HAL
open science

The Correlation Problem in Sensor Fusion in a Possibilistic Framework

Michel Grabisch, Henri Prade

To cite this version:

Michel Grabisch, Henri Prade. The Correlation Problem in Sensor Fusion in a Possibilistic Framework. International Journal of Intelligent Systems, 2001, 16 (11), pp.1273-1283. 10.1002/int.1059 . halshs-00273189

HAL Id: halshs-00273189

https://shs.hal.science/halshs-00273189v1

Submitted on 14 Apr 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Correlation Problem in Sensor Fusion in a Possibilistic Framework

Michel GRABISCH, Henri PRADE

Abstract

This paper addresses the correlation problem which is central in sensor fusion, from the viewpoint of possibility theory. This problem aims at separating pieces of information pertaining to different objects and to gather those which are likely to pertain to the same object. We present two different views of the problem, one based on similarity relations, while the other discusses the problem in a logical framework.

1 Introduction

This paper addresses the *correlation* (or *association*) problem in sensor fusion and can be seen as a complement to the introductory paper about the same topic by Ayoun and Smets [1] in this special issue. The problem is to decide which observations are pertaining to the same object in a scene. As already said [1], this is a central topic in most of practical sensor fusion problems.

We present here solutions based on fuzzy relations and possibility theory, that is, which require a relatively low level of information quality, in contrast to probabilistic solutions more widely used in such applications. In particular, the approaches we present here may work with only qualitative data, or a mix of numerical and qualitative data.

We assume that the reader has some familiarity with possibility theory and fuzzy sets (see [6] for a thorough introduction on the topic). Section 2 presents the problem, and gives the main working assumptions and notations. Section 3 presents the approach based on fuzzy relations, more precisely on proximity relations. See [10, 15] for a more detailed presentation of this approach, together with a larger view of the data fusion problem, and an application to battlefield intelligence. Section 4 presents a logical view of the same problem.

2 The correlation problem

We begin by introducing some vocabulary and notations. Let us call *unit* any existing object in the scene, for which a track history exists. We denote by $E_1, \dots, E_p$ the list of units which are present in the scene. We do not take here into account the time aspect, which is not important for the correlation problem.

Each unit E_k is described by a set of attributes $X_1, \dots, X_n$, taking values in universes $U_1, \dots, U_n$ respectively. Each universe can be an interval of the real line (or even R^q), or a finite set, with or without an ordering relation on it.

Knowledge on E_k is described by a possibility distribution on the Cartesian product of $U_1, \ldots, U_n$, assuming for simplicity that the attributes are non-interactive, so that knowing the marginal distributions is sufficient. We denote by π_i^k the possibility distribution of the values of attribute X_i concerning unit E_k.

We suppose to receive s observations $O_1, \ldots, O_s$ coming from the different sensors. These observations are described by the same attributes $X_1, \ldots, X_n$, and are modelled by a set of marginal possibility distributions π_i^j, $i = 1, \ldots, n$, $j = 1, \ldots, s$. This notation points out that the information gathering system is such that we are sure that the pieces of information represented by π_i^j for $i = 1, \ldots, n$ pertain to the same observation j which is thus relative to a unique item.

The correlation step consists in the assignment of observations to existing units. This implies that the system is able to regroup all observations which are believed to belong to the same unit, to create new units, and to merge or to delete existing units if necessary. In other words, it amounts at creating the "best" partition of $E_1, \ldots, E_p, O_1, \ldots, O_s$, in the sense that the subsets in the partition should contain only similar elements, and elements belonging to different subsets should be dissimilar as much as possible.

As already announced, we address this problem with two different approaches. The fuzzy relation approach is presented first.

3 The approach based on fuzzy relations

The present approach is based on the fact that *similarity* seems to be the central concept in the process of partitioning. We divide this process in two steps:

- Constructing a proximity matrix for observations and units,
- Exploiting the proximity matrix in order to partition the objects in it.

We detail the two steps below. Before, we introduce basic material on fuzzy relations.

3.1 Basic concepts on fuzzy relations

We begin by explaining briefly some concepts of fuzzy relations, in order to be self-contained. In the rest of the section, $\wedge$ and $\vee$ respectively denote min and max.

