Iterative estimation of Dempster-Shafer's basic probability assignment: application to multisensor image segmentation*

Fabien Salzenstein
Université de Strasbourg
Laboratoire Phase
23 rue du Loess
Strasbourg, France

Abdel-Ouahab Boudraa*
IRENav, Ecole Navale
29200 Brest-Armées, France
E-mail: boudra@ecole-navale.fr

Abstract. Basic probability assignment (BPA) definition remains a difficult problem to apply Desmpter-Shafer evidence theory to practical applications such as in image processing. A new iterative approach of multisensor data fusion based on the Dempster-Shafer framework is proposed. BPAs, modeled by a Gaussian distribution, are estimated iteratively and in an unsupervised way using the fused image and the source images. Data fusion is performed at the pixel level. Results on synthetic and real images are presented to illustrate the effectiveness of the proposed fusion scheme. Limitations of the method are discussed.
[DOI: 10.1117/1.1737373]

Subject terms: data fusion; Dempster-Shafer evidence theory; basic probability assignment; image segmentation.

Paper 030124 received Mar. 18, 2003; revised manuscript received Sep. 2, 2003 and Jan. 2, 2004; accepted for publication Nov. 3, 2003.

1 Introduction

Data fusion deals with the combination of information from different sensors aiming at an improved accuracy of the classification and for better decision. The information provided by one sensor is usually limited and sometimes of low accuracy. The use of multiple sensors is an alternative to improve accuracy and provide the user with additional information of increased reliability about the environment in which the sensors operates. With the development of new imaging sensors in many fields such as remote sensing, robotics, nondestructive evaluation, medical imaging, machine vision, target tracking, and airborne surveillance, image fusion has emerged as a promising new research area. Data fusion can take place at different stages: signal, pixel, feature, and symbolic levels.[1] In this paper, we focus on the pixel-level fusion process. To perform pixel-level fusion successfully, all input images must be exactly spatially registered, i.e., the pixel positions of all input images must correspond the same location in real world. Different strategies have been developed for data fusion. The frameworks used for data management are Bayesian inference, Dempster-Shafer (DS) theory[2,3] and fuzzy logic inference.[4] DS theory makes inferences from incomplete and uncertain knowledge, provided by different independent knowledge sources. A first advantage of DS theory is its ability to deal with ignorance and missing information. In particular, it provides explicit estimation of imprecision and conflict between information from different sources and can deal with any unions of hypotheses[5] (clusters). This is particularly useful to represent ''mixed'' pixels in classification problems. The main limitation of Bayesian inference is that it cannot model imprecision about uncertainty measurement. The degree of belief we have on a union of clusters (without being able to discriminate between them) should be shared by all the simples hypotheses, thus penalizing the good one. DS theory handles uncertain and incomplete information through the definition of two dual nonadditive measures: plausibility and belief. These measures are derived from a density function, m, called Basic probability assignment (BPA) or mass function. This probability assigns evidence to a proposition (hypothesis). The derivation of the BPA is the most crucial step since it represents the knowledge about the application as well as the uncertainty incorporate in the selected information source. BPA definition remains a difficult problem to apply DS theory to practical applications such in image processing. For example, BPA can be derived, at pixel level, from probabilities,[6–8] from the distance to cluster centers,[9] or from fuzzy membership functions.[10] In this paper, BPAs modeled by a normal distribution are estimated in unsupervised way and iteratively using the output fused image and the sensor images. The number of the clusters of the image is supposed known.

2 DS Theory Basics

DS theory is developed as an attempt to generalize probability theory. This theory is suitable to reason with uncertainty and has been developed to overcome the limitations of conventional probability theory by distinguishing between uncertainty and ignorance. In the case of probability theory, uncertainty about an event is measured by a single value (probability) and imprecision about uncertainty measurement is assumed to be null. DS theory is suited for the combination of information from different sensors. In the

*A part of this work has been presented at the Sixth International Symposium on Signal Processing and Its Applications (ISSPA), Kuala Lumpur, Malaysia, 2001.

