# Adaptive Fuzzy Evidential Reasoning for Automated Brain Tissue Segmentation

**Hongwei Zhu**
Dept. of Systems Design Engineering
University of Waterloo
200 University Ave. W., Waterloo, ON, N2L 3G1
Canada
h4zhu@engmail.uwaterloo.ca

**Otman Basir**
Dept. of Systems Design Engineering
University of Waterloo
200 University Ave. W., Waterloo, ON, N2L 3G1
Canada
obasir@uwaterloo.ca

**Abstract –** *This paper presents an adaptive fuzzy evidential reasoning approach, for segmenting multi-modality MR brain images. A novel fuzzy evidence structure model is proposed under the assumption that each information source provides two types of evidence: probabilistic evidence and fuzzy evidence. A new information measure, called hybrid entropy, is employed for evaluating the overall uncertainty contained in a fuzzy evidence structure. For adaptive reasoning, two discounting strategies are included. To handle conflict between the probabilistic evidence and the fuzzy evidence, local discounting takes into account Kullback-Leibler distance between the two types of evidence. Global discounting takes into account source quality, in terms of Shannon entropy and hybrid entropy, for dealing with conflict of sources. To demonstrate its effectiveness, the approach is applied to segmenting multi-modality MR brain images. It is concluded that the proposed approach performs better than K-mean clustering, majority voting, fuzzy set operators, and Bayesian approach.*

**Keywords:** Dempster-Shafer evidence theory, adaptive fuzzy evidential reasoning, discounting, MRI, brain tissue segmentation.

## 1 Introduction

Brain tissue segmentation is an important step in many medical imaging applications, including volume analysis, visualization of regions of interest, and diagnosis of brain diseases. Segmenting MR images has been found a quite hard problem due to the existence of image noise, partial volume effects, the presence of smoothly varying intensity inhomogeneity, and large amounts of data to be processed. To handle these difficulties, a large number of approaches have been studied, including fuzzy logic methods [1], neural networks [2], Markov random field methods with the maximum expectation [3], statistical methods [4], and data fusion methods [5], to name a few.

As one typical data fusion problem, the segmentation of multi-modality brain MR images aims at achieving improved segmentation performance by taking advantage of redundancy and complementariness in information provided by multiple sources. There have existed many data fusion methodologies, which are capable of reasoning under various types of uncertainty. Typical ones include probability theory based approaches, fuzzy set and fuzzy logic theory based approaches, and Dempster-Shafer evidence theory based approaches [6]. As a hybrid body of these three paradigms, fuzzy Dempster-Shafer evidential reasoning (FDS) has been introduced and mainly studied in the domain of artificial intelligence and approximate reasoning [7–9]. FDS is a generalized version of the Dempster-Shafer evidence theory with fuzzy sets serving as focal elements. FDS possesses a quite pleasing framework for managing randomness, fuzziness and conflict. However, when FDS is applied to real-word applications, the following difficulties often arise: 1) how to effectively formulate evidence as fuzzy evidence structures; 2) how to reason about different types of uncertainty in a consistent and coherent manner, and how to effectively cope with conflicts.

For dealing with these difficulties, a novel fuzzy evidence structure model is first proposed under the assumption that every information source to be fused provides two types of evidence: probabilistic evidence in terms of posteriori probabilities, and fuzzy evidence in terms of fuzzy sets or fuzzy rules. An adaptive fuzzy evidential reasoning scheme is then designed by discounting evidence in two manners: local discounting and global discounting. An efficient implementation scheme is further developed, which simplifies the adaptive fuzzy evidential reasoning and reduces computational costs. For automated segmentation of brain tissues, each of three modality MR images, namely T1, T2, and PD, is treated as an information source equipped with a fuzzy evidence structure. To obtain the probabilistic evidence, the K-mean clustering algorithm is first applied to each modality MR image. By means of knowledge about pixel intensities of brain tissues, each cluster is further associated with one of three brain tissues: white matter (WM), grey matter (GM), and cerebrospinal fluid (CSF). Further, posteriori probabilities are derived by the Parzen window technique and the Bayes rule [10], based on pixel intensities of clusters. For the determination of the fuzzy evidence, fuzzy focals are established by the Mamdani fuzzy inferencing scheme in which a set of fuzzy rules formulate the domain knowledge about spatial similarities between neighboring pixels. A local discounting factor is then adaptively determined by the Kullback-Leibler distance measure between the probabilistic evidence and the fuzzy evidence. A global discounting factor is determined by the average Shannon entropy and the average hybrid entropy of the investigated modality. To examine the performance of the adaptive approach, experiments are carried out on the Simulated Brain Database (SBD) [11] for normal data sets. Comparisons have been made between the adaptive approach and other relevant approaches, such as K-mean clustering on both single-

modality images and the stacked vector form of three modality images, majority voting, fuzzy set operators, and Bayesian approach. Results have shown that the proposed approach outperforms all the others, in terms of segmentation accuracies and robustness to the variation of the probabilistic evidence.

The remainder of this paper is organized as follows. The Dempster-Shafer evidence theory and fuzzy evidential reasoning are briefly reviewed in Section 2. Section 3 presents an adaptive fuzzy evidential reasoning approach for automated brain tissue segmentation. Experiments and results are provided in Section 4, and conclusions are finally presented in Section 5.

## 2 Review of DSET and FDS

The Dempster-Shafer evidence theory (DSET) provides a pleasing framework to aggregate evidence of multiple information sources [6]. DSET deals with uncertainty and imprecision in three levels: 1) representing evidence by focal elements and masses, 2) combing evidence by the Dempster's rule, and 3) making decisions.

A *mass function* is a mapping $m : 2^{\Omega} \rightarrow [0,1]$, satisfying $\sum_{A\subseteq\Omega} m(A) = 1$ and $m(\emptyset) = 0$, where $\Omega = \{\omega_1, \cdots, \omega_c\}$ is called a frame of discernment. $2^{\Omega}$ is the corresponding power set. Subset $A$ with non-zero mass is called a *focal element* (focal for short). Focal elements and their masses construct an *evidence structure*, expressed in the form: $\{(A, m(A))|A \subseteq \Omega, m(A) > 0\}$. The Dempster's rule combines multiple sources which are assumed independent of each other. Commonly used evidential measures in DSET are belief, plausibility, and commonality, based on which final decisions can be made. The belief measure is defined by $\text{Bel}(A) = \sum_{B\subseteq A} m(B)$; the commonality measure is defined by: $Q(A) = \sum_{A\subseteq B} m(B)$, where $A \subseteq \Omega, B \subseteq \Omega$.

