Recursive Classification of Multiple Objects Using Discordant and Non-Specific Data

Branko Ristic[†] and **Philippe Smets**[‡]

Abstract

The problem of multiple object classification based on discordant and non-specific data is considered. A general methodology for solving this problem is suggested and a suboptimal single-scan algorithm, referred to as the the global nearest neighbour (GNN), is implemented. The exact global dissimilarity measure, which is minimised by the GNN algorithm, is derived within the framework of the belief function theory. This measure, based on the plausibility of the global assignment, is related to the degree of conflict as understood in the transferable belief model interpretation of the belief function theory. The performance of the GNN algorithm was analysed by Monte Carlo simulations using different variants of the basic algorithm. One of the variants considered was the Bayesian GNN classifier. The results of this study suggest that the GNN classifier based on the exact global dissimilarity measure performs by far the best of the considered alternatives.

Keywords: Classification, belief function theory, non-specific evidence, data association.

Submitted to IEEE Trans. on Pattern Analysis and Machine Intelligence

May 11, 2004

[†] Employed by DSTO, ISRD - 200 Labs, PO Box 1500, Edinburgh SA 5111, Australia. Currently on study leave at IRIDIA, Université libre de Bruxelles, Av. F. Roosevelt 50, CP 194/6, 1050 Bruxelles, Belgium; tel +32 (2) 650 27 36; email: branko@iridia.ulb.ac.be

[‡] IRIDIA, Université libre de Bruxelles, Av. F. Roosevelt 50, CP 194/6, 1050 Bruxelles, Belgium; tel: +32 (2) 344 82 96; email: psmets@ulb.ac.be

1 Introduction

This paper deals with the classification of an unknown number of objects, using the reports that describe object classes in an uncertain manner. The reports are characterised by very small probability of detection, meaning that on average a small fraction of objects is detected and their characteristics reported at the time of sensing the environment. A sequence of independent reports (possibly from multiple sensors) is at disposal and the goal is to determine the number and classes of objects.

The available reports for multi-object classification are characterised by two types of uncertainty [1], [2]. The first type corresponds to the uncertainty due to randomness. For example, if we were to select at random a ball from an urn with 4 red, 2 blue and 14 green balls, there is a non-zero probability of selecting either of the three colours (all three are possible). This type of uncertainty is therefore referred to as [1] the discord between the alternatives, and is well known in the scientific and engineering literature: the primary tool for dealing with the discord is the probability theory. Suppose we consider a universe with a finite set of exclusive and exhaustive alternatives (hypotheses) Θ. The probability density function (pdf) in this case is a weighted sum of Dirac delta functions. If all but one weights in the pdf are equal to zero, the discord is zero, and there is no uncertainty (i.e. the pdf describes a certain event). The highest discord corresponds to the case where all weights are equal. The measure of discord is the Shannon entropy.

The second type of uncertainty is related to vagueness or imprecision of reports such as the statement S: "The true hypothesis belongs to subset A of the universe Θ", where the cardinality of subset A (denoted as $|A|$) is greater than 1. In the extreme case, when $|A| = |\Theta|$, statement S is an expression of total ignorance. This type of uncertainty is referred to as *non-specificity* [1] or vagueness [2]. A few mathematical theories have been developed recently to deal with non-specificity in addition to the discord, notably the belief function theory [3] and the random set theory [4]. The probability theory is not formally equipped to deal with non-specificity. Most probabilists argue, however, that non-specificity can be dealt appropriately within the framework of the probability theory by using equi-probability or the uniform density. This approach would then replace statement S (above) by a statement $\hat{S}$: "The probability density function over Θ is a weighted sum of Dirac delta functions, where the weights of the elements in A are equal to $1/|A|$ and the others are zero". We will see in the paper that the described approach may lead to the loss of performance, which should not be surprising: it replaces a non-specific (but correct) statement with a specific but possibly incorrect statement. A generally accepted measure of non-specificity was defined by Dubois and Prade [5].

The described problem of multi-object classification can be of importance for air surveillance systems when objects are closely spaced so that their positional measurements are of no value[1], but their attribute reports (obtained for example by processing electromagnetic emissions) are available for classification. The problem has many similarities with the classical multi-object tracking problem [6], because sensor reports are not labelled and the determination of the measurement origin (the so called data association problem) has to be resolved. Similar issues have been encountered in mine detection [7], multi-sensor allocation for submarine detection [8] and intelligence clustering [9].

The paper is organised as follows. Section 2 formulates the problem of multiple object classification. Section 3 presents a review of the belief function theory which provides an

[1] Due to finite resolution and/or measurement errors.

adequate framework for dealing with discordant and non-specific pieces of evidence. Section 4 describes an adopted solution to multi-object classification. Section 5 is devoted to dissimilarity for the global assignment. Some numerical results of multi-object classification are reported and discussed in Section 6. The conclusions are presented in Section 7.

2 Problem Formulation

Suppose there is an unknown number M of objects in a given volume of interest. The collection of objects is denoted by $O = \{1, 2, \ldots, M\}$. Each object $i \in O$ belongs to one of the predefined class categories. The exclusive and exhaustive set of class categories (the universal set) is given by:

$$\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}. \tag{1}$$

For simplicity we assume that the number of objects M is constant. The class of object $i \in O$ is denoted as $\theta(i) \in \Theta$.

At our disposal are sensors that are capable of measuring attributes or features that relate to object classes. In the context of the air surveillance application, for example, we could monitor a group of flying objects (in the surveillance volume of interest), and for each detected object we could measure its characteristic features (e.g. its shape, kinematic behaviour or its electromagnetic emissions). These features in general lead to discordant and non-specific sensor reports. For example, an IFF sensor may report: "the object is not friendly, with the probability of 0.7". This report is discordant because its probability is less than 1, but it is also non-specific, because a non-friendly object may be either neutral or hostile[2]. Sensors in general detect objects with probability $P_D < 1$ and in addition may

[2]Furthermore, a universal set that consists of three primary identification classes, i.e. {friendly, neutral

report false alarms with probability P_{FA}. In this paper, however, we will restrict ourselves to the $P_{FA} = 0$ case.

A sequence of independent sensor reports (from possibly multiple sensors) is available for classification. These reports are indexed by $k = 1, 2, \ldots,$ where k can be interpreted as the time index of a report. Suppose a sensor providing a report at time k detected $D_k \leq M$ objects, denoted as $o_{kj_k}, j_k = 1, \ldots, D_k$. For each detected object it provides a partial (discordant and/or non-specific) knowledge about the actual class to which object o_{kj_k} belongs.

This partial knowledge represents a weighted opinion expressed over the subsets of the universal set Θ, and is denoted as $m_k^{\Theta}\{\theta(o_{kj_k})\}$, or simply m_{kj_k}. For convenience, the weights are scaled to 1, i.e.:

$$\sum_{A \subseteq \Theta} m_{kj_k}(A) = 1. \tag{2}$$

Note that for $A \subseteq \Theta$, $m_{kj_k}(A)$ is the part of the unit mass of the sensor opinion that supports set A, but due to a lack of further information, does not support any strict subset of A. Function m_{kj_k} is a mapping of the power set over Θ (denoted as 2^{Θ}) to $[0, 1]$. We will see in the next section that function m_{kj_k} plays a central role in the belief function theory.

A sensor report at k represents a collection:

$$R_k = \{m_{kj_k}, j_k = 1, \ldots, D_k\}. \tag{3}$$

We further assume that no single object is detected and reported more than once in R_k. Conversely, each detection originates from a single object only.

and hostile}, is in practice only a coarsening of the universal set of all aircraft platforms. In this case, the cardinality of a set of non-friendly objects is much higher than 2.

The problem is to determine the number of objects M and the class of each object $\theta(i)$, $i = 1, \ldots, M$. We want this to be done in a recursive manner, that is, sequentially at each $k = 1, 2, \ldots,$ as the reports arrive. Note that it is possible that all objects are of the same class, and inversely that all objects are of different classes.

3 A review of the belief function theory

3.1 Representation of partial knowledge

We adopt the framework and terminology of the belief function theory [3] as interpreted by the transferable belief model (TBM) [10, 11], in order to provide a mathematical description of discordant and/or non-specific knowledge about a class. In this framework, function m_{kj_k} is referred to as the basic belief assignment (bba), and $m_{kj_k}(A)$ is interpreted as a fraction of unit mass of belief which is allocated by a sensor in report R_k specifically to $A \subseteq \Theta$, in relation to the class of detected object o_{kj_k}. The subsets A of Θ such that $m_{kj_k}(A) > 0$ are referred to as the focal sets of this bba. The set of all focal sets of m_{kj_k} is denoted by $\mathcal{F}_{kj_k}$. The universal set in the terminology of belief function theory is referred to as the *frame of discernment*.

