

# Improved dynamic object detection within evidential grids framework

Abderraouf Hadj Henni, Angel Soriano, Rafael Lopez, Nacim Ramdani

▶ **To cite this version:**

Abderraouf Hadj Henni, Angel Soriano, Rafael Lopez, Nacim Ramdani. Improved dynamic object detection within evidential grids framework. 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), Aug 2019, Vancouver, Canada. pp.1080-1086, 10.1109/COASE.2019.8843016 . hal-02317994

**HAL Id: hal-02317994**

**https://cnrs.hal.science/hal-02317994**

Submitted on 29 Oct 2019

**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.

# Improved dynamic object detection within evidential grids framework

Abderraouf HADJ HENNI[1], Angel SORIANO[2], Rafael LOPEZ[2], Nacim RAMDANI[1]

***Abstract*— The deployment of autonomous robots/vehicles is increasing in several domains. To perform tasks properly, a robot must have a good perception about its environment while detecting dynamic obstacles. Recently, evidential grids have attracted more interest for environment perception since they permit more effective uncertainty handling. The latest studies on evidential grids relied on the use of thresholds for information management *e.g.* the use of a threshold, for the conflict characterized by the mass of empty set, in order to detect dynamic objects. Nevertheless, the mass of empty set alone is not consistent in some cases. Also, the thresholds used were chosen either arbitrary or tuned manually without any computational method. In this paper, first the conflict is composed of two parameters instead of mass of empty set alone, and dynamic objects detection is performed using a threshold on the evolution of this conflict pair. Secondly, the paper introduces a general threshold along with a mathematical demonstration to compute it which can be used in different dynamic object detection cases. A real-time experiment is performed using the RB1-BASE robot equipped with a RGB-D camera and a laser scanner.**

## I. Introduction

Mobile robots are more and more used to accomplish tasks autonomously even in environment with distinct moving obstacles. Such robots can perform tasks in outdoor environment *e.g.* autonomous vehicles, or in indoor environment for logistic tasks in warehouses, for health-care in hospitals, and for user's comfort in smart-buildings. To accomplish the missions safely, the robot must perceive its environment properly in order to avoid collision with the surrounding static and moving obstacles. Occupancy grid [1] is a common technique used for environment perception where the environment area is represented by a grid of several cells. Each cell represents a portion of the navigation area for which we need to estimate its state. The latter can be estimated by computing an occupancy percentage obtained by combining several sources of information using Bayesian methods [1], fuzzy methods [2] or evidential methods [3]. In recent years, evidential methods have received more interests due to the advantages of evidence theory (also known as Dempster-Shafer Theory DST [4]) especially for handling uncertainty, ignorance, and for modeling conflict between sources.

*Moras et al* used evidential grids in several vehicle environment perception works [5] [6] [7] where mobile objects detection relied on analyzing a conflict based on the mass of the empty set and obtained from temporal fusion. Unfortunately, such conflict may also occur even if there is no moving object in case where faulty data from Lidar scanner are provided [8]. Therefore, relying on conflict measurement with only one sensing modality to detect moving obstacles is not consistent since such conflict can be obtained if the sensor is faulty. Consequently, adding other sensing modalities is essential for consistent conflict analysis in dynamic objects detection.

Indeed, recent evidential grid studies have raised the interest of multi-mode sensing for environment perception applications [9] [10] [11]. In [9] authors evoked that in [5] [6] the laser scanner is used only to increase accuracy of the navigation without taking into account the mission of the robot, therefore, they have considered a multi-modal sensing (RGB-D and audio) to better consider the robot's mission. In [10] authors used a stereo-camera to enhance the laser's detection highlighting that mapping is not the only robot's/vehicle's task but other complementary tasks could be considered using evidential grids. Nevertheless, either in multi-sensing evidential grids works [9] [10] or in *Moras et al* [5] [6] mono-sensing works, the dynamic/static object distinction relied on thresholds which are taken arbitrary or set by manual tuning and no demonstration of evaluating them was provided.

In this paper, the conflict is composed of a pair of parameters which are the mass of empty set and the distance between betting commitment as defined in [12]. We introduce then the notions of the evolution of the conflict pair and the evolution of the mass of empty set, and we show through an example that the evolution of the conflict pair can discern between some cases where the evolution of mass of empty set alone cannot. The dynamic objects are detected by analyzing the evolution of the conflict pair w.r.t to a threshold supported by a mathematical proof. The evolution exploits the conflict pair obtained between two sensing modalities before temporal fusion, and the conflict pair obtained after temporal fusion.

The paper is organized as follows. The concept of evidential grids is explained in II and the construction of these grids is illustrated in III. Section IV details the new conflict based environment perception algorithm. The real-time experiments are shown in V while section VI concludes this paper.

## II. Belief functions and evidential grids

### A. Belief functions and evidence theory

*Overview:* Evidence theory [4] (or Dempster Shafer Theory DST) is a theory that uses belief functions to model the

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 823887. A partial support of the first author was provided by the French CoCaPs project issued from FUI N° 20.

[1]Univ. Orleans, INSA-CVL, PRISME, EA 4229, F45072, Orleans, FRANCE. (abderraouf.hadj-henni, nacim.ramdani)@univ-orleans.fr

[2]R&D Department Robotnik Automation, SLL 46988 Valencia, SPAIN (asoriano, rlopez)@robotnik.es

knowledge about a hypothesis and/or uncertain hypothesis (*i.e.* union of hypothesis). This theory was further extended by Smets [13] within the transferable belief model TBM framework. The main common aspects between evidence theory DST and TBM are the frame of discernment (FoD) $\Omega = \{h_1, h_2, ..., h_n\}$ which contains a set of hypothesis, and the corresponding power set $2^\Omega = \{\phi, h_1, h_2, ..., h_n, \{h_1 \cup h_2\}, ..., \Omega\}$ which contains uncertainties and ignorance (*i.e.* union of hypotheses and $\Omega$ *respectively*), in addition to the initial single hypotheses of $\Omega$. Finally, the mass function $m(A)$ where $A \in 2^\Omega$ represents the belief over an element of the power set $2^\Omega$. The mass functions take values in $[0, 1]$ and should satisfy $\sum_{A \in 2^\Omega} m(A) = 1$.

