



HAL
open science

Fusion of possibilistic sources of evidences for pattern recognition

Anas Dahabiah, John Puentes, Basel Solaiman

► **To cite this version:**

Anas Dahabiah, John Puentes, Basel Solaiman. Fusion of possibilistic sources of evidences for pattern recognition. *Integrated Computer-Aided Engineering*, 2010, 17 (2), pp.117-130. 10.3233/ICA-2010-0333 . hal-00486761

HAL Id: hal-00486761

<https://hal.science/hal-00486761v1>

Submitted on 9 Jun 2021

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.

Fusion of possibilistic sources of evidences for pattern recognition

Anas Dahabiah^{a,b}, John Puentes^{a,b} and Basel Solaiman^{a,b}

^a*Institut Telecom; Telecom Bretagne, Département Image et Traitement de l'Information, Brest, France*

^b*INSERM, U650, Laboratoire de Traitement de l'Information Médicale, Brest, France*

Abstract. Information processing in modern pattern recognition systems is becoming increasingly complex due to the flood of data and the need to deal with different aspects of information imperfection. In this paper a simple and efficient possibilistic evidential method is defined, taking account of data heterogeneity, combined with proportional conflict redistribution to include information conflict, paradox, and scarcity, within a fusion framework. It ponders information constraints and updating for dynamic fusion, and appropriately considers training set elements imperfection, class set continuity, and system output information scalability, encompassing a significant range of issues encountered in current databases. One example of knowledge sources processing with those constraints is given to explain the main processing phases, followed by suitable application instances in satellite and medical image recognition.

Keywords: Possibility theory, Proportional conflict Redistribution, Hybrid Dezert-Smarandache Model, Pattern Recognition, Satellite Images, Gastroenterology endoscopic images.

1. Introduction

Many problems related to information processing in pattern recognition systems have been addressed and pointed out in the literature [25], like information imperfection, heterogeneity, conflict, paradox, scarcity, constraints, dynamism, etc. [6], [14], [17], [19], [22], [27-29]. In general, these problems have been approached separately [15], using specific methods according to particular restricted conditions [6], [17]. Nevertheless, in real practical applications, it is often very likely that all these aspects take place simultaneously, given the significant amount of processed data, besides the complexity of systems, the sought tasks and purposes [6], [14]. Consequently, this paper proposes a method adequate for managing all the aforementioned points under a conjoined framework.

None of the proposed theoretical models available today like for example, fuzzy set, probability, possibility, evidence, etc., can fit or appropriately deal with all types of information elements imperfection,

due to their multiform [25]. On the contrary, these models are complementary (not concurrent), and each of them aims at describing a given aspect of imperfection. Accordingly, we intend to make use of several models combined within the scheme of possibility theory, in a pattern recognition context to handle a significant set of imperfection elements.

Our latest work [1], addressed some information processing problems like imperfection, heterogeneity, conflict, scarcity and scalability, ignoring other comparably important issues like information paradox, contradiction, continuity, constraints and updating (dynamism). As a result, we search to ameliorate our approach in such a way that all the aforementioned elements are considered.

Information Imperfection can take place when labeling the patterns of the training set (examples) via an automatic system [4], [22], [28], or when assigning imprecise, ambiguous, evidential, probabilistic, uncertain, or missing values to pattern features

[1-3]. Unlike the first type of imperfection that has rarely been studied [28], pattern feature imperfection appears to be markedly disregarded in the literature. This unsolved question inevitable leads to considerable difficulties in measuring the similarity between the observed and the training set patterns [2, 3], which is a pivotal task in recognition [7], [9, 10], due to the inherent limitations and application constraints of reported similarity measures [2]. Yet recent works [7], [9] underline the importance and the need of a general measure, capable to take account of all relevant aspects of imperfection, simultaneously in a rather simple and fast way. Such claim relies on the ground that similarity measurement represents the corner stone in a significant number of data mining tasks and techniques, and is motivated by the fact that similarity measures proposed until now, only consider some aspects of imperfection, disregarding the rest [6], [14], [17]. For instance, some approaches merely take account of imprecision [17], whereas others only address the ambiguity [4].

Information Heterogeneity is another important aspect to be considered in pattern recognition, because patterns generally consist of information elements having different measuring scales (binary, qualitative, quantitative, ordinal, etc.) [6], [13, 14], [26]. As a result, various data conversion approaches have been proposed to transform all the scales to one type [13], [26]. Nonetheless, these techniques have been implemented arbitrarily and did not demonstrate how to deal with real complex patterns, containing a remarkable number of different attributes, especially in the presence of imperfection. Recent applied works [6], [14] have stressed the importance of assessing information imperfection and heterogeneity together under a unified context, when making similarity measures. On the contrary, those methods have only examined the heterogeneity, without including ordinal or imperfect information elements. We show in this paper that possibilistic similarity measures dealing with both of the aforementioned points, play a vital role to improve similitude appraisal, between the observed objects and the examples.

Information Conflict and Scarcity are two other important prospects to be studied when managing pattern recognition information. The first issue could take place when a pattern is close to several training examples from different categories, whereas the second arises when a pattern is far away from any pattern in the training set (outlier). In this case, the most probably pattern belongs to an unknown class,

for which no information has been gathered in the training set, and as a consequence should be rejected. These two points have been very well discussed in [23], [28] using evidence theory.

Information Scalability is a complementary and interesting issue in the system we are looking for. Owing to the fact that pattern recognition could be used systematically as a pre-processing step, preceding other data mining tasks like rule extraction, outlier and anomaly detection, our system output information elements must be flexibly transformable to other forms suitable for the next steps. It was clearly illustrated in [2], [27] that if it is feasible to build an evidential partition of the information using belief masses, then the transformation into other types of partitions like possibilistic, fuzzy or hard partition is direct. Therefore, we aim at producing our results in this form, by using instead an enhanced version of belief masses that apply the model proposed by Dezert and Smarandache [11], [12].

Information Continuity or Vagueness is also an important characteristic to be considered in order to assure the simplicity and the efficiency of the proposed method [11]. In [1] we supposed that the classes of our system were exclusive or could be converted to exclusive elements, by finding an appropriate refinement of the class set. That is, we assumed the association of observed patterns with one and only one class. Consequently, the belief masses were defined upon the power class set according to Shafer's model [11], assuming that there were not intersections between classes, i.e. for any two different classes ω_i and ω_j from the class set Ω : $\omega_i \cap \omega_j = \emptyset$ and therefore no belief can be distributed to this set, $m(\omega_i \cap \omega_j) = 0$.

However, in a large spectrum of practical problems the categories intrinsic nature is only vague and imprecise. Given that these classes cannot be identified or precisely separated, further refinement of the associated categories is either very complex or not feasible at all. Many problems involving fuzzy continuous and relative concepts described in natural language and having no absolute interpretation like tallness/smallness, pleasure/pain, cold/hot, etc., enter in these types. Consequently, in order to guarantee the generality and the integrity of the proposed system, we use Dezert-Smarandache model (DSm), in which belief masses are defined on the hyper power

set that contains all the classes, their intersections, and their unions. The DS_m model includes the Dempster Shafer model (DS), a particular case characterized by empty intersections, implying that the hyper power set is reduced to the power set. Additionally, it has been reported that the DS model could give sometimes counter-intuition results, particularly when the conflicting mass takes significant values [11, 12].

