Estimation of Mass Functions in Dempster-Shafer Evidence Theory using Fuzzy Clustering and Spatial Information for Gray Level Based Image Fusion

Layachi BENTABET[1,3], Sylvie JODOUIN[1], and Abdel-Ouahab BOUDRAA[2,3]

[1] *Départment de Mathématiques et d'Informatique, Sherbrooke University, Canada.*

[2] *Départment Signal, Ecole Navale, Lanvéoc Poulmic, 29240 Brest-Naval, France.*

[3] *CREATIS, CNRS UMR 5515, 69621 Villeurbanne, France.*

Abstract

In this paper a new multisensor information method for image segmentation is presented. Main characteristics of the method are the use of fuzzy logic theory to handle the ambiguity connected with pixel classification and the automatic estimation of mass functions using pixel membership degrees and spatial neighborhood information. Results on synthetic and real images are presented in order to demonstrate the effectiveness of the proposed method.

I. Introduction

The recent advent of multisensor imaging systems enables one to extract far more detailed information than ever before from the observed data. These systems provide a large amount of data that must be interpreted as a whole in order to make correct decisions. Thus, data fusion has been developed to utilize the full potential of the observed data. Data fusion is an important process in the areas of environmental systems, surveillance, automation, medical imaging and robotics. The main approaches to data fusion are statistical methods, Dempster-Shafer evidence theory, neural networks and fuzzy logic.

Data fusion is a technique allowing to simultaneously take into account heterogeneous data coming from different sources in order to get an optimal estimation of objects under investigation. Whatever the application, the goal of data fusion is to reduce uncertain and inaccurate information by combining both redundant and complementary data. A few mathematical formalisms (probability, fuzzy logic, possibilities, evidence theory, etc.) are available to perform a measure of the uncertainty and the inaccuracy [1]. Among these, evidence theory, also called Dempster-Shafer (DS) evidence theory [2], is a powerful and flexible mathematical tool for handling uncertain, imprecise and incomplete information. That is at least for three reasons. Firstly, by representing the uncertainty and the inaccuracy of a body of knowledge via the notion of evidence, belief can be committed to one single hypothesis (singleton) or a composite hypothesis (union of hypotheses). Secondly, the evidence combination rule of the DS theory provides an interesting operator to integrate multiple information from different sources. Finally, decision on the optimal hypothesis choice can be made in a flexible and rational manner.

In evidence theory, the determination of mass functions is a crucial step of fusion process. In image segmentation problems, mass functions are determined directly or indirectly from gray level histograms [3]-[5]. In spite of these works, the optimal determination of mass functions in practical applications remains a persistent problem.

This paper addresses the problem of mass functions estimation in DS evidence theory. The idea is to assign a mass function value to each pixel using its fuzzy membership degrees and contextual information. Each pixel is characterised by a membership grade, generated by the Fuzzy C-Means (FCM) clustering algorithm, to a given cluster or class or still hypothesis (throughout the paper, we will use these three terms in an indifferent manner). Simple and composite hypotheses (clusters) are derived from the calculated pixels membership degrees as well as the corresponding mass functions. Membership degrees are derived from gray level values and corrected taking into account pixels spatial neighborhood. Once the mass function values of each image to be fused are estimated, DS combination rule and decision are applied to obtain the final segmented image.

II. Dempster-Shafer evidence theory

DS theory [2],[6] makes inferences from incomplete and uncertain knowledge, provided by different independent knowledge sources. The theory allows strengthening or erosion of beliefs by combining additional sources of evidence, even in the presence of partly contradictory evidence. The DS theory contains the Bayesian theory of partial belief as a special case. In DS theory, there is a fixed set of N mutually exclusive and exhaustive elements, called the frame of discernment, which is symbolized by $\Theta = \{H_1, H_2, \ldots, H_N\}$. The representation scheme Θ defines the working space for the desired application since it consists of all propositions for which the information sources can provide evidence. Information sources can distribute mass values on subsets of the frame of discernment, $A_i \in 2^{\Theta}$. An information source assigns mass values only to those hypotheses, for which it has direct evidence. That is, if an information source can not distinguish between two propositions A_i and A_j, it assigns a mass value to the set including both propositions ($A_i \cup A_j$).