Let us consider a universe Ω. A *fuzzy relation* R or *valued binary relation* on Ω is a fuzzy set on $\Omega \times \Omega$, i.e. a function $R : \Omega^2 \mapsto [0, 1]$ assigning to every pair of elements in Ω a number expressing the strength of the relation between them. By convention, $R(a, b) = 1$ indicates the strongest relation between a and b, while $R(a, b) = 0$ means that there is no relation between a and b.

The following properties of fuzzy relations will be useful in the sequel:

- R is *reflexive* if $R(a, a) = 1$, $\forall a \in \Omega$,
- R is *symmetric* if $R(a, b) = R(b, a)$, $\forall a, b \in \Omega$,

- R is $\top$-*transitive* for a given t-norm $\top$ if

$$R(a,c)\top R(c,b) \leq R(a,b), \tag{1}$$

where a t-norm is any associative, commutative binary operation on $[0,1]$, increasing in each place, and such that $1\top 1 = 1$ and $0\top a = 0$.

All these definitions are extensions of classical ones on binary relations.

A fuzzy relation being reflexive, symmetric and $\top$-transitive is called a $\top$-*similarity relation.* It is the extension of the classical notion of equivalence relation. A fuzzy relation being only reflexive and symmetric is called a *proximity relation.*

Let R, S be two fuzzy relations on Ω, and let $\top$ be a t-norm. The $\top$-*composition* of R and S is a fuzzy relation on Ω defined by:

$$(R \circ_\top S)(a,b) := \sup_{c\in\Omega}(R(a,c)\top S(c,b)). \tag{2}$$

Again, this is an extension of the classical notion of composition of binary relations.

Let R be a fuzzy relation. The *transitive closure* of R is the fuzzy relation defined by:

$$\hat{R}(a,b) = \sup_m R^m(a,b), \quad \forall a,b, \in \Omega \tag{3}$$

where $R^m := \underbrace{R \circ_\top R \circ_\top \cdots \circ_\top R}_{m \text{ times}}$. As the name suggests, it is easy to verify that $\hat{R}$ is $\top$-transitive, and that $\hat{R} = R$ if and only if R is $\top$-transitive. Moreover, it can be shown that $\hat{R}$ is the intersection of all $\top$-transitive relations S containing R, i.e. such that $S(a,b) \geq R(a,b)$ for every $a,b \in \Omega$. $\hat{R}$ is also the nearest $\top$-transitive relation containing R, in the sense of the Hamming distance defined by:

$$d(R,S) = \sum_{a,b\in\Omega} |R(a,b) - S(a,b)|. \tag{4}$$

Let R be a fuzzy relation on Ω, not necessarily $\top$-transitive. For a given $\alpha \in [0,1]$, one can compute the classical binary relation R_α defined by $R_\alpha(a,b) := 1$ if $R(a,b) \geq \alpha$, and 0 otherwise. Considering R_α, which is not necessarily transitive, one can look for transitive subrelations of R_α, i.e. classical binary relations defined on a subset of Ω which are transitive (such relations always exist, since any subrelation defined on a singleton of Ω will do). In order to avoid redundancies, only *maximal transitive subrelations* are of interest, i.e. transitive subrelations defined on a subset A of Ω, such that there is no $B \subset \Omega$ which contains A and for which R_α restricted to B is transitive. Finding all such subrelations is not an easy task (see [11] for an algorithm).

3.2 Constructing a proximity matrix

The aim of this step is to build a proximity matrix for all the observations $O_1, \ldots, O_s$ and present units $E_1, \ldots, E_p$. The object k will be denoted o_k when we do not wish to distinguish between unit and observation. The set of

all objects will be denoted $\mathcal{O} := \{o_1, \ldots, o_{s+p}\}$. The word *proximity* is more suitable than *similarity* as it will be explained below.

Let us denote M this matrix, and $m_{kl} \in [0,1]$ one of its terms, $k, l = 1, \ldots, s+p$. The term m_{kl} represents the degree of proximity between the concerned objects (observations or units) k and l. The value $m_{kl} = 1$ indicates the maximum of proximity.