Information Constraints and Updating can be regarded as a fundamental issue for the development of reliable modern pattern recognition systems, because they are directly related to the properties and the characteristics of pattern classes. Problems without intersection or union between some classes are representative examples. For instance, if Ω is a set of three different pathologies $\Omega = \{\omega_1, \omega_2, \omega_3\}$, and according to specialized knowledge, it is inconceivable that a patient has two pathologies ω_1 and ω_2 at the same time, given the observed completely different symptoms, then the information - $\omega_1 \cap \omega_2 = \emptyset$ or $m(\omega_1 \cap \omega_2) = 0$ - must be taken into account by the recognition system in the reasoning phase. This issue was neither included in our previous work [1].

When the classes are exclusive, then $\omega_1 \cap \omega_2 = \omega_1 \cap \omega_3 = \omega_2 \cap \omega_3 = \emptyset$.

The belief masses of these intersections are considered in this case as conflicting masses that should be redistributed to the non empty sets. The hybrid DS_m model proposed in [11, 12] represents an appropriate framework to deal with class sets submitted to intersection and/or union constraints. When there are only exclusivity constraints, this model is equivalent to Shafer's model implemented in [11], and when no assumption or constraint is applied to the system it is called free DS_m model. Thanks to this property, hybrid DS_m is adapted in our work with a slight modification related to the redistribution of total and partial conflicts (section 2).

Indeed, the most interesting property of this model is its capability to give coherent and robust results in dynamic fusion [11], for information updating, when the class data set changes through time with revision of the available knowledge. For instance, let us suppose that in the previous example and according to new evidence, we knew that our

database has no object belonging to the class ω_3 ($\omega_3 = \emptyset$). In this case, we can be sure that the hybrid DS_m model is capable to efficiently distribute the new conflicting mass $m(\omega_3)$ to the other non-empty sets.

Information Paradox or Contradiction is a very important issue related to training set elements carefully pondering in pattern recognition systems. At a practical level, the examples of a training set can never be ideally selected, and sometimes different training patterns give incoherent information about the observed pattern class. This contradiction can be materialized as seen in [1] by the conflicting mass. Nevertheless, instead of distributing this mass to all the belief masses of all the examples, Dezert and Smarandache proved that it is more reasonable and logic to only redistribute the conflict to the examples involved in it, proportionally to their belief masses [12]. Accordingly, they established five different versions of proportional conflict redistribution approach (*PCR1*, *PCR2*, ..., *PCR5*), showing that this redistribution (except *PCR1*) satisfies the three main properties of a good combination rule: the coherence, the commutativity, and the neutral impact of the vacuous belief assignments.

In addition to these properties, *PCR6* proposed by Martin and Osswald is quasi-associative, stable in terms of decision, and more coherent than the other versions, when there are more than two evidence sources [5]. For two bodies of belief, *PCR6* coincides with *PCR5* [12]. We propose the use of the *PCR* approach at the output of our system, since it can properly work for any degree of conflict, for each DS_m model (Shafer's model, free DS_m model or the hybrid DS_m models), and for every type of fusion (static or dynamical).

The proposed general simple method, essentially based on possibility theory and Dezert-Smarandache proportional conflict redistribution, is described in the following section, supported by a concrete illustrative example in section 3. This technique is then analyzed in section 4, regarding two compatible application instances, related to satellite and medical image recognition. Finally, and as a way of conclusion, these points will be discussed and analyzed in the last section.

2. Possibilistic evidential pattern recognition system

Suppose that a set of observed patterns $\Theta = \{A_j^1, A_j^2, \dots, A_j^i, \dots, A_j^n\}$, is to be classified depending on a training pattern set $T = \{A_k^1, A_k^2, \dots, A_k^l, \dots, A_k^t\}$ into the class set $\Omega = \{\omega_1, \omega_2, \dots, \omega_c, \dots, \omega_N\}$ (fig. 1). Given that A represents a multivariate feature vector of the considered pattern, its variables could have imperfect (imprecise, ambiguous, uncertain, or missing) values, and could be heterogeneous (quantitative, qualitative, ordinal, etc.). We will also suppose that the class membership of the training pattern is assessed by an expert or by another automatic system as a possibility distribution [1-2]. Specifically, each example A_k^l for $l=1,2,\dots,t$ is associated with a possibilistic label modeled as:

$$[\mu_1(\omega_1) \quad \mu_1(\omega_2) \quad \dots \quad \mu_1(\omega_c) \quad \dots \quad \mu_1(\omega_N)]$$

in such a way that:

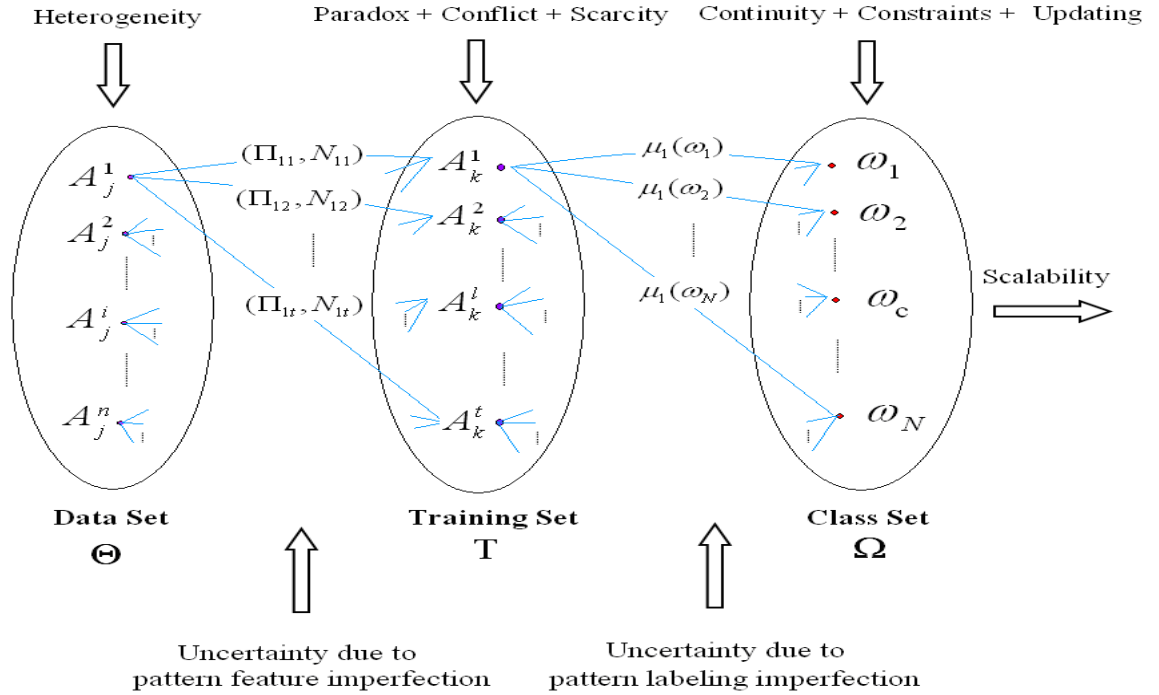


Fig. 1. A pattern recognition system.

$$\mu \in [0, 1] \text{ and } \sum_{c=1}^N \mu_l(\omega_c) = 1 \quad (1)$$

In the first phase, the similarity between the observed patterns and all the patterns of the training set is calculated via the necessity degree that the two patterns are similar (lower bound) and the possibility degree of resemblance (the upper degree) as we proposed in [1]. These degrees are the average of inter-attribute necessity and possibility degrees that take into account, if required, specific physicians' points of view, to adjust the similarity of each attribute between the patterns modeled by the tolerance function. Even if, sometimes necessity and possibility degrees are aggregated using the minimum or the maximum operator [16], this strategy does not work for our purpose. In those cases the value of the most dissimilar attribute will dominate, determining the final computation result, regardless of the other attributes, for which similarity values may be extremely close. Therefore, we suggest the average as a fair estimation that considers the values of all the attributes.

