

# Analysis and Comparison of Probability Transformations for Fusing Sensors with Uncertain Detection Performance

Edwin El-Mahassni, Samuel Davey and Jonathan Legg

ISR Division  
Defence Science and Technology Organisation  
Edinburgh, SA, 5111, Australia  
{Edwin.El-Mahassni, Samuel.Davey, Jonathan.Legg}@dsto.defence.gov.au

**Abstract.** In a recent paper by Davey, Legg and El-Mahassni a way of fusing sensors with uncertain performance was outlined using the Transferable Belief Model (TBM) theory. It showed that if the target prior was uncertain, then the resulting fused mass was also uncertain. That is, some belief mass was assigned to the case that the presence or absence of a target was unknown. Various methods have been proposed to transform an uncertain belief function into a probability mass. This paper analyses the relationship between an important subset of these methods and compares the resulting probability masses with those obtained via Bayesian methods using random priors.

**Key words:** belief functions, probability transformations, sensor fusion

## 1 Introduction

Sensor fusion is the combination of sensor data obtained from different sources to obtain more complete or accurate information than that available from any individual source. Two or more sensors with complementary capabilities may be used together to achieve a more reliable target state estimate. The scenario addressed by this paper is that of fusing two similarly performing sensors to enhance an operator's confidence in his/her decision making. The most popular methods for combining information from multiple sensors are Bayesian in nature, including networks and Kalman filters. Bayesian networks can make predictions based on a small number of observations [1]. They are models that represent variables and their probabilistic relationships.

A second method which has also attracted attention in the area of sensor fusion is that of Dempster-Shafer Theory (DST) [2, 3]. DST may be thought of as a generalisation of probability theory; rather than assigning probability mass to an exhaustive collection of mutually-exclusive events, DST allows belief mass to be assigned to sets of events when the evidence or data is not able to distinguish between these events. Thus the DST is better suited for representing ambiguity and in particular can represent some measure of total ignorance by

assigning a non-zero mass to the universal set. An advantage of using DST is that it allows for random variables with uncertain probability mass, for example an event that occurs with probability at least  $p$ , or a prior between  $p_1$  and  $p_2$ .

In [4], the Transferable Belief Model (TBM)[6, 7], a variant of DST, was used to fuse two sensors at the decision level (is there a target or not). It was found that uncertainty in the prior probability of a target being present lead to uncertainty in the fused belief mass. In [4] this was resolved by using the pignistic transform. However, alternative methods for transforming uncertain belief functions into probabilities have been proposed. Here, we extend the work of [4] and analyse several probability transformations that satisfy the minimal criterion of [5].

This paper also presents a Bayesian solution to the fusion problem when the sensor parameters are uncertain. This is achieved through treating the parameters as random variables with known priors. The Bayesian solution is compared with the TBM solution. The paper is divided in the following way: in Section 2 we give a short introduction to evidential reasoning and its application to sensor fusion when uncertainty exists. In Section 3 we describe and analyse some of the methods that transform belief masses into probabilities. In Section 4, we derive the Bayesian solution when target, detection and false alarm uncertainty exists. In Section 5, we provide a graphical example comparing the TBM probabilities and the Bayesian probabilities. Finally in Section 6, we provide some concluding remarks.

## 2 Applying Evidential Reasoning to Sensor Fusion

In DST, the *Basic Belief Assignment* (BBA) function (which is sometimes referred to as belief mass),  $m$ , defines a mapping of the power set to the unit interval. Note that  $m(\emptyset) = 0$  is usually assumed for DST. For a given set  $A \in \wp(X)$ , where  $\wp(X)$  is the powerset of a set  $X$ , the value  $m(A)$  represents the proportion of all relevant and available evidence that supports the claim that a particular element of  $X$  belongs to the set  $A$ , but to no particular subset of  $A$  if  $A$  is not a singleton. We remark that conceptually,  $m$  can represent, but is not limited to, probability. The sum of all the masses is always 1.

