



**HAL**  
open science

## Fusion of evidential occupancy grids for cooperative perception

Federico Camarda, Franck Davoine, Véronique Cherfaoui

► **To cite this version:**

Federico Camarda, Franck Davoine, Véronique Cherfaoui. Fusion of evidential occupancy grids for cooperative perception. 13th Annual International Conference on System of Systems Engineering (SoSE 2018), Jun 2018, Paris, France. pp.284-290, 10.1109/SYSOSE.2018.8428723 . hal-01938394

**HAL Id: hal-01938394**

**<https://hal.science/hal-01938394v1>**

Submitted on 28 Nov 2018

**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.

# Fusion of evidential occupancy grids for cooperative perception

Federico Camarda, Franck Davoine, Véronique Cherfaoui

**Abstract**—In the field of autonomous navigation, perception of the driving scene is one of the essential elements. Existing solutions combine on-board exteroceptive sensors, and are capable of understanding some of the near vehicle’s dynamic surrounding environment. Furthermore, the perception capability of each vehicle can be enhanced by wireless information sharing if vehicles in the neighborhood transmit pertinent information. The primary benefit of such approach is to enable and improve the performance–safety of cooperative autonomous driving. The introduction of such a vehicle-to-vehicle communication leads to think of the global architecture as a system of systems. In this work, we address the task of evidential occupancy grid fusion so that a given vehicle can refine and complete its occupancy grid with the help of grids received from other near vehicles. The communication channel is supposed to be ideal, noiseless with infinite capacity. We focus on the fusion framework itself, using the theory of belief functions for reasoning with uncertainties on the relative poses of the vehicles and on the exchanged sensor measurement data. We evaluate the fusion system with real data acquired on public roads, with two connected vehicles.

**Index Terms**—autonomous driving, cooperative perception, belief functions, lidar, fusion of evidential occupancy grids.

## I. INTRODUCTION

Nowadays, driver assistance systems have entered the market with successes like adaptive cruise control and lane keeping assistance. Despite this, much research and development efforts are still necessary to go toward fully autonomous driving with complex tasks to perform, safely (e.g. overtaking and lane changing on motorways, or crossing intersections in urban areas). The development of these systems relies on functions like multi-sensory perception, communication or adaptive learning. Vehicles so-called smart or intelligent have variable capabilities in terms of self-localization, perception of the driving environment or even prediction of trajectories of other traffic participants. They can take advantage of perceptual or intentional information exchanged with other road users (e.g. vehicles, pedestrians or bicycles) to augment their field of view and situational awareness of the dynamic traffic scene [1][2]. Furthermore, vehicles can exploit information provided by roadside units, by the infrastructure, or high-level context given by digital maps. These different sources when combined can provide sufficient details to build a representation of the driving environment which can be so-called object-based [3][4] or grid-based. In this way, vehicle-to-vehicle and vehicle-to-infrastructure based collaborative perception interacting with various other components for multi-vehicle

localisation and decision making, defines a complex system of systems [5].

The purpose of this work is to improve the perception capabilities of the vehicle, making the best of information received from other vehicles present in the driving scene via wireless communications. It is shown how occupancy grids built on Lidar sensor measurements can be exchanged and fused ultimately achieving a better representation of the urban scene. The feasibility of this work is validated by our implementation and experimental results involving two connected vehicles driving in different driving scenarios.

### A. Occupancy grids

This paradigm of occupancy grids, introduced in [6], is significantly different from other object-based representations. The environment is defined as a two dimensional grid, where the vehicle can be located. The grid is divided in cells. Each cell represents a part of the space and contains occupancy related information that has been deduced by raw data coming from a sensor. A sensor model has to be defined to properly interpret the raw data and synthesize it in a grid with coherency [7].

In the simplest example of binary occupancy grids, in accordance with an agreed standard, each cell can contain value 0 if the space that represents is free and value 1 if it is occupied. More complex models referring to the same meaning can contain the probability for each cell to be free or occupied according to a certain sensor; these are probabilistic occupancy grids. Other grids for other uses can contain different information, such as elevation of the ground.

With occupancy grid, the level of understanding of the scene will be in general lower and further elaborations will be necessary to deduct other characteristics such as object detection and labeling, estimation of speeds and trajectories. On the other hand, the navigable space is directly represented and for the realization of a meaningful grid it is possible to use multi-grid approaches.

### B. Case of study

The main goal of this work is the study and the development of cooperative perception techniques conceived for the urban scene understanding. The idea is to exploit wireless information exchange among intelligent vehicles to expand everyone’s field of view and detect obstacles that would be otherwise hidden to some. More specifically, the objective is to make the vehicle able to:

- 1) Create a representation of its near environment using evidential occupancy grids.
- 2) Maintain and update this representation continuously.
- 3) Enrich this representation with the point of view of other vehicles executing the same application.