To handle conflict of information sources, a discounting scheme has been introduced in DSET such that the belief function Bel is modified by $\text{Bel}(A) = (1-\alpha) \times \text{Bel}(A),\ \forall A \subset \Omega$, and $\text{Bel}(\Omega) = 1$. Correspondingly, mass functions are usually modified in the following manner [6, 12, 13]:

$$m^{\alpha}(A) = \begin{cases} (1-\alpha) \times m(A), & \text{if } A \subset \Omega \\ \alpha + (1-\alpha) \times m(\Omega), & \text{if } A = \Omega \end{cases} \quad (1)$$

where $\alpha \in [0,1]$ denotes a discounting factor, and $m^{\alpha}(A)$ denotes the discounted mass of $m(A)$. The larger is $\alpha$, the more masses are discounted from non frame of discernment focals, while the more mass is assigned to the frame of discernment $\Omega$.

Fuzzy evidential reasoning (FDS) is derived from DSET, in which focal elements are replaced by fuzzy sets [7]. Similarly as DSET, fuzzy focal elements and their masses construct a *fuzzy evidence structure*, denoted by

$$\text{FS} = \{(A, m(A), \mu_A(x)) \,|A \subseteq \Omega, m(A) > 0, x \in U\} \quad (2)$$

where $A$ denotes a fuzzy focal element with membership $\mu_A(x)$ for its element $x \in U$; $U = \{x_1, x_2, \cdots, x_u\}$ denotes the universe of discourse. One way to combine multiple fuzzy evidence structures is to apply Dempster's rule to derived crisp evidence structures as follows by two steps [14, 15].

1) Given a piece of fuzzy evidence: $(A, m(A), \mu_A(x))$, fuzzy focal $A$ is decomposed into its $\alpha$-cuts using the resolution identity principle of the fuzzy set theory: $A = \bigcup_{\alpha_j} A'_{\alpha_j}$, where $\bigcup$ denotes the set union operation (max); $\alpha_j = \mu_A(x_j)$, and $A'_{\alpha_j}$ denotes a fuzzy set associated with $A$'s $\alpha$-cut: $A_{\alpha_j} = \{x_1, \cdots, x_j\}$ at level $\mu_A(x_j)$. $\mu_A(x_1) \geq \mu_A(x_2) \cdots \geq \mu_A(x_u)$ is assumed here. $A'_{\alpha_j}$ has two membership values: $\mu_{A'_{\alpha_j}}(x) = \mu_A(x_j)$ if $x \in A_{\alpha_j}$; 0 otherwise. $j = 1, \cdots, u$.

2) Having obtained $\alpha$-cuts, the mass of fuzzy focal $A$ is distributed to its $\alpha-$cuts by

$$m(A_{\alpha_j}) = \frac{m(A)}{\max_x \mu_A(x)} \times (\mu_A(x_j) - \mu_A(x_{j+1})) \quad (3)$$

where $\max_x \mu_A(x)$ is the maximum membership in fuzzy focal $A$. This scheme for mass assignment is similar to that suggested by [14].

## 3 The proposed approach

### 3.1 A fuzzy evidence structure model

For a multi source data fusion problem, a fuzzy evidence structure model is proposed through the following two steps, in light of the assumption that each source provides two types of evidence: probabilistic evidence in terms of posteriori probabilities, and fuzzy evidence in terms of a set of fuzzy sets or fuzzy rules.

- Step 1: Form an evidence structure for each source using the Proportional Difference Evidence Structure scheme (PDES) developed in our previous work [16]. PDES is similar to the possibility distribution to mass function conversion scheme suggested in [17]. For details of PDES, please refer to [16].

- Step 2: Convert the crisp evidence structure to a fuzzy evidence structure. This is realized through replacing crisp focal elements of the resultant evidence structure by available fuzzy sets or consequent fuzzy sets from fuzzy inference systems.

To facilitate understanding the essence of the adaptive fuzzy evidential reasoning scheme to be developed, a property of PDES is presented below.

**Property 1** *Given $n$ probabilistic evidential sources, in terms of posteriors:*

$$PE_i = \{p(\omega_1^i|X_i), \cdots, p(\omega_c^i|X_i)\}$$

*ranked in a non-increasing order and priors:*

$$PE_0 = \{p(\omega_1^0), \cdots, p(\omega_c^0)\}$$

*ranked in a non-decreasing order, where $X_i$ denotes the observation of source $i$ for $i = 1, \cdots, n$ and $\{\omega_1^{i'}, \cdots, \omega_c^{i'}\}$ is a permutation of the frame of discernment $\Omega = \{\omega_1, \cdots, \omega_c\}$ for $i' = 0, \cdots, n$, the DSET based reasoning using the maximum commonality decision rule (MCD: $\omega = \arg\max_{\omega_j} Q(\{\omega_j\})$) on the combined evidence structure is equivalent to the Bayesian approach using the maximum a posterior principle (MAP), if the following conditions are satisfied, where $j = 1, \cdots, c$:*

*(1) the sources are independent of each other;*

*(2) the evidence structure associated with each source is constructed by the PDES scheme:*

$ES_i =$

$$\left\{ \left( \{\omega_1^i, \cdots, \omega_j^i\}, \frac{p(\omega_j^i|X_i) - p(\omega_{j+1}^i|X_i)}{p(\omega_1^i|X_i)} \right) \middle| \ j \right\} \quad (4)$$

*(3) the prior evidence structure is constructed by:*

$ES_0 =$

$$\left\{ \left( \{\omega_1^0, \cdots, \omega_j^0\}, \frac{p^{-(n-1)}(\omega_j^0) - p^{-(n-1)}(\omega_{j+1}^0)}{p^{-(n-1)}(\omega_1^0)} \right) \middle| \ j \right\}$$

We refer to the Dempster-Shafer evidential reasoning applied to the above PDES derived evidence structures as PDES associated DSET. Notice that by definition, $\{\omega_1^i, \cdots, \omega_j^i\}$ should not be considered as a focal if $p(\omega_j^i|X_i) = p(\omega_{j+1}^i|X_i)$, and would be excluded from $\text{ES}_i$. For illustration purpose, we use Eq. (4) since including 0 masses does not influence the results of reasoning from a mathematical point of view [16]. So is for $p^{-(n-1)}(\omega_j^0) = p^{-(n-1)}(\omega_{j+1}^0)$ in $\text{ES}_0$.

