Controlling Remanence in Evidential Grids Using Geodata for Dynamic Scene Perception

Marek Kurdej, Julien Moras, Véronique Cherfaoui, Philippe Bonnifait

Controlling Remanence in Evidential Grids Using Geodata for Dynamic Scene Perception[☆]

Marek Kurdej[a,**], Julien Moras[a,*], Véronique Cherfaoui[a,*], Philippe Bonnifait[a,*]

[a] *Université de Technologie de Compiègne – CNRS, Heudiasyc UMR 7253, BP 20529, 60205 COMPIÈGNE CEDEX, France*

Abstract

This article proposes a perception scheme in the field of intelligent vehicles. The method exploits prior map knowledge and makes use of evidential grids constructed from the sensor data. Evidential grids are based on occupancy grids and the formalism of the Dempster–Shafer theory. Prior knowledge is obtained from a geographic map which is considered as an additional source of information and combined with a grid representing sensor data. Since the vehicle environment is dynamic, stationary and mobile objects have to be distinguished. In order to achieve this objective, evidential conflict information is used for mobile cell detection. As well, an accumulator is introduced and used as a factor for mass function specialisation in order to detect static cells. Different pieces of information become obsolete at different rates. To take this property into account, contextual discounting is employed to control cell remanence. Experiments carried out real-world data recorded in urban conditions illustrate the benefits of the presented approach.

Keywords: dynamic perception, belief functions theory, spatio-temporal fusion, contextual fusion, data remanence, intelligent vehicles

[☆] This article is an elaborated version of [1], which was presented in the 2nd International Conference on Belief Functions, Compiègne, France in 2012.

[*] Corresponding author

[**] Principal corresponding author

Email addresses: `marek.kurdej@hds.utc.fr` (Marek Kurdej), `julien.moras@hds.utc.fr` (Julien Moras), `veronique.cherfaoui@hds.utc.fr` (Véronique Cherfaoui), `philippe.bonnifait@hds.utc.fr` (Philippe Bonnifait)

URL: `http://www.hds.utc.fr/~kurdejma` (Marek Kurdej), `http://www.hds.utc.fr/~morasjul` (Julien Moras), `http://www.hds.utc.fr/~vberge` (Véronique Cherfaoui), `http://www.hds.utc.fr/~bonnif` (Philippe Bonnifait)

1. Introduction

The research field of this paper is the domain of intelligent vehicles, which have been a hot topic for at least a few decades. The need to perceive and understand the surrounding environment is crucial for multiple tasks, such as navigation, planning and other Advanced Driver Assistance Systems (ADAS), which aim at improving safety in road and urban environments. The incontestable appeal of possible applications collides with numerous theoretical and practical problems. A difficulty resulting from robot perception is that the sensors which are at one's disposal have partial view of the scene and provide imperfect data. Multi-sensor data fusion is one way to improve the field of view of a perception system and to improve the quality of data. Moreover, the availability of more and more detailed and precise geographic databases gives an opportunity to incorporate prior map knowledge into a perception system. Geodata can be fused with sensor data in order to infer more refined information about the vehicle environment. The semantic information from this data can be used to perceive and to analyse the dynamics of different objects in the scene. The contribution of this paper is to propose a new perception scheme based on the fusion of data coming from embedded sensors and a digital map and on managing the remanence of scene objects in an evidential grid representation. The formalism of Dempster–Shafer theory is used to manage uncertainties.

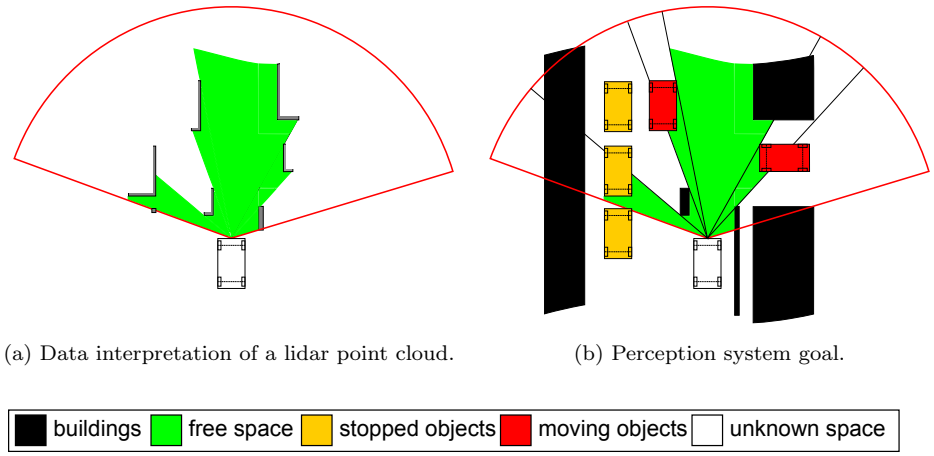
This article is organised as follows. Section 2 presents the background and the general idea of the proposed method. Tools and basic notions, such as evidential grids, are explained in Section 3. Further, details on the information fusion are given in Section 4. Different aspects are presented: from the spatial fusion of grids and combination of prior knowledge through the temporal fusion, accumulation method and specialisation until contextual discounting. Section 5 serves as a detailed description of the experimental setup and of the data used in tests. It also includes a description of the sensor model, which is independent of the presented method. Section 6 presents the results which show how the proposed method addresses the occultation problem and various dynamics of scene. Finally, Section 7 concludes the paper and presents ideas for future work.

2. Problem statement

An efficient perception scheme for intelligent vehicles has to address many issues in order to provide desired outcomes. This is possible using multi-sensor data fusion and/or temporal fusion. Temporal fusion exploits the fact that the vehicle moves and so the data from one sensor can be combined at different moments, therefore at different positions. In this way, the field of view can be enlarged, so that the limitations of the sensors become less stringent. As seen in Figure 1a, raw sensor information is semantically poor and needs more elaborate processing to achieve useful results. Figure 1b shows how a perception system, using available sensors, is able to distinguish various types of objects. Similarly, the use of multiple data sources and fusion algorithms improves the accuracy

and the integrity of data. Having accurate data is of course important, but it is essential to maintain data integrity. This means that such data are complete and consistent, and uncertainties are quantified.

Another important question is the notion of scene dynamics: static environments are much less demanding than the dynamic ones, and a robust method has to deal with this problem explicitly. The complexity of a typical scene is in general high, so that it is necessary to simplify the perception system and to add more semantic information.



(a) Data interpretation of a lidar point cloud.

(b) Perception system goal.

Figure 1: Bird's eye view of a scene example.

2.1. Method description

In this article, we propose a method which tackles these issues. The main idea is to accumulate temporal information and to incorporate prior knowledge from maps in a grid which covers a part of the environment. The intention is similar to the Simultaneous Localisation and Mapping, Moving Object Tracking (SLAMMOT) problem [2, 3], where moving objects are detected, with the difference that no tracking is performed. One can remark that moving versus static object detection is not provided by any optical sensor, but is the result of the fusion process. Indeed, our perception system does not include any sensor like a radar, which uses Doppler effect.

The difficulty of dealing with the dynamics of the scene arises from two main reasons. The system carrier (robot, vehicle) is a moving actor in interaction with other objects in the scene. These objects themselves can be mobile: momentarily stopped or on the move. Temporal fusion and data accumulation serve a double purpose. On the one side, they allow to filter the sensor noise and, on the other hand, to conserve some pieces of information. Preserved information can, for instance, concern the zones of vehicle environment that are not subject

to occultations. An important assumption is made: the scene dynamics is limited. It means that one can fix a forgetting factor which bounds the process of information conservation. This factor is closely attached to and acts on the data remanence: adapting this parameter changes the persistence of a given stimulus after its disappearance. The term stimulus corresponds to sensor data, and the persistence represents the time during which these data are present in the perception grid.

2.2. Geodata exploitation

The principal novelty of the method is the use of meta-knowledge obtained from a digital map. This approach is at odds with the SLAMMOT methods, because it uses maps prepared beforehand. More and more maps are accessible nowadays. They are even more precise, accurate and complete than ever. The proposed method tries to enhance the dynamic perception of the environment by considering maps as an additional source of information on a par with other sources, e.g. sensors. So far, the prior information obtained from maps can be used to predict the trajectories of dynamic objects [4]. In contrast, this article focuses on maps as prior knowledge serving the purposes of perception. The proposed idea is to control different perception dynamics in the same scene, thanks to the map. For example, perceived buildings should not be forgotten quickly. On the contrary, mobile objects with a short "lifetime" in the scene should be updated rapidly, so discarded (almost) as soon as they disappear. In this context, "lifetime" means the time that an object spends in the environment represented by a single cell of perception grid. As a consequence, this method allows to manage the occulted zones and objects. It could as well be used in a closed-loop localisation system, perceived landmarks would be associated to the reference map data. A similar approach is presented by Hentschel et al., where a GPS receiver is coupled with a laser-based sensor and a 2D reference map containing static line features [5]. Some other researchers have already used maps and geodata for various purposes. Several works have successfully exploited such data for mobile navigation [6, 7]. Cappelle, Dawood et al. defined a 3D city model as a source of prior knowledge for vision-based perception and navigation [6, 8].

2.3. Use of the Dempster–Shafer theory

The Dempster–Shafer theory (DST) proposed by Dempster [9] and developed, among others, by Shafer and Smets [10, 11, 12] gained its popularity thanks to various interesting properties. The most important reasons that convinced the authors to use the DST are as follows. The vehicle environment contains many occultations and barely observed or non-observed zones. The representation of the unknown, inherent to the DST and missing in the theory of probability permits to handle this notion in a way. Fusion operators defined in the DST are able to manage uncertainties and conflict between information sources. Recent work of Klein et al. has shown by introducing a particular conflict criterion that a proper conflict analysis may be helpful to identify singular,

or outlying, information sources [13]. There has been also substantial research work on data association problems, such as multi-target tracking, which exploits conflict management [14].

In a perception system, it is desirable to have a tool to manage different levels of detail, since the obtained information cannot be always interpreted clearly and precisely. In the DST, this tool is at the core of the theory. A frame of discernment can be as refined as the most detailed data obtained from the sensors, but still, it remains possible and easy to combine information which is more general by affecting masses on non-singletons. There are also well-established methods to deal with multiple frames of discernment [15]. Such management of levels of detail would be impossible or at least difficult with an accumulation schema. There are already some works which take advantage of the theory of evidence in the context of mobile perception [16, 17]. In other domains, the Dempster–Shafer theory has been used as well, e.g., for visual tracking. Klein et al. presented a hierarchical combination scheme that makes use of existing fusion rules and source classification with respect to their reliability and precision [18].