Matrix M is obtained by aggregating proximity matrices $M_i = [m^i_{kl}]_{1 \leq k,l \leq s+p}$ for attribute i, $i = 1, \ldots, n$. The aggregation may be done by a weighted minimum [7] :

$$m_{kl} = \bigwedge_{i=1}^{n} [(1 - w_i) \vee m^i_{kl}], \tag{5}$$

where the w_i's are weights satisfying $\vee_i w_i = 1$, expressing the importance of attribute i to decide if two objects are close or not. Remark that w_i should not reflect the confidence or uncertainty pertaining to the measurement of attribute i, since this should be already included in π^k_i. The importance rather reflects here how imperative it is that the observed values of an attribute be similar, if they pertain to the same object. Indeed if the attribute may vary with time, similarity is no longer absolutely compulsory. This way of aggregation expresses the fact that two objects are close if and only if there is no important attribute for which the objects are not close. Of course, other ways of aggregation might be considered in particular problems .

The proximity degree m^i_{kl} is defined as the degree of intersection between the respective distributions, i.e.

$$m^i_{kl} := \sup_{x_i \in U_i} (\pi^k_i(x_i) \wedge \pi^l_i(x_i)) \tag{6}$$

where π^k_i is the possibility distribution of object k. It amounts at evaluating how much the pieces of information k and l are consistent w.r.t. attribute i.

This way of computing similarities may appear not enough discriminant in many situations. In order to solve ambiguities, on can use a second comparison index, which reflects the portion of the attribute domain where the two possibility distributions agree on the fact that a particular element of the universe is, or is not, a possible value for the attribute. Considering that the universe U_i is discretized at $x^1_i, \ldots, x^q_i$, the degree of agreement $\mu_{kl}(x^j_i)$ on x^j_i between the two distributions is computed as a multiple-valued logic equivalence of the form $s \leftrightarrow t = (\neg s \vee t) \wedge (\neg t \vee s)$, where $\neg$ denotes negation, namely:

$$\mu_{kl}(x^j_i) = (\pi^k_i(x^j_i) \vee (1 - \pi^q_i(x^j_i))) \wedge ((1 - \pi^k_i(x^j_i)) \vee \pi^q_i(x^j_i)) \tag{7}$$

This in fact defines the fuzzy set of agreement on U_i. Considering now the vector $[\mu_{kl}(x^1_i) \cdots \mu_{kl}(x^q_i)]$, we can order pairs of distributions by ordering the corresponding vectors using, e.g., a leximin ordering. For two vectors a, b, the leximin ordering is defined by:

$$a \succ_{\text{lex}} b \Leftrightarrow \exists k \leq n, a_{(i)} = b_{(i)}, \forall i \in \{1, \ldots, k-1\}, \text{ and } a_{(k)} > b_{(k)}, \tag{8}$$

with $a_{(1)} \leq a_{(2)} \leq \cdots \leq a_{(n)}$ and similarly for b. This ordering can be used for discriminating between pairs having the same value for m^i_{kl}. Another mean

would be to compute a representative value for the fuzzy set of agreement, which can be used as a second degree of proximity. Such a refinement might be used for refining the clustering process described in the next section.

Squared matrices as M, M_i can be viewed as fuzzy relations (see section 3.1). A reflexive fuzzy relation will correspond to a matrix with only 1's on the diagonal, while a symmetric relation will correspond to a symmetric matrix, etc. By construction, the matrices M_i and M are reflexive, symmetric, but not $\wedge$-transitive. According to the standard terminology [9], these are proximity matrices, and not *similarity* matrices (see section 3.1).

3.3 Exploiting the proximity matrix

In this step, we exploit the proximity matrix M in order to get a partition $\Pi = \{P_1, \ldots, P_r\}$ of the objects, so that objects inside a given P_l are believed to be similar and to refer to the same unit.

We can think of two different approaches for solving this problem: either we perform a kind of step-by-step clustering of close objects, or we try to obtain a similarity matrix (i.e. $\wedge$-transitive) from M (we remark that anyway, to any partition corresponds a transitive matrix). We detail both approaches hereafter, and a third one.