By Step 2, corresponding to Eq. (4) a fuzzy evidence structure can be obtained as:

$\text{FS}_i =$

$$\left\{ \left( A_j^i, \frac{p(\omega_j^i|X_i) - p(\omega_{j+1}^i|X_i)}{p(\omega_1^i|X_i)}, \mu_{A_j^i}(\omega_l^i) \right) \middle| \ j, l \right\} \quad (5)$$

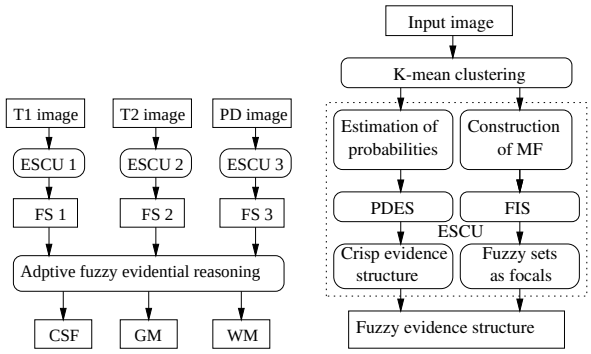
where $A_j^i$ is a fuzzy set associated with source $i$ and its membership is denoted by $\mu_{A_j^i}(\omega_l^i)$ for elements in its support set $A_j^{0i} = \{\omega_1^i, \cdots, \omega_j^i\}$. $j = 1, \cdots, c$, and $l = 1, \cdots, j$. Correspondingly, a piece of fuzzy evidence is represented as:

$$\left( A_j^i, \frac{p(\omega_j^i|X_i) - p(\omega_{j+1}^i|X_i)}{p(\omega_1^i|X_i)}, \mu_{A_j^i}(\omega_l^i) \right) \quad (6)$$

Notice that it is assumed here that $\Omega$ is identical with $U$ in the fuzzy evidence structure form given in Eq. (2). The motivation of using the above fuzzy evidence structure model is to show that the use of both fuzzy evidence and probabilistic evidence through fuzzy evidential reasoning is able to produce better data fusion performance than the PDES associated DEST (i.e., equivalently Bayesian approach due to Property 1). Next we adopt this fuzzy evidence structure model to the brain tissue segmentation problem.

## 3.2 Constructing evidence structures

Three modalities, namely T1, T2 and PD images, are treated as distinct information sources, with pixel intensity denoted by $x_i$ ($i = 1, 2, 3$). The data fusion based segmentation paradigm is given in Fig. 1 (a), where ESCU denotes the evidence structure construction unit as sketched in Fig. 1 (b). MF stands for membership function, and FS stands for a fuzzy evidence structure in form (5). Given a MR image, the K-mean algorithm is first employed to cluster pixel intensities and each obtained cluster is further associated with one of three brain tissues: CSF ($\omega_1$), GM ($\omega_2$), and WM ($\omega_3$) by means of domain knowledge. Therefore $\Omega = \{\text{CSF}, \text{GM}, \text{WM}\}$. Given a pixel of source $i$, a fuzzy evidence structure is constructed based on probabilistic evidence and fuzzy evidence. The probabilistic evidence is represented by posteriori probabilities $p^i = \{p(\omega_1^i|x_i), p(\omega_2^i|x_i), p(\omega_3^i|x_i)\}$, which are obtained by the Parzen window technique and the Bayes rule. Notice that $\{\omega_1^i, \omega_2^i, \omega_3^i\}$ is a permutation of $\Omega$ satisfying $p(\omega_1^i|x_i) \geq p(\omega_2^i|x_i) \geq p(\omega_3^i|x_i)$. Associated with Property 1, the prior probabilities of classes are assumed to be proportional to the numbers of pixels within three resultant clusters of T1 images, since K-mean clustering applied to T1 images yields the best single-modality based result, as will be justified by later experiments. The fuzzy evidence in terms of a fuzzy set $A^i = \sum_{j=1}^{c} \mu_j^i / \omega_j^i$, where $c = 3$, is constructed by a fuzzy inference system (FIS) which takes into account the spatial similarity of neighboring pixels. For details about how to obtain the probabilistic evidence and the fuzzy evidence, please refer to [18].




(a) The overall paradigm (b) ESCU

Fig. 1: Paradigm of brain tissue segmentation.

Having $p^i$ and $A^i$ for source $i$, a fuzzy evidence structure is constructed in Eq. (5), where $\mu_{A_j^i}(\omega_l^i) = \mu_{A^i}(\omega_l^i)$. Further, two discounting schemes are designed below for adaptive fuzzy evidential reasoning.

## 3.3 Adaptive fuzzy evidential reasoning

### *3.3.1 K-L distance guided local discounting*

Recall that the decomposition of fuzzy evidence structures is required before the Dempster's rule can be employed for the fuzzy evidential reasoning. Following the notation of discounting as used in DSET to deal with conflict of sources, we deal with the dissonance during the decomposition, by taking advantage of the Kullback-Leibler distance between the two types of evidence. To specify, we rewrite here $p^i = \left\{p(\omega_1^i|x_i), p(\omega_2^i|x_i), p(\omega_3^i|x_i)\right\}$ representing the posteriori probabilities, and

$$A^i = \left\{ \frac{\mu_1^i}{\omega_1^i}, \frac{\mu_2^i}{\omega_2^i}, \frac{\mu_3^i}{\omega_3^i} \right\}$$

representing a fuzzy set. $\mu_j^i$'s are first normalized by ${\mu'}_j^i = \frac{\mu_j^i}{\mu_1^i + \mu_2^i + \mu_3^i}$ for $j = 1, 2, 3$, and a distance between $p^i$ and ${\mu'}^i = \left\{ {\mu'}_1^i, {\mu'}_2^i, {\mu'}_3^i \right\}$ is then defined as

$$D({\mu'}^i, p^i) = D({\mu'}^i || p^i) + D(p^i || {\mu'}^i)$$

where $D(\cdot||\cdot)$ denotes the Kullback-Leibler distance measure. The discounting is realized by the following two steps.

1) Modify membership values of fuzzy focal element $A_j^i$

$$\mu''_{A_j^i}(\omega_l^i) =$$

$$r(D(\mu'^i, p^i)) \times (\mu_{A_j^i}(\omega_l^i) - \min \mu_{A_j^i}) + \min \mu_{A_j^i} \quad (7)$$

where

$$r(D(\mu'^i, p^i)) = \frac{1}{1 + \exp(-K_i \times (D(\mu'^i, p^i) - D_{i0}))}$$

and $\min \mu_{A_j^i}$ is the minimum membership value of elements in fuzzy focal $A_j^i$. $r$ is defined by a unipolar sigmoidal function mapping $D_i(\mu'^i, p^i)$ to range $[0, 1]$. Parameter $K_i > 0$ controls the variation speed of the function and $D_{i0}$ ensures $r(D_{i0}) = 0.5$.

