

Map-aided Fusion Using Evidential Grids for Mobile Perception in Urban Environment

Marek Kurdej, Julien Moras, Véronique Cherfaoui, Philippe Bonnifait

Abstract Evidential grids have been recently used for mobile object perception. The novelty of this article is to propose a perception scheme using prior map knowledge. A geographic map is considered an additional source of information fused with a grid representing sensor data. Yager's rule is adapted to exploit the Dempster-Shafer conflict information at large. In order to distinguish stationary and mobile objects, a counter is introduced and used as a factor for mass function specialisation. Contextual discounting is used, since we assume that different pieces of information become obsolete at different rates. Tests on real-world data are also presented.

1 Introduction

Autonomous driving has been an important challenge in recent years. Navigation and precise localisation aside, environment perception is an important on-board system of a self-driven vehicle. The level of difficulty in autonomous driving increases in urban environments, where a good scene understanding makes the perception subsystem crucial. There are several reasons that make cities a demanding environment. Poor satellite visibility deteriorates the precision of GPS positioning. Vehicle trajectories are hard to predict due to high variation in speed and direction. Also, the sheer number of mobile objects poses a problem, e.g. for tracking algorithms.

On the other hand, more and more detailed and precise geographic databases become available. This source of information has not been well examined yet, hence our approach of incorporating prior knowledge from digital maps in order to improve perception scheme. A substantial amount of research has focused on the mapping problem for autonomous vehicles, e.g. Simultaneous Localisation and Mapping (SLAM) approach, but the use of maps for perception is still understudied.

Marek Kurdej e-mail: marek.kurdej@hds.utc.fr · Véronique Cherfaoui · Julien Moras · Philippe Bonnifait
UMR CNRS 6599 Heudiasyc, University of Technology of Compiègne, France

In this article, we propose a data fusion method based on Dempster–Shafer theory [8] taking into account meta-knowledge obtained from a digital map. We show the advantage of including prior knowledge into an embedded perception system of an autonomous car. The vehicle environment is modelled by 2D occupancy grids proposed in [2]. This paper describes a robust and unified approach to a variety of problems in spatial representation using the theory of probability. The theory of evidence was not combined with occupancy grids until recently to build environment maps for robot perception [7]. Only recent works take advantage of the theory of evidence in the context of mobile perception [4]. Some works use 3D city model as a source of prior knowledge for localisation and vision-based perception [1], whereas our method uses maps for scene understanding.

This article is organised as follows. In section 2, we describe the details of the method. Section 3 presents the results and section 4 concludes the paper.

2 Multi-grid fusion approach

This section presents the proposed perception schemes. The grid construction method is described in section 2.2 and all data processing steps are detailed in section 2.4. Figure 1 presents a general overview of our approach.

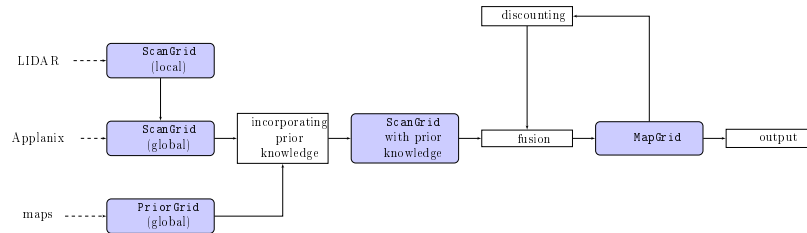


Fig. 1 Method overview (lidar: laser scanner, Applanix: inertial measurement unit).

2.1 Heterogeneous data sources

There are three sources in our perception system: vehicle pose, lidar range scanner point cloud and vector maps. The vehicle pose comes from the Applanix system based on a GPS, an odometer and an IMU. The system is supposed to provide precise and integral positioning. Our main source of information about the environment is an IBEO Alaska XT lidar able to provide a cloud of about 800 points 10 times per second. The digital maps that we use were provided by the French National Geographic Institute (IGN) and contain 3D building models as well as the road surface.

We also performed successful tests with freely available OpenStreetMap project 2D maps [6], but here we limited the use to building data. We assume the maps to be precise and accurate.

