus also the orientation is based on a particle filter, which uses the CV model, resulting in difficulties when driving around curves. Hereby, the orientation in the direction of the boundary of the curve is estimated. In this case the tracking can better track the orientation due to the used CTRV model. For this reason, the object hypothesis is held as an additional hypothesis.
The presented approach is based on the existence of an object hypothesis of tracking. This is also necessary for the initialization. However, after the cells have been associated to the hypothesis, the selected hypothesis can be predicted for the next time step and included in the calculation as a possible hypothesis.

D. Selection of Object Hypothesis

In the presented approach, two further hypotheses were generated for one object hypothesis. Now a criterion is required for the selection of the more precise hypothesis as to reality. In general, such an evaluation is difficult, as some assumptions have already been made for measurements, movements or objects for the state estimation.

In the used tracking approach, point objects are tracked based on a single measurement. Therefore, no orientation can be made directly, but rather by movement and model assumptions. In addition, the object is abstracted to a point. In the grid-based approach, as described in the previous section, the form of the objects is discretized by cells and thus a set of cells represents an object. This produces a more realistic representation than a feature-based approach. Consequently, the underlying cells are examined for the selection of the object hypothesis.

Finally, the number of dynamic cells below an object hypothesis is the decisive criterion. Cells that are regarded as static are excluded because false positive hypotheses are so filtered.

IV. Evaluation

The fusion presented in this paper is implemented with C++ and is evaluated by real experimental sequences. The Ulm University own test vehicles are used for this research. A qualitative evaluation is carried out on a sequence of a test area with up to three test vehicles, so that the results can be compared to ground truth data. Each test vehicle is equipped with a differential global positioning system (DGPS) sensor, which ensures maximum accuracy. The ego vehicle is equipped with an ibeo LUX four layer laser scanner and an LRR in the center of the front. The DOGMa uses the laser scanner data. Input for the LMB tracking is either only the LRR data or the ibeo data in combination with a box fitting and the LRR data.

To illustrate the sequences, the routes of the reference vehicles and the ego vehicle are shown in Fig. 6. The sequences show complex situations in which the vehicles drive out or into the sensor area. Further, the vehicles obscure each other and the measurements were carried out at a long distance. In the first sequence, no route of the ego vehicle is drawn, because it is stationary. The graphics show the chronological course of the routes via color scale. The positions are given in ego-stationary coordinates.

The results of the qualitative evaluation are noted in Table I. The columns are for the respective RMSE for the vehicles V1 and V2 as well as the tracking time of the objects. For each sequence, the values were evaluated for LMB tracking with LRR only and in combination of LRR and ibeo laser scanner. The results for position and orientation are rounded to two and for time to one decimal places. Furthermore, there is a line marked with symbol %, in which the percentage improvement of the error as well as the percentage extension for the track time are entered.

The developed fusion requires an object hypothesis at the beginning. After the cells have been associated with the hypothesis, the fusion algorithm can hold the hypothesis over a longer period of time. This can be seen in the last columns of Table I. The duration of tracking is determined by accumulating the times, if the time difference between two adjacent time steps is less than 100 ms. This means that the hypotheses must be continuous.

Thus, there are more time steps with hypotheses from the fusion than from the tracking. To enable now a comparison of position and orientation between the two methods, only the hypotheses are compared if a hypothesis of both methods exists, with a corresponding time tolerance. The differences in the number of measurements can be seen in Fig. 7.

Up to 3.8 times more was achieved in the duration of object tracking. The graphs of Fig. 7 are shown in detail. In this figure the error due to the euclidean distance to ground truth data of sequence #1 is plotted. It can be seen that the results of the tracking based on LRR and ibeo data are inferior than with the LRR alone, because there is not enough data available to perform a good box fitting. Tracking with LRR benefits from the long range of the sensor and model assumptions. It is sufficient for the fusion algorithm if an object hypothesis appears temporarily, as is the case for some time step, see Figs. 7c and 7d in the tail part.