DS theory, there is a fixed set of $(q+1)$ mutually exclusive and exhaustive elements, called the frame of discernment, which is symbolized by

$$\Theta=\{H_0,H_1,\ldots,H_q\}.$$

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, $A_i \in 2^\Theta$ [Eq. (1)]. In this paper, Θ is the set of all the pixel classes and 2^Θ contains all the possible unions of pixel classes. Thus, not only single classes (singletons) but also any union of classes can be represented. An information sources assigns mass values only to those hypotheses, for which it has direct evidence; i.e., if an information source can not distinguish between two propositions, it assigns a mass value to the set including both propositions $(H_i \cup H_j)$. 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.

$$0 \leqslant m(A_i) \leqslant 1, \tag{1}$$

where $m(A_i)$ represents the degree of belief that is exactly committed to the set A_i (and not to any proper subset of A_i). For example, if $m(A_i)=0.45$, then a 45% portion of one's total belief is assigned to exactly A_i. When the mass affected to a compound hypothesis $H_i \cup H_j$ is nonzero, it means that we have an option not to make the decision between H_i or H_j but rather leave the pixel in the $H_i \cup H_j$ class. In particular, assigning a nonnull mass to Θ enables us to not classify some pixels, for which there is a global ignorance. The mass distribution has to fulfill the following conditions:

$$m(\varnothing)=0, \tag{2}$$

$$\sum_{A_i \in 2^\Theta} m(A_i)=1. \tag{3}$$

Condition (2) reflects the fact no belief ought to be committed to $\varnothing$ and condition (3) reflects the convention that one's total belief has measure one.[3] Mass distribution from different information sources, $m_j(j=1,\ldots,d)$, are combined with Dempster's orthogonal rule, Eq. (4). The result is a new distribution, $m(A_k)=(m_1 \oplus m_2 \oplus \cdots \oplus m_d)(A_k)$, which incorporates the joint information provided by the sources.

$$m(A_k)=(1-K)^{-1} \times \sum_{A_1 \cap A_2 \ldots A_d = A_k} \left[\prod_{1 \leqslant j \leqslant d} m_j(A_j) \right], \tag{4}$$

$$K=\sum_{A_1 \cap A_2 \ldots A_d = \varnothing} \left[\prod_{1 \leqslant j \leqslant d} m_j(A_j) \right] \tag{5}$$

where $A_k \neq \varnothing$. Often K is interpreted as a measure of conflict between the different sources [Eq. (5)] and is introduced as a normalization factor. The larger is K the more the sources are conflicting and the less sense has their combination. The factor K indicates the amount of evidential conflict. If $k=0$, this shows complete compatibility, and if $0<K<1$, it shows partial compatibility. Finally, the orthogonal sum does not exist where $K=1$. In this case, the sources are totally contradictory, and it is no longer possible to combine them. In cases of sources that are highly conflicting, the normalization used in the Dempster combination rule can lead to mistakes, since it artificially increases the masses of the compromise hypotheses.[11] One may suggest, as in Ref. 11, that the conflict comes from the fact that the ''true'' assumption has been forgotten (in the set of hypotheses). From a mass distribution, numerical values can be calculated that characterize the uncertainty and the support of certain hypotheses. Belief [Eq. (6)] measures the minimum or necessary support, whereas plausibility [Eq. (7)] reflects the maximum or potential support for that hypothesis. These two measures span an uncertainty interval $[\mathrm{Bel}(A_i),\mathrm{Pls}(A_i)]$ for this hypothesis. This interval is called ''belief interval'' and the length of this belief interval gives a measurement of the imprecision about the uncertainty value. These two measures, derived from mass values, are respectively defined from 2^Θ to [0,1]:

$$\mathrm{Bel}(A_i)=\sum_{A_j \subseteq A_i} m(A_j), \tag{6}$$

$$\mathrm{Pls}(A_i)=\sum_{A_j \cap A_i \neq \varnothing} m(A_j), \tag{7}$$

where $\mathrm{Bel}(A_i)$ is the total belief committed to A_i, that is, the belief that the truth is in A_i. $\mathrm{Pls}(A_i)$ can be seen as the amount of evidence that does not refute A_i. $\mathrm{Bel}(A_i)$ and $\mathrm{Pls}(A_i)$ are sometimes referred to as lower and upper probability functions. Relations (6) and (7) imply that $\mathrm{Bel}(\cdot)$ and $\mathrm{Pls}(\cdot)$ are dual measures related by

$$\mathrm{Pls}(A_i)=1-\mathrm{Bel}(\neg A_i). \tag{8}$$

According to Eq. (8), $\mathrm{Pls}(A_i)$ is the plausibility value that is the ''unbelief'' of the complementary hypothesis $\neg A_i$ and can be interpreted as the maximum uncertainty value of A_i $[\mathrm{Bel}(A_i)<\mathrm{Pls}(A_i)]$.

