Heteroceptive Sensing for Autonomous Cars

Tiago Pereira Alexandre
tiago.p.alexandre@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

November 2019

Abstract

This work presents a multi-sensor data fusion approach that aims at providing in real-time the traffic sign the vehicle is passing by. A camera-based traffic signs recognition system (TSRS) is fused with a topological map holding the location of traffic signs, in the context of Dempster-Shafer Theory. Using the discounting scheme of Dempster-Shafer, one proposes the computation of information confidence measure by taking into account the reliability of each source. A topological map holding traffic signs is required to be updated to cover traffic signs and road network changes, it is unable to support temporary signs and may be inaccurate under some circumstances. On the other hand, a camera-based TSRS can detect temporary traffic signs, but it also can be very inaccurate, whether due to bad visibility conditions or to other adverse situations. As such, the fusion of information from both sources can provide more accurate traffic signs information. Both the Dempster-Shafer Fusion and the Graph Based Topological Map are implemented, whereas the TSRS is simulated. The obtained results clearly show that the proposed fusion has a greater accuracy than each of the systems as standalone solutions.
Keywords: Dempster-Shafer, Fusion, Traffic Signs Recognition System, Topological Map

1. Introduction

The autonomous navigation process relies on information that comes from sensors and other types of sources, depending on the accurate interpretation of the sensing data to perceive the environment. In order to have a safe navigation, one has to ensure the reliability of this information. Many autonomous car accidents reported are due to faulty sensor data or failure of the algorithms to analyse sensor data accurately [14]. Sensors may be prone to several sources of error, such as environmental noise or sensor's drift. Fusing data from different sensors and when possible cross-checking with information gathered from external sources, allows the vehicle to have a more complete perception of the environment.

The work proposed in this dissertation aims at providing reliable environment information, namely traffic signs, by fusing data from both a TSRS and a topological map. One implements a Graph Based Topological Map holding information about traffic signs, and a Multi-criteria Fusion method based on the Dempster-Shafer Theory. Both the Graph Based Topological Map and a TSRS can be inaccurate under some circumstances, so the use of either system as a standalone solution can lead to the incorrect evaluation of traffic signs. Considering a topological map, many reasons can lead to a wrong selection of a traffic sign, such as vehicle positioning errors, representation errors of the database, modifications of the road network since the realization of the map, road network details not contained in the map, wrong selection of the candidate road during map-matching, and so on. Likewise, considering a TSRS, several factors mainly related to the driving context can lead to incorrect detections/recognitions, such as obstacles covering traffic signs, presence of decals containing traffic signs images, presence of several signs for a given situation (intersections, highway, etc.) [8], the performance of the recognition system itself, and also high speed driving makes the detection difficult because of a reduced measurement distance [8]. Both systems have their own weak and strong spots, allowing one to explore the redundancy and the complementarity of the data provided by them. Thus, the fusion of their data allows to produce more consistent and accurate information than provided by any of these individual systems.

The Dempster-Shafer Theory is not only able to deal with uncertainties and inaccuracies, as it can also deal with ambiguities and conflicts between sources [8], making it ideal to this type of classification problems.

2. Related work

Several fusion approaches that combine data from digital maps with camera based systems have been developed in recent years.
Gehrig *et al.* [6] use information from a digital map to narrow down the search space of the camera-based system when locating traffic signs. Digital maps containing the type and position of traffic signs are used in [7], and a Dempster-Shafer fusion between this information and a camera-based system is implemented to find the most reliable solution (traffic sign). In the system proposed by [1], a digital map and a TSRS provide probabilities for each speed limit, and information from both sources is fused in a Bayesian approach. A Dempster-Shafer fusion of a digital map and a speed signs recognition system is implemented in [8] to improve the accuracy of the vehicle position estimation. Unlike the approaches referred above, the system proposed in [11] is able to explicitly respect environmental situations influencing the sensors, while fusing digital maps and a speed signs recognition system Also, a parallel implementation of a system for traffic sign recognition with digital map fusion on emerging multicore processors and graphics processing units can be found in [12].
It should be noted that most approaches address specifically to speed signs, whereas fusion approaches addressing all types of traffic signs seem somehow poorly explored.

3. Case Study

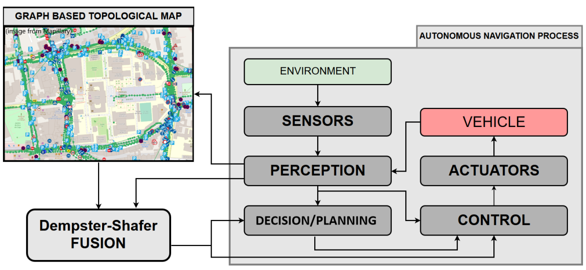


Figure 1: Interaction between the proposed system and the Autonomous Navigation Process.

The Graph Based Topological Map more than a database containing traffic signs information, it is a real-time system capable of matching the vehicle's position to the logical model of the real world (Map structure) and pointing the traffic signs the vehicle is passing by, using only GPS data. The Dempster-Shafer fusion, also known as The Belief Theory, is a reasoning for fusing data from several sources, which is particularly adapted to deal with uncertainties and inaccuracies on the one hand, and ambiguities and conflicts between sources on the other.

For the proposed method, evidence (a mass function m) is assigned to each traffic sign pointed by the TSRS and the Map. Traffic signs pointed by the Map, are assigned with evidence based on criteria regarding both the traffic sign itself and the reliability of the map. For the TSRS evidence assigned to the recognized sign is based on the recognition score and a criterion given by the visibility conditions (which model the reliability of the TSRS). Finally, using Dempster-Shafer Theory, traffic signs pointed by both sources are fused to obtain a reliable decision on the effective traffic sign present ahead on the road, as well as the confidence level in that sign.