On the hypothesis of a given precise ego-position estimation, occupancy grids built on information coming from a Lidar sensor and wireless communication will be used as representation of the vehicle surrounding environment. Moreover, in order to facilitate information fusion and discounting, the specificity of the belief functions theory has been exploited.

An architecture based on cooperative perception has been proposed, implemented, and studied and it is presented in the next sections.

## II. BELIEF FUNCTIONS

Hereafter is presented the framework of belief functions theory [8][9][10][11], and how we plan to use it to generate evidential occupancy grids.

### A. Uncertain information representation

Let  $\Omega$  be the set of all the possible states of the system under analysis. It can be referred as universe or frame of discernment. The power set  $2^\Omega$  is the set of all subsets of  $\Omega$  (including the empty set) and it contains all and only the states in which a given proposition linked to the system can be true. To represent this uncertain information, mass functions are used assigning a belief mass to each element of the power set.

Formally, mass function, also called basic belief assignment (BBA) or basic probability assignment, over a frame of discernment  $\Omega$  is a function  $m: 2^\Omega \rightarrow [0, 1]$  verifying:

$$m(\emptyset) = 0, \quad \sum_{A \subseteq \Omega, A \neq \emptyset} m(A) = 1 \quad (1)$$

As example, in the specific case of occupancy grid, the universe of the all possible states simply splits in two hypothesis: the considered segment of space can be either free  $F$  or occupied  $O$ , therefore the frame of discernment is defined as  $\Omega = \{F, O\}$ . We can give an illustrative belief distribution represented by:

$$m = \begin{bmatrix} \emptyset & O & F & \Omega \\ m(\emptyset) & m(O) & m(F) & m(\Omega) \end{bmatrix} \quad (2)$$

Each mass quantity under each hypothesis shows the belief committed to that hypothesis.

A mass function  $m$  for which  $m(\emptyset) = 0$  is called normal or regular. Otherwise, the mass function is said to be unnormalized. Then, the quantity  $m(\emptyset)$  can be interpreted as the degree of support of the hypothesis that the true class  $\omega$  is actually outside of our frame  $\Omega$ .

The information encoded by a mass function can be represented in other ways. The notions of belief, plausibility and pignistic probability are other major constructs of the belief functions theory and they play important roles in many aspects of evidential reasoning.

### B. Discounting

In the theory of belief functions, knowledge about the reliability of a source of information can be handled by a discounting factor. A discounting factor is used to weaken a mass function by transferring some masses to the ignorance state  $m(\Omega)$ .

For a factor  $\alpha \in [0, 1]$ , the discounted mass function  ${}^\alpha m$  is defined as:

$${}^\alpha m(A) = (1 - \alpha)m(A), \quad \forall A \subsetneq \Omega \quad (3)$$

$${}^\alpha m(\Omega) = (1 - \alpha)m(\Omega) + \alpha \quad (4)$$

If  $\alpha = 0$ , the information is considered reliable and is kept as it is. On the other hand, if  $\alpha = 1$ , the information is totally unreliable and leads to the vacuous mass function,  $m(\Omega) = 1$ . Smets [10] showed that this discounting can be derived by interpreting  $1 - \alpha$  as the degree of belief that the information is reliable.

### C. Combination rules

The combination rules enables to fuse information from different sources. The sources of information should be defined in the same frame of discernment to use the following rules.

1) *Conjunctive rule*: Given two mass functions  $m_1$  and  $m_2$  induced by two independent sources of information and defined on the same frame of discernment  $\Omega$ , it is possible to combine them using conjunctive rule defined as:

$$\forall A \in 2^\Omega, (m_1 \odot m_2)(A) = \sum_{B \cap C = A | B, C \subseteq \Omega} m_1(B)m_2(C) \quad (5)$$

2) *Dempster's rule*: The Dempster's rule normalizes the mass function distributing the conflict mass into other propositions. The Dempster's rule of combination, or orthogonal sum, computes a new mass function  $m_1 \oplus m_2$  defined as:

$$(m_1 \oplus m_2)(\emptyset) = 0 \quad (6)$$

$$(m_1 \oplus m_2)(A) = \frac{1}{1 - K} \sum_{B \cap C = A} m_1(B)m_2(C), \quad A \neq \emptyset \quad (7)$$

where:

$$K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C) \quad (8)$$

This value measures the conflict between the two mass functions. The combination rule is valid if and only if  $K < 1$ , otherwise,  $m_1$  and  $m_2$  are incompatible and cannot be combined. Dempster's rule is commutative, associative and has the vacuous mass function as unique neutral element.

