Mathematical Modeling of Mass Extraction from Real-Life Measurements for Dempster-Shafer Theory

Vlad-Ion IANCU

School of Engineering and Sciences,
Jacobs University Bremen, Germany
v.iancu@jacobs-university.de

Abstract. From the introduction of the Dempster-Shafer Theory, numerous improvements have been made to it, mostly related to computation time or to the behavior in high conflicting situations. However, the mass extraction from a sensor measurement has remained largely ignored although there are quite a few challenges in this operation. This paper proposes a mathematical modeling for extracting masses in a typical sensor fusion problem.
Keywords: basic assignment probability, mass function, Dempster rule, sigmoid function, Gaussian distribution
Math. Subject Classification 2000: 68T30; 68T37

1 Introduction

The Dempster-Shafer theory is a mathematical theory proposed by Arthur P. Dempster and Glenn Shafer [1], [2] as a generalization of the Bayesian theory of subjective probability. Its main application is in sensor data fusion where 2 distinct operations must be performed: first, extract a degree of belief, commonly referred as mass, from each sensor and second, combine all the data in a single mass assignment. Although the latter has been analyzed and improved, both from an accuracy and from a computation time point of view, in numerous papers [3], the mass extraction from the sensor data has not been studied as thoroughly, even though is equally important.

Consider $\theta_1, \theta_2, ..., \theta_n$ to be the exhaustive and exclusive states under consideration, referred to as elementary hypotheses. Then the set $\Theta = \{\theta_1, \theta_2, ..., \theta_n\}$ is called the frame of discernment.

Definition 1. *A basic belief assignment (BBA), also called mass, is a function*

$$m : 2^{\Theta} \to [0, 1],$$

where 2^{Θ} *stands for the power set of* Θ *, with the following properties:*

$$m(\emptyset) = 0$$

$$\sum_{A \subseteq 2^\Theta} m(A) = 1.$$

It is important not to confuse the notion of mass with that of probability, since they are fundamentally different. The mass of a set A expresses the proportion of evidence that agrees with the fact that the actual state belongs to A, but makes no claims about the mass of a subset of superset of A. Unlike probability, $m(A \cup B)$ is completely independent of $m(A)$ and $m(B)$, as long as

$$m(A \cup B) + m(A) + m(B) \leq 1,$$

otherwise property 2 from Definition 1 would be violated .

From the BBA assignment, a probability interval can be computed:

$$bel(A) \leq P(A) \leq pl(A), \forall A \in 2^\Theta$$

where

$$bel(A) = \sum_{B \in 2^\Theta, B \subseteq A} m(B)$$

is called the *belief* of set A and expresses the proportion of the total data that supports, at least in part, set A, while

$$pl(A) = \sum_{B \in 2^\Theta, B \cap A \neq \emptyset} m(B)$$

is called the *plausibility* of set A and expresses the proportion of the total data that does not directly contradict the set A.

Suppose we have 2 sets of data from 2 different sensors, modeled as 2 BBA assignments $m_1, m_2 : 2^\Theta \rightarrow [0, 1]$. This information can be combined into a new mass m . There were quite a few combination rules developed since the introduction of the Dempster-Shafer Theory but for the purpose of this paper we will use the original rule developed by Dempster, which is still the most popular despite some criticism on how it deals with conflicting information: all conflicting masses are ignored via a normalization factor:

$$m(\emptyset) = 0,$$

$$m(A) = \frac{\sum_{X,Y \in 2^\Theta, X \bigcap Y = A} m_1(X) m_2(Y)}{1 - k_{12}}$$

where

$$k_{12} = \sum_{X,Y \in 2^\Theta, X \cap Y = \emptyset} m_1(X) m_2(Y)$$

is called the *conflict degree.* The greater the conflict degree, the worse the combination formula performs. Notice that for the extreme value $k_{12} = 1$, the combination fails completely, since the denominator becomes 0.

2 Calculating the mass

Now that the theory behind Dempster-Shafer is understood, the sole problem that we have is transforming the sensor data into a mass assignment. We will work with a theoretical model of a robot equipped with an array of sensors. The "world" around the robot is modeled as a square grid of size $2n \times 2n$. The robot is in the center and has negligible dimensions. For each cell, we have 2 elementary hypotheses: θ_1-full (there is an obstacle in the cell) and θ_2-empty. The frame of discernment is $\Theta = \{\theta_1, \theta_2\}$ and the power set $2^{\Theta} = \{\emptyset, \theta_1, \theta_2, \theta_1 \cup \theta_2\}$, but since we deal with the Dempster rule of combination, we will neglect the empty set, since it always has a mass/belief/plausibility of 0 and does not affect the computation.