2) Apply the decomposition scheme (3) to the modified memberships in Eq. (7). It can be shown that the above adjustment of memberships followed by the decomposition scheme in Eq. (3) is essentially equivalent to a discounting scheme, similar to that of Eq. (1) if the derived set of crisp evidence is viewed as a special evidence structure. In the context of Eq. (1), the discounting factor can be computed as:

$$\alpha_i^j = 1 - \frac{r(D(\mu'^i, p^i)) \times \max \mu_{A_j^i}}{r(D(\mu'^i, p^i)) \times (\max \mu_{A_j^i} - \min \mu_{A_j^i}) + \min \mu_{A_j^i}}.$$

$\max \mu_{A_j^i}$ is the maximum membership value of elements in fuzzy focal $A_j^i$.

This discounting scheme is pixel based, and the discounting factor is affected by the distance between the probabilistic evidence and the fuzzy evidence. Therefore, we call this discounting the pixel-level local discounting.

Let us investigate the nature of this discounting. If $\mu'^i$ is equal to $p^i$, i.e., $D(\mu'^i, p^i) = 0$, $r$ is approaching 0. Therefore, all obtained memberships $\mu''_{A_j^i}$ approach $\min \mu_{A_j^i}$. As a result of applying the decomposition scheme (3) to the locally discounted evidence structure, FDS is almost degraded to DSET on the PDES derived evidence structures (for the probabilistic evidence). Thus, FDS is nearly equivalent to the Bayesian approach due to Property 1. On the other hand, if $\mu'^i$ is largely different from $p^i$, with a large value of $D(\mu'^i, p^i)$, $r$ approaches 1. This indicates less discounting on the fuzzy evidence structure. Therefore, that the masses of fuzzy focals are redistributed to their $\alpha$-cuts is mainly controlled by their membership values. This would be reasonable, since the probabilistic evidence tends to be more sensitive to image noise than the fuzzy evidence. Recall that the fuzzy evidence is obtained by taking into account eight neighboring pixels of a central pixel, while the probabilistic evidence is determined from one pixel intensity by the Parzen window technique. A large value of $D(\mu'^i, p^i)$ signifies that it is rather important to use the fuzzy evidence.

Furthermore, to make the local discounting scheme effective, $K_i$ and $D_{i0}$ need to be determined. Our strategy is to make them source sensitive. Give a source $i$, the K-L based distances of all pixels within ROI are calculated, and the minimum and the maximum ones are identified, say $\min D_i$ and $\max D_i$. $D_{i0}$ is set to $(\min D_i + \max D_i)/2$, and $K_i$ is decided by two assumed mappings: $r(\min D_i) = r_l$, and $r(\max D_i) = 1 - r_l$, where $0 < r_l < 0.5$. Solving the two equations yields a common value $K_i = -\left(\log(\frac{1}{r_l} - 1)\right)/(r_l - D_{i0})$. In the experiments presented in the next section, we empirically set $r_l = 0.1$.

### *3.3.2 Global discounting*

The above discounting scheme happens during the decomposition and it takes into account the dissonance between the two types of evidence inside each fuzzy evidence structure. As usual, evidence structures would be of different quality and may conflict each other. One way to handle this issue in the Dempster-Shafer evidence theory is to take the discounting scheme presented in Section 2. Therefore, in addition to the above local discounting, we develop a second discounting scheme with discounting factors determined by two information measures: Shannon entropy and hybrid entropy. Given probabilistic evidence $p^i$ for source $i$, Shannon entropy $H(p^i)$, which quantifies the amount of uncertainty contained in $p^i$, is computed by

$$H(p^i) = \sum_{\omega_j \in \Omega} -p(\omega_j|x_i) \log_2 p(\omega_j|x_i). \quad (8)$$

Hybrid entropy is a new information measure which quantifies the overall uncertainty contained in a fuzzy evidence structure [5]. For fuzzy evidence structure FS $= \{(A_1, m_1, \mu_{A_1}(x)), \cdots, (A_c, m_c, \mu_{A_c}(x))\}$, hybrid entropy FH(FS) is defined as [5]

$$\text{FH(FS)} = -\sum_{j=1}^{c} m_j \log_2(m_j(1 - \tilde{F}_j)) \quad (9)$$

$$\text{with } \tilde{F}_j = \frac{1}{|A_j^0|} \sum_{l=1}^{|A_j^0|} \frac{\mu_{A_j \cap \bar{A}_j}(x_{jl})}{\mu_{A_j \cup \bar{A}_j}(x_{jl})}$$

where $\tilde{F}_j$ denotes the fuzzy entropy of fuzzy set $A_j$ [19], and $x_{jl}$ denotes the $l$th element in $A_j$. $A_j^0$ denotes the support set of $A_j$, and $|A_j^0|$ is the cardinality of $A_j^0$. Hybrid entropy possesses following properties (where $\mu_{jl} = \mu_{A_j}(x_{jl})$, and $\mu_{jl}^* = \mu_{A_j^*}(x_{jl})$; $l = 1, \cdots, |A_j^0|$):

1) FH(FS) is *minimum* if $m_j = 0$ or 1, and $A_j$ is a crisp set, i.e., $\mu_{jl} = 0$ or 1;

2) FH(FS) is *maximum* if $m_j = \frac{1}{c}$ and $A_j$ is the most fuzzy set, i.e., $\mu_{jl} = 0.5$;

3) FH(FS) $\geq$ FH(FS$^*$), if $\tilde{F}_j \geq \tilde{F}_j^*$, where $\tilde{F}_j^*$ is a sharpened version of $\tilde{F}_j$; $\tilde{F}_j$ and $\tilde{F}_j^*$ are associated with FS and FS$^*$ respectively; $\tilde{F}_j$ and $\tilde{F}_j^*$ have the same associated masses. That $\tilde{F}_j^*$ of fuzzy set $A_j^*$ is a sharpened version of $\tilde{F}_j$ of fuzzy set $A_j$ is defined by [20]

$$0 \leq \mu_{jl}^* \leq \mu_{jl} \leq 0.5 \quad \text{for } 0 \leq \mu_{jl} \leq 0.5$$
$$0.5 \leq \mu_{jl} \leq \mu_{jl}^* \leq 1 \quad \text{for } 0.5 \leq \mu_{jl} \leq 1$$