3. Dynamic perception using evidential grids

In the proposed approach evidential grids are used to represent 2D environment. The grids used by the method are called "perception grids". In general, these are occupancy grids like those proposed by Elfes [19] for robot navigation. Initially, they were used merely for indoors robotics and could differentiate between free and occupied space only. The original idea used purely probabilistic scheme, whereas Pagac et al. [20] proposed evidential grids based on DST [9]. The basic concept motivating this article is to develop this approach. Refining the frame of discernment enabling the storage of additional information and adding supplementary data makes this solution a promising approach towards scene understanding. Among many advantages of the grids, the most important are presented below.

Object-based approaches, which consist in detecting, recognising and tracking objects in the scene, although useful and semantically rich, are difficult to manage except in the case of well-known environments. In opposition, grid-based approaches can handle any kind of environment. Grids can also serve as a basis for further processing steps, which can be object-based. Evidential grids use the theory of evidence and benefit from its properties like natural representation of the unknown and well-developed theoretical tools. Last but not least, the use of evidential grids allows for the fusion of multiple sensors in a straightforward manner. A grid can be constructed for each data source, and all grids can be combined together into one `SensorGrid` (cf. Section 3.5) before further processing. In this article, only one lidar sensor is used, hence there is no need for a multi-sensor fusion step. Yet, it is worth noting that this step is easily realisable.

Despite multiple benefits, occupancy and evidential grids present some challenges. Various computational problems, such as large needs for computing

power are to be mentioned. They can be however diminished on highly efficient modern architectures, exploiting, for instance, many-core systems by parallel computation. Other issues can arise due to necessary coordinate transformations and the transit between local (robot-centred) and global (world-centred) reference frames.

An occupancy grid models the world using a tessellated representation of spatial information. In general, it is a multidimensional spatial lattice with cells storing some stochastic information. In case of an evidential grid, each cell represents a box (a part of environment) $X \times Y$ where $X = [x_-, x_+]$, $Y = [y_-, y_+]$ stores a mass function $m_G^{\Omega}\{X, Y\}$. In this notation, $m_G^{\Omega,(t)}\{X,\ Y\}(A)$ is the mass on Ω of element A for the grid G at time t and at position $\{X,\ Y\}$. Some parts of this notation will be omitted in the following when no risk of confusion exists.

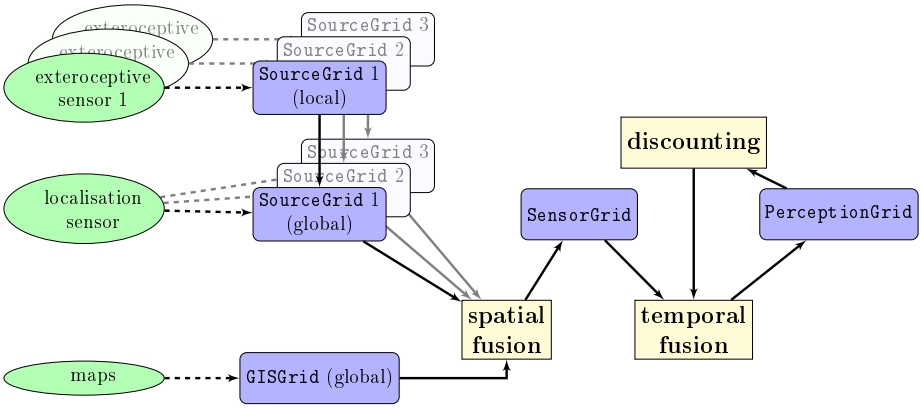


Figure 2: Method overview.

3.1. Map-aided perception system architecture

The high complexity of a perception system deserves a profound consideration on the architecture of the system. There are few works on the architecture of systems using grid-based approach, but some insightful studies of similar topics have been realised by Durekovic and Smith [21] and by Benenson and Parent [22]. Among many foreseeable architectures, the chosen one is considered to be the simplest and the most effective.

Figure 2 presents a general overview of our approach. Starting on the leftmost side, the diagram shows the data sources. The method employs different evidential grids for the storage of: prior information, sensor acquisitions and fusion result. Every processing step will be detailed in Section 4, which contains as well the description of the evidential grids and the fusion method.

All system inputs mentioned on Figure 2 are necessary in the proposed perception system:

- at least one exteroceptive sensor,

- localisation sensor (GNSS receiver, proprioceptive sensor, IMU),
- vector map containing geometric information (roads, buildings).

The presented approach is based on the hypothesis that all information sources are available. If multiple sensors are at hand, the method is adapted to use them and takes advantage of this supplementary information.

An exteroceptive sensor gives a partial view of the vehicle environment. The sensor is assumed to distinguish free and occupied space and model it in 2D x, y or 3D x, y, z coordinates. The coordinates can be relative to the robot or world-referenced. A typical exteroceptive sensor capable of satisfying this assumption is a lidar (laser range scanner), a radar, or a stereo camera system.

A proprioceptive sensor like an Inertial Measurement Unit (IMU) or an odometer hybridised with a GNSS receiver are needed to provide the vehicle pose. Provided pose is assumed to be reliable, accurate and precise. The pose should be globally referenced and is needed to situate the vehicle in the environment. A hypothesis is made that the pose reflects accurately the real state of the vehicle.

Lastly, our method tries to exploit at large the information contained in geographical maps, so we assume that the maps are sufficiently detailed and contain valuable and accurate data. At the minimum, the map data have to contain buildings and road surface description.

3.2. `PerceptionGrid`

One of the evidential grids used in the system is the `PerceptionGrid`. It has been introduced to store the results of information fusion. `PerceptionGrid` is as well the output of the perception system and could be used in further steps, e.g. for navigation. Mass functions of each cell use the frame of discernment (FOD) $\Omega_{PG} = \{F,\ I,\ M,\ S,\ U\}$. The choice of such a FOD is motivated by the objectives to be achieved. Respective classes represent: free space F, mapped infrastructure (buildings) I, mobile moving objects M, temporarily stopped objects S and unmapped infrastructure U. Ω_{PG} is a common, most refined, frame used during data processing when the frames of discernment in question differ. As the `PerceptionGrid` retains the result of information fusion, the need to store previous data disappears.

3.3. `GISGrid`

`GISGrid` is intended to contain all the data exploited from maps. It allows to perform a contextual information fusion incorporating some meta-knowledge about the environment. `GISGrid` uses the FOD $\Omega_{GG} = \{B,\ R,\ T\}$. Class B corresponds to the area occupied by buildings, R to the road surface, and class T models intermediate space, e.g. pavements, where mobile and stationary objects as well as small urban infrastructure can be present. This grid is created, for instance, by projecting map data onto a two-dimensional world-referenced grid. This is the step where the meta-information from maps is included. For example, this meta-knowledge can ban the existence of mobile objects where

buildings are present and, conversely, it indicates the possibility to find these objects on roads. The exact construction method of the `GISGrid` depends on available geodata. It will be presented in Section 5.2.

The FOD Ω_{GG} is different from the common frame Ω_{PG}. Some rules in the theory of evidence, such as Dempster's rule, do not allow the direct combination of BBAs expressed on different frames of discernment. It is then necessary to express every belief assignment on a common frame of discernment before the combination. In this article, the mapping r_{GG} is used when needed:

$$\begin{aligned} r_{GG} &: 2^{\Omega_{GG}} \rightarrow 2^{\Omega_{PG}} \\ r_{GG}\left(\{B\}\right) &= \{I\} \\ r_{GG}\left(\{R\}\right) &= \{F,\ S,\ M\} \\ r_{GG}\left(\{T\}\right) &= \{F,\ U,\ S,\ M\} \\ r_{GG}(A) &= \bigcup_{\theta \in A} r_{GG}(\theta) \qquad (1) \\ & \qquad \forall A \subseteq \Omega_{GG} \text{ and } A \notin \{\{B\},\ \{R\},\ \{T\}\} \end{aligned}$$

and the mass is transferred as follows:

$$m_{GG}^{\Omega_{PG}}\left(r_{GG}\left(A\right)\right) = m_{GG}^{\Omega_{GG}}\left(A\right) \qquad \forall A \subseteq \Omega_{GG} \qquad (2)$$

The mapping r_{GG} indicates that, for instance, building information B fosters mass transfer to class I. On the road surface R, the existence of free space F as well as stopped S and moving M objects is possible. Lastly, on the intermediate area T, the existence of mapped infrastructure I can be excluded, but the presence of the other classes cannot.

3.4. `SourceGrid`s

For each exteroceptive sensor, such as a lidar, an evidential grid called `SourceGrid` is created. In our case, only one grid of this kind is used, but the system architecture permits the use of multiple sensors. For each acquisition, a new grid is computed.

Each cell of the `SourceGrid` stores a mass function m_{S_i} defined on the frame of discernment Ω_{S_i}. Typically for a lidar, $\Omega_{S_i} = \{F,\ O\}$, where F refers to the free space and O to the occupied space. The basic belief assignment depends on the model of the actual sensor. Details about the sensor model used in this article will be given in Section 5.3.

The frame of discernment Ω_{S_i} is distinct from Ω_{PG} and a common frame for all sources has to be found. Hence, a refining r_{S_i} is defined as stated in equation 3.

$$
\begin{aligned}
r_{S_i} &: 2^{\Omega_{S_i}} \rightarrow 2^{\Omega_{PG}} \\
r_{S_i}(\{F\}) &= \{F\} \\
r_{S_i}(\{O\}) &= \{I, U, S, M\} \\
r_{S_i}(A) &= \bigcup_{\theta \in A} r_{S_i}(\theta) \\
&\qquad \forall A \subseteq \Omega_{S_i} \text{ and } A \notin \{\{F\},\ \{O\}\}
\end{aligned} \tag{3}
$$

r_{S_i} makes it possible to perform the fusion of `SourceGrid`$_i$ containing instantaneous grid obtained from sensor i with other grids. Equation 4 expresses the refined mass function.

$$
m_{S_i}^{\Omega_{PG}}\left(r_{S_i}(A)\right) = m_{S_i}^{\Omega_{S_i}}(A) \qquad \forall A \subseteq \Omega_{S_i} \tag{4}
$$

3.5. `SensorGrid`

The evidential `SensorGrid` contains the combined instantaneous information from all sensors and the prior knowledge from the maps. This grid is globally referenced and uses the same frame of discernment Ω_{PG} as the `PerceptionGrid`. `SensorGrid` is the output of the spatial fusion and the input of the temporal fusion, both described in the following sections.

