



HAL
open science

2.5D Evidential Grids for Dynamic Object Detection

Hind Laghmara, Thomas Josso-Laurain, Christophe Cudel, Jean-Philippe Lauffenburger

► **To cite this version:**

Hind Laghmara, Thomas Josso-Laurain, Christophe Cudel, Jean-Philippe Lauffenburger. 2.5D Evidential Grids for Dynamic Object Detection. 22th International Conference on Information Fusion (FUSION) 2019, Jul 2019, Ottawa, Canada. 10.23919/FUSION43075.2019.9011417. hal-04131742

HAL Id: hal-04131742

<https://hal.science/hal-04131742>

Submitted on 16 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/337171728>

2.5D Evidential Grids for Dynamic Object Detection

Conference Paper · July 2019

DOI: 10.23919/FUSION43075.2019.9011417

CITATIONS

5

READS

632

4 authors:



Hind Laghmara

Institut National des Sciences Appliquées de Rouen

8 PUBLICATIONS 71 CITATIONS

SEE PROFILE



Thomas Josso-Laurain

Université de Haute-Alsace

30 PUBLICATIONS 148 CITATIONS

SEE PROFILE



Christophe Cudel

Université de Haute-Alsace

51 PUBLICATIONS 219 CITATIONS

SEE PROFILE



Jean-Philippe Lauffenburger

Université de Haute-Alsace

87 PUBLICATIONS 613 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Autonomous vehicle [View project](#)



Navigation Aided Intelligent Cruise Control (NAICC) [View project](#)

2.5D Evidential Grids for Dynamic Object Detection

Hind Laghmara, Thomas Laurain, Christophe Cudel and Jean-Philippe Lauffenburger

IRIMAS EA7499, Université de Haute-Alsace

Mulhouse, France

firstname.name@uha.fr

Abstract—Perception is a crucial and challenging part of Intelligent Transportation Systems. One of the main issues is to keep up with the moving objects in complex and dynamic environments. This paper proposes a method for dynamic object detection using Evidential 2.5D Occupancy Grids. The approach is based on a map representation for occupation modeling and navigable area definition. At each time step, a local grid is derived from the sensor data. Belief Theory is then retained to perform a grid fusion over time in order to keep track of the moving objects in the grid. The description of the dynamic behavior of objects in a scene is related to the conflict issued after the temporal fusion. Finally, the construction of the objects themselves is realized with a segmentation based on Density-Based Spatial Clustering of Applications with Noise (DBSCAN). In order to validate the efficiency of the proposed approach, experimental results are provided based on the KITTI dataset. Performances are evaluated through comparison with the ground truth.

Index Terms—Occupancy Grids, Belief Theory, Dynamic Object Detection, Autonomous Vehicles, LiDAR.

I. INTRODUCTION

This paper focuses on the perception of Intelligent Transportation Systems for which the objective is to model the vehicle’s local environment according to data issued from multiple sensors. This paper aims to achieve such task considering that the vehicle’s pose is known. The surrounding environment representation is based on Occupancy Grids (OG) as it indicates two main features: the navigable space as well as the location of obstacles which may be static or dynamic. OGs can be constructed based on the measures of the objects’ distance from the ego-vehicle which can be given by exteroceptive sensors such as LiDARs (Light Detection And Ranging), Radars or stereo vision.

The main challenges considering this issue are the uncertainty and imprecision of information as well as the complexity that lies behind modeling a dynamic scene. Multiple Object Tracking (MOT) is the application allowing to handle this issue which includes the detection, association and tracking of dynamic objects sequentially. OGs have proven to be effective with a limited computation complexity for local environment modelling taking account of temporal data. The aim in this paper is to improve OGs to insure a robust detection of dynamic objects. Due to the high uncertainty in such task, the use of Dempster-Shafer theory is convenient for autonomous systems’ applications.

Some of the main references in the literature to treat the detection of dynamic objects with grid-based solutions are reviewed in Section II-A. A survey is also presented in [1] where a 2.5D (two dimensions plus the elevation information obtained by averaging the height of all points that fall into a given cell) approach is used for the determination of moving cells.