Under the TBM variant of DST, mass is allowed to be assigned to the empty set,  $m(\emptyset)$ . One interpretation of the mass assigned to the empty set is that it is the degree of conflict between combined BBAs. The TBM combination rule does not contain the normalising term customary for DST and is given by

$$m_{1,2}(C) = \sum_{A \cap B = C} m_1(A) m_2(B). \quad (1)$$

The TBM was used in [4] to fuse detection level data from two sensors and derive the belief that a target is present or absent given the sensor outputs and uncertain knowledge of the sensor performance.

In [4], it was assumed that, for sensor  $i$ , the probability of detection, false alarm and prior probability of target being present are respectively given and

bounded by

$$d_{\text{MIN}}^i \leq P_{\text{D}}^i \leq d_{\text{MAX}}^i, \quad f_{\text{MIN}}^i \leq P_{\text{FA}}^i \leq f_{\text{MAX}}^i \quad \text{and} \quad t_{\text{MIN}} \leq P_t \leq t_{\text{MAX}}.$$

Letting  $D = d_{\text{MAX}}^1 d_{\text{MAX}}^2$  and  $F = f_{\text{MAX}}^1 f_{\text{MAX}}^2$ , [4] showed that the normalised belief that a target is present given detections from both sensors is given by

$$\begin{aligned} m(1) &= D t_{\text{MAX}} - D F (t_{\text{MAX}} - t_{\text{MIN}})/C, \\ m(0) &= F (1 - t_{\text{MIN}}) - D F (t_{\text{MAX}} - t_{\text{MIN}})/C, \\ m(x) &= D F (t_{\text{MAX}} - t_{\text{MIN}})/C, \end{aligned} \tag{2}$$

where  $C = t_{\text{MIN}}(1 - D) + (1 - F) + (t_{\text{MAX}} - t_{\text{MIN}})(1 - D)(1 - F)$  is the mass assigned to the empty set and  $m(x)$  reflects the uncertainty and is the mass assigned to the case that we do not know if there is a target.

When a decision is made using TBM, BBAs need to be transformed into probabilities. In order to do this the uncertain mass,  $m(x)$ , should be distributed between the target present and target absent cases. Note that if the target prior is known,  $t_{\text{MIN}} = P_t = t_{\text{MAX}}$ , and so  $m(x) = 0$ . A certain probability of target presence is obtained by simply normalising  $m(0)$  and  $m(1)$  even though the sensor performance is still uncertain. This solution is equivalent to the Bayesian solution for known sensor priors where these have been replaced with their upper bound values.

### 3 Probabilities from Belief Function Models

Different methods have been proposed to transform belief functions into probability masses, many of which were explored in [5]. Here we consider those which satisfy the *minimal criterion* of [5], which states that the result of applying the transform to the vacuous belief function should be a uniform probability. The vacuous belief function is one describing complete uncertainty. Three such transforms were identified in [5]: the Aggregate Uncertainty (AU); the pignistic transformation; and the Plausibility transformation. Each is described in detail in [5]. In addition, we will consider the generalized pignistic transformation [8, 9] since it also satisfies the minimal criterion.

We will consider only the case of a binary frame of discernment,  $X = \{A, B\}$ , since this is appropriate for the decision fusion example. Assume that after using the TBM for sensor fusion, the masses are normalized to force the empty set mass to be zero. Thus there are three masses,  $m(A)$ ,  $m(B)$  and  $m(A \cup B)$ . Without loss of generality, we let  $m(A) \geq m(B)$ . The general requirement of the transform is to map these three masses into binary probabilities,  $p(A)$  and  $p(B)$ .

#### 3.1 Aggregate Uncertainty Approach

The Aggregate Uncertainty method distributes the mass  $m(A \cup B)$  between  $m(A)$  and  $m(B)$  and then normalises the two to arrive at a probability mass.

The AU selects the proportion of  $m(A \cup B)$  distributed to each in such a way as to maximise the entropy of the resulting probability mass [3]. For a binary frame, this process is straightforward: if  $m(A) \geq 0.5$ , the transform yields  $P_1(A) = m(A)$  and  $P_1(B) = m(B) + m(A \cup B)$ , otherwise  $P_1(A) = P_1(B) = 0.5$ .