We will further simplify notation by dropping the subscripts of the bba and $\mathcal{F}$. Some special cases of bba's are as follows:

- **Certain bba** represents perfect knowledge of the class – it has only one focal set whose cardinality equals 1 (a singleton).
- **Vacuous bba** represents total ignorance – it has only one focal set, the universal set Θ.

- **Bayesian bba** represents a discordant but fully specific knowledge, i.e. all focal sets are singletons. This bba corresponds to a probability function.
- A **categorical non-specific bba** represents a fully accordant (non-conflicting) but non-specific knowledge – it has only one focal set A, such that $1 < |A| < |\Theta|$. Note that statement S in the Introduction is an example of this type of the bba.

A bba with a zero mass given to the empty set, i.e. $m(\emptyset) = 0$, is said to be normalised. In this paper we will assume that sensors are supplying partial knowledge concerning the class membership of detected objects as normalised bba's.

A bba m can be equivalently represented by two non-additive fuzzy measures [1], a belief function $bel : 2^\Theta \to [0,1]$ and a plausibility function $pl : 2^\Theta \to [0,1]$, defined as:

$$bel(A) \triangleq \sum_{\emptyset \neq B \subseteq A} m(B), \quad \forall A \subseteq \Theta \tag{4}$$

$$pl(A) \triangleq \sum_{B \cap A \neq \emptyset} m(B) = bel(\Theta) - bel(\bar{A}) \qquad \forall A \subseteq \Theta \tag{5}$$

respectively ($\bar{A}$ is a complement of A). According to (4), a normalised bba is characterised by $bel(\Theta) = 1$. The interpretation of these two functions is as follows: $bel(A)$ represents the total belief that is committed to A without also being committed to $\bar{A}$; $pl(A)$ corresponds to the total belief which does not contradict A.

3.2 Combination of partial knowledge

Suppose we have two bba's m_1 and m_2 on the same frame of discernment Θ, representing two distinct pieces of evidence (sensor reports) about the object class. The joint impact of these two pieces of evidence can be expressed by the bba:

$$m_{1 \textcircled{\cap} 2}(A) = (m_1 \textcircled{\cap} m_2)(A) = \sum_{B,C \subseteq \Theta : B \cap C = A} m_1(B) \cdot m_2(C) \tag{6}$$

for all $A \subseteq \Theta$. Operation $\textcircled{\cap}$, referred to as the conjunctive rule of combination, is both commutative and associative. It is important to know that the conjunctive combination of two normalised bba's m_1 and m_2 may result in a bba $m_{1 \textcircled{\cap} 2}$ with non-zero value of mass given to the empty set. This quantity, denoted as

$$m_{1 \textcircled{\cap} 2}(\emptyset) \triangleq (m_1 \textcircled{\cap} m_2)(\emptyset) = \sum_{B \cap C = \emptyset} m_1(B) m_2(C), \tag{7}$$

is known as the degree of conflict between m_1 and m_2 because it represents the degree of disagreement between the two sources of evidence. The normality condition may be imposed on $m_{1 \textcircled{\cap} 2}$ by setting $m_{1 \textcircled{\cap} 2}(\emptyset)$ to zero and by dividing each mass $m_{1 \textcircled{\cap} 2}(A)$, where $A \neq \emptyset$, by $1 - m_{1 \textcircled{\cap} 2}(\emptyset)$. The conjunctive rule of combination with normalisation is referred to as the Dempster's combination and is denoted as $\oplus$.

3.3 Pignistic probability

The pignistic probability is the result of a mapping of a belief function m to a Bayesian belief function, that is a probability function, denoted as $BetP$. For the singletons $\theta_i \in \Theta$ of the resulting Bayesian bba we have [11]:

$$BetP(\{\theta_i\}) = \sum_{\theta_i \in A \subseteq \Theta} \frac{1}{|A|} \frac{m(A)}{1 - m(\emptyset)}. \tag{8}$$

The pignistic transformation (8) is linear and has some other useful properties [11], such as that $bel(\theta_i) \leq BetP(\theta_i) \leq pl(\theta_i)$, if $m(\emptyset) = 0$. Note that any Bayesian bba is invariant to the pignistic transformation. The transformation of statement S (a categorical non-specific bba) to Statement $\hat{S}$ (a Bayesian bba) in the Introduction was carried out using the pignistic transform. *BetP* is the probability measure that we use for decision making (betting) and hence its name (*pignus* means a bet or a wage in Latin).

3.4 Vacuous extension and the product space

Given a bba m^X (superscript here denotes the domain, that is the universal set on which the bba is defined), its vacuous extension on space $X \times Y$, denoted $m^{X\uparrow X\times Y}$ is given by

$$m^{X\uparrow X\times Y}(C) = \begin{cases} m^X(A) & \text{iff } C = A \times Y, A \subseteq X \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Given two bba's m_1^X and m_2^Y, their conjunctive combination on $X \times Y$ can be obtained by combining their vacuous extensions on $X \times Y$ using (6). Formally:

$$m_1^X \circledcap m_2^Y = m_1^{X\uparrow X\times Y} \circledcap m_2^{Y\uparrow X\times Y} \tag{10}$$

We thus obtain

$$(m_1^X \circledcap m_2^Y)(C) = \begin{cases} m_1^X(A)\, m_2^Y(B) & \text{iff } C = (A, B), A \subseteq X, B \subseteq Y \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

4 Multi-Object Classification

4.1 General Methodology

The proposed general methodology for multi-object classification using discordant and non-specific object description is shown in Figure 1. The most difficult part is to determine at each k the origin of all detected objects up to time k. This block, referred to as *data*

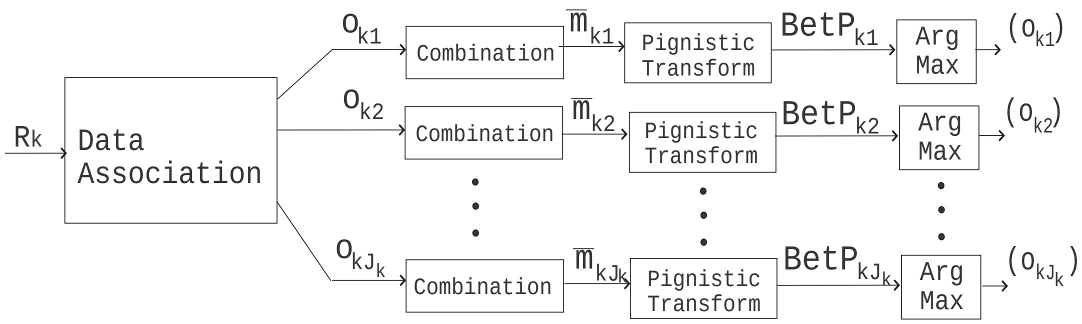

Figure 1: *A scheme for multi-object classification*

association, aims to partition the accumulated set of sensor reports $\{R_j; j = 1, \dots, k\} = \{m_{\ell,j_\ell}, \ell = 1, \dots, k, j_\ell = 1, \dots, D_\ell\}$ into clusters, each cluster defining an object. A cluster of bba's concerning the same object and taking into account sensor reports $R_1, R_2, \dots, R_k$ is denoted as $\mathcal{M}_{kj} = \{m_{1j_1}, m_{2j_2}, \dots, m_{kj_k}\}$, where $j = 1, \dots, J_k$ and $j_\ell \in \{0, 1, 2, \dots, D_\ell\}$. In this notation $j_\ell = 0$ means that object j has not been detected at scan ℓ and $m_{\ell 0}$ is then a vacuous bba. A partitioning hypothesis thus forms J_k clusters. Number J_k is an estimate of the number of objects M.

An optimal algorithm for data association would need to enumerate all possible partitioning hypotheses and then to rank them based on some criterion to select the best.

Once data association is completed, the clustered measurements are combined by the Dempster's rule of combination in order to determine the object class. The result of this combination is a (possibly non-specific) bba

$$\bar{m}_{kj} = m_{1j_1} \oplus m_{2j_2} \oplus \cdots \oplus m_{kj_k}. \tag{12}$$

concerning the class of object o_{kj}, $j = 1, \dots, J_k$.