3.3.1 Hierarchical clustering of the objects

The first approach is to perform a hierarchical clustering of the objects by a linkage method (see e.g. Miyamoto [13]). The initial partition Π_0 is the finest one, i.e. each element of the partition is a single object. This partition is thus $\Pi_0 = \{\{o_1\}, \ldots, \{o_{s+p}\}\}$. The next step is to find the two closest objects, i.e. such that m_{kl} is maximum. Let o_{k_0}, o_{l_0} be these two objects, we withdraw them from the partition Π_0 and introduce a new element $P_{r'} = \{o_{k_0}, o_{l_0}\}$. In other words, we have merged the two closest objects. In the new partition obtained Π_1, we perform again the same operation: we look for the two closest elements in the partition and merge them. More specifically, at an arbitrary step i, if we denote by P_{k_0}, P_{l_0} the two closest elements in Π_i, we have:

$$\Pi_{i+1} = (\Pi_i \setminus \{P_{k_0}, P_{l_0}\}) \cup \{P_{k_0} \cup P_{l_0}\}. \tag{9}$$

This operation is repeated until the coarsest partition is obtained, i.e. $\{\{o_1, \ldots, o_{s+p}\}\}$.

The remaining problem is to compute the proximity $s(P_k, P_l)$ between two elements P_k, P_l of a partition, i.e. between two subsets of objects. This can be done by several methods [13], among which:

- Single linkage

$$s(P_k, P_l) = \bigvee_{o_i \in P_k, o_j \in P_l} m_{ij}, \tag{10}$$

- Complete linkage

$$s(P_k, P_l) = \bigwedge_{o_i \in P_k, o_j \in P_l} m_{ij}, \tag{11}$$

- Average linkage

$$s(P_k, P_l) = \frac{1}{|P_k| \cdot |P_l|} \sum_{o_i \in P_k, o_j \in P_l} m_{ij}, \tag{12}$$

where $|P_k|$ indicates the cardinal of P_k. Clearly, in our case, the most natural way would be the complete linkage, which means that two subsets are close to each other when all their corresponding elements are close.

Storing all the partitions obtained during the process together with the level of proximity of the closest elements, we obtain a kind of tree with threshold levels, called a dendrogram (see e.g. [11, 13]), which gives all possible merging of objects according to the level of proximity. Fixing beforehand a level of proximity gives a unique partition of the objects.

3.3.2 Closest $\wedge$-transitive matrix

The second approach consists in looking for the closest (in some sense) $\wedge$-transitive matrix of M. Every such matrix defines a dendrogram, and we are naturally looking for the closest one to M. There are several methods to do this, we describe briefly the main ones below (see [9, 8] for more details, and see 3.1 for definitions).

- Computation of the transitive closure of M, denoted $\hat{M}$. As said in section 3.1, $\hat{M}$ is the smallest $\wedge$-transitive matrix above M, and the nearest one is the sense of the Hamming distance. However, it is known that $\hat{M}$ may considerably overestimate some similarity values, and this is not desirable in our case.

- Computation of all maximal transitive lower matrices, denoted $\check{M}$. There is not a unique maximal transitive lower matrix in general, and as far as the authors know, there is no algorithm to find them all, except in the case of proximity relations where a general algorithm has been found by Leclerc [12]. See Fodor and Roubens [8] for an algorithm finding some of the maximal solutions in the general case. Usually, there exists solutions which are nearer to M than $\hat{M}$ in the sense of the Hamming distance. However, the computation is rather complex.

3.3.3 Maximal transitive subrelations

A third approach is based on the computation of all maximal transitive subrelations (submatrices) of M (see section 3.1 for definitions), for a given threshold. Doing this, we obtain a covering of the set of objects $\mathcal{O}$, i.e. a set of subsets $C_1, \ldots, C_m$ such that $\cup_{i=1}^m C_i = \mathcal{O}$. The subsets C_i correspond to the domains of the different subrelations.

In general, this covering is not a partition since subsets C_i may intersect themselves, and thus this does not define groups of objects to be merged. In order to obtain a partition, we compute the compacity of each group, defined as the average proximity among objects inside the group:

$$\kappa(C_i) := \frac{1}{|C_i|} \sum_{1 \leq j < k \leq |C_i|} m_{jk}. \tag{13}$$

The group with highest compacity, denoted C_0, is selected as the first element of the partition. Objects belonging to C_0 are withdrawn from all C_i's, and the same process is repeated, until the partition is completed.

