The effective use of the DSmT for multi-class classification

Nassim Abbas [a, *], Youcef Chibani [a], Arnaud Martin [b] and Florentin Smarandache [c]

[a] Faculty of Electronics and Computer Science, University of Science and Technology Houari Boumediene, Algeria

[b] IUT of Lannion, Rennes 1 University, France

[c] Department of Mathematics, University of New Mexico, USA

Abstract

The extension of the Dezert-Smarandache theory (DSmT) for the multi-class framework has a feasible computational complexity for various applications when the number of classes is limited or reduced typically two classes. In contrast, when the number of classes is large, the DSmT generates a high computational complexity. This paper proposes to investigate the effective use of the DSmT for multi-class classification in conjunction with the Support Vector Machines using the One-Against-All (OAA) implementation, which allows offering two advantages: firstly, it allows modeling the partial ignorance by including the complementary classes in the set of focal elements during the combination process and, secondly, it allows reducing drastically the number of focal elements using a supervised model by introducing exclusive constraints when classes are naturally and mutually exclusive. To illustrate the effective use of the DSmT for multi-class classification, two SVM-OAA implementations are combined according three steps: transformation of the SVM classifier outputs into posterior probabilities using a sigmoid technique of Platt, estimation of masses directly through the proposed model and combination of masses through the Proportional Conflict Redistribution (PCR6). To prove the effective use of the proposed framework, a case study is conducted on the handwritten digit recognition. Experimental results show that it is possible to reduce efficiently both the number of focal elements and the classification error rate.

Keywords: *Handwriting digit recognition; Support Vector Machines; Dezert-Smarandache theory; Belief assignments; Conflict management*

1. Introduction

Nowadays a large number of classifiers and methods of generating features is developed in various application areas of pattern recognition [Duda, 2001], [Cher, 2007]. Nevertheless, it failed to underline the incontestable superiority of a method over another in both steps of generating features and classification. Rather than trying to optimize a single classifier by choosing the best features for a given problem, researchers found more interesting to combine the recognition methods [Rahm, 2003], [Cher, 2007]. Indeed, the combination of classifiers allows exploiting the redundant and complementary nature of the responses issued from different classifiers.

Researchers have proposed various approaches for combining classifiers increasingly numerous and varied, which led the development of several schemes in order to treat data in different ways [Rahm, 2003], [Cher, 2007]. Generally, three approaches for combining classifiers can be considered: parallel approach, sequential approach and hybrid approach [Cher, 2007]. Furthermore, these ones can be performed at a class level, at a rank level, or at a measure level [Xu, 1992], [Jain, 2000], [Ruta, 2000], [Abba, 2012c].

In many applications, various constraints do not allow an efficient joint use of classifiers and feature generation methods leading to an inaccurate performance. Thus, an appropriate operating method using mathematical approaches is needed, which takes into account two notions: uncertainty and imprecision of the responses of classifiers. In general, the most theoretical advances which have been devoted to the theory of probabilities are able to represent the uncertain knowledge but are unable to model easily the information which is imprecise, incomplete, or not totally reliable. Moreover, they often lead to confuse both concepts of uncertainty and imprecision with the probability measure. Therefore, new original theories dealing with uncertainty and imprecise information have been introduced, such as the fuzzy set theory [Zade, 1968], evidence theory [Dem, 1967], [Shaf, 1976], possibility theory [Dubo, 1988] and, more recently, the theory of plausible and paradoxical reasoning [Smar, 2004], [Smar, 2006], [Smar, 2009].

The evidence theory initiated by Dempster and Shafer termed as Dempster-Shafer theory (DST) [Demp, 1967], [Shaf, 1976] is generally recognized as a convenient and flexible alternative to the bayesian theory of subjective probability [Shaf, 1990]. The DST is a powerful theoretical tool which has been applied in many kinds of applications [Smet, 1999] for the representation of incomplete knowledge, belief updating and for the combination of evidence [Prov, 1992], [Dubo, 1992] through the Dempster-Shafer's combination rule. Indeed, it offers a simple and direct representation of ignorance and has a low computational complexity [Rusp, 1992] for most practical applications.

