A Method of Spatial Unmixing Based on Possibilistic Similarity in Soft Pattern Classification

Bassem Alsahwa, Basel Solaiman, Eloi Bosse, Shaban Almouahed, Didier Gueriot

Fuzzy Inf. Eng. (2016) 8: 295–314

Available online at www.sciencedirect.com
ScienceDirect
Fuzzy Information and Engineering
http://www.elsevier.com/locate/fiae

ORIGINAL ARTICLE

A Method of Spatial Unmixing Based on Possibilistic Similarity in Soft Pattern Classification

B. Alsahwa · B. Solaiman · É. Bossé · S. Almouahed · D. Guériot

Received: 30 January 2016/ Revised: 18 June 2016/
Accepted: 30 July 2016/

Abstract This paper proposes an approach for pixel unmixing based on possibilistic similarity. The approach exploits possibilistic concepts to provide flexibility in the integration of both contextual information and a priori knowledge. Possibility distributions are first obtained using a priori knowledge given in the form of learning areas delimitated by an expert. These areas serve for the estimation of the probability density functions of different thematic classes also called endmembers. The resulting probability density functions are then transformed into possibility distributions using Dubois-Prade's probability-possibility transformation. The pixel unmixing is then performed based on the possibilistic similarity between a local possibility distribution estimated around the considered pixel and the obtained possibility distributions representing the predefined endmembers in the analyzed image. Several possibilistic similarity measures have been tested to improve the discrimination between endmembers. Results show that the proposed approach represents an efficient estimator of the proportion of each endmember present in the pixel (abundances) and achieves higher

B. Alsahwa (✉) · B. Solaiman (✉) · S. Almouahed (✉) · D. Guériot (✉)
Image & Information Processing Department (iTi), Télécom Bretagne, Technopôle Brest Iroise CS 83818, 29238 Brest Cedex, France
emails: Bassem.Alsahwa@telecom-bretagne.eu
Basel.Solaiman@telecom-bretagne.eu
Shaban.Almouahed@telecom-bretagne.eu
Didier.Gueriot@telecom-bretagne.eu
Corresponding Author: É. Bossé (✉)
Expertises Parafuse Inc., 7-1030 Blvd Graham, Mont-Royal, H3P 2G2, Canada
email: ebosse861@gmail.com
Peer review under responsibility of Fuzzy Information and Engineering Branch of the Operations Research Society of China.

http://dx.doi.org/10.1016/j.fiae.2016.11.004

classification accuracy. Performance analysis has been conducted using synthetic and real images.

Keywords Spatial unmixing · Endmembers · Possibilistic similarity · Contextual information

1. Introduction

An accurate and reliable image classification is a crucial task in many applications such as content based image retrieval, medical and remote-sensing image analysis, computer vision and robotics, web image search, visual tracking, and scene interpretation. An important difficulty related to this task stems from the existence of mixed pixels usually called 'mixels'. These mixels contain a mixture of more than one class of different thematic classes also called endmembers contained in the analyzed scene. Endmembers, as illustrated in Fig.1, correspond to macroscopic objects in the scene, such as water, soil, metal, vegetation, etc. Unmixing is critical for image analysis. In [1], they describe unmixing in the following way: "Spectral unmixing is the procedure by which the measured spectrum of a mixed pixel is decomposed into a collection of constituent spectra, or endmembers, and a set of corresponding fractions, or abundances, that indicate the proportion of each endmember present in the pixel."

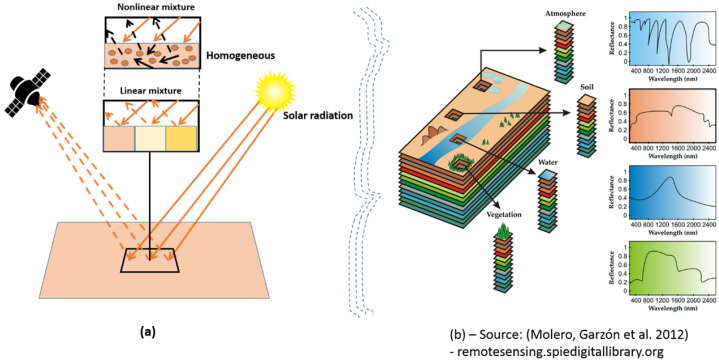

Fig. 1 a) Linear and non-linear pixel mixtures;
b) Concept of hyperspectral imaging [2]

Mixels arise mainly due to the limitation in spatial and spectral resolving capacity of the sensor being used. Spectral measurement might be the result of some composite of the individual spectra, a mixel, if the sensor spatial resolution is low enough that disparate materials can jointly occupy a single pixel, for instance, in remote sensing performing wide-area surveillance at high altitude. Mixels might also result from

distinct materials combined into a homogeneous mixture that is independent of the spatial resolution of the sensor (e.g. sand grains on a beach).

There is a vast and rich literature on methods of unmixing for image processing. Amongst it, from time to time, appear papers that present surveys covering in whole or in parts that large field. For instance, in [1], they present an important review on linear models and non-linear modelling approaches. In [3], they define also a taxonomy of algorithms for hyperspectral unmixing. They consider 'unmixing' as a special case of the generalized inverse problem that estimates parameters describing an object using an observation(s) of a signal that has interacted with the object before reaching the sensor. More recently, in [4], they present a quite elaborated discussion on the mixture problem through techniques named spectral mixture analysis (SMA). SMA models a mixed spectrum as a linear or nonlinear combination (see Fig.1) of its spectral endmembers weighted by their subpixel fractional cover where SMA provides abundances by model inversion. The reader can benefit from the following papers to get a more complete survey on spectral unmixing [5-10]. This is beyond the scope of this paper to present a throughout analysis of unmixing approaches.

In the majority of the hyperspectral literature, unmixing algorithms are rather based on the spectral signatures of each individual pixel and do not incorporate the spatial information. However, some studies, not a lot, present methods to integrate spatial information found in a hyperspectral data cube [9, 10, 12, 13]. The two types of resolution, spatial and spectral, have an inextricable relationship to one another [14] where the higher spectral variability of local areas of the analyzed scene becomes apparent as the spatial resolution becomes finer. Therefore, using advanced sensors with higher spatial resolving power may not necessarily enable improved classifications when the pixel-based images classification systems are used.

The originality of our proposed approach resides in it can work on very few images even on a single image. There are situations with limited means where one does not have access to the richness of data provided by hyperspectral or multispectral imagers but we rather have only few images. The development of methods of pixel unmixing by endmembers becomes very important for image analysis where subpixel detail is valuable and more accurate classification results are needed. Especially in situations where no spectral information is available (i.e., monochromatic sensor with limited capacity) as pointed by the dashed rectangle of Fig.2.