*Fusion:* The TBM offers several combination rules in addition to the DS combination rule of evidence theory. Nevertheless, in this paper we will focus only on the DS rule. Two mass functions from different sources can be combined using the DS rule as follows:

$$m_{1\oplus 2}(A) = m_1(A) \oplus m_2(A) = \frac{m_{1\cap 2}(A)}{1 - m_{1\cap 2}(\phi)} \quad (1)$$

With: $m_{1\oplus 2}(\phi) = 0$. $m_{1\cap 2}(\phi) = \sum_{A\cap B=\phi} m_1(A) \times m_2(B)$

$m_{1\cap 2}(A) = \sum_{B\cap C=A\neq\phi} m_1(C) \times m_2(B)$. $\forall A, B, C \in 2^\Omega/\phi$

*Discounting:* The discounting allows to weight an information source w.r.t to other source(s). Shafer's discounting [4] consists in reducing the mass functions of the corresponding source (except the mass over $\Omega$) as follows:

$$\begin{cases} {}^\alpha m(A) = (1-\alpha) \times m(A) \quad \forall A \in 2^\Omega/\Omega. \\ {}^\alpha m(\Omega) = (1-\alpha) \times m(\Omega) + \alpha \end{cases} \quad (2)$$

Where ${}^\alpha m$ is the discounted mass and $\alpha$ is the discounting factor and $(1 - \alpha)$ is the confidence in the source.

*Pignistic transform:* In contrary to evidence theory, within the TBM framework we can construct a probability function from mass functions using the pignistic transform as follows:

$$BetP(D) = \frac{1}{1 - m(\phi)} \sum_{D \in C \subseteq \Omega} \frac{m(C)}{\mid C \mid} \quad (3)$$

Where: $\mid C \mid$ is the cardinal of $C$ *i.e.* is the number of elements in $C$. $D$ is a *singleton* i.e. $\mid D \mid = 1$.

*Distance between betting commitments and conflict pair:*

In evidence theory DST, the mass of empty set $m_{1\cap 2}(\phi)$ is considered as a measure of conflict. Nevertheless, several studies *e.g.* [12] have shown that this mass of empty set alone cannot represent a consistent measure of conflict in some cases. Therefore, the author of [12] presented then an alternative measure of the conflict by using a pair of parameters as follows:

$$\langle m_{1\cap 2}(\phi), DifBetP \rangle \quad (4)$$

Where $DifBet$ is the distance between betting commitments with: $DifBet = max(|BetP_1(D) - BetP_2(D)|)$.

$D$ is a singleton of $\Omega$. $BetP_1$ and $BetP_2$ correspond to the pignistic probability from source 1 and source 2 respectively.

Note that in our case, the pair (4) is not used to represent only the conflict, but it is further used to get a more consistent analysis of the *evolution* of conflict.

### B. Evidential grid

An occupancy grid consists in dividing the environment perception into several cells. The state of each cell can be represented by a set of hypothesis *e.g.* free or occupied. Nevertheless, when the state of a cell is not certain, it can be represented by an uncertain state corresponding to the union of several hypothesis *e.g.* free $\cup$ occupied thanks to the evidence theory DST or its extensions such as TBM.

Hence, in evidential grid each cell can be represented by the following FoD: $\Omega = \{F, O\}$ leading to the power set $2^\Omega = \{\phi, F, O, \Omega\}$. Where $F, O$ and $\Omega$ corresponds to free, occupied and unknown respectively. A mass function can be assigned then to each element of $2^\Omega$.

## III. Perception and prediction evidential grids

In our case, two sensing modalities are used which are a laser scanner and a RGB-D camera. After building an evidential grid from each sensor, the two grids are combined to get a perception grid. This latter is combined further with the prediction grid during temporal fusion.

### A. From sensors data to evidential grids

The RB1-base robot we use is compatible with the robot operating system ROS. Therefore, we exploit some existing ROS packages to get an occupancy grid map from each sensor (we use $gmapping$[1] to get the map from the laser scanner and both $pointcloud_to_laserscan$[2] and $gmapping$ to get map from the RGB-D camera). The parameters of $gmapping$ are the same for both sensors in order to obtain identical grid size and resolution.

Each occupancy grid map is converted to an evidential grid as follows:

For each cell, construct the following mass functions:

$$\begin{array}{ll} {}^\alpha m_{s_i}(D) = 1 - \alpha_{s_i} & {}^\alpha m_{s_i}(\overline{D}) = 0 \\ {}^\alpha m_{s_i}(\phi) = 0 & {}^\alpha m_{s_i}(\Omega) = \alpha_{s_i} \end{array} \quad (5)$$

$$\text{With: } D = \begin{cases} O & \text{if cell is occupied.} \\ F & \text{if cell is free.} \end{cases}$$

$s_i$ represents the considered sensor (*i.e.* $s_1$ or $s_2$) and $\alpha_{s_i}$ is the discounting on the considered sensor.