Nevertheless, this theory presents some weaknesses and limitations mainly when the combined evidence sources become very conflicting. Furthermore, the Shafer's model itself does not allow necessary holding in some fusion problems involving the existence of the paradoxical information. To overcome these limitations, a recent theory of plausible and paradoxical reasoning, known as Dezert-Smarandache theory (DSmT) in the literature, was elaborate by Jean Dezert and Florentin Smarandache for dealing with imprecise, uncertain and paradoxical sources of information. Thus, the main objective of the DSmT was to introduce combination rules that would allow to correctly combining evidences issued from different information sources, even in presence of conflicts between sources or in presence of constraints corresponding to an appropriate model (free or hybrid DSm models [Smar, 2004]). The DSmT has proved its efficiency in many current pattern recognition application areas such as remote sensing [Corg, 2003], [Maup, 2004], [Elha, 2011], [Zhun, 2012], identification and tracking [Pann, 2008], [Pann, 2009], [Kech, 2009], [Sun, 2010], [Deze, 2010], [Pann, 2011], biometrics [Sing, 2008], [Vats, 2009a], [Vats, 2009b], [Vats, 2010], computer vision [Garc, 2008], [Khod, 2010], [Deze, 2011], robotics [Huan, 2006], [Li, 2006a], [Li, 2006b], [Li, 2007], [Li, 2008], [Huan, 2009] and more recently handwritten recognition applications [Abba, 2012a], [Abba, 2012b], [Abba, 2012c] as well as many others [Smar, 2004], [Smar, 2006], [Smar, 2009].

The use of the DSmT for multi-class classification has a feasible computational complexity for various applications when the number of classes is limited or reduced typically two classes [Abba, 2012a]. In contrast, when the number of classes is large, the DSmT generates a high computational complexity closely related to the number of elements to be processed. Indeed, an analytical expression defined by Tombak *et al.* [Tomb, 2001] shows that the number of elements to be processed follows the sequence of Dedekind's numbers [Dede, 1897], [Comt, 1974]: 1,2,5,19,167,7580,7828353,... For instance, if the number of classes belonging to discernment space is 8, then the number of elements to be deal in DSmT framework is $\approx 5.6 \times 10^{22}$. Hence, it is not easy to consider the set of all subsets of the original classes (but under the union and the intersection operators) since it becomes untractable for more than 6 elements in the discernment space [Deze, 2004a]. Thus, Dezert and Smarandache [Deze, 2004b] proposed a first work for ordering all elements generated using the free DSm model for matrix

calculus such as made in DST framework [Kenn, 1992], [Smet, 2002]. However, this proposition has limitations since in practical applications it is more appropriate to only manipulate the focal elements [Deno, 2001], [Djik, 2004], [Mart, 2006], [Abba, 2012c].

Hence, few works have already been focused on the computational complexity of the combination algorithms formulated in DSmT framework. Djiknavorian and Grenier [Djik, 2004] showed that there's a way to avoid the high level of complexity of DSm hybrid (DSmH) combination algorithm by designing a such code that can perform a complete DSmH combination in very short period of time. However, even if they have obtained an optimal process of evaluating DSmH algorithm, first some parts of their code are really not optimized and second it has been developed only for a dynamic fusion. Martin [Mart, 2009] further proposed a practical codification of the focal elements which gives only one integer number to each part of the Venn diagram representing the discernment space. Contrary to the Smarandache's codification [Deze, 2004a] used in [Deze, 2004c] and the proposed codes in [Djik, 2004], author thinks that the constraints given by the application must be integrated directly in the codification of the focal elements for getting a reduced discernment space. Therefore, this codification can drastically reduce the number of possible focal elements and so the complexity of the DST as well as the DSmT frameworks. A disadvantage of this codification is that the complexity increases drastically with the number of combined sources especially when dealing with a problem in the multi-class framework. To address this issue, Li *et al.* [Li, 2011] proposed a criterion called evidence supporting measure of similarity (ESMS), which consists in selecting, among all sources available, only a subset of sources of evidence in order to reduce the complexity of the combination process. However, this criterion has been justified for only a two-class problem. Nowadays, the complexity of reducing both the number of combined sources and the size of the discernment space are research challenges that still need to be addressed.