#### D. Evidential occupancy grids

The framework of belief functions natively presents tools of information fusion, ignorance representation, and information weakening. We have to present the complete perception model yet but it is already evident that these capabilities are perfectly suitable for our purpose: we want to fuse grids of others with ours, to represent ignorance when an area still has not been observed and to reduce the value given to an evidence when it becomes obsolete.

For these reason our grids will be evidential occupancy grids, that is to say that each cell will contain our belief of it to be *Free*, *Occupied* or a measurement of our lack of knowledge.

Each cell  $(i, k)$  will not contain a single probability like in probabilistic occupancy grids but a mass function  $m_{i,k}$  defined over  $\Omega = \{F, O\}$ :

$$m_{i,k} = \begin{bmatrix} \emptyset & O & F & \Omega \\ m_{i,k}(\emptyset) & m_{i,k}(O) & m_{i,k}(F) & m_{i,k}(\Omega) \end{bmatrix}$$

For example, here is a cell certainly occupied:

$$m_{i,k} = \begin{bmatrix} \emptyset & O & F & \Omega \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

A cell certainly free:

$$m_{i,k} = \begin{bmatrix} \emptyset & O & F & \Omega \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

A cell we have no information of:

$$m_{i,k} = \begin{bmatrix} \emptyset & O & F & \Omega \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

A cell of which a certain sensor told us that it is *Free* with a partial belief:

$$m_{i,k} = \begin{bmatrix} \emptyset & O & F & \Omega \\ 0 & 0 & 0.7 & 0.3 \end{bmatrix}$$

### III. GENERATION AND FUSION OF EVIDENTIAL OCCUPANCY GRIDS

This section presents the perception framework and describes the role played by the different occupancy grids in the global perception system. As we can see in figure 1, at each time  $t$  a new *sensor acquisition* is activated on the *EgoVehicle* to create a new *Scangrid* which updates the cumulative *Mapgrid* and is shared with others to the network. In addition, the *EgoVehicle*'s *Mapgrid* can be enhanced with the help of *Sockgrids* transmitted by near *OtherVehicles*, if any.

In the schematic representation, and in this work in general, all the terms with the prefix *Ego* are used to refer to the "first person" vehicle whose *Mapgrid* results centered on. Equivalently, the prefix *Other* refers to one (or more) other vehicle in the same urban scene which is sharing its perception data. *EgoPose* and *OtherPoses* therefore describe each vehicle's position and orientation, they are defined as  $[x, y, \theta]$  where  $\theta$  is the heading angle of the vehicle.

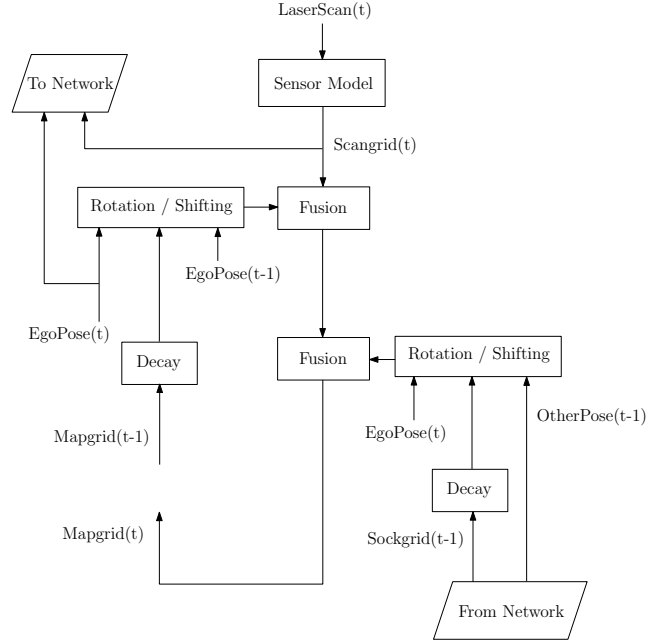


Fig. 1: Illustration of the occupancy grid fusion framework.

#### A. Scangrid

The sensor used in this work is a multi-layer, multi-echo laser scanner SICK LD-MRS400001S01 that is geared towards rough outdoor environments. It provides a point cloud (collection of measurements performed by the Lidar and presented as points in space) with a scanning frequency of 12.5 Hz. We describe hereafter the procedure to properly interpret Lidar scans into a well constructed occupancy grid named *Scangrid*. Its construction is composed of two steps.