2.2 Occupancy grids

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 our case, each cell representing a box (a part of environment) $X \times Y$ where $X = [x_-, x_+]$, $Y = [y_-, y_+]$ stores a mass function.

- **ScanGrid (SG) construction:** In order to process the lidar data, an evidential occupancy grid is computed when a new scan arrives, this grid is called `ScanGrid`. Each cell of this grid stores a mass function on the frame of discernment (FOD) $\Omega_{SG} = \{F, O\}$, where F refers to the free space and O – to the occupied space. The basic belief assignment, which reflects the sensor model, is described in [4].
- **MapGrid (MG):** To store the results of information fusion, an occupancy grid MG has been introduced with a FOD $\Omega_{MG} = \{F, C, N, S, V\}$. Respective classes represent: free space F , mapped infrastructure (buildings) C , non-mapped infrastructure N , temporarily stopped objects S and mobile (moving) V objects. Ω_{MG} is a common frame used for information fusion. By using MG as a cumulative information storage, we are not obliged to aggregate preceding `ScanGrids`.
- **PriorGrid (PG) context representation:** PG allows us to perform a contextual information fusion incorporating some meta-knowledge about the environment. This grid uses the same frame of discernment Ω_{MG} as MG. The grid is obtained by projection of map data, buildings and roads, onto a 2D grid with global coordinates.

We define two sets of polygons defining the 2D position of buildings and road surface by, respectively, $\mathcal{B} = \left\{ b_i = \begin{bmatrix} x_1 x_2 \dots x_{m_i} \\ y_1 y_2 \dots y_{m_i} \end{bmatrix}, i \in [0, n_B] \right\}$ and $\mathcal{R} = \left\{ r_i = \begin{bmatrix} x_1 x_2 \dots x_{m_i} \\ y_1 y_2 \dots y_{m_i} \end{bmatrix}, i \in [0, n_R] \right\}$, $\mathcal{B} \cap \mathcal{R} = \emptyset$. Then, we attribute the mass to each cell $\{X, Y\}$ of the `PriorGrid` in the following way:

We note that $B = \{C\}$, $R = \{F, S, V\}$, $T = \{F, N, S, V\}$ for convenience and readability only. A denotes all other strict subsets of Ω . These aliases characterise the meta-information inferred from geographic maps. For instance, on the road surface R , we *encourage* the existence of free space F as well as stopped S and moving V objects. Analogically, building information B fosters mass transfer to C . Lastly, T denotes the intermediate area, e.g. pavements, where mobile and stationary objects as well as small urban infrastructure can be present. Note that neither buildings nor roads are present, so we exclude existence of mapped infrastructure C , but we cannot omit other classes. Also, we define a level of con-

fidence β for each map source, possibly different for each context. Let $\tilde{x} = \frac{x_- + x_+}{2}$, $\tilde{y} = \frac{y_- + y_+}{2}$.

$$\begin{aligned}
m_{PG}\{X, Y\}(B) &= \begin{cases} \beta_B & \text{if } (\tilde{x}, \tilde{y}) \in b_i \\ 0 & \text{otherwise} \end{cases} & \forall i \in [0, n_B] \\
m_{PG}\{X, Y\}(R) &= \begin{cases} \beta_R & \text{if } (\tilde{x}, \tilde{y}) \in r_i \\ 0 & \text{otherwise} \end{cases} & \forall i \in [0, n_R] \\
m_{PG}\{X, Y\}(T) &= \begin{cases} 0 & \text{if } (\tilde{x}, \tilde{y}) \in b_i \vee (\tilde{x}, \tilde{y}) \in r_j \\ \beta_T & \text{otherwise} \end{cases} & \forall i \in [0, n_B], \forall j \in [0, n_R] \\
m_{PG}\{X, Y\}(\Omega) &= \begin{cases} 1 - \beta_B & \text{if } (\tilde{x}, \tilde{y}) \in b_i \\ 1 - \beta_R & \text{if } (\tilde{x}, \tilde{y}) \in r_i \\ 1 - \beta_T & \text{otherwise} \end{cases} & \forall i \in [0, n_B], \forall j \in [0, n_R] \\
m_{PG}\{X, Y\}(A) &= 0 & \forall A \subsetneq \Omega \text{ and } A \notin \{B, R, T\}
\end{aligned} \tag{1}$$