In many pattern recognition applications, the classes belonging to the discernment space are naturally and then mutually exclusive such as in biometrics [Sing, 2008], [Vats, 2009a], [Vats, 2009b], [Vats, 2010] and handwritten recognition applications [Abba, 2012a], [Abba, 2012b], [Abba, 2012c]. Hence, several classification methods have been proposed as template matching techniques [Deng, 1999], [Fang, 2003], [Guo, 2001], minimum distance classifiers [Fang, 2001], [Sabo, 1997], support vector machine (SVM) [Just, 2005], hidden Markov Models (HMMs) [Just, 2001], [Just, 2005], [Coet, 2004], neural networks [Kaew, 1999], [Quek, 2002]. In various pattern recognition applications, the SVMs have proved their performance from the mid-1990s comparatively to other classifiers [Cher, 2007]. The SVM is based on an optimization approach in order to separate two classes by an hyperplane. In the context of multi-class classification, this optimization approach is possible [Weston, 1998] but requiring a very costly duration. Hence, two preferable methods of multi-class implementation of SVMs have been proposed for combining several binary SVMs, , which are One Against All (OAA) and One Against One (OAO), respectively [West, 1998], [Guer, 1999], [Hsu, 2001]. The former is the most commonly used implementation in the context of multi-class classification using binary SVMs, which constructs n SVMs to solve a n-class problem [Bott, 1994]. Each SVM is designed to separate a simple class θ_i from all the others, i.e., from the corresponding complementary class $\overline{\theta_i} = \bigcup_{\substack{0 \le j \le n-1 \\ j \ne i}} \theta_j$. In contrast, the OAO implementation is designed to separate two simple classes θ_i and

θ_j ($i \neq j$), which requires $n \times (n-1)/2$ SVMs. Hence, various decision functions can be used such as the Decision Directed Acyclic Graph (DDAG) [Huan, 2002] since it has the advantage to eliminate all possible unclassifiable data.

Generally, the combination of binary classifiers is performed through very simple approaches such as voting rule or a maximization of decision function coming from the classifiers. In this context, many combination operators can be used, especially in the DST framework [Martin, 2007]. Still in the same vein, some works have been tried out the combination of binary classifier originally from SVM in the DST framework [Aregui, 2007], [Quost, 2007a]. For instance, the pairwise approach has been revisited by Quest *et al.* [Quost, 2005], [Quost, 2006], [Quost, 2007a], [Quost, 2007b] in the framework of the DST of belief functions for solving a multi-class problem. In [Hu, 2005], the combination method based on DST has been used by Hu *et al.* for combining multiple multi-class probability SVM classifiers in order to deal with distributed multi-source multi-class problem [Hu, 2005]. Martin and Quidu proposed an original approach based on DST [Martin, 2008] for combining binary SVM classifiers using OAO or OAA strategies, which provides a decision support helping experts for seabed characterization from sonar images. Burger *et al.* [Burger, 2006] proposed to apply a belief-based method for SVM fusion to hand shape recognition. Optimizing the fusion of the sub-classifications and dealing with undetermined cases due to uncertainty and doubt have been investigated by other works [Burger, 2008], through a simple method, which combines the fusion methods of belief theories with SVMs. Recently, one regression based approach [Laanay, 2010] has been proposed to predict membership or belief functions, which are able to model correctly uncertainty and imprecision of data.

In this work, we propose to investigate the effective use of the DSmT for multi-class classification in conjunction with the SVM-OAA implementation, which allows offering two advantages: firstly, it allows modeling the partial ignorance by including the complementary classes in the set of focal elements, and then in the combination process, contrary to the OAO implementation which takes into account only the singletons, and secondly, it allows reducing drastically the number of focal elements from $Dedekind(n)$ to $2 \times n$. The reduction is performed through a supervised model using exclusive constraints. Combining the outputs of SVMs within DSmT framework requires that the outputs of SVMs must be transformed into membership degree. Hence, several methods of estimating of mass functions are proposed in both DST and DSmT frameworks, these ones can be directly explicit through special functions or indirectly explicit through transfer models [Demp, 1967], [Deno, 1997], [Smet, 1994], [Dubo, 1994], [Appr, 1999]. In our case, we propose a direct estimation method based on a sigmoid transformation of Platt [Plat, 1999]. This allows us to satisfy the OAA implementation constraint.