3 Basic Probability Assignment

We assume that we have a set of images of the same scene that have to be fused in order to classify the scene into C classes of interest. Let d be the set of images to be fused and defined as random vectors $\mathbf{Y}^j=[y_s^j]$, $j=1,2,\ldots,d$; $s=1,2,\ldots,N$. Note that y_s^j is gray-level value at site s of jth image, and N is the number of sites of each image. For image segmentation Θ denotes the set of hypotheses about pixel class. Let Ω be the set of C classes: $\Omega=\{\omega_1,\omega_2,\ldots,\omega_C\}$ and let Ω^* be the power set of Ω: $\Omega^*=\{A_1,A_2,\ldots,A_{2^C}\}$. The result of the fusion process is a segmented image $X=[x_s] \in \Omega^N$ into C nonoverlapping classes. An advantage to apply DS theory to classification problems is that not only single classes (simple hypotheses)

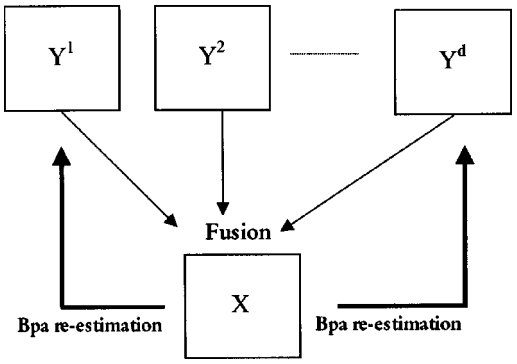

Fig. 1 Proposed fusion scheme.

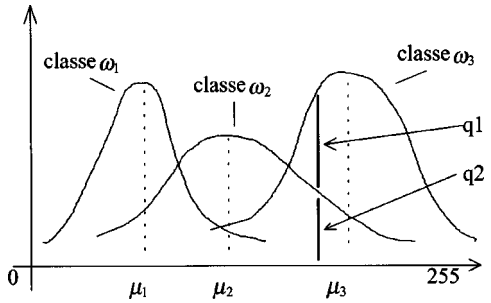

Fig. 3 Gray-level probability density functions of three pixel classes; the estimation of BPAs in the first approach for a three-classes-output fused image.

but also union of classes (compound hypotheses) can be represented. If there is no ambiguity between classes (i.e., nonignorance about classes), a null mass is affected to their union. This means that the different classes are well discriminated in the image and each pixel $\mathbf{y}_s^j$ is affected to only one class. Conversely, when two or more classes are not distinguishable by a given sensor, a nonnull mass is affected to their union. We assume that the distribution of feature vector $\mathbf{y}_s^j$ conditionally to a class ω_i is a Gaussian distribution:

$$P[\mathbf{y}_s^j | x_s \in \omega_i] = \frac{1}{\sigma_i^j \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\mathbf{y}_s^j - \mu_i^j}{\sigma_i^j} \right)^2 \right], \tag{9}$$

where μ_i^j and σ_i^j represent the mean and the standard deviation of the i'th class corresponding to the j'th sensor respectively and $\mathbf{y}_s^j \in \Omega^*$. BPAs are estimated for each pixel, $\mathbf{y}_s^j$, and for the d images over all hypothesis according the scheme shown in Fig. 1. Let us now define the singleton or simple hypothesis H_p, which corresponds to class $\omega_p \in \Omega$. This hypothesis is constructed as an intersection of the compound hypothesis of sensor images, for the same area. Thus an ideal classes output image is defined as follows. Let us consider two sensor images Y^1 and Y^2, as shown in Fig. 2. Each pixel of the sensor image belongs respectively to the following sets: $\mathbf{y}_s^1 \in \{A^1, B^1 \cup C^1\}$ and $\mathbf{y}_s^2 \in \{A^2 \cup B^2, C^2\}$. In the resulting fused image X, the pixels belong to the set $\Omega = \{A, B, C\}$, where $A = A^1 \cap (A^2 \cup B^2)$, $B = (B^1 \cup C^1) \cap (A^2 \cup B^2)$, and $C = (B^1 \cup C^1) \cap C^2$. In this paper, the BPAs are directly estimated from the Gaussian distribution of gray-level values of the pixels conditionally to the corresponding class.

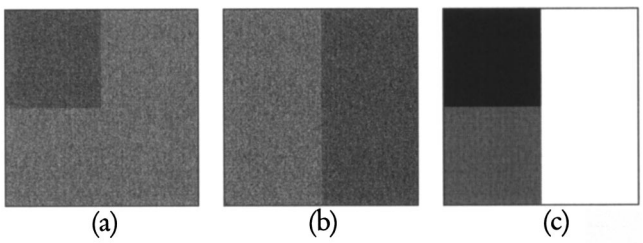

Fig. 2 Output of two sensors: (a) image of the first sensor, (b) image of the second sensor, and (c) "ground-truth" image.