The advantage of this method is that it does not overestimate proximity values, but as the threshold varies, it does not lead to a dendrogram.

4 The logical view

This section intends to provide a logical discussion of the correlation problem, and to offer some bridge between the fuzzy relation-based approach surveyed in section 3, and the method using belief functions proposed in [1]. We mainly consider representation issues, but not the algorithmic problems.

4.1 The possibilistic logic setting

Let K be a possibilistic logic base [4], $K = \{(p_j, \alpha_j), j = 1, \ldots, r\}$. Each pair (p_j, α_j) is made of a classical proposition p_j, and a level $\alpha_j \in [0, 1]$, which reflects how much it is held as certain that p_j is true. In other words, a pair (p_j, α_j) has to be understood as it is certain at level α_j that proposition p_j is true for some object under consideration. As usual, the literals appearing in the propositional formulas p_j of K, induce a finite set of interpretations W. The semantics of K is a possibility distribution which rank-orders the possible interpretations generated by the language according to their plausibility. Namely,

$$\forall w \in W, \pi_K(w) = \min_{j=1,\ldots,r} \max(\mu_{[p_j]}(w), 1 - \alpha_j) \tag{14}$$

where $[p_j] = \{w, w \models p_j\}$, i.e., the sets of interpretations w which make p_j true, and $\mu_{[p_j]}(w) = 1$ if $w \in [p_j]$ and $\mu_{[p_j]}(w) = 0$ otherwise. The equation (14) describes the fuzzy set of models of K. It expresses that w is a model of K at degree 1 if it is a model of each p_j, i.e., an interpretation that makes each p_j true. Otherwise, w is all the less possible as a model of K, as w makes false a formula p_j having a high degree of certainty α_j.

The degree of inconsistency of K is defined by

$$\mathrm{Inc}(K) = 1 - \max_w \pi_K(w) \tag{15}$$

It can be shown that $\mathrm{Inc}(K) = 0$ iff $K^* = \{[p_j], j = 1, \ldots, r\}$ is consistent in the classical sense. $\mathrm{Inc}(K)$ can be also directly computed from K, put in clausal form, by using an extended resolution rule, without calculating π_K [4]. Note that $1 - \mathrm{Inc}(K) = \max_w \pi_K(w)$ is a consistency degree, similar to m^i_{kl} given by (6), but here computed for the whole set of pieces of information in K, rather than on a pair of observations pertaining to the same attribute.

4.2 Partitioning in consistent subparts

If K is inconsistent, it may be due to the fact that the contents of K pertains to the description of different objects. Then we have to look for a partition of K into $K_1, \ldots, K_q$ such that $\forall i = 1, \ldots, q$ $\mathrm{Inc}(K_i) = 0$. Such a partition obviously exists: $K_1 = \{p_1\}, \ldots, K_r = \{p_r\}$ with $q = r$. However a principle

of parsimony suggests that the number of objects should remain as small as possible, i.e., q, the number of elements in the partition is as small as possible, keeping the constraint $\forall i, \mathrm{Inc}(K_i) = 0$. This corresponds to the idea of complete linkage expressed by (11). So we are led to look for maximal consistent subbases of K, from which it is possible to obtain partitions of K with q minimal.

For a given partition, its degree of acceptability is defined by

$$\mathrm{Acc}(K_1, \ldots, K_q) = 1 - \max_i \mathrm{Inc}(K_i). \tag{16}$$

Note that $\mathrm{Acc}(K_1, \ldots, K_q) = 1 \Leftrightarrow \forall i, \mathrm{Inc}(K_i) = 0$.

Example: Let us consider the following base.

$K = \{(p, 0.7), (p \wedge u, 0.8), (\neg v \vee \neg p, 1), (v, 0.6), (\neg p \wedge s, 0.6), (\neg s \vee \neg u, 1), (s, 0.9)\}$.