The inter-attribute necessity and possibility degrees of similarity between the attribute value a_{ff} given in A_j^i , $\forall i \in \{1, 2, \dots, n\}$ modeled by its possibility distribution (represented on the axis y) $\pi_{A_j, a_{ff}}(a_{ff}, y)$ and the value a_{fk} given in A_k^l , $A_{j, a_{ff}}$

$\forall l \in \{1, 2, \dots, t\}$ provided by its possibility distribution (represented on the axis x) $\pi_{A_k, a_{fk}}(x, a_{fk})$,

$\forall f \in \{1, 2, \dots, S\}$ (S is the number of attributes in each pattern), are calculated as follows.

Supposing that D is the definition domain of the considered attribute ($U = D \times D$) and that Λ is the tolerance function associated to this attribute, the conjoint possibility distribution π_D is calculated by:

$$\pi_D(a_{ff}, a_{fk}) = \min(\pi_{A_j, a_{ff}}(x), \pi_{A_k, a_{fk}}(y)) \quad (2)$$

In this case, the inter-attribute necessity and possibility degrees of similarity, N_f and Π_f , can be calculated respectively as [15]:

$$N_f(a_{ff}, a_{fk}) = \text{Inf}_{u \in U} [\max(\Lambda(u), 1 - \pi_D(u))] \quad (3)$$

$$\Pi_f(a_{ff}, a_{fk}) = \text{Sup}_{u \in U} [\min(\Lambda(u), \pi_D(u))] \quad (4)$$

We consider that if the value of an attribute is given in a pattern and is unassigned in the other (the case of missing values), it is completely possible that these values are similar $\Pi_f = 1$, but we are entirely uncertain, $N_f = 1$.

After calculating the degrees of similarity between $A_j^i \in \Theta$ and $A_k^l \in T$, denoted as N_{il} and Π_{il} , $\forall i \in \{1, 2, \dots, n\}$ and $\forall l \in \{1, 2, \dots, t\}$, the basic belief assignments of the observed patterns are calculated according to the following algorithm:

For each $A_j^i \in \Theta$, $\forall i \in \{1, 2, \dots, n\}$

For each $A_k^l \in T$, $\forall l \in \{1, 2, \dots, t\}$

$$\begin{aligned} M_{N_i}^l &= \begin{bmatrix} m_i^l(\omega_1) \\ N \\ m_i^l(\omega_2) \\ N \\ \cdot \\ \cdot \\ m_i^l(\omega_N) \\ N \end{bmatrix} = \begin{bmatrix} N_{il}\mu_l(\omega_1) \\ N_{il}\mu_l(\omega_2) \\ \cdot \\ \cdot \\ N_{il}\mu_l(\omega_N) \end{bmatrix} \\ M_{\Pi_i}^l &= \begin{bmatrix} m_i^l(\omega_1) \\ \Pi \\ m_i^l(\omega_2) \\ \Pi \\ \cdot \\ \cdot \\ m_i^l(\omega_N) \\ \Pi \end{bmatrix} = \begin{bmatrix} \Pi_{il}\mu_l(\omega_1) \\ \Pi_{il}\mu_l(\omega_2) \\ \cdot \\ \cdot \\ \Pi_{il}\mu_l(\omega_N) \end{bmatrix} \end{aligned} \quad (5)$$

Where $M_{N_i}^l$ represents the necessity-based evidence, and $M_{\Pi_i}^l$ represents the possibility-based evidence, provided by the example $A_k^l \in T$, concerning the membership of the observed pattern $A_j^i \in \Theta$ to all the classes of Ω .

As we might remark, each training pattern plays the role of an item of evidence (witness) influenced by similarity necessity and possibility degrees, considering that the higher is the similarity of the observed pattern to it, the more is our belief that it belongs to its classes with the same degree.

For each element X from the hyper power set of the class set (Dedekind's lattice) denoted D^Ω , we apply the conjunctive consensus combination rule as:

$$m_c(X) = \sum_{Y_1 \cap \dots \cap Y_t = X} \prod_{l=1}^t m_l(Y_l) \quad (6)$$

Where $Y_l \in D^\Omega$ is the response of the witness l (evidence source), and $m_l(Y_l)$ the associated belief function.

Then *PCR6* rule proposed by Martin and Osswald [5] is applied in the following manner:

$$m_{PCR6}(X) = m_c(X) + \quad (7)$$

$$\sum_{l=1}^t m_l(X)^2 \sum_{\substack{\bigcap_{k=1}^{l-1} Y_{\sigma_l(k)} \cap X = \emptyset \\ (Y_{\sigma_l(1)}, \dots, Y_{\sigma_l(l-1)}) \in (D^\Omega)^{l-1}}} \left(\frac{\prod_{j=1}^{l-1} m_{\sigma_l(j)}(Y_{\sigma_l(j)})}{m_l(X) + \sum_{j=1}^{l-1} m_{\sigma_l(j)}(Y_{\sigma_l(j)})} \right)$$

where σ_l counts from 1 to l avoiding l :

$$\begin{cases} \sigma_l(j) = j & \text{if } j < l, \\ \sigma_l(j) = j+1 & \text{if } j \geq l, \end{cases} \quad (8)$$

As Y_l is a focal element of the witness l , then

$$m_l(X) + \sum_{j=1}^{l-1} m_{\sigma_l(j)}(Y_{\sigma_l(j)}) \neq 0.$$

This rule is applied twice: initially to the necessity-measure-based testimony of the examples in the training set, and then to the possibility-based testimony. Afterwards, the two results are combined using DSm model. Thereafter, the generalized pignistic probability transformation is applied to take a decision about the class of the pattern under consideration [11]. This transformation is more general than the pignistic probability utilized in our previous work and is given as follows:

$$\forall A \in D^\Omega, \quad BetP\{A\} = \sum_{X \in D^\Omega} \frac{C_M(X \cap A)}{C_M(X)} m(X) \quad (9)$$

where $C_M(X)$ denotes DSm cardinal of the proposition X for DSm model M of the problem [18].

3. An illustrative example of knowledge sources treatment

Let us suppose that our system consists of three elements $\{A_j, A_k^1, A_k^2\}$ where A_j is the observed

object, while A_k^1 and A_k^2 are the training patterns. Each object is described by three attributes $\{a_1, a_2, a_3\}$ as follows: a_1 is a nominal binary attribute that takes its values in the set {positive, negative}; a_2 is an imprecise quantitative measure; and a_3 is characterized according to other measures or automatic systems like the one that we proposed in [4], or by a probability distribution. The values assigned to these three attributes in the three aforementioned patterns are depicted in fig. 2-a. In order to estimate the similarity between these three objects – which is a fundamental step in any recognition system – we assume that experts' viewpoints about the resemblance between each pair of variables can be modeled by tolerance functions [1-3]. As schematized in fig. 2-a, the true/false tolerance function is designed for the first attribute. Such representation means that, in the expert's opinion, the two values of this attribute are identical, only if they are considered completely similar, elsewhere they are totally dissimilar. Otherwise for the second attribute, similarity is modeled by f_2 (fig. 2-a). According to f_2 , two values of an attribute are totally similar if the difference between them is null. This similarity decreases when the difference increases up to ± 30 . Beyond these values, the similarity is equal to 0. Regarding the last attribute, we suppose that the expert did not have a particular opinion to formulate the similarity viewpoint between two attributes. Nevertheless, similarity necessity and possibility degrees between the values of this attribute must be considered and computed even if the expert is not capable to indicate or to express an opinion, due to knowledge insufficiency, problem complexity, need of other evidences or proofs, etc. We suppose in this apprehensible example that even if the expert (a physician, for instance) is incapable of describing the similarity between two attributes, estimated as probability distributions, obtained for example from a patient record database, the similarity between those attributes can be calculated. Since tolerance functions are a way to personalize the process by taking account of experts' similarity standpoints, it is not appropriate to consider those values as weights. In this case, similarity degrees can be easily evaluated using conventional equations, after transforming the probability distributions to possibility distributions, through an adapted method like Dubois-Prade transformation [24].