4. Spatio-temporal fusion

4.1. Spatial fusion

As stated before, the proposed method uses maps in order to ameliorate the perception scheme. The maps are the source of prior information which can be used to gain more insight about the vehicle environment. The fusion of the information from `GISGrid` with the sensor data stored in the `SourceGrid`s (cf. Figure 2) is performed on a cell-by-cell basis. At first, all grids are transformed in order to use common frame of discernment Ω_{PG}. Conjunctive rule of combination (denoted $\textcircled{\cap}$) is used to combine the `SourceGrid`s, indexed from 1 to N_s. Then, a normalised version of this operator is applied to the result and to the `GISGrid`, as expressed by equation 5.

$$
m_{SG}^{\Omega_{PG},(t)} = \left(\bigcirc\!\!\!\!\bigcap_{i=1}^{N_s} m_{S_i}^{\Omega_{PG},(t)} \right) \oplus m_{GG}^{\Omega_{PG}} \tag{5}
$$

The conjunctive normalised rule of combination (also called Dempster's rule and denoted $\oplus$) was chosen because the geodata from maps and the sensor data are considered to be independent. Furthermore, the sources are supposed reliable, even if errors are possible. At the end of this stage, the resulting grid, `SensorGrid`, is the combination of the sensor data from `SourceGrid`s with the prior knowledge from `GISGrid`.

4.2. Temporal fusion

The temporal fusion serves the role of combining current sensor acquisition with preceding perception result. The sensor information input has been already combined with prior information as described before. The general form of information fusion operation is expressed by Equation 6.

$$m_{PG}^{(t)} = {}^{\alpha}m'^{(t-1)}_{PG} \circledast m_{SG}^{(t)} \tag{6}$$

Fusion steps combine the `PerceptionGrid` from preceding epoch ${}^{\alpha}m'^{(t-1)}_{PG}$ with the `SensorGrid` from current epoch $m_{SG}^{(t)}$ using fusion operator $\circledast$ which belongs to the family of conjunctive operators. `PerceptionGrid` used for the fusion operation is the result of the process, including the conflict analysis[1] and mass function specialisation denoted by the apostrophe in m'_{PG}, as described in the following paragraphs. Notation ${}^{\alpha}m$ represents the discounting operation as presented in Section 4.3.

4.2.1. Conflict analysis

To exploit dynamic characteristics of the scene, we propose the analysis of inflicted conflict masses. The idea presented in [23] is used here to manage conflict masses. This need arises from the fact that the environment is dynamic. Some authors have elaborated different conflict management to detect changing areas [24]. Two types of conflict are therefore distinguished.

In the proposed fusion scheme, $\emptyset_{FO}$ denotes the conflict induced when a free cell in `PerceptionGrid` is fused with an occupied cell in `SensorGrid`. Analogically, $\emptyset_{OF}$ indicates the conflict mass caused by an occupied cell in `PerceptionGrid` fused with a free cell in `SensorGrid`. Conflict masses are given by:

$$\begin{aligned} m_{PG}^{(t)}(\emptyset_{OF}) &= m_{PG}^{(t-1)}(O) \cdot m_{SG}^{(t)}(\{F\}) \\ m_{PG}^{(t)}(\emptyset_{FO}) &= m_{PG}^{(t-1)}(\{F\}) \cdot m_{SG}^{(t)}(O) \end{aligned} \tag{7}$$

where $m(O) = \sum_{A \subseteq \{I,M,S,U\}} m(A)$. In the ideal case where no noise is present in the input data, conflicts $\emptyset_{FO}, \emptyset_{OF}$ represent, respectively, appearance and disappearance of an object.

4.2.2. Static cell characterisation

Mobile object detection is a crucial issue in dynamic environments. To meet this need, an accumulator ζ is introduced. Secondly, a mass function specialisation using ζ is performed to distinguish temporarily stopped objects from those that are moving.

---

[1] Actually, ${}^{\alpha}m'^{(t-1)}_{PG}$ represents the result of the prediction model: grid from time $t-1$, is used to obtain the predicted state of the world at time t. Hence, both mass functions m_{SG} and m_{PG} describe the same instant of environment state, and so the conflict analysis is justified.

ζ is defined in each cell in order to include temporal information on the cell occupancy. For this purpose, a gain $\delta \in [0, 1]$ and an decrement-to-increment ratio γ have been chosen. Section 5.4 explains what factors influence the computation of these parameters and sheds some light on their physical interpretation.

$$\zeta^{(t)} = \zeta^{(t-1)} + \delta \left[m_{PG}(O) \cdot (1 - m_{\emptyset}) - \gamma \cdot (1 - m_{PG}(O)) \right] \tag{8}$$

Value of ζ is consequently clamped into range [0, 1] so that it could be used in a specialisation matrix.

$$\zeta^{(t)} = \max \left[0,\ \min \left(1,\ \zeta^{(t)} \right) \right] \tag{9}$$

Accumulator ζ behaviour is described in Section 4.2.4. ζ brings a piece of evidence about a more specific set, here the static classes. ζ values are used to specialise mass functions in `PerceptionGrid` using equation 10. Masses on elements of $m_{PG}^{(t)}$ are transferred to $m'^{(t)}_{PG}$ according to specialisation matrix $S^{(t)}$ as presented by equation 10. It is noteworthy to mention that $S^{(t)}(A, B)$ represents the ratio of the mass attributed to set B that will be transferred to set A.

$$m'^{(t)}_{PG}(A) = \sum_{B \subseteq \Omega_{PG}} S^{(t)}(A, B) \cdot m_{PG}^{(t)}(B) \qquad \forall A \subseteq \Omega_{PG} \tag{10}$$

A specialisation matrix $S^{(t)}$ is used to do the mass transfer. Matrix $S^{(t)}$ is identically zero except for the following elements:

$$\begin{cases} S^{(t)}(A \setminus \{M\},\ A) & = \zeta^{(t)} \\ S^{(t)}(A,\ A) & = 1 - \zeta^{(t)} \end{cases} \qquad \forall A \subseteq \Omega_{PG} \text{ and } A \ni \{M\}$$

$$S^{(t)}(A,\ A) = 1 \qquad \forall A \subseteq \Omega_{PG} \text{ and } A \not\ni \{M\} \tag{11}$$

The idea behind the specialisation matrix and the accumulator is that moving objects are differentiated from static or stopped objects. The mass attributed to sets $\{U, S, M\}$ or $\{S, M\}$ will be transferred to $\{U, S\}$ or $\{S\}$, respectively. Additionally, the value of the transferred mass is proportional to the time that the cell in question stayed occupied.

4.2.3. Fusion rule

The fusion rule $\circledast$ from Equation 6 is based on the conjunctive rule of combination, but it has been influenced by Yager's fusion operator [25]. An effort has been made to adapt it for mobile object detection. Yager's operator has the advantage of not attributing more mass than given by the sources to any class except for the unknown Ω (which is the case with normalised rules).

As indicated above, some modifications to the conjunctive rule have to be performed in order to distinguish between moving and stationary objects. These changes consist in transferring the mass corresponding to a newly appeared object to the class of moving objects M as described by equation 12. Fusion rule $\circledast$

has no longer the commutative nor associative properties, but the temporal fusion is performed sequentially and the order is imposed.

$$
\begin{aligned}
(m_1 \circledast m_2)(A) &= \left(m_1 \textcircled{\cap} m_2\right)(A) && \forall A \subsetneq \Omega \text{ and } A \neq M \\
(m_1 \circledast m_2)(\{M\}) &= \left(m_1 \textcircled{\cap} m_2\right)(\{M\}) + \left(m_1 \textcircled{\cap} m_2\right)(\emptyset_{FO}) \\
(m_1 \circledast m_2)(\Omega) &= \left(m_1 \textcircled{\cap} m_2\right)(\Omega) + \left(m_1 \textcircled{\cap} m_2\right)(\emptyset_{OF}) \\
(m_1 \circledast m_2)(\emptyset_{FO}) &= 0 \\
(m_1 \circledast m_2)(\emptyset_{OF}) &= 0
\end{aligned} \tag{12}
$$

The above mentioned steps conduct to the construction of a `PerceptionGrid`, which is the system output. Such a `PerceptionGrid` contains rich information on the environment state. It includes the knowledge on mobile and static cells divided in classes.

4.2.4. Temporal fusion behaviour analysis

The temporal fusion step described in the previous paragraphs is the core part of the proposed fusion scheme. This operation is also a complex one and requires additional comments. Figures 3 and 4 present the behaviour of the proposed fusion scheme in different contexts: road and intermediate space, respectively.[2]

In both figures, parts (a) show the evolution of the mass function of the `SourceGrid`, i.e. the sensor acquisition. Performed simulation models a single cell that stays free for 5 iterations, or sensor cycles. Then, the cell stays occupied until iteration no. 24 and, finally, becomes free again. This situation can be interpreted as an appearance of an object which stays for a longer time and leaves its place in the end.

Parts (b) represent the mass function of the `GISGrid`, so the prior information obtained from maps. In the road context, the major part of the mass is attributed to the FMS class (free space, moving or stopped object), shown as dotted red line. In contrast, in the intermediate space context, the affected class is the FMSU (free space, moving or stopped object, or unmapped infrastructure), depicted by dash-dot blue line.

Parts (c) and (d) consitute the most important part of the comparison. They present the evolution of the `PerceptionGrid` mass function with different values of parameter δ.

Looking at Figures 3c and 3d, one can observe that the evolution of free space mass F (solid green line) follows a well-known pattern. During the iterations 1 to 5, as the sensor detects the cell as free repeatedly, the F mass augments due to conjunctive rule behaviour. When the cell is detected as occupied in

[2]Not all the subsets of the frame of discernment have been shown on these figures, so the sum of visible masses may be less than 1.

the 6th cycle, the pieces of information in `SourceGrid` and `PerceptionGrid` become contradictory and create a conflict, which is transferred to moving objects class M (dotted red line with plus signs). The mass of movable (moving or stopped) class MS (dotted black line with squares) is affected as well due to the discounting operation.

The cell stays occupied until iteration 25. During this period, accumulator ζ increases progressively, which causes the mass transfer from MS to stopped class S (dotted magenta line with diamonds)[3]. The M mass diminishes as the cell stays occupied and finally approaches zero.

When the cell gets free again at iteration 25, a high peak in unknown $\Omega = FIMSU$ mass value is observed. This behaviour arises because of the transfer of $\emptyset_{OF}$ conflict mass to Ω set[4] . At the same moment, the accumulator ζ starts to decrease. In the next cycles, the free mass F grows steadily and the other (occupied) classes diminish rapidly. ζ drops to zero at a much faster rate than it grew before. Such a behaviour is of course attended and should be interpreted as if a stopped object started to move.