3.1. Graph Based Topological Map

One implements a Graph Based Topological Map, which allows the insertion of all sorts of information about the environment, such as traffic signs, road hazards, traffic jams, accidents, crowded zones, and so on. However, given that the target of this thesis is to provide reliable traffic signs information, only traffic signs are inserted.
The Graph Based Topological Map is structured as a simple directed and unweighed graph, where each node represents a road or a segment of it and each edge represents a connection between roads, which can be one-way or two-way. Each node holds a list containing the set of traffic signs present on the corresponding road.
The Graph Based Topological Map is a real-time system capable of matching the vehicle's position to the logical model of the real world, based on GPS coordinates and velocity components. When the system starts running, which may correspond to the start of the car, it has to determine the initial position and direction of the vehicle on the map structure (map-matching), but then while on the same road/node information ahead come sequentially, without the need to be constantly performing map-matching.
Most road networks are made up of similar roads sharing the same characteristics, but some unique ones can also be found. Those unique roads may require personalized encoding. It may be for example the insertion of geographically referenced information, or even creating and adding functionalities to the system.
The implementation of the proposed system is based on the assumption that the driver acts in compliance with Portuguese road legislation.

One implements a map containing only roads from Instituto Superior Técnico campus. Even though the campus has just a few roads, it is a good model to test the system and to prove that it would also work if adapted for every other scenario.

3.1.1. Map Structure


```json
{
    "roads": [
      {
        "road_name" : "IST6",
        "connections" : [
            {
                "road" : "IST3",
                "orientation" : "East"
            }
        ],
        "polygon_coord" : [
            [38.73544601220729, -9.139795877938942]
        ],
        "infos1" : [
            {
                "info" : "end",
                "info_type" : "point",
                "orientation" : "East",
                "coordinates" : [38.736228, -9.139439],
                "radius" : 3,
                "missed" : 0
            }
        ],
        "infos2" : null
      }
    ]
}
```


Figure 2: Illustrative example of a Graph Based Topological Map structure.

Each node of the graph consists of a **road_name**, a vector **connections**, a vector **polygon** and two information lists (**infos1** and **infos2**). The **road_name** is just the label of the road/node. A **connection** is made up by a road/node and the orientation of that connection. For example, if road *IST6* has the connection *[IST3;East]*, it means that somewhere in road *IST6*, the vehicle may turn to road *IST3* by turning East. The first step in map-matching is to find out which road/node the vehicle is on. A **polygon** encloses a given road, and if the current position's coordinates fall within that polygon the vehicle is on that road. Each node has one or two lists of road information (**infos1** and **infos2**), depending if it corresponds to a one-way or two-way road, respectively. Each traffic sign, way-point and other information about the road are saved in these lists, where the sequence of information respects the order in which it is found on the road.
Each road information belongs to a linked list of information. These lists are doubly linked lists, where information nodes contain references to both the next and the previous information nodes. An information node consists of its label **info**, the **info_type**, the **orientation** of the vehicle when it is supposed to detect the corresponding information, references to both the next and previous information, information's geodetic **coordinates**, a **radius** which is the distance the vehicle must be from the information so the system points the next one, and the variable **missed** that saves the number of times that a given information was not detected by the TSRS. An information is of type **sign** when it corresponds to a traffic sign, textbfroad mark if it is a road mark, **intersection** if it is a auxiliary information indicating the vehicle is passing by an intersection, or **point** if it is an auxiliary information used by the system, such as a way-point.

3.1.2. Map Workflow

Here one explains the overall functioning of the Graph Based Topological Map.

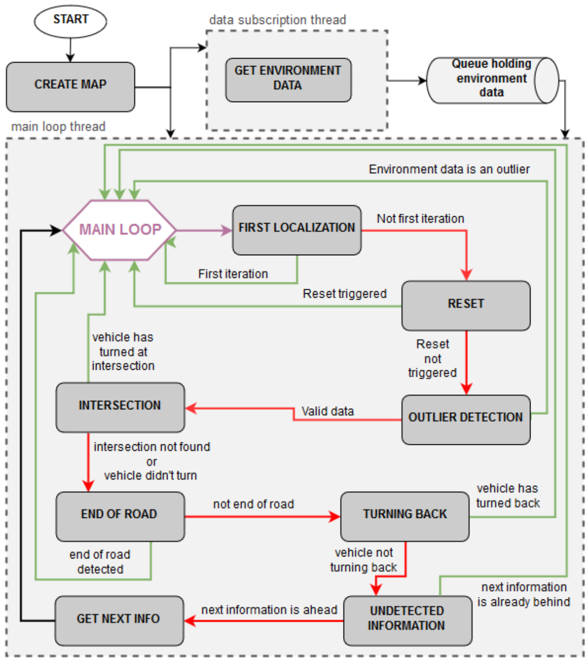


Figure 3: Graph Based Topological Map workflow.