Suppose we have a measurement of a cell, given by a sensor. We need to model the measurement as a set of 3 masses: $\{m(\theta_1), m(\theta_2), m(\theta_1 \cup \theta_2)\}$, which correspond to cell full, cell empty and cell status unknown, respectively.

2.1 Mass of $\theta_1 \cup \theta_2$

We will start with the mass for the unknown situation, since it's simpler to compute. The sensor's accuracy decreases with distance between the robot and the cell: the mass of $m(\theta_1 \cup \theta_2)$ is close to 0 at small distances and grows to 1 as the cell distance approaches infinity. With this information in mind, we will model $m(\theta_1 \cup \theta_2)$ of a cell using a sigmoid function, more precisely a Gompertz function:

$$f(x) = exp(ln(Z)exp(-Gx)),$$

where Z is the value at 0, while G expresses the growth rate. Both Z and G are small positive numbers, the exact value depending on the characteristics of the sensor. For a cell of vertices

$$\{(x, y), (x + 1, y), (x, y + 1), (x + 1, y + 1)\},$$

we define

$$f(cell_{(x,y)}) = exp(ln(Z)exp(-G\sqrt{(x + 1/2)^2 + (y + 1/2)^2})) \qquad (1)$$

since we consider the distance to a cell, to be the distance to its center (see Fig. 1 and Fig. 2).

2.2 Mass of θ_1 and θ_2

Suppose the sensor measures a cell to be full. There is an incertitude due to the distance and also due to the distance finding method. This is usually done by using 2 identical sensors with a small distance in between and finding the depth of an object by the difference between the 2 sensor's measurements (similar to how human vision works). This is still an area of research and

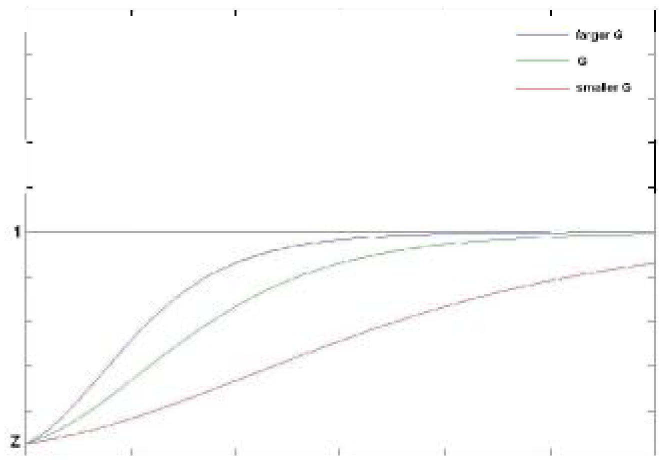

Fig. 1. $m(\theta_1 \cup \theta_2)$ as a function of the cell distance, from different growth parameters

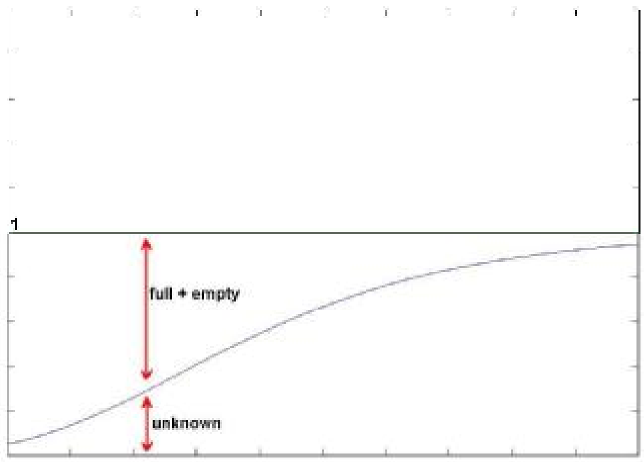

Fig. 2. $m(\theta_1 \cup \theta_2)$ and $m(\theta_1) + m(\theta_2)$

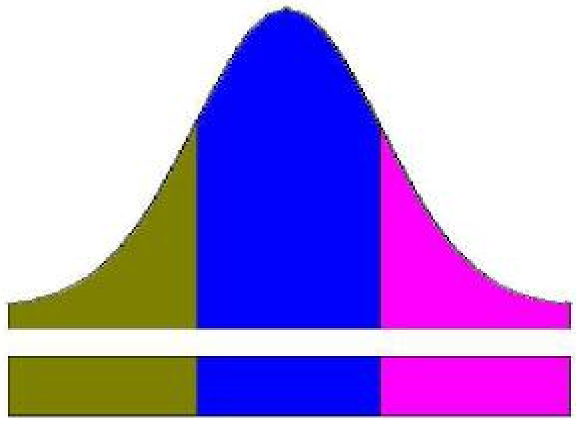

Fig. 3. For each cell is given a mass according to the area under the Gaussian curve