For visualization purposes, sections of the DOGMa from sequence #1 at different distances with object hypotheses from tracking (red), fusion (green) and ground truth data (blue) are shown in Fig. 8. For the object tracking only the LRR data are used as input. Fig. 8 is schematically drawn, so that the overlap of the hypothesis is better recognizable. In the background lies the DOGMa, so that the cells and
their amount can be seen as well. In this figures not all grid

cells are below the ground truth box, because the DOGMa
itself contains inaccuracies. So an error is accumulated
by the sensor measurement, the measurement model as well
as by the discretization and a small difference due to the
processing time. Fig. 8a shows a section of the near area,
therefore many grid cells are occupied. In this area the object
tracking in combination of LRR and ibeo laser scanner is
well (see at beginning of Fig. 7). In contrast, only a few
cells are occupied in the far distance (see Fig. 8b). The LRR
is still provides measurements here, so that hypotheses are
also generated by tracking (red). In combination with the
ibeo laser data, hypotheses are no longer generated here
because the data is too sparse (see at end of Fig. 7). For
this reason, the tracking with combination of LRR and ibeo
laser data is worse for this case. This would not create any
hypotheses and therefore no red boxes would be displayed.
For comparison, this is shown in Fig. 7. At the beginning
there are object hypotheses for both variants and at the end,
i.e. in the distance, there are only hypotheses for tracking
with LRR.

### Table 1

| Sequences | RMSE of Position [m] | RMSE of Orientation [rad] | Duration of Tracking [s] |
|-----------|----------------------|---------------------------|-------------------------|
|           | V 1                  | V 2                       | V 1                     | V 2                       |
| #1 LRR    |                      |                           |                         |
| O_T       | 1.12                 | 2.08                      | 0.26                    | 0.21                      |
| O_P       | 0.75                 | 1.63                      | 0.18                    | 0.20                      |
| %         | 33.04                | 21.63                     | 30.77                   | 4.76                      |
| #1 LRR & ibeo |                  |                           |                         |
| O_T       | 1.34                 | 1.10                      | 0.17                    | 0.22                      |
| O_P       | 0.68                 | 0.59                      | 0.13                    | 0.21                      |
| %         | 49.25                | 46.36                     | 23.53                   | 4.55                      |
| #2 LRR    |                      |                           |                         |
| O_T       | 0.51                 | 3.18                      | 0.02                    | 0.22                      |
| O_P       | 0.20                 | 0.23                      | 0.02                    | 0.02                      |
| %         | 60.78                | 92.77                     | 0.00                    | 90.91                     |
| #2 LRR & ibeo |                  |                           |                         |
| O_T       | 0.43                 | 0.76                      | 0.02                    | 0.06                      |
| O_P       | 0.19                 | 0.48                      | 0.02                    | 0.01                      |
| %         | 55.81                | 36.84                     | 0.00                    | 83.33                     |
| #3 LRR    |                      |                           |                         |
| O_T       | 0.54                 | 1.29                      | 0.15                    | 0.21                      |
| O_P       | 0.40                 | 0.41                      | 0.09                    | 0.11                      |
| %         | 25.93                | 68.22                     | 60.00                   | 47.62                     |
| #3 LRR & ibeo |                  |                           |                         |
| O_T       | 0.16                 | 0.84                      | 0.15                    | 0.09                      |
| O_P       | 0.38                 | 0.68                      | 0.09                    | 0.07                      |
| %         | 67.24                | 19.05                     | 60.00                   | 22.22                     |

**Fig. 8.** Schematic DOGMa fragment with object hypothesis from tracking (red) with only LRR sensor, fusion (green) and ground truth (blue) from sequence #1.

**Fig. 7.** Comparison of the results for the position and orientation error in sequence #1 with both sensor variations for tracking.
The evaluations result in a reduction of the error in the position by up to 92% and in the orientation by up to 90%. This high improvement is due to the fact that only sparse information is available in some object hypotheses and thus position or orientation are not correctly estimated. This applies in particular to sequence #2. This can be seen in detail for the orientation in Fig. 9. In contrast to sequence #1 the tracking with LRR and ibeo laser scanner is now better and keeps the object hypothesis longer and more accurate.