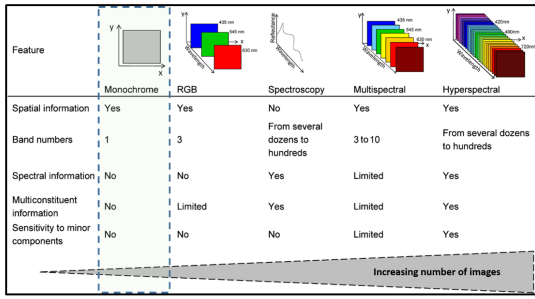

Fig. 2 Tabling features to compare monochrome, RGB, spectroscopy, multispectral, and hyperspectral data, adapted from [11]

This paper is organized as follows. In the next section, a brief review of basic concepts of possibility theory is introduced along with different possibilistic similarity functions to quantify the similarity between classes or endmembers. Our proposed approach is detailed in the subsequent section followed by Section 4 that is devoted to experimental results obtained using synthetic and real images.

2. A Review of Possibility Theory

Possibility theory is a relatively new theory devoted to handle epistemic uncertainty, i.e., uncertainty in the context where the available knowledge is only expressed in an ambiguous form (ex. knowledge expressed by an expert). This theory was introduced by Zadeh in 1978 as an extension of fuzzy sets and fuzzy logic theory to express the intrinsic fuzziness of natural languages as well as uncertain information [15] and further developed by Dubois and Prade [16, 17]. It is well established that probabilistic reasoning, based on the use of a probability measure, constitutes the optimal approach dealing with uncertainty. In the case where the available knowledge is ambiguous and encoded as a membership function into a fuzzy set defined over the decision set, the possibility theory transforms each membership value into a possibilistic interval of possibility and necessity measures [18]. The use of these two dual measures in possibility theory makes the main difference from the probability theory. Besides, possibility theory is not additive in terms of beliefs combination and makes sense on ordinal structures [19]. In the following subsections, the basic concepts of a possibility distribution, the dual possibilistic measures (i.e., possibility and necessity measures), the probability-possibility transformation and the possibilistic decision rules are briefly recalled.

2.1. Possibilistic Knowledge Representation

Let us consider an exclusive and exhaustive universe of discourse $\Omega = \{C_1, C_2, \cdots, C_M\}$ formed by M mutually exclusive elementary decisions (e.g., decisions, thematic classes, endmembers, hypothesis, etc), and let $\{\pi_{C_m}\}, m = 1, \cdots, M$ be a set of M possibility distributions defined on Ω. Exclusiveness means that one and only one decision may occur at one time, whereas exhaustiveness states that the occurring decision certainly belongs to Ω. Possibility theory is based on the notion of possibility distribution denoted by π, which maps elementary decisions from Ω to the interval $[0, 1]$, thus encoding "our" state of knowledge or belief, on the possible occurrence of each element $C_m \in \Omega$. The value $\pi(C_m)$ represents to what extent it is possible for C_m to be the unique occurring decision. This value $\pi(C_m)$ encodes our state of knowledge, or belief, about the real world and represents the possibility degree for C_m to be the unique occurring element. In this context, two extreme cases of knowledge are given:

- Complete knowledge: $\exists!\, C_m \in \Omega,\ \pi(C_m) = 1$ and $\pi(C_n) = 0,\ \forall\, C_n \in \Omega,\ C_n \neq C_m$.
- Complete ignorance: $\forall\, C_m \in \Omega,\ \pi(C_m) = 1$ (all elements from Ω are considered as totally possible).

$\pi(\cdot)$ is called a normal possibility distribution if at least one element C_{m_0} from Ω such that $\pi(C_{m_0}) = 1$.

2.2. Possibility Distributions Estimation Based on $Pr - \pi$ Transformation

A crucial step in possibility theory applications is the determination of possibility distributions. Recall that a possibility distribution encodes our state of knowledge about the real world. Nevertheless, the appropriate estimation of the possibility distribution shape and shape's parameters is a difficult task. Two approaches are generally used for the estimation of a possibility distribution. The first approach consists on using standard forms predefined in the framework of fuzzy set theory for membership functions (i.e., triangular, Gaussian, trapezoidal, S-Shape, etc.), and tuning the shape-parameters using a manual or an automatic tuning method. The second estimation approach of possibility distributions is based on the use of statistical data and is conducted in the following two consecutive steps:

- Using statistical data, an uncertainty function describing the uncertainty inherent to the statistical data is estimated first (e.g. histogram, probability density function, basic belief function, etc.).
- The estimated uncertainty function is then transformed into a possibility distribution. In the framework of possibility theory, the probability-possibility transformations ($Pr-\pi$ transformations) are frequently used for the implementation of this step.

In this study, the available expert's knowledge is expressed through the definition of small learning areas representing different endmembers, i.e., statistical data. The second estimation approach will then be used. Several $Pr-\pi$ transformations are proposed in the literature. Dubois et al. [20] suggested that any $Pr-\pi$ transformation of a probability density function, Pr, into a possibility distribution, π, should be guided by the two following principles:

- The probability-possibility consistency principle. This principle is expressed by Zadeh [15] as: "what is probable is possible". Dubois and Prade formulated this principle by indicating that the induced possibility measure Π should encode upper probabilities:

$$\Pi(A) \geq \Pr(A), \forall A \subseteq \Omega. \tag{1}$$

- The preference preservation principle ensuring that any $Pr - \pi$ transformation should satisfy the relation:

$$\Pr(A) < \Pr(B) \Leftrightarrow \Pi(A) < \Pi(B), \forall A, B \subseteq \Omega. \tag{2}$$

Verifying these two principles, a $Pr - \pi$ transformation turning a probability distribution Pr, defined by probability values $Pr(\{C_m\})$, $C_m \in \Omega$, $m = 1, 2, \cdots, M$, into a possibility distribution π, defined by $\Pi(\{C_m\})$, $C_m \in \Omega$, $m = 1, 2, \cdots, M$ has

been suggested by Dubois et al. [20]. This transformation, called symmetric $Pr - \pi$ transformation, is defined by:

$$\pi(C_m) = \Pi(\{C_m\}) = \sum_{j=1}^{M} \min\left[\Pr(\{C_m\}), \Pr\left(\left\{C_j\right\}\right)\right]. \tag{3}$$

This transformation is being used in our study to transform the probability distributions into possibility distributions. The reason standing behind this choice is due to good performance of that symmetric transformation (3) provides in pattern recognition and classification [21, 22].