### 3.2 The Plausibility Transformation

Also known as the Bayesian approximation method, it has been argued that this transformation is the most consistent technique for translating Dempster's rule of combination to the Bayesian probability domain. For a binary frame, if  $x \in X$ , the probabilities given by this transformation become

$$P_2(x) = \frac{m(x) + m(A \cup B)}{1 + m(A \cup B)}.$$

### 3.3 The Pignistic Transformation

The pignistic transform moves the belief mass from the union elements of the power set and distributes it equally amongst the singleton members. For a binary frame, if  $x \in X$ , the transform gives

$$P_3(x) = m(x) + \frac{1}{2}m(A \cup B). \quad (3)$$

For a binary frame, it can be shown that the Baroni-Vicig transformation [10] is equivalent to the pignistic transformation.

### 3.4 The Generalised Pignistic Transformation

This transformation aims to maximise the Probabilistic Information Content (PIC). Here it suffices to say that maximization of the PIC is equivalent to minimization of Shannon's entropy [8, 9, 11]. For a binary frame, if  $x \in X$ , the transformation is given by

$$P_4(x) = \frac{m(x)(1 + 2\varepsilon) + \varepsilon m(A \cup B)}{m(A) + m(B) + 2\varepsilon},$$

for some small arbitrary  $\varepsilon \geq 0$ .

The entropy of  $P_4(x)$  is minimised if  $\varepsilon = 0$  [8, 9]. However, if  $m(B)$  is zero and  $m(A \cup B)$  is not, then the transform gives  $p(A) = 1$  despite the uncertainty, which is inappropriate.

### 3.5 Analysis of Probability Transformations

We now present a relationship between these probability transformations for a binary frame and  $m(A) \geq m(B)$ .

**Theorem 1.**

*If  $X = \{A, B\}$  and  $m(A) \geq m(B)$  then  $P_1(A) \leq P_2(A) \leq P_3(A) \leq P_4(A)$ .*

*Proof.* For the case when  $m(A) \geq 0.5$ , then showing  $P_1(A) \leq P_2(A)$  is equivalent to proving that

$$m(A) \leq \frac{1 - m(B)}{1 + m(A \cup B)}$$

$m(A) + m(A)m(A \cup B) + m(B) \leq 1$ , so that  $P_1(A) \leq P_2(A)$  when  $m(A) \geq 0.5$ . If  $m(A) < 0.5$ ,  $m(A) \geq m(B)$  then  $P_1(A) = P_1(B) = 0.5$ . In such a case we note that  $0.5(1 + m(A \cup B)) \leq m(A) + m(A \cup B)$  so that  $P_1(A) \leq P_2(A)$ .

Similarly, after some manipulation we can also see that when  $m(A) \geq m(B)$

$$m(A) + m(A \cup B) \leq \left[ m(A) + \frac{1}{2}m(A \cup B) \right] [1 + m(A \cup B)].$$

This means that  $P_2(A) \leq P_3(A)$ . And lastly, we note that  $P_3(A) \leq P_4(A)$  because it can easily be shown that

$$\left[ m(A) + \frac{1}{2}m(A \cup B) \right] [m(A) + m(B) + 2\varepsilon] \leq m(A)[1 + 2\varepsilon] + \varepsilon m(A \cup B).$$

Finally, from above the following lemma can be established. It provides a lower bound on the probability values for all the transformations.

**Lemma 1.** *For all the probability transformations listed above we have  $p(A) \geq m(A)$  and  $p(B) \geq m(B)$ .*

As mentioned in section 2, when the prior probability of a target is known, then there is no mass assigned to the uncertain event  $A \cup B$  and so all of the transforms are equivalent.

## 4 Bayesian Solution

Section 3 establishes a relationship between a number of different transforms, but it does not provide any advice regarding which, if any, is preferred. One potential way to discriminate between them is to compare the transforms with the result of Bayesian analysis. Is it possible, for example, to craft a Bayesian prior to give the same result as the pignistic transform?