In the quantitative evaluations, which were carried out on real sequences, it was shown that object hypotheses can be held by the developed fusion algorithm over a longer period of time and the positioning as well as orientation become more accurate. This fusion approach improves the results in general significantly. In addition, any dynamic objects can be tracked since no object assumption was made for the fusion itself. In the implementation, this algorithm is decoupled from the environment models, so that no adjustments are necessary. The fusion was presented with a vehicle tracking, since the developed fusion is independent of object models, it can be used for pedestrians or other dynamic objects.

In future work, various movement models can be used for predicting the previous hypotheses. In addition, an estimate of the extent of objects can be made on the basis of the DOGMa, as suggested in [7].

### ACKNOWLEDGMENT

In the context of the research of the second author within the thesis to obtain the Bachelor of Engineering, this work is based on that.

In addition, this research is accomplished within the project "UNICARagil" (FKZ 16EMO00290). We acknowledge the financial support for the project by the Federal Ministry of Education and Research of Germany (BMBF).

### REFERENCES

[1] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietmayer, “The Labeled Multi-Bernoulli Filter,” 2014 IEEE Transactions on Signal Processing, vol. 62, no. 12, pp. 3246–3260, 2014.
[2] A. Elfes, “Using Occupancy Grids for Mobile Robot Perception and Navigation,” Computer, vol. 22, no. 6, pp. 46–57, 1989.
[3] S. Thrun, W. Burgard, and D. Fox, Probabilistic Robotics, ser. Intelligent robotics and autonomous agents series. Cambridge, Mass.: MIT Press, 2006.
[4] 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,” CoRR, 2016.
[5] G. Tanzmeister, J. Thomas, D. Wollherr, and M. Buss, “Grid-based Mapping and Tracking in Dynamic Environments using a Uniform Evidential Environment Representation,” in 2014 IEEE International Conference on Robotics and Automation (ICRA), pp. 6000–6005.
[6] M. E. Bouzouara and U. Hofmann, “Fusion of Occupancy Grid Mapping and Model Based Object Tracking for Driver Assistance Systems using Laser and Radar Sensors,” in 2010 IEEE Intelligent Vehicles Symposium. IEEE, 2010, pp. 294–300.
[7] F. Gies, A. Danzer, and K. Dietmayer, “Environment Perception Framework Fusing Multi-Object Tracking, Dynamic Occupancy Grid Maps and Digital Maps,” in IEEE Intelligent Transportation Systems Conference. Piscataway, NJ: IEEE, 2018, pp. 3859–3865.
[8] D. Nuss, M. Stuebler, and K. Dietmayer, “Consistent Environmental Modeling by use of Occupancy Grid Maps, Digital Road Maps, and Multi-Object Tracking,” in 2014 IEEE Intelligent Vehicles Symposium (IV), pp. 1371–1377.
[9] S. Steyer, G. Tanzmeister, and D. Wollherr, “Object Tracking Based on Evidential Dynamic Occupancy Grids in Urban Environments,” in 2017 IEEE Intelligent Vehicles Symposium. Piscataway, NJ: IEEE, 2017, pp. 1064–1070.
[10] A. Vatavu, et al., “Environment Estimation with Dynamic Grid Maps and Self-Localizing Tracklets,” in 2018 IEEE Intelligent Transportation Systems Conference. Piscataway, NJ, 2018, pp. 3370–3377.
[11] R. Schubert, E. Richter, and G. Wanielik, “Comparison and Evaluation of Advanced Motion Models for Vehicle Tracking,” in 2008 International Conference on Information Fusion, 2008, pp. 1–6.
[12] K. Granström, M. Baum, and S. Reuter, “Extended Object Tracking: Introduction, Overview and Applications,” CoRR, 2017.