The derivation of the mass distribution is the most crucial step since it represents the knowledge about the actual application as well as the uncertainty incorporated in the selected information source. The mass distribution for all the hypotheses has to fulfill the following conditions:

$$0 \le m(A_i) \le 1$$
$$m(\phi) = 0 \qquad (1)$$
$$\sum_{A_i \in 2^\Theta} m(A_i) = 1$$

If $m(A_i) > 0$, A_i is called a focal element.

Mass distributions m_1, m_2 from two different information sources are combined with Dempster's orthogonal rule. The result is a new distribution, $m = m_1 \oplus m_2$, which carries the joint information provided by the two sources:

$$m(A_i) = (1-K)^{-1} \times \sum_{Ap \cap Aq = A_i} m_1(A_p) m_2(A_q) \qquad (2)$$

where

$$K = \sum_{Ap \cap Aq = \phi} m_1(Ap) m_2(Aq)$$

K is often interpreted as a measure of conflict between the different sources and it is introduced in Equation (2) as a normalisation factor. The larger K is, the more the sources are conflicting and the less sense their combination. From a mass distribution, two functions can be evaluated that characterize the uncertainty about the hypothesis A_i. The belief function *Bel* measures the minimum uncertainty value about A_i whereas plausibility *Pls* reflects the maximum uncertainty value about this hypothesis. These two measures span an uncertainty interval [$Bel(A_i)$, $Pls(A_i)$], which is called "belief interval". The length of this interval gives a measurement of imprecision about the uncertainty value. Belief and plausibility functions are defined from 2^Θ to [0, 1]:

$$Bel(A_i) = \sum_{A_j \subseteq A_i} m(A_j) \qquad (3)$$

$$Pls(A_i) = \sum_{A_j \cap A_i \neq \phi} m(A_j) \qquad (4)$$

These measures, which have been sometimes referred to as lower and upper probability functions, have the following properties:

$$Pls(A_i) = 1 - Bel(\overline{A_i}) \qquad (5)$$
$$Bel(A_i) \le Pls(A_i) \qquad (6)$$

where $\overline{A_i}$ is the complementary of A_i

III. Image segmentation

To fuse different images using DS theory, we first determine mass functions for all pixels of each image, and then we combine the values of these functions through the DS combination rule. Final segmented image is achieved using a decision strategy of maximum belief or plausibility.

In the proposed DS theory based segmentation method, appropriate determination of mass functions plays a crucial role since affectation of a pixel to a cluster is directly yielded from the estimated mass functions. Moreover, the mass functions should be generated by taking into consideration the pixels' properties (fuzziness for example, as explained below). That is rather different from many DS fusion applications reported in the literature, in which the determination of mass functions is not driven by image data. In all cases, mass function determination is itself subject to introducing subjective factors in fusion process. It would then be fundamental to reduce this subjectivity to a minimum.

In the present study, the method of generating mass functions is based on the concept of fuzzy logic. Mass functions are derived from fuzzy membership degrees. For each pixel is associated a membership function using FCM algorithm. Owing to this unsupervision character of the FCM, the determination of mass functions is also unsupervised and completely data-driven.

The interests of using fuzzy logic to deal with uncertainty attached to pixels are as follows. Generally, the boundaries of real data structures (clusters) are not well defined (fuzzy) and the transition from one cluster to another is not abrupt but a smooth one. This means that pixel information comes from several clusters. As pointed out in [7], image processing bears some fuzziness in nature. Therefore, there are intrinsic ambiguity and vagueness in clustering process. To cope with this situation, a fuzzy approach such as FCM is an interesting pixel clustering model owning to its capability to handle ambiguity and vagueness attached to pixel gray level value and which come from the imperfections of imaging systems. Indeed, compared to the conventional mathematics, in fuzzy logic there are no well-defined boundaries and the transition between full membership and no membership is gradual. Also, fuzzy segmentation relies on the fact that each pattern is characterized by a degree of membership to a given cluster.

This enables us to distinguish between patterns forming a main structure of the cluster and those being outliners [8].