Next one can apply the pignistic transform and find its maximum in order to make a decision about the class of o_{kj} as:

$$\theta(o_{kj}) = \arg\max_{\theta_i} \left\{ BetP(\{\theta_i\}) \right\}. \tag{13}$$

In order to further illustrate the data association problem consider the following example.

Example 1. Suppose that the frame of discernment is $\Theta = \{\square, \triangle, \diamondsuit, \bigcirc\}$. Let the detected objects and sensor reports be as shown in Figure 2, with $k = 3$ scans, $P_D < 1$. All sensor reports in this example consist of certain bba's. The correct partition of bba's in the

R_1	R_2	R_3
$m_{11}(\square)=1$	$m_{21}(\diamondsuit)=1$	$m_{31}(\bigcirc)=1$
$m_{12}(\bigcirc)=1$	$m_{22}(\square)=1$	$m_{32}(\square)=1$
$m_{13}(\triangle)=1$	$m_{23}(\triangle)=1$	$m_{33}(\diamondsuit)=1$
	$m_{24}(\bigcirc)=1$	

Figure 2: *An illustration of data association problem for multiple object classification (certain bba's)*

accumulated set of reports $\{R_1, R_2, R_3\}$, i.e. the correct data association hypothesis, in this case is as follows:

$$\begin{aligned}
\mathcal{M}_{31} &= \{m_{11}, m_{22}, m_{32}\} \text{ for } \square \\
\mathcal{M}_{32} &= \{m_{12}, m_{24}, m_{31}\} \text{ for } \bigcirc \\
\mathcal{M}_{33} &= \{m_{13}, m_{23}, m_{30}\} \text{ for } \triangle \\
\mathcal{M}_{34} &= \{m_{10}, m_{21}, m_{33}\} \text{ for } \diamondsuit.
\end{aligned}$$

With discordant and non-specific measurements, however, the correct partition will not be so obvious.

The number of possible hypotheses in the data association grows exponentially with k; this is a well known NP hard problem in the tracking community [6], [12]. Practical algorithm for data association are therefore based on approximations and one such an approximation, known as the global nearest neighbour (GNN) algorithm, is presented next.

4.2 GNN algorithm

4.2.1 Overview

The GNN algorithm [6] is a single-scan algorithm which selects only the best hypothesis at scan k and discards the rest. Suppose this best hypothesis forms J_k clusters $\mathcal{M}_{kj} = \{m_{1j_1}, m_{2j_2}, \dots, m_{kj_k}\}$, where $j = 1, \dots, J_k$ and $j_\ell \in \{0, 1, 2, \dots, D_\ell\}$. Here the assumption is that all bba's of cluster $\mathcal{M}_{kj}$ have a common origin, which is object o_{kj}. The class of object o_{kj} is described by bba $\bar{m}_{kj} = \oplus_{\nu=1}^{k} m_{\nu j_\nu}$.

When a new report arrives at $k+1$, $R_{k+1} = \left\{m_{k+1,j_{k+1}}; j_{k+1} = 1, \dots, D_{k+1}\right\}$, the GNN algorithm builds an *assignment matrix*, whose elements are dissimilarity measures between the description of the existing (current) objects at k, $\{\bar{m}_{kj}; j = 1, \dots, J_k\}$ and the measurements in report R_{k+1}. The choice of an appropriate dissimilarity measure will be discussed in Section 5. An example of the assignment matrix is given in Table 1, with $J_k = 4$ and $D_{k+1} = 3$.

The goal of the GNN algorithm is to find the overall (global) minimum dissimilarity assignment between the currently existing objects and the measurements of the new report. This problem is known as the two-dimensional (2D) assignment problem, and has a long history [6]. Several optimal algorithmic solutions, such as the Hungarian, Munkres, JVC and auction algorithm, have been proposed for 2D assignment. These solutions only differ in the

Table 1: *Example of an assignment matrix (optimal solution in bold)*

Current at k	Detected $m_{k+1,1}$	$m_{k+1,2}$	$m_{k+1,3}$
$\bar{m}_{k1}$	0.15	**0.12**	0.18
$\bar{m}_{k2}$	0.35	0.10	**0.12**
$\bar{m}_{k3}$	**0.17**	0.15	0.20
$\bar{m}_{k4}$	0.35	0.25	0.30

computational efficiency, not in the final assignment. Note that 2D assignment algorithms treat the global dissimilarity as the *sum* of the pairwise dissimilarities.

The solution to the 2D assignment problem may not be trivial and this is illustrated by the example of Table 1. Here, the globally optimal solution (indicated by bold numbers in the table) pairs $m_{k+1,1}$ with $\bar{m}_{k3}$, despite the fact that $m_{k+1,1}$ is more similar to $\bar{m}_{k1}$. Likewise, $m_{k+1,2}$ is paired with $\bar{m}_{k1}$, although $m_{k+1,2}$ is "closer" to $\bar{m}_{k2}$.

Once the global assignment is decided (using one of the mentioned algorithms), the bba's of detected objects at $k+1$ are combined with bba's of assigned existing objects at time k, using the Dempster's rule. In the example of Table 1 one would perform the combination as follows:

$$\begin{aligned}
\bar{m}_{k+1,1} &= \bar{m}_{k1} \oplus m_{k+1,2} \\
\bar{m}_{k+1,2} &= \bar{m}_{k2} \oplus m_{k+1,3} \\
\bar{m}_{k+1,3} &= \bar{m}_{k3} \oplus m_{k+1,1} \\
\bar{m}_{k+1,4} &= \bar{m}_{k4}.
\end{aligned}$$

4.2.2 The number of objects

When two objects belong to the same class, in general it does not imply that they *are* the same object. Due to the lack of additional information, we will adopt the parsimonious approach which asserts that the solution to the global assignment problem should involve the smallest possible number of objects. In this way equality of classes becomes equality of objects.

For the global assignment, the parsimonious approach will always try to assign all measurements at $k+1$ to the existing objects at k. This, however, may not be always possible, because for some measurement to object pairings, the degree of conflict could be equal to 1.0. These associations are referred to as incompatible as their corresponding bba's cannot be combined by the Dempster's rule (the effect of normalisation would be to divide by zero). This process of eliminating some associations based on their incompatibility is known as *gating* in the tracking literature [6]. Thus, we define a compatibility matrix $\boldsymbol{\kappa} = [\kappa_{ij}]$, which is of the same dimension as the assignment matrix, with elements

$$\kappa_{ij} = \begin{cases} 1 & \text{if } m_i \bigcirc\!\!\!\!\cap\, m_j(\emptyset) < 1 - \epsilon \\ 0 & \text{otherwise,} \end{cases} \tag{14}$$

where ϵ is a tuning parameter.

The number of objects, based on the parsimonious approach and the compatibility matrix, is then given by:

$$\hat{M}_{k+1} = J_k + D_{k+1} - \min\left(\sum_i \max_j(\kappa_{ij}), \sum_j \max_i(\kappa_{ij})\right) \tag{15}$$

For the assignment matrix of Table 1, the compatibility matrix is a matrix of ones and $\hat{M}_{k+1} = J_k = 4$.

What remains in the proposed method of GNN for multi-object classification is to choose the appropriate dissimilarity measure in the association matrix. This is the subject of the next section.

5 Dissimilarity measures for the GNN algorithm

5.1 TBM solution

The basic assumption about our sensors is that two bba's, if they originate from the same class objects, are similar to one another and dissimilar otherwise. The more similar two bba's are, the more plausible it is that they originate from the same class objects. Thus we will adopt a similarity measure based on the plausibility that measurement pairs originate from the same class. A justification of this choice is given in Appendix.