When the system starts running the data needed to construct the Map is uploaded, the graph structure is created, and then two threads start running. Data subscription thread is in charge of subscribing updated GPS data from a publisher and putting it in a queue so the main loop thread can access it.
The main loop thread keeps track of the vehicle path while providing traffic signs information in real-time. Here, in each loop iteration the system gets the most recent GPS data available. In figure (3), each grey block inside the main loop thread is a module that evaluates a specified condition, and if it is true one breaks the current main loop iteration (green arrows on figure (3)), otherwise the next module is executed (red arrows on figure (3)). When the system starts it has no information about the vehicle's location in the Map structure, so the **first localization module** performs the map-matching in order to point the information the vehicle is going to pass by. The **reset module** resets the navigation process whenever state of the vehicle does not match the expected, that is when the distance travelled by the vehicle

towards an information is much greater than expected. The **outliers detection module** consists of a filter to detect outliers, and here a point is perceived as an outlier when it is significantly distanced from its estimated location. When the next information ahead is of type intersection, the **intersection module** detects if the vehicle turns at the intersection, and points the next information regardless of whether the vehicle turns at it. If there are no more information in the current list, it means the vehicle has reached the end of the current road, and the **end of road module** checks which road the vehicle is going to. The **turning back module** detects when a vehicle turns back while on the same road (U-turn), and if that happens it points the new next information ahead. The **undetected information module** detects if the current next information is still ahead or not, by comparing relative distances between road information and the vehicle. The **get next info module** computes the distance between the vehicle and the current next information and if that distance is shorter than the **radius** of the information, this module points the next information ahead.

3.1.3. Results

3.1.3.1 Real data tests

After collecting raw data from a single GPS antenna, corresponding to different routes in Intituto Superior Técnico campus, the performance of the system is evaluated by performing offline tests. The performance of the system depends on the number of detected outliers (outliers), the number of resets performed (resets), the maximum elapsed time between iterations (max_time), the number of times an iteration took longer than supposed (n_iter), and the number of missed traffic signs (missed). Table (1) displays the results of each test.

Table 1: Performance of the proposed system with raw GPS data as input.

test	outliers	resets	missed	max_time	n_iter
1	452	1	0	0.49 seg	10/4102
2	406	1	0	0.38 seg	4/3957
3	114	2	0	0.23 seg	3/2715

The performance of the Topological Map is optimal if the system is able to correctly point all traffic signs the vehicle passes by. For each test performed, every information (traffic sign) was pointed by the system in the correct location, so its performance is optimal for these tests. For the first and second tests, the system only had to reset once. For both tests the system reset while turning from $IST3$ to $IST1$. That happens because of $ITS1$'s wide width. In this specific case that does not cause the system to miss any sign. The main loop is designed so that each iteration takes exactly 0.1 seconds. This time interval is considered appropriate for this specific application. However, sometimes the implemented system takes more than 0.1 seconds to perform an iteration. This only happens for about 0.2% of iterations, and the maximum elapsed time between iterations that one has registered is around 0.5 seconds. For non-high speeds it is not a problem, although for high velocities 0.5 seconds may imply a significant travelled distance and the system may miss a traffic sign if present. Nevertheless, the frequency with which this happens is very low, not compromising the reliability of the proposed system.

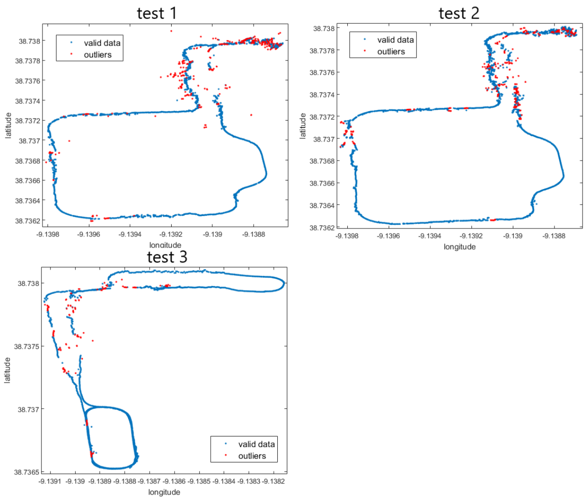


Figure 4: Plots of the raw data used in each test in table 1.

Considering the good performance of the system for the available tests, the proposed Topological Map is reliable. Nevertheless, much more testing (at least hundred of kilometres would have to be covered) would be needed to be possible to draw conclusions with high confidence about the reliability of the proposed system.

3.1.3.2 Performance variation with faulty input

In order to infer about the quality and the risk of failure of the proposed Map, one subjects the proposed system to noisy inputs. A single data set is used for testing the system, and in each test a different amount of noise is added to the original data set. The noise values added to each coordinate of the data set are Gaussian-distributed, and higher values of standard deviation σ imply more noise.
The results of each test are in table (2). The first row, $\sigma = 0$, corresponds to the original raw data from a single GPS antenna. The performance of the system is given by the rate of correct information, while the opposite is given by the rate of wrong

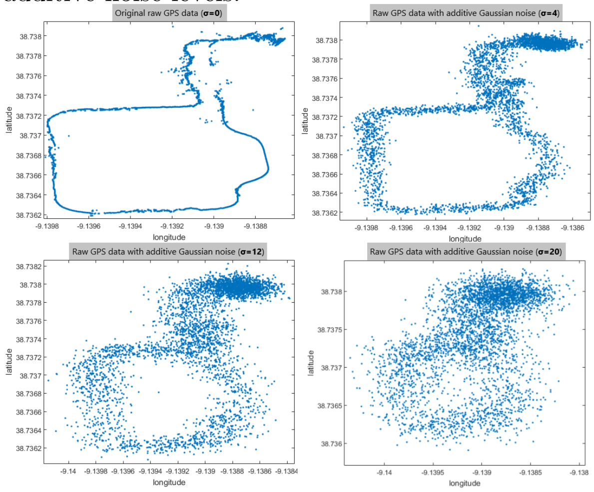
or not given information. For $\sigma = 0$, one assumes a performance of 100% as the system correctly points all information. From the test corresponding to $\sigma = 0$ one acquires the interval of time when an information is supposed to be pointed by the Map system. Then, for the other tests, $\sigma > 0$, one checks whether information is pointed by the Map in the correct time interval or not, by comparing with the results from $\sigma = 0$.