However, membership functions are generated solely from the images' gray levels. In order to improve the robustness of the segmentation, spatial information should be taken into account in the segmentation paradigm. To achieve that, the final mass function assigned to a pixel is yielded by combining membership functions of its neighbouring pixels. Once the spatial information based mass functions are determined for each of the images to be fused, the DS combination rule and decision are applied to obtain the final segmented image.

III. 1 Fuzzy Clustering

Let $X=\{x_1,\dots, x_M\}$ be a finite unlabeled data set and C be the number of clusters; $2\leq C\leq M$ and let $\Re^{C\times M}$ the set of all real $C\times M$ matrices. A fuzzy C-Partition of X is represented by a matrix $U=[\mu_{ik}]\in\Re^{C\times M}$, where $\mu_{ik}=\mu_i(x_k)$ represents the degree of membership of x_k to cluster i, and verifies the following constraints:

$$\mu_{ik}\in[0,1]\quad 1\leq i\leq C;\quad 1\leq k\leq M$$
$$\sum_{i=1}^{C}\mu_{ik}=1\quad 1\leq k\leq M \tag{7}$$
$$\sum_{k=1}^{M}\mu_{ik}>0\quad 1\leq i\leq C$$

U is used to describe the clusters of X, and a good partition U of X is yielded by the minimization of the FCM objective functional [9]:

$$J_q(U,V:X)=\sum_{k=1}^{M}\sum_{i=1}^{C}(\mu_{ik})^q\|x_k-v_i\|_A^2 \tag{8}$$

where $q\in[1,+\infty[$ is a weighting exponent called the fuzzifier, and $V=(v_1, v_2,\dots, v_C)$ the vector of the cluster centers. $\|x\|_A=\sqrt{x^TAx}$ is any inner product norm where A is any positive definite matrix. Approximate optimization of J_q by the FCM algorithm is based on iteration through the following necessary conditions for its local extrema:

FCM Theorem ([9]): Assume $q\geq 1$ and $\|x_k-v_i\|_A^2>0,\ 1\leq i\leq C,\ 1\leq k\leq M$. (U, V) may minimize J_q only if:

$$\mu_{ik}=\left[\sum_{j=1}^{C}\left(\frac{\|x_k-v_i\|_A}{\|x_k-v_j\|_A}\right)^{\frac{2}{(q-1)}}\right]^{-1} \tag{9}$$

$$v_i=\frac{\sum_{k=1}^{M}(\mu_{ik})^q\cdot x_k}{\sum_{k=1}^{M}(\mu_{ik})^q} \tag{10}$$

The FCM algorithm consists of iterations alternating between Equations (9) and (10). This algorithm converges to either a local minimum or a saddle point of J_q.

In the context of gray level image segmentation, a fast version of the FCM can be used [10]. The proposed algorithm is based on one-dimensional attribute such as the gray level. Let H be the histogram of image of L-levels, where L is the number of gray levels. Each pixel, l, has a feature that lies in the discrete set:

$$X=\{0, 1,\dots, L\text{-}1\} \tag{11}$$

In this case relation (8) can be written as follows:

$$J_q(U,V:L)=\sum_{l=0}^{L-1}\sum_{i=1}^{C}(\mu_{il})^q\cdot H(l)\cdot(l-v_i)^2 \tag{12}$$

where $\mu_{il}=\mu_i(l)$ represents the membership degree of gray level l to cluster i.

Thus, the computation of membership degrees of $H(l)$ pixels is reduced to that of pixel with l as gray level value. The algorithm is outlined in the following steps:

Step1) Select the number of clusters C, $2\leq C\leq L$, the termination criteria $\varepsilon>0$, the value of $q>1$ and initialize the partition matrix U according to:

$$\sum_{i=1}^{C}\mu_{il}=1\quad l=0,1,\dots,L-1$$

Step2) Compute the prototypes of the clusters and update the partition U as follows :

$$v_i=\frac{\sum_{l=0}^{L-1}(\mu_{il})^q\cdot H(l)\cdot l}{\sum_{l=0}^{L-1}(\mu_{il})^q}\quad i=1,\dots,C \tag{13}$$

$$\tilde{\mu}_{il}=\left[\sum_{j=1}^{C}\left(\frac{(l-v_i)}{(l-vj)}\right)^{\frac{2}{(q-1)}}\right]^{-1} \tag{14}$$

$$\text{Step3)}\ E = \sum_{i=1}^{C}\sum_{l=0}^{L-1} | \tilde{\mu}_{il} - \mu_{il} |$$