Parts (c) and (d) present how the change in the accumulator gain δ impacts the mass evolution. Higher value of δ accelerates the classification of a cell as a stopped one, class S. Similarly, higher value of γ ratio would accentuate the drop of the accumulator value when a stopped object leaves the cell (iteration 25 onwards).

The above description is valid for both Figures 3 and 4 with minor differences. Firstly, in the intermediate space, classes stopped S and unmapped infrastructure U cannot be distinguished by the fusion itself and the difference in values comes mostly from the discounting. Namely, the behaviour of masses MS and S in the road context corresponds to the classes MSU, SU and U in the intermediate space.

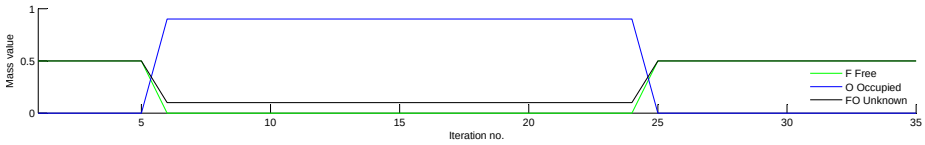
4.3. Discounting

The belief function theory makes it possible to model one's opinion about the reliability of an information source using discounting. The point is that more reliable sources get assigned heavier weights than the less reliable ones. The result of discounting of the basic belief assignment (bba) m is a new bba ${}^{\alpha}m$ (both defined on the FOD Ω), which is obtained by:

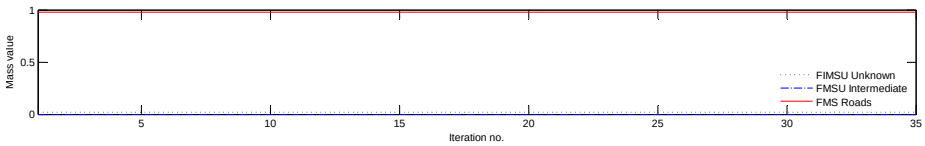
$$\begin{aligned} {}^{\alpha}m(A) &= (1-\alpha)\cdot m(A) && \forall A \subseteq \Omega, A \neq \Omega \\ {}^{\alpha}m(\Omega) &= (1-\alpha)\cdot m(\Omega) + \alpha \end{aligned} \tag{13}$$

[3] The mass transfer from class MS to S is just one example. Actually, all the classes that contain class M are affected and a part of their masses get transferred to their subset without the M class, e.g., from MSU to SU.

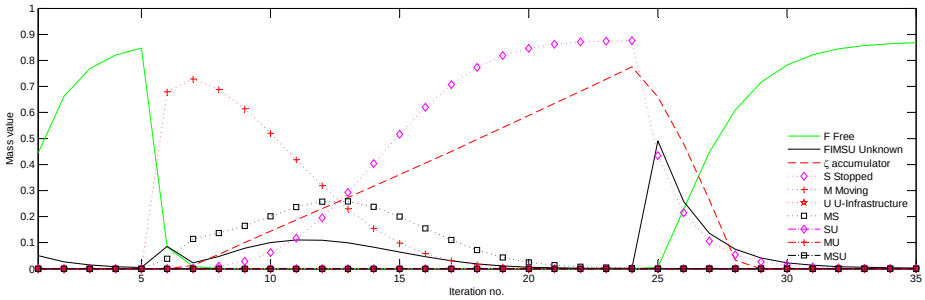
[4] One could argue that the fusion rule is unnecessarily non-symmetrical here, and the $\emptyset_{OF}$ conflict could be transferred to the free class F. However, it seems prudent to postpone the growth of the free mass, especially in the case of the presented application. Namely, an intelligent vehicle should compensate for aberrant input data obtained from not completely reliable sources.



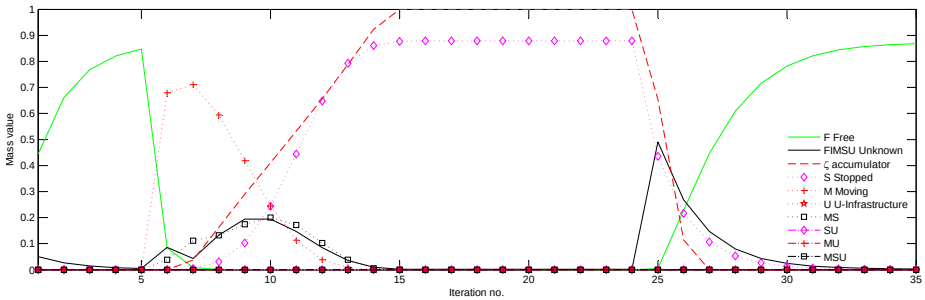
(a) `SourceGrid`.



(b) `GISGrid`.

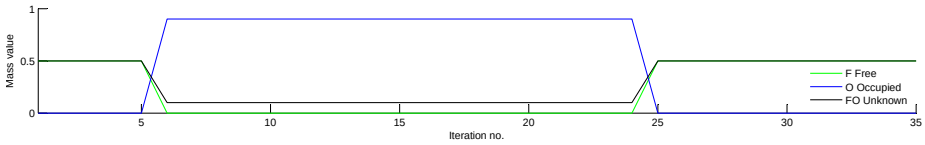


(c) `PerceptionGrid`. $\gamma = 5, \delta = 0.05$

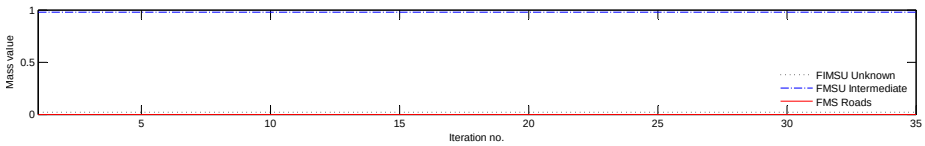


(d) `PerceptionGrid`. $\gamma = 5, \delta = 0.15$

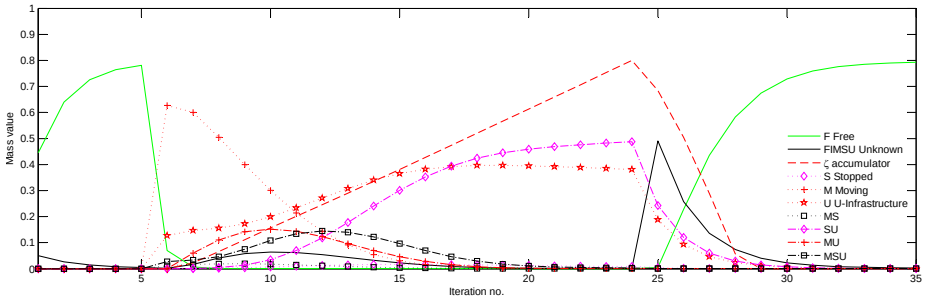
Figure 3: Fusion rule behaviour in the road context.



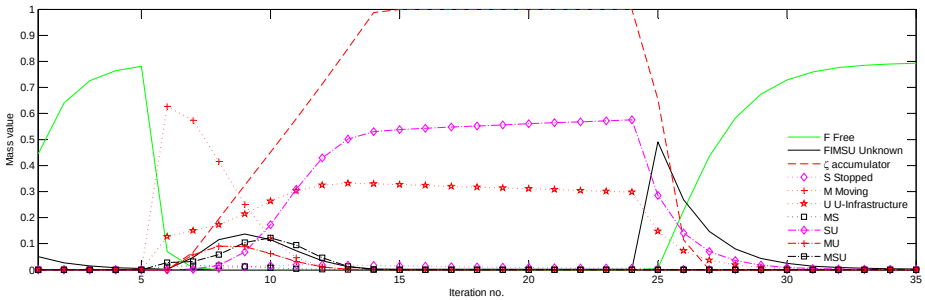
(a) `SourceGrid`.



(b) `GISGrid`.



(c) `PerceptionGrid`. $\gamma = 5, \delta = 0.05$



(d) `PerceptionGrid`. $\gamma = 5, \delta = 0.15$

Figure 4: Fusion rule behaviour in the intermediate space context.

Factor α, called the *discounting factor* may be considered as the level of distrust in this particular sensor.

4.3.1. Temporal discounting

The trust assigned to the information on the `PerceptionGrid` depends on the time elapsed since the acquisition of the data because the environment is dynamic. The term *temporal discounting* denotes this particular discounting operation [26]. Then, temporal discounting can be used to partially "forget" information which is no longer valid. In the frame of mobile perception, the environment changes rapidly and discounting becomes indispensable to avoid keeping obsolete information.

Discounting factor α still models the level of reliability assigned to the source, but it serves also another purpose. In temporal discounting, α represents the speed with which information becomes obsolete. This process is often called "information ageing". Some authors proposed that α is a function of elapsed time $\Delta t = t_{\text{current}} - t_{\text{acquisition}}$ and a remanence characteristic ρ of the event E (information arrival) [26], expressed as a time value in seconds:

$$\alpha = 1 - \exp \frac{\Delta t}{-\rho(E)} \tag{14}$$

Often data arrives at regular intervals and the processing starts instantly, so α can be fixed to a constant value.

4.3.2. Contextual discounting

Mercier et al. in [27, 28] introduced and further developed *contextual discounting* — a type of discounting that makes it possible to adapt the forgetting rate to the context. When more detailed information regarding the confidence ascribed to the sources is available, contextual discounting permits to model this fact. Together with temporal discounting, it can be modelled that different pieces of information become obsolete at different rates. Since meta-knowledge of the robot environment states that some objects (like buildings) do not change rapidly, whereas other do (mobile objects: cars, pedestrians), the contextual discounting process takes these facts into account.

To perform contextual discounting, the frame of discernment Ω has been partitioned[5], each element of partition containing classes with similar dynamics (similar changing rate). We have based the choice of the partition Θ on common sense and divided Ω using the coarsening $\Theta = \{\theta_{static}, \theta_{dynamic}\}$. This choice can be different depending on, among other things, the meta-knowledge possessed about the environment. In the presented method, factor $\alpha_{dynamic}$ was attributed to the free space as well as stopped and moving objects $\theta_{dynamic} = \{F, S, M\}$. Discount rate α_{static} was assigned to the static objects $\theta_{static} = \{I, U\}$.

[5]However, it is not necessary when using the contextual discounting as described in [28].