5.1.1 Single assignment

We focus first on a simple case with a single assignment. We consider two objects, o_i and o_j, and their respective normalised bba's, $m^{\Theta}\{o_i\}$ and $m^{\Theta}\{o_j\}$, based on distinct pieces of evidence. The notation $m^{\Theta}\{o_i\}$ is adopted here to emphasize that this bba concerns object o_i; note that the domains of two bba's, $m^{\Theta}\{o_i\}$ and $m^{\Theta}\{o_j\}$ are identical. The plausibility that the two objects belong to the same class must be evaluated on the product space $\Theta^2 = \Theta \times \Theta$ using the conjunctive combination of the vacuous extensions of $m^{\Theta}\{o_i\}$ and $m^{\Theta}\{o_j\}$ [13]. From (11) it follows:

$$m^{\Theta^2}\{o_i, o_j\} = m^{\Theta}\{o_i\}^{\uparrow\Theta^2} \bigcirc\!\!\!\!\cap\; m^{\Theta}\{o_j\}^{\uparrow\Theta^2}. \tag{16}$$

Thus we have:

$$m^{\Theta^2}\{o_i,o_j\}(C) = \begin{cases} m^{\Theta}\{o_i\}(A)\cdot m^{\Theta}\{o_j\}(B) & \text{iff } C=(A,B), A\subseteq\Theta, B\subseteq\Theta \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

This bba represents the joint belief about the actual class of each object.

The plausibility that the two objects o_i and o_j belong to the same class, i.e. $\theta(o_i)=\theta(o_j)$, according to (5) represents a sum of the masses $m^{\Theta}\{o_i\}(A)\cdot m^{\Theta}\{o_j\}(B)$, where $A\cap B\neq\emptyset$. Formally:

$$\begin{aligned} pl^{\Theta^2}\{o_i,o_j\}\big(\theta(o_i)=\theta(o_j)\big) &= \sum_{A\cap B\neq\emptyset} m^{\Theta}\{o_i\}(A)\, m^{\Theta}\{o_j\}(B) && (18)\\ &= 1-\sum_{A\cap B=\emptyset} m^{\Theta}\{o_i\}(A)\, m^{\Theta}\{o_j\}(B) && (19)\\ &= 1-\sum_{A\cap B=\emptyset} m_i^{\Theta}(A)\, m_j^{\Theta}(B) && (20) \end{aligned}$$

where in (20) we used for $m^{\Theta}\{o_i\}$ and $m^{\Theta}\{o_j\}$ a shortened notation m_i^{Θ} and m_j^{Θ}, respectively. Based on the relationship (5) for normalised bba's, $pl(A)=1-bel(\bar{A})$, we observe from (20) that:

$$bel^{\Theta^2}\{o_i,o_j\}\big(\theta(o_i)\neq\theta(o_j)\big) = \sum_{A\cap B=\emptyset} m_i^{\Theta}(A)\, m_j^{\Theta}(B). \tag{21}$$

Suppose we had assumed that two objects o_i and o_j belong to the same class. Then we would have combined the two bba's by the conjunctive rule of combination and the mass given to the empty set would then be equal to:

$$m^{\Theta}_{i \bigcirc\!\!\!\!\cap\, j}(\emptyset) = \sum_{A\cap B=\emptyset} m_i^{\Theta}(A) m_j^{\Theta}(B). \tag{22}$$

From (21) and (22) it follows that

$$bel^{\Theta^2}\{o_i,o_j\}\big(\theta(o_i)\neq\theta(o_j)\big) = m^{\Theta}_{i \bigcirc\!\!\!\!\cap\, j}(\emptyset), \tag{23}$$

that is

$$pl^{\Theta^2}\{o_i, o_j\}\big(\theta(o_i) = \theta(o_j)\big) = 1 - m^{\Theta}_{i \bigcirc\!\!\!\!\cap\, j}(\emptyset). \tag{24}$$

Equation (23) provides a meaning of the mass given to the empty set in the framework of the TBM: *the mass given to the empty set by the conjunctive combination rule applied to two normalised belief functions is equal to the belief that the two objects do not belong to the same class* [13].

5.1.2 Global assignment of multiple objects

First we consider the case where $P_D = 1$, meaning that the number of detected objects is constant. For this case we prove the following theorem.

Theorem 5.1 *Suppose there are two collections of objects,* $\mathbb{O}_1 = \{o_{1i} : i = 1, \ldots, n\}$ *and* $\mathbb{O}_2 = \{o_{2j} : j = 1, \ldots, n\}$. *One's belief regarding the class of object* $o_{1i} \in \mathbb{O}_1$ *and* $o_{2j} \in \mathbb{O}_2$ *is quantified by bba* $m^{\Theta}\{o_{1i}\} = m^{\Theta}_{1i}$ *and* $m^{\Theta}\{o_{2j}\} = m^{\Theta}_{2j}$*, respectively. The global assignment of objects in* $\mathbb{O}_1$ *to the objects of* $\mathbb{O}_2$ *is specified by an assignment vector* $V = \big[v_1, \ldots, v_n\big]^T$*, where* v_i *means that* $\theta(o_{1i}) = \theta(o_{2v_i})$*, for* $i = 1, \ldots, n$*. The plausibility of global assignment* V *is then given by:*

$$pl^{\Theta^{2n}}\Big(\theta(o_{1i}) = \theta(o_{2v_i}) : i = 1, \ldots, n\Big) = \prod_{i=1}^{n}\Big(1 - m^{\Theta}_{1i \bigcirc\!\!\!\!\cap\, 2v_i}(\emptyset)\Big). \tag{25}$$

Proof. The plausibility of a global assignment must be evaluated on the product space Θ^{2n} by the conjunctive combination of vacuous extensions of bba's m^{Θ}_{1i} and m^{Θ}_{2i}, $i = 1, \ldots, n$. The joint bba on Θ^{2n} is given by:

$$m^{\Theta^{2n}}\{o_{11}, \ldots, o_{2n}\} = \bigcirc\!\!\!\!\cap_{i=1,2} \bigcirc\!\!\!\!\cap_{j=1,\ldots,n} m_{ij}^{\Theta\uparrow\Theta^{2n}}. \tag{26}$$

Since the assignment is specified by n pairs, the plausibility that $\theta(o_{1i}) = \theta(o_{2v_i})$, for $i = 1, \ldots, n$ is expressed as:

$$\begin{aligned}
pl^{\Theta^{2n}}\Big(\theta(o_{1i}) = \theta(o_{2v_i}) : i = 1, \ldots, n\Big) &= \sum_{i=1,\ldots,n; A_i, B_{v_i} \subseteq \Theta; A_i \cap B_{v_i} \neq \emptyset} \left(\prod_{i=1}^{n} m_{1i}^{\Theta}(A_i)\, m_{2v_i}^{\Theta}(B_{v_i}) \right) \\
&= \prod_{i=1}^{n} \sum_{A_i \cap B_{v_i} \neq \emptyset} m_{1i}^{\Theta}(A_i)\, m_{2v_i}^{\Theta}(B_{v_i}) \\
&= \prod_{i=1}^{n} pl^{\Theta^2}\{o_{1i}, o_{2v_i}\}\big(\theta(o_{1i}) = \theta(o_{2v_i})\big) \\
&= \prod_{i=1}^{n} \big(1 - m_{1i \circledcap 2v_i}^{\Theta}\{o_i\}(\emptyset)\big). \qquad (27)
\end{aligned}$$

The step from line 1 to line 2 above is based on the relation

$$\sum_{w_1, w_2} f(w_1) g(w_2) = \sum_{w_1} f(w_1) \sum_{w_2} g(w_2).$$

□

Note that $1 - pl^{\Theta^{2n}}\Big(\theta(o_{1i}) = \theta(o_{2v_i}) : i = 1, \ldots, n\Big)$ has the same form as the "meta-conflict" which Schubert [14] used as a criterion for clustering non-specific data.