If $(E < \varepsilon)$ Stop. Else goto Step2)

III.2 Integration of Spatial Information

Like the popular k-means and ISODATA clustering algorithms, FCM does not exploit spatial information. These methods work well for uncontaminated data and may show ambiguous results under noisy environment because of the random selection of the initial membership degrees. Consequently, estimation of mass functions derived from membership degrees might be hampered, yielding segmentation errors. Thus, for robust determination of mass functions membership degree of each pixel is re-estimated using its spatial neighborhood.

Let V_{mn} be the neighborhood of the pixel located at (m,n) in the image and $l(m,n)$ be its gray level value. The membership grade, designated by $\mu_i^{m,n}=\mu_i(l(m,n))$, of the central pixel $l(m,n)$ is obtained by combining the membership degrees of all the pixels belonging to V_{mn}. Note that $\mu_i^{m,n}$ or $\mu_i(l(m,n))$ represents the membership degree of the pixel $l(m,n)$ to cluster i. We have $\mu_i^{m,n}=\mu_i(l(m,n))=\mu_{il}$, if the explicit indication of the position (m,n) is omitted. Two methods have been investigated to combine membership degrees:

a) ***Mean Value Method***:

$$\overline{\mu}_i^{m,n} = \frac{\sum\limits_{m',n' \in V_{mn}} \left(\mu_i^{m',n'}\right)}{|V_{mn}|}, \qquad 1 \le i \le C, \qquad (15)$$

where $|V_{mn}|$ is the cardinal of V_{mn}.

b) ***Median Value Method***:

The membership grade is calculated as the median value of the membership degrees of neighboring pixels. Thus for simple hypotheses we have

$$\overline{\mu}_i^{m,n} = \underset{m',n' \in V_{mn}}{Median} \left(\mu_i^{m',n'}\right), \qquad 1 \le i \le C, \qquad (16)$$

Final mass functions associated to the pixel being considered are then derived from the above spatially filtered membership functions, as discussed in the following section.

IV. Mass Functions Estimation

In this paper our study is limited to two sensors. Let I_1 and I_2 be the two images to be fused. We make assumption that the two images are well mapped. Recall that the mass functions are calculated using the spatially filtered membership degrees.

The first step in computing mass functions using FCM is to represent each gray level l of the images with a function of membership μ_{il} to class i. To do that, FCM is applied to I_1 and I_2 . Then, the pieces of evidence in DS theory are represented by the membership functions (or averaged membership functions).

The hypotheses being considered in DS formulation are the following: ϕ (whose mass is null), simple hypotheses H_i and composite hypotheses $H_r \cup \ldots \cup H_s$. In the present study, we consider only double hypotheses $H_r \cup H_s$ for composite hypotheses. Indeed, when performing clustering using solely gray level information, the intersection occurs only between two curves $\mu_r(l)$ and $\mu_s(l)$ corresponding to two consecutive centroids v_s and v_r. The intersection point corresponds to the highest degree of ambiguity between hypotheses H_r and H_s. Another argument can also be quoted. It resides in the fact that, gray levels presenting a high level of ambiguity in their memberships correspond to pixels located in the edges between two clusters, such as r and s, on the image. Thus, for these pixels the ambiguity is dealt with two neighbored clusters.

A double hypothesis is created or not according to the following strategy:

- Assign a non-null mass to $H_r \cup H_s$ if H_r and H_s are not discriminated in the images (not distinguishable by sensors). That corresponds to case there is an ambiguity between H_r and H_s.

- Assign a null mass to $H_r \cup H_s$ if H_r and H_s are discriminated in the image. That reflects the situation where there is no or little membership of l to both clusters r and s at the same time.