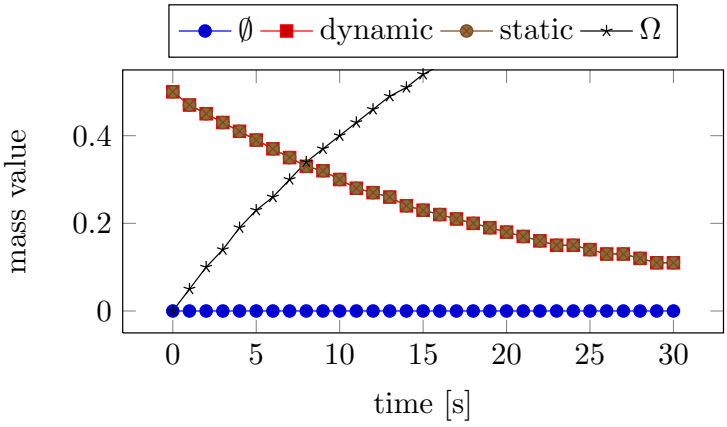


Figure 5: Effect of uniform temporal discounting on different classes. Evolution of m_{PG} as a function of time; $\alpha = 0.05$. The masses on the dynamic and static classes are superposed as they decrease at the same rate.

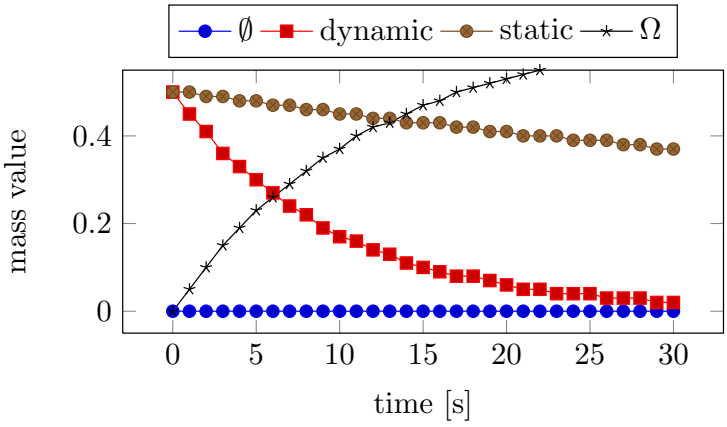


Figure 6: Effect of contextual temporal discounting on different classes. Evolution of m_{PG} as a function of time; $\alpha_{\text{dynamic}} = 0.01$, $\alpha_{\text{static}} = 0.1$. The mass on the dynamic class decreases more rapidly than the mass on the static class.

Next, a mass function has to be created for each element θ of the partition Θ and its corresponding discount factor α as follows:

$$\begin{aligned} m_\theta(\emptyset) &= 1-\alpha \\ m_\theta(\theta) &= \alpha \\ m_\theta(A) &= 0 \qquad\qquad \forall A \subseteq \Omega, A \neq \emptyset, A \neq \theta \end{aligned} \tag{15}$$

The discounted mass function is then computed using the disjunctive combination (denoted $\textcircled{\cup}$) of the input mass function and each function m_θ. By definition of the contextual discounting, factor α_θ is responsible for discounting the masses on classes being the complement of θ.

In our case, the role of the contextual discounting is to control the remanence of different classes. Higher remanence is attributed to static, slowly evolving classes and lower level of persistence can be assigned to rapidly changing, dynamic contexts. Remanence level ρ, if known, can be used to calculate discount factor α for the complement of θ using some decreasing function $f : \mathcal{R} \to [0, 1]$ like $\alpha^c = \frac{1}{1+\rho}$. Nevertheless, discount factors α can be learnt using the learning algorithm presented in [27]. In this case, learnt factors α would serve to quantify the remanence ρ.

The discounted mass function ${}^{\alpha}m_{PG}$ is calculated using:

$${}^{\alpha}m^{(t)}_{PG} = m^{(t)}_{PG} \textcircled{\cup} m_{\theta_{\text{static}}} \textcircled{\cup} m_{\theta_{\text{dynamic}}} \tag{16}$$

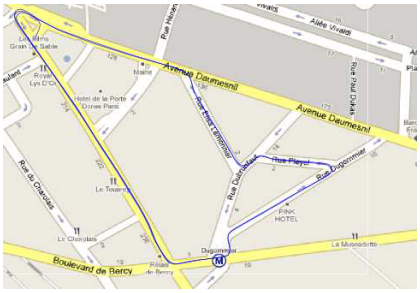
Figure 6 presents the evolution of various mass functions when contextual discounting is applied. The example uses partition $\Theta = \{\text{dynamic, static}\}$ and discount factors $\alpha_{\text{dynamic}} = 0.01$, $\alpha_{\text{static}} = 0.1$. To compare, a classical discounting, presented in Figure 5, behaves differently. With uniform discounting, all masses are being forgotten at the same rate. Contextual discounting allows to slowly forget classes of high remanence and to discount more rapidly less remanent classes.

5. Experimental setup

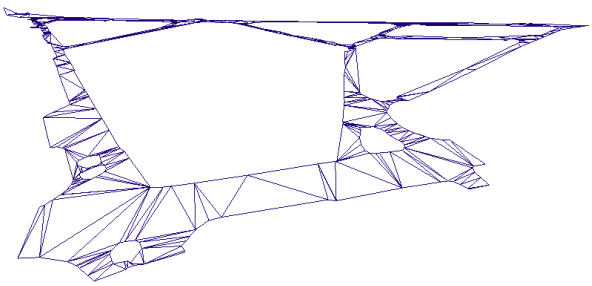
5.1. Data

The data used in this article were acquired thanks to the collaboration with the French National Geographic Institute (IGN). Experiments took place in the 12th district of Paris and the overall length of the test trajectory was approximately 3 kilometres. Other experiments have been performed using the PACPUS platform [29] of the Heudiasyc laboratory and Carmen vehicule (shown in Figure 8). The Applanix sensor based on a GPS, an odometer and an IMU provided one of the system entries, namely the vehicle pose. The pose given by Applanix is supposed precise and of high confidence. As the exteroceptive sensor, an IBEO Alaska XT lidar was used. It provided a point cloud of 800 impacts at a frequency of $10Hz$.

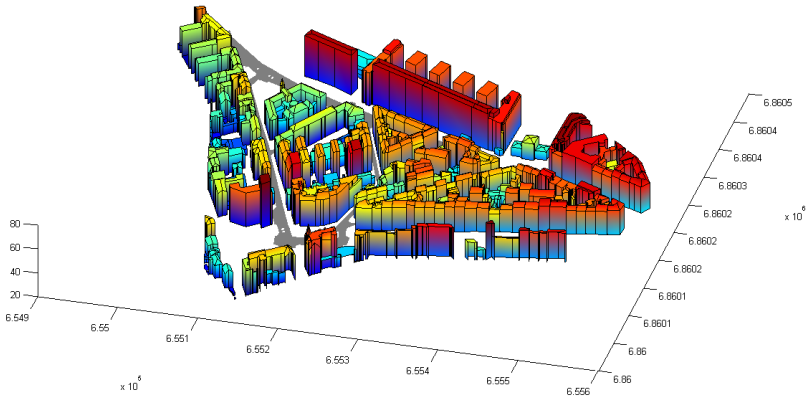
The vector maps were provided by the IGN and contain 3D models of road surfaces and buildings (see Figure 7). Successful tests were also performed with



(a) Vehicle trajectory.



(b) Road surface polygons.



(c) Buildings.

Figure 7: Visual representation of the dataset.

free editable 2D maps from the *OpenStreetMap* project [30], but their use was limited to building data. The maps were accurate and up-to-date.

5.2. `GISGrid` *construction*

The map data may be represented by two sets of polygons $\mathcal{B}$ and $\mathcal{R}$, see equations 17 and 18. Each polygon b_i is described by m_i vertices in 2D Cartesian coordinates. x coordinate denotes longitude and y coordinate indicates latitude of a vertex. A polygon is composed of segments (x_1, y_1)–$(x_2, y_2), (x_2, y_2)$–(x_3, y_3), $\ldots, (x_{m_{i-1}}, y_{m_{i-1}})$–$(x_{m_i}, y_{m_i}), (x_{m_i}, y_{m_i})$–$(x_{m_1}, y_{m_1})$.

- Buildings:

$$\mathcal{B} = \left\{ b_i = \begin{bmatrix} x_1 x_2 \ldots x_{m_i} \\ y_1 y_2 \ldots y_{m_i} \end{bmatrix}, i \in [0, n_B] \right\} \tag{17}$$

Figure 8: Test vehicle Carmen with the lidar sensor in front.

- Road surface:

$$\mathcal{R} = \left\{ r_i = \begin{bmatrix} x_1 x_2 \dots x_{m_i} \\ y_1 y_2 \dots y_{m_i} \end{bmatrix}, i \in [0, n_R] \right\} \tag{18}$$

The polygons satisfy the condition:

$$\mathcal{B} \cap \mathcal{R} = \emptyset$$

A level of confidence β is defined for each map source and is possibly different for each context[6]. Let $\tilde{x} = \frac{x_- + x_+}{2}$, $\tilde{y} = \frac{y_- + y_+}{2}$, then:

$$m_{GG}\{X,Y\}(B) = \begin{cases} \beta_B & \text{if } (\tilde{x},\tilde{y}) \in b_i \\ 0 & \text{otherwise} \end{cases} \tag{19}$$
$$\forall i \in [0, n_B]$$

$$m_{GG}\{X,Y\}(R) = \begin{cases} \beta_R & \text{if } (\tilde{x},\tilde{y}) \in r_i \\ 0 & \text{otherwise} \end{cases} \tag{20}$$
$$\forall i \in [0, n_R]$$

$$m_{GG}\{X,Y\}(T) = \begin{cases} 0 & \text{if } (\tilde{x},\tilde{y}) \in b_i \vee (\tilde{x},\tilde{y}) \in r_j \\ \beta_T & \text{otherwise} \end{cases} \tag{21}$$
$$\forall i \in [0, n_B], \forall j \in [0, n_R]$$

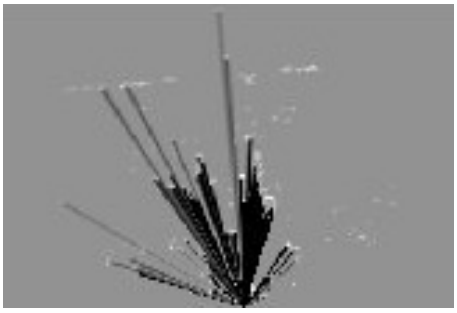
$$m_{GG}\{X,Y\}(\Omega) = \begin{cases} 1-\beta_B & \text{if } (\tilde{x},\tilde{y}) \in b_i \\ 1-\beta_R & \text{if } (\tilde{x},\tilde{y}) \in r_i \\ 1-\beta_T & \text{otherwise} \end{cases} \tag{22}$$
$$\forall i \in [0, n_B], \forall j \in [0, n_R]$$

[6] In our case, however, $\beta_B = \beta_R = \beta_T = \beta$.