3.1 *First Approach*

For the choice of BPAs of two classes H_p and $H_{p'}$, the following strategy is used[7]

$$m(y_s^j, H_p) = (P[y_s^j | x_s \in \omega_p] - P[y_s^j | x_s \in \omega_{p'}])/M, \tag{10}$$

$$m(y_s^j, H_p \cup H_{p'}) = P[y_s^j | x_s \in \omega_{p'}]/M, \tag{11}$$

$$M = P[y_s^j | x_s \in \omega_p], \tag{12}$$

$$p = \arg\{ \max_{1 \le i \le C} P[y_s^j | x_s \in \omega_i] \}, \tag{13}$$

$$p' = \arg\{ \max_{i \ne p} P[y_s^j | x_s \in \omega_i] \}, \tag{14}$$

where M is a normalization factor; $m(y_s^j, H_p)$ represents the degree of belief, at the pixel located at site s and provided by the j'th source (image), that is exactly committed to the hypothesis H_p. It is easy to see, according to condition (3), that $m(y_s^j, H_p) + m(y_s^j, H_p \cup H_{p'})$ is equal to 1. Also $m(y_s^j, H_p \cup H_{p'})$ measures the overlapping between the H_p and $H_{p'}$ classes. The more important this mass value, the more difficult it is to make a decision between H_p or $H_{p'}$. The method is illustrated in Fig. 3 ($p = 3, p' = 2$), where the quantities q_1 and q_2 represent $m(y_s^j, H_3)$ and $m(y_s^j, H_2 \cup H_3)$, respectively, for a three-classes-output fused image. Figure 3 represents a gray-level probability density function of three pixel classes. In this approach, only two adjacent classes are taken into account to estimate BPA.

3.2 *Second Approach*

To improve the first approach a conjunction of more than two hypothesis is taken into account. Let $(\mu_{k_1}^j, \mu_{k_2}^j, \ldots, \mu_{k_L}^j)$ and $(\sigma_{k_1}^j, \sigma_{k_2}^j, \ldots, \sigma_{k_L}^j)$ denote, respectively, the means and standard deviations of the L hypotheses $H_{k_1}, H_{k_2}, \ldots, H_{k_L}$, according to gray-level distribution. The mean and standard deviation of the distribution repre-

senting the conjunction of L hypotheses are denoted, respectively, by $\mu_L^j=(\mu_{k_1}^j+\mu_{k_2}^j+\ldots+\mu_{k_L}^j)/L$ and $\sigma_L^j=\max(\sigma_{k_1}^j,\sigma_{k_2}^j,\ldots,\sigma_{k_L}^j)$:

$$m\left(y_s^j,\bigcup_{i=1}^{L}H_{k_i}\right)=\frac{1}{\sigma_L^j\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{y_s^j-\mu_L^j}{\sigma_L^j}\right)^2\right]. \tag{15}$$

For more than L hypotheses a null mass is affected.

3.3 *Spatial Information*

Since the proposed method operates only on the image histogram, the calculated BPAs of Eqs. (10), (11), and (15) are thus sensitive to noise.[10] To reduce the influence of the noise, spatial information is taken into account for each pixel. More precisely, for a current pixel y_s^j the $m(y_s^j,)$ values are re-estimated using the mass values of its neighborhood V_s centered at s and including the site s as follows:

$$\tilde{m}(y_s^j,H_i)=\frac{\Sigma_{y_t\in V_s}m(y_t^j,H_i)}{|V_s|}\quad\forall i\in\{1,\ldots,C\}, \tag{16}$$

$$\tilde{m}(y_s^j,H_i\cup H_k)=\frac{\Sigma_{y_t\in V_s}m(y_t^j,H_i\cup H_k)}{|V_s|}\quad\forall i\neq k\in\{1,\ldots,C\}, \tag{17}$$

$$\tilde{m}\left(y_s^j,\bigcup_{i=1}^{L}H_{k_i}\right)=\frac{\sum_{y_t\in V_s}m\left(y_t^j,\bigcup_{i=1}^{L}H_{k_i}\right)}{|V_s|}, \tag{18}$$