The paper is organized as follows. Section 2 reviews the Proportional Conflict Redistribution (PCR6) rule based on DSmT. Section 3 describes the combination methodology for multi-class classification using the SVM-OAA implementation. Experiments conducted on the dataset of the isolated handwritten digits are presented in section 4. The last section gives a summary of the proposed combination framework and looks to the future research direction.

2. Review of PCR6 combination rule

In pattern recognition, the multi-class classification problem is generally formulated as a n-class problem where classes are associated to patterns classes, namely $\theta_0, \theta_1, \ldots,$ and θ_n. Hence, the parallel combination of two classifiers, namely

information sources S_1 and S_2, respectively, is performed through the PCR6 combination rule based on the DSmT. For n-class problem, a reference domain also called the discernment space should be defined for performing the combination, which is composed of a finite set of exhaustive and mutually exclusive hypotheses.

In the context of the probabilistic theory, the discernment space, namely Θ, is composed of n elements as: $\Theta = \{\theta_0, \theta_1, \ldots, \theta_n\}$, and a mapping function $m \in [0, 1]$ is associated for each class, which defines the corresponding mass verifying $m(\varnothing) = 0$ and $\sum_{i=0}^{n} m(\theta_i) = 1$. In Bayesian framework, combining two sources of information by means of the weighted mean and consensus based rules seems effective for non-conflicting responses <mark>[Bloc, 2003]</mark>, <mark>[Fren, 1985]</mark>, <mark>[Cook, 1988]</mark>, <mark>[Cook, 1991]</mark>. In the opposite case, an alternative approach has been developed in DSmT framework to deal with (highly) conflicting imprecise and uncertain sources of information <mark>[Smar, 2009]</mark>. Example of such approaches is PCR6 rule.

The main concept of the DSmT is to distribute unitary mass of certainty over all the composite propositions built from elements of Θ with $\cup$ (Union) and $\cap$ (Intersection) operators instead of making this distribution over the elementary hypothesis only. Therefore, the *hyper-powerset* D^Θ is defined as:

1. $\varnothing, \theta_0, \theta_1, \ldots, \theta_n \in D^\Theta$.

2. If $A, B \in D^\Theta$, then $A \cap B \in D^\Theta$ and $A \cup B \in D^\Theta$.

3. No other elements belong to D^Θ, except those obtained by using rules 1 or 2.

The DSmT uses generalized basic belief mass, also known as the generalized basic belief assignment (gbba) computed on hyper-powerset of Θ and defined by a map $m(.): D^\Theta \to [0, 1]$ associated to a given source of evidence which can support paradoxical information, as follows: $m(\varnothing) = 0$ and $\sum_{A \in D^\Theta} m(A) = 1$. The combined masses m_{PCR6} obtained from $m_1(.)$ and $m_2(.)$ by means of the PCR6 rule <mark>[Smar, 2006]</mark>, <mark>[Smar, 2009]</mark> are defined as:

$$m_{PCR6}(A_i) = \begin{cases} 0 & \text{if } A_i \in \Phi, \\ m_\wedge(A_i) + \sum_{k=1}^{2} m_k^2(A_i) L_k & \text{otherwise.} \end{cases} \tag{1}$$

Where

$$L_k = \sum_{\substack{Y_{\sigma_k(1)} \cap A_i \in \Phi \\ Y_{\sigma_k(1)} \in D^\Theta}} \frac{m_{\sigma_k(1)} Y_{\sigma_k(1)}}{m_k(A_i) + m_{\sigma_k(1)} Y_{\sigma_k(1)}} \tag{2}$$

$\Phi = \{\Phi_M, \emptyset\}$ is the set of all relatively and absolutely empty elements, Φ_M is the set of all elements of D^Θ which have been forced to be empty in the hybrid model M defined by the exhaustive and exclusive constraints, $\emptyset$ is the empty set, the denominator $m_k(A_i) + m_{\sigma_k(1)} Y_{\sigma_k(1)}$ is different to zero, and where $\sigma_k(1)$ counts from 1 to 2 avoiding k, i.e.:

$$\sigma_k(1) = \begin{cases} 2 & \text{if } k = 1, \\ 1 & \text{if } k = 2. \end{cases} \tag{3}$$