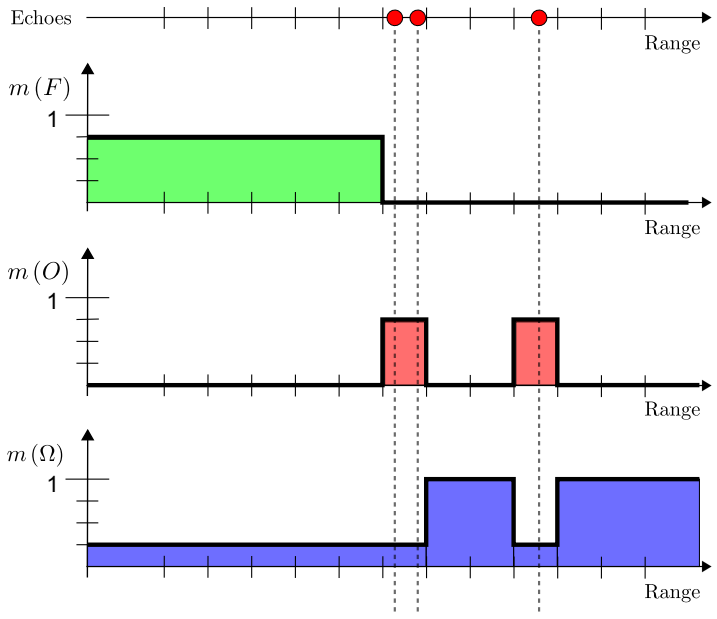
5.3. Lidar sensor model

This section describes the way in which the data obtained from the sensor are transformed into the `SourceGrid`. If another exteroceptive sensor is used, one has to define an appropriate model. The model used here is based on the one described in [17].

(a) Camera view of the scene.



(b) `SourceGrid` projected into a Cartesian frame. Colour code: white – occupied, black – free, grey – unknown.



(c) Example of one angular segment of a lidar acquisition. Red dots represent laser impacts, diagrams show mass attribution.

Figure 9: Multi-echo lidar sensor model.

5.3.1. `SourceGrid` construction

Once the sensor model is defined, it can be used to build a `SourceGrid` from each lidar scan. A lidar is a polar sensor, so a polar grid model is used to compute the occupancy of the cells. The sensor precision is higher than the grid resolution and the grid is sensor-centred.

Each row of the polar `SourceGrid` corresponds to one angular sector. In one sector, several echoes are possible due to the following reasons:

- several echoes from one direction were received;
- projection of 4 layers on 2D plane provided echoes located at different distances;
- lidar angular resolution is higher than the grid resolution, i.e. several lidar directions are projected on the same column (angular segment) of the polar grid.

It is worth noticing that this grid sensor model takes into account the sensor multi-echo capabilities. This means that the sensor can provide several echoes for one angle and one layer. Each angular sector $\Phi = [\phi_-, \phi_+]$ of the grid is processed independently. We define a basic belief assignment (bba) for each cell computed from the sensor scan points. One cell is a box $R \times \Phi$, $R = [r_-, r_+]$. Let

$$Z_\Phi = \left\{ z_i = \begin{bmatrix} r_i \\ \phi_i \end{bmatrix}, \phi_i \in \Phi, i \in [0, n] \right\}$$

be the set of n scan points in the angular sector Φ. Then, the bbas of the `SourceGrid` are defined as illustrated by Figure 9c and given by equations:

$$m_S\left\{\Phi, R\right\}(O) = \begin{cases} \mu_O & \text{if } r_i \in R \\ 0 & \text{otherwise} \end{cases} \qquad \forall i \in [0, n] \tag{23}$$

$$m_S\left\{\Phi, R\right\}(F) = \begin{cases} \mu_F & \text{if } min(r_i) < r_+ \\ 0 & \text{otherwise} \end{cases} \qquad \forall i \in [0, n] \tag{24}$$

$$m_S\left\{\Phi, R\right\}(\Omega) = \begin{cases} 1 - \mu_F & \text{if } min(r_i) < r_+ \\ 1 - \mu_O & \text{if } r_i \in R \\ 1 & \text{otherwise} \end{cases} \qquad \forall i \in [0, n] \tag{25}$$

Values μ_O, μ_F represent the confidence in the measurement for the given sensor [17, Section III.B]. An example of a `SourceGrid` where multi-echoes are observed is given in Figure 9b.

5.3.2. From `SourceGrid` to `SensorGrid`

Each `SourceGrid` can be projected onto the `SensorGrid` using the pose provided by the proprioceptive sensor. Figure 10 illustrates the general idea of the process. The polar `SourceGrid` is converted to Cartesian coordinates using a bi-linear interpolation. Next, a transformation of the Cartesian `SourceGrid` is

applied in order to obtain a world-referenced grid. This transformation consists of one rotation and one translation. The rotation is done with a bi-linear transformation, because one cell may be partially projected on many cells. Bi-linear transformation can interpolate values, so, in the transformed cell, masses are set to mean values of the neighbourhood of the polar cell. Such a method can cause a phenomena of edge smoothing, but well chosen grid size renders this effect negligible. In our implementation, this transformation and interpolation have been performed using the image processing library OpenCV [31].

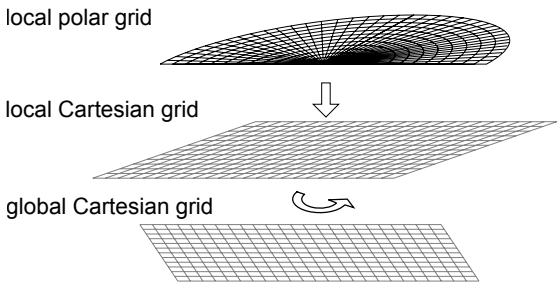


Figure 10: Transformations: from local polar grid to global Cartesian grid.

5.4. Parameters

The size of the grid cell in the occupancy grids was set to 0.5 m, which is sufficient to model a complex environment with mobile objects. The discount rates α describing the speed of information becoming obsolete were defined empirically. We have defined the map confidence factor β by ourselves and set it to 0.98, but ideally, it should be given by the map provider. β describes data currentness (age), errors introduced by geometry simplification and spatial discretisation. β can also be used to depict the localisation accuracy. Parameters δ (accumulator gain) and γ (decrement-to-increment ratio) used for static object detection determine the sensitiveness of this detection process.

To compute the values δ and γ, let consider an object of a length L moving at a speed V_{min} which is the minimum speed one would be able to detect. This configuration is illustrated by Figure 11.

The exteroceptive sensor provides scans with a constant frequency f. One can compute the maximum number of sensor cycles i_{point} necessary for a point of an object to pass through a cell of size Δ as follows:

$$i_{\text{point}} = \frac{\Delta \cdot f}{V_{min}} \tag{26}$$

The object of length L occupies at least $\frac{L}{\Delta}$ cells, so, analogically to Equation 26, we can compute the number of cycles i_{object} during which this object occupies one cell:

$$i_{\text{object}} = \frac{L \cdot f}{V_{min}} \tag{27}$$

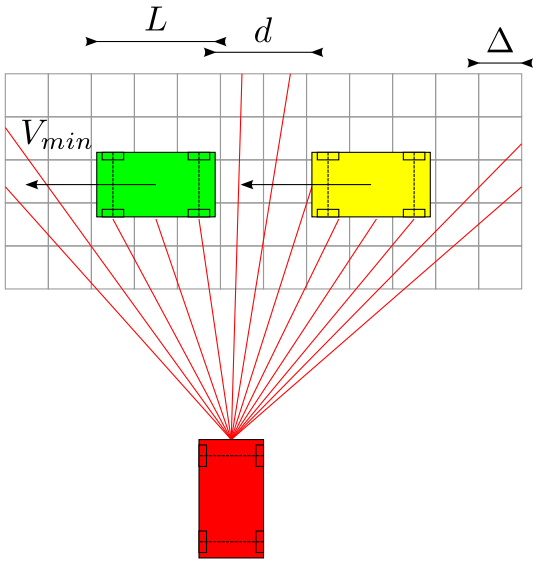


Figure 11: Illustration of geometrical distances used for the parameter computation.

δ is computed in such a manner that the accumulator ζ grows up to 1 in i_{object} cycles.

$$\delta \cdot i_{\text{object}} = 1 \tag{28}$$

And then, from Equations 27 and 28:

$$\delta = \frac{V_{min}}{L \cdot f} \tag{29}$$

For a typical moving car at a minimal speed $V_{min} = 1$ m/s with a length $L = 3$ m and the sensor scan frequency $f = 15$ Hz, one obtains $\delta = 0.02$.

Coefficient γ is determined in a similar way to δ, but in this case we consider the inter-object distance d between two successive objects. The fact that the accumulator ζ should reach 0 before the following object enters into this cell implies:

$$\delta \cdot \gamma = \frac{V_{min}}{d \cdot f} \tag{30}$$

So we obtain:

$$\gamma = \frac{L}{d} \tag{31}$$

Setting an inter-obstacle distance to $d = 0.5$ m, one computes decrement-to-increment ratio $\gamma = 6$.

Table 1 lists all the parameter values used for the construction of `SourceGrid` (see Section 5.3.1) and the fusion operation (see Section 4).

Parameter name	Value
Map confidence level β	0.98
Cell size Δ	0.5 m
Accumulator gain δ	0.02
Decrement-to-increment ratio γ	6
Free space confidence μ_F	0.7
Occupied space confidence μ_O	0.8

Table 1: Parameter values used in the experimental setup.