2.3. Possibilistic Similarity

The concept of similarity is a very important topic for many applications. Any system that needs to analyze or organize automatically a set of data or knowledge must use a similarity operator to estimate relations and resemblances that exist in data [23]. The issue of comparing imperfect pieces of information depends on the way they are represented. In the case of possibility theory, comparing uncertain pieces of information is to compare their possibility distributions. Hence, a similarity measure is a quantification of the amount of similarity between two possibility distributions.

Considering the expert's predefined set of M endmembers contained in the analyzed image, $\Omega = \{C_1, C_2, \cdots, C_M\}$, a set of M possibility distributions can be defined as follows:

$$\pi_{C_m} : D_m \rightarrow [0, 1], \tag{4}$$

$$x(P) \rightarrow \pi_{C_m}(x(P)), \tag{5}$$

where D refers to the definition domain of the observed feature $x(P)$ of the pixel P. For each class C_m, $\pi_{C_m}(x(P))$ associates each pixel $P \in J$, of an image J observed through a feature $x(P) \in D$, with a possibility degree of belonging to the class C_m, $m = 1, \cdots, M$. Different possibilistic similarity and distance functions "Sim" can be defined between the two possibility distributions π_{C_m} and π_{C_n} of two endmembers C_m and C_n of the set Ω. The behavior of these functions can be studied in order to obtain a better discrimination between classes C_m and C_n. To this end, calculating a similarity matrix S informs on such inter-classes behavior and help in the choice of a suitable measure in a given context:

$$S = \begin{bmatrix} Sim\left(\pi_{C_m}, \pi_{C_m}\right) & Sim\left(\pi_{C_m}, \pi_{C_n}\right) \\ Sim\left(\pi_{C_n}, \pi_{C_m}\right) & Sim\left(\pi_{C_n}, \pi_{C_n}\right) \end{bmatrix}. \tag{6}$$

2.3.1. Possibilistic Similarity Functions

This subsection reviews some existing possibilistic similarity and distance functions that are the most frequently used in literature. Recall that one considers an exclusive and exhaustive universe of discourse $\Omega = \{C_1, C_2, \cdots, C_M\}$ formed by M mutually

exclusive elementary decisions, and let $\{\pi_{C_m}\}$, $m = 1, \cdots, M$ be a set of M possibility distributions defined on Ω.

- Information closeness: This similarity measure was proposed by Higashi and Klir [24] based on the information variation measure G:

$$G(\pi_{C_m}, \pi_{C_n}) = g(\pi_{C_m}, \pi_{C_m} \vee \pi_{C_n}) + g(\pi_{C_n}, \pi_{C_m} \vee \pi_{C_n}), \tag{7}$$

where $g(\pi_{C_m}, \pi_{C_n}) = U(\pi_{C_n}) - U(\pi_{C_m})$. The operator $\vee$ is taken as maximum operator and U is the non-specificity measure defined as in (8). Given an ordered possibility distribution π such that $1 = \pi_1 \geq \pi_2 \geq \cdots \geq \pi_M$, the U of π is formulated as:

$$U(\pi) = \left[\sum_{i=1}^{M} (\pi_i - \pi_{i+1}) \log_2 i\right] + (1 - \pi_1) \log_2 M, \tag{8}$$

where $\pi_{M+1} = 0$ by convention. Hence, the similarity measure based on the Information closeness is given by:

$$\mathrm{Sim}_G(\pi_{C_m}, \pi_{C_n}) = 1 - \frac{G(\pi_{C_m}, \pi_{C_n})}{G_{Max}}. \tag{9}$$

- Minkowski distance: Since possibility distributions are often represented as vectors, the most popular metrics for possibility distributions are induced by the Minkowski norm (Lp) which is used in vector spaces.

$$\mathrm{L}_p(\pi_{C_m}, \pi_{C_n}) = \sqrt[p]{\sum_{i=1}^{|D|} \left|\pi_{C_m}(x_i) - \pi_{C_n}(x_i)\right|^p}. \tag{10}$$

Two particular cases of (10) are often investigated: L_1-norm (Manhattan distance), and L_2-norm (Euclidean distance). They are given by the following expressions:

$$\mathrm{L}_1(\pi_{C_m}, \pi_{C_n}) = \sum_{i=1}^{|D|} \left|\pi_{C_m}(x_i) - \pi_{C_n}(x_i)\right|, \tag{11}$$

$$\mathrm{L}_2(\pi_{C_m}, \pi_{C_n}) = \sqrt{\sum_{i=1}^{|D|} \left|\pi_{C_m}(x_i) - \pi_{C_n}(x_i)\right|^2}. \tag{12}$$

These cases of Minkowski distance can be transformed into similarity measure by the following:

$$\mathrm{Sim}_p(\pi_{C_m}, \pi_{C_n}) = 1 - \frac{\mathrm{L}_p}{\sqrt[p]{|D|}}. \tag{13}$$

- Information affinity: This similarity measure was proposed by Jenhani et al. [25].

$$\mathrm{Sim}_{IA}\left(\pi_{C_m}, \pi_{C_n}\right) = 1 - \frac{\kappa \mathrm{L}_p\left(\pi_{C_m}, \pi_{C_n}\right) + \lambda \mathrm{Inc}\left(\pi_{C_m}, \pi_{C_n}\right)}{\kappa + \lambda}, \tag{14}$$

where $k > 0$ and $\lambda > 0$, $Inc(\pi_{C_m}, \pi_{C_n})$ represents the inconsistency degree between π_{C_m} and π_{C_n} defined as follows:

$$\mathrm{Inc}\left(\pi_{C_m}, \pi_{C_n}\right) = 1 - \max\left(\min\left(\pi_{C_m}, \pi_{C_n}\right)\right). \tag{15}$$

- Similarity index [25]:

$$\mathrm{Sim}_{SI}\left(\pi_{C_m}, \pi_{C_n}\right) = \min\left\{\alpha\left(\pi_{C_m}, \pi_{C_n}\right), \alpha\left(1 - \pi_{C_m}, 1 - \pi_{C_n}\right)\right\}, \tag{16}$$

where:

$$\alpha\left(\pi_{C_m}, \pi_{C_n}\right) = \frac{\sum_{i=1}^{|D|} \pi_{C_m}(x_i) \times \pi_{C_n}(x_i)}{\sum_{i=1}^{|D|} \left\{\max\left(\pi_{C_m}(x_i), \pi_{C_n}(x_i)\right)\right\}^2}. \tag{17}$$

It is worth noticing that this list should not be considered as complete depending on the application or the type of images that have to be analyzed. In the following subsection, a process allowing the selection of the most "suitable" similarity measure is proposed and evaluated using synthetic images.