Therefore, giving a fuzzy evidence structure FS$_i$ for a pixel of source $i$, hybrid entropy FH(FS$_i$) can be computed by Eq. (9). For considering source quality over all pixels of ROI of

brain, the average Shannon entropy and the average hybrid entropy are taken into consideration. Therefore, we obtain the average Shannon entropy $H_i$ and the average hybrid entropy $\mathrm{FH}_i$ with $i = 1, 2, 3$ corresponding to T1, T2, and PD images. Obviously, the smaller average hybrid entropy an image has, the less uncertain the image is, and the more distinguishable its classes are. During the computation of the average hybrid entropy, we set $-m_j \log_2(m_j(1-\tilde{F}_j)) = 5$ if $-m_j \log_2(m_j(1-\tilde{F}_j)) > 5$. This essentially introduces an upper bound to $-m_j \log_2(m_j(1-\tilde{F}_j))$, and makes the discounting manageable since $-m_j \log_2(m_j(1-\tilde{F}_j))$ in Eq. (9) is positive infinity if $\tilde{F}_j = 1$. The rationale behind this consideration can be justified by numerical computations. $-m_j \log_2(m_j(1-\tilde{F}_j)) = 5$ indicates that whatever values $m_j$ takes in the range $(0, 1]$, $\tilde{F}_j$ approximately approaches 1, indicating that the corresponding fuzzy focal is almost most fuzzy. For example, if $m_j = 1$, $-m_j \log_2(m_j(1-\tilde{F}_j)) = 5$ indicates $\tilde{F}_i = 0.9688$; if $m_j = 0.5$, $-m_j \log_2(m_j(1-\tilde{F}_j)) = 5$ indicates $\tilde{F}_j = 0.9980$; if $m_j = 0.1$, $-m_j \log_2(m_j(1-\tilde{F}_j)) = 5$ implies nearly $\tilde{F}_j = 1$.

Based on $H_i$'s and $\mathrm{FH}_i$'s, a discounting factor is finally defined for source $i$ by

$$\alpha_i = \frac{0.2H_i}{\log_2 3} \times \frac{\mathrm{FH}_i - \min \mathrm{FH}}{\max \mathrm{FH} - \min \mathrm{FH}} \tag{10}$$

where $\max$ FH and $\min$ FH denote respectively the maximum and the minimum of FH1, FH2, and FH3.

Eq. (10) shows that among the three modalities, the best quality modality is discounted only by the average Shannon entropy $H_i$. Denominator $\log_2 3$, which is the maximum Shannon entropy for a three class problem, is used to normalize average Shannon entropy $H_i$. Parameter $0.2$ is set empirically, which constrains the maximum value of the discounting factor. The smaller is $H_i$, the less uncertain is the corresponding probabilistic evidence; Thus, the smaller is $\alpha_i$, and the less discounted is this information source. Modalities other than the best one is further discounted by a ratio of the differences of average hybrid entropies. The smaller $\mathrm{FH}_i$ is, the less discounted is the modality.

In contrast to the previous local discounting, this discounting takes into account information summarized from all pixels in ROI. Thus, we refer it as global discounting.

### 3.4 Efficient implementation

We now calculate the commonalities of singleton class sets for an resultant evidence structure. As the local discounting is employed, the commonality of a singleton class set for the piece of fuzzy evidence in Eq. (6) is

$$Q_j^i(\{\omega_l^i\}) = \beta_{jl}^i m(A_j^i) \delta\left(\omega_l^i \in A_j^{0i}\right) \tag{11}$$

where $m(A_j^i) = (p(\omega_j^i|X_i) - p(\omega_{j+1}^i|X_i))/p(\omega_1^i|X_i)$, and $\beta_{jl}^i = \mu''_{A_j^i}(\omega_l^i)/\max \mu''_{A_j^i}$ is referred to as an expansion ratio ($\mu''_{A_j^i}(\omega_l^i)/\max \mu''_{A_j^i}$ is not less than $\mu_{A_j^i}(\omega_l^i)/\max \mu_{A_j^i}$ since $0 \le r \le 1$ in Eq. (7)). $\mu''_{A_j^i}(\omega_l^i)$ is computed by Eq. (7), and $\max \mu''_{A_j^i}$ denotes the maximum of $\mu''_{A_j^i}(\omega_l^i)$ for $\omega_l^i \in A_j^i$. $A_j^{0i}$ denotes the support set of $A_j^i$. Further fuzzy evidence structure $\mathrm{FS}_i$ is discounted with discounting factor $\alpha_i$ in Eq. (10), thus the commonality in Eq. (11) equivalently becomes $Q_j^i(\{\omega_l^i\}) =$

$$\begin{cases} \beta_{jl}^i(1-\alpha_i)m(A_j^i)\delta\left(\omega_l^i \in A_j^{0i}\right), & \text{if } j \neq c \\ \beta_{jl}^i(1-\alpha_i)m(A_c^i) + \alpha_i, & \text{if } j = c \end{cases} \tag{12}$$

The case $j \neq c$ can be deduced directly from the part of "if $A \subset \Omega$" in Eq. (1); All derived crisp pieces of evidence do not include the frame of discernment $\Omega$ as their focals; Hence, the discounting process directly produces the proportional commonalities of singleton class sets. Likewise, case $j = c$ in Eq. (12) is also due to discounting scheme (1). For clarification, this part can be rewritten as:

$$\begin{aligned} Q_j^i(\{\omega_l^i\}) = &\frac{\mu''_{A_c^i}(\omega_l^i) - \min \mu''_{A_c^i}}{\max \mu''_{A_c^i}} \times (1-\alpha_i)m(A_c^i) \\ &+\alpha_i + \frac{\min \mu''_{A_c^i}}{\max \mu''_{A_c^i}} \times (1-\alpha_i)m(A_c^i), \end{aligned}$$

where the first part corresponds to that contributed to by non $\Omega$ focal elements, while the remaining parts are associated with focal $\Omega$.

The resultant commonality part takes a simple computational form, and it is conceptually interpretable. Eq. (12) shows that the commonalities are adjusted in two manners: multiplying factor $(1-\alpha_i)$ and expansion ratio $\beta_{jl}^i$, as well as adding baseline commonality $\alpha_i$ for all singleton class sets in the case $j = c$.

By definition of commonality, the final commonality $Q^i(\{\omega_l^i\})$ for fuzzy evidence structure given in Eq. (5) is the sum of $Q_j^i(\{\omega_l^i\})$ for $j = 1, \cdots, c$, therefore

$$Q^i(\{\omega_l^i\}) = \alpha_i + \sum_{j=1}^{c} \beta_{jl}^i (1-\alpha_i) m(A_j^i) \delta\left(\omega_l^i \in A_j^{0i}\right)$$

Applying the Dempster's rule to combine all the $n$ sources as well as the prior source $E_0$ produces the final combined commonalities [6]:

$$Q(\{\omega_l\}) = K \prod_{i=0}^{n} Q^i(\{\omega_l\}).$$

$K$ is a normalization factor, a constant that is determined during the combination, and it is independent of $\omega_l$. $Q^0(\{\omega_l\}) = p^{-(n-1)}(\omega_l)$ where $p(\omega_l)$ is the prior probability of $\omega_l$ due to Property 1. Final decision is made using the maximum commonality decision rule (MCD).