This paper will be based on the same representation to include a tri-dimensional modeling of the environment. The main contribution of this approach is that the classification of dynamic objects in a 2.5D grid is done according to an evidential fusion of multi-grids. The second aim was to extract an object-level representation according to the detected dynamic cells for tracking purposes. The objective is to achieve the dynamic detections on an object-level rather than on a cell-level which is commonly done in literature. The objects are built based on the clustering of mobile cells using the DBSCAN algorithm. The third contribution of this work is a quantitative evaluation according to a measure of average precision of the detection results based on a KITTI dataset for comparison purposes.

The paper is structured as follows: Section II covers a survey on multiple object detection based on occupancy grids as well as the definition of a 2.5D representation. Section III introduces the different steps allowing the transition from a grid-level representation to the dynamic object detection. This includes the definition of an evidential grid as well as mobile cells labeling (cf. Fig. 1). The segmentation algorithm is also specified in order to extract dynamic objects. Section IV illustrates the dataset used for evaluation as well as a quantitative result analysis. Section V concludes the paper.

II. 2.5D GRID MAPS

A. Related Work

An OG is a representation which employs a multidimensional tessellation of the space into cells where each one stores a knowledge of its state of occupancy [2]. Today, there is a large use of OGs due to the availability of powerful resources to handle their computational cost. The construction of a grid has been applied in multiple dimensions (2D, 2.5D and 3D) [1] using different sensor technologies like 2D radars, 2D or 3D LiDARs and stereo-vision. In this representation, each cell-state is described according to a chosen formalism. The most common one is the Bayesian framework which

was adopted first by Elfes [2] followed by many extensions as the well-known Bayesian Occupancy Filter (BOF) [3, 4]. The latter estimates the dynamics of the grid cells using the Fast Clustering and Tracking Algorithm in order to ensure MOT [5].

Other works suggested a formalism based on Dempster-Shafer (or Evidence) Theory. It has been applied in [6] based on a 2D occupancy grid using a ring of ultrasonic transducers and a sensor scanner. Moras *et al.* proposed a similar approach also used for mobile object detection based on an inverse sensor model [7, 8]. The latter is realized according to the conflict analysis by a temporal evidential fusion of multiple grids. Extending Moras *et al.*'s work, contextual discounting is applied in [9] to control cell remanence.

Some references study the dynamics of the environment at the cell level to avoid the inconsistencies of the object representation [10]. Tanzmeister *et al.* [11] also estimate the static and dynamic characteristics at the grid cell level and use a particle filter for obtaining the cell velocity distribution. Honer *et al.* [12] focus on the classification of stationary and dynamic elements based on an evidential semantic 2D grid map with a five state cell configuration: each cell can either be free, dynamic, static, occupied or unknown. However, the update of cells is done according to a combination table heuristically determined.

The above literature review and especially [1] show that most of the works consider a two-dimensional grid for the environment representation even when 3D sensors are providing the data to build the map. In fact, 3D solutions like voxel grids or octomaps can generate high complexity and computation load when applied to real-time applications like autonomous navigation. An interesting tradeoff remains in 2.5D occupancy grid which are known to be memory efficient and at the same time store elevation data. In the particular context of autonomous driving and ITS, in which the elevation variation of the terrain is limited in the local area in which the vehicles are driving, 2.5D representations are of real interest and are retained here.

In this work, the objective is to consider an object-oriented tracking which necessitates an efficient object detection module. The idea is to consider the tri-dimensional sensor data issued by a Velodyne LiDAR to built at each time step a 2.5D grid for which the elevation is attributed to each cell. Sections II-B and III describe the approach illustrated in Fig. 1.

B. Building a 2.5D Grid

The pre-processing step from Fig. 1 required to build a 2.5D grid is derived from [1]. The grid is composed of discrete cells in which the object height above the ground level is stored. This representation can describe the elevated objects from the ground which can correspond to dynamic or static objects.

Building the 2.5D grid includes defining the covering area as well as its resolution, which corresponds to the dimensions of each cell. An example of a 2.5D grid map is shown in Fig. 2 where the resolution is $0.4 \times 0.4m$. The grid covers 40m in

front, 20m behind and 20m along right and left sides of the vehicle.