2.3.2. Evaluation of the Similarity between Two Classes

A 100×100 synthetic image composed of two thematic classes has been generated in order to evaluate the similarity between two classes. The intensity of the pixels from C_1 and C_2 are generated as two Gaussian distributions $G(m_1 = 110,\ \sigma_1 = 10)$ and $G(m_1 = 120,\ \sigma_1 = 10)$. Fig.3 illustrates the obtained synthetic image. The evaluation principle of the similarity between the two classes selects the possibilistic similarity function for which similarity matrix is the closest to the identity matrix I_2 in terms of the Euclidean distance. In the considered case, where only two classes are involved, this distance D is summarized by the following measure $(i, j \in \{0, 1\})$:

$$D = \sqrt{\sum_{i,j} \left[S(i, j) - I_2(i, j)\right]^2}. \tag{18}$$

Lower is the distance value D, better is the discrimination power (between classes) of that similarity function. The distance value D is computed for each similarity function by first varying the mean value of the generated pixels of class C_2 and then the standard deviation, while maintaining a fixed value for mean and standard deviation of class C_1. Fig.3 shows the evolution of the measured D with respect to means or standard deviations. From Fig.3, the similarity function called "Similarity Index"

$\mathrm{Sim}_{SI}(\pi_{C_m}, \pi_{C_n})$ function [25] sounds like the most suitable to describe the similarity between two classes. Indeed, the possibilistic similarity function Sim_{SI} tends to the identity matrix faster than the others when the values $m_2 - m_1$ and $\sigma_2 - \sigma_1$ increase.

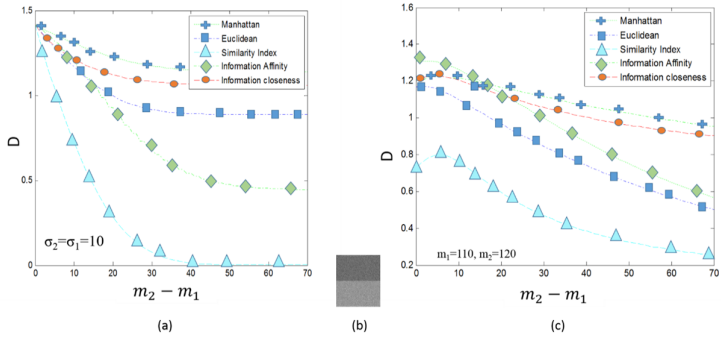

Fig. 3 Estimating the power of discrimination D between classes of different similarity functions: (a) Evolution of the measure D as a function of the difference of deviations between classes C_1 and C_2; (b) Synthetic image; (c) Evolution of the measure D as a function of the mean difference between classes C_1 and C_2

3. A Possibilistic Approach of Pixel Unmixing by Endmembers

In this paper, an approach based on possibilistic similarity is proposed for pixel unmixing by endmembers. This approach exploits possibilistic concepts to provide flexibility, on one side, in representing limited information (contextual information and a priori knowledge), and on the other side, in the integration of both contextual information and a priori knowledge. For instance, in pixel-based image segmentation context, several methods have been proposed in the literature to cope with limited initial a priori knowledge. The method of image segmentation by region growing [26] consists on selecting, manually, few "seeds" designating anchor points for the initialization for the segmentation of regions contained in the image. According to a criterion of similarity measure, a region grows iteratively in merging pixels similar and adjacent to starting seeds. Fan et al. [27] presented an extensive and comparative study on seeded region growing approaches. These semantic image segmentation approaches, according to [27], suffer from two main problems: pixel sorting orders for labeling and automatic seed selection.

These approaches are called 'semantic' because they involve high-level knowledge of image components in the seeds selection procedure as in our proposed approach. One of the methods which has also been used to compensate the limitation of initial prior knowledge is the semisupervised fuzzy pattern matching method [21] based on transforming different class histograms (established using the small size learning sets) into possibility distributions using the Dubois-Prade $Pr - \pi$ transformation. The use of reference class patterns (serving for the decision of attributing a new pattern to the different predefined classes) is applied on each new pattern. When a new pattern

is “accepted” as one of the predefined classes, then, this pattern is used to adjust the class histogram, and hence, the class possibility distribution. This method does not exploit the spatial context in the classification process.

Alsahwa et al. [22] also presented a method to enrich the limited initial prior knowledge. This method is based on transforming different class probability density functions into possibility distributions using the Dubois-Prade $Pr - \pi$ transformation and on using spatial context to evaluate the decision of attributing a new pattern to the different predefined classes and hence, updating their possibility distributions. Possibility distributions are first obtained using a prior knowledge given in the form of learning areas delimitated by an expert. These areas serve for the estimation of the probability density functions of different endmembers. The resulting probability density functions are then transformed into possibility distributions using Dubois-Prade’s probability-possibility transformation [20, 28].

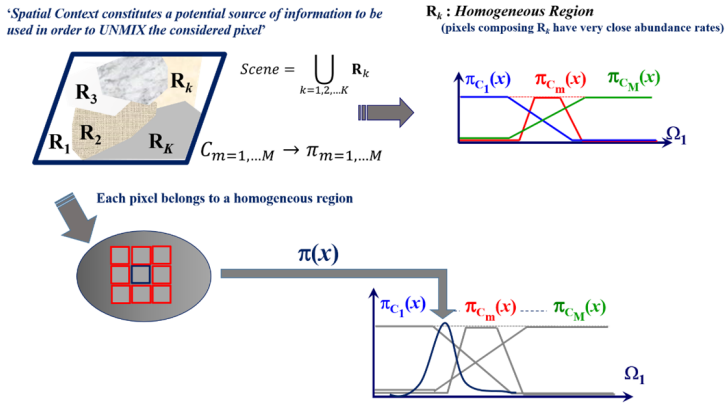

Fig. 4 A possibilistic approach of pixel unmixing

The estimation of these M possibility distributions forms the first step in the proposed approach (Fig.4). The second step consists in determining the pixels’ similarity to the M predefined classes of the analyzed image J by first estimating the local possibility distribution around the pixel of interest P_0. Secondly, by measuring the similarity Sim_{SI} between this local possibility distribution and each of the M estimated possibility distributions. Fig.5 shows the proposed approach in the case of the synthetic image composed of two classes.