Table 2: System's performance, number of resets performed, numbers of outliers detected and additive noise variation.

Noise	Outliers	Resets	Performance
$\sigma = 0$	474	1	100%
$\sigma = 1$	1520	4	86%
$\sigma = 2$	2700	8	80%
$\sigma = 4$	3150	8	75%
$\sigma = 6$	3330	15	68%
$\sigma = 8$	3275	20	60%
$\sigma = 10$	3370	23	53%
$\sigma = 14$	3330	23	40%
$\sigma = 20$	3325	27	40%
$\sigma = 30$	3220	56	25%
$\sigma = 40$	2900	180	15%
$\sigma = 50$	2100	370	5%
$\sigma = 60$	1570	520	0%

As the amount of additive noise increases the performance of the system also decreases. As data loses quality the system becomes unable to correctly point the traffic signs the vehicle is passing by and the reset is triggered by the incapacity of the system to identify the correct path of the vehicle. So, the number of resets performed by the system also increases when the noise increases. The more noise is added to original data, the more outliers the system detects. For values of standard deviation between $\sigma = 3$ and $\sigma = 30$ the number of detected outliers tends to stabilize because the system discards almost all the position coordinates, although for standard deviations greater than $\sigma = 30$ the number of detected outliers starts to decrease. That happens because when facing large standard deviations, most position coordinates fall outside the polygons enclosing the roads. The system faces this situation most of the time for standard deviations greater than $\sigma = 30$ and most position coordinates are not even evaluated by the outliers detection module. When running the original raw data set corresponding to $\sigma = 0$, the system is able to correctly point all the traffic signs ahead. The reset is only triggered once, while turning from $IST3$ to $IST1$, as explained in the previous subsection. For standard deviations σ greater than 0, the reset is mostly triggered because of faulty data. For standard deviations below $\sigma = 5$, the system is able to correctly point all the traffic signs the vehicle passes by. Nevertheless, the performance is not 100% . As the standard deviation increases the system starts to point more often some traffic signs where/when it is not supposed to.

Figure 5: Plots from different tests, with different additive noise levels.



For standard deviations greater than $\sigma = 10$ the performance of the Graph Based Topological Map stays below 50%, only achieving 0% for standard deviations greater than $\sigma = 50$ (approximately). The Topological Map is supposed to be a reliable source of traffic signs information, and its fusion with a TSRS most overcome the performance of the TSRS as a standalone system. Having said that, the performance of the proposed Map must be close to 100% in order to its deployment to be beneficial. Although the proposed system is quite noise tolerant, the ideal would be for coordinates to be filtered and/or fused with inertial information in order to achieve a better accuracy.

3.2. Dempster-Shafer Fusion of Topological Map and Traffic Signs Recognition System

Data provided by both sources is fused in the context of Dempster-Shafer Theory, also known as Belief Theory.

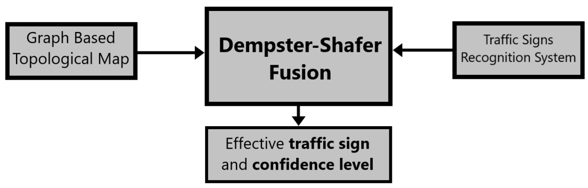


Figure 6: Fusion overview.

The Dempster-Shafer Theory, introduced by Dempster [2] and mathematically formalized by

Shafer [15], can be viewed as a generalization of the Bayesian Theory [5]. It allows the combination of evidence from different sources in order to compute a degree of belief in an information, by taking into account all the available evidence. The degree of belief in an information is given by a mass function m, which may not have the mathematical properties of probabilities. For example, given two mutually exclusive events A and B, $mass(A, B) = mass(A) + mass(B)$ does not have to be true. This theory also allows the evaluation of conflict between sources.

3.2.1. Formal definition

Here, one presents the formal definition of Dempster-Shafer Theory as formalized by Shafer in [15] ([4]).

Belief Theory requires the definition of a state space θ (1) known as the frame of discernment. It is an exhaustive set of mutually exclusive states, composed of all the k possible decisions d to a given problem.

$$\theta = \{d_1, d_2, ..., d_k\} \tag{1}$$

Based on the frame of discernment, it is defined a mass functions referential corresponding to its power set. It is the set of all disjunctions of the decisions d denoted:

$$2^\theta = \{\emptyset, d_1, \cdots, d_k, \{d_1, d_2\}, \{d_{1,3}\}, \cdots, \theta\} \,. \tag{2}$$

A mass function, also known as basic belief assignment function, corresponding to the degree of belief that can be assigned to a given decision, assigns to each value of the power set (2) a mass value. Considering a source S_j, the respective mass function m_j satisfies:

$$m_j : 2^\theta \to [0, 1] \,, \tag{3}$$

$$m_j(\emptyset) = 0 \,, \tag{4}$$

$$\sum_{x \in 2^\theta} m_j(x) = 1 \,, \tag{5}$$

where the mass value of the empty set being equal to zero means that the assumption of a *closed world* is made.