In the case when $P_D < 1$ we deal with two collections of objects $\mathbb{O}_1 = \{o_{1i} : i = 1, \ldots, n_1\}$ and $\mathbb{O}_2 = \{o_{2j} : j = 1, \ldots, n_2\}$ with possibly $n_1 \neq n_2$. The smallest number of objects then must be $n = \max\{n_1, n_2\}$. Suppose for the argument sake that $n_1 < n_2$. Then we add $n - n_1$ "unobserved" objects to $\mathbb{O}_1$ so that $|\mathbb{O}_1| = n$. Added objects are represented by vacuous bba's, so that the pairwise assignment of objects in $\mathbb{O}_2$ to the added objects of $\mathbb{O}_1$ has plausibility equal to 1.0 (i.e. no conflict). The likewise procedure applies if $n_2 < n_1$. After addition of objects to either $\mathbb{O}_1$ or $\mathbb{O}_2$, the cardinality of these two sets will be equal to n and we can again apply Theorem 5.1. A similar procedure of adding objects to $\mathbb{O}_1$ and/or $\mathbb{O}_2$ is necessary if the compatibility matrix has all elements in some rows or columns equal to

Table 2: *The basic belief assignments quantifying one's belief about the class of objects in* $\mathbb{O}_1$ *and* $\mathbb{O}_2$*; e.g.* $m_{11}^{\Theta}(\{\theta_1, \theta_2\}) = 0.7$ *and* $m_{23}^{\Theta}(\{\theta_4, \theta_5\}) = 0.2$

$\mathbb{O}_1$				$\mathbb{O}_2$					
o_{11}		o_{12}		o_{21}		o_{22}		o_{23}	
$\mathcal{F}_{11}$	m_{11}^{Θ}	$\mathcal{F}_{12}$	m_{12}^{Θ}	$\mathcal{F}_{21}$	m_{21}^{Θ}	$\mathcal{F}_{22}$	m_{22}^{Θ}	$\mathcal{F}_{23}$	m_{23}^{Θ}
θ_1, θ_2	.7	θ_3, θ_4	1.	θ_3, θ_5	.6	$\theta_1, \theta_4, \theta_5$	1.	θ_1, θ_2	.8
θ_2	.3			Θ	.4			θ_4, θ_5	.2

zero. The following example illustrates the case where $n_1 \neq n_2$, with the full compatibility between the objects in $\mathbb{O}_1$ and $\mathbb{O}_2$.

Example 2. Suppose $\Theta = \{\theta_1, \ldots, \theta_5\}$, $n_1 = 2$ and $n_2 = 3$. Table 2 presents five collected bba's m_{si}^{Θ}, $s = 1, 2, i = 1, ..., n_s$, and their respective focal sets $\mathcal{F}_{si}$. Table 3 presents the results of six possible conjunctive combinations of the bba's specified by Table 2. Table 4 presents for each combination the value of the plausibility that two objects belong to the same class (the values of 1 minus the mass given to $\emptyset$ by the conjunctive combination rule). The goal is to select an assignment which maximises the product of plausibilities, while keeping the total number of objects as small as possible. There are at least $n = \max\{n_1, n_2\} = 3$ objects, so we must add one object, o_{13}, to $\mathbb{O}_1$. This added object is represented by a vacuous bba for $m_{13}^{\Theta}\{o_{13}\}$. Hence $m_{13 \bigcirc\!\!\!\!\cap\, 2j}^{\Theta}(\emptyset) = 0$ for $j = 1, 2, 3$ and thus the entries in row o_{13} of Table 4 are all ones. From Table 4 we observe that the two best assignments with 3 objects are achieved by associating o_{11} with o_{23}, o_{12} with o_{21} and o_{13} with o_{22} (the product is 0.80, assignment vector $V = (3, 1, 2)$) or associating o_{11} with o_{23}, o_{12} with o_{22} and o_{13} with o_{21} (the product is 0.80, assignment vector $V = (3, 2, 1)$).

Table 3: *Pairwise conjunctive combination of the bba's in* $\mathbb{O}_1$ *and* $\mathbb{O}_2$*. Focal sets are represented as binary numbers; e.g.* 10100 *and* 11001 *represent the sets* $\{\theta_3, \theta_5\}$ *and* $\{\theta_1, \theta_4, \theta_5\}$*, respectively. Masses given to the conflicts are in bold.*

			$\mathbb{O}_2$					
			o_{21}		o_{22}		o_{23}	
			$\mathcal{F}_{21}$	m_{21}	$\mathcal{F}_{22}$	m_{22}	$\mathcal{F}_{23}$	m_{23}
			10100	0.60	11001	1.00	00011	0.80
$\mathbb{O}_1$			11111	0.40			11000	0.20
	$\mathcal{F}_{11}$	m_{11}	$\mathcal{F}$	$m_{11\textcircled{\cap}21}$	$\mathcal{F}$	$m_{11\textcircled{\cap}22}$	$\mathcal{F}$	$m_{11\textcircled{\cap}23}$
o_{11}	00011	0.70	00000	**0.60**	00000	**0.30**	00000	**0.20**
	00010	0.30	00010	0.12	00001	0.70	00010	0.24
			00011	0.28			00011	0.56
	$\mathcal{F}_{12}$	m_{12}	$\mathcal{F}$	$m_{12\textcircled{\cap}21}$	$\mathcal{F}$	$m_{12\textcircled{\cap}22}$	$\mathcal{F}$	$m_{12\textcircled{\cap}23}$
o_{12}	01100	1.00	00100	0.60	01000	1.00	00000	**0.80**
			01100	0.40			01000	0.20

A full justification for the use of the proposed TBM similarity measure based on the plausibility of the global assignment is illustrated in Appendix for the example in Table 2.

Theorem 5.1 provides the global *similarity* measure for a 2D assignment V. The goal of GNN algorithm is to find the assignment which would maximise (25). Note that the right-hand-side of (25) represents a *product* of pairwise plausibilities. Since the standard 2D assignment algorithms (Munkres, auction, etc) minimise the global dissimilarity considered as a *sum* of pairwise dissimilarities, we define an additive global dissimilarity measure of an assignment based on (25) as follows:

$$\mathcal{D} = -\log\left[pl^{\Theta^{2n}}\Big(\theta(o_{1i}) = \theta(o_{2v_i}) : i = 1, \ldots, n\Big)\right] \quad (28)$$

$$= \sum_{i=1}^{n} d_{iv_i} \quad (29)$$

Table 4: *The plausibility that object* o_{1i} *and object* o_{2j} *belong to the same class;* o_{13} *is the unobserved object added to* $\mathbb{O}_1$*. The best association vector with* $n = 3$ *objects is either* $(3,1,2)$ *(meaning that* o_{11} *is associated with* o_{23}*,* o_{12} *with* o_{21} *and* o_{13} *with* o_{22}*) or* $(3,2,1)$*.*

	$\mathbb{O}_2$		
$\mathbb{O}_1$	o_{21}	o_{22}	o_{23}
o_{11}	0.40	0.70	0.80
o_{12}	1.00	1.00	0.20
o_{13}	1.00	1.00	1.00

where

$$d_{ij} = -\log\Big(1 - m^{\Theta}_{1i \circledcap 2j}(\emptyset)\Big). \tag{30}$$

Thus in creating an assignment matrix (as the one in Table 1) we use (30) to fill up its entries. In this way we can apply one of the standard 2D assignment algorithms to solve the data association.

5.2 Other proposed measures

The additive pairwise dissimilarity measure between m_i^{Θ} and m_j^{Θ}, defined by (30), has a theoretical justification via the plausibility of class equivalence, as explained earlier.

Several alternative pairwise dissimilarity (or distance) measures have been proposed recently in the framework of the belief function theory.

- Jousselme et al. [15] define a distance measure between m_i and m_j as follows:

$$d_{ij} \triangleq \sqrt{\frac{1}{2}\big(< m_i, m_i > + < m_j, m_j > -2 < m_i, m_j > \big)} \tag{31}$$

  where

$$< m_i, m_j > = \sum_{A \in \mathcal{F}} \sum_{B \in \mathcal{F}} m_i(A) m_j(B) \frac{|A \cap B|}{|A \cup B|} \tag{32}$$

is referred to as the scalar product between two bba's.

- The Euclidean distance is also defined by (31), but the scalar product takes the form:

$$< m_i, m_j >= \sum_{A \in \mathcal{F}} \sum_{B \in \mathcal{F}} m_i(A) m_j(B). \tag{33}$$

For Bayesian bba's, the Euclidean and the Jousselme distances are identical.

- Tessem [16] proposed a distance measure to quantify the quality of belief function approximations. His distance measure between m_i and m_j is defined as:

$$d_{ij} \triangleq \max_{\theta_\ell} |BetP_i(\theta_\ell) - BetP_j(\theta_\ell)| \tag{34}$$

where $BetP_i$ is the pignistic probability corresponding to bba m_i.

- Blackman and Popoli [6, p.620] argue that the distance between two bba's must in some way include the ignorance factor. Following [17] they define a pairwise additive dissimilarity measure as:

$$d_{ij}^B \triangleq -2\log[(1 - K_{ij})/(1 - K_{mij})] + I_i + I_j - I_j I_j \tag{35}$$

where $K_{ij} = m^{\Theta}_{1i \textcircled{\cap} 2j}(\emptyset)$, $K_{mij} = \max\{K_{ii}, K_{jj}\}$, and $I_i = \sum_{A \in \mathcal{F}} m_i(A)(|A|-1)/(|\Theta|-1)$ is the ignorance factor. Since distance d_{ij}^B can have negative values, we discard it as being unsuitable for our application.