where $|V_s|$ is the cardinal of the set V_s. Before the estimated BPAs are combined, using the orthogonal sum of Eq. (4), the K value is calculated according to Eq. (19). Since fusion is performed at the pixel level, to be able to correctly fuse the d images, a precise geometrical correspondence between the d images is required to ensure that each d-uplet of pixels in the d images does represent the same physical point in the object. The proposed geometrical correspondence is based on an approach using labels configuration of a given reference image i_r. The labels of each image are permuted and the corresponding mass functions estimated so that the calculated average conflict [Eq. (20)] between sensors is minimum:

$$K(s)=\sum_{A_1\cap A_2,\ldots,\cap A_d=\varnothing}\left[\prod_{1\leqslant j\leqslant d}m_j(y_s^j,A_i)\right], \tag{19}$$

$$\overline{K}(i_r)=\sum_{s=1}^{N}K(s)/N. \tag{20}$$

The best image reference is given by

$$i_{rb}=\arg[\min_{1\leqslant i_r\leqslant d}\overline{K}(i_r)]. \tag{21}$$

Once the conflict is calculated, BPAs are combined using the orthogonal sum. Since the BPAs are derived from Gaussian probabilities, the simple hypotheses are never null.[5] Thus, a decision rule such as maximum of belief over all hypotheses will always favor compound hypotheses. In this case, a decision rule taken over simple hypotheses is preferred to compound ones:

$$x_s^n\in\omega_p\Leftrightarrow\mathrm{Bel}(y_s,H_p)=\sup_{1\leqslant i\leqslant C}\mathrm{Bel}(y_s,H_i). \tag{22}$$

The fusion resulting image is $X^n=[x_s^n]$, where n denotes the n'th iteration. Thus, at the $(n+1)$ iteration the couple of images (X^n,Y^j) is used to reestimate the statistical parameters of the Gaussian model. The maximum-likelihood estimates for mean and variance, at $(n+1)$'th iteration, are given by

$$\mu_i^j(n+1)=\frac{\Sigma_{s=1}^{N}y_s^j(1_{[x_s^n\in\omega_i]})}{\Sigma_{s=1}^{N}1_{[x_s^n\in\omega_i]}}, \tag{23}$$

$$[\sigma_i^j(n+1)]^2=\frac{\Sigma_{s=1}^{N}[y_s^j-\mu_i^j(n+1)]^2(1_{[x_s^n\in\omega_i]})}{\Sigma_{s=1}^{N}1_{[x_s^n\in\omega_i]}}, \tag{24}$$

where the indicator function $1_{[.]}$ is defined as follows:

$$1_{[x_s^n\in\omega_i]}=\begin{cases}1 & \text{if } x_s^n\in\omega_i\\ 0 & \text{otherwise}\end{cases} \tag{25}$$

The proposed iterative data fusion method is summarized as follows:

3.3.1 *Algorithm*

Fix C and L values.
Initialize the parameters $(\mu_i(0),\sigma_i(0))$ of each image histogram.
Fix the number of iterations $iter_{max}$.
for $1\leqslant n\leqslant iter_{max}$
Step 1)
$K_{min}\leftarrow 1$.
for $1\leqslant s\leqslant N$ **and** $1\leqslant j\leqslant d$
Compute $m(y_s^j,)$ according to (10) and (11) /*First Approach*/
Compute $m(y_s^j,)$ according to (15) /*Second Approach*/
Step 2)
for $1\leqslant s\leqslant N$ **and** $1\leqslant j\leqslant d$
Compute $\tilde{m}(y_s^j,)$ according to (16), (17) /*First Approach*/
Compute $\tilde{m}(y_s^j,)$ according to (18) /*Second Approach*/
Step 3)
for $1\leqslant i_r\leqslant d$
Choose the reference i_r
for $1\leqslant s\leqslant N$
Compute $K(s)$ according to (19)
Compute $\overline{K}(i_r)$ according to (20)