1) *From point cloud to polar grid*: The intrinsic angular nature of the device brings us to make use of a polar representation of the measured information. We aim first to build an ego-centered polar evidential occupancy grid. Its abscissa  $\theta$  goes from the minimum to the maximum angle covering the whole aperture. Its ordinate  $d$  goes from 0 to the maximum range of the maximum possible measurable distance of the device. These two dimensions are clearly discretized both according to the Lidar precision, respectively to the angular resolution and the precision on the distance.

Let's now analyze how the device presents its scans and how can we take advantage of it. At a data level, each cloud is a set of point and each point  $p$  is characterized by:

- $(x, y, z)$  or equivalently a measured distance  $d$  from the sensor to the obstacle.
- Angle  $\theta$  of the mirror at the moment of the measurement. The abscissa of the polar grid.
- Layer  $l \in \{0, 1, 2, 3\}$ .
- Echo  $e \in \{0, 1, 2\}$ .
- Flags. There can be different but we are only interested in discriminating points belonging to the ground or not. Therefore,  $f \in \{\text{ground}, \overline{\text{ground}}\}$ , where *ground* refers to each point of impact that did not occur on the ground ( $\overline{\text{ground}} \equiv \text{Not-Ground}$ ).

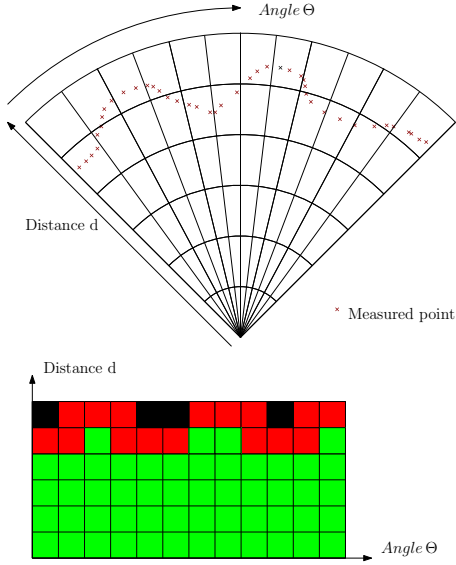


Fig. 2: From points to polar grid. Cells containing at least one point are labeled as *Occupied* (= red). The space between the observer and an *Occupied* cell is labeled as *Free* (= green). Remaining cells are unexplored, therefore *Unknown* (= black).

If for a given pair  $(\theta, l)$  no corresponding point exist in the point cloud then nothing has been detected at that location, at that range.

With 4-layer scans we may have a 3D representation but as we want to lean on 2D grids we perform a projection of the measurements and we suppose that obstacles are vertical and completely in contact with the ground. According to all of this, we can observe for each angle  $\theta$  that:

- Ground points tell us up to where the space is free.
- “No points” tell us that the space is free up to the maximum range.
- A single Not-Ground point tells us that the space is free up to the measured distance and immediately afterwards there is an obstacle occupying the space.
- More than one Not-Ground point tells us that the space is free up to the minimum of the measured distances and immediately afterwards there is an obstacle occupying the space. Also for each point of this kind we can confirm the detection of an obstacle but we are not able to conclude on the navigability of the space in between them.

The building procedure of the evidential polar grid from a laser scan is based on these observations and on the algorithm proposed by [12].

2) *From polar grid to Cartesian grid*: With a polar model as the one we just presented we will not be able to immediately create and fuse evidential occupancy grids. For these reasons, we need to compute a transformation from polar to Cartesian coordinate system. In the Cartesian representation we will assume that the position of the observer is at the bottom of the grid. Several transformation methods exist; we have to choose the one with the least loss. We chose the bi-linear interpolation algorithm, as proposed in [12]. An example of transformation is shown in figure 3.

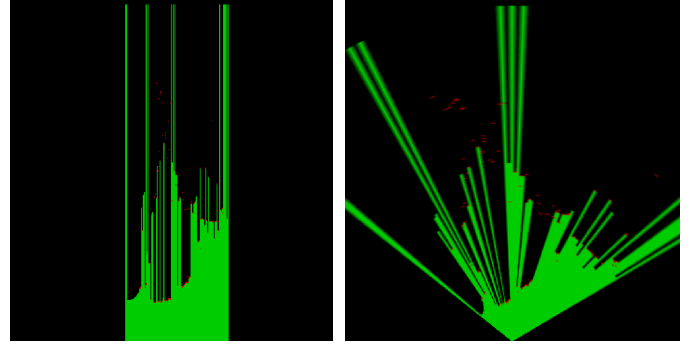


Fig. 3: On the left, a polar grid. On the right, its equivalent Cartesian grid.