Obviously, the PDES associated adaptive fuzzy evidential reasoning is parallel to the Bayesian approach if the above resultant commonality is compared with the joint posteriori probability in the Bayesian approach. It is trivial to see that if all fuzzy focals are crisp sets and there is no local discounting and global discounting employed, the normalized commonalities are just the posteriori probabilities resulted from the Bayesian approach. The proposed FDS differs from the Bayesian approach in that fuzzy evidence has been exploited in two adaptive manners, for achieving improved data fusion performance.

Table 1: Summary of examined methods

| ID | Method | Description |
|---|---|---|
| 1 | KT1 | K-mean clustering applied to T1 image |
| 2 | KT2 | K-mean clustering applied to T2 image |
| 3 | KPD | K-mean clustering applied to PD image |
| 4 | KVec | K-mean clustering applied to a vector image composed of T1, T2, and PD |
| 5 | MV | Majority voting applied to results of KT1, KT2, and KPD |
| 6 | Fuz-min | MMD applied to fuzzy min intersection of fuzzy sets derived from three modalities |
| 7 | Bayes | MAP applied to joint posteriori probabilities by the Bayes rule |
| 8 | FDS1 | FDS reasoning with local discounting, decision making by MCD |
| 9 | FDS2 | FDS reasoning with both local and global discounting, decision making by MCD |

## 4 Experimental results

The performance of the adaptive fuzzy evidential reasoning is evaluated for brain tissue segmentation based on the Simulated Brain Database (SBD), a benchmark developed by the McConnell Brain Imaging Centre [11]. In SBD, T1, T2, PD weighted images have been already registered, and corresponding ground truth images are available. In this study, two normal MR image data sets are simulated using different parameters, and the two obtained data sets are referred to as NM-SBD1 and NM-SBD2. In NM-SBD1, the slice thickness is 1mm; the noise level is $3\%$; and the level of intensity non-uniformity is set to $20\%$. In NM-SBD2, the slice thickness is 1mm; the noise level is $9\%$; and the level of intensity non-uniformity is set to $40\%$. A large number of slices are examined for these two data sets, and results for 31 MR images from slice 80 to slice 110 in the axial direction are reported here since these slices are located almost in the middle part along the axial axis, and have an appropriate portion for every brain tissue.

Different relevant methods are examined and compared for segmenting (classifying) CSF, GM, and WM. They are summarized in Table 1, where MMD denotes the maximum membership decision rule, MCD the maximum commonality decision rule, and MAP the maximum a posterior principle. The nine methods tabulated in Table 1 can be categorized into two groups, 1) probability independent methods: KT1, KT2, KPD, KVec, MV, and Fuz-min, and probability dependent methods: Bayes, FDS1, and FDS2; or 2) non-fusion methods: the four direct K-mean clustering approaches, and fusion based methods: MV, Fuz-min, Bayes, FDS1, and FDS2. Notice that in addition to the fuzzy min operator, other operators are also examined, such as max, average, and product, for combining the fuzzy evidence of three modalities. Results have suggested that the fuzzy min operator performs best among these four operators. Fuz-min is thus selected as the representative for comparison.

Performances of all examined methods are evaluated using the overall segmentation (classification) accuracy $\mathcal{A}$, which is calculated as the ratio of the number of correctly labelled pixels to the total number of examined pixels of three considered classes (CSF, GM, and WM in ROI of brain). To compare two methods having accuracy $\mathcal{A}_1$ and $\mathcal{A}_2$, error reduction ratio $ERR$ is employed, which is defined by $ERR = ((1-\mathcal{A}_2)-(1-\mathcal{A}_1))/(1-\mathcal{A}_1)$.

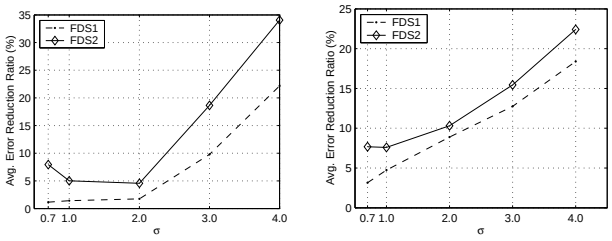
Performances of probability dependent methods rely on probabilities estimated by the Parzen window technique. We have empirically tested a number of values of $\sigma$ in different ranges, and experimental results have shown that the method Bayes produces almost the best overall accuracies at $\sigma = 2.0$ for NM-SBD1 and NM-SBD2. Therefore, to make fair and meaningful comparisons, we set $\sigma$ to 5 values around 2, i.e., $\sigma = 0.7, 1, 2, 3, 4$, and report the corresponding results. $\sigma$ is a parameter used for the determination of the width of Gaussian kernel functions in the Parzen window technique. Segmentation performances are summarized in Table 2, in terms of average overall accuracy $\bar{\mathcal{A}}$, minimum overall accuracy $\min \mathcal{A}$, and maximum overall accuracy $\max \mathcal{A}$ for the 31 slices. In Table 2, the performances of Bayes, FDS1, and FDS2 are the best ones among the examined 5 values of $\sigma$. That is $\sigma = 2$ for NM-SBD1, and $\sigma = 3$ for NM-SBD2. The following observations are obtained from this performance table.

1) When K-mean clustering is applied to single-modalities, KT1 performs best, and the accuracy ranges from $95.04\%$ to $96.88\%$ with average $96.14\%$ for NM-SBD1, and from $78.87\%$ to $82.51\%$ with average $80.70\%$ for NM-SBD2. KPD performs worst. However, KVec outperforms KT1 further, with accuracies from $95.36\%$ to $97.09\%$, and average $96.38\%$ for NM-SBD1, from $80.89\%$ to $85.63\%$ and average $83.76\%$ for NM-SBD2. This suggests that the stacked vector form of three modality pixel intensities benefits the K-mean clustering results, compared to the use of each single-modality.

2) Regarding the performances of the fusion based methods for NM-SBD1, Fuz-min is the worst, and MV is the next worst, both even worse than KT1. However, Bayes, FDS1, and FDS2 at $\sigma = 2$ have very good performances. They are all better than KT1 and KVec, in terms of the minimum, the average, and the maximum segmentation accuracies. The two best methods are FDS2 and FDS1. FDS2 produces overall accuracy from $96.11\%$ to $97.58\%$ with average $97.04\%$, and FDS1 from $95.95\%$ to $97.51\%$ with average $96.95\%$. When NM-SBD2 is examined, the above observations on NM-SBD1 hold except that Fuz-min performs quite well, slightly worse than FDS2 at $\sigma = 3.0$ in terms of the average overall accuracy. This is due to the fact that NM-SBD2 is very noisy and largely variant in intra class pixel intensities. As a result, estimated probabilities by the Parzen window technique are not able to be accurate no matter what values of $\sigma$ are selected, and the performance of Bayes, FDS1, and FDS2 would be thus highly impacted. Differently, method Fuz-min has nothing to do with the estimated probabilities while it performs like a filtering process by taking into account eight neighboring pixels. Anyhow, for NM-SBD2, FDS2 resulted in the highest average overall accuracies $86.33\%$.