### B. Perception grid

The perception grid is obtained by fusing the two sensor's evidential grids, cell by cell, using the DS combination rule:

$$m_{perc}(A) = {}^\alpha m_{s1}(A) \oplus {}^\alpha m_{s2}(A). \forall A \in 2^\Omega/\phi \quad (6)$$

$$m_{perc}(\phi) = m_{s1\oplus s2}(\phi) = 0$$

[1] $http://wiki.ros.org/gmapping$

[2] $http://wiki.ros.org/pointcloud_to_laserscan$

A perception conflict pair is computed then to asses the disagreement between the two sensors:

$$\langle m_{s1\cap s2}(\phi), DifBetP_{perc}(D)\rangle. D \in \{F, O\} \quad (7)$$
$$DifBetP_{perc} = max(|BetP_{s1}(D) - BetP_{s2}(D)|).$$

## IV. Our environment perception algorithm

### A. Temporal fusion

As in previous related works [5] [9] [10], no evolution model of a cell is considered. Therefore, the grid map obtained at $t-1$ is discounted using (2) with $\alpha_{pred} = 0.2$ in order to get the prediction $m_{pred}$ as follows:

$$\begin{cases} m_{pred}(D) = (1-\alpha_{pred}) \times m_{t-1}(D) \quad \forall D \in \Omega. \\ m_{pred}(\Omega) = (1-\alpha_{pred}) \times m_{t-1}(\Omega) + \alpha_{pred} \\ m_{pred}(\overline{D}) = 0. \end{cases} \quad (8)$$

Where $m_{t-1}$ is the a posteriori mass function $m_{a_post}$ defined in (9) and obtained at the previous step $t-1$. $m_{pred}(\overline{D}) = 0$ since if the state of the cell was $D$ at $t-1$ then $m_{t-1}(\overline{D}) = 0$.

A temporal fusion is performed then by merging the perception grid with the prediction grid, cell by cell, using DS combination rule leading to the following *a posteriori*:

$$m_{a_post}(A) = m_{perc}(A) \oplus m_{pred}(A). \forall A \in 2^{\Omega}/\phi \quad (9)$$
$$m_{a_post}(\phi) = m_{perc\oplus pred}(\phi) = 0$$

Again, a conflict pair is computed to asses the conflict obtained between the perception and the prediction during the temporal fusion as follow:

$$\langle m_{temp}(\phi), DifBetP_{a_post}\rangle \quad (10)$$

Where $|D| = 1$ and:

$$m_{temp}(\phi) = m_{perc\cap pred}(\phi).$$
$$DifBetP_{a_post} = max(|BetP_{perc}(D) - BetP_{pred}(D)|).$$

### B. Evolution of the conflict pair

In order to detect dynamic objects, we analyze the evolution of the conflict pair. Indeed, when an object enters (*resp* leaves) a $cell_i$ at time $t$, and when sensors agree on the same state occupied (*resp* free), hence, the perception conflict pair (7) will be the smallest. Nevertheless, if the $cell_i$ was free (*resp* occupied) a $t-1$, the *a posteriori* conflict pair (9) will be the highest. In such a case, the evolution of the conflict pair will be significant. Therefore, we define the evolution of the conflict pair:

$$Evol(\langle m(\phi), DifBetP\rangle) =$$
$$\langle m_{s1\cap s2}(\phi) - m_{temp}(\phi), DifBetP_{perc} - DifBetP_{a_post}\rangle \quad (11)$$

This evolution is further compared to a threshold in order to detect dynamic objects.

### C. Environment perception and dynamic object detection

The environment perception consists in choosing for each cell a state $A \in 2^{\Omega}/\phi$ after the temporal fusion. This choice depends on the quality of the temporal fusion along with the obtained mass functions.

Indeed, if the perception and prediction disagree during temporal fusion, then we should drop the a posteriori mass functions. In such a case, two possible ways are derived: 1) the sensors agree on same state $A$, then we trust the perception. 2) the sensors do not agree on $A$ then we do not take any risk and we put $m_{a_post}(\Omega) = 1$.

Now, we have to fix the three thresholds corresponding to the following: the threshold for conflict between perception and prediction $(\varepsilon_1, \varepsilon_2)$, the threshold for agreement between the two sensors $(\eta_1, \eta_2)$, and finally the threshold on the evolution of the conflict pair from which we detect dynamic obstacles $(\delta_1, \delta_2)$.

For the threshold $(\varepsilon_1, \varepsilon_2)$, we use the computational approach presented in [14] to define the value of *a posteriori* conflict from which we drop the temporal fusion results. We consider that the *a posteriori* conflict occurs when $m_{perc}(D) \geqslant 0.5$ and $m_{pred}(\overline{D}) \geqslant 0.5$. According to (5), (6) and (8) we get for $D \in \{F, O\}$:

$$m_{perc}(D) \geqslant 0.5 \Rightarrow m_{perc}(\overline{D}) = 0 \text{ and } m_{perc}(\Omega) \leqslant 0.5$$
$$m_{pred}(\overline{D}) \geqslant 0.5 \Rightarrow m_{pred}(D) = 0 \text{ and } m_{pred}(\Omega) \leqslant 0.5$$

This leads to $(\varepsilon_1, \varepsilon_2) = (0.25, 0.25)$ according to computation method of [14].

Following a similar reasoning, we define the threshold $(\eta_1, \eta_2)$ representing the highest value that perception conflict pair is allowed to take when sensors are considered agree. The computation led to $(\eta_1, \eta_2) = (0, 0.25)$ as shown in (14) in *appendix VI*.

The issue now is how to define the threshold $(\delta_1, \delta_2)$ on the evolution of the conflict pair since the computation approach in [14] concerns rather the conflict pair and not the evolution of conflict pair. Therefore, a novel mathematical proof is provided in *appendix VI* that illustrates how we have obtained $(\delta_1, \delta_2) = (-0.375, -0.625 + \frac{\alpha_{pred}}{2})$. In our case, since $\alpha_{pred} = 0.2$, the threshold for conflict evolution become then $(\delta_1, \delta_2) = (-0.375, -0.525)$.