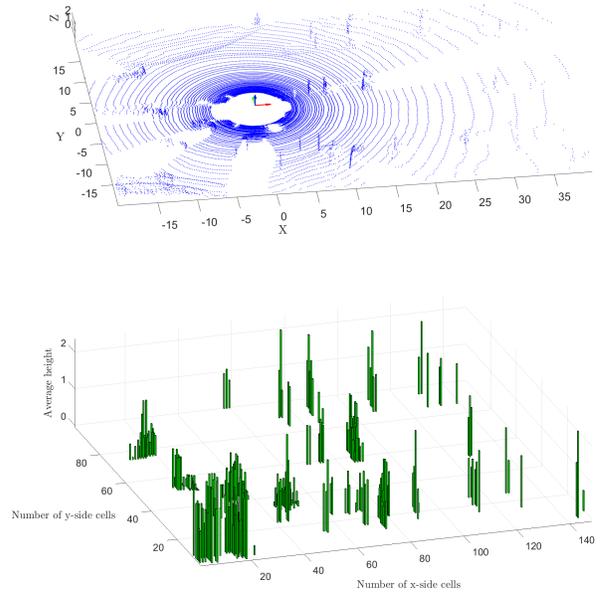


Fig. 2. Top: 3D LiDAR point cloud from KITTI, Bottom: Corresponding 2.5D grid with $0.4 \times 0.4m$ cells.

In order to consider the elevation of objects, it is necessary to determine all measures that correspond to the ground. Several approaches as [13] treat this point because it can induce errors when investigating the occupancy. Such cases are very frequent when the road is uneven or tilted. In this work, the method presented in [1] is employed. It consists in evaluating the variance of height of the points which correspond to a cell. If this value is larger than a threshold, the average height is verified to surpass a defined value in order to make sure it belongs to the ground surface and not any other planar surface. This is equivalent to the following statement:

$$G(i, j) = \begin{cases} 0 & \text{if } \sigma_{i,j}^2 < tr_{\sigma} \text{ and } \mu_{i,j} < tr_{\mu} \\ \mu_{i,j} & \text{otherwise} \end{cases}, \quad (1)$$

where $\mu_{i,j}$ and $\sigma_{i,j}^2$ are the average height and its variance in cell with index (i, j) . The thresholds tr_{σ} and tr_{μ} defined in [1] are respectively equal to $2cm$ and $30cm$. G is the resulting 2.5D grid which will be further used for object detection.

III. FROM AN EVIDENTIAL GRID TO THE OBJECT LEVEL

A. Modeling an Evidential Grid

Extending the probability theory, the Belief Theory offers an adequate representation of the data and source imperfections and thus is appropriate for perception in ITS. It offers a wide range of fusion operators handling these properties according to the application.

In this work, the solution from [7] is adapted to the 2.5D grid. Moras *et al.* suggest an approach based on the conflict

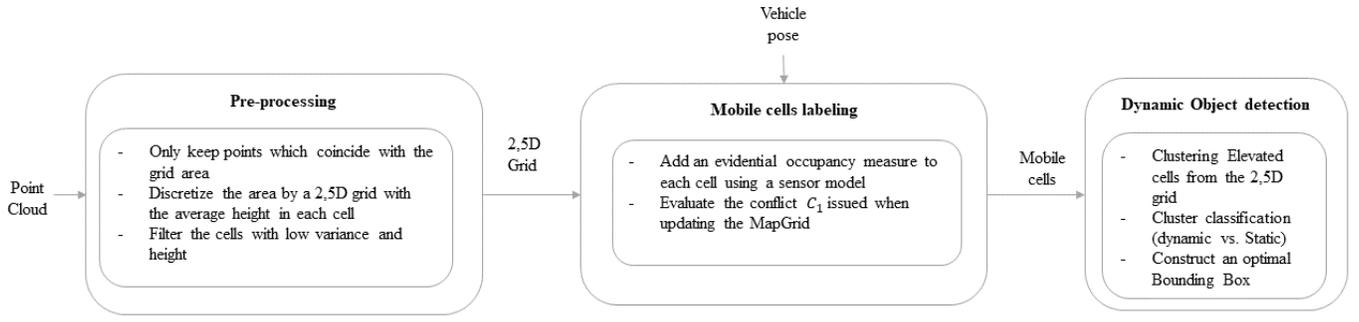


Fig. 1. Dynamic object detection with an evidential 2.5D grid.