NM-SBD1 and NM-SBD2 have opposite natures, in terms

of the reliability of the probabilistic evidence and the fuzzy evidence. In NM-SBD1, the probabilistic evidence is more reliable than the fuzzy evidence while the probabilistic evidence is less reliable than the fuzzy evidence in NM-SBD2. The better performances achieved by FDS1 and FDS2 than other methods on these two data sets would justify the effectiveness of the adaptive fuzzy evidential reasoning. Considering that the Bayesian approach with the MAP decision rule is theoretically optimal in minimizing the error rate, it would not be surprising to see good performance achieved by method Bayes than the K-mean clustering. Further, the contextual information in terms of an overall fuzzy set that is resulted from a FIS, is expected in FDS1 and FDS2 to make up the impreciseness of probabilities through the adaptive FDS schemes for better segmentation performance.




(a) Avg. ERR, NM-SBD1 (b) Avg. ERR, NM-SBD2

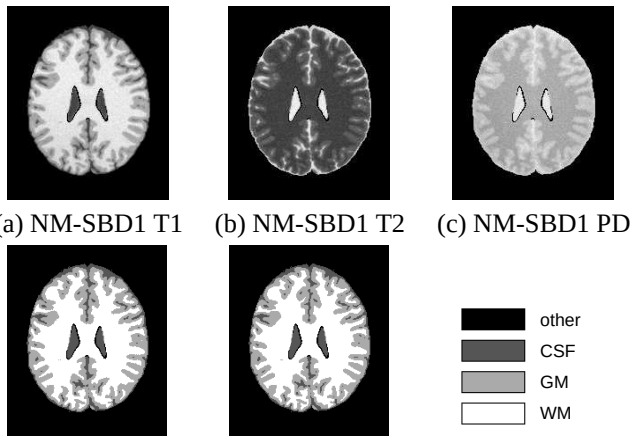
Fig. 2: Comparisons of Bayes, FDS1, and FDS2

In this experiment, we are mainly interested in comparing adaptive fuzzy evidential reasoning with Bayesian. Due to the space constraints, slice-wise performance curves, in terms of segmentation accuracies and error reduction ratios, are not presented for all five settings of $\sigma$. Instead, average overall segmentation accuracies, and average overall ERR's over the 31 slices are respectively summarized in Table 3. For clear comparison between FDS1 and FDS2, their error reduction ratios versus Bayes are plotted in Fig. 2.

First let us take a look at results on NM-SBD1. The best overall accuracies are all obtained at $\sigma = 2.0$ for Bayes, FDS1, and FDS2, which are respectively $96.90\%$, $96.95\%$, and $97.04\%$. Fig. 2 (a) and Table 3 show that the improvement of FDS1 over Bayes is very limited at the three smallest values of $\sigma$: 0.7, 1.0, and 2.0. In fact, FDS1 performs slightly better than Bayes, and the average ERR's are less than $2\%$. Among the 31 slices there are 4, 5, and 6 slices for which FDS1 performs worse than Bayes, respectively at $\sigma = 0.7$, 1.0, and 2.0. The small improvement would not undermine our confidence about the effectiveness of FDS1, since we accept the fact that the Bayesian approach minimizes the error rate ideally, and that $\sigma = 0.7, 1, 2$ can produce accurate enough probability estimations for NM-SBD1. With $\sigma$ increasing, the average ERR of FDS1 vs. Bayes reaches $9.78\%$ and $22.18\%$, respectively at $\sigma = 3.0$ and $\sigma = 4.0$. Furthermore, regarding ERR of FDS2 vs. Bayes, obviously FDS2 produces much improved performance than FDS1 over Bayes, as can be clearly justified by Table 3 and Fig. 2 (a). Even in the best case for Bayes with $\sigma = 2.0$, the average ERR of FDS2 vs. Bayes is $4.57\%$. Average ERR increases with $\sigma$ departing away from $\sigma = 2.0$ when FDS2 is applied to NM-SBD1. This is different from FDS1 in which ERR decreases as $\sigma$ decreases. This is clearly reflected by the shapes of the two curves in Fig. 2 (a). In this sense, FDS2 is more robust and effective than FDS1. Similarly as FDS1, as $\sigma$ increases, ERR increases also, but much larger than that of FDS1. For instance, at $\sigma = 3.0$ and 4.0, average ERR reaches $18.64\%$ and $34.02\%$ due to FDS2 while they are respectively $9.78\%$ and $22.18\%$ for FDS1.

As far as NM-SBD2 is concerned, similar observations as above apply for Bayes, FDS1, and FDS2. However, for the 5 values of $\sigma$, the probabilistic evidence is less reliable than the fuzzy evidence for NM-SBD2. Thus, introducing the fuzzy evidence through FDS1 and FDS2 would definitely produce better performance than Bayes, which takes into account merely the probabilistic evidence. This is exactly verified by Fig. 2 (b), in which bigger error reduction ratios exist in contrast to Bayes for both FDS1 and FDS2 at small values of $\sigma$. From an overall point of view, the larger is $\sigma$, the larger are the improvements achieved by FDS1/FDS2 over Bayes.

One example segmentation result by FDS2 is presented in Fig. 3 (d), together with the three modality images of slice 99 of NM-SBD1. From the viewpoint of human visual inspection, pixels in every tissue type are perfectly segmented.




(a) NM-SBD1 T1 (b) NM-SBD1 T2 (c) NM-SBD1 PD

(d) Classified (f) Ground truth (g) Legend map

Fig. 3: An example set of MR images at slice 99.

## 5 Conclusions

In this paper, we present an adaptive fuzzy evidential reasoning based data fusion scheme, and apply it to automated brain tissue segmentation, based on multi-modality MR images. A novel fuzzy evidence structure model is proposed in the presence of probabilistic evidence and fuzzy evidence; The local discounting and global discounting schemes are developed for adaptive fuzzy evidential reasoning; An efficient implementation scheme is developed.