Thus, the term $m_\wedge(A_i)$ represents a conjunctive consensus, also called DSm Classic (DSmC) combination rule <mark>[Smar, 2006]</mark>, <mark>[Smar, 2009]</mark>, which is defined as:

$$m_\wedge(A_i) = \begin{cases} 0 & \text{if } A_i = \emptyset, \\ \sum_{(X,Y \in D^\Theta, X \cap Y = A_i)} m_1(X) m_2(Y) & \text{otherwise.} \end{cases} \tag{4}$$

3. Methodology

The proposed combination methodology shown in Fig. 1 is composed of two individual systems using SVMs classifiers. Each one is trained using its own source of information providing two kinds of complementary features, which are combined through the PCR6 rule. In the following, we give a description of each module composed our system.
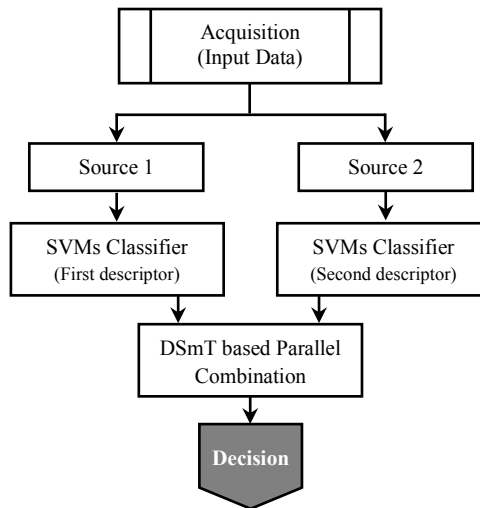

Figure 1. Structure of the combination scheme using SVM and DSmT

3.1. Classification based on SVM

The classification based on SVMs has been used widely in many pattern recognition applications as the handwritten digit recognition <mark>[Cher, 2007]</mark>. The SVM is a learning method introduced by Vapnik *et al*. <mark>[Vapn, 1995]</mark>, which tries to find an optimal hyperplane for separating two classes. Its concept is based on the maximization of the distance of two points

belonging each one to a class. Therefore, the misclassification error of data both in the training set and test set is minimized. Basically, SVMs have been defined for separating linearly two classes. When data are non linearly separable, a kernel function K is used. Thus, all mathematical functions, which satisfy Mercer's conditions, are eligible to be a SVM-kernel [Vapn, 1995]. Examples of such kernels are sigmoid kernel, polynomial kernel, and Radial Basis Function (RBF) kernel. Then, the decision function $f : \mathrm{R}^p \to \{-1, +1\}$, is expressed in terms of kernel expansion as:

$$f(x) = \sum_{k=1}^{Sv} \alpha_k y_k K(x, x_k) + b \tag{5}$$

where α_k are Lagrange multipliers, Sv is the number of support vectors x_k which are training data, such that $0 \le \alpha_k \le C$, C is a user-defined parameter that controls the tradeoff between the machine complexity and the number of nonseparable points [Huan, 2002], the bias b is a scalar computed by using any support vector.

Finally, for a two-class problem, test data are classified according to:

$$x \in \begin{cases} class\ (+1) & \text{if } f(x) \succ 0 \\ class\ (-1) & \text{otherwise} \end{cases} \tag{6}$$

The extension of the SVM for multi-class classification is performed according the One Against-All (OAA) [Cort, 1995]. Let a set of N training samples which are separable in n classes $\{\theta_0, \theta_1, \ldots, \theta_{n-1}\}$, such that $\{(x_k^i, y_k^i) \in \mathrm{R}^p \times \{\pm 1\}; k = 1, .., N; i = 1, . n\}$. The principle consists to separate a class from other classes. Consequently, n SVMs are required for solving n class problem.