### B. Mapgrid

Whilst the *Scangrid* is defined as the base element of the perception in this work, the *Mapgrid* is a cumulative evidential occupancy grid centered on the ego-vehicle that keeps information coming from subsequent laser scan. When possible, information comes from other vehicles and it is identified as *Sockgrid*, a *Scangrid* from a third party point of view.

The name *Sockgrid* comes from the implementation where that particular grid is produced over data coming, in fact, from a socket - and indeed from the network. Therefore, fusion is performed in two different occasions according to the origin of the upcoming new information.

1) *Fusion with Scangrid*: This operation deals with the maintenance of an updated representation of the environment. In the case of a memoryless system, we would simply discard the old *Mapgrid* and replace it with the last *Scangrid*. We want instead a proper map - *Mapgrid* - that keeps memory of past information while being constantly updated with new data, along with the vehicle movement. We present hereinafter the algorithm for the update of  $Mapgrid(t-1)$  with  $Scangrid(t)$ , resulting in  $Mapgrid(t)$ :

- 1) We apply a discounting to  $Mapgrid(t-1)$  according to the elapsed time
- 2) We compute the transformation between  $EgoPose(t)$  and  $EgoPose(t-1)$
- 3) We apply this transformation (rotation and translation) to decayed  $Mapgrid(t-1)$
- 4) We compute the cell-by-cell fusion using the Dempster’s rule

Figure 4 shows the evolution of a *Mapgrid* of a single vehicle which is steady and is detecting an object in movement. The movement is given by cell where  $m_1 \odot m_2(\emptyset)$  is high (in blue in figure 4). On the left,  $Mapgrid(t-1)$  with the object in its initial position. On the right, the result of the fusion of  $Mapgrid(t-1)$  and  $Scangrid(t)$  which is  $Mapgrid(t)$ .

2) *Fusion with Sockgrid*: Handling a *Sockgrid* (*Scangrid* from another vehicle) is similar but presents some differences. Now the idea is not to transform our *Mapgrid* but to properly prepare the *Sockgrid* to be fused. Hypothesizing that, with the *Sockgrid*, it has been communicated the corresponding pose

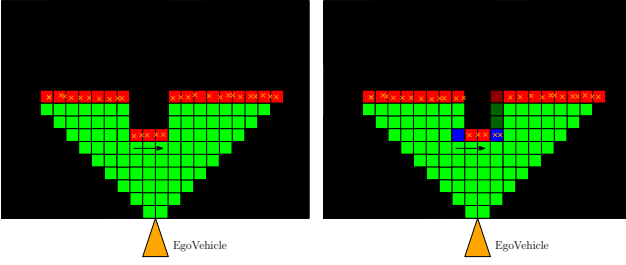


Fig. 4: An example of *Mapgrid* progression with observation of an object in movement from a stationary observer. Red cells represent, in this example, another vehicle moving along a wall and orthogonally to the *EgoVehicle*.

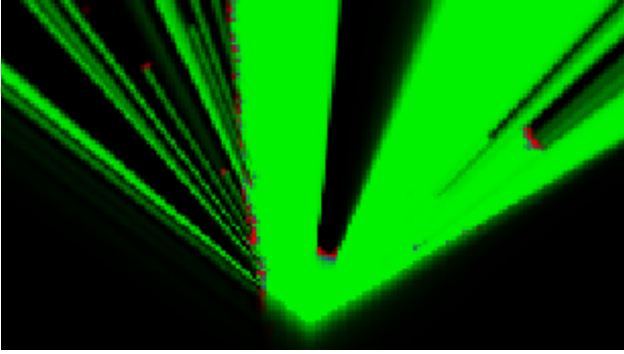


Fig. 5: Effective *Mapgrid* presenting objects in movement, therefore conflict mass presented as blue cells.

and timestamp, the algorithm proposed to enrich the grid of the ego-vehicle with that of another is the following:

- 1) We apply a discounting to  $Sockgrid(t-1)$  according to its timestamp and confidence given
- 2) We compute the transformation between  $EgoPose(t)$  and  $OtherPose(t-1)$
- 3) We apply this transformation (rotation and translation) to decayed  $Sockgrid(t-1)$
- 4) We compute the cell-by-cell fusion using the Dempster's rule

#### IV. EXPERIMENTS

The designed system has been implemented in C++ under the framework ROS (Robot Operating System [13]). In order to test its consistency without the actual effects of a wireless communication between vehicles, real on-the-road data has been acquired and used as an offline test bench. This was possible because all the streams of data have been acquired with synchronized clocks at the microsecond level taking advantage of the GNSS receiver connected to each calculator. The acquisitions involved two vehicles configured as follows:

- *EgoVehicle* (Renault Zoe Grey): back position, Lidar LD-MRS on the front, dashcam, Novatel Span estimator, RTK-GPS.
- *OtherVehicle* (Renault Zoe White): front position, Lidar LD-MRS on the front, dashcam, Novatel Span estimator.