Finally, when a dynamic object is detected, it can impact the considered cell in two ways by either rending the cell recently occupied or recently free. The whole environment perception algorithm is summarized in Algorithm 1.

### D. Interest of the conflict pair evolution and its threshold:

*Interest of the conflict pair:*

| case | Sensor 1 | Sensor 2 | Pred | Evol($m_\phi$, DifBet) |
|---|---|---|---|---|
| I | m(O)= 0.95<br>m(Ω)= 0.05 | m(Ω) = 1 | m(F)= 0.8<br>m(Ω)= 0.2 | ( - 0.76, **- 0.4** ) |
| II | m(O)= 0.95<br>m(Ω)= 0.05 | m(O)= 0.95<br>m(Ω)= 0.05 | m(F)= 0.8<br>m(Ω)= 0.2 | ( - 0.79, **- 0.89** ) |

TABLE I: Interest of the pair $Evol(m(\phi), DifBet)$

**Algorithm 1** Environ. perception & dynamic object detection

**For each cell of the grid do:**
**Construct:** $m_{s1}(A)$, $m_{s2}(A)$, $m_{pred}(A)$. $\forall A \in 2^{\Omega}/\phi$
**1)** Fuse the sensors evidential grids and get:
$m_{perc}(A)$, $m_{s1\cap s2}(\phi)$, $DifBet_{perc}$. $\forall A \in 2^{\Omega}/\phi$
**2)** Perform temporal fusion and get:
$m_{a_post}(A)$, $m_{temp}(\phi)$, $DifBet_{a_post}$. $\forall A \in 2^{\Omega}/\phi$
**3)** Compare the *a post* conflict to the threshold $(\varepsilon_1, \varepsilon_2)$ :
**if** $(m_{temp}(\phi), DifBet_{a_post}) < (\varepsilon_1, \varepsilon_2)$ **then**
Keep: $m_{a_post}(A) = m_{Perc}(A) \oplus m_{Pred}(A)$
**else**
**if** $(m_{perc}(\phi), DifBet_{perc}) \leq (\eta_1, \eta_2)$ **then**
$m_{a_post}(A) = m_{Perc}(A)$
**else**
$m_{a_post}(\Omega) = 1$
**end if**
**4)** Deduce the state of the cell:
**if** $m_{a_post}(D) > 0.5$ **then**
State of the cell is $D$ with $D \in \{O, F\}$
$m_{a_post}(D) = 1$
**if** $Evol(m(\phi), DifBet) \leq (\delta_1, \delta_2)$ **then**
Dyn object $\Rightarrow$ cell is *recently* $D$. $D \in \{O, F\}$
**endif**
**else**
State of the cell is unknown *i.e.* $m_{a_post}(\Omega) = 1$
**end if**

The advantage of using the evolution of the conflict pair instead of the evolution of $m(\phi)$ alone is because that the latter may not discern between some cases. Indeed, let's consider the example with the two cases shown in TableI.

We can see that in case I, the sensors do not agree since $m_{s1}(O) > 0.5$ but $m_{s2}(O) = 0 < 0.5$ in contrary to case II where sensors agree. Since a dynamic obstacle is considered when both sensors agree and when the prediction disagree with perception, hence, only case II corresponds to a dynamic obstacle case while case I corresponds to an uncertain case.

If we rely only on the evolution of $m(\phi)$ alone, there will be no difference between the two cases I and II since in both cases $Evol(m(\phi)) < (\delta_1 = -0.37)$. However, when we consider the evolution of the pair we have $Evol(\langle m(\phi), DifBetP\rangle) < (\delta_1, \delta_2)$ only in the case II. Therefore, the two cases I and II are distinguished only when using the evolution of the conflict pair, and for such a reason we should use the evolution of the pair instead of the evolution of $m(\phi)$ alone.

*Interest of the threshold* $(\delta_1, \delta_2)$*:*

Note that the threshold on the evolution of conflict pair $(\delta_1, \delta_2)$ is based on the two following conditions: a) both sensors agree *i.e.* $m_{s1}(D) > 0.5$ et $m_{s2}(D) > 0.5$. And b) the perception and the prediction disagree *i.e.* $m_{perc}(D) > 0.5$ and $m_{Pred}(\overline{D}) > 0.5$ (as shown in step 4 of the *appendix*) where $D \in \{O, F\}$ and $s1$, $s2$ correspond to sensor 1, sensor 2 respectively. Hence, one can say that we can use these two conditions directly in order to detect the dynamic obstacles.

Unfortunately, relying on these two conditions is not a good manner for some applications. Indeed, if the robot has to detect different moving obstacles $O_i$ with $\Theta = \{O_1, ..., O_n, F\}$ (*e.g.* detect dynamic obstacles with $\Theta = \{Human, Robot, Other, F\}$ using two different vision sensors $s1$ and $s2$). In such case, one have to: a) cite all cases when sensors agree *i.e.* all cases when $m_{s1}(H) > 0.5$, $m_{s2}(H) > 0.5$ with $H = O_i$ or $F$. And b) cite all cases when the perception disagree with prediction *i.e.* cases of $m_{perc}(F) > 0.5$, $m_{pred}(O_i) > 0.5$ and cases of $m_{perc}(O_i) > 0.5$, $m_{pred}(F) > 0.5$ for all $O_i \in \Theta/F$ in order to detect the moving objects. Therefore, it is clear that relying on the two conditions become exponentially complex w.r.t the cardinal of $\Theta$. Moreover, the conditions have to be adapted when the number of considered obstacles $O_i$ change.

Fortunately, the threshold on the evolution of the conflict pair $(\delta_1, \delta_2)$ can be used with any number of obstacles by comparing the conflict pair evolution $Evol(m(\phi), DifBet)$ of each obstacle w.r.t to the same threshold $(\delta_1, \delta_2)$. Therefore, using this threshold is more easier and more general for different dynamic objects detection applications.