2.3 Incorporating prior knowledge

The frame of discernment Ω_{SG} used in SG is distinct from Ω_{MG} , so in order to enable the fusion of SG and MG we define a refining $r_{SG} : 2^{\Omega_{SG}} \rightarrow 2^{\Omega_{MG}}$ such that $r_{SG}(\{F\}) = \{F\}$, $r_{SG}(\{O\}) = \{C, N, S, V\}$, $r_{SG}(A) = \bigcup_{\theta \in A} r_{SG}(\theta)$. The refined mass function can be expressed as $m_{SG}^{\Omega_{MG}}(r_{SG}(A)) = m_{SG}^{\Omega_{SG}}(A)$, $\forall A \subseteq \Omega_{SG}$. Then, Dempster's rule is applied in order to exploit the prior information included in PriorGrid:

$$m_{SG,t}^{\Omega_{MG}} = m_{SG,t}^{\Omega_{MG}} \oplus m_{PG}^{\Omega_{MG}} \tag{2}$$

2.4 Temporal fusion

Computing conflict masses

We use the idea from [5] to distinguish between two types of conflict, which arise from the fact that the environment is dynamic. We denote \emptyset_{FO} the conflict induced when a free cell in MG is fused with an occupied cell in SG. Similarly, \emptyset_{OF} indicates the conflicted caused by an occupied cell in MG fused with a free cell in SG. In an error-free case, these conflicts represent, respectively, the disappearance and the appearance of an object. Conflict masses are calculated using the formu-

las: $m_{MG,t}(\emptyset_{OF}) = m_{MG,t-1}(O) \cdot m_{SG,t}(F)$, $m_{MG,t}(\emptyset_{FO}) = m_{MG,t-1}(F) \cdot m_{SG,t}(O)$,
 where $m(O) = \sum_A m(A)$, $\forall A \subseteq \{C, N, S, V\}$.

MapGrid specialisation using a counter

Mobile object detection is an important issue in dynamic environments. We propose the introduction of a counter ζ in each cell in order to include temporal information on the cell occupancy. For this purpose, incrementation and decrementation steps $\delta_{inc} \in [0, 1]$, $\delta_{dec} \in [0, 1]$, as well as threshold values γ_O , γ_\emptyset have been defined.

$$\begin{aligned} \zeta^{(t)} &= \min\left(1, \zeta^{(t-1)} + \delta_{inc}\right) && \text{if } m_{MG}(O) \geq \gamma_O \text{ and } m_{MG}(\emptyset_{FO}) + m_{MG}(\emptyset_{OF}) \leq \gamma_\emptyset \\ \zeta^{(t)} &= \max\left(0, \zeta^{(t-1)} - \delta_{dec}\right) && \text{if } m_{MG}(\emptyset_{FO}) + m_{MG}(\emptyset_{OF}) > \gamma_\emptyset \end{aligned}$$

Otherwise $\zeta^{(t)}$ rests unchanged. Using ζ values, we impose a specialisation of mass functions in MG using the equation:

$$m'_{MG,t}(A) = S(A, B) \cdot m_{MG,t}(B) \quad (3)$$

where specialisation matrix $S(\cdot, \cdot)$ is defined as:

$$\begin{aligned} S(A \setminus \{V\}, A) &= \zeta && \forall A \subseteq \Omega_{MG} \text{ and } \{V\} \in A \\ S(A, A) &= 1 - \zeta && \forall A \subseteq \Omega_{MG} \text{ and } \{V\} \in A \\ S(A, A) &= 1 && \forall A \subseteq \Omega_{MG} \text{ and } \{V\} \notin A \\ S(\cdot, \cdot) &= 0 && \text{otherwise} \end{aligned} \quad (4)$$

Fusion rule

An important part of the method consists in fusing a discounted and specialized MG (see section 2.5 and preceding paragraph) with a SG combined with prior knowledge (see section 2.3).

$$m_{MG,t} = {}^\alpha m'_{MG,t-1} \otimes m'_{SG,t} \quad (5)$$

The fusion rule \otimes is a modified Yager's rule [10] adapted to mobile object detection. There are of course many different rules that could be used, but in order to distinguish between moving and stationary objects some modifications had to be included. These modifications consist in transferring the mass corresponding to a newly appeared object \emptyset_{FO} to the class of moving objects V as described by the equation 6. Symbol \odot denotes the conjunctive fusion rule.