Each proposition x that satisfies $m_j(x) > 0$ is a focal element.

Closely related to mass function, a belief function summing up all evidence that supports x is computed:

$$Bel(x) = \sum_{y \in 2^\theta : y \subseteq x} m(y) \,. \tag{6}$$

It is also computed a function that sums up all evidence that does not disagree with x, the plausibility function:

$$\begin{aligned} Pl(x) &= \sum_{y \in 2^\theta : y \cap x \neq \emptyset} m(y) \\ &= 1 - Bel(\bar{x}) \,. \end{aligned} \tag{7}$$

The belief and plausibility functions values of 1 are assigned to the frame of discernment (1). On the other hand, its mass function can take any value between 0 and 1. The mass function of the frame of discernment can be used to infer about the reliability of a source. The known high reliability of a source in a certain situation entails a low mass value $m(\theta)$.

Belief and plausibility measures allow one to define an interval containing the exact probability of a set belonging to the power set (2):

$$Bel(x) \le P(x) \le Pl(x) \quad , x \in 2^\theta \,. \tag{8}$$

The next step is the combination of the sets of mass assignments from each source. This fusion stage between the sources is based on the Dempster-Shafer conjunctive combination rule, which reinforces the belief on decisions for which the sources are in agreement, and in contrary attenuates it when the sources are in conflict. The normalized representation of the combined mass function is defined as follows:

$$\begin{aligned} & m(x) = \frac{1}{1-K} \sum_{B_1 \cap \cdots \cap B_m = x} \prod_{j=1}^{m} m_j(B_j) \quad , x \neq \emptyset \,, \\ & m(\emptyset) = 0 \,, \end{aligned} \tag{9}$$

where

$$K = \sum_{B_1 \cap \cdots \cap B_m = \emptyset} \prod_{j=1}^{m} m_j(B_j) \tag{10}$$

is a measure of the amount of conflict between sources. If two sources do not contradict, then $K = 0$, but if they completely contradict each other, $K = 1$ and the orthogonal sum (9) does not exist. As K approaches 1, the results from Dempster's Rule are often inaccurate [5].

After the combination, a decision rule must be applied in order to obtain the final decision.

3.2.2. Discounting scheme

The discounting scheme of Mercier *et al.* [9] is used for taking into account the reliability of a source in the mass assignments. The less reliable the source is, the more weakened the belief function is.

Given a discounting rate $\alpha \in [0, 1]$, the quantity $1 - \alpha$ corresponds to the degree of belief in the source's reliability R:

$$
\begin{aligned}
m(R) &= 1-\alpha\,,\\
m(R,NR) &= \alpha\,,
\end{aligned} \tag{11}
$$

where $\{R, NR\}$ is the set of possible values of reliability, with R meaning that the source is reliable and NR meaning that it is not.
All information given by a non-reliable source is then assigned to the ignorance θ. Once applied the discount, the resulting mass function ${}^{\alpha}m_{source}$ is given by:

$$
\begin{aligned}
{}^{\alpha}m_{source}(x) &= (1-\alpha)m_{source}(x) \quad ,x \in 2^{\theta}\,,\\
{}^{\alpha}m_{source}(\theta) &= (1-\alpha)m_{source}(\theta)+\alpha\,.
\end{aligned} \tag{12}
$$

3.2.3. Multi-criteria Fusion

Firstly, data is collected from both sources, the TSRS and Topological Map. The Map provides a set of nearby signs that are assigned with a non-zero mass value, the focal points. The TSRS only provides one focal point which is the traffic sign recognized by the algorithm, alongside with the level of confidence in that recognition, the *score*. The mass values assigned to the focal points from each source are based on criteria C. This concept, was first introduced by [8] and developed by [13].
The frame of discernment θ represents the set of traffic signs which can be determined by the sources. This frame is defined in order to model the traffic signs present in Instituto Superior Técnico campus, assuming these as all the possible existing signs.
Not all disjunctions belonging to the power set 2^{θ} are assigned with evidence. Concerning the traffic signs recognition system, only one focal point is defined, which is the traffic sign recognized by the system. To consider possible inaccuracies of the Map, nearby signs are assigned with evidence. A radius of 25 meters centered at vehicle's position is defined, and all the signs within that radius are considered focal points, even if they do not belong to the current road.
The mass m_{cam} assigned to a focal point depends on two other mass values, m_{TS} which concerns the evidence of the focal point and m_{rel} which concerns the reliability of the source. One extracts from the TSRS directly the sign identified by an image processing algorithm and its confidence level, *score*, related to the performed detection. This confidence level corresponds to the evidence $m_{TS_{cam}}$ assigned to the focal point $x \in 2^{\theta}$, while the lack of evidence $(1 - score)$ is assigned to the ignorance θ:

$$
\begin{aligned}
m_{TS_{cam}}(x) &= score(x)\,,\\
m_{TS_{cam}}(\theta) &= 1-score(x)\,.
\end{aligned} \tag{13}
$$

The reliability of the source (camera) is highly dependent on visibility conditions [11]. The proposed system acquires online information about weather conditions in order to assign values to three criteria C_{vis} used for computing the mass value $m_{rel_{cam}}$. C_{vis1} concerns cloudiness, C_{vis2} concerns the rain volume and C_{vis3} concerns light conditions. The discounting rate α, used in the discounting scheme (3.2.2), is given by a weighting equation (14) applied to criteria C_{vis}. Low visibility conditions imply a high rate α. One finds $k_1 = 0.2$, $k_2 = 0.4$ and $k_3 = 0.4$ empirically.