## V. Real-time experiments

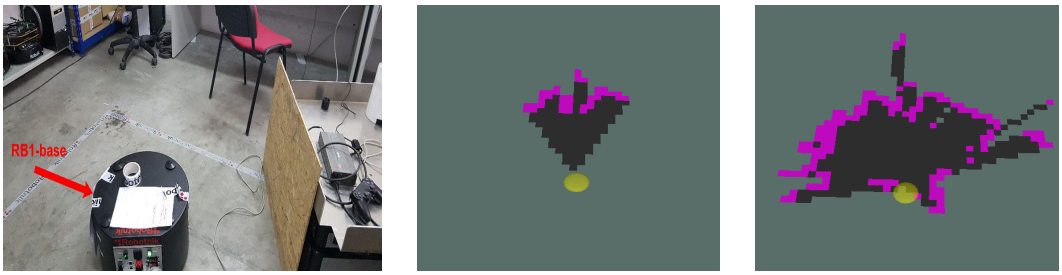
We should note that since our algorithm relies on temporal fusion, the evidential grids obtained at $t-1$ and at time $t$ must be in the same coordinate framework. This condition can be satisfied relying on one of the two following hypotheses. The first hypothesis assumes that the robot's position and orientation at time $t$ are well estimated using existing Simultaneous Localization and Mapping (SLAM) algorithms as assumed in [10] or by using precise positioning systems (*e.g.* Applanix data as in [5]). The second hypothesis consists in performing tests with non-moving robot as in [9]. Since we are rather focused on environment perception and grids fusion in this paper, we perform the tests without moving the robot as in [9]. Nevertheless, the algorithm could be used with a moving robot if the grids at $t-1$ and at time $t$ are adjusted to the same framework relying on a good robot's localization.

### A. Setup:

In order to test our algorithm, we have implemented it on the RB1-base[3] robot using a ROS package. The RB1-base shown in *Fig.1.a)* is equipped with the HOKUYO UTM-30lx laser scanner and the Orbbec Astra RGB-D camera. The ROS package contains two gmapping launch files to get the grid maps from the laser data and from the RGB-D camera's converted data, along with a third file containing our algorithm. Both sensor grid maps have same size of $160\times 160$ cells and resolution of $0.15(m)$/cell's edge. *Fig.1.(a)* shows the RB1-base robot and its surrounding environment.

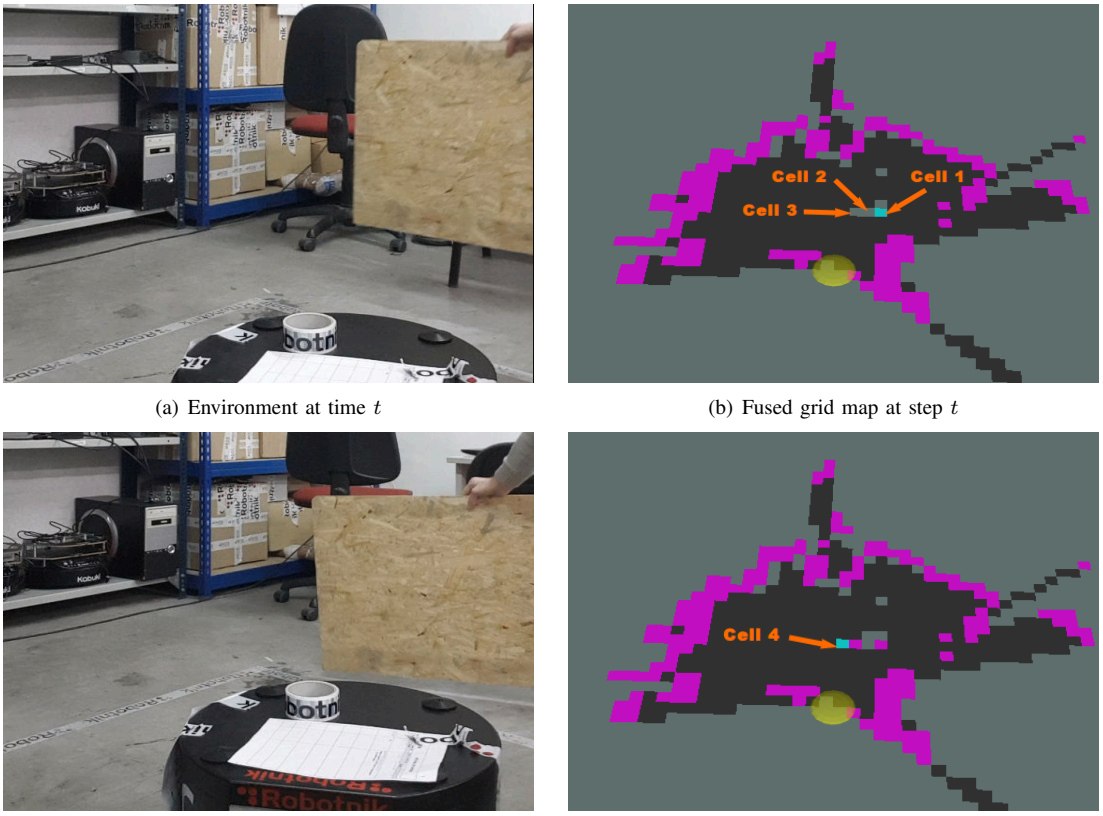
The HOKUYO laser has a horizontal Field of View FoV of $270^{o}$ (a part of this FoV is masked in our robot) while the horizontal FoV of the RGB-D camera is about $60^{o}$. Nevertheless, note that in addition to the horizontal FoV,

[3]https://www.robotnik.eu/logistics/portfolio/rb-1-base/




(a) RB1-base and its environment (b) Grid map from the RGB-D (c) Grid map from the laser

Fig. 1: Robot's environment and the corresponding grid maps from each sensor. Grey, purple and black colors represent unknown, occupied and free respectively. The yellow circle represents the landmark *i.e.* the robot.