A third unequipped vehicle also took part to the acquisition in order to have control and predictable movements of another

entity of the scene. This fleet has been driven through public roads, to build a dataset complete of several scenarios and urban situations.

#### A. Performance evaluation

Without an actual ground truth at disposal, it is difficult to define a quantitative measure of the quality of the fusion. The performance depends indeed directly on the estimation of the relative pose between *EgoPose* and *OtherPose*. As these two poses come from two different independent estimators, two different uncertainties are involved. The *EgoPose* is measured at a RTK-level of precision, its uncertainty magnitude can not be compared to the uncertainty on *OtherPose*. Therefore, this last indicator is used to measure the error coming with every fusion and it is computed taking advantage of *EgoVehicle*'s Lidar and position, as follows:

- Before sharing the *Sockgrid*, *OtherVehicle* adds to the grid a mask representing itself with a shape of "red rectangle". This geometric form is not just a display of the vehicle; those cells result red because in the *OtherVehicle*'s *Scangrid* that space has been labeled as *Occupied* by properly assigning mass. Similarly, it is reasonable to suppose that the space around the car is free, therefore it is labeled as *Free*. This area of the grid, where knowledge has been added according to our observations and not according to the sensor measurement, is identified as a set of cells named *OV* (short for *OtherVehicle*) (an example is shown at the bottom right of figures 7 and 8);
- After reception, *EgoVehicle* performs the fusion between its *Mapgrid* and *Sockgrid*. Given that during the dataset acquisition *EgoVehicle* always kept its back position w.r.t. the other car, it is likely that *OtherVehicle* is detected by its Lidar and it already appears (totally or in part) in its *Mapgrid*;
- Completing the fusion with Dempster's Rule, the sum of the amount of conflict reported in each cell belonging to the set *OV* indicates the discrepancy between *OtherPose*'s estimation and its actual measurement via *EgoVehicle*'s laser scan. This *ConflictError* is therefore an indicator of the incompatibility between the two evidential occupancy grids fused and it is computed at each time instant as:

$$ConflictError = \sum_{cell_{i,j} \in OV} m_{i,j}(\emptyset)$$

The thus defined quantity varies accordingly to the amount of conflict reported in the information fusion nearby the *OtherVehicle*. It can not be considered a formal and properly modeled measurement of the error, but it can discriminate coherent fusion (low conflict) and poor fusion (high conflict).

#### B. Results

Updated *Mapgrid*, *Scangrid*, and *Sockgrid* are shown on figures 7 and 8. Each figure displays important results of the study in a three columns format used during this work to evaluate the outcome of the designed system. In these three columns are presented, from the left:

- *EgoVehicle*'s *Mapgrid*, resulting from subsequent fusions;
- *EgoVehicle*'s camera and *Scangrid*;
- *OtherVehicle*'s camera and *Scangrid* (or *Sockgrid*).

This representation matches with its *ConflictError* and easily allows to identify instants of coherent fusion.

The resulting *ConflictError* evolution during a four-minute data acquisition is displayed in figure 6. For a better assessment, please follow the trend of the error along with its corresponding updated grid, which is displayed in video format and available on Youtube<sup>1</sup>.

Two remarkable instants of the execution are hereinafter described.

In figure 7 (timestamp 210s), both vehicles are detecting, from two different points of view, the same slope along the street presented in the top left of each grid as an oblique red line. Nevertheless, in the final grid (first column from the left, so-called *EgoVehicle*'s *Mapgrid*), that unique slope appears in the representation of the environment as two different red lines, therefore as two different obstacles. Qualitatively, we deduce with a graphical evaluation of the result that the two grids who took part in the fusion were not properly aligned. This incompatibility is caused by a poor estimation of the relative pose, which is crucial element of the algorithm described in section III-B2. Quantitatively, we state the same outcome by reporting an accordingly high value of the indicator *ConflictError*.

In figure 8 (timestamp 170s), instead, the estimated poses are consistent with the performed Lidar measurements and the reported *ConflictError* results comparatively smaller. The edge of the road, the horizontal barrier in the top part of the each grid, is in fact well aligned in the final grid, where the fusion process did not lead to any double representation of this obstacle.