$$\begin{aligned}
(m_1 \otimes m_2)(A) &= (m_1 \oplus m_2)(A) && \forall A \subseteq \Omega \wedge A \neq V \\
(m_1 \otimes m_2)(V) &= (m_1 \oplus m_2)(V) + (m_1 \oplus m_2)(\emptyset_{FO}) \\
(m_1 \otimes m_2)(\Omega) &= (m_1 \oplus m_2)(\Omega) + (m_1 \oplus m_2)(\emptyset_{OF}) \\
(m_1 \otimes m_2)(\emptyset_{FO}) &= 0 \\
(m_1 \otimes m_2)(\emptyset_{OF}) &= 0
\end{aligned} \tag{6}$$

All the above steps allow us to construct a `MapGrid` containing reach information on the environment state, including the knowledge on mobile and static objects.

2.5 Contextual discounting

Information discounting allows to *forget* information which is no longer valid. Discounting parameter α serves to model the speed with which information becomes obsolete. Thanks to the contextual discounting [3], we make use of more detailed information regarding the confidence we have in the source in various contexts. We noticed that different pieces of information become obsolete with different speed. Hence, the coarsening used is $\Theta = \{\theta_{static}, \theta_{dynamic}, \theta_{free}\}$, with $\theta_{static} = \{C, N\}$, $\theta_{dynamic} = \{S, V\}$, $\theta_{free} = \{F\}$, and discount rates $\alpha = \{\alpha_{static}, \alpha_{dynamic}, \alpha_{free}\}$. We assign higher discount rates (lower confidence) to rapidly changing contexts such as free space, stopped and moving objects, and lower rates to the static context. The discounted mass function is obtained by the disjunctive combination of the input mass function m_{MG} and mass functions for each element of the partition Θ .

$$\alpha m_{MG,t} = m_{MG,t} \cup m_{static} \cup m_{dynamic} \cup m_{free} \tag{7}$$

where each mass function m_l ($l = static, dynamic, free$) is defined by $m_l(\theta_l) = \alpha_l$, $m_l(\emptyset) = 1 - \alpha_l$, $m_l(A) = 0$, $\forall A \subseteq \Omega \wedge A \notin \{\emptyset, \theta_l\}$.

3 Results

3.1 Setup

The data set used for our experiments was acquired in cooperation with IGN in Paris. The overall length of the trajectory was about 3 km. 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, but they can be learnt from data, as proposed in [3]. We have defined the map confidence factor β by ourselves, but ideally, it should be given by the map provider. β describes data currentness

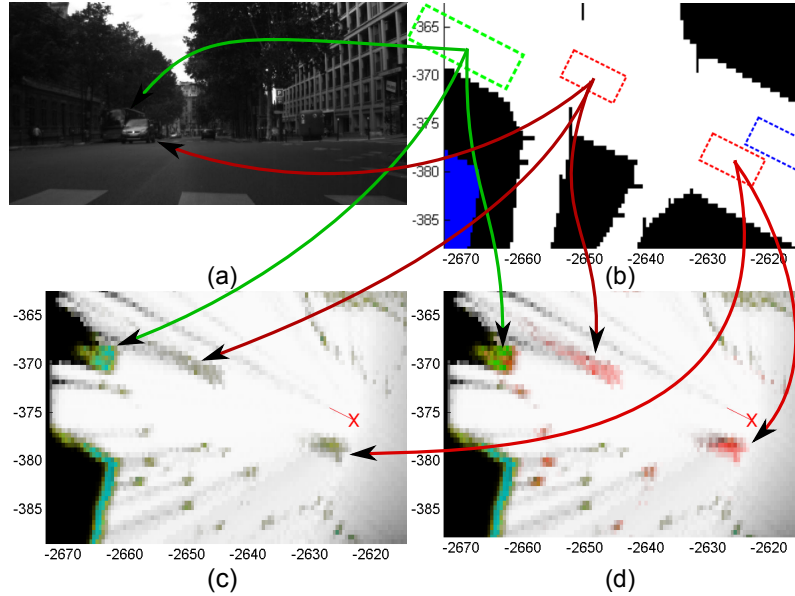


Fig. 2 (a) Scene. (b) PG. (c) MG without prior information. (d) MG with prior map knowledge.