All the measured similarity values are transformed into percentages as the following:

$$a_k = \mathrm{Sim}_{SI}\left(\pi_{C_k}, \pi_{P_0}\right) \Big/ \sum_{m=1}^{M} \mathrm{Sim}_{SI}\left(\pi_{C_m}, \pi_{P_0}\right), \tag{19}$$

where a_k is supposed to be the “abundance” of k^{th} predefined thematic class in the considered pixel P_0 and $\sum \mathrm{Sim}_{SI}$ serves as a normalizing factor. It is worth noticing

that high overlapping case (high discrimination complexity) between the predefined thematic classes can be treated with our proposed approach. In the case of low overlapping (low discrimination complexity), the "abundance" of a predefined thematic class in the considered pixel P_0 is roughly inversely proportional to the distance between the pixel vector and the mean of that class [29].

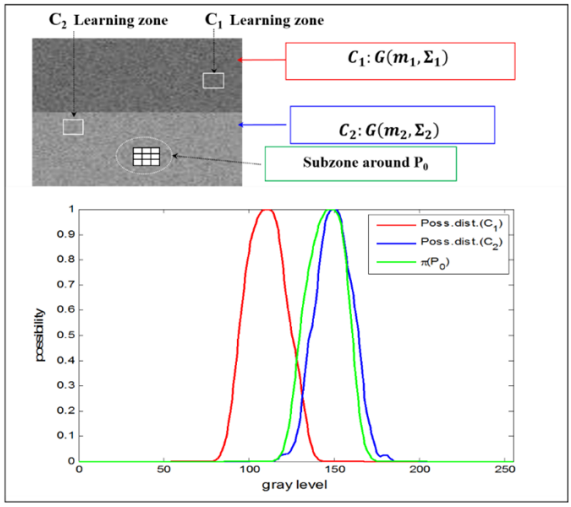

Fig. 5 Synthetic image, possibility distributions of classes C_1, C_2 and the local possibility distribution in a subzone around the pixel of interest P_0

The simplest and most widely used approach, the linear mixture model [30] is used in the proposed unmixing approach. This model is based on the assumption that a linear combination exists between the pixel brightness and the M predefined thematic class. The spectral reflectance of a pixel is the sum of the spectral reflectance from the predefined endmembers weighted by their relative "abundances":

$$\mathrm{B} = \sum_{i=1}^{M} a_i \times B_i, \tag{20}$$

where B is the brightness value (i.e., the realization of the random variable measured by the sensor) of the considered pixel P_0, B_i is the brightness value of the i^{th} predefined thematic class, and a_i is its abundance in the considered pixel P_0. There are two constraints on the abundances that should be satisfied: the abundances must all be non-negative to be meaningful in a physical sense $a_i \geq 0$ [31], and must sum up to one ($\sum a_i = 1$). In the case of multispectral or hyperspectral images composed of N bands, (20) is calculated in each band. If $N > M$ (number of the predefined endmembers), a least squares procedure can be used to obtain the best abundances' estimation.

A classification step is conducted at the end of the proposed approach. This step consists in the process of assigning a class to the considered pixel P_0 by determining the nearest class via the similarity function Sim_{SI} used to measure the similarity between this pixel's local possibility distribution and the possibility distributions of each of the M classes.

4. Experimental Results

4.1. Performance on Synthetic Data

In many applications, collecting mixed pixels and determining their exact abundances of the predefined thematic classes is a very difficult task. Therefore, a 550×550 pixels synthetic image, given in Fig.6, is generated. This image is composed of eleven sectors. The first and second sectors are assumed to contain two "pure" thematic classes generated by two Gaussian distributions $G(m_1 = 100,\ \sigma_1 = 15)$ and $G(m_2 = 150,\ \sigma_2 = 15)$. Pixels of sectors from three to eleven are generated as a linear mixture of the first and second pure classes.

The abundances of class C_1 and class C_2 in these mixed pixels is varying incrementally by 10%. For instance, the abundance of class C_1 in the third sector is 10% (resp. abundance of class C_2 is 90%) and in the fourth sector 20% (resp. abundance of class C_2 is 80%), etc. 20×20 pixel learning zones positioned by the expert (as being representative areas of the considered thematic classes) are also illustrated on the generated image. The effectiveness evaluation of the proposed approach is studied through two perspectives (next two sections): an estimation of classes' abundances in the mixed pixels and the evaluation of the improvement in overall classification accuracy.

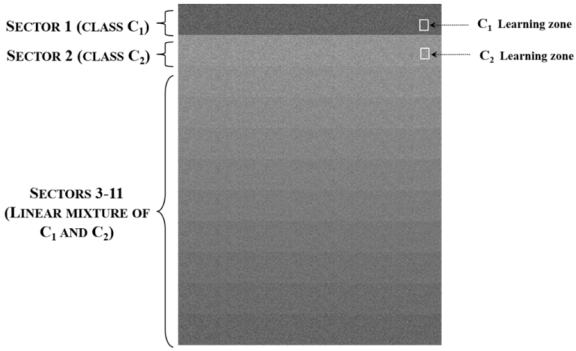

Fig. 6 Synthetic image composed of two classes and their learning zones

4.2. Estimation of Classes's Abundances in the Mixed Pixels

Using the learning zones, the initial estimation of the class probability density functions are established based on the kernel density estimation (KDE) approach. The application of the $Pr - \pi$ Dubois-Prade's transformation allows obtaining the possibility

distributions for each class in the analyzed image. A 3×3 pixel window centred on each pixel (see Fig.4) is considered as the local spatial possibilistic context and then local probability density functions are established based on the KDE approach. The application of the $Pr - \pi$ Dubois-Prade's transformation allows obtaining the local possibility distributions.

Abundances of the predefined endmembers in each sector, from three to eleven, can be estimated from the possibilistic similarity values. In each of these sectors, the proposed approach, using the possibilistic similarity measure Sim_{SI}, is applied on all its pixels and their possibilistic similarity values of each endmember are calculated. The results obtained, in terms of abundance maps, are given in Fig.7. A visual inspection of the abundance maps shows a linear variation in the abundance values corresponding to both classes. A quantitative analysis step of abundances maps is conducted to confirm the visual inspection. The mean and the standard deviation of the possibilistic similarity values for each class are given in Table 1.