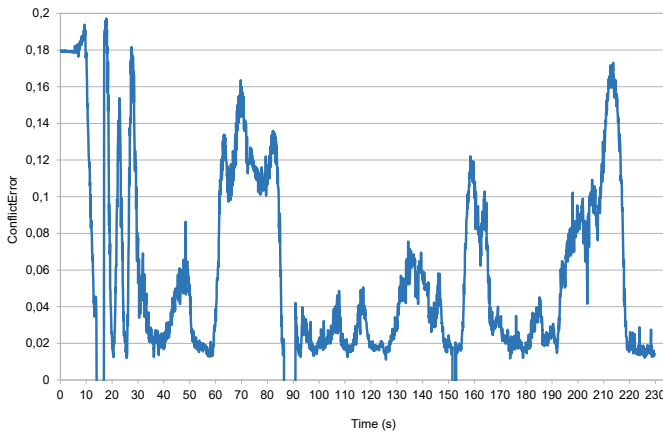


Fig. 6: *ConflictError* is not computed when the *OtherVehicle* is out of *EgoVehicle*'s field of view (14s – 16s). Its value results high when the two fused grid are poorly aligned (60s – 80s) and low when their combination is coherent (170s – 190s).

## V. CONCLUSIONS AND FUTURE WORKS

The occupancy grid fusion framework we propose in this paper has been successfully implemented in a prototype version. There is still room for optimization in the algorithms

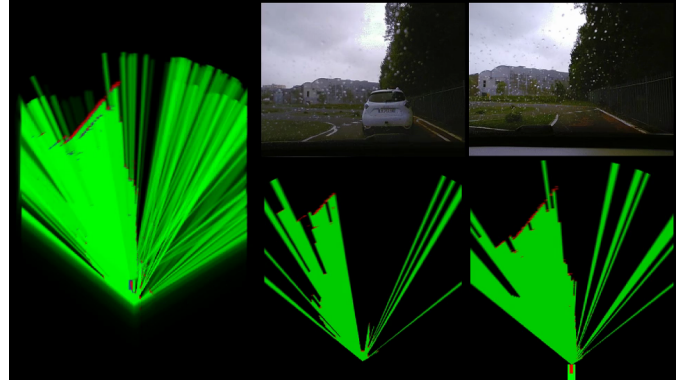


Fig. 7: Poor position estimation leads to fusion of unaligned grids and to double detection of an obstacle (top left of the grid, timestamp 210s).

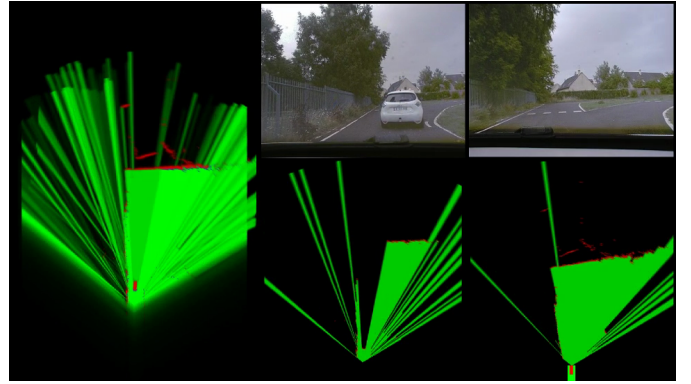


Fig. 8: Consistent fusion in the presence of a barrier detected by both vehicles (top of the grid, timestamp 170s).

implementation but the current application satisfies the three requirements listed in section I. The results of its execution on the dataset acquired with multiple vehicles are promising and we make different observations.

The current implementation is computationally too heavy to run online. In order to allow an adequate rate of *Mapgrid* updating it is necessary to replay the acquired dataset at 0.5 times its original speed. With this performance, it is obviously not possible to design on-road tests with live application running on more than two vehicles.

We also conclude that the quality of the resulting grids strongly depends on the position estimation accuracy. The fusion of two grids is preceded by a transformation based on the relative pose, therefore on two independent position estimation, one per vehicle. If these two are poorly estimated and there is a significant discrepancy from the truth, the fusion results accordingly poorly aligned and conflictual.

However, at a sufficiently slow rate of execution and with sufficiently precise positions, the results are consistent with the surrounding environment. The indicator *ConflictError* has proven to be effective in identifying instants of coherent fusion when our application can factually enhance the overall system of system understanding of the urban scene.