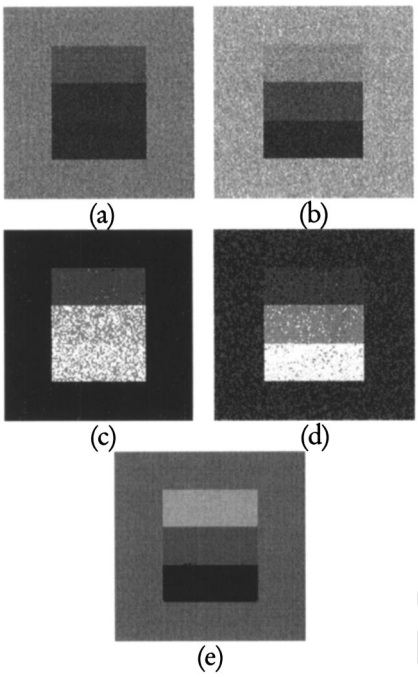

Fig. 4 Two images simulating (a) strong and (b) weak x ray acquisitions, respectively; (c) fuzzy segmentation of a weak x-ray image; (d) fuzzy segmentation of a strong x-ray image; and (e) output fused image.

```
  Find i_rb according to (21)
  If K̄(i_rb) ≤ K_min
    K_min ← K̄(i_rb)   /*Geometrical Correspondence*/
    Goto Step 1)
  EndIf
Step 4)
  Decision Making according to (22)
  Update the parameters (μ_i(n), σ_i(n)) according
  to (23) and (24)
}
```

4 Results

Two synthetic images ($d=2$) with different contrast are used to illustrate the proposed method. Two images, corrupted by Gaussian noise, simulating weak and strong x ray acquisitions are shown in Figs. 4(a) and 4(b), respectively. Each image contains four regions ($C=4$). In the first image [Fig. 4(a)], one region (smallest thickness) is confused with the background and in the second one [Fig. 4(b)], the greatest thickness is underexposed and the thicker regions are not well distinguished. The two images are combined to provide a classification of the image into four classes. The fused output image is shown in Fig. 4(e), where the four regions are well identified and discriminated. Note the very low noise effect on the segmentation result [Fig. 4(e)]. This result illustrates that the informations complementarity provided by the two sensors are well exploited by the proposed scheme and the adequate mass modeling of the information associated to the different hypotheses. The result is obtained with $|V_s|$ set to 9 and with $\text{iter}_{\max}=10$. Our method has been compared with a classical segmentation method, where each image is segmented separately using the fuzzy

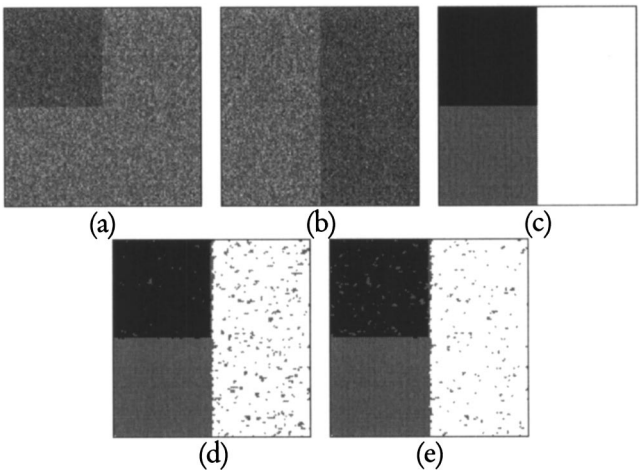

Fig. 5 Two synthetic images where the contrast has been reversed: (a) image of the first sensor; (b) image of the second sensor; (c) ''ground-truth'' image; and (d) and (e) output fused images obtained with 5 and 20 iterations, respectively.

c-means (FCM) clustering algorithm.[12] Fuzzy segmented images, shown in Figs. 4(c) and 4(d), demonstrate that none of the images provides complete and reliable information compared to DS fusion approach [Fig. 4(e)]. Figure 5 shows two synthetic images [Figs. 5(a) and 5(b)], where the contrast has been reversed and the ''ground-truth'' image is known [Fig. 5(c)]. Each sensor identifies only one class. Figures 5(d) and 5(e) show that the three regions are well brought out. The fusion result shown in Figs. 5(d) and 5(e) are performed with 5 and 20 iterations corresponding to 4.1 and 3.7% of pixels misclassification, respectively. To show the effect of the spatial information on fusion result, different neighborhood sizes are tested. Figure 6(a) shows output fused image obtained without taking the spatial information into account. One can notice the high number of misclassified pixels. This result demonstrates the interest of using neighborhood information in estimating BPAs. The effect of three neighborhoods of 9, 25, and 49 pixels is illustrated

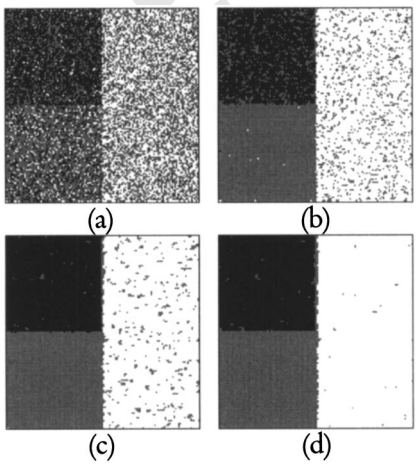

Fig. 6 Effect of the neighborhood size on the fusion result: (a) output fused image without pixel spatial information; and (b), (c), and (d) output fused images obtained with neighborhoods of 9, 25, and 49 pixels, respectively.