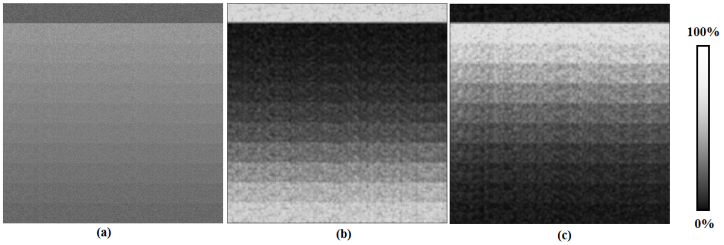

Fig. 7(a) Synthetic image composed of two classes; (b) abundance map of class C_1 and (c) abundance map of class C_2

Table 1: Abundances of the predefined endmembers in each sector.

	C_1(10%) C_2(90%)	C_1(20%) C_2(80%)	C_1(30%) C_2(70%)	C_1(40%) C_2(60%)	C_1(50%) C_2(50%)	C_1(60%) C_2(40%)	C_1(70%) C_2(30%)	C_1(80%) C_2(20%)	C_1(90%) C_2(10%)
mean(C_1)	0.14	0.20	0.28	0.39	0.50	0.61	0.72	0.79	0.87
std(C_1)	0.07	0.08	0.09	0.10	0.10	0.09	0.08	0.08	0.06
mean(C_2)	0.86	0.80	0.72	0.61	0.50	0.39	0.28	0.21	0.13
std(C_2)	0.06	0.07	0.09	0.10	0.10	0.10	0.09	0.07	0.06

Results show that the abundances of the predefined endmembers in the mixed pixels can be estimated with a reasonable accuracy from possibilistic similarity values. This estimation is in conformity with the values used in synthetic image generation. For instance, it can be estimated that the fifth sector contains about 28% of class C_1 and 72% of class C_2 (Fig.7) while the used values in synthetic image generation are 30% of class C_1 and 70% of class C_2. The small values of standard deviation constitute another indication confirming that this estimation is quite consistent to the values used in synthetic image generation.

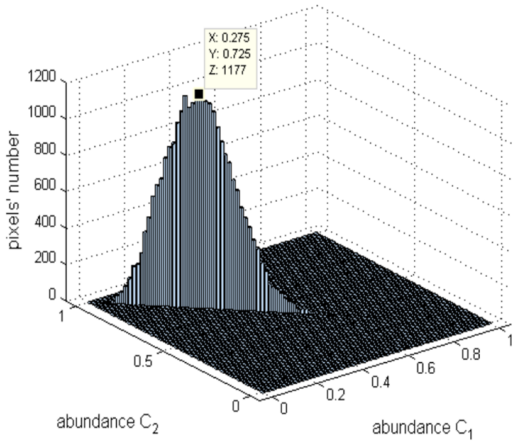

Fig. 8 Abundances of the predefined endmembers in the mixed pixels of the fifth sector

4.3. Evaluation of the Improvement in Overall Classification Accuracy

The above synthetic image (Fig.6) is classified using the proposed approach and the conventional Bayesian approach, respectively. The classification recognition rate is then calculated in order to compare the classification results of the two approaches.

Table 2: Classification recognition rate of the predefined thematic classes in each sector calculating first by the proposed approach and second by the Bayesian approach.

	Recognition rate (%)								
	C_1(10%) C_2(90%)	C_1(20%) C_2(80%)	C_1(30%) C_2(70%)	C_1(40%) C_2(60%)	C_1(50%) C_2(50%)	C_1(60%) C_2(40%)	C_1(70%) C_2(30%)	C_1(80%) C_2(20%)	C_1(90%) C_2(10%)
Our approach(C_1)	0	1	1	11	49	93	99	100	100
Our approach(C_2)	100	99	99	89	51	7	1	0	0
Bayesian approach(C_1)	2	4	12	27	51	28	88	95	99
Bayesian approach(C_2)	98	96	88	73	49	72	12	5	1

Results show an overall improvement in classification accuracy using the proposed approach. This improvement has reached 17% in some cases (e.g. C_1(40%) and C_2(60%)). In addition to this improvement in terms of the classification accuracy, the estimation of the classes' abundances in the mixed pixels (Sec. 4.2) enables the assessing of the classification accuracy which, in turns, may contribute to the interpretation of the analyzed scene. For instance, the classification of the third sector is 100% class C_1 with a small deviation of the assignment to its pixels (about 14% of the class C_2) while the classification result of the fourth sector is also about 100% class C_1 but with a bigger deviation of the assignment to its pixels (about 20% of class C_2).

It is important to note that this assessment of accuracy cannot be obtained using the conventional pixel-based images classification systems.

4.4. Experimental Results Using Real Medical Images

This section presents results on the application of our proposed approach on real images. In certain types of images, the description of thematic classes by an expert is a difficult task. Consequently, it makes the interpretation process more difficult. For example, in the case of mammography, the difficulty of interpreting these images is mainly due to the variety of tissues density, complicated structures of the breast, the great diversity existing in tumor areas in terms of type, shape, contours, etc. Thus, the expert often needs an interpretation support system to perform a preliminary segmentation of these images. Unfortunately, no segmentation method can ensure a fully reliable segmentation that gives a clear idea on tumor areas [32].

The performance of our proposed approach for highlighting the content of regions of interest (i.e., over-density areas in the mammary gland) is evaluated using mammographic images. Fig.9 illustrates the difficulty of interpreting such images with low contrast and highly textured. The image is first segmented using our proposed approach to have a first idea of regions of interest (Fig.10).

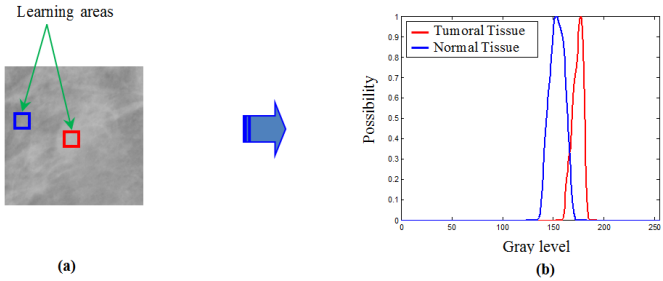

Fig. 9 (a) Mammographic test image (with learning areas) (b) possibility distributions

The abundance map corresponding to the class of tumor tissue obtained by our approach is given in Fig.10. The image is superimposed on the original image. Each region identified in the segmentation process can be described by a vector which components are the abundance rates histograms of thematic classes (present in the analyzed scene) and calculated on the set of pixels of the region (see Fig.11).