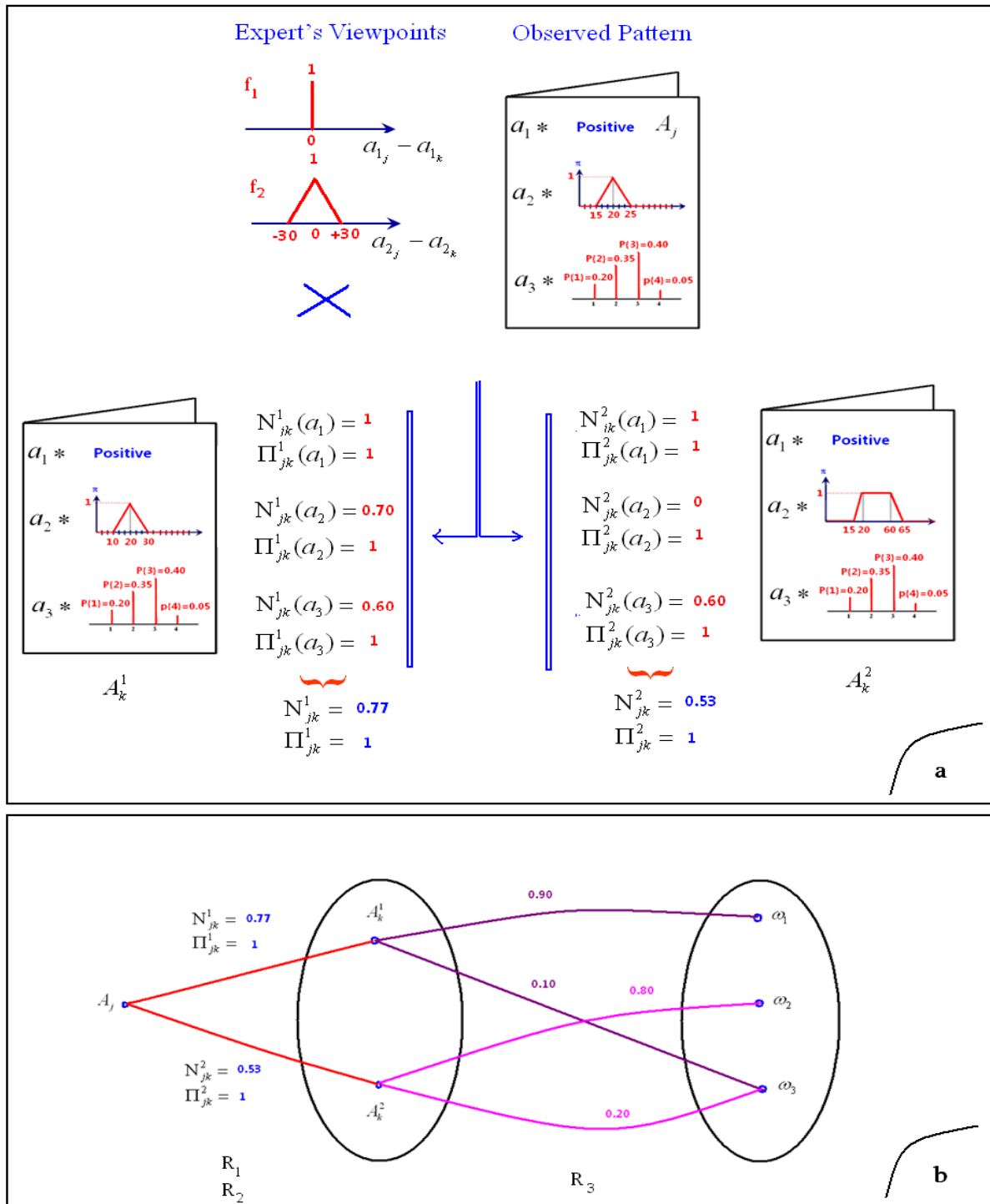


Fig. 2. Illustrative example: a- the observed object and the training patterns, with their assigned values, inter-variable necessity and possibility degrees, and the global degrees. b- the pattern recognition system.

The values of inter-attribute similarities modeled by the similarity possibility degrees and the similarity necessity degrees are shown in fig. 2-a. These values have been averaged in order to calculate the inter-pattern similarity between A_j and A_k^l , ($\forall l \in \{1, 2\}$), modeled by N_{jk}^l and Π_{jk}^l [1].

Measuring the similarity between these patterns that contain heterogeneous values (qualitative, quantitative, etc.), and imperfect information elements (imprecise, probabilistic, etc), is extremely complex and uncertain using the conventional and the prior measures and approaches (section 1), particularly when experts' points of view must also be taken into account. On the contrary, this calculation is completely straightforward and efficient applying the proposed framework.

It can be noticed that the observed object is close to both training patterns, A_k^1 which is mostly assigned to ω_1 , and A_k^2 which is approximately assigned to ω_3 . Therefore, the paradox resulting from these two information elements must be considered, in addition to information uncertainty, when assigning the classes to the training patterns.

We also assume that there are three classes, $\Omega = \{\omega_1, \omega_2, \omega_3\}$, and according to expert's knowledge, neither ω_1 and ω_3 , nor ω_2 and ω_3 can occur at the same time. Nonetheless, ω_1 and ω_2 can happen simultaneously. These information constraints resulting from our knowledge about the nature of classes must be pondered in the process (fig. 2-b).

Applying the given values of the computed similarity necessity and possibility degrees on the one hand, the class membership values of the first training pattern ($\mu_{A_k^1}(\omega_1) = 0.90$, $\mu_{A_k^1}(\omega_2) = 0$, and $\mu_{A_k^1}(\omega_3) = 0.10$) and those of the second training object ($\mu_{A_k^2}(\omega_1) = 0$, $\mu_{A_k^2}(\omega_2) = 0.80$, and $\mu_{A_k^2}(\omega_3) = 0.20$), on the other hand, to equation 5, we obtain:

$$\underline{m_1^l(\omega_1)} = 0.77 \times 0.90 = 0.693,$$

$$\overline{m_1^l(\omega_1)} = 1 \times 0.90 = 0.90;$$

$$\underline{m_1^l(\omega_2)} = 0.77 \times 0 = 0, \quad \overline{m_1^l(\omega_2)} = 1 \times 0 = 0;$$

$$\underline{m_1^l(\omega_3)} = 0.77 \times 0.10 = 0.077,$$

$$\overline{m_1^l(\omega_3)} = 1 \times 0.10 = 0.10;$$

$$\underline{m_2^l(\omega_1)} = 0.53 \times 0 = 0, \quad \overline{m_2^l(\omega_1)} = 1 \times 0 = 0;$$