When the parameters of a prior distribution are unknown, the Bayesian approach is to treat these parameters as random variables. Such random parameters are referred to as hyperparameters and their distributions as hyperpriors [12]. Within the sensor fusion context, we will now treat the probability of detection, the probability of false alarm and the probability of target presence as hyperparameters with known hyperpriors. That is,  $P_D^i$  is a random variable with known distribution (hyperprior)  $p(P_D^i)$ .

Let  $t$ ,  $s^1$  and  $s^2$  be binary indicator variables that denote the presence (or absence) of a target and sensor reports respectively.

The target is present with probability  $P_t$ . This probability is unknown, but it has a known hyperprior,  $p(P_t)$ . The mean target prior is  $E\{P_t\}$  and the hyperprior is zero outside of the region  $[t_{\text{MIN}}, t_{\text{MAX}}]$ . Explicitly,

$$p(t|P_t) = \begin{cases} 1 - P_t & t = 0, \\ P_t & t = 1. \end{cases} \quad (4)$$

The sensor output is a function of the sensors' probability of detection, probability of false alarm, and the presence of a target, i.e.

$$p(s^i|t, P_D^i, P_{FA}^i) = \begin{cases} 1 - P_{FA}^i & t s^i = 00, \\ P_{FA}^i & t s^i = 01, \\ 1 - P_D^i & t s^i = 10, \\ P_D^i & t s^i = 11, \end{cases} \quad (5)$$

where again the probability of detection,  $P_D^i$ , and probability of false alarm,  $P_{FA}^i$ , are unknown but have known hyperpriors,  $p(P_D^i)$  and  $p(P_{FA}^i)$ .

**Theorem 2.** *Let  $p(t|s^1, s^2)$  be the conditional probability of a target being present ( $t = 1$ ) or not ( $t = 0$ ), given two sensor outputs. Then, the Bayesian solution when target, false alarm and detection priors are unknown is given by*

$$p(t|s^1, s^2) \propto p(t|E\{P_t\}) \prod_{i=1}^2 p(s^i|t, E\{P_D^i\}, E\{P_{FA}^i\}) \quad (6)$$

*Proof.* The fused output,  $p(t|s^1, s^2)$ , is derived through the use of Bayes' Rule, the law of total probability and conditional independence:

$$\begin{aligned} p(t|s^1, s^2) &\propto p(t, s^1, s^2) \\ &= \int_0^1 \int_0^1 \int_0^1 \int_0^1 \int_0^1 p(t, s^1, s^2, P_t, P_D^1, P_D^2, P_{FA}^1, P_{FA}^2) \\ &\quad dP_t dP_D^1 dP_D^2 dP_{FA}^1 dP_{FA}^2, \\ &= \int_0^1 \int_0^1 \int_0^1 \int_0^1 \int_0^1 p(P_t) p(t|P_t) \\ &\quad \times \prod_{j=1}^2 p(P_D^j) p(P_{FA}^j) p(s^j|t, P_D^j, P_{FA}^j) dP_t dP_D^1 dP_D^2 dP_{FA}^1 dP_{FA}^2, \\ &= \left\{ \int_0^1 p(P_t) p(t|P_t) dP_t \right\} \\ &\quad \times \prod_{j=1}^2 \left\{ \int_0^1 \int_0^1 p(P_D^j) p(P_{FA}^j) p(s^j|t, P_D^j, P_{FA}^j) dP_D^j dP_{FA}^j \right\} \quad (7) \end{aligned}$$

The first term in (7) can be simplified by substituting (4)

$$\begin{aligned} \int_0^1 p(P_t) p(t|P_t) dP_t &= \begin{cases} \int_0^1 p(P_t) (1 - P_t) dP_t & t = 0, \\ \int_0^1 p(P_t) P_t dP_t & t = 1, \end{cases} \\ &= \begin{cases} 1 - E\{P_t\} & t = 0, \\ E\{P_t\} & t = 1. \end{cases} \\ &= p(t|E\{P_t\}). \end{aligned} \quad (8)$$