6 Numerical simulations

Numerical simulations have been carried out in order to quantify the performance of the proposed method for multiple object classification. The performance analysis has been

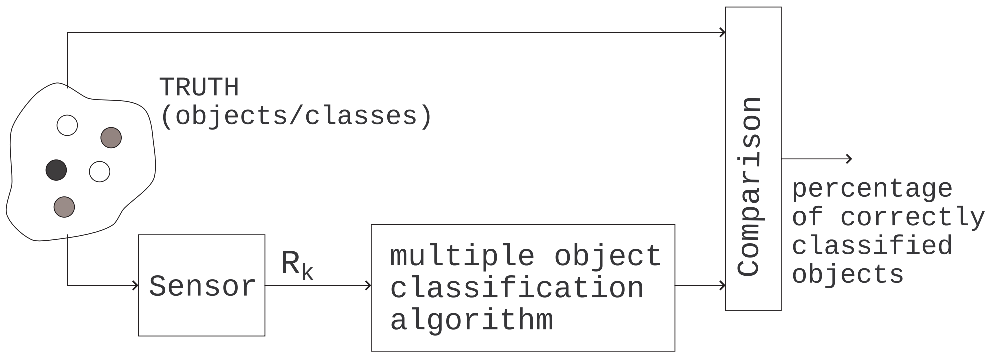

Figure 3: *Performance analysis scheme*

carried out using a number of Monte Carlo trials in the general scheme shown in Figure 3. Two types of sensors are available for collecting a sequence of reports R_k. Sensor type 1 is providing non-specific bba characterisation of detected objects with cardinality being the sensor parameter. Sensor type 2 is providing purely Bayesian bba's (that is probability functions). The multiple object classification algorithm in Figure 3 has 5 different versions, all based on the GNN algorithm and the global dissimilarity measure given by (29).

Version C-dist is using in (29) the TBM pairwise dissimilarity measure, defined by (30);

Version J-dist is using in (29) the Jousselme distance defined by (31) and (32);

Version E-dist is using in (29) the Eucledian distance, defined by (31) and (33);

Version T-dist is using in (29) the Tessem's distance (34);

Version Pignistic is somewhat different from the others, since it first applies the Pignistic transform (8) to the bba's of report R_k in order to transform them into the Bayesian bba's. Subsequently it applies the GNN algorithm for multi-object classification, using the Euclidean distance in (29). This version is effectively a Bayesian classifier, it is applying the probability theory to solve the problem, even when the sensor reports are non-specific pieces of evidence (recall statements S and $\hat{S}$ in Introduction). If the probability theory is indeed capable of dealing with non-specific measurements, this version of multi-object classification algorithm should not be worse then any other in the described experimental set-up.

Measurements from sensor 1 and 2 are generated in a random manner in order to carry out Monte Carlo simulations. Munkres algorithm has been applied to perform the 2D assignment.

6.1 Sensor 1: non-specific data

The bba's from Sensor 1 are always simple support functions, with 95% of the belief mass given to $A \subset \Theta$, and the remaining 5% to the universal set Θ. Subset A always contains the true class of a detected object, i.e. if o_i is detected and its class is $\theta(o_i) = \theta_n$ then $\theta_n \in A$. The cardinality of subset A is fixed and greater than 1; this is a sensor parameter. Other elements of A are selected at random with a uniform density.

Figure 4 shows the classification results for the five variations of the GNN algorithm (described above). The parameters used in simulations are as follows:

$$P_D = 0.3, \quad |\Theta| = 6, \quad O = \{1, 2, 3\} \quad \theta(1) = \theta_2,\ \theta(2) = \theta_3,\ \theta(3) = \theta_5. \tag{36}$$

The cardinality of focal set $A \subset \Theta$ was fixed at 3. The performance curves in Figure 4 represent the average percentage of correctly classified objects (PCCO), with averaging done over 500 Monte Carlo runs. The abscissa is index $k = 1, 2, \ldots$.

From Figure 4 we observe that version C-dist performs the best, followed by J-dist, T-dist and the Pignistic version. By far the worst is the performance of the E-dist version.

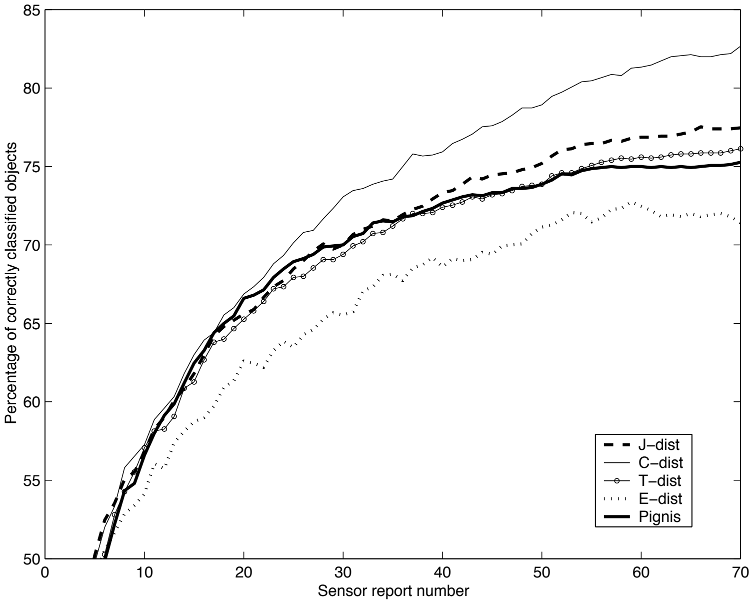

Figure 4: *Multi-object classification performance using non-specific sensor reports with cardinality* 3

The results confirm that the TBM similarity measure based on the plausibility of object-measurement pairings is the best for data association. We also note that the Bayesian classifier (represented by the version "Pignistic") is inadequate: the equi-probabilistic transformation of a non-specific piece of evidence to a specific piece of evidence is just an approximation which results in a degraded performance.

Similar results to those shown in Figure 4 were obtained for cardinality of focal set A equal to 2, 4, 5, etc. The cardinality only affects the rate of increase of the PCCO: the larger the cardinality the smaller the rate of increase.

6.2 Sensor 2: specific discordant data

The bba's from Sensor 2 are Bayesian bba's (that is probability distributions), simulated using the Dirichlet distribution [18]. The probability density of the Dirichlet distribution for variables $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N]$ is determined by a parameter vector $\mathbf{u} = [u_1, \ldots, u_N]$. This vector defines the mean and the variance of the distribution. Denoting $u_0 = \sum_{i=1}^{N} u_i$, we have:

$$E\{\theta_i\} = \frac{u_i}{u_0} \qquad \text{var}\{\theta_i\} = \frac{(u_0 - u_i)u_i}{u_0^2(u_0+1)} \tag{37}$$

for $i = 1, \ldots, N$. If object $o_\ell \in O$ has been detected, the components of parameter vector $\mathbf{u}$ are selected as follows:

$$u_i = \begin{cases} U_T & \text{if } \theta(o_\ell) = \theta_i \\ U_F & \text{otherwise} \end{cases} \tag{38}$$

where $U_T > U_F$ (otherwise the sensor would be worthless). In simulations we use $U_T = 5$ and $U_F = 1.8$. This type of sensor is sometimes specified by a confusion matrix [19]. In our case, the confusion matrix would be diagonal with the diagnosticity factor (the ratio of the diagonal and off-diagonal values) equal to $U_T/U_F \approx 2.78$.

Figure 5 shows the classification results obtained using the same parameters as before, specified by (36). We compare the same five variations of the GNN algorithm, using the PCCO as the performance measure. Note, however, that for sensor 2 which supplies the Baysian bba's, the J-dist, E-dist and the Pignistic version of the GNN algorithm become identical. The PCCO curves are obtained by averaging over 500 Monte Carlo runs.

From Figure 5 we conclude that once again the version C-dist performs the best, while this time the worst performance is obtained using the Tessem's distance. The result is somewhat surprising because the bba's reported by sensor 2 are Baysian bba's, and hence one would expect the probabilistic approach (version Pignistic) to perform at least as well as the C-dist version. However, it does not – even in this case it is advantageous (although to a lesser degree than in the case with Sensor 1) to use the plausibility (conflict) as a similarity (dissimilarity) measure in data association.