$$\underline{m_2^l(\omega_2)} = 0.53 \times 0.80 = 0.42,$$

$$\overline{m_2^l(\omega_2)} = 1 \times 0.80 = 0.80;$$

$$\underline{m_2^l(\omega_3)} = 0.53 \times 0.20 = 0.110,$$

$$\overline{m_2^l(\omega_3)} = 1 \times 0.20 = 0.200;$$

Accordingly:

$$m_1^l(\omega_1) = [0.693, 0.90],$$

$$m_1^l(\omega_2) = 0$$

$$m_1^l(\omega_3) = [0.077, 0.10]$$

$$m_2^l(\omega_1) = 0$$

$$m_2^l(\omega_2) = [0.42, 0.80]$$

$$m_2^l(\omega_3) = [0.110, 0.20].$$

Supposing that information conflict and paradox must be precisely taken into account, we implement the proportional conflict redistribution approach in the hybrid DSm model. Even so, as these belief masses are admissible, the imprecise version of PCR5 [12] can be applied as well to facilitate the calculation.

Working with sets, and supposing that there is not any constraint that dominates the system, the following masses are obtained for the free DS_m model:

$$m_{FreeDSm}^I(\omega_1) = 0, m_{FreeDSm}^I(\omega_2) = 0,$$

$$m_{FreeDSm}^I(\omega_3) = [0.0085, 0.02]$$

$$m_{FreeDSm}^I(\omega_1 \cap \omega_3) = [0.0762, 0.18]$$

$$m_{FreeDSm}^I(\omega_2 \cap \omega_3) = [0.0323, 0.08]$$

$$m_{FreeDSm}^I(\omega_1 \cap \omega_2) = [0.2911, 0.72]$$

$$m_{FreeDSm}^I(\omega_1 \cap \omega_2 \cap \omega_3) = 0$$

Presuming that part of the partial conflict redistributed to ω_1 , ω_2 , and ω_3 , are x_{ω_1} , x_{ω_2} and x_{ω_3} respectively, then the partial conflicting mass $m_{FreeDSm}^I(\omega_1 \cap \omega_3) = [0.0762, 0.18]$ will be redistributed to ω_1 , and ω_3 as follows:

$$\frac{x_{\omega_1}}{[0.693, 0.90]} = \frac{x_{\omega_3}}{[0.11, 0.20]} = \frac{[0.0762, 0.18]}{[0.8030, 1.10]}$$

Whence,

$$x_{\omega_1} = \frac{[0.0528, 0.1620]}{[0.8030, 1.100]} = [0.048, 0.2]$$

And

$$x_{\omega_3} = \frac{[0.0084, 0.036]}{[0.8030, 1.10]} = [0.00764, 0.045]$$

Once again, the proportional conflict redistribution approach is applied to the second partial conflict mass given by the second information constraint as follows:

$$\frac{x_{\omega_2}}{[0.42, 0.80]} = \frac{x_{\omega_3}}{[0.077, 0.10]} = \frac{[0.0323, 0.08]}{[0.4970, 0.90]}$$

Whence

$$x_{\omega_2} = \frac{[0.0136, 0.064]}{[0.497, 0.90]} = [0.0151, 0.1288]$$

And

$$x_{\omega_3} = \frac{[0.0025, 0.0080]}{[0.4970, 0.90]} = [0.0028, 0.0161].$$

As a result:

$$m_{PCR}^I(\omega_1) = [0.048, 0.20],$$

$$m_{PCR}^I(\omega_2) = [0.0151, 0.1288],$$

$$m_{PCR}^I(\omega_3) = [0.0189, 0.0811],$$

$$m_{PCR}^I(\omega_1 \cap \omega_3) = 0,$$

$$m_{PCR}^I(\omega_2 \cap \omega_3) = 0,$$

$$m_{PCR}^I(\omega_1 \cap \omega_2) = [0.2911, 0.720],$$

$$m_{PCR}^I(\omega_1 \cap \omega_2 \cap \omega_3) = 0.$$

We can notice by closely looking at the result obtained above that the greatest belief will be given to the occurrence of both ω_1 and ω_2 , and not to one of them, which is very reasonable and logical. Contrarily, the fusion methods proposed in [1] cannot deal with this type of problem taking account of the imposed constraints, the imprecise belief masses of sources, and the composed hypotheses like $\omega_1 \cap \omega_2$, since their power class sets are closed on unions only. Furthermore, the traditional fusion approaches, Yager, Shafer's model, etc., cannot fairly redistribute the conflicting mass to the elements involved in the paradox, like the proportional conflict redistribution approach applied in this conflict scenario.

If we assume the exclusivity of the classification system as Shafer's model does by adding the constraint $\omega_1 \cap \omega_2 = \emptyset$ to the other two constraints, then the partial conflicting mass $m_{FreeDSm}^I(\omega_1 \cap \omega_2) = [0.2911, 0.72]$ will be redistributed as follows:

$$\frac{x_{\omega_1}}{[0.693, 0.90]} = \frac{x_{\omega_2}}{[0.42, 0.80]} = \frac{[0.2911, 0.72]}{[1.113, 1.70]}$$

$$x_{\omega_1} = \frac{[0.2017, 0.6480]}{[1.113, 1.70]} = [0.1187, 0.5838]$$

And

$$x_{\omega_2} = \frac{[0.1223, 0.5760]}{[1.113, 1.70]} = [0.0719, 0.5175]$$

As a result:

$$m_{PCR}^I(\omega_1) = [0.1667, 0.7838],$$

$$m_{PCR}^I(\omega_2) = [0.087, 0.6463],$$

$$m_{PCR}^I(\omega_3) = [0.0189, 0.0811];$$

$$m_{PCR}^I(\omega_1 \cap \omega_3) = 0,$$

$$m_{PCR}^I(\omega_2 \cap \omega_3) = 0,$$

$$m_{PCR}^I(\omega_1 \cap \omega_2) = 0,$$

$$m_{PCR}^I(\omega_1 \cap \omega_2 \cap \omega_3) = 0.$$

While applying Shafer's model in such type of problems leads to:

$$m_{Shafer}(\omega_3) = 1, \quad m_{Shafer}(X) = 0, \\ \forall X \in 2^\Omega / \{\omega_3\}.$$

Nonetheless, this result is illogical, since the belief has been totally given to the element ω_3 , which has received a very small amount of it from the examples in the training set.

This example illustrates a simple, coherent and unified processing scheme, to treat multiple types of information imperfection and limitations, as well as user personalization and system's constraints. Two

straightforward envisioned utilization instances are discussed in the next section.

4. Image recognition applications

Thanks to its simplicity, flexibility, and generality, the previous example can be extended to a significant number of pattern recognition problems, where all or some of the aforementioned information limitations and challenges are encountered.

This section introduces two interesting applications in which our technique can be exploited analogously, following equivalent phases of reasoning and calculation, to recognize similar objects, in satellite and medical images. Other types of pattern identification could be extrapolated from these two examples, likewise.

In remote sensing images, the objective is to assign each pixel to a particular class (vegetation, crops, rivers, etc.) depending on the fusion of several information sources, like the spectral bands of the satellite image, or the a priori knowledge of the expert about the studied area. In fig. 3, we give an example in which the objective is to classify the pixels of 512×512 LANDSAT images, to one of nine different categories of vegetation, based on its spectral bands MSS.