3.2. Classification Based On DSmT

The proposed classification based on DSmT is presented in Fig. 2, which is conducted into three steps: i) estimation of masses, ii) combination of masses through the PCR6 combination rule and iii) decision rule.
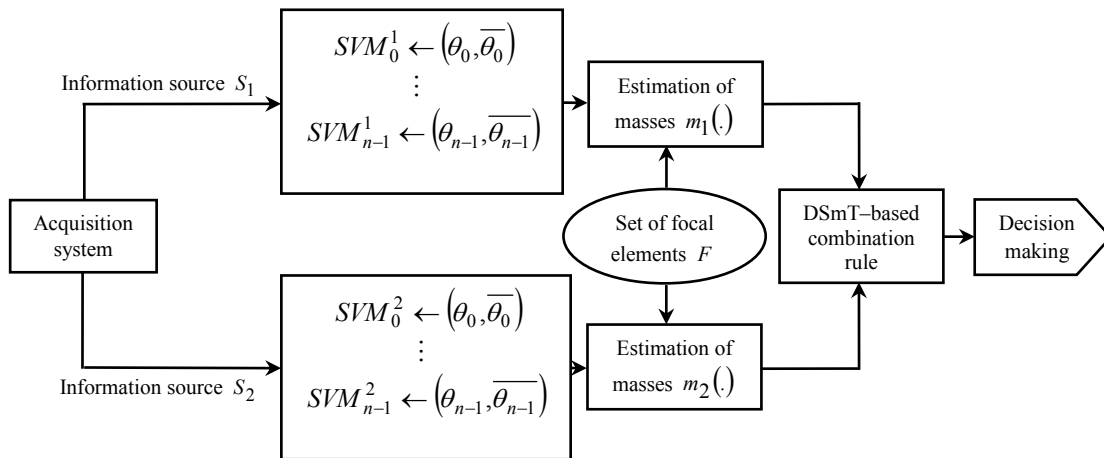

Figure 2. DSmT-based parallel combination for multi-class classification

7

3.2.1. *Estimation of Masses*

The difficulty of estimating masses is increased if one assigns weights to the composed classes [Lowr, 1991]. Therefore, transfer models of the mass function have been proposed whose the aim is to distribute the initial masses on the simple and compound classes associated to each source. Thus, the estimation of masses is performed into two steps: i) assignment of membership degrees for each simple class through a sigmoid transformation proposed by Platt [Plat, 1999], ii) estimation of masses of simple classes and their complementary classes using a supervised model, respectively.

- *Calibration of the SVM outputs:* Although, standard SVM is very discriminative classifier, its output values are not calibrated for appropriately combining two sources of information. Hence, an interesting alternative is proposed in [Plat, 1999] to transform the SVM outputs into posterior probabilities. Thus, given a training set of instance-label pairs $\{(x_k, y_k), k = 1, \ldots, N\}$, where $x_k \in \mathrm{R}^p$ and $y_k \in \{-1, +1\}$, the unthresholded output of an SVM is a distance measure between a test pattern and the decision boundary as given in (5). Furthermore, there is no clear relationship with the posterior class probability $P(y = +1|x)$ that the pattern x belongs to the class $y = +1$. A possible estimation for this probability can be obtained [Plat, 1999] by modeling the distributions $P(f|y = +1)$ and $P(f|y = -1)$ of the SVM output $f(x)$ using Gaussian distribution of equal variance and then compute the probability of the class given the output by using Bayes' rule. This yields a sigmoid allowing to estimate probabilities:

$$\hat{P}(y = +1|x) = \frac{1}{1 + \exp(A \times f(x) + B)} \tag{7}$$

Parameters A and B are tuned by minimizing the negative log-likelihood of the training data:

$$-\sum_{k=1}^{N} t_k \log(Q_k) + (1 - t_k) \log(1 - Q_k) \tag{8}$$

where $Q_k = \hat{P}(y_k = 1|x)$ and $t_k = \dfrac{y_k + 1}{2}$ denotes the probability target.

- *Supervised Model:* Denoting $m_1(.)$ and $m_2(.)$ the gbba provided by two distinct information sources S_1 (First descriptor) and S_2 (Second descriptor), F is the set of focal elements for each source, such that $F = \{\theta_0, \theta_1, \ldots, \theta_{n-1}, \overline{\theta_0}, \overline{\theta_1}, \ldots, \overline{\theta_{n-1}}\}$, the classes θ_i are separable (One relatively to its complementary class $\overline{\theta_i}$) using the SVM-OAA multi-class implementation corresponding to different singletons of the patterns assumed to be known. Therefore, each compound element $A_i \notin F$ has a mass m_1 equal to zero, on the other hand, the mass of the complementary element $\overline{\theta_i} = \bigcup_{\substack{0 \leq j \leq n-1 \\ j \neq i}} \theta_j$ is