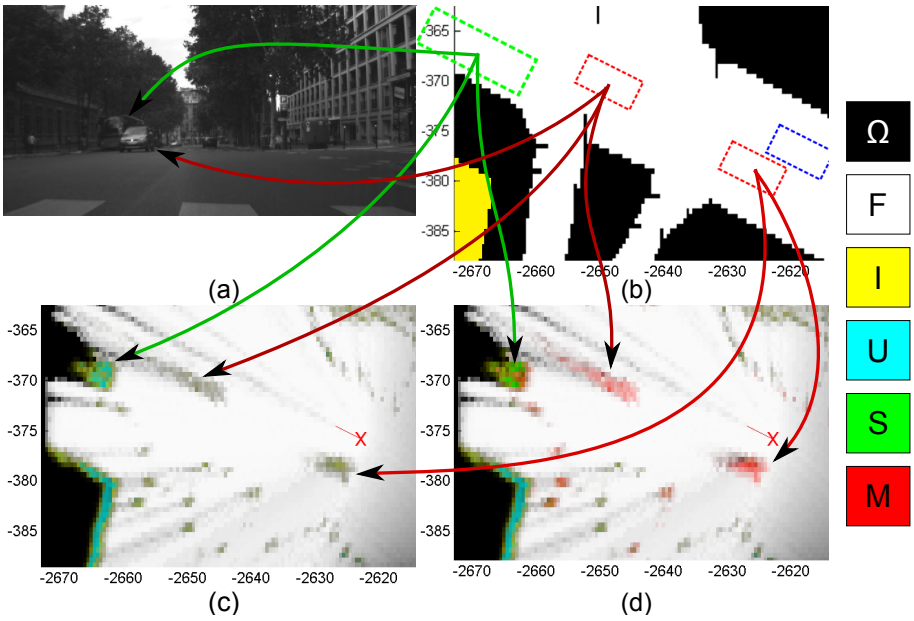
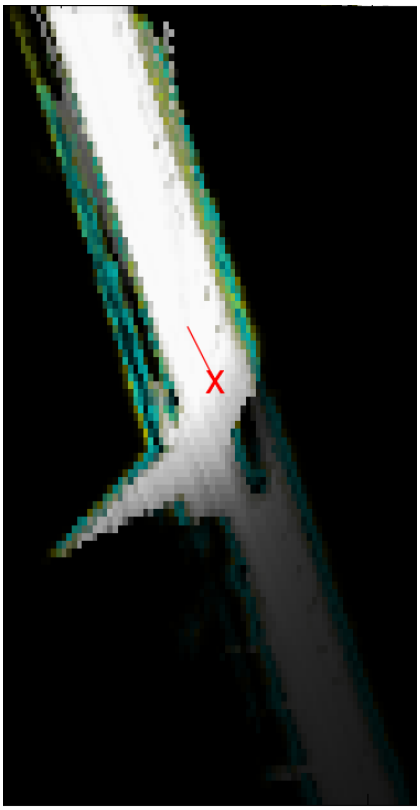
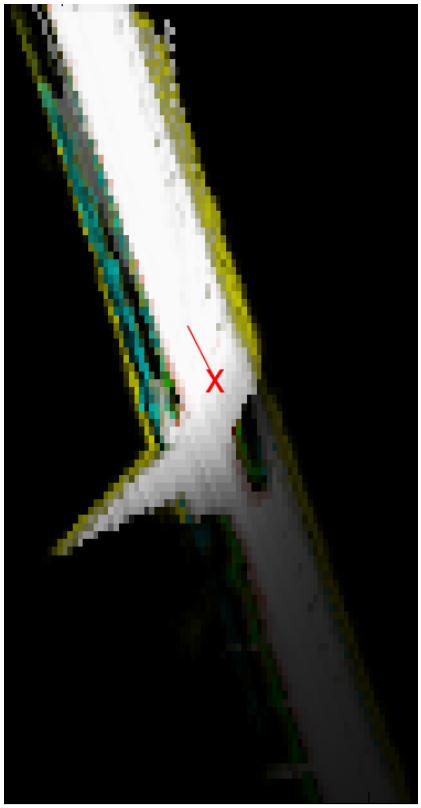


Figure 12: (a) Scene. (b) `GISGrid` with superposed vehicle outlines. (c) `PerceptionGrid` without prior information. (d) `PerceptionGrid` with prior map knowledge. Right: colour code.

(a) Camera acquisition



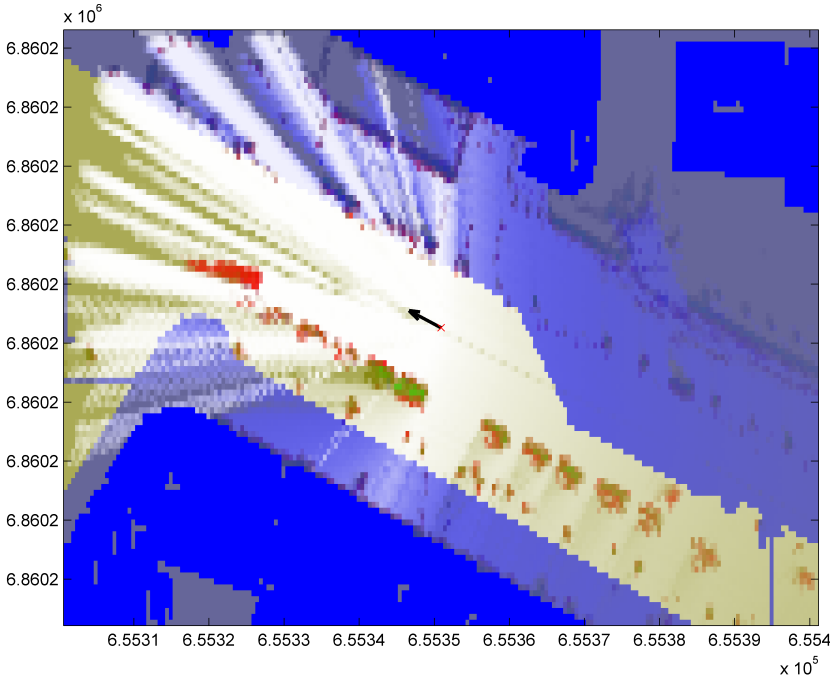
(b) `PerceptionGrid` without map



(c) `PerceptionGrid` with map

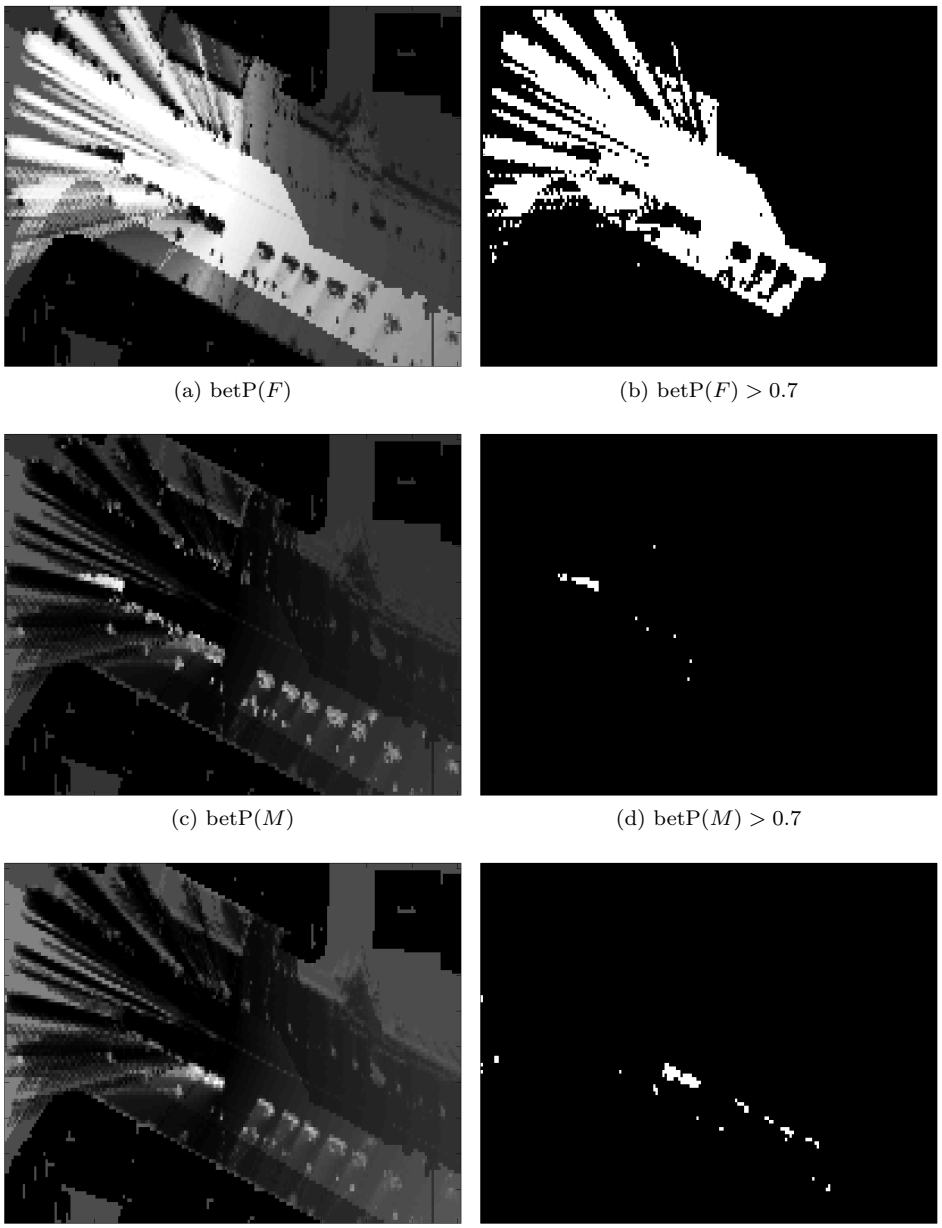
Figure 13: `PerceptionGrid` comparison and scene snapshot. Using classical (uniform) discounting. Colour code as in Figure 12.

(a) Scene representation by a camera image.



(b) `PerceptionGrid` – pignistic probability.

Figure 14: Camera acquisition and the pignistic probabilities of `PerceptionGrid`. Coordinates in meters in ENU frame.



(a) betP(F) (b) betP(F) > 0.7

(c) betP(M) (d) betP(M) > 0.7

(e) betP(S) (f) betP(S) > 0.35

Figure 15: `PerceptionGrid`. Left column: pignistic probability (betP) for different classes. Right column: a simple decision rule example – threshold on pignistic probability.

6. Results

To assess the performance of our method, two cases have been considered. Firstly, when maps are present and prior information can be exploited. Secondly, when no maps are available and only mobile and static detection is done. A comparison of perception results for these two cases has then been shown. In this way, we show the interest of using a map-aided approach to the perception problem. All tests have been performed on real-world data recorded in an urban environment in Paris, France.

The perception results for a particular instant of the tested approach are presented on Figure 12. The visualisation of the `PerceptionGrid` has been obtained by attributing a colour to each class with proportional to the mass value and calculating the mean colour. The presented scene contains two cars, one visible in the camera image and one invisible, both going in the direction opposite to the test vehicle, and a bus parked on the road edge. Bus and car positions are marked on the grids by green and red boxes, respectively. The position of the test vehicle is shown as a blue box. Classes of Ω_{PG} are represented by different colours as described at the right side of Figure 12. `GISGrid` in Figure 12b visualises the prior knowledge obtained from maps by showing the position of the road surface, in white, and buildings, in blue.