To accomplish this task, the training dataset is constructed by choosing the most significant and representative regions of pixels (samples) with known labels in the images, and by extracting the characterizing features of each class (called the spectral signature) that allow to distinguish it from the others [4].

Actually, the selected samples that play a fundamental role in the fusion process can be heterogeneous, i.e. their composing pixels may belong to several classes, and this heterogeneity must be considered during the fusion. For instance, in fig. 3-a each sample corresponds to the spectral responses of every object in the ground of a rectangle of 57 meters \times 79 meters, and consequently it may have more than one class.

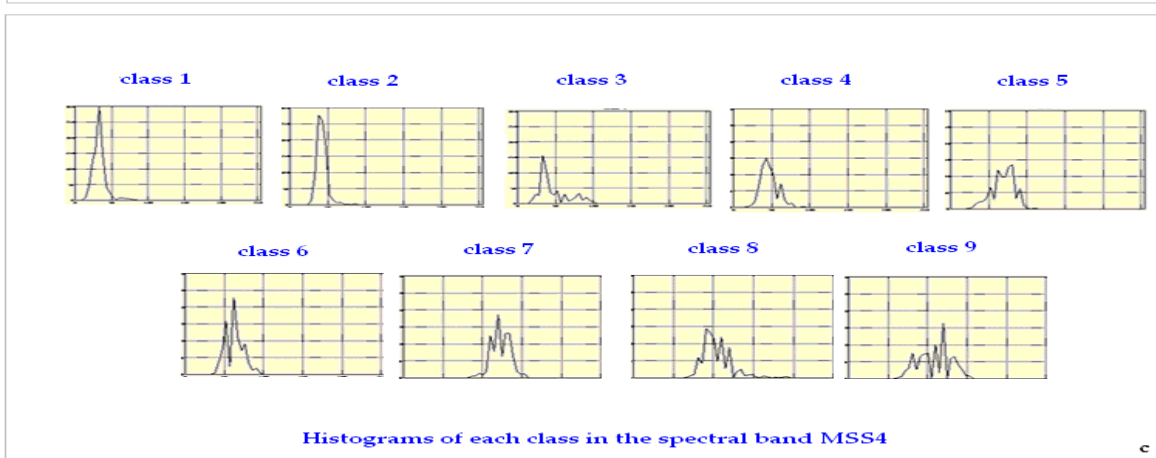
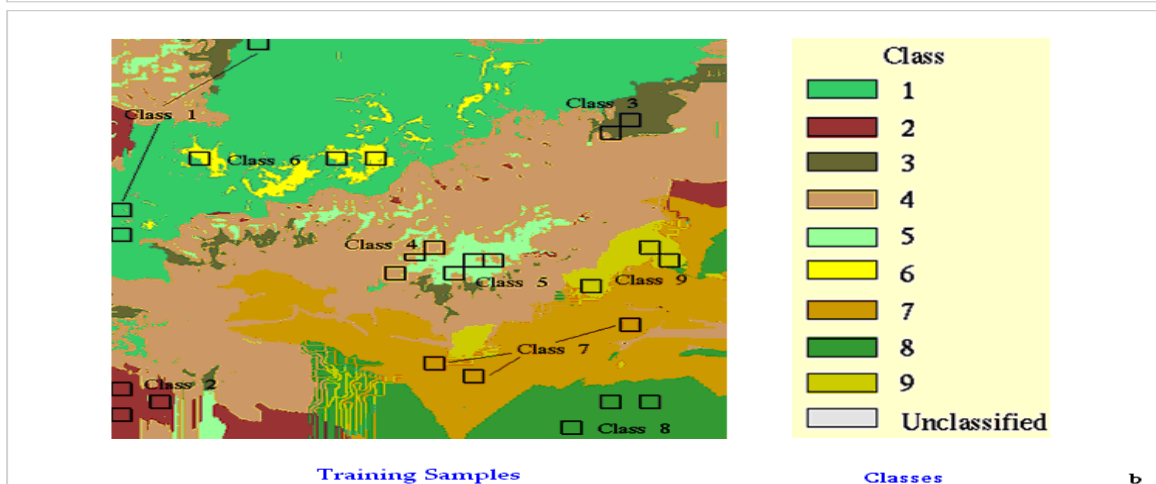
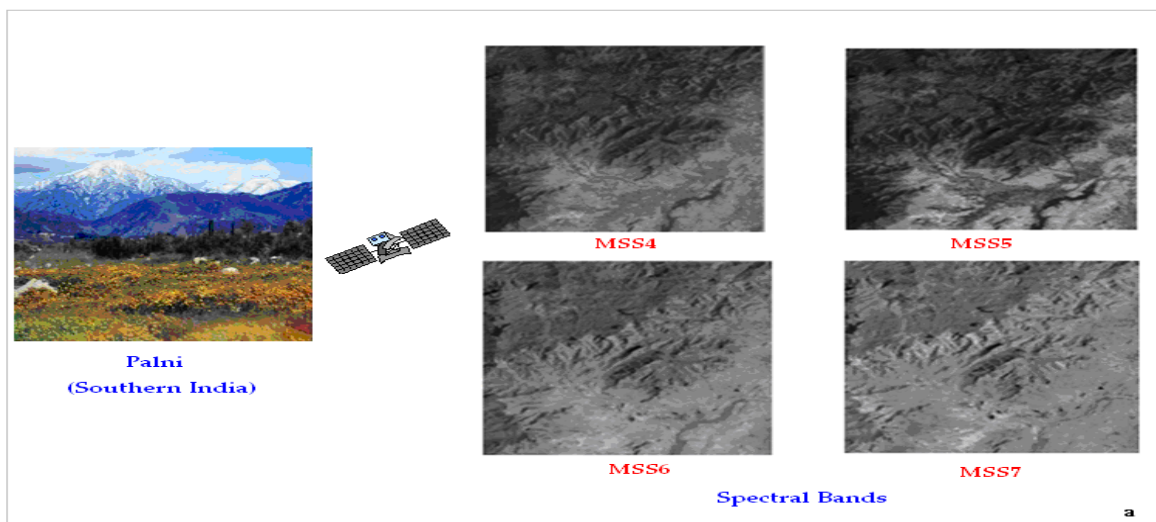


Fig. 3. Information processing in satellite image classification [23], a) four spectral bands of Palni, b) the training samples and the classes, c) the class histograms.

Such problem is found when examining the samples in the centre of the image (fig. 3-b), where it can be remarked that class 4 and class 5 coexist simultaneously in the majority of them. Moreover, the two samples of class 3 in the upper right corner are completely heterogeneous and their membership degrees to the classes which are closely related to the resolution of the spectral band MSS of LANDSAT, must be different too. Accordingly, one can deduce that pixel regions can be assigned to more than one class with different weights. Hence, this system cannot be treated within Shafer's model that tends to exaggerate the classes exclusivity principle [11].

By analyzing the histograms of the nine system classes (fig 3-c), for the spectral band *MSS4* (fig. 3-a), it can be remarked that some of them are spread out, of low amplitude, and are embedded among the others. This is the case of classes 3 and 4, characterized by a poorly spread out histogram that can be merged with classes 1 and 2, whose histogram is very narrow. Consequently, these classes (3 and 4) will be deficiently recognized in the absence of other information sources. Actually, the paradox and the conflict between the classes resulting from each band, can be properly pondered using DSm model and the PCR approach.