We may think of each proposition as describing possible locations of objects. The pieces of information $(\neg v \vee \neg p, 1)$ and $(\neg s \vee \neg u, 1)$ express that p and v on the one hand and u and s on the other hand are mutually exclusive locations. We can check that $\mathrm{Inc}(K) = 0.8$. Let us consider the partition $K = \{K_1, K_2\}$ with:

$$\begin{aligned} K_1 &= \{(p, 0.7), (p \wedge u, 0.8), (\neg v \vee \neg p, 1)\} \\ K_2 &= \{(\neg p \wedge v, 0.6), (\neg p \wedge s, 0.6), (\neg p \wedge s, 0.9), (\neg s \vee \neg u, 1)\} \end{aligned}$$

Clearly, $\mathrm{Inc}(K_1) = 0 = \mathrm{Inc}(K_2)$ and $\mathrm{Acc}(K_1, K_2) = 1$. In such a case we may conclude on the presence of two objects, one in the location $p \wedge u$ with certainty 0.8 (since $(p \wedge u, 0.8)$ subsumes $(p, 0.7)$), and one in location $\neg p \wedge v \wedge s$ with certainty 0.6 after combining the pieces of information in K_2. $\square$

Thus it is theoretically possible to rank-order all the partitions of K according to their level of acceptability, and for those having the same level of acceptability to refine the ranking by considering the partitions of smallest cardinality q first. However, as pointed out by the following very simple example, these different criteria may not be enough to discriminate between two partitions. Let us take $K' = \{p, \neg p, s\}$ (where s is consistent with both p and $\neg p$), then the two partitions $(\{p, s\}, \{\neg p\})$ and $(\{p\}, \{s, \neg p\})$ are candidates. Is the information s refining the location of the object in p, or of the one in $\neg p$? We may use the following heuristic principle for deciding. As suggested in the previous section, the only way for discriminating between them in such a case is to compare the cardinalities of the sets of models of $p \wedge s$ and of $\neg p \wedge s$.

4.3 Dealing with the correlation problem

As the first example already suggests it, it is important to distinguish between two parts into K, namely, the part L with pertains to observations, and the part I which rather describes our general knowledge about the observed world (e.g., mutual exclusiveness of locations). Thus, let $K = L \cup I$ (with $L \cap I = \emptyset$)

and what we have to do is to partition L into $L_1, \ldots, L_r$ in such a way that $\mathrm{Inc}(L_i \cup I) = 0$ for all i.

In practice the formulas in L can be sorted according to the attribute to which the proposition pertains. Then for each attribute j we may look for a partition such that $\mathrm{Acc}(L_i^j \cup I) = 1$ where L_i^j gathers propositions pertaining to j only, and then build a partition of L by gathering the pieces of information pertaining to the attributes of the same object. So this approach is just the logical counterpart of (5)-(6)-(11), when all the attributes have the same importance (i.e. $\forall i, w_i = 1$ in (5)). Note that any possibility distribution π_i^j considered in Section 3 can be represented by a finite set of possibilistic formulas (see, e.g. [14]), once the possibility scale has been discretized if necessary in a finite number of levels. This explains the close relationship between the approaches of Sections 3 and 4.

Since possibility distributions are consonant belief structures [17], the approach discussed here is also similar to the one proposed by Schubert [16] using belief functions and presented in [1], as suggested by the following elementary example:

$$L = \{(a, 0.7), (a, 0.8), (b, 0.4), (b, 0.4), (b, 0.9)\}$$

$$I = \{(\neg a \vee \neg b, 1)\}$$

Here L represents five observations, the two first ones say that "there is something in a", the three last ones "something in b", while I says that a and b are mutually exclusive. Clearly the only fully acceptable partition here is $L_1 = \{(a, 0.7), (a, 0.8)\}$ and $L_2 = \{(b, 0.4), (b, 0.4), (b, 0.9)\}$. In standard possibilistic logic, there is no reinforcement effect, only the strongest piece of information is taken into account, e.g., L_1 is equivalent to $\{(a, 0.8)\}$ and L_2 to $\{(b, 0.9)\}$. However reinforcement effects can be also introduced in the symbolic setting of possibilistic logic for data fusion processes (see [2] for details). It amounts to change $\wedge$ into a product or other operations in (6) and in (14).

Besides, in constrast with what is done in Section 3 where descriptions of already observed items (the E_k's) and new observations (the O_l's) are dealt with together, the logical view may suggest to deal first with the new observations, before matching them with the descriptions of already known objects, in order to identify the observations or to decide if some of them correspond to new objects (which is rather a diagnostic problem, see [5]).