Therefore, a non-normalized mass $\tilde{m}$ of hypothesis A_i is evaluated depending on the simple or double type hypothesis. The final mass m is obtained after normalisation with respect to the entire set 2^{Θ} , that is:

$$m(A_i) = \frac{\tilde{m}(A_i)}{\sum\limits_{A_j \in 2^{\Theta}} \tilde{m}(A_j)} \qquad (17)$$

such that:

$$\sum_{A_j \in 2^{\Theta}} m(A_j) = 1 \qquad (18)$$

The non-normalized mass functions $\tilde{m}$ is computed from the filtered membership functions $\overline{\mu}_i^{m,n}$ following two cases:

a) ***Simple hypotheses:***

Masses of simple hypotheses H_i are directly obtained from the filtered membership functions $\overline{\mu}_i^{m,n}$ of the gray level l(m,n) to cluster i as follows:

$$\tilde{m}(A_i) = \tilde{m}(H_i) = \overline{\mu}_i^{m,n}, \quad 1 \le i \le N \qquad (19)$$

b) ***Double hypotheses:***

If there is a high ambiguity in assigning a gray level l(m,n) to cluster r or s, that is $|\overline{\mu}_r(l) - \overline{\mu}_s(l)| < \xi$ where ξ is a thresholding value, then a double hypothesis is formed and its associated mass is calculated according to the following formula:

$$\tilde{m}(A_i) = \tilde{m}(H_r \cup H_s) = \Im(\overline{\mu}_r(l), \overline{\mu}_s(l)) \qquad (20)$$

$\Im$ denotes a mass function operator. It is evident that the more $\overline{\mu}_r(l)$ or $\overline{\mu}_s(l)$ is important, the more the mass to their union. At the same time, the closer the values of $\overline{\mu}_r(l)$ and $\overline{\mu}_s(l)$ are, the greater is the ambiguity between the two hypotheses forming their union, and consequently, the greater is the mass to assign to the double hypothesis $H_r \cup H_s$. Therefore, to determine the function $\Im$, two constraints should be taken into account: a) For a given gray level l, the mass assigned to the double hypothesis $H_r \cup H_s$ should be proportional to the membership degrees of l(m,n) to each of hypotheses H_r and H_s; b) The mass assigned to $H_r \cup H_s$ should be inversely proportional to the (absolute) difference between the membership degrees of $l(m,n)$ to hypotheses H_r and H_s. That has led us to formulate the problem by considering $\Im$ as a new variable, which represents the surface of a triangle that we will construct below. The surface of such a triangle will depend both on the amplitudes of the membership functions of l to clusters r and s, and on the difference between these amplitudes.

Figure 1 shows how the triangle is constructed and how the mass of double hypotheses $H_r \cup H_s$ is derived from the surface of the triangle. The vertical axis of Figure 1 represents the membership degrees. The unit of the horizontal one has no importance on the mass computation. In practical implementations, the horizontal axis was considered as a normalized one, which therefore has no unit. The two dotted triangles represent two so-called triangular membership functions corresponding to classes r and s, between which ambiguity is to be studied. The two triangles are isosceles, and have the same length for their bases, which is equal to the two times the distance between the vertexes of the two triangles. The heights of the triangles are equal to $\overline{\mu}_r(l)$ and $\overline{\mu}_s(l)$, respectively. The overlapping surface, as represented by the shaded region S in Figure 1, of the two triangles then represents the membership of a gray level to the composite hypothesis $H_r \cup H_s$ in the new reference system. It can be regarded as the level of ambiguity associated to the classification of a pixel to one or the other cluster. Therefore, the mass value attributed to the double hypothesis $H_r \cup H_s$ can be directly calculated from the surface S.

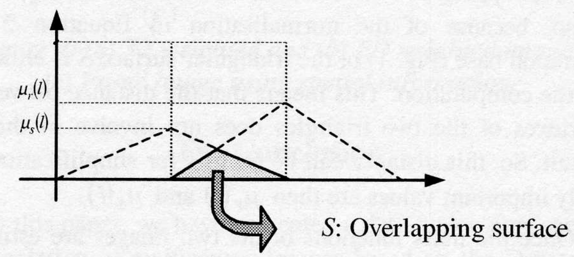

Figure 1: Construction of triangular membership functions and determination of the overlapped surface.