Similarly, the second term in (7) is simplified by substituting (5) to get

$$\begin{aligned} \int_0^1 \int_0^1 p(P_D^i) p(P_{FA}^i) p(s^i|t, P_D^i, P_{FA}^i) dP_D^i dP_{FA}^i &= \begin{cases} 1 - E\{P_{FA}^i\} & t s^i = 00, \\ E\{P_{FA}^i\} & t s^i = 01, \\ 1 - E\{P_D^i\} & t s^i = 10, \\ E\{P_D^i\} & t s^i = 11, \end{cases} \\ &= p(s^i|t, E\{P_D^i\}, E\{P_{FA}^i\}). \end{aligned} \quad (9)$$

It is very interesting, and perhaps unexpected, that the fused probability of a target given multi-sensor data depends only on the means of the hyperpriors. The implication is that we could solve this problem using the Bayesian method for an unspecified hyperprior constrained only by its mean.

It is intuitive that (6) extends directly to an arbitrary number of sensors simply by appropriately changing the domain of the product. For the special case of  $N$  identical sensors where  $M \leq N$  sensors report a target, the result resembles a binomial distribution.

For the example of two sensors reporting a detection, the probability of a target being present is given by

$$p(t = 1|1, 1) = \frac{E\{P_t\}E\{P_D^1\}E\{P_D^2\}}{E\{P_t\}E\{P_D^1\}E\{P_D^2\} + (1 - E\{P_t\})E\{P_{FA}^1\}E\{P_{FA}^2\}}. \quad (10)$$

Let the Bayesian probability distribution be denoted by  $P_5$ . An obvious lemma when all the marginals are known can now be stated.

**Lemma 2.** *If  $t_{\text{MIN}} = t_{\text{MAX}}$ ,  $d_{\text{MIN}}^i = d_{\text{MAX}}^i$ ,  $i = 1, 2$ , then*

$$P_1(j|(1, 1)) = P_2(j|(1, 1)) = P_3(j|(1, 1)) = P_4(j|(1, 1)) = P_5(j|(1, 1)),$$

*that is, the TBM solution is the same as the Bayesian solution, independent of the belief to probability transformation method used.*

Recall that when the target probability is known then all the probability transformations are equivalent. For this case, denote the TBM derived probability value of an event  $A$  by  $P^*(A)$  and establish the following resulting lemma which indicates when the Bayesian solution will yield a higher or lower value than the probability transformations when the target prior probability is known.

**Lemma 3.** *When  $t_{\text{MIN}} = t_{\text{MAX}}$ , then*

$$\frac{D}{F} \begin{matrix} \geq \\ < \end{matrix} \frac{E\{P_{\text{D}}^1\}E\{P_{\text{D}}^2\}}{E\{P_{\text{FA}}^1\}E\{P_{\text{FA}}^2\}} \iff P^*(t|(1,1)) \begin{matrix} \geq \\ < \end{matrix} P_5(t|(1,1))$$

The dependence between the parameters and the probability values is more complicated when the target prior is uncertain and depends on the transformation method. For example, under the pignistic transform, Lemma 3 becomes:

**Lemma 4.**

$$\frac{t_{\text{MAX}}D - \frac{1}{2}(t_{\text{MAX}} - t_{\text{MIN}})DF}{(1 - t_{\text{MIN}})F - \frac{1}{2}(t_{\text{MAX}} - t_{\text{MIN}})DF} \begin{matrix} \geq \\ < \end{matrix} \frac{E\{P_t\}E\{P_{\text{D}}^1\}E\{P_{\text{D}}^2\}}{(1 - E\{P_t\})E\{P_{\text{FA}}^1\}E\{P_{\text{FA}}^2\}} \\ \iff P_3(t|(1,1)) \begin{matrix} \geq \\ < \end{matrix} P_5(t|(1,1))$$

While there is some relationship between the transforms and the Bayesian solution, it is too complicated to draw any intuitive conclusion.