$$
\alpha = k_1 \times C_{vis1} + k_2 \times C_{vis2} + k_3 \times C_{vis3} \tag{14}
$$

The mass value assigned to the reliability of the source $m_{rel_{cam}}$ ($m(R)$ in (11)), decreases as the visibility also decreases:

$$
m_{rel_{cam}} = 1-\alpha\,. \tag{15}
$$

To obtain the final mass value $m_{cam}(x)$ (${}^{\alpha}m_{source}$ in (12)), the discounting scheme (3.2.2) is applied:

$$
\begin{aligned}
m_{cam}(x) &= m_{TS_{cam}}(x)\times m_{rel_{cam}}\,,\\
m_{cam}(\theta) &= m_{TS_{cam}}(\theta)\times m_{rel_{cam}} + (1-m_{rel_{cam}})\,.
\end{aligned} \tag{16}
$$

The mass m_{map} assigned to a focal point pointed by the Topological Map also depends on m_{TS} and m_{rel}. The definition of seven criteria based on information extracted from the Map allow the assignment of a mass value to each focal point. These criteria are subdivided into two classes, C_{TS} which the system uses for computing $m_{TS_{map}}$ and C_{rel} for computing $m_{rel_{map}}$. The four criteria C_{TS} concern the evidence of a traffic sign. C_{TS1} assigns a value of 1 to the current traffic sign ahead pointed by the Map or to multiple signs if they are side by side or in another configuration that sets equal probabilities of those signs being detected in a given moment by the TSRS. To the other nearby signs lower values are assigned. Here, the greater the distance between a given sign and the vehicle position the lower the assigned value. C_{TS2} assigns a value of 1 to signs belonging to the current road, while signs from nearby roads are assigned with a value of 0.2. C_{TS3} assigns a value of 1 to signs having the same orientation has the vehicle, while the others with a value of 0.2. C_{TS4} assigns lower values to signs with a large **missed** value. These criteria allow the computation of the mass value $m_{TS_{map}}$ assigned to each focal point $x \in 2^{\theta}$:

$$
\begin{aligned}
m_{TS_{map}}(x) =& w_1 \times C_{TS1} + w_2 \times C_{TS2} +\\
& w_3 \times C_{TS3} + w_4 \times C_{TS4}\,,\\
m_{TS_{map}}(\theta) =& 0\,.
\end{aligned} \tag{17}
$$

In equation (17), the choice of each coefficient w is empiric and based on the influence of each criterion in the evaluation, being $w_1 = 0.5$, $w_2 = 0.3$,

$w_3 = 0.1$ and $w_4 = 0.1$. After computing the masses of all signs, they are normalized so their sum add to 1, as imposed by (5). The other three criteria, C_{rel}, concern the reliability of the map itself, that is the trust in the ability of the system to select the correct traffic sign. C_{rel1} concerns the confidence in the positioning tool, and the higher the standard deviation of vehicle coordinates the lower the assigned value. Considering C_{rel2}, when the travelled distance is greater than the distance between consecutive signs, lower values are assigned. C_{rel3} assigns values depending on the number of outliers detected consecutively. These criteria allow the computation of the mass value m_{rel} ($m(R)$ in (11)) assigned to the reliability of the map system:

$$m_{rel_{map}} = w_5 \times C_{rel1} + w_6 \times C_{rel2} + w_7 \times C_{rel3}\,. \quad (18)$$

Again, the choice of each coefficient w is empiric and based on the influence of criterion in the evaluation of the reliability, being $w_5 = 0.35$, $w_6 = 0.55$ and $w_7 = 0.1$. Finally, the discounting scheme (3.2.2) allows to determine the final mass value m_{map} (${}^{\alpha}m_{source}$ in (12)), of each focal point $x \in 2^{\theta}$ of the Map:

$$\begin{aligned} m_{map}(x) &= m_{TS_{map}}(x) \times m_{rel_{map}}\,, \\ m_{map}(\theta) &= 1 - m_{rel_{map}}\,. \end{aligned} \quad (19)$$

After the combination step (9), a decision rule (20) gives the effective traffic sign x_{eff}, based on the approach proposed in [11]. The effective traffic sign is the one with the highest belief (Bel). A hypothesis is considered credible if its confidence level is higher than a given threshold c_{max} and if the conflict between sources is not higher than the threshold $c_{conflict}$.

$$x_{\text{eff}} := \begin{cases} \text{unknown}, & K > c_{conflict} \\ \text{unknown}, & max_{x \in \theta} Bel(x) < c_{max} \\ arg\, max_{x \in \theta} Bel(x), & \text{otherwise} \end{cases} \quad (20)$$

For high conflict situations the proposed fusion may produce counter-intuitive results ([10] cited in [5]). The constants c_{max} and $c_{conflict}$ prevent the possible wrong selection of a traffic sign in case of high conflicts and/or little evidence for a particular sign. In this application, $c_{max} = 0.5$ and $c_{conflict} = 0.65$.

3.2.4. Results

No TSRS is used for performing the tests, instead, several outputs of this system are simulated to recreate different possible situations. Thus, situations when both sources are in conflict (medium and high), when one of the sources is not available, and also a situation when the output of one of the sources (TSRS) does not remain constant, allow to conclude whether the proposed fusion yields better results than each source as a standalone system.
A focal point x is assigned to the recognition performed by the TSRS, and a mass value $m_{TS_{cam}}(x)$ corresponding to the hypothetical *score* of the recognition algorithm is assigned to this focal point x. All the other mass values used in the tests are output by the implemented system.
Even though the system is not tested with a real TSRS and the fact that there is not enough data to conduct a thorough quantitative evaluation, the several tests performed show promising results.