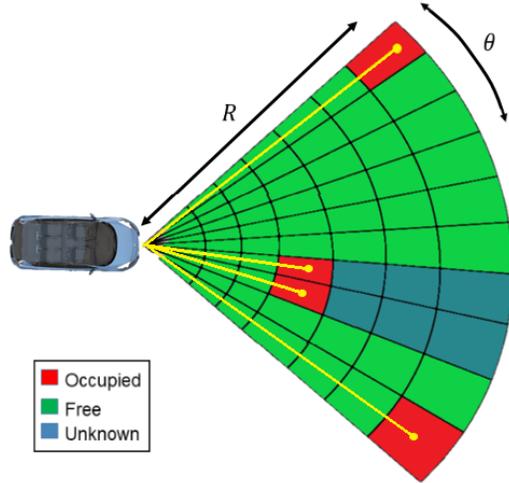


Fig. 3. Polar representation of an occupancy map according to LiDAR data where R is the range of a cell and θ its angular sector [14].

appearing during the temporal grid fusion for mobile object detection and navigable space determination. For that, a frame of discernment is defined to include the states of a cell considering it to be Free (F) or Occupied (O). The discernment frame is then $\Omega = \{F, O\}$.

The referential power set contains all possible combinations of the discernment frame hypotheses: $2^\Omega = \{\emptyset, F, O, \{F, O\}\}$. To express the belief in each state, a mass function $m(\cdot)$ is defined to respectively express conflict $m(\emptyset)$, Free state $m(F)$, Occupied state $m(O)$ and the unknown state $m(\{F, O\})$.

B. Inverse Sensor Model

Basically, a *sensor model* is how the *mass function* of a state according to a measure is calculated. This basic belief assignment (*bba*) also includes the reliability of the source. In this application, the considered sensor is a 3D multi-echo LiDAR provided by Velodyne. The input data will include ranges r_i and angles θ_i of each laser beam or point p_i as shown in Fig. 3.

According to this set of data, a Scan Grid (SG) in polar coordinates is constructed. Each row of this SG corresponds

to an angular sector $\Theta = [\theta^-, \theta^+]$ for which a cell is defined in $R \times \Theta$. The range of a cell is $R = [r^-, r^+]$ which means that each cell is defined by a pair on which a mass is attributed as $m\{\Theta, R\}$. The masses corresponding to each proposition $A \in \Omega$ are found hereby [7]:

$$m\{\Theta, R\}(\emptyset) = 0 \quad (2)$$

$$m\{\Theta, R\}(O) = \begin{cases} 1 - \mu_F & r_i \in R \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

$$m\{\Theta, R\}(F) = \begin{cases} 1 - \mu_O & \text{if } r^+ < \min(r_i) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

$$m\{\Theta, R\}(\Omega) = \begin{cases} \mu_F & r_i \in R \\ \mu_O & \text{if } r^+ < \min(r_i) \\ 1 & \text{otherwise} \end{cases} \quad (5)$$

where μ_F and μ_O respectively correspond to the probability of false alarm and missed-detection of the sensor. For simplicity reason, these mass functions will be noted $m(\emptyset)$, $m(O)$, $m(F)$ and $m(\Omega)$.

C. Combination of Evidential Grids

The construction of a SG is sequentially done to translate the sensor's data. However, the temporal propagation of the knowledge and uncertainties provided by every point cloud given by the sensor requires a fusion process between the current SG and the result of the previous fusion. The complete description of the environment resulting from such a combination provides a Map Grid (MG). This update allows to detect the consistencies of data as well as some cases of conflict. Fig. 4 illustrates the process of building and updating a MG

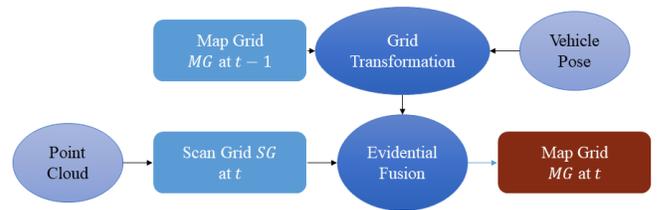


Fig. 4. Map Grid Construction.

using the sensor point cloud provided at a time t . It is the outcome of a combination of a SG built at t according to (2) and a transformed MG built at $t - 1$.