Considering fig. 3-c histograms, it can be pointed out that the classes are roughly divided into two main groups: the first one contains classes 1, 2, ..., 6 whose histograms have rather low grey levels (from 0 to 100), while the second consists of the other classes for which histograms have higher grey levels (from 60 to 200). This is because samples of the first group correspond to forest or savannah situated in mountains, while the second group corresponds to cultures and thickets found in plains. Adding this knowledge in the form of intersection and union constraints can notably enhance the performance of the system.

The proposed possibilistic framework appears as the most appropriate for this system, to estimate similarity between histograms of the training and the observed patterns, or the resemblance of other features between different information sources [21]. In the first case histograms are normalized before applying an appropriate probability-possibility distribution transformation on which our possibilistic similarity measuring can be carried out [16]. In the second case, an equivalent operation can be carried out even if the attributes are described by means of heterogeneous

and imperfect information elements [3]. Another important issue in this type of applications is information updating, because information sources can frequently change through time. For instance, some elements can spread (like the towns), others can be transformed (i.e. a forest replaced by a field), while some features change their appearance during summer and winter (vegetation) [23].

In the medical domain, the same issues can be found, but in other forms. Fig. 4 shows an example of lesions observed in gastroenterology images, which are described by an expert physician. Those described characteristics have different measure scales (information heterogeneity) and can have imprecise or missing data (information imperfection) [8]. Given the descriptions complexity, in some cases it is not feasible to solve pattern recognition problems applying the conventional similarity measure framework, particularly when the physician's viewpoint must be considered in the calculation [2, 3].

Again, the proposed possibilistic fusion framework leads to significantly improved results, in agreement with ground truth. Previously formulated algorithms have materialized different aspects of our possibilistic framework through iterative-projection-optimization-based representational visualization models, validating their concordance with the experts' diagnosed patterns or classes [1-3], [20]. Fig. 5 presents two scaling-based visualization models, and two graphical representational models, applied to an image dataset containing 18 described images of 5 different lesions $\Omega = \{P_1, P_2, P_3, P_4, P_5\}$ structured in the following manner according to the ground-truth provided by the specialist [8]: $P_1 = \{O_1, O_2\}$ corresponds to the "Dilated Lumen" pathology; $P_2 = \{O_3, O_4, O_5, O_6, O_7, O_8, O_9, O_{10}\}$ conforms with the description of "Esophagus Stenosis"; $P_3 = \{O_{11}, O_{12}, O_{13}, O_{14}\}$ is a set of images that represent the "Extrinsic Compression" pathology; $P_4 = \{O_{15}\}$ describes the "Web Shape" pathology; $P_5 = \{O_{16}, O_{17}, O_{18}\}$ is a set of images on which the "Ring Shape" pathology is visible.

The results displayed in fig. 5 represent strong similarity relations existing between objects that belong to the same pathology. Namely, an object that corresponds to a given pathology is more similar to any other object of the same class, than to objects of other pathologies. For instance, fig. 5-b depicts those relations using possibilistic similarity, as a closed

This can be assured using the proportional conflict redistribution approach used in our system, as explained in the above numeric example (section 3).

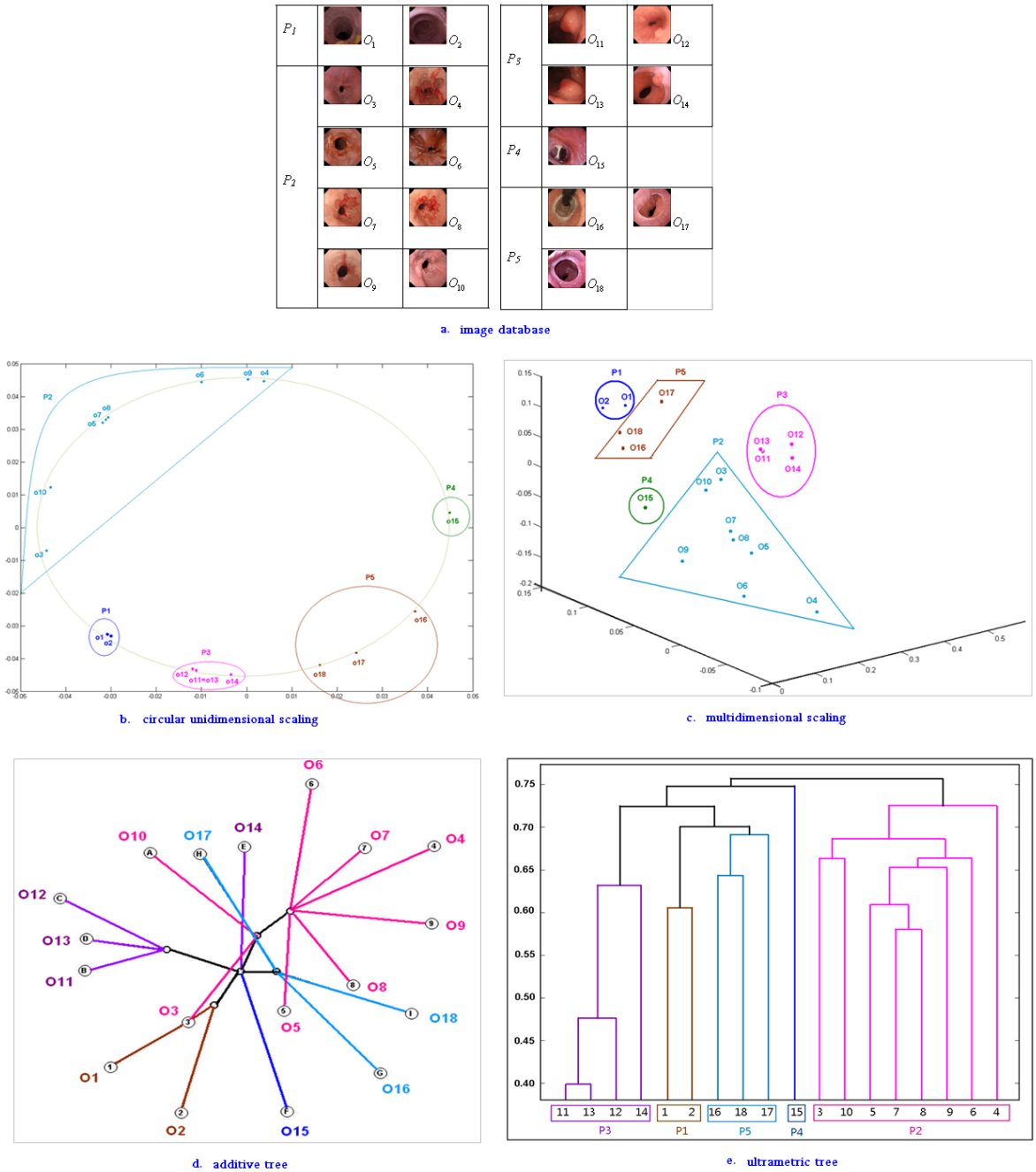


Fig. 5. Possibilistic similarity representational models in a gastroenterology image dataset. a) The medical image database, b) the circular unidimensional scaling model, c) the multidimensional scaling model, d) the additive tree representation, and e) the ultrametric tree representation..