To further demonstrate the behavior of the entity S for the quantification of ambiguity, let us assume that $\Theta = \{H_r, H_s\} \Rightarrow \overline{\mu}_r(l) + \overline{\mu}_s(l) = 1$. From the membership function curves of H_r and H_s, it can be seen that the highest ambiguity between the hypotheses r and s is obtained when $\overline{\mu}_r(l) = \overline{\mu}_s(l) = 0{,}5$. Also, from Figure 1, it is clear that for $\overline{\mu}_r(l) + \overline{\mu}_s(l) = 1$, S is maximum when $\overline{\mu}_r(l) = \overline{\mu}_s(l) = 0{,}5$. Therefore, we conclude that the maximum of overlapped surface S corresponds to the highest ambiguity case.

The above conclusion can be extended to more general cases $\overline{\mu}_r(l) + \overline{\mu}_s(l) = \alpha = cte$. Intuitively, ambiguity becomes higher as $|\overline{\mu}_r(l) - \overline{\mu}_s(l)|$ decreases. Meanwhile, for $\overline{\mu}_r(l) + \overline{\mu}_s(l) = \alpha = cte$, the surface S increases when $|\overline{\mu}_r(l) - \overline{\mu}_s(l)|$ decreases, and if α increases, S increases too. The two extreme cases are: a) $\overline{\mu}_r(l) = \overline{\mu}_s(l) = 0{,}5$, which yields a maximum ambiguity, and b) $\overline{\mu}_r(l) = 1$, $\overline{\mu}_s(l) = 0$, which corresponds to the case of no ambiguity between the two classes. Thus, S depends on ($\overline{\mu}_r(l) + \overline{\mu}_s(l)$) and ($\overline{\mu}_r(l)$-$\overline{\mu}_s(l)$) in such a way to correctly model the existing ambiguity between r and s. Another condition that S must satisfy is that its value has to be confined to the interval [0,1].

To do that, S is normalized with the maximum value S_{max} of S. From the membership functions of hypotheses H_r and H_s, the mass value assigned to the composite hypothesis $H_r \cup H_s$ is then given by:

$$\tilde{m}(H_s \cup H_r) = \Im(\overline{\mu}_r(l), \overline{\mu}_s(l)) = \frac{S(\overline{\mu}_r(l), \overline{\mu}_s(l))}{2.S_{max}} \qquad (21)$$

The choice of triangular shapes for the membership functions is made in order to reduce the computation complexity. On the other hand, in fuzzy clustering, each class is characterised by a center (a gray level in the present study) with μ =1. Therefore, when moving away from this center, the membership degree to this class decreases, and at the same time the membership degree to the neighbouring class increases until arriving at its core which has a μ = 1. Hence, an overlapping of 50 % between the two classes is produced. Also, because of the normalisation in Equation 21, the common base (Fig. 1) of the triangular surface S is eliminated in the computation. This means that the distance between the vertexes of the two triangles does not involve in the final result. So, this distance can be set to 1 for simplification. The only important values are then $\mu_r(l)$ and $\mu_s(l)$.

Once the mass functions of the two images are estimated, their combination is performed using the orthogonal sum (Eq. 2). The decisional procedure for classification purpose consists in choosing one of maximum plausibility, maximum belief or pignistic probability as the most likely hypothesis H_i. Decision-making is carried out on simple hypotheses which represent the classes in the images. If we accept the composite hypotheses as final results in the decisional procedure, the obtained segmentation results would be more reliable but with a decreased precision.

V. Results

The proposed data fusion method is first tested on synthetic images. Two images, corrupted by Gaussian noise, simulating weak and strong X-rays acquisitions are shown in Figure 2. Each image contains four regions (C=4). In the first (left) image, one region (smallest thickness) is confused with the background and in the second one (right) the greatest thickness is under-exposed and the thicker regions are not well distinguished. The aim here is to exploit, through using the proposed data fusion technique, the redundant and complementary information of the two images in order to correctly segment the image in four regions.