A visual analysis of the obtained results shows the spatial distribution of abundance values of tumor tissue while highlighting areas with high abundance values. Such analysis allows the expert to provide a better idea about the content and location of the tumor specially and that is a very important result: the small regions of interest as well as the extent of tumor tissues. This is very valuable information to the clinician from a finer interpretation of the analyzed images.

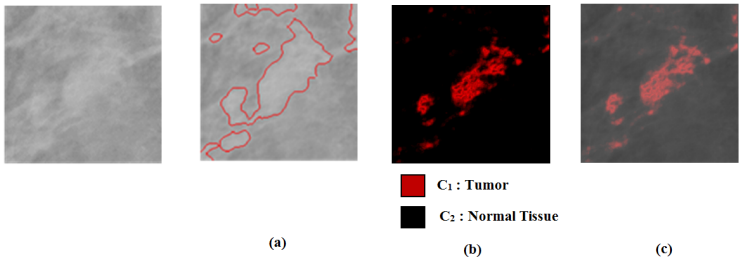

Fig. 10 (a) Mammographic test image segmented by the proposed approach (b) abundance map of Tumor and (c) superimposition of abundance map on the original image

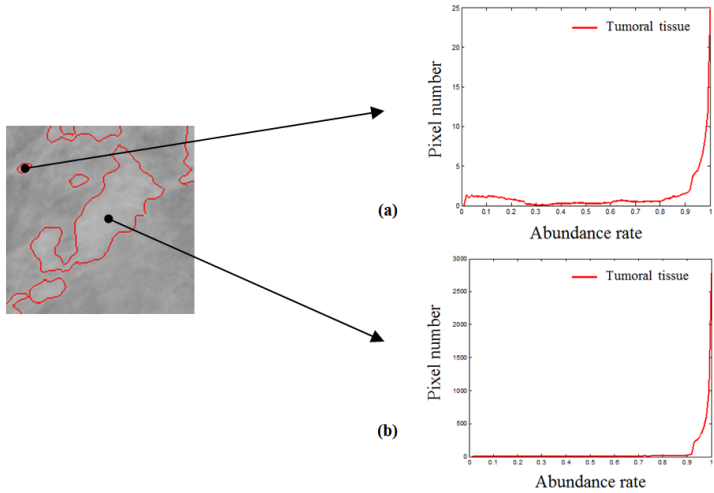

Fig. 11 (a) and (b) abundances histogram of the tumor tissue class in the suspected region

4.5. Experimental Results Using Real Remote Sensing Images

In this section, the performance of the proposed approach is evaluated using a remote sensing image (SPOT: Earth Observation Satellite). This image is a multi-spectral image with three spectral bands (Fig.12) and is composed of four thematic classes (Fig.13) [33].

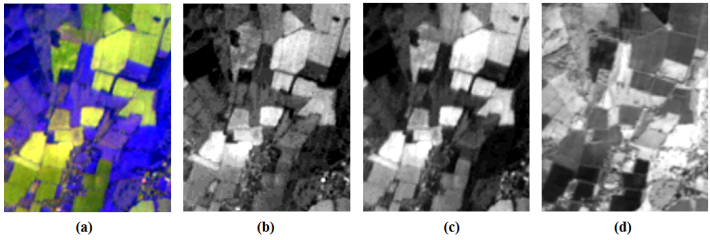

Fig. 12 (a) SPOT image composed of four classes (b) first band (c) second band and (d) third band

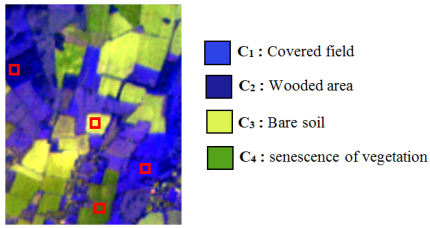

Fig. 13 SPOT image composed of four classes (with learning areas)

For each source of information (i.e., spectral band in this case), the possibility distributions (estimated from the learning areas) produce the abundance maps for all thematic classes. For each thematic class, the final abundance map is obtained as the average of abundance maps obtained from different sources of information. The results obtained by applying the proposed approach on the SPOT image, in terms of abundance maps are given in Fig.14.

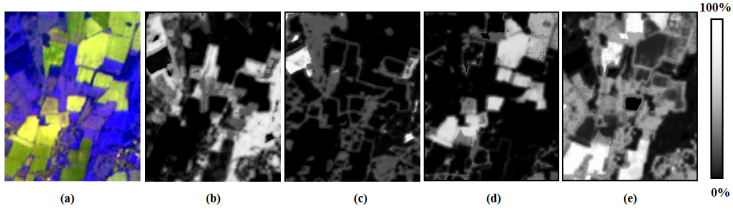

Fig. 14 (a) SPOT image (b) abundance map of class C_1 (c) abundance map of class C_2 (d) abundance map of class C_3 and (e) abundance map of class C_4

A visual analysis of these abundance maps shows an overlap, on one side, between the classes C_1 (covered field), C_2 (wooded area) and C_4 (senescence of vegetation), and on the other side, between the C_3 class (bare soil) and class C_4 (senescence of vegetation). On the other hand, little overlap is found between the class C_3 (bare soil)

and two classes C_1 and C_2. Fig.15 shows a selected region and the four histograms representing the statistical distributions of abundance rates of different classes in this region.

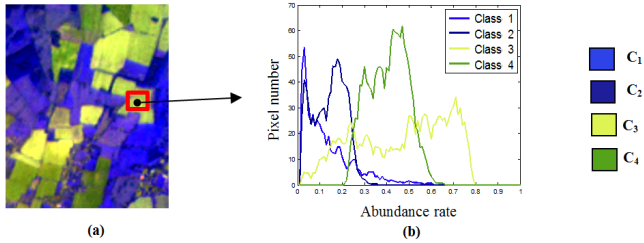

Fig. 15 (a) SPOT image (with selected region) and (b) abundances histograms of the four classes in the selected region

These histograms can confirm/deny the pure aspect of an identified region and provide essential information on its qualitative (i.e., thematic classes) and quantitative (i.e., abundance rates of thematic classes) content. Note that the visual inspection of the region shown in Fig.15, suggesting that it is a mixed region with only the thematic content: C_3 and C_4, however, we note the presence of significant abundance rate of the other two thematic classes C_1 and C_2.