The grid transformation is applied with respect to the new pose of the vehicle at t in order to guarantee that the information is expressed in the current coordinate system of the vehicle. This operation is realized by a spatial transformation for which the aim is to associate to each cell new coordinates. Algorithm 1 describes the approach.

Algorithm 1 Grid Transformation to new vehicle coordinates

Require: Previous Map Grid MG_{t-1} , Rotation matrix R , Translation Vector T .

Ensure: Build a transformed Map Grid $MG_{t-1,tr}$

Initialize $MG_{t-1,tr}$ cells with $m_{MG_{t-1,tr}}(\Omega) = 1$

for each cell with index (p, q) **do**

 Apply a change of coordinates $(p, q) = R \times (p, q) + T$

 Calculate the new indices

$(p_{new}, q_{new}) = \min(|\text{ceil}(p, q)|, |\text{floor}(p, q)|) \times \text{sign}(i, j)$

if $(p_{new} \times q_{new}) > 0$ **then**

$MG_{t-1,tr}(p_{new}, q_{new}) = MG_{t-1}(p, q)$

end if

end for

This update is done according to an evidential multi-grid fusion. This is the crucial point of the grid-based object detection process as it allows the temporal update of the map grid and also the evaluation of the state of cells. Among the various operators in Belief Theory, Dempster's rule of combination is used:

$$m_{MG_t} = m_{MG_{t-1,tr}} \oplus m_{SG_t} \quad (6)$$

where $m_{MG_{t-1,tr}}$ and m_{SG_t} are resp. the mass function of the transformed MG and SG at time t . The operator is defined as:

$$(m_1 \oplus m_2)(A) = K \sum_{\forall B, C \in 2^\Theta, B \cap C = A, A \neq \emptyset} m_1(B).m_2(C) \quad (7)$$

where

$$K^{-1} = 1 - \sum_{\forall B, C \in 2^\Theta, B \cap C = \emptyset} m_1(B).m_2(C) \quad (8)$$

The resulting masses $m_{MG_t}(A)$ define the state of each cell which depends on the previous state and the new measures. The resulting masses according to each state are found as follows [7]:

$$\begin{aligned} m_{MG_t}(O) &= m_{SG_t}(O)m_{MG_{t-1,tr}}(O) + m_{SG_t}(\Omega) \\ &\quad m_{MG_{t-1,tr}}(O) + m_{SG_t}(O)m_{MG_{t-1,tr}}(\Omega) \\ m_{MG_t}(F) &= m_{SG_t}(F)m_{MG_{t-1,tr}}(F) + m_{SG_t}(\Omega) \\ &\quad m_{MG_{t-1,tr}}(F) + m_{SG_t}(\Omega)m_{MG_{t-1,tr}}(F) \\ m_{MG_t}(\Omega) &= m_{SG_t}(\Omega)m_{MG_{t-1,tr}}(\Omega) \\ m_{MG_t}(\emptyset) &= m_{SG_t}(O)m_{MG_{t-1,tr}}(F) + m_{SG_t}(F) \\ &\quad m_{MG_{t-1,tr}}(O) \end{aligned} \quad (9)$$

with $m_{MG_t}(\emptyset)$ being the combined mass expressing the conflict. Basically, this property shows the discordance between the knowledge expressed at $t - 1$ and t . The reason for which a conflict appears is that when a cell changes its state from F to O or vice-versa. Therefore, the detection of this conflict can lead to the evaluation of the dynamic cells. The conflict allows to label the occupied cells which change their state according to two types of conflict:

$$\begin{aligned} C_1 &= m_{SG_t}(O).m_{MG_{t-1,tr}}(F) \quad \text{from } F \text{ to } O \\ C_2 &= m_{SG_t}(F).m_{MG_{t-1,tr}}(O) \quad \text{from } O \text{ to } F \end{aligned} \quad (10)$$

where $m_{MG_t}(\emptyset) = C_1 + C_2$.