(age), errors introduced by geometry simplification and spatial discretisation. β can also be used to depict the localisation accuracy. Other parameters, such as counter steps δ_{inc} , δ_{dec} and thresholds γ_O , γ_0 used for mobile object detection determine the sensitiveness of mobile object detection and were set by manual tuning.

3.2 Impact of prior knowledge

The results for a particular instant of the approach tested on real-world data are presented on figure 2. The visualisation of the MG has been obtained by calculating the pignistic probability of each class [9]. The presented scene contains two cars (only one is visible in the camera image) 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 test vehicle position is shown as a blue box. Different classes of Ω_{MG} are represented by different colours: F – white, C , N – blue, S – green and V – red. PG on figure 2(b) shows the position of the road space (white) and buildings (blue).

The principal advantage gained by using map knowledge is richer information on the detected objects. A clear difference between a moving object (red, car) and a stopped one (green, bus) is visible. Also, stopped objects are distinct from infrastructure when prior map information is available (cf. figures 2(c) and 2(d). In addition, thanks to the prior knowledge, stationary objects (cyan) such as infras-

structure are distinguished from stopped objects on the road. Grids make noticeable the effect of discounting, as information on the environment behind the vehicle is being forgotten. On the other hand, the parked bus is still in evidence despite being occluded by the passing car.

4 Conclusion and perspectives

A new mobile perception scheme based on prior map knowledge has been introduced. Geographic information is exploited to reduce the number of possible hypotheses delivered by an exteroceptive source. A modified fusion rule taking into account the existence of mobile objects has been defined. Furthermore, the variation in information lifetime has been modelled by the introduction of contextual discounting. In the future, we anticipate removing the hypothesis that the map is accurate. This approach will entail considerable work on creating appropriate error models for the data source. Moreover, we envision differentiating the free space class into two complementary classes to distinguish navigable and non-navigable space. This will be a step towards the use of our approach in autonomous navigation. Another perspective is the use of reference data to validate the results, choose the most appropriate fusion rule and learn algorithm parameters. We envision using map information to predict object movements. It rests also a future work to exploit fully the 3D map information.

Acknowledgements This work has been supported by ANR (French National Agency) CityVIP project under grant ANR-07_TSFA-013-01.

References

1. Cappelle C. et al.: Virtual 3D City Model for Navigation in Urban Areas. In: *J. Intell. Robot. Syst.*, Springer (2011)
2. Elfes, A.: Using Occupancy Grids for Mobile Robot Perception and Navigation. In: *Computer*, 22(6), pp. 46–57 (1989)
3. Mercier, D., Quost, B., Denoeux, T.: Refined modeling of sensor reliability in the belief function framework using contextual discounting. *J. Inf. Fusion*, 9(2), pp. 246–258 (2008)
4. Moras, J., Cherfaoui, V., Bonnifait, P.: Credibilist Occupancy Grids for Vehicle Perception in Dynamic Environments. *IEEE Int. Conf. Robot. Autom.*, pp. 84–89 (2011)
5. Moras, J., Cherfaoui, V., Bonnifait, P.: Moving Objects Detection by Conflict Analysis in Evidential Grids. *Int. Veh. Symp.*, pp. 1120–1125, Baden-Baden, Germany (2011)
6. OpenStreetMap project. <http://www.openstreetmap.org>. (Cited 9 Nov 2011)
7. Pagac, D., Nebot, E. M., Durrant-Whyte, H.: An evidential approach to map-building for autonomous vehicles. In: *IEEE Trans. Robot. Autom.*, 14(4), pp. 623–629 (1998)
8. Shafer, G.: *A Mathematical Theory of Evidence*. Princeton University Press (1976)
9. Smets, P.: Decision making in the tbm : the necessity of the pignistic transformation. *Int. J. Approx. Reason.*, 38(2) pp. 133–147 (2005)
10. Yager, R.R.: On the Dempster-Shafer framework and new combination rules. *Information sciences*, 4 pp. 93–138 (1987)