current methods are far from perfect. With this in mind, for a cell that has been measured to be full, we also distribute $m(\theta_1)$ to adjacent cells, using a Gaussian distribution (see Fig. 3).

Notice that we are working on a bi-dimensional grid and we need to rotate the curve around it's mean, which is the center of the cell. The variance will depend on the incertitude (through constant K) , so on the function we used for the unknown mass. Thus, consider a cell $cell_{(x_0,y_0)}$ that has been measured to be full. A cell $cell_{(x_1,y_1)}$ will have a $m(\theta_1)$ due to $cell_{(x_0,y_0)}$ of :

$$m(\theta_1)_{cell(x_1,y_1)} = (1 - f(cell_{(x_0,y_0)}))\frac{\int_{x_1}^{x_1+1}\int_{y_1}^{y_1+1} g(x,y)dxdy}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)dxdy}$$

where

$$g(x,y) = exp(\frac{-((x-x_0)^2+(y-y_0)^2)}{2K^2 exp(2ln(Z)exp(-G\sqrt{(x_0+1/2)^2+(y_0+1/2)^2}))}). \qquad (2)$$

The value $m(\theta_2)_{cell(x_1,y_1)}$ is just the remaining mass up to 1:

$$m(\theta_2)_{cell(x_1,y_1)} = (1 - f(cell_{(x_0,y_0)}))(1 - \frac{\int_{x_1}^{x_1+1}\int_{y_1}^{y_1+1} g(x,y)dxdy}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)dxdy}).$$

Suppose we have a cell $cell(x_0, y_0)$. Two full cells $cell(x_1, y_1)$ and $cell(x_2, y_2)$ will each give it a mass assignment:

$$\{m_1(\theta_1), m_1(\theta_2), m_1(\theta_1 \cup \theta_2)\}$$

and

$$\{m_2(\theta_1), m_2(\theta_2), m_2(\theta_1 \cup \theta_2)\}.$$

What will then be the mass assignment of $cell(x_0, y_0)$? The most reasonable solution is to combine the 2, into a new mass assignment using the Dempster rule of combination. Since the rule is associative and commutative there is no problem in extending this method to any number of full cells. However, this method has a flaw, from a computational point of view. To exemplify, consider the following situation. Suppose , we are working with a 200×200 grid, or 40000 cells. If a quarter of these cells contain an obstacle, this means that we have 10000 gaussian mass distributions and to compute the mass assignment of any of the total 40000 cells we need to combine 10000 measurements, so in total $4 * 10^8$ Dempster combinations. If we consider the fact that the Gaussian function decreases rapidly with the distance from the mean, this means that a full cell will have a significant influence only on the masses of nearby cells. This gives us the idea to "cut" the Gaussian at a threshold:

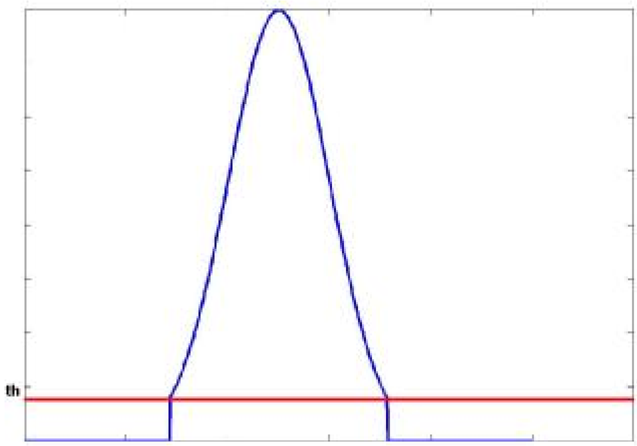

Fig. 4. For values smaller than a threshold we "cut" the gaussian

The solution to

$$\frac{exp(\frac{-(x-\mu)^2}{2\sigma^2})}{\sigma\sqrt{2\pi}} = th$$

is

$$x = \mu \pm \sigma\sqrt{-2ln(\sqrt{2\pi} \cdot th \cdot \sigma)}$$

where th represents the threshold from the Figure 4. Notice that we also need to rescale the function to keep the area under (or volume for our occupancy grid application) equal to 1. So the new modeling function becomes:

$$h(x) = \begin{cases} \frac{h_1(x)}{\int_{\mu-a}^{\mu+a} h_1(y)dy} & if \quad x \in (\mu - a, \mu + a) \\ 0 & otherwise \end{cases}$$

where

$$h_1(x) = exp(-\frac{(x-\mu)^2}{2K^2 exp(2ln(Z)exp(-G\mu))})$$

and

$$a = K \cdot exp(ln(Z)exp(-G\mu))\sqrt{-2ln(\sqrt{2\pi} \cdot th \cdot K \cdot exp(ln(Z)exp(-G\mu)))}$$

The mass assignments of $cell(x_i, y_i)$ due to $cell_(x_0, y_0)$ being measured full, becomes:

$$m(\theta_1)_{cell(x_i,y_i)} = (1 - f(cell_{(x_0,y_0)}))\frac{\int_{x_i}^{x_i+1}\int_{y_i}^{y_i+1} g(x,y)dxdy}{\int\int g(x,y)dxdy} \qquad (3)$$

where $g(x, y)$ is given by (2) and the integral from the denominator is taken over all cells with

$$\sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2} < K \cdot exp(c(x_0, y_0))\sqrt{-2ln(\sqrt{2\pi} \cdot th \cdot K) - 2c(x_0, y_0)}$$

and

$$c(x_0, y_0) = ln(Z)exp(-G\sqrt{(x_0 + 1/2)^2 + (y_0 + 1/2)^2}.$$

Note that $m(\theta_1)_{cell(x_i,y_i)} = 0$ outside the this region.

Again, $m(\theta_2)_{cell(x_i,y_i)}$ is just the remaining mass up to 1:

$$m(\theta_2)_{cell(x_i,y_i)} = (1 - f(cell_{(x_0,y_0)}))(1 - \frac{\int_{x_i}^{x_i+1}\int_{y_i}^{y_i+1} g(x,y)dxdy}{\int\int g(x,y)dxdy}) \qquad (4)$$

with the denominator integral taken over the same region.

3 Algorithm

The previous results can be implemented with the following algorithm.

1. For all sensors do steps 2-4.

2. For all cells set $m(\theta_1 \cup \theta_2)_{cell(x,y)} = f(cell_{(x,y)}))$, with f defined by (1)

3. For all cells $cell(x, y)$ measured to be "full" set the surrounding cells $cell(x_i, y_i)$ with

$$\sqrt{(x - x_i)^2 + (y - y_i)^2} < K \cdot exp(c(x, y))\sqrt{-2ln(\sqrt{2\pi} \cdot th \cdot K) - 2c(x, y)} \qquad (5)$$

and

$$c(x,y) = ln(Z)exp(-G\sqrt{(x+1/2)^2+(y+1/2)^2} :$$

3.1 If the cell $cell(x_i, y_i)$ does not have a mass assignment for θ_1 and θ_2 then the values $m(\theta_1)_{cell(x_i,y_i)}$ and $m(\theta_2)_{cell(x_i,y_i)}$ are computed using the relations (3) and (4), respectively, and the denominator integral is taken over all cells that satisfies the relation (5).

3.2 The value $m(\theta_1 \cup \theta_2)_{cell(x_i,y_i)}$ is kept as it as.

3.3 Otherwise combine the above measurement with the one that is already assigned to the cell using Dempster's rule.

4. For cells that have not been assigned a mass for θ_1 and θ_2: keep $m(\theta_1 \cup \theta_2)$ and set $m(\theta_1) = 0$ and $m(\theta_2) = 1 - m(\theta_1 \cup \theta_2)$.

5. For all cells combine the mass assignments from all sensors using Dempster's rule.

4 Conclusions

A mathematical model of mass extraction from real-life measurements to be used with Dempster-Shafer Theory is presented in this paper.

References

[1] **A. Dempster**, Uper and lower probabilities induced by multivalued mapping, Annals of Mathematical Statistics, 38, 1967, 325-339

[2] **G. Shafer**, A Mathematical Theory of Evidence, Princeton Univ. Press, NJ, 1976

[3] **F. Smarandache, J. Dezert (Editors)**, Advances and Applications of DSmT for Information Fusion, American Research Press, Rehoboth, 2004