The advantage of using map knowledge is richer information on the detected objects. A difference between moving cells (red, car) and stopped ones (green, bus) is clearly visible. Also, stopped objects are distinct from infrastructure when prior map information is available (cf. Figures 12c and 12d). In addition, thanks to the prior knowledge, stationary cells, in cyan, modelling infrastructure are distinguished from stopped cells (road objects).

Figure 12 shows the effect of temporal discounting, which is particularly visible on the free space behind the vehicle. On the other hand, the parked bus is still in evidence despite being occluded by the passing car. Masses attributed to grid cells are being discounted, so the mass on the free class F diminishes gradually.

In Figure 13, there is a clear difference between the perception of buildings when the map data is available and when it is not. When no map is present, buildings are confused with barriers and bicycles. With prior information from the maps, the proposed approach easily distinguishes between the two.

Following paragraphs use the pignistic probability that has been introduced by Smets in [32] as a method of transforming masses into probabilities. This mechanism is used in order to show an example of a decision rule that can be executed on an evidential grid.

$$\mathrm{betP}(A) = \sum_{\substack{B \subseteq \Omega \\ B \neq \emptyset}} \frac{m(B)}{1 - m(\emptyset)} \cdot \frac{|A \cap B|}{|B|} \tag{32}$$

Equation 32 is valid for any A being a subset of the frame of discernment.

In the above results, we have filtered only the cells that have been reached by the sensor at least once in order to put forward the sensor information. Below,

Figures 14 and 15 show all the cells. Both these figures present the same scene. Ahead at the left side, one can observe a moving white car, whereas closer at the left, a column of vehicles parked on the road surface. Figure 14a shows the scene captured by a FishEye camera. On Figure 14b, we present the singleton pignistic probabilities of the masses contained in `PerceptionGrid`. Vehicle position and its direction are represented on this figure by a red cross and a black arrow, respectively. We have changed the colour code, so that the classes of interest could be easily visible in the figure. The corresponding colours are as follows: F – white, I, U – blue, M – red and S – green. Other colours are the result of the fact that the pignistic probabilities (and their corresponding colours) are mixed together.

To better understand the information contained in the `PerceptionGrid`, we display, in Figure 15, the pignistic probabilities at a grid level of a few classes of interest. The left column contains images which are a visualisation of the pignistic probability for a given class. The right-column images give an example of a simple binary decision rule based on a threshold of the value of pignistic probability. These figures highlight the ability of the method to distinguish different classes.

One can spot the effect of the discounting, especially, in Figures 15a and 15b. Namely, the space far behind the ego-vehicle is no more recognised as free, since it has been observed for a long time. Figures 15d and 15f demonstrate the performance of the proposed approach in classifying moving cells M and stopped cells S. One can remark as well a few outliers due to the sensor noise.

Obtained results constitute only the first level of a perception system. Fully exploiting these data would mean performing further processing on the resulting grid. Clustering the cells into more meaningful object-level information and tracking these objects would be the next step towards the scene understanding.

7. Conclusion and perspectives

The article has presented a mobile perception scheme based on prior knowledge from vector maps. Map data allows to infer more refined information about the environment. A modified fusion rule taking into account the existence of mobile objects has been defined.

The remanence can be controlled using classical tools of the theory of belief functions. Contextual discounting has been used in order to represent the variation in information lifetime of objects present in the environment. At the end, there are few parameters to tune: mainly remanence level for each context. Moreover, they are easy to interpret and can be learnt given reference data.

Consequently, one of the perspectives for future work is the use of reference data to learn algorithm parameters, to choose the most appropriate fusion rule and to quantitatively validate the results. It is envisioned that the hypothesis of accurate maps will be removed. Considerable work on creating appropriate error models for the data source will be needed. Moreover, we anticipate distinguishing between navigable and non-navigable free space. This step will

entail the introduction of an additional class in the frame of discernment of the `PerceptionGrid`. Such an improvement will be a step towards the use of our approach for navigation system in autonomous vehicles. Map information will be used to predict object movements. Lastly, more work is to be done to fully explore and exploit 3D map information.

Acknowledgements

The dataset used comes from CityVIP project ANR-07_TSFA-013-01. This work has been supported by the French Ministry of Defence DGA (Direction Générale de l'Armement) with a Ph.D. grant delivered to Marek Kurdej.

Bibliography

[1] M. Kurdej, J. Moras, V. Cherfaoui, P. Bonnifait, Map-aided Fusion Using Evidential Grids for Mobile Perception in Urban Environment, in: International Conference on Belief Functions, Springer, Compiègne, 2012, pp. 343–350.

[2] D. F. Wolf, G. S. Sukhatme, Mobile Robot Simultaneous Localization and Mapping in Dynamic Environments, Autonomous Robots 19 (2005) 53–65.

[3] S. Thrun, W. Burgard, D. Fox, Probabilistic Robotics (Intelligent Robotics and Autonomous Agents), MIT Press, Cambridge, Massachusetts, USA, 2005.

[4] T. Gindele, S. Brechtel, J. Schröder, R. Dillmann, Bayesian Occupancy Grid Filter for Dynamic Environments Using Prior Map Knowledge, IEEE Intelligent Vehicles Symposium (2009) 669–676.

[5] M. Hentschel, O. Wulf, B. Wagner, A GPS and laser-based localization for urban and non-urban outdoor environments, in: IEEE International Conference on Intelligent Robots Systems, 2008, pp. 149–154. doi:10.1109/IROS.2008.4650585.

[6] C. Cappelle, M. E. El Najjar, F. Charpillet, D. Pomorski, Virtual 3D City Model for Navigation in Urban Areas, Journal of Intelligent Robotic Systems 66 (3) (2012) 377–399. doi:10.1007/s10846-011-9594-0.

[7] M. Hentschel, B. Wagner, Autonomous Robot Navigation Based on OpenStreetMap Geodata, in: International IEEE Annual Conference on Intelligent Transportation Systems, Madeira Island, 2010, pp. 1645–1650. doi:10.1109/ITSC.2010.5625092.

[8] M. Dawood, C. Cappelle, M. E. E. Najjar, M. Khalil, D. Pomorski, Vehicle geo-localization based on IMM-UKF data fusion using a GPS receiver, a video camera and a 3D city model, in: IEEE Intelligent Vehicles Symposium, 2011, pp. 510–515.

[9] A. P. Dempster, A Generalization of Bayesian Inference, Journal of the Royal Statistical Society 30 (1968) 205–247.

[10] G. R. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.

[11] P. Smets, What is Dempster-Shafer's model?, John Wiley & Sons, 1994.

[12] P. Smets, Decision Making in the TBM: the Necessity of the Pignistic Transformation, International Journal of Approximate Reasoning 38 (2) (2005) 133–147.

[13] J. Klein, O. Colot, Singular sources mining using evidential conflict analysis, International Journal of Approximate Reasoning 52 (9) (2011) 1433–1451. doi:10.1016/j.ijar.2011.08.005.

[14] A. Ayoun, P. Smets, Data association in multi-target detection using the transferable belief model, International Journal of Intelligent Systems 16 (10) (2001) 1167–1182. doi:10.1002/int.1054.

[15] R. Kruse, F. Klawonn, Mass distributions on L-fuzzy sets and families of frames of discernment, in: Advances in the Dempster-Shafer Theory of Evidence, John Wiley & Sons, Inc., New York, NY, 1994, pp. 239–250.

[16] J. Moras, V. Cherfaoui, P. Bonnifait, A lidar Perception Scheme for Intelligent Vehicle Navigation, in: International Conference on Control, Automation, Robotics and Vision, IEEE, Singapore, 2010, pp. 1809–1814.

[17] J. Moras, V. Cherfaoui, P. Bonnifait, Credibilist Occupancy Grids for Vehicle Perception in Dynamic Environments, in: IEEE International Conference on Robotics and Automation, Shanghai, 2011, pp. 84–89.

[18] J. Klein, C. Lecomte, P. Miché, Hierarchical and conditional combination of belief functions induced by visual tracking, International Journal of Approximate Reasoning 51 (4) (2010) 410–428. doi:10.1016/j.ijar.2009.12.001.

[19] A. Elfes, Using Occupancy Grids for Mobile Robot Perception and Navigation, Computer Journal 22 (6) (1989) 46–57.

[20] D. Pagac, E. M. Nebot, H. Durrant-Whyte, An evidential approach to map-building for autonomous vehicles, IEEE Transactions on Robotics and Automation 14 (4) (1998) 623–629. doi:10.1109/70.704234.

[21] S. Durekovic, N. Smith, Architectures of Map-Supported ADAS, IEEE Intelligent Vehicles Symposium (2011) 207–211.

[22] R. Benenson, M. N. Parent, Design of an urban driverless ground vehicle, IEEE International Conference on Intelligent Robots Systems (2008) 16–21.

[23] J. Moras, V. Cherfaoui, P. Bonnifait, Moving Objects Detection by Conflict Analysis in Evidential Grids, IEEE Intelligent Vehicles Symposium (2011) 1120–1125.

[24] E. Ramasso, C. Panagiotakis, M. Rombaut, D. Pellerin, Belief Scheduler based on model failure detection in the TBM framework. Application to human activity recognition, International Journal of Approximate Reasoning 51 (7) (2010) 846–865. doi:10.1016/j.ijar.2010.04.005.

[25] R. R. Yager, On the Dempster-Shafer framework and new combination rules, International Journal of Information Sciences 41 (2) (1987) 93–137. doi:10.1016/0020-0255(87)90007-7.

[26] V. Cherfaoui, T. Denœux, Z. L. Cherfi, Confidence Management in Vehicular Network, in: Y. Zhang, M. H (Eds.), Confidence Management in Vehicular Network, CRC Press, 2010, pp. 357–378.

[27] D. Mercier, B. Quost, T. Denœux, Refined modeling of sensor reliability in the belief function framework using contextual discounting, Information Fusion 9 (1) (2006) 246–258.

[28] D. Mercier, E. Lefèvre, F. Delmotte, Belief Functions Contextual Discounting and Canonical Decompositions, International Journal of Approximate Reasoning 53 (2) (2012) 146–158. doi:10.1016/j.ijar.2011.06.005.

[29] (Heudiasyc UMR CNRS 7253), PACPUS plaftorm.
URL http://www.hds.utc.fr/pacpus

[30] OpenStreetMap.
URL http://www.openstreetmap.org

[31] OpenCV vision library.
URL http://opencv.org/

[32] P. Smets, R. Kennes, The Transferable Belief Model, Artificial Intelligence 66 (1994) 191–234.