3.2.4.1 Medium conflict

This example presents a situation of medium conflict where the sources disagree, though not entirely. Here, the Map does not assign the highest mass value to the same focal point the TSRS recognizes. The Map system points two focal points, *End of reserved parking* and *Crosswalk*. The Map assigns the highest mass to the *End of reserved parking*, which is the traffic sign closest to the vehicle, while the TSRS only recognizes the sign *Crosswalk*, which is further. Tables (3) and (4) show the results obtained.

Table 3: Mass assignments for the various focal points.

	$m_{TS_{map}}$	$m_{rel_{map}}$	m_{map}	$m_{TS_{cam}}$	$m_{rel_{cam}}$	m_{cam}	m
End of reserved parking	0.66	0.94	0.62	0	0.99	0	**0.14**
Crosswalk	0.34		0.32	0.91		0.90	**0.85**
θ	0		0.06	0.09		0.1	**0.01**

Table 4: Believes and plausibilities associated with Table 3.

	Bel_{map}	Pl_{map}	Bel_{cam}	Pl_{cam}	Bel	Pl
End of reserved parking	0.62	0.68	0	0.1	**0.14**	0.15
Crosswalk	0.32	0.38	0.90	1	**0.85**	0.86
θ	1	1	1	1	**1**	1

The conflict between sources is given by the value $K = 0.56$ which corresponds to medium conflict. The conflict indicator K is lower than the threshold $c_{conflict}$, so for this medium conflict situation the effective traffic sign is known, and it is also given by the focal point x with the highest belief $Bel(x)$. The effective traffic sign is the *Crosswalk* because even though the Map assigns a low mass value to it, the TSRS assigns a mass value of 90% to it, which in turn leads to a high final belief. Also, the final belief in *End of reserved parking* is very low because only the Map assigns a mass value to it. The system outputs the traffic sign *Crosswalk* with a confidence level of 85%.

3.2.4.2 High conflict

This example presents a situation when both sources fully disagree. Here, each source points a different focal point. The Map system points a single focal point, *Stop*. The TSRS, on the other hand, did not recognize the *Stop* sign. Instead, the TSRS wrongly recognized a *Prohibit* sign due a mismatch. Tables (5) and (6) show the results obtained.

Table 5: Mass assignments for the various focal points.

	$m_{TS_{map}}$	$m_{rel_{map}}$	m_{map}	$m_{TS_{cam}}$	$m_{rel_{cam}}$	m_{cam}	m
Prohibit	0		0	0.95		0.79	**0.27**
Stop	1	0.90	0.90	0	0.83	0	**0.65**
θ	0		0.10	0.05		0.21	**0.08**

Table 6: Believes and plausibilities associated with Table 5.

	Bel_{map}	Pl_{map}	Bel_{cam}	Pl_{cam}	Bel	Pl
Prohibit	0	0.10	0.79	1	**0.27**	0.35
Stop	0.9	1	0	0.21	**0.65**	0.73
θ	1	1	1	1	**1**	1

The conflict between sources is given by the value $K = 0.71$. Although the focal point x with the highest belief $Bel(x)$ corresponds to the effective sign, the system output is unknown because the conflict K is greater than $c_{conflict}$.

3.2.4.3 Only one source available

It is assumed that for some reason the Map is unavailable, so no information is extracted from it. In this circumstance all evidence is assigned to ignorance, that is $m_{map}(\theta) = 1$.
When only one source is available, the system falls back to interpret the mass function of the active source, the TSRS, instead of the fused one, because all evidence is assigned to ignorance of the Map. In this type of situation, the conflict is zero, so if the belief (Bel) of the traffic sign with the highest belief (Bel) is greater than $c_{max} = 0.5$, the system points it as the effective traffic sign.

3.2.4.4 Unstable source

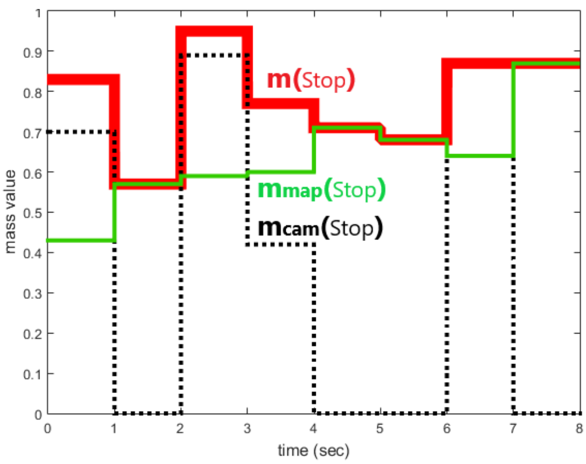


Figure 7: Fused mass m, in a situation where the TSRS output is unstable.

This example is a situation where the output of the TSRS does not remain constant. The TSRS is simulated so that the recognition of a given traffic sign is intermittent. The TSRS recognizes a *stop* sign for a while ($m_{cam}(stop) \neq 0$), then it does not recognize any sign ($m_{cam}(stop) = 0$), and so on. Here, the Map points the *stop* sign uninterruptedly. The result of the proposed fusion remains constant for the 8 seconds and corresponds to the correct traffic sign, as it is possible to check in figure (7). Whenever the TSRS does not recognize the *stop* sign, the fusion falls back to interpret the mass function of the Map instead the TSRS. The proposed method may output a stable result even if one of the sources provides unstable information. In situations like this when the output of the TSRS is unstable, its deployment as a standalone solution may cause the autonomous system to also be unstable.