(a) Environment at time $t$ (b) Fused grid map at step $t$

(c) Environment at time $t+1$ (d) Fused grid map at step $t+1$

Fig. 2: Experiment results. Blue cyan color corresponds to recently occupied cell by moving object.

the RGB-D camera has a vertical FoV which can reach $50^{o}$ which is not the case for the laser scanner. Consequently, the RGB-D camera can detect objects which are not within the laser's vertical FoV.

We discount both sensors with $0.05$ *i.e.* $\alpha_{si} = 0.05$ in (5) since we consider that both sensors are highly and equally reliable. The prediction is discounted with $0.2$ *i.e.* $\alpha_{pred} = 0.2$ in (8). Note that the values of the discounting factors are taken arbitrary, and may change depending on the reliability of each source, in contrary to the values of the thresholds on the conflict which are computed with a mathematical proof.

### B. Scenario, results, and discussion

*Scenario:* For the experiment scenario, we have considered a moving person carrying a wooden panel. The advantage of the latter is that when it is carried, it will be visible only by the RGB-D and not by the laser scanner since the panel's position will be out of the laser's vertical FoV which makes the panel a different moving object compared to the moving person. At time $t$, the moving person is in the FoV of both sensors (precisely at the left limit of the RGB-D's FoV), however, the wooden panel is only visible for the RGB-D since its position is higher than the vertical FoV of the laser which corresponds to the case of *Fig.2.(a)*. At the next step $t+1$, the person stopped moving and started approaching

the wooden panel to the ground which made the panel in the FoV of both sensors as shown in *Fig.2.(c)*.

*Results:* From *Fig.2.(b)*, we can see that at time $t$, the moving person is detected in cell 1 which is colored with blue *cyan i.e.* recently occupied since both sensors have detected the entrance of the person inside cell 1. Nevertheless, the neighbor cells (cell 2 and 3) are colored with *grey i.e.* unknown because the wooden panel is not within the laser's FoV which led to disagreement between sensors. Therefore, cells 2 and 3 are *grey* instead of blue *cyan*, nevertheless, these two cells are not navigable since they are not free.

From *Fig.2.(d)*, we can see that at $t+1$ the cell 1 has became *purple* (*i.e.* occupied instead of recently occupied) since it was already occupied by the person at time $t$. Also, the cell 3 has became *purple* since the wooden panel is now within the FoV of both sensors, however, it is not considered recently occupied since it was not free at the previous time $t$. For cell 2, we notice that it stayed *grey* because it was not occupied yet in the laser's map which may be caused by a *gmapping* limit due to low laser intensity in this cell 2 at $t+1$. Finally, we can see that a new cell (cell 4) has became blue *cyan i.e.* recently occupied because when the person bent to approach the panel to the ground, she moved the panel to the left and the panel entered cell 4 leading to a recently detected object by both sensors in cell 4.

*Discussion:* Based on these results, it is clear that the thresholds used in this paper are consistent since several cases have been distinguished. For example, we could discern between cell 1 and cells 2 and 3 at time $t$ thanks to the perception conflict threshold $(\eta_1, \eta_2)$. Moreover, we could identify the moving object *i.e.* recently occupied cell and already occupied cell (*e.g.* the case of cell 1 and cell 4 at time $t+1$) thanks to the threshold on the evolution of conflict pair $(\delta_1, \delta_2)$. Keep in mind that at time $t+1$, both cell 1 and cell 4 are not navigable since none of them is *black i.e.* free, nevertheless, differentiating the state of these cells may help the robot to make decisions depending on its mission.

Finally, note that the prediction was discounted more than the perception since the prediction is obtained by a simple discounting of the previous results. Nevertheless, a more consistent prediction can be inspired from the one used in [15] if the motion model of the dynamic obstacles is known.

## VI. Conclusion

In this paper, we have shown the advantage of using a pair of parameters to represent the conflict since the evolution of this pair can discern between some cases where the evolution of empty set alone cannot. Also, we have defined a threshold on the evolution of the conflict pair for mobile objects detection. This threshold is computed mathematically instead of choosing it arbitrary or tuning it manually. Moreover, we have emphasized that this threshold could be more easier and general for different dynamic object detection applications. Finally, a real-time experiment showed the consistency of using the threshold on the evolution of the conflict pair and the presented algorithm for dynamic object detection.

## Appendix

This appendix shows how the threshold on the evolution of the conflict pair is computed trough the 6 steps below

1) Let us consider the following mass functions obtained from the sensors $s_1$, $s_2$ for all $D \in \{F, O\}$ as follows:

$$\begin{array}{lll} m_{s1}(D) = a_1. & m_{s1}(\overline{D}) = b_1. & m_{s1}(\Omega) = 1 - a_1 - b_1 \\ m_{s2}(D) = a_2. & m_{s2}(\overline{D}) = b_2. & m_{s2}(\Omega) = 1 - a_2 - b_2 \end{array}$$

We consider that two sensors agree on a state $D$ if $m_{s1}(D) > 0.5$ and $m_{s2}(D) > 0.5$ which leads to: $a_1, a_2 \in [0.5, 1] \Rightarrow b_1, b_2 < 0.5$.