Figure 3a shows the segmentation result obtained without use of the spatial neighborhood information. It is observed that the four regions are well brought out, showing that the complementary information provided by two images was well exploited by the fusion algorithm. This demonstrates that the calculated mass functions gave a good modeling of the available information associated to the different hypotheses. However, it is also seen that within the segmented regions, many artifacts are present, reflecting the influence of noise present in the initial images on final segmentation. After incorporation of spatial neighborhood information in the calculation of mass functions, we obtain the segmentation results given in Fig. 3b, in which the image corresponds to the case when averaging is used, and in Fig. 3c where a median operation is used. In both cases, improvement in segmentation is significant. In particular, the averaging method has led a very good homogeneity in the segmented regions. This demonstrates the interest of introducing spatial neighborhood information at the level of mass function calculation for the DS fusion based segmentation. Our method has been compared with a classical segmentation method where each image is segmented separately using FCM algorithm. Fuzzy segmented images, shown in Figures 3d and 3e, demonstrate that none of the two images provides complete and reliable information compared to the DS fusion approach (Figures 3a-c). We have noted that 23,94 % and 34,94 % of pixels have been mis-segmented for the strong and weak X-rays images, respectively. Segmentation errors have been largely reduced when exploiting simultaneously the two images through the use of DS fusion approach including spatial information. Indeed only 0,32 % of pixels have been mis-segmented. This big performance difference between these two types of segmentation approaches can also be easily assessed by visually comparing the segmentation results.

Figure 4 shows another example of segmentation using the DS based fusion and spatial neighborhood information. This is a typical example in which raw data are by essence multispectral, and data fusion appears as a natural technique for exploiting the redundant and complementary information contained in these two images corresponding to the same object. The raw image pair shown in Figures. 4a and 4b represents a single dual echo magnetic resonance sequence. The image in Figure 4a is a T2-weighted image and that in Figure 4b a proton density (PD) weighted image. The two images correspond to the same slice of a human brain. The fused image is shown in Figure 4c. The segmentation result shows the presence of four classes: gray matter, white matter, multiple sclerosis (MS) lesions and cerebrospinal fluid (CSF), background. The size and shape of brain tissues are well segmented. It is also observed that the segmented regions are rather homogeneous which makes it possible to do an accurate measurement of brain tissues volumes.

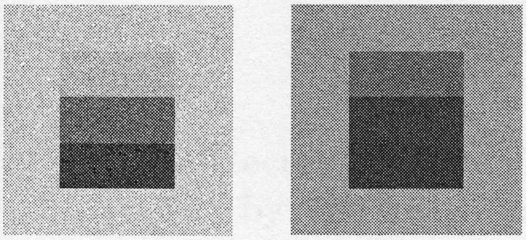

Figure 2: Two images simulating Strong (left) and Weak (right) X-rays acquisitions respectively.

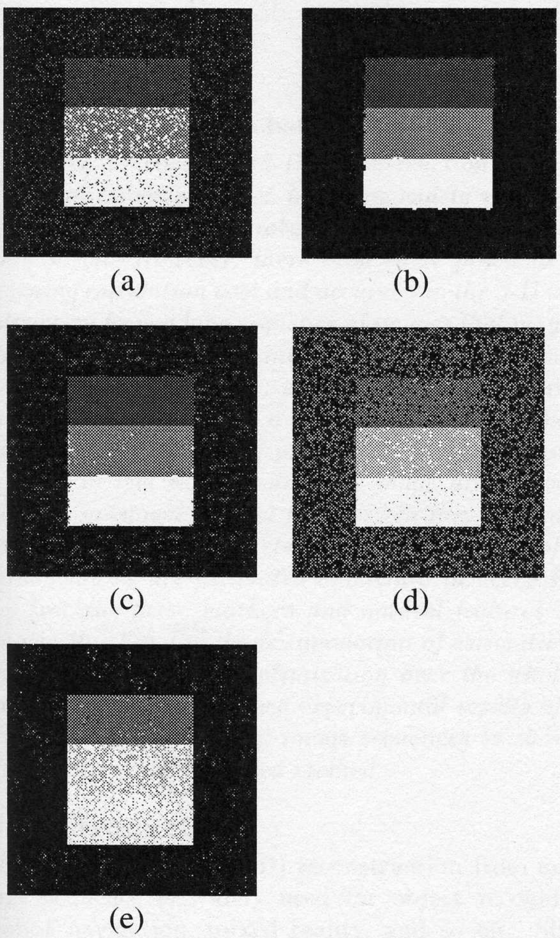

Figure 3: (a) Fused image without spatial information. (b) Fused image using membership degrees averaging. (c) Fused image using median value of the membership degrees. Fuzzy segmentation of Strong (d) and Weak (e) X-rays images