3.2.4.5 Three sources

Although the proposed fusion only combines two information sources, the Dempster-Shafer theory allows the combinations of more. Dempster's combination rule is commutative and associative, allowing the combination of several sources to be done sequentially instead globally [3]. Here, the order in which sources are combined does not matter [3].
For instance, information from the depth sensor of a Kinect camera can be fused with information from the TSRS and the Map. The Map can contain the side of the road on which a sign is, and the Kinect camera can also detect the location (left or right) of a detected sign. The redundancy of the information from the three sources makes this fusion possible.

4. Conclusions

In this work one proposes a system that fuses information from a topological map and a camera-based traffic sign recognition system, in order to provide more reliable information than any of these systems as standalone solutions.

Firstly, the implementation of the Graph Based Topological Map, which is a real-time system that points the traffic signs nearby while following the path described by the vehicle, is carried out. Then, the fusion process relies on Dempster-Shafer Theory, which takes into account the uncertainties and imprecision of both sources, and also computes the confidence level of the fused information. The mass functions assigned to the each information are adjusted based on the defined criteria such that a source has less influence on the fusion result if it is known to be less reliable.

Tests using real GPS data show the validity of the proposed Graph Based Topological Map. Also, the proposed fusion is able to correctly determine the effective traffic sign in situations where either the Topological Map or the TSRS as standalone systems would fail.

5. Future Work

Regarding the Graph Based Topological Map (3.1), one proposes five modifications. A more robust and advanced map-matching algorithm could be implemented. The outliers detection module perceives a pair of coordinates as an outlier when it is significantly distanced from its estimated location, based on a fixed radius. The use of a dynamic radius whose value would be the greater the worse the quality of the GPS signal, would improve the performance of the proposed module. The Map system should be able to acquire the location of construction sites and other situations that increase the likelihood of temporary traffic signs. The main modification would be the expansion of the information map to support all the desired roads. The manual insertion of road information to the Map structure is an exhaustive task, so an algorithm to automatically create the Map structure containing all the desired information would be a plus. Online map data could be acquired to create and update the map.

Considering the Dempster-Shafer Fusion (3.2), one proposes two improvements. The deployment of an image-based algorithm capable of estimating visibility conditions would improve the estimation of the source's accuracy loss. If the TSRS detects a temporary traffic sign in a zone where temporary traffic signs are most likely to exist, the information provided by the Map should be less valued as the Map cannot provide temporary signs.

References

[1] C. Bahlmann, M. Pellkofer, J. Giebel, and G. Baratoff. Multi-modal speed limit assistants: Combining camera and gps maps. In *2008 IEEE Intelligent Vehicles Symposium*, pages 132 – 137, 07 2008.

[2] A. Dempster. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, 38(2):57–72, 01 1967.

[3] J. Dezert, A. Tchamova, and F. Dambreville. On the mathematical theory of evidence and dempster's rule of combination. *Archive ouverte HAL*, 05 2011.

[4] T. Fine. Review: Glenn shafer, a mathematical theory of evidence. *Bulletin of The American Mathematical Society - BULL AMER MATH SOC*, 83, 01 1977.

[5] B. G. Foley. A dempster-shafer method for multi-sensor fusion. Master's thesis, Wright-Patterson Air Force Base, Ohio - Air Force Institute of Technology, 2012.

[6] S. Gehrig, S. Wagner, and U. Franke. System architecture for an intersection assistant fusing image, map, and gps information. In *IEEE Intelligent Vehicles Symposium, Proceedings*, pages 144 – 149, 07 2003.

[7] H. Janssen and W. Niehsen. Vehicle surround sensing based on information fusion of monocular video and digital map. In *IEEE Intelligent Vehicles Symposium, Proceedings*, pages 244 – 249, 07 2004.

[8] J.-P. Lauffenburger, B. Bradai, M. Basset, and F. Nashashibi. Navigation and speed signs recognition fusion for enhanced vehicle location. In *IFAC Proceedings Volumes*, volume 41, pages 2069 – 2074, 01 2008. 17th IFAC World Congress.

[9] D. Mercier, B. Quost, and T. Denœux. Contextual discounting of belief functions. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, pages 552–562, 07 2005.

[10] R. Murphy. Dempster-shafer theory for sensor fusion in autonomous mobile robots. *IEEE Transactions on Robotics and Automation*, 14:197 – 206, 05 1998.

[11] D. Nienhüser, T. Gumpp, and J. Zöllner. A situation context aware dempster-shafer fusion of digital maps and a road sign recognition system. In *2009 IEEE Intelligent Vehicles Symposium*, pages 1401 – 1406, 07 2009.

[12] K. Par and O. Tosun. Real-time traffic sign recognition with map fusion on multicore/many-core architectures. *Acta Polytechnica Hungarica*, 9, 01 2012.

[13] A.-S. Puthon, F. Nashashibi, and B. Bradai. Improvement of multisensor fusion in speed limit determination by quantifying navigation reliability. In *Conference Record - IEEE Conference on Intelligent Transportation Systems*, pages 855 – 860, 10 2010.

[14] S. Saxena. Reliable navigation for autonomous vehicles in connected vehicle environments by using multi-agent sensor fusion. Master's thesis, School of Computer Science - Carnegie Mellon University, 2017.

[15] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.