Note that the sensors can provide only one state for each cell as shown in (5). Hence, if we consider that both sensors agree on $D$, the mass over $\overline{D}$ will be null which leads to the following mass functions:

$$\begin{array}{lll} m_{s1}(D) = a_1 & m_{s1}(\overline{D}) = 0. & m_{s1}(\Omega) = 1 - a_1 \\ m_{s2}(D) = a_2. & m_{s2}(\overline{D}) = 0. & m_{s2}(\Omega) = 1 - a_2 \end{array}$$

Hence, it is clear that when sensors agree we get:

$$m_{s1 \cap s2}(\phi) = 0 \quad \forall\, a_1, a_2 \in [0.5, 1] \tag{12}$$

2) Now let us compute the $DifBet_{Perc}$. We have:

$$\begin{array}{l} Bet_{s1}(D) = a_1 + \frac{1-a_1}{2} = \frac{1+a_1}{2} \\ Bet_{s2}(D) = a_2 + \frac{1-a_2}{2} = \frac{1+a_2}{2} \end{array}$$

$DifBet_{Perc} = max|Bet_{s1}(D) - Bet_{s2}(D)| = |\frac{a_1 - a_2}{2}| \quad \forall\, a_1, a_2 \in [0.5, 1]$. $\Rightarrow$ It is clear that:

$$0 \leq DifBet_{perc} \leq 0.25 \tag{13}$$

Since: $max(DifBet_{perc}) = |\frac{max(a_1) - min(a_2)}{2}| = |\frac{min(a_1) - max(a_2)}{2}| = \frac{0.5}{2} = 0.25$. This leads to:

$$(0,0) \leq \underbrace{(m_{s1 \cap s2}(\phi), DifBet_{perc})}_{(\eta_1, \eta_2)} \leq (0, 0.25) \tag{14}$$

3) Also, when the sensors agree we have:

$m_{perc}(D) = m_{s1 \oplus s2}(D) = (a_1 \times a_2) + (a_1 \times m_{s2}(\Omega)) + (a_2 \times m_{s1}(\Omega)) = a_1(a_2 + m_{s2}(\Omega)) + (a_2 \times m_{s1}(\Omega)) = a_1 + a_2 - a_1 a_2$. Note that: $\frac{\partial m_{perc}(D)}{\partial a_1} = 1 - a_2 \geq 0$ and $\frac{\partial m(D)}{\partial a_2} = 1 - a_1 \geq 0 \ \forall a_1, a_2 \in [0.5, 1]$. $\Rightarrow m_{perc}(A)$ is increasing in sense of $a_1$ and $a_2$. Hence, $min(m_{perc}(D)) = m_{perc}(D)_{a_1 = 0.5 = a_2} = 0.75$ and $max(m_{perc}(D)) = m_{perc}(D)_{a_1 = 1 = a_2} = 1$. We get then:

$$0.75 \leq m_{perc}(D) \leq 1 \tag{15}$$

Moreover, $m_{perc}(\Omega) = m_{s1 \oplus s2}(\Omega) = (1 - a_1) \times (1 - a_2)$. $\Rightarrow Bet_{perc}(D) = m_{perc}(D) + \frac{m_{perc}(\Omega)}{2} = a_1 + a_2 - a_1 a_2 + \frac{(1-a_1) \times (1-a_2)}{2} = \frac{a_1 + a_2 - a_1 a_2 + 1}{2} = \frac{m_{perc}(D) + 1}{2}$.

From (15) we have: $0.75 \leq m_{perc}(D) \leq 1 \Rightarrow \frac{0.75+1}{2} \leq \frac{m_{perc}(D)+1}{2} = Bet_{perc}(D) \leq \frac{1+1}{2}$. Consequently:

$$0.875 \leq Bet_{perc}(D) \leq 1 \tag{16}$$

4) The prediction is obtained by discounting the previous results (*i.e.* previous cell's state) with $\alpha_{pred} \in ]0, 1[$ .

If the cell was $\overline{D}$ at $t-1$, the prediction will be then: $m_{pred}(\overline{D}) = 1-\alpha_{pred}$ and $m_{pred}(\Omega) = \alpha_{pred}$

The conflict between prediction and perception is considered when perception supports a state $D$ and when the prediction supports the opposite state $\overline{D}$ *i.e.* when:

$$m_{perc}(D) > 0.5 \quad \text{and} \quad m_{pred}(\overline{D}) > 0.5 \; D \in \{F, O\}.$$

Indeed, a state is considered when its mass is geater than 0.5 since as evoked in [7], this ensures to chose the most likely level without considering the part of unknown.

However, from (15) we have: $0.75 \leq m_{perc}(D) \leq 1$. Hence, the conflict between perception and prediction can be considered in our case when:

$$0.75 \leq m_{perc}(D) \leq 1 \text{ and } 0.5 < m_{pred}(\overline{D}) \leq 1-\alpha_{pred}$$

Since: $m_{temp}(\phi) = m_{perc \cap Pred}(\phi)$. We obtain:

$$0.375 < m_{temp}(\phi) \leq 1-\alpha_{pred} \tag{17}$$

5) Also, since in such a case we have $(m_{pred}(\overline{D}) = 1-\alpha_{pred}) > 0.5$, we get then:

$$m_{pred}(D) = 0 \text{ and } m_{pred}(\Omega) = \alpha_{pred} < 0.5.$$

This implies that $\alpha_{pred} \in ]0, 0.5[$ in order to consider that prediction is in conflict with the *agreed* sensors. Therefore: $Bet_{pred}(D) = m_{pred}(D) + \frac{m_{pred}(\Omega)}{2} = 0 + \frac{\alpha_{pred}}{2} = \frac{\alpha_{pred}}{2}$.

$$\Rightarrow Bet_{pred}(D) = \frac{\alpha_{pred}}{2} \quad \text{with } \alpha_{pred} \in ]0, 0.5[ \tag{18}$$