In the future, it will be interesting to:

- Take better advantage of the *ConflictError* value. A

<sup>1</sup><https://youtu.be/dndqe1F-kfs>

- possible use is as a threshold in a go/no-go condition on the feasibility of the fusion;
- Implement a method of grid alignment, such as the one proposed in [14] and [15], to face the uncertainty on the pose estimations;
  - Enlarge the frame of discernment replacing the class  $\{O\}$  with  $\{S, D\}$  to consider both static and dynamic cells, and implement a solution able to detect moving cells such as the one proposed in [16].

#### ACKNOWLEDGMENT

This work was carried out and funded in the frameworks of the Equipex ROBOTEX (Reference ANR-10-EQPX-44-01), the Labex MS2T (Reference ANR-11-IDEX-0004-02) and the challenge DAPAD. It was supported by the French Government, through the program “Investments for the future” managed by the National Agency for Research.

#### REFERENCES

- [1] S.-W. Kim, Z. J. Chong, B. Qin, X. Shen, Z. Cheng, W. Liu, and M. H. Ang, “Cooperative perception for autonomous vehicle control on the road: Motivation and experimental results,” in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, 2013, pp. 5059–5066.
- [2] S.-W. Kim, W. Liu, M. H. A. Jr., E. Frazzoli, and D. Rus, “The impact of cooperative perception on decision making and planning of autonomous vehicles,” *IEEE Intelligent Transportation Systems Magazine*, pp. 39–50, Fall 2015.
- [3] N. El Zoghby, V. Cherfaoui, and T. Denoeux, “Evidential distributed dynamic map for cooperative perception in vanets,” in *Intelligent Vehicles Symposium Proceedings, 2014 IEEE*. IEEE, 2014, pp. 1421–1426.
- [4] M. Vasic, D. Mansolino, and A. Martinoli, “A system implementation and evaluation of a cooperative fusion and tracking algorithm based on a gaussian mixture phd filter,” in *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*. Ieee, 2016, pp. 4172–4179.
- [5] J. Straub, W. Amer, C. Ames, K. R. Dayananda, A. Jones, G. Miryala, N. Olson, N. Rockenback, F. Slaby, S. Tipparach, S. Fehring, D. Jedyak, H. Lou, D. Martin, M. Olberding, A. Oltmanns, B. Goenner, J. Lee, and D. Shipman, “An internetworked self-driving car system-of-systems,” in *2017 12th System of Systems Engineering Conference (SoSE)*, June 2017, pp. 1–6.
- [6] A. Elfes, “Using occupancy grids for mobile robot perception and navigation,” *Computer*, vol. 22, no. 6, pp. 46–57, 1989.
- [7] C. Yu, V. Cherfaoui, and P. Bonnifait, “An evidential sensor model for velodyne scan grids,” in *Control Automation Robotics & Vision (ICARCV), 2014 13th International Conference on*. IEEE, 2014, pp. 583–588.
- [8] A. P. Dempster, “Upper and lower probabilities induced by a multivalued mapping,” *The annals of mathematical statistics*, pp. 325–339, 1967.
- [9] G. Shafer, *A mathematical theory of evidence*. Princeton university press, 1976, vol. 42.
- [10] P. Smets, “Decision making in the tbm: the necessity of the pignistic transformation,” *International Journal of Approximate Reasoning*, vol. 38, no. 2, pp. 133–147, 2005.
- [11] P. Smets and R. Kennes, “The transferable belief model,” *Artificial intelligence*, vol. 66, no. 2, pp. 191–234, 1994.
- [12] J. Moras, V. Cherfaoui, and P. Bonnifait, “Moving objects detection by conflict analysis in evidential grids,” in *Intelligent Vehicles Symposium (IV), 2011 IEEE*. IEEE, 2011, pp. 1122–1127.
- [13] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, “Ros: an open-source robot operating system,” in *ICRA workshop on open source software*, vol. 3, no. 3.2. Kobe, Japan, 2009, p. 5.
- [14] H. Li and F. Nashashibi, “A new method for occupancy grid maps merging: Application to multi-vehicle cooperative local mapping and moving object detection in outdoor environment,” in *Control Automation Robotics & Vision (ICARCV), 2012 12th International Conference on*. IEEE, 2012, pp. 632–637.
- [15] G. Trehard, E. Pollard, B. Bradai, and F. Nashashibi, “On line mapping and global positioning for autonomous driving in urban environment based on evidential slam,” in *Intelligent Vehicles Symposium (IV), 2015 IEEE*. IEEE, 2015, pp. 814–819.
- [16] G. Tanzmeister and D. Wollherr, “Evidential grid-based tracking and mapping,” *IEEE Transactions on Intelligent Transportation Systems*, no. 99, pp. 1–14, 2016.