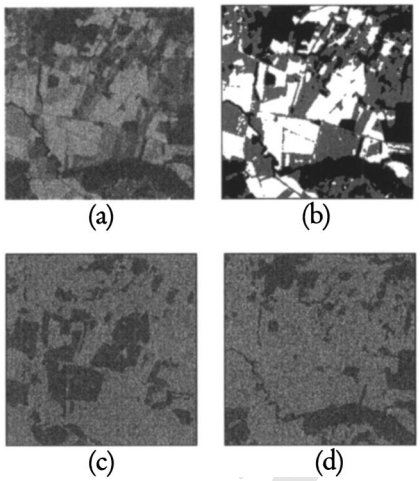

Fig. 7 Simulation of the output of two sensors using a real satellite image: (a) SPOT image of the Aquitaine region, (b) segmentation result of the SPOT image into three regions, and simulated output of the first (c) and the second (d) sensors.

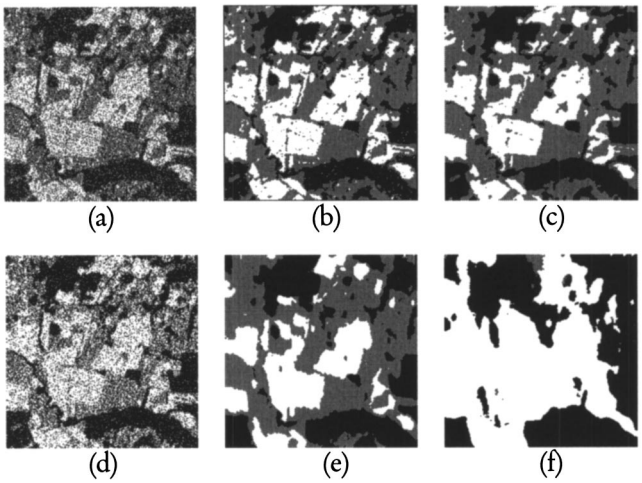

Fig. 9 Effect of the neighborhood size on the fusion result. Output fused images obtained by using (a) the first and (d) the second approaches without taking into account the spatial information. Output fused images obtained with neighborhoods of 9 and 25 pixels obtained by using, respectively, (b) and (c) the first and (e) and (f) the second approaches.

in Figs. 6(b), 6(c), and 6(d), respectively. Note the significant improvement of the segmentation. In particular, as the neighborhood size increases, the homogeneity of the segmented regions also increases. These results are obtained with only five iterations. Figure 7(a) shows a real satellite image of Aquitaine region (SouthWest France) acquired by the SPOT system [Center National d'Etudes Spaciales (CNES) Toulouse]. Figure 7(b) shows the result of segmentation of this image [Fig. 7(a)] into three regions (R_1,R_2,R_3) using clustering Expectation Maximization (EM) algorithm. To simulate the output of two sensors we add a Gaussian noise to the segmented image [Fig. 7(b)]. For the first sensor, we add a noise so that the gray-level distributions of regions R_1 and R_2 have the same statistics $(\mu_1=\mu_2=3,\ \sigma_1^2=\sigma_2^2=1)$. For region R_3 the statistics are $(\mu_3=1,\ \sigma_1^2=1)$. Thus, regions R_1 and R_2 are not distinguishable by the first sensor [Fig. 7(c)]. For the second sensor, only R_1 is discriminated $(\mu_1=3,\ \sigma_1^2=1)$. Regions R_2 and R_3 are merged $(\mu_2=\mu_3=1,\ \sigma_2^2=\sigma_3^2=1)$ [Fig. 7(d)]. Finally, each sensor identifies only two regions. To fuse the outputs of the two sensors, the image shown in Fig. 7(b) is used as the ''ground-truth'' image. Figures 8(a) and 8(b) show that the three regions are well brought out by the first[7] and the second approaches. The first approach provides the details well (roads, boundaries between the fields), whereas the second one tends to keep the homogeneous parts of the image. This can be explained by the fact that the second approach gives importance to compound hypothesis and so eliminates the small areas. Furthermore, this approach includes more combinations of all hypothesis than the first one. The fusion result shown in Figs. 8(a) and 8(b) are performed with 10 iterations corresponding respectively to 8.7 and 16.6% of pixels misclassification. To show the effect of the spatial information on fusion result, different neighborhood sizes are tested. Figures 9(a) and 9(d) show the output fused image obtained without taking into account the spatial information obtained by the first and the second approaches, respectively. We can notice the high number of misclassified pixels. This result demonstrates the interest of using neighborhood information in estimating BPAs. The effect of two neighborhoods of 9 and 25 pixels is illustrated in Figs. 9(b) and 9(c) and 9(e) and 9(f) for both approaches, respectively. Note the significant improvement of the segmentation in the case of the first approach, despite the fact that the second one does not give good results for the 25 neighbors. In all cases, a small neighborhood (9 pixels) should be a good compromise to keep the local information. Actually as the neighborhood size increases, the homogeneity of the segmented regions also increases. These results were obtained with only 10 iterations.