From (16) and (18) we conclude that when perception is in conflict with the agreed sensors for, $DifBet_{a_post} = max|Bet_{perc}(D) - Bet_{pred}(D)|$ become:

$$(0.875 - \frac{\alpha_{pred}}{2}) \leq DifBet_{a_post}(D) \leq (1 - \frac{\alpha_{pred}}{2}) \tag{19}$$

From (17) and (19) we get:

$$(0.375,\, 0.875 - \frac{\alpha_{pred}}{2}) \leq (m_{temp}(\phi), DifBet(D)_{a_post}) \leq (1-\alpha_{pred},\, 1 - \frac{\alpha_{pred}}{2}) \tag{20}$$

6) Now, since we have from (14) the range of the perception conflict $(m_{s1\cap s2}(\phi), DifBet_{perc})$ when the sensors agree, and we have from (20) the range of $(m_{temp}(\phi), DifBet_{a_post})$ when prediction is in conflict with the agreed sensors, hence, let's find ranges of $Evol(\, m_{\oplus}(\phi), DifBet\,)$ for this case. The evolution of the conflict pair corresponds to the evolution of each parameter of the pair *i.e.* $Evol(\, m_{\oplus}(\phi), DifBet(D)\,) = (\, Evol(m(\phi)),\, Evol(DifBet)\,)$ where:

$$Evol(m(\phi)) = m_{s1\cap s2}(\phi) - m_{temp}(\phi).$$
$$Evol(DifBet) = DifBet_{perc} - DifBet_{a_post}$$

From (12) and (17) the $Evol(m(\phi))$ is then:

$$-(1-\alpha_{pred}) \leq Evol(m(\phi)) \leq -0.375 \tag{21}$$

Also, we have from (13) and (19) :

$0 \leq DifBet_{perc} \leq 0.25$ and $(0.875 - \frac{\alpha_{pred}}{2}) \leq DifBet_{a_post} \leq (1 - \frac{\alpha_{pred}}{2})$ with $\alpha_{pred} \in ]0, 0.5[$.

This implies $max(Evol(DifBet)) = -0.625 + \frac{\alpha_{pred}}{2}$ and $min(Evol(DifBet)) = 0 - (1 - \frac{\alpha_{pred}}{2}) = \frac{\alpha_{pred}}{2} - 1$ .

$$\Rightarrow \frac{\alpha_{pred}}{2} - 1 \leq Evol(DifBet) \leq -0.625 + \frac{\alpha_{pred}}{2} \tag{22}$$

Consequently, from (21) and (22) we obtain:

$$(\alpha_{pred} - 1, \frac{\alpha_{pred}}{2} - 1) \leq Evol(m(\phi), DifBet) \leq (-0.375, -0.625 + \frac{\alpha_{pred}}{2}) \tag{23}$$

The range in (23) corresponds to the values that $Evol(m_{\oplus}(\phi), DifBet)$ can have when sensors agree on a state $D$ and when prediction is in conflict by supporting the opposite state $\overline{D}$. Consequently, $(-0.375, -0.625 + \frac{\alpha_{pred}}{2})$ represents the threshold for dynamic obstacle detection.

## REFERENCES


[1] A. Elfes. Using occupancy grids for mobile robot perception and navigation. *Computer*, 22(6):46–57, jun 1989.

[2] G. Oriolo, G. Ulivi, and M. Vendittelli. Fuzzy maps: A new tool for mobile robot perception and planning. *Journal of Robotic Systems*, 14:179–197, 1997.

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

[4] Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[5] J. Moras, V. Cherfaoui, and P. Bonnifait. Credibilist occupancy grids for vehicle perception in dynamic environments. In *2011 IEEE International Conference on Robotics and Automation*. IEEE, may 2011.

[6] M Kurdej, J Moras, V Cherfaoui, and P Bonnifait. Map-aided fusion using evidential grids for mobile perception in urban environment. In *Belief Functions: Theory and Applications*, pages 343–350, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[7] J. Moras, V. Cherfaoui, and P. Bonnifait. Evidential grids information management in dynamic environments. In *17th International Conference on Information Fusion (FUSION)*, pages 1–7, July 2014.

[8] G. Tanzmeister and D. Wollherr. Evidential grid-based tracking and mapping. *IEEE Transactions on Intelligent Transportation Systems*, 18(6):1454–1467, 2017.

[9] Q Labourey, O Aycard, D Pellerin, M Rombaut, and C Garbay. An evidential filter for indoor navigation of a mobile robot in dynamic environment. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 286–298, Cham, 2016. Springer International Publishing.

[10] M. Valente, C. Joly, and A. de La Fortelle. Fusing laser scanner and stereo camera in evidential grid maps. *CoRR*, abs/1805.10046, 2018.

[11] E Capellier, F Davoine, V Frémont, J Ibañez-Guzmán, and Y Li. Evidential grid mapping, from asynchronous LIDAR scans and RGB images, for autonomous driving. In *21st IEEE Int Conf on Intelligent Transportation Systems*, Maui, Hawaii, United States, November 2018.

[12] W. Liu. Analyzing the degree of conflict among belief functions. *Artificial Intelligence*, 170(11):909 – 924, 2006.

[13] P. Smets and R. Kennes. The transferable belief model. *Artificial Intelligence*, 66:191234, 1994.

[14] A Hadj Henni, D Pasquier, and N Betrouni. A transferable belief model decision support tool over complementary clinical conditions. In *Bioinformatics and Biomedical Engineering, LNCS vol 10814*, pages 409–420, Cham, 2018. Springer.

[15] A. Hadj Henni, R. Ben Bachouch, O. Bennis, and Nacim Ramdani. Enhanced multiplex binary PIR localization using the transferable belief model. *IEEE Sensors Journal*, page In press, 2019. 10.1109/jsen.2019.2918844.