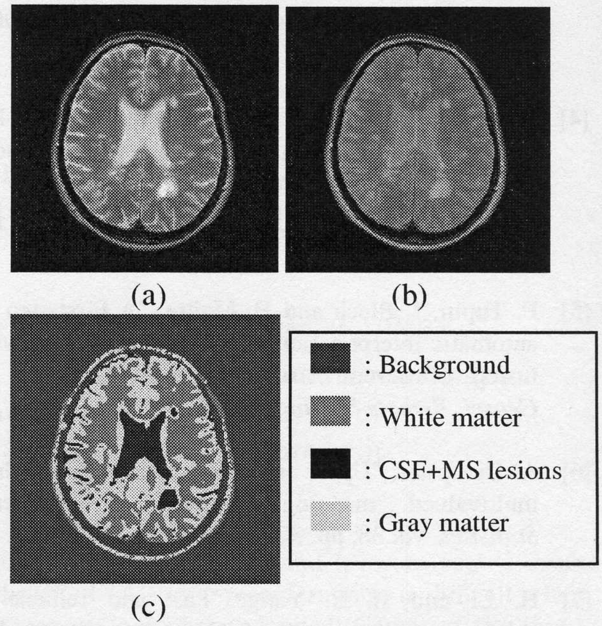

Figure 4: (a) T2 weighted and (b) PD weighted images. (c) Fused image using spatial information.

IV. Conclusion

In this paper, we have presented a data fusion approach to segmentation of multisource images based on the Dempster-Shafer evidence theory. Within this framework, we have investigated in particular the possibility of automatically determining mass functions using the concept of fuzzy logic. To each gray level, we have associated a mass function value so that uncertainty and imprecision are handled at the pixel level. On the other hand, in the calculation of mass functions, we have also introduced spatial neighborhood information in order to take into account the spatial correlation between neighbouring pixels that exists in any physical images. The obtained results have shown the generic and robust character of the approach in the sense that only gray level and spatial neighborhood properties of the images were involved. The comparison study has also illustrated that performance in segmentation was significantly improved when using several images representing the same object.

References

[1] I. Bloch and H. Maître, "Fusion de données en traitement d'images: modèles d'information et de décision," *Traitement du Signal*, vol. 11, no. 6, pp. 435-446, 1994.

[2] G. Shafer, A mathematical theory of evidence, Princeton University, 1976.

[3] I. Bloch, "Some aspects of Dempster-Shafer evidence theory for classification of multi-modality medical imaging taking partial volume effect into account,"

Pattern Recognition Letters, vol. 17, no. 8, pp 905-919, 1996

[4] S. Le Hégarat-Mascle, I. Bloch and D. Vidal-Madjar, "Application of Dempster-Shafer evidence theory to unsupervised classification in multi-source remote sensing," *IEEE Trans. Geosci. Remote Sensing,* vol. 35, pp. 1018-1032, 1997.

[5] F. Tupin, I. Bloch and H. Maître, "A First step toward automatic interpretation of SAR images using evidential fusion of several structure detectors," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 1327-1334, 1999.

[6] A. Dempster, "Upper and lower probabilities induced by multivalued mapping," *Annals of Mathematical Statistics*, vol.38, pp. 325-339, 1967.

[7] H. Li and H. S. Yang, "Fast and reliable image enhancement using fuzzy relaxation technique," *IEEE Trans. Inform. Theory,* vol. 19, pp. 1276-1281, 1989.

[8] W. Pedrycz, "Fuzzy sets in pattern recognition: methodology and methods," *Pattern Recognition*, vol. 23, pp. 121-146, 1990.

[9] J. C. Bezdek, "A convergence theorem for fuzzy isodata clustering algorithm," *IEEE Trans. Pattern Analysis and Machine intelligen*ce, vol. 2, pp. 130-142, 1980.

[10] A. Boudraa and P. Clarysse, "Fast fuzzy grey level image segmentation method," *Medical Biological Engineering Computing*, vol. 35, Suppl. Part. 2, pp. 804, 1997.