different from zero, which represents the mass of the partial ignorance. The same reasoning is applied to the classes issued from the second source S_2 and $m_2(.)$. Hence, both gbba $m_1(.)$ and $m_2(.)$ are given as follows:

$$m_b(\theta_i) = \frac{\hat{P}_b(\theta_i|x)}{Z_b}, \forall \theta_i \in F \tag{9}$$

$$m_b(\bar{\theta}_i) = \frac{\sum_{\substack{j=0 \\ j \neq i}}^{n-1} \hat{P}_b(\theta_j|x)}{Z_b}, \forall \bar{\theta}_i \in F \tag{10}$$

$$m_b(A_i) = 0, \forall A_i \in \Phi = D^\Theta \setminus F \tag{11}$$

where $Z_b = \sum_{j=0}^{n-1} \hat{P}_b(\theta_j|x)$ represent the normalization factors introduced in the axiomatic approach in order to respect the mass definition, $\hat{P}_b$ are the posterior probabilities issued from the first source $(b=1)$ and the second source $(b=2)$, respectively. They are given for a test pattern x as follows:

$$\hat{P}_b(\theta_i|x) = \frac{1}{1 + \exp(A_{ib} \times f_{ib}(x) + B_{ib})}. \tag{12}$$

A_{ib} and B_{ib} are the parameters of the sigmoid function tuned by minimizing the negative log-likelihood during training for each class of patterns θ_i, and $f_{ib}(x)$ is the i-th output of binary SVM classifier SVM_i^b issued from the source S_b, such that $i = 0, 1, \ldots, n-1$ and $b \in \{1, 2\}$.

In summary, the masses of all elements $A_j \in D^\Theta$ allocated by each information source S_b $(b=1,2)$ are obtained according the following steps:

1. Define a frame of discernment $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$.
2. Classify a pattern x through the SVM-OAA implementation.
3. Transform each SVM output to the posteriori probability using Eq. (12).
4. Compute the masses associated to each class and its complementary using Eq. (9) and Eq. (10), respectively.

3.2.2. Combination of masses

In order to manage the conflict generated from the two information sources S_1 and S_2 (i.e. both SVM classifications), the combined masses are computed as follows:

$$m_c = m_1 \oplus m_2 \tag{13}$$

where $\oplus$ defines the PCR6 combination rule as given in (1). Hence, in the context of some application of pattern recognition area, such as handwritten digit recognition, we take as constraints the propositions ($\theta_i \cap \theta_j = \varnothing$, $\forall \theta_i, \theta_j \in \Theta$), such that $i \neq j$, which allow separating between each two classes belonging to Θ.

Therefore, the hyper power set D^Θ is reduced to the set F as $F = \left\{ \theta_0, \theta_1, \ldots, \theta_{n-1}, \overline{\theta_0}, \overline{\theta_1}, \ldots, \overline{\theta_{n-1}} \right\}$, which defines a particular case of the Shafer's model. Thus, the conflict $K_c \left(\in [0,1] \right)$ measured between two sources is defined as:

$$K_c = \sum_{\substack{A_k, A_l \in F \\ A_k \cap A_l \in \Phi}} m_1(A_k) \times m_2(A_l) \tag{14}$$

where $\Phi = D^\Theta \setminus F$ is the set of all relatively and absolutely empty elements, $m_1(.)$ and $m_2(.)$ represent the corresponding generalized basic belief assignments provided by two information sources S_1 and S_2, respectively.

3.2.3. Decision rule

A membership decision of a pattern to one of the simple classes of Θ is performed using the statistical classification technique. First, the combined beliefs are converted into probability measure using a new probabilistic transformation, called Dezert-Smarandache probability (DSmP), that maps a belief measure to a subjective probability measure <mark>[Smar, 2009]</mark> defined as:

$$DSmP_\varepsilon(\theta_i) = m_c(\theta_i) + \left(m_c(\theta_i) + \varepsilon \right) \sum_{\substack{A_j \in 2^\Theta \\ A_j \supset \theta_i \\ C_M(A_j) \geq 2}} \frac{m_c(A_j)}{\displaystyle\sum_{\substack{A_k \in 2^\Theta \\ A_k \subset A_j \\ C_M(A_k)=1}} m_c(A_k) + \varepsilon \, C_M(A_j)} \tag{15}$$