## 5 Examples

In this section we provide some graphical examples which help to visualize the relationships between the different transformations and the Bayesian solution. In all cases, we assume that the false alarm and detection parameters of both sensors are the same.

First consider the case of a known target prior,  $P_t = 0.5$ . Assume that the hyperpriors for the probability of detection and the probability of false alarm are both uniform, so

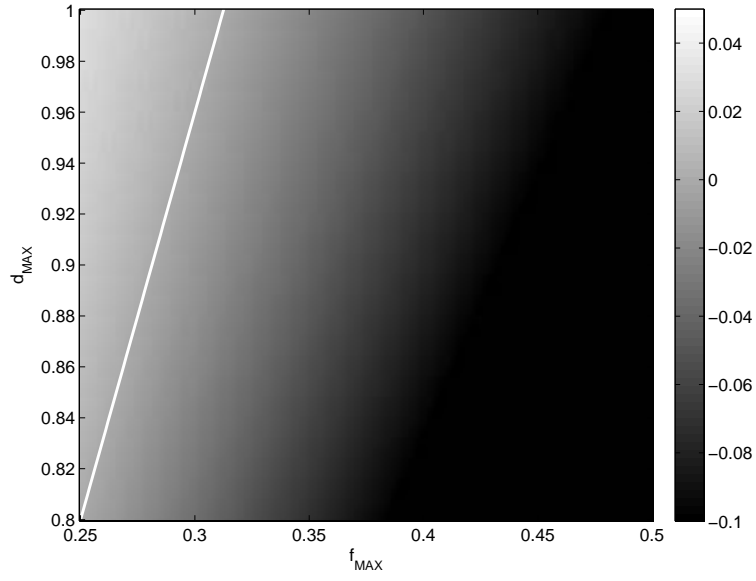
$$p(P_{\text{D}}^i) = \begin{cases} \frac{1}{d_{\text{MAX}} - d_{\text{MIN}}} & d_{\text{MIN}} \leq P_{\text{D}}^i \leq d_{\text{MAX}}, \\ 0 & \text{otherwise,} \end{cases}$$

with  $E\{P_{\text{D}}^i\} = \frac{1}{2}(d_{\text{MAX}} + d_{\text{MIN}})$ , and similarly for  $p(P_{\text{FA}}^i)$ .

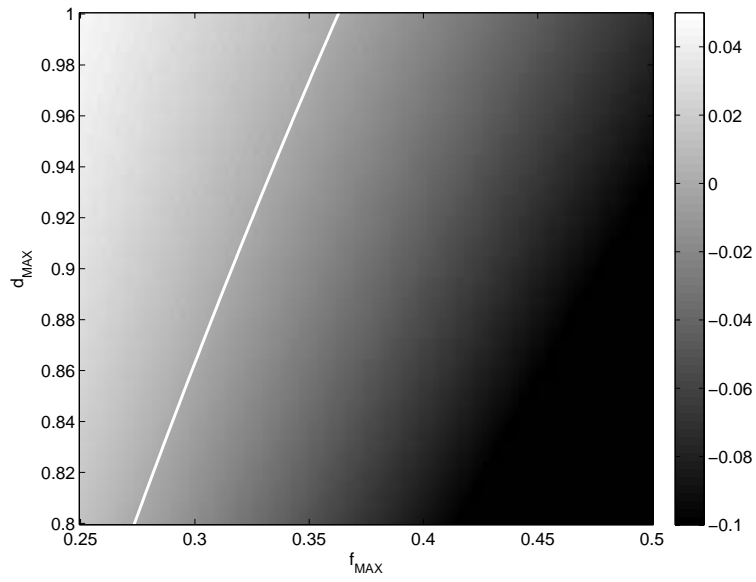
Figure 1 shows the difference between the TBM probability of a target and the Bayesian probability of a target as  $d_{\text{MAX}}$  and  $f_{\text{MAX}}$  are varied. Both  $d_{\text{MIN}}$  and  $f_{\text{MIN}}$  were also varied to maintain fixed values of  $E\{P_{\text{D}}\} = 0.8$  and  $E\{P_{\text{FA}}\} = 0.25$ . A white line is plotted showing the values where the two probabilities are equal, as defined by Lemma 3. For this example, when the uncertainty in the probability of detection is greater than the uncertainty in the probability of false alarm, the TBM fused probability of a target is higher than the Bayesian fused probability and when the probability of false alarm uncertainty is greater, the TBM fused probability is lower.