Dempster's operator implies a normalization of conflict at fusion considering its absorbing property. Basically, if the conflict is included in the next combination, it induces loss of information because $m_{MG_t}(\emptyset)$ increases at each fusion. Therefore, the updated grid contains no conflict. It is only preserved to classify the mobile cells to be studied for dynamic object extraction.

D. Clustering for Dynamic Object Detection

In order to attain the object level representation from the mobile cells, a clustering is applied to group those cells related to the same object in the grid. For that, the partitioning method must be unsupervised considering that the number of objects to be found is unknown. The well-known Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm is used [15]. It is based on the estimated density of measures for partitioning clusters. This algorithm uses two main parameters: the minimal distance ϵ and the minimum number of points $minPts$ which must reside within a radius ϵ to be included in a cluster. This algorithm is convenient because it is simple and can handle aberrant or noisy values while clustering. However, it can have some issues when clusters have different densities.

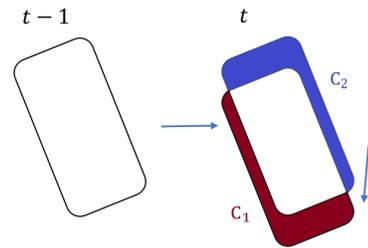


Fig. 5. Appearance of conflict due to the displacement of objects. Conflict C_1 informs about newly occupied cells whereas C_2 describes transitions between occupied to free.

The clustering algorithm is applied to a set of cells which should necessarily be occupied (i.e. non-zero elevation) and for which the conflict $m_{MG_t}(\emptyset)$ is later used to classify the set of the resulting clusters. The partial conflict C_2 informs about cells changing state from occupied to free at a time t . The cells affected with this conflict do not belong to a given object and hence do not provide any knowledge about the

object's presence. That is why, it is trivial to only consider the partial conflict C_1 to determine the location of the object at time t . However, exclusively clustering the C_1 -labeled cells is not informative enough to obtain a complete representation of the shape of the dynamic object. On the grid, the displacement of objects is only visible at the perimeter, the conflict is mostly located at the boundaries of objects as shown in Fig. 5. For that, clustering is applied to detect both classes of static and dynamic objects according to the elevation measure on the 2.5D grid. Afterwards, a classification of these clusters is made according to those which partially contain conflictual cells as mentioned in Fig. 1.

IV. EXPERIMENTAL RESULTS

The presented approach is applied to real data and has been tested offline. The validation is done at the grid-level as well as the object-level according to the ground truth (GT) for qualitative and quantitative evaluation.

A. Dataset and Performance Evaluation

The used data is extracted from the KITTI database [16]. It is a widely used dataset in research for autonomous driving as it provides a large set of images, GPS/IMU recordings and laser scans raw data as well as labeled scenes. In this study, sequence 17 from the raw data set is used considering that annotations with the GT on detected objects as well as the vehicle pose are available. A total of 114 frames are used for which 59 contain annotated moving cars. The used data for this approach are the point cloud recorded by a Velodyne HDL-64 which is characterized by 64 horizontal layers and a 360° horizontal field of view and a 26.9° vertical field of view. The GPS data is also used to obtain the vehicle's pose. The images are not exploited for this application but are used for visualization.

3D object detection benchmarks offer various criteria for performance evaluation purposes. The most common measure is the precision. It is the proportion of all example above the rank which are from the positive class [17]:

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

where TP and FP respectively stand for True Positive and False Positives.

This metric is calculated according to the overlap of candidate detections with the GT. For the computation, the bounding box of a detected object is compared to the GT bounding box. A correct detection has an overlap area a_o which exceeds 50% to be considered True and can be calculated as follows:

$$a_o = \frac{area(B_p \cap B_{gt})}{area(B_p \cup B_{gt})} \quad (12)$$

where B_p and B_{gt} are respectively the candidate bounding box and the GT bounding box.