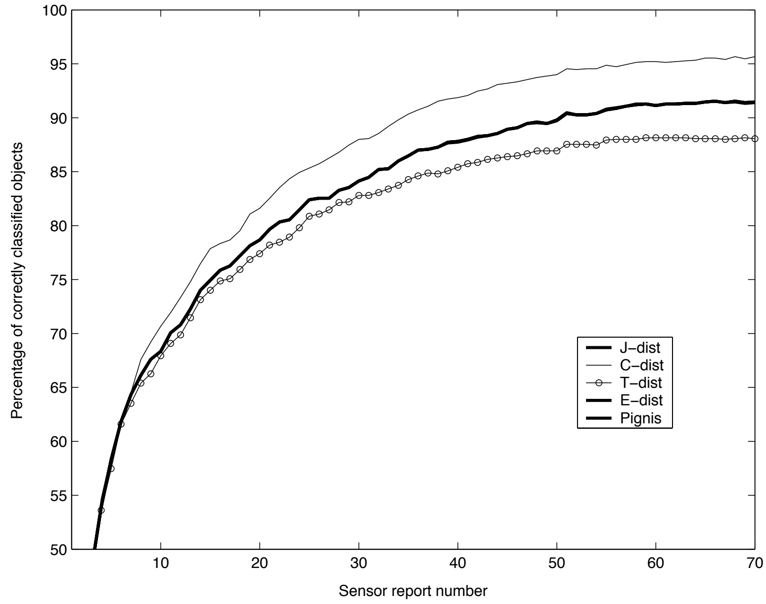

Figure 5: *Multi-object classification performance using specific discordant evidence reports*

We have further observed (figures not presented) that if we increase the diagnosticity factor of Sensor 2, all considered algorithms exhibit similar performance. Higher diagnosticity essentially means smaller discordance and therefore easier classification.

Similarly, for the classification of a single object using discordant but specific data, all considered algorithms perform the same. This observation is in the stark contrast to that reported in [19], because the authors of [19] carried out an unfair comparison: they did not provide the same input data to the Bayesian and to the belief function based classifier. From this unfair comparison they drew wrong conclusions[3]; for further explanation see [20].

[3]In [19], the Bayesian classifier was supplied with the Bayesian bba's (specific discordant data), while the belief function classifier was given nonspecific discordant data (simple support functions with two focal sets).

7 Conclusions

The paper studied the problem of multiple object classification based on discordant and non-specific data. A general methodology for solving this problem is suggested based on the transferable belief model, a model developed to represent uncertainty using the belief functions. A suboptimal single-scan algorithm, referred to as the the global nearest neighbour (GNN), is adopted for data association. The exact dissimilarity measure of a global assignment in the GNN context was derived. The performance of the GNN algorithm using several alternative dissimilarity measures was analysed by Monte Carlo simulations. One of the variants considered was the Bayesian GNN classifier. The results of this study suggest that the belief function theory based GNN classifier, which maximises the plausibility of the global assignment, performs by far the best among the considered alternatives. It appears that when dealing with non-specific and highly discordant data, multiple object classification using the Bayesian probabilistic approach is not appropriate.

Future work will extend the presented concepts to a multi-scan multi-object classification algorithm (as opposed to the single-scan GNN algorithm), via the concept of multidimensional data assignment [6, Sec.7.3],[21].

A A justification of the TBM similarity measure

There are many possible ways to define a global similarity measure in addition to the one given by the right-hand side of eq.(25). Some seemingly viable solutions, for example, are:

$$\mathcal{S}_1 = 1 - \sum_{i=1}^{n} m^{\Theta}_{1i \bigcirc\!\!\!\!\cap\, 2v_i}(\emptyset) \quad \text{or} \quad \mathcal{S}_2 = 1 - \max_i \left\{ m^{\Theta}_{1i \bigcirc\!\!\!\!\cap\, 2v_i}(\emptyset) \right\}.$$

By working out in detail the example of Table 2 on a Θ^5 space, we derive the score of the assignment which happens to be exactly the same as the similarity measure of eq.(25), proposed by Theorem 5.1. In this way we justify the choice of the global similarity measure based on the plausibility of a global assignment. We could derive the same result for the general case, but the equations would become very messy and difficult to follow. We feel that working out an example is more readable and convincing.

Let $S = \{1, 2\}$ and $\mathcal{F}_{si} = \{T_{si\nu} : \nu = 1, ..., N_{si}\}$ be the set of focal sets of m_{si}^{Θ}, where $s = 1, 2,\ i = 1, \ldots, n_s$, and m_{si}^{Θ} is the bba reported by sensor s about i-th object that it detected. Let $\mathcal{T}_{Sl} = \{T_{si\nu_{si}} : s \in S, i = 1, ..., n_s\} \in \Theta^{n_1+n_2}$ be a set of focal elements, one per bba. The five bba's from Example 2 are produced by distinct pieces of evidence and we can thus combine their vacuous extension on Θ^5 by the conjunctive combination rule:

$$m_S^{\Theta^5} = \textcircled{\cap}_{s=1,2} \textcircled{\cap}_{i=1,\ldots,n_s} m_{si}^{\Theta \uparrow \Theta^5}.$$

Consider the product term

$$m_S^{\theta^5}(\mathcal{T}_{Sl}) = \prod_{s \in S} \prod_{i=1}^{n_s} m_{si}^{\Theta}(T_{si\nu_{si}}) \tag{39}$$

where $S = \{1, 2\}$ and $l = 1, 2..., L = \prod_{s \in S} \prod_{i=1}^{n_s} N_{si} = 8$.

This mass $m_S^{\Theta^5}(\mathcal{T}_{Sl})$ supports that the actual classes of the objects o_{si} belong to $T_{si\nu_{si}}$, respectively:

$$\theta(o_{si}) \in T_{si\nu_{si}};\ \ s \in S, i = 1, \ldots, n_s.$$

The bba $m_S^{\Theta^5}$ is the joint belief about the actual classes of the five objects before considering they can be associated.

Next we define a compatibility matrix at the focal set level. Fix l in $\mathcal{T}_{Sl} = \{T_{si\nu_{si}} : s \in S, i = 1, ..., n_s\}$ where $T_{si\nu_{si}}$ is a focal element of m_{si}^{Θ}. A compatibility matrix of the focal sets $C_{Sl} = [c_{ij}]$ is a $n_1 \times n_2$ matrix where $c_{ij} \in \{0, 1\}$ and $c_{ij} = 1$ if $T_{1i\nu_{1i}} \cap T_{2j\nu_{2j}}$ and zero otherwise.

Table 5 presents the 8 masses computed from relation (39), and for each of them their corresponding compatibility matrices. In all cases except the fifth with mass 0.036, we need $n = 3$ objects. For the fifth case, $n = 4$ objects are needed as o_{11} is not compatible with any of the three objects o_{2j}.

With 3 objects, we can generate 6 assignment vectors. For the 4 object case, we need 3 assignment vectors (the other correspond just to a permutation of objects o_{13} and o_{14} which do not exist in $\mathbb{O}_1$, and the only focal set of their bba is Θ. The set of possible assignment vectors based on $n = 3$ objets and those relevant for the case with $n = 4$ objects are listed in Table 6.