Next we change the target prior to be random and uniformly distributed on  $[0.3, 0.7]$ . As before, the detection probability hyperprior and false alarm probability hyperprior were varied but constant means were maintained. Figure 2 shows the difference between the pignistic transform probabilities and the Bayesian probabilities. The general shape of the function is the same as for a fixed  $P_t$ , but it has been shifted to the right. This means that for a particular false alarm hyperprior, equivalence is achieved for a smaller  $d_{\text{MAX}}$  when the target prior is uncertain.





**Fig. 1.** TBM probability - Bayes Probability,  $P_t$  known



**Fig. 2.** TBM probability - Bayes Probability,  $P_t$  random

## 6 Concluding Remarks

In this paper, we investigated several Dempster-Shafer mass probability transformations and analysed how they related to one another in the context of decision fusion in a dual sensor example. We showed that if the target prior is fixed, then an intuitive relationship can be established between the probability transformations and the Bayesian solution. However, even in this simplified case, the two are only equivalent in special cases. There is no apparent relationship between the TBM and the Bayesian solutions or particular form of prior that makes the two equivalent. This is because the TBM belief depends on the boundary conditions of the parameter hyperpriors, whereas the Bayesian solution presents an overall average result. Potential work could concentrate on determining in which situations it would be preferable to use a particular transformation over another.

## References

1. Niedermeyer, D.: An Introduction to Bayesian Networks and Their Contemporary Applications, *Innovations in Bayesian Networks. Studies in Computational Intelligence (SCI)*, Holmes, D. and Jain, L. (eds) **156**, 117–130 (2008)
2. Dempster, A. P.: A Generalisation of Bayesian Inference. *J. of the Royal Statistical Society, Series B* **30**, 205–247 (1968).
3. Shafer, G. *A Mathematical Theory of Evidence*, Princeton University Press, 1976.
4. Davey, S., Legg, J. and El-Mahassni.: Fusing Sensors with Uncertain Detection Performance. Proc. 5th. International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 267–272 (2009)
5. Oxenham, M., Challa, S. and Morelande, M.: Fusion of Disparate Identity Estimates for Shared Situation Awareness in a Network-Centric Environment. *Information Fusion*, **7**, 395–417 (2006)
6. Smets, Ph.: The Combination of Evidence in the Transferable Belief Model. *IEEE Trans. PAMI*, **12** 447–458 (1990)
7. Smets, Ph.: The nature of the unnormalized beliefs encountered in the transferable belief model. Proc. 8th Conf. on Uncertainty in AI, Dubois D., Wellmann M.P., D'Ambrosio B. and Smets Ph. (eds), Morgan Kaufmann Publ., San Mateo, California, 292–297 (1992)
8. Dezert, J. and Smarandache, F.: A New Probabilistic Transformation of Belief Mass Assignment. 11th. International Conference in Information Fusion (2008)
9. Dezert, J. and Smarandache, F.: Advances and Applications of DSMT for Information Fusion. *Collected Works*, **3**, Smarandache, F. and Dezert, J. (eds), Perfect Paperback, (2009)
10. Baroni, P. and Vicig, P.: Transformations from Imprecise to Precise Probabilities, Proc. of the 7th European Conf. on Symb. and Quant. Approaches to Reasoning with Uncertainty (ECSQARU 2003), 37–49 (2003)
11. Sudano, J.: The System Probability Information Content (PIC) Relationship to Contributing Components, Combining Independent Multi-Source Beliefs, Hybrid and Pedigree Pignistic Probabilities. Proc. 5th Conf. in Information Fusion, **2**, 1277–1283 (2002).
12. Bernardo, J. M. and Smith, A.F.M.: *Bayesian Theory* (2000) Wiley.