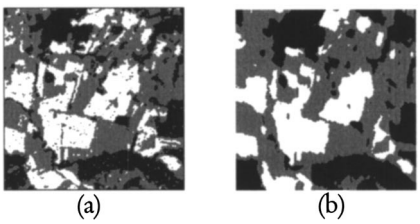

Fig. 8 Fusion result of the simulated output images using (a) the first and (b) the second methods.

5 Conclusions

An iterative unsupervised multisensor classification method using DS theory was described. BPAs, modeled as a Gaussian distribution, were estimated in an unsupervised way at each iteration using the output fused image and the source images. The results obtained on synthetic and real images were encouraging. A comparison of the two approaches and the effect of the neighborhood sizes was presented. The results showed the interest of the contextual information and to take into account more or less compound hypothesis. Extensive tests on real data are necessary to evaluate the method. Investigation of decision rules other than maximum belief are also necessary. A Gaussian model of the

data is widely accepted, but it is not the best choice in all applications. For example, in sonar imaging, an image histogram is better modeled as a Rayleigh distribution than as a Gaussian one. Thus, the proposed fusion scheme presents limitations when information sources can not be described correctly by a Gaussian model.

Acknowledgment

The authors would like to thank Pr. Isabelle Bloch from Ecole nationale supérievre des télécoms (ENST)-Paris for her helpful comments and suggestions.

References

1. R. C. Luo and M. G. Kay, ''Data fusion and sensor integration: state of the art 1990s,'' in M. A. Abibdi and R. C. Gonzalez, Eds., *Data Fusion in Robotics and Machine Intelligence*, Academic Press, San Diego (1992).
2. A. Dempster, ''Upper and lower probabilities induced by a multivalued mapping,'' *Ann. Math. Stat.* **38**, 325–339 (1967).
3. G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, NJ (1976).
4. G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic. Theory and Applications*, Prentice Hall PTR, Upper Saddle River, NJ (1995).
5. S. Le Hegarat-Mascle, I. Bloch, and D. Vidal-Madjar, ''Application of DS evidence theory to unsupervised classification in multiple remote sensing,'' *IEEE Trans. Geosci. Remote Sens.* **35**(4), 1018–1031 (1997).
6. T. Lee, J. A. Richards, and R. H. Swain, ''Probabilistic and evidential approaches for multisource data analysis,'' *IEEE Trans. Geosci. Remote Sens.* **25**, 283–293 (1987).
7. F. Salzenstein and A. O. Boudraa, ''Unsupervised multisensor data fusion approach,'' in *Proc. 6th ISSPA*, Vol. 1, pp. 152–155, Kuala Lumpur, Malaysia (2001).
8. H. Rasoulian, W. E. Thompson, L. F. Kazda, and R. Parra-Loera, ''Application of the mathematical theory of evidence to the image cueing and image segmentation problem,'' *Proc. SPIE* **1310**, 199–206 (1990).
9. I. Bloch, ''Some aspect of Dempster-Shafer evidence theory for classification of multimodality medical images taking partial volume effect into account,'' *Pattern Recogn. Lett.* **17**(8), 905–916 (1996).
10. L. Bentabet, S. Jodouin, and A. O. Boudraa, ''Estimation of mass functions in Dempster-Shafer theory using fuzzy clustering and spatial information,'' in *Proc. 14th Int. Conf. on Vision Interface*, pp. 149–156, Toronto, Canada (2001).
11. P. Smets, ''The combination of evidence in the transferable belief model,'' *IEEE Trans. Pattern Anal. Mach. Intell.* **12**, 447–458 (1990).
12. J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York (1981).

Biographies and photographs of the authors not available.

PROOF COPY 014406JOE