Table 7 presents for each mass and each assignment vector, the focal set of the 3 or 4 final objects. For example, consider assignment number 3 (assignment vector $V(3) = (2, 1, 3)$, in shortened notation 213). The mass $m(1) = 0.336$, for assignment 3, is given to $\theta(o_{s1}) = \theta_3$ and $\theta(o_{s2}) = \theta_1$ and $\theta(o_{s3}) = \theta_1$ or θ_2. Similarly mass $m(5) = 0.144$, given assignment 7 (4123) is given to $\theta(o_{s1}) = \theta_3$ and $\theta(o_{s2}) = \theta_1$ or θ_4 or θ_5 and $\theta(o_{s3}) = \theta_1$ or θ_2 and $\theta(o_{s4}) = \theta_2$. Mass $m(1) = 0.336$ under assignment 1 (123) is allocated to the empty set. The bottom row of Table 7 represents the conflict corresponding to each of the nine association vectors, and computed by summing the masses $m(i)$ corresponding to the empty sets. For example, the conflict of $V(1) = (1, 2, 3)$ is a sum $0.336 + 0.084 + 0.144 + 0.036 = 0.60$. It is important to note that associations $V(4) = (3, 1, 2)$ and $V(6) = (3, 2, 1)$ are characterised by the smallest amount of conflict among those with $n = 3$ objects. This, however, is exactly the same result as the one obtained by Theorem 5.1 and described in Example 2; the appendix

Table 5: *The eight compatibility matrices with their masses and the list of assignment vectors $V(\ell)$ induced by them and listed in Table 6*

$m = .336$	θ_3, θ_5	$\theta_1, \theta_4, \theta_5$	θ_1, θ_2		$V(\ell)$
θ_1, θ_2	0	1	1		3 4 6
θ_3, θ_4	1	1	0		7 8
$m = .084$	θ_3, θ_5	$\theta_1, \theta_4, \theta_5$	θ_4, θ_5		
θ_1, θ_2	0	1	0		3 5
θ_3, θ_4	1	1	1		7 8 9
$m = .224$	Θ	$\theta_1, \theta_4, \theta_5$	θ_1, θ_2		
θ_1, θ_2	1	1	1		1 3 4 6
θ_3, θ_4	1	1	0		7 8
$m = .056$	Θ	$\theta_1, \theta_4, \theta_5$	θ_4, θ_5		
θ_1, θ_2	1	1	0		1 2 3 5
θ_3, θ_4	1	1	1		7 8 9
$m = .144$	θ_3, θ_5	$\theta_1, \theta_4, \theta_5$	θ_1, θ_2		
θ_2	0	0	1		1 2 3 5
θ_3, θ_4	1	1	0		7 8 9
$m = .036$	θ_3, θ_5	$\theta_1, \theta_4, \theta_5$	θ_4, θ_5	Θ	
θ_2	0	0	0	1	
θ_3, θ_4	1	1	1	1	7 8 9
$m = .096$	Θ	$\theta_1, \theta_4, \theta_5$	θ_1, θ_2		
θ_2	1	0	1		1 4 6
θ_3, θ_4	1	1	0		7 8
$m = .024$	Θ	$\theta_1, \theta_4, \theta_5$	θ_4, θ_5		
θ_2	1	0	0		1 2
θ_3, θ_4	1	1	1		7 8 9

Table 6: *The overall score for each assignment vector* $V(\ell)$

ℓ	$\prod_{j=1}^{3}(1 - m_{jv_j}(\emptyset))$	assignment vector $V(\ell)$			
1	0.40	1	2	3	
2	0.08	1	3	2	
3	0.70	2	1	3	
4	0.80	3	1	2	
5	0.14	2	3	1	
6	0.80	3	2	1	
7	1.00	4	1	2	3
8	1.00	4	2	1	3
9	0.20	4	3	1	2

explains the origin of the criterion used in (25).

Acknowledgments

The authors would like to thank Prof. T. Denoeux for useful technical discussions.

References

[1] G. J. Klir and M. J. Wierman, *Uncertainty-based information: Elements of generalized information theory.* New York: Physica-Verlag, 2nd ed., 1999.

[2] R. Kruse, E. Schwecke, and J. Heinsohn, *Uncertanty and vagueness in knowledge based systems.* Berlin: Springer-Verlag, 1991.

[3] G. Shafer, *A mathematic theory of evidence.* Princeton University Press, 1976.

[4] I. R. Goodman, R. P. S. Mahler, and H. T. Nguyen, *Mathematics of Data Fusion.* Kluwer, 1997.

Table 7: *For each possible association vector $V(1)$ to $V(9)$, and each focal set on Θ^3 or Θ^4, we present the values (indexes of the $\theta_n, n = 1, \ldots, 5$) of each of the three or four final objects o_{Sj}. A blanks triple or quadruple in column means an empty set*

		Assignment vector (see Table 6)								
i		1	2	3	4	5	6	7	8	9
$m(i)$	o_{Sj}	123	132	213	312	231	321	4123	4213	4312
	o_{S1}			3	3		35	3	35	
1	o_{S2}			1	145		4	145	4	
.336	o_{S3}			12	12		12	12	12	
	o_{S4}							12	12	
	o_{S1}			3		35		3	35	35
2	o_{S2}			1		1		145	4	145
.084	o_{S3}			45		4		45	45	4
	o_{S4}							12	12	12
	o_{S1}	12		34	34		12345	34	12345	
3	o_{S2}	4		1	145		4	145	4	
.224	o_{S3}	12		12	12		12	12	12	
	o_{S4}							12	12	
	o_{S1}	12	12	34		12345		34	12345	12345
4	o_{S2}	4	145	1		1		145	4	145
.056	o_{S3}	45	4	45		4		45	45	4
	o_{S4}							12	12	12
	o_{S1}				3		35	3	35	
5	o_{S2}				145		4	145	4	
.144	o_{S3}				2		2	12	12	
	o_{S4}							2	2	
	o_{S1}							3	35	35
6	o_{S2}							145	4	145
.036	o_{S3}							45	45	4
	o_{S4}							2	2	2
	o_{S1}	2			34		12345	34	12345	
7	o_{S2}	4			145		4	145	4	
.096	o_{S3}	12			2		2	12	12	
	o_{S4}							2	2	
	o_{S1}	2	2					34	12345	12345
8	o_{S2}	4	145					145	4	145
.024	o_{S3}	45	4					45	45	4
	o_{S4}							2	2	2
conflict		0.60	0.92	0.30	0.20	0.86	0.20	0.00	0.00	0.80

[5] D. Dubois and H. Prade, "A note on measures of specificity for fuzzy sets," *Intern. Journal General Systems*, vol. 4, pp. 279–283, 1985.

[6] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems.* Artech House, 1999.

[7] N. Milisavljevic and I. Bloch, "Sensor fusion in anti-personnel mine detection using a two-level belief function model," *IEEE Trans. Systems, Man and Cybernetics – Part C: Appl. and Reviews*, vol. 33, no. 2, pp. 269–283, 2003.

[8] A. Ayoun and P. Smets, "Data association in multi-target detection using the transferable belief model," *Intern. Journal of Intelligent Systems*, vol. 16, no. 10, pp. 1167–1182.

[9] J. Schubert, "Managing inconsistent intelligence," in *Proc. 3th Int. Conf. on Information Fusion*, (Paris, France), 2000.

[10] P. Smets and R. Kennes, "The transferable belief model," *Artificial Intelligence*, vol. 66, pp. 191–234, 1994.

[11] P. Smets, "The transferable belief model for quantified belief representation," in *Handbook of Defeasible Reasoning and Uncertainty Management Systems* (D. M. Gabbay and P. Smets, eds.), vol. 1, pp. 267–301, Kluwer, Doordrecht, The Netherlands, 1998.

[12] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association.* Academic Press, 1988.

[13] T. Denoeux and M.-H. Masson, "EVCLUS: Evidential clustering of proximity data," *IEEE Trans. Systems, Man and Cybernetics – Part B*, vol. 34, pp. 95–109, Feb. 2004.

[14] J. Schubert, "On non-specific evidence," *Intern. Journal of Intelligent systems*, vol. 8, pp. 711–725, 1993.

[15] A.-L. Jousselme, D. Grenier, and Éloi Bossé, "A new distance between two bodies of evidence," *Information Fusion*, vol. 2, pp. 91–101, 2001.

[16] B. Tessem, "Approximations for efficient computation in the theory of evidence," *Artificial Intelligence*, vol. 61, pp. 315–329, 1993.

[17] H. E. Stephanou and S.-Y. Lu, "Measuring consensus effectiveness by a generalised entropy criterion," *IEEE Trans. Pattern Analysis and machine Intelligence*, vol. 10, pp. 544–554, July 1988.

[18] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*. Boca Raton: Chapman & Hall/CRC, 1995.

[19] D. M. Buede and P. Girardi, "A target identification comparison of Bayesian and Dempster-Shafer multisensor fusion," *IEEE Trans. Systems, Man and Cybernetics – Part A: Systems and Humans*, vol. 27, no. 5, pp. 569–577, 1997.

[20] F. Delmotte and P. Smets, "Target identification based on the transferable belief model interpretation of Dempster-Shafer model," *IEEE Trans. Systems, Man and Cybernetics - Part A*, 2004. (Accepted for publication).

[21] A. B. Poore, N. Rijavec, M. Liggins, and V. C. Vannicola, "Data association problems posed as multidimensional assignment problems: Problem formulation," in *Proc. SPIE*, vol. 1954, pp. 552–563, 1993.