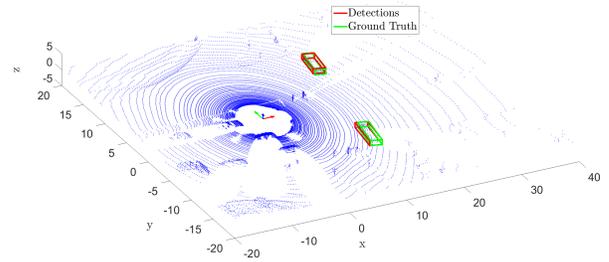
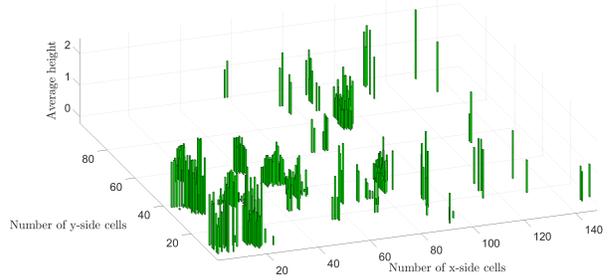


Fig. 6. Top: Frame 40 of Sequence 17, Middle: The corresponding 2.5D grid, Bottom: Comparison of the detection results on frame 40 with the ground truth.

B. Results

In the following section, the perception results will be illustrated according to the 2.5D grid, the evidential occupancy measures and the detected objects. The results are confronted with the available GT-based objects.

Fig. 6 shows an example of results according to the image captured by the camera facing the front of the vehicle. The corresponding 2.5D grid found below the image expresses the average height of elevated objects in the scene. In this view, the car's position is approximately $x = 50, y = 50$ and heading to the right. It can be seen that this map contains voxels describing moving objects (2 cars) as well as static ones like numerous traffic signs or static vehicles behind the ego-car. The grid fusion allows to determine which among these voxels belong to dynamic objects. The bottom figure, exclusively, shows the results of dynamic objects found in the range of the camera view. It provides a comparison between the bounding boxes resulting from the fusion and clustering process and the GT bounding boxes. We choose not to display the objects detected behind the ego-vehicle since no annotations are available in the GT.

The position of these objects is found according to the evaluation of conflict in the corresponding MG. The conflict C_1 allows to observe the cells changing state from Free to

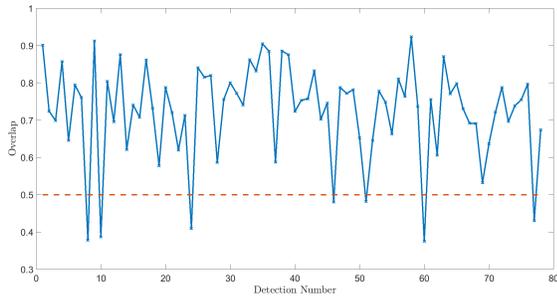


Fig. 7. Overlap of the detected objects with the ground truth showing the rate above which a detection is eligible.

Occupied and the voxels which contain a non-zero value of C_1 are grouped to define the moving objects. Considering that we only detect cars in this sequence, the parameters of DBSCAN are $minPts = 4$ and $\epsilon = 5$ in grid coordinates. This algorithm is advantageous for this application because it discards the measures which can be considered noisy. This allows to optimize the number of relevant clusters. The extracted clusters are labeled according to bounding boxes containing the exact number of LiDAR data belonging to the cluster. The use of 3D bounding boxes allows to evaluate the results according to the known locations of GT. The total AP is found to be 91.23% with an overlap illustrated in Fig. 7. It can be noticed that most detected objects overlap with true objects at a rate varying between 65% – 90%. Note that the number of false alarms is very low due to the property of the clustering algorithm which only considers dense groups of measurement. The noisy or distant data do not belong to any object.

V. CONCLUSION

The approach presented in this paper aims at the detection of multiple objects based on LiDAR data according to an evidential 2.5D grid. The main contributions of this paper is the use of an evidential elevation map and the evaluation of conflict for the determination of mobile objects. Another contribution is the clustering to have an object-level representation for tracking purposes. The detection of dynamic objects is evaluated according to ground truth given by a set of annotations of KITTI dataset. The approach is shown to be efficient according to its high average precision. The perspectives for future work are the identification of the clustering algorithm parameters in order to identify many classes of objects. Extending these first results by testing the approach on more complex scenarios including occluded objects is a second work perspective. Furthermore, a comparative study of this work with the state of the art results will be performed.

ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support from Fondation Wallach (Mulhouse) in the context of the Project SIMPHA (Solution Innovantes pour la Mobilité individualisée et durable des séniors et Personnes présentant un Handicap).

REFERENCES

- [1] A. Asvadi, P. Peixoto, and U. Nunes, “Detection and tracking of moving objects using 2.5d motion grids,” in *18th International Conference on Intelligent Transportation Systems*, Las Palmas, Spain, 2015.
- [2] A. Elfes, “Using occupancy grids for mobile robot perception and navigation,” *Computer*, vol. 22, no. 6, pp. 46–57, Jun. 1989.
- [3] C. Coué, C. Pradalier, C. Laugier, T. Fraichard, and P. Bessiere, “Bayesian Occupancy Filtering for Multitarget Tracking: an Automotive Application,” *International Journal of Robotics Research*, vol. 25, no. 1, pp. 19–30, Jan. 2006.
- [4] A. Broggi, S. Cattani, M. Patander, M. Sabbatelli, and P. Zani, “A full-3d voxel-based dynamic obstacle detection for urban scenario using stereo vision,” in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Oct 2013, pp. 71–76.
- [5] K. Mekhnacha, Y. Mao, D. Raulo, and C. Laugier, “The “fast clustering-tracking” algorithm in the bayesian occupancy filter framework,” in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Aug 2008, pp. 238–245.
- [6] D. Pagac, E. M. Nebot, and H. F. Durrant-Whyte, “An evidential approach to map-building for autonomous vehicles,” *IEEE Transactions on Robotics and Automation*, vol. 14, no. 4, pp. 623–629, 1998.
- [7] J. Moras, V. Berge-Cherfaoui, and P. Bonnifait, “Moving objects detection by conflict analysis in evidential grids,” in *IEEE Intelligent Vehicles Symposium (IV)*, Baden-Daden, Germany, 2011, pp. 1122–1127.
- [8] J. Moras, V. Cherfaoui, and P. Bonnifait, “Credibilist occupancy grids for vehicle perception in dynamic environments,” in *IEEE International Conference on Robotics and Automation*, May 2011, pp. 84–89.
- [9] M. Kurdej, J. Moras, V. Cherfaoui, and P. Bonnifait, “Controlling remanence in evidential grids using geodata for dynamic scene perception,” *International Journal of Approximate Reasoning*, vol. 55, no. 1, Part 3, pp. 355–375, 2014.
- [10] R. Danescu, F. Oniga, and S. Nedeveschi, “Modeling and tracking the driving environment with a particle-based occupancy grid,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1331–1342, Dec 2011.
- [11] G. Tanzmeister and D. Wollherr, “Evidential grid-based tracking and mapping,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 6, pp. 1454–1467, June 2017.
- [12] J. Honer and H. Hettmann., “Motion state classification for automotive lidar based on evidential grid maps and transferable belief model.” in *21st International Conference on Information Fusion (FUSION)*, Oxford, U.K, Jul. 2018.
- [13] L. Wang and Y. Zhang, “Lidar ground filtering algorithm

- for urban areas using scan line based segmentation,” *Computing Research Repository*, vol. abs/1603.00912, 2016.
- [14] K. Jo, S. Cho, C. Kim, P. Resende, B. Bradai, F. Nashashibi, and M. Sunwoo, “Cloud update of tiled evidential occupancy grid maps for the multi-vehicle mapping,” *Sensors*, vol. 18, no. 12, 2018.
- [15] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A density-based algorithm for discovering clusters in large spatial databases with noise.” AAAI Press, 1996, pp. 226–231.
- [16] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, “Vision meets robotics: The kitti dataset,” *International Journal of Robotics Research (IJRR)*, 2013.
- [17] M. Everingham, L. V. Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “The pascal visual object classes (voc) challenge,” *International Journal of Computer Vision*, 2010.