Lastly, if by construction it is known that several pieces of information pertain to the same object although they are somewhat conflicting, the level of inconsistency of the greatest internal conflict of this kind provides a threshold θ strictly greater than 0 which indicates that the degree of acceptability of the partition we are looking for should be $1 - \theta$ rather than 1, since an inconsistency at level θ should be tolerated (see Ayoun and Smets [1] for similar concerns).

5 Concluding remarks

This note has advocated possibility theory as a qualitative setting which may be sufficient in practice for solving the correlation problem in sensor fusion. Two related modelings, the fuzzy relation-based approach and the possibilistic logic view have been presented. Note that here the idea of similarity used in section 3 does not refer to the idea of distance but is just a matter of consistency.

The problem could be also contrasted with the handling of inconsistent knowledge bases in a reasoning perspective [3], where it is assumed that inconsistency is due to accumulation of information coming from multiple (non fully reliable) sources, rather than to the fact that the information may refer to several objects.

References

[1] A. Ayoun and P. Smets. Associating measurements from multiple sensors. *Int. J. of Intelligent Systems*, this issue.

[2] S. Benferhat, D. Dubois and H. Prade. From semantic to syntactic approaches to information combination in possibilistic logic. *Aggregation and Fusion of Imperfect Information*, (B. Bouchon-Meunier, ed.) Physica Verlag, 141–161, 1998.

[3] S. Benferhat, D. Dubois and H. Prade. Some syntactic approaches to the handling of inconsistent knowledge bases: A comparative study. Part 2: The prioritized case. *Logic at Work*, (E. Orlowska, ed.) Physica Verlag, 473–511, 1999.

[4] D. Dubois, J. Lang and H. Prade. Automated reasoning using possibilistic logic: semantics, belief revision, and variable certainty weights. *IEEE Transactions on Knowledge and Data Engineering*, 6:64–71, 1994.

[5] D. Dubois, M. Grabisch, O. de Mouzon and H. Prade. A possibilistic framework for causal diagnosis under uncertainty. *Int. J. of General Systems*, to appear.

[6] D. Dubois and H. Prade. *Possibility Theory.* Plenum Press, 1985.

[7] D. Dubois and H. Prade. Weighted minimum and maximum operations in fuzzy set theory. *Information Sciences*, 39:205–210, 1986.

[8] J. Fodor and M. Roubens. Structure of transitive valued binary relations. *Mathematical Social Science*, 30:71–94, 1995.

[9] J.C. Fodor and M. Roubens. *Fuzzy Preference Modelling and Multi-Criteria Decision Aid.* Kluwer Academic Publisher, 1994.

[10] M. Grabisch and P. Savéant. Uncertainty modeling and its linguistic expression in data fusion systems. In *Int. IEEE Conf. on Fuzzy Systems (FUZZ-IEEE'98)*, pages 921–926, Anchorage, Alaska, May 1998.

[11] A. Kaufmann and E. Pichat. *Méthodes Mathématiques Non Numériques et Leurs Algorithmes.* Masson, 1977.

[12] B. Leclerc. Caractérisation, construction et dénombrement des ultramétriques supérieures minimales. *Stat. Anal. Données*, 11:26–50, 1986.

[13] S. Miyamoto. *Fuzzy Sets in Information Retrieval and Cluster Analysis.* Kluwer Academic, 1990.

[14] J. Moura Pires and H. Prade. Logical analysis of fuzzy constraint satisfaction problems. *Proc. on the 7th Conf. on Fuzzy Systems (FUZZ-IEEE'98), World Congress on Computational Intelligence*, Anchorage, May 1998, 1:857–862, 1998.

[15] P. Savéant, M. Grabisch, and F. Planchon. Handling uncertainty in data fusion systems with possibility theory. *IEEE Tr. on Fuzzy Systems*, submitted.

[16] J. Schubert. *Cluster-based specification techniques in Dempster-Shafer theory for an evidential intelligence analysis of multiple target tracks.* Ph.D. Thesis, Royal Inst. of Technology, Stockholm, Sweden, 1994.

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