Experimental results for automated brain tissue segmentation have demonstrated that the adaptive fuzzy evidential reasoning outperforms other methods: K-mean clustering, majority voting, fuzzy set operation, and Bayesian approach. Especially, no matter how accurate the probabilistic evidence is, the adaptive fuzzy evidential reasoning always better performs than the Bayesian approach, in terms of segmentation accuracies and robustness to the variation of probabilistic evidence.

It is worth mentioning that the adaptive fuzzy evidential reasoning scheme can be certainly applied to other applications, as long as information sources can provide the two types of

Table 2: Summary of overall accuracies (%)

| Data set | Method | KT1 | KT2 | KPD | KVec | MV | Fuz-min | Bayes | FDS1 | FDS2 |
|---|---|---|---|---|---|---|---|---|---|---|
| NM-SBD1 | $\min \mathcal{A}$ | 95.04 | 79.69 | 73.56 | 95.36 | 93.91 | 93.16 | 95.87 | 95.95 | 96.11 |
| | $\bar{\mathcal{A}}$ | 96.14 | 87.73 | 80.11 | 96.38 | 95.74 | 94.78 | 96.90 | 96.95 | 97.04 |
| | $\max \mathcal{A}$ | 96.88 | 90.64 | 86.13 | 97.09 | 96.47 | 95.44 | 97.43 | 97.51 | 97.58 |
| NM-SBD2 | $\min \mathcal{A}$ | 78.87 | 67.77 | 49.34 | 80.89 | 78.28 | 84.59 | 81.85 | 83.99 | 84.48 |
| | $\bar{\mathcal{A}}$ | 80.70 | 71.21 | 52.51 | 83.76 | 80.05 | 86.11 | 83.83 | 85.89 | 86.33 |
| | $\max \mathcal{A}$ | 82.51 | 73.52 | 55.56 | 85.63 | 81.40 | 87.73 | 85.44 | 87.09 | 87.61 |

Table 3: Summary of performance (%)

| | | average overall accuracies | | | | | average ERR, FDS vs. Bayes | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Source | Method $\setminus \sigma$ | 0.7 | 1 | 2 | 3 | 4 | 0.7 | 1 | 2 | 3 | 4 |
| NM-SBD1 | Bayes | 96.43 | 96.61 | 96.90 | 95.76 | 92.32 | | | | | |
| NM-SBD1 | FDS1 | 96.47 | 96.66 | 96.95 | 96.18 | 94.10 | 1.15 | 1.41 | 1.74 | 9.78 | 22.18 |
| NM-SBD1 | FDS2 | 96.71 | 96.77 | 97.04 | 96.56 | 95.03 | 7.94 | 5.01 | 4.57 | 18.64 | 34.07 |
| NM-SBD2 | Bayes | 82.42 | 83.01 | 84.12 | 83.83 | 81.91 | | | | | |
| NM-SBD2 | FDS1 | 82.97 | 83.81 | 85.53 | 85.89 | 85.24 | 3.15 | 4.73 | 8.89 | 12.75 | 18.39 |
| NM-SBD2 | FDS2 | 83.78 | 84.30 | 85.75 | 86.33 | 85.97 | 7.68 | 7.59 | 10.31 | 15.45 | 22.42 |

evidence as required in this study. This would be one part of the future work.

## References


[1] Y. Hata, S. Kobashi, and S. Hirano. Automated segmentation of human brain MR images aided by fuzzy information granulation and fuzzy inference. *IEEE Trans. SMC, Part C*, 30: 381-395, 2000.

[2] D. Goldberg-Zimring, A. Achiron, and S. Miron. Automated detection and characterization of multiple sclerosis lesions in brain MR images. *Magnetic Resonance Imaging*, 16: 311-318, 1998.

[3] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based tissue classification of MR images of the brain. *IEEE Trans. Medical Imaging*, 18: 897-908, 1999.

[4] Y. Wang, T. Adali, J. Xuan, and Z. Szabo. Magnetic resonance image analysis by information theoretic criteria and stochastic models. *IEEE Trans. Infomation Technology in Biomedicine*, 5:150-158, 2001.

[5] H. Zhu and O. Basir. Fuzzy evidential reasoning applied to multimodality MR brain image segmentation with source selection. In *Proc. 7th Joint Conf. Info. Sci.*, 2003.

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

[7] J. Yen. Generalizing the Dempster-Shafer theory to fuzzy sets. *IEEE Trans. SMC*, 20: 559-570, 1990.

[8] R.R. Yager. Generallized probabilities of fuzzy events from fuzzy belief structures. *Inform. Sci.*, 28:45-62, 1982.

[9] R.R. Yager and D.P. Filev. Including probabilistic uncertainty in fuzzy logic controller modelling using Dempster-Shafer theory. *IEEE Trans. SMC,*, 25: 1221-1230, 1995.

[10] R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern Classification*, John Wiley & sons, Chichester, New York, 2001.

[11] D.L. Collins, A.P. Zijdenbos, V. Kollokian, J.G. Sled, N.J. Kabani, C.J. Holmes, and A.C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Trans. Medical Imaging*, 17: 463-468, 1998.

[12] P. Smets. Data fusion in transferable belief model. In *Proc. 3rd Int. Conf. Information Fusion*, pages 21-33, 2000.

[13] J.W. Guan and D.A. Bell. Approximate reasoning and evidence theory. *Information Sciences*, 96:207-235, 1997.

[14] R.R. Yager. On the normalization of fuzzy belief structures, *Int. J. Approximate Reasoning*, 14: 127-153, 1996.

[15] L.M. Zouhal, and T. Denoeux. Generalizing the evidence-theoretic K-NN rule to fuzzy pattern recognition, In *Proc. 2nd Int. ICSC Symposium on Fuzzy Logic and Applications*, pages 294-300, 1997.

[16] H. Zhu and O. Basir. A scheme for constructing evidence structures in Dempster-Shafer evidence theory for data fusion. In *Proc. 5th IEEE Int. Symposium on Computational Intelligence in Robotics and Automation*, pages 16-20, 2003.

[17] D. Dubois and H. Prade. On several representations of an uncertainty body of evidence. In M. M. Gupta, E. Sanchz, editors, *Fuzzy Informaton and Decision Processes*, pages 167-181, North-Holland, New-York, 1982.

[18] H. Zhu and O. Basir. Automated brain tissue segmentation and MS lesion detection using fuzzy and evidential reasoning. In *Pro. 10th IEEE Int. Conf. Electronics, Circuits and Systems*, 2003.

[19] X. Shang and W. Jiang. A note on fuzzy information measures. *Pattern Recognition Letters*, 18: 425-432, 1997.

[20] A. De Luca and S. Termini. A definition of non probabilistic entropy in the setting of fuzzy set theory. *Information and Control*, 20: 301-312, 1972.