5. Conclusion

In this study, a pixel-based unmixing approach is developed based on a possibility theory that enables the integration of contextual information and a priori knowledge. Indeed, one of the key points of the proposed approach is to characterize the pixel taking into account its neighborhood. This is done through the creation of local possibility distributions. Another key point of our approach has been to propose a classification method based on the similarity between class possibility distribution and local possibility distribution and not on the membership degrees of parameters extracted from the local window. Results obtained on a synthetic image compared to the results obtained using a Bayesian approach are promising. Information about pixel's content of the predefined endmembers are made available and increased classification accuracy has been achieved. Preliminary results of real images seems promising. Future work will generalize and validate these results on various types of images (synthetic and real) and study the sensibility to learning (e.g. number of predetermined endmembers).

Acknowledgments

Authors would like to thank referees for their helpful comments.

References

[1] N. Keshava, J.F. Mustard, Spectral unmixing, IEEE Signal Processing Magazine 19 (2002) 44-57.

[2] J.M. Molero, E.M. Garzon, I. Garcia, A. Plaza, Anomaly detection based on a parallel kernel RX algorithm for multicore platforms, Journal of Applied Remote Sensing 6(1) (2012) 623-632.

[3] N. Keshava, J.P. Kerekes, D.G. Manolakis, G.A. Shaw, Algorithm taxonomy for hyperspectral unmixing, International Society for Optical Engineering 4049 (2000) 42-63.

[4] B. Somers, G.P. Asner, L. Tits, P. Coppin, Endmember variability in spectral mixture analysis: A review, Remote Sensing of Environment 115(7) (2011) 1603-1616.

[5] Y. Altmann, N. Dobigeon, S. McLaughlin, J.Y. Tourneret, Nonlinear spectral unmixing of hyperspectral images using Gaussian processes, IEEE Transactions on Signal Processing 61 (2013) 2442-2453.

[6] D.G. Goodenough, T. Han, A. Dyk, Comparing and validating spectral unmixing algorithms with AVIRIS imagery, Canadian Journal of Remote Sensing 34 (2008) S82-S91.

[7] Z. Guo, T. Wittman, S. Osher, L1 unmixing and its application to hyperspectral image enhancement, SPIE Defense, Security, and Sensing (2009) doi: 10.1117/12.818245.

[8] C. Quintano, A. Fernandez-Manso, Y.E. Shimabukuro, G. Pereira, Spectral unmixing, International Journal of Remote Sensing 33 (2012) 5307-5340.

[9] C. Shi, L. Wang, Incorporating spatial information in spectral unmixing: A review, Remote Sensing of Environment 149 (2014) 70-87.

[10] A. Zare, Spatial-spectral unmixing using fuzzy local information, IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2011, pp.1139-1142.

[11] Q. Li, X. He, Y. Wang, H. Liu, D. Xu, F. Guo, Review of spectral imaging technology in biomedical engineering: Achievements and challenges, Journal of Biomedical Optics 18(10) (2013) 100901.

[12] A. Zare, P. Gader, Piece-wise convex spatial-spectral unmixing of hyperspectral imagery using possibilistic and fuzzy clustering, IEEE International Conference on Fuzzy Systems, 2011, pp.741-746.

[13] A. Zare, K. Ho, Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing, IEEE Signal Processing Magazine 31 (2014) 95-104.

[14] P.M. Mather, B. Tso, Classification Methods for Remotely Sensed data, CRC press, 2009.

[15] L. Zadeh, Fuzzy sets as the basis for a theory of possibility, Fuzzy Sets and Systems 1 (1978) 3-28.

[16] D. Dubois, H. Prade, Possibility Theory: An Approach to Computerized Processing of Uncertainty, Springer Science & Business Media, 2012.

[17] D. Dubois, H. Prade, Possibility theory: Qualitative and quantitative aspects, Quantified Representation of Uncertainty and Imprecision, Springer Netherlands, 1998, pp.169-226.

[18] D. Dubois, Fuzzy Sets and Systems: Theory and Applications, Academic Press, Orlando, 1980.

[19] D. Dubois, H. Prade, When upper probabilities are possibility measures, Fuzzy Sets and Systems 49 (1992) 65-74.

[20] D. Dubois, H. Prade, S. Sandri, On possibility/probability transformations, Fuzzy Logic, Springer Netherlands, 1993, pp.103-112.

[21] M.S. Mouchaweh, Semi-supervised classification method for dynamic applications, Fuzzy Sets and Systems 161 (2010) 544-563.

[22] B. Alsahwa, B. Solaiman, S. Almouahed, É Boss, D. Guriot, Iterative refinement of possibility distributions by learning for pixel-based classification, IEEE Transactions on Image Processing 25 (2016) 3533-3545.

[23] G. Bisson, La similarité: une notion symbolique/numérique. Apprentissage symbolique-numérique: Eds Moulet, Brito, Cepadues Edition, 2000.

[24] M. Higashi, G.J. Klir, On the notion of distance representing information closeness: Possibility and probability distributions, International Journal of General Systems 9 (1983) 103-115.

[25] I. Jenhani, N.B. Amor, Z. Elouedi, S. Benferhat, K. Mellouli, Information affinity: A new similarity measure for possibilistic uncertain information, Symbolic and Quantitative Approaches to Reasoning with Uncertainty, Springer, 2007, pp.840-852.

[26] A. Mencattini, G. Rabottino, M. Salmeri, R. Lojacono, E. Colini, Breast mass segmentation in mammographic images by an effective region growing algorithm, Lecture Notes in Computer Science 5259 (2008) 948-957.

[27] J. Fan, G. Zeng, M. Body, M.S. Hacid, Seeded region growing: An extensive and comparative study, Pattern Recognition Letters 26 (2005) 1139-1156.

[28] D. Dubois, L. Foulloy, G. Mauris, H. Prade, Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. Reliable Computing 10 (2004) 273-297.

[29] F. Wang, Fuzzy supervised classification of remote sensing images, IEEE Transactions on Geoscience

and Remote Sensing 28 (1990) 194-201.

[30] J.B. Adams, M.O. Smith, P.E. Johnson, Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 site, Journal of Geophysical Research: Solid Earth (1978-2012) 91 (1986) 8098-8112.

[31] N. Keshava, A survey of spectral unmixing algorithms, Lincoln Laboratory Journal 14 (2003) 55-78.

[32] W. Eziddin, Segmentation itérative d'images par propagation de connaissances dans le domaine possibiliste, Thèse, Télécom Bretagne, France, 2012.

[33] Z. Liu, J. Dezert, G. Mercier, Q. Pan, Dynamic evidential reasoning for change detection in remote sensing images, IEEE Transactions on Geoscience and Remote Sensing 50 (2012) 1955-1967.