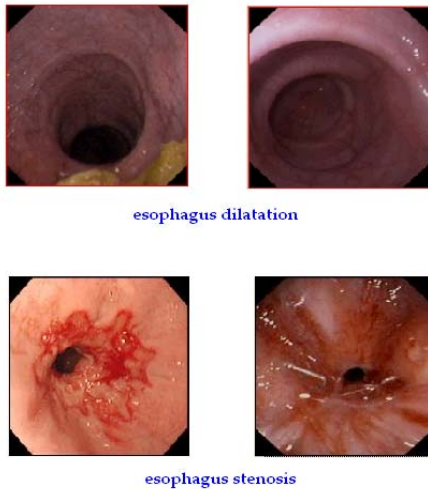


Fig. 6. Two exclusive esophagus lesions (the dilatation and the stenosis), given as an example of information constraints [8].

5. Discussion and conclusion

A general efficient technique to easily manage information challenges in pattern recognition systems has been proposed. The method is essentially based on possibility theory to deal with pattern information element heterogeneity and imperfection, as well as the proportional conflict redistribution approach, in order to take account of information conflict, paradox, and scarcity. This latest approach is applied in the framework of hybrid DS_m model with the goal of pondering the information constraints and frame, concerning the categories in the class set, and information updating when another constraint must be considered due to new knowledge elements (dynamic fusion). We showed also that this model can take into consideration information imperfection associated to the training set element classes, information continuity linked to the class set, and information scalability related to the total output of the system. Despite the proposed method simplicity, it has been able to successfully provide a clear insight to process, within a fusion framework, several complex information problems, commonly encountered in real practical databases.

A corresponding simplified example has been given to numerically present the proposed technique, analyzing patterns between an observed object and two training patterns, permitting to easily follow the different steps and to interpret the results. It also fa-

cilitated the understanding of how to exploit system information sources, by overcoming their inherent limitations and problems.

The applications in satellite and medical image recognition systems, stress the importance and convenience of the proposed approach, by underlying the needs it is intended to fulfill. In fact, the same steps of processing and calculation used in the numeric example can be applied to these applications or to any other domain.

Acknowledgment

The authors are very grateful to C. Le Guillou and J-M. Cauvin from CHU Brest University Hospital, for the access to the gastroenterology database.

References

- [1] A. Dahabiah, J. Puentes, B. Solaiman, Imperfect Pattern Recognition Using the Fuzzy Measure Theory. Springer LNCS, **5788** (2009), 101-108.
- [2] A. Dahabiah, J. Puentes, B. Solaiman, Possibilistic Evidential Clustering. WSEAS AIKED, (2009), 212-217.
- [3] A. Dahabiah, J. Puentes, B. Solaiman, Possibilistic Ordination-based Analysis of an Imperfect Database, IEEE AICCSA, (2009), 199-204.
- [4] A. Dahabiah, J. Puentes, B. Solaiman, Venous Thrombosis Supervised Image Indexing and Fuzzy Retrieval, IEEE EMBC, (2007), 4528-4531.
- [5] A. Martin, C. Osswald, Une Nouvelle Règle de Combinaison Répartissant le Conflit - Applications en Imagerie Sonar et Classification de cibles Radar, Revue Traitement de Signal, **24** (2), (2007), 71-82.
- [6] A. Zemirline, L. Lecornu, B. Solaiman, A New Hybrid Fusion Method for Diagnostic Systems, IPMU, Chapter 28 (2006), 389-401.
- [7] B. Bouchon-Meunier, M. Rifqi, M-J. Lesot, Similarities in Fuzzy Data Mining: From a Cognitive View to Real-World Applications. Springer LNCS, **5050** (2008) 349-367.
- [8] C. Le Guillou, J-M. Cauvin, From Endoscopic Imaging and Knowledge to Semantic Formal Images, Springer, LNCS, **4370** (2007), 189-201.
- [9] E.L. Rissland, AI and Similarity, IEEE Intelligent Systems, **21** (3) (2006), 39-49.
- [10] E. Pecalska, R. Duin, The Dissimilarity Representation for Pattern Recognition, World Scientific Publishing, Chapters 1, 5, 6, and 9, (2005).
- [11] F. Smarandache, J. Dezert, Advances and Applications of DS_mT for Information Fusion, American Research Press Rehoboth, **1**, (2006), 1-141.
- [12] F. Smarandache, J. Dezert, Advances and Applications of DS_mT for Information Fusion, American Research Press Rehoboth, **2**, (2006), 1-106.
- [13] G. Gan, C. Ma, J. Wu, Data Clustering: Theory, Algorithms, and Applications, "Scale Conversion", ASA-SIAM Series on Statistics and Applied Probability, (2007), 25-41.

- [14] G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, C. Roux, Multimodal Information Retrieval Based on DSMT. Application to Computer-Aided medical Diagnosis. *Advances and Applications of DSMT for Information Fusion*, **3**, edited by F. Smarandache, J. Dezert Chapter 18, (2009), 471-502.
- [15] G.J. Klir, *Uncertainty and Information: Foundations of Generalized Information Theory*, Wiley-Interscience, (2005), Chapters 1 and 2.
- [16] H. Rakoto, J. Hermosillo, M., Ruet, Integration of Experience Based Decision Support in Industrial Processes, *IEEE Conference on Systems, Man, Cybernetics*, **7** (2002), 1-6.
- [17] J. Diaz, M. Rifqui, B. Bouchon-Meunier, A Similarity Measure between Basic Belief Assignments, *IEEE Information Fusion Conference*, (2006), 1-6.
- [18] J. Dezert, F. Smarandache, M. Daniel, The Generalized Pignistic Transformation, *Proc. 7th International Conference on Information Fusion*, (2004), 384-391.
- [19] K. Touil, M. Zribi, M. Benjelloun, Application of Transferable Belief Model to Vehicle Navigation System, *Integrated Computer-Aided Engineering*, **14** (1), (2007), 93-105.
- [20] L. Hubert, P. Arabie, Iterative Projection Strategies for the Least-Squares Fitting of Tree Structures to Proximity Data, *British Journal of Mathematical & Statistical Psychology*, **48** (2), (1995), 281-317.
- [21] L. Roux, J. Desachy, Satellite Image Classification Based on Multi-Source Information-Fusion with Possibility Theory, *IGARSS IEEE*, **2** (1994), 885-887.
- [22] L.M. Zouhal, T. Denoeux, An Evidential-Theoretic KNN Rule with Parameter Optimization, *IEEE transactions on Systems, Man, and Cybernetics*, **28** (5), (1998), 263-271.
- [23] M. Zehana, *Connaissances Structurelles et Interprétation d'Images Satellitaires*. Ph.D. Dissertation, Paul Sabatier University, Toulouse, France, 1995.
- [24] M.H. Masson, T. Denoeux, Inferring a Possibility Distribution from Empirical Data, *Fuzzy Sets and Systems*, **157** (3), (2006), 319-340.
- [25] Ph. Smets, Uncertainty management in information systems from needs to solutions, "Imperfect Information: Imprecision and uncertainty", Kluwer Academic Publishers, (1997), 225-254.
- [26] S.S. Stevens, On the Theory of Scales of Measurements, *American Association for the Advancement of Science*, **103** (2684), (1946), 677-680.
- [27] T. Denoeux, M.H. Masson. EVCLUS: Evidential Clustering of Proximity Data, *IEEE Transactions on Systems, Man and Cybernetics*, **34** (1), (2004), 95-109.
- [28] T. Denoeux, Neural Network Classifier Based on Dempster-Shafer Theory, *IEEE Transactions on Systems, Man, and Cybernetics*, **30** (2), (2000), 131-150.
- [29] X. Tao, Y. Li, R., Nayak, A Knowledge Retrieval Model Using Ontology Mining and User Profiling, *Integrated Computer-Aided Engineering*, **15** (4), (2008), 313-329.