where $i = \{0, 1, \ldots, 9\}$, $\varepsilon \geq 0$ is a tuning parameter, M is the Shafer's model for Θ, and $C_M(A_k)$ denotes the DSm cardinal of A_k <mark>[Smar, 2004]</mark>. Therefore, the maximum likelihood (ML) test is used for decision making as follows:

$$x \in \theta_i \text{ if } DSmP_\varepsilon(\theta_i) = \max \left\{ DSmP_\varepsilon(\theta_j), 0 \leq j \leq 9 \right\} \tag{16}$$

where x is the pattern test characterized by both descriptors, which are used during the feature generation step, and ε is fixed to 0.001 in the decision measure given by (15).

4. Experimental results

4.1. Database description and performance evaluation

For evaluating the effective use of the DSmT for multi-class classification, we consider a case study conducted on the handwriting digit recognition application. For this, we select a well-known US Postal Service (USPS) database that contains normalized grey-level handwritten digit images of 10 numeral classes, extracted from US postal envelopes. All images are segmented and normalized to a size of 16×16 pixels. There are 7291 training data and 2007 test data where some of them are corrupted and difficult to classify correctly (Fig. 3). The partition of the databse for each class according tranining and testing is reported in table 1.
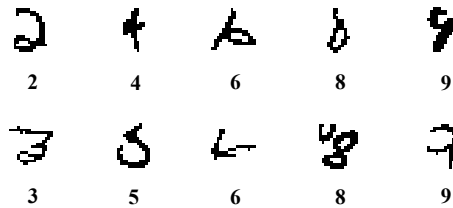

Figure 3. Some samples with their alleged classes from USPS database.

Table 1. Partitioning of the USPS dataset

Classes	0	1	2	3	4	5	6	7	8	9
Training	1194	1005	731	658	652	556	664	645	542	644
Testing	359	264	198	166	200	160	170	147	166	177

For evaluating performances of the handwritten digit classification, a popular error is considered, which is the *Error Rate per Class (ERC)* and *Mean Error Rate (MER)* for all classes. Both errors are expressed in %.

3.2. Pre-processing

The acquired image of isolated digit should be processed to facilitate the feature generation. In our case, the pre-processing module includes a binarization step using the method of Otsu [Otsu, 1979], which eliminates the homogeneous background of the isolated digit and keeps the foreground information. Thus, we use the processed digit without unifying size image for recognition process.

3.3. Feature Generation

The objective of the feature generation step is to underline the relevant information that initially exists in the raw data. Thus, an appropriate choice of the descriptor improves significantly the accuracy of the recognition system. In this study, we use a collection of popular feature generation methods, which can be categorized into background features [Brit, 2004], [Cava, 2006], foreground features [Brit, 2004], [Cava, 2006], geometric features [Cher, 2007], and uniform grid features [Fava, 1996], [Abba, 2011].

3.4. Validation of SVM Models

The SVM model is produced for each class according the used descriptor. Hence, the training dataset is partitioned into two equal subsets of samples, which are used for training and validating each binary SVM, respectively. Thus, the validation phase allows finding the optimal hyperparameters for the ten SVM models. In our case, the RBF kernel is selected for the experiments. Furthermore, both the regularization and RBF kernel parameters (C,σ) of each SVM are tuned experimentally during the training phase in such way that the misclassification error of data in the training subset is zero and the validation test gives a minimal error during validation phase for each SVM separating between a simple class and its complementary class.

Table 2 shows an example of the optimal parameters, which are obtained during both training and validation phases by using the UG-SVMs classifier. The parameters n and m define the number of the lines (vertical regions) and columns (horizontal regions) of the grid, respectively, which have been optimized during the validation phase for each SVM model. Therefore, these all parameters are used afterwards during the testing phase. *ERCs* and *ERCc* are the Error Rates per Class for simple and complementary classes, respectively. As we can see, the choice of the optimal size of the uniform grid and hyperparameters of each SVM should be tuned carefully in order to produce a reduced error.

Table 2. Optimal parameters of the UG-SVMs classifier

Parameters	SVM Classifier									
	0	1	2	3	4	5	6	7	8	9
n	7	2	8	5	4	7	7	8	8	7
m	5	3	3	6	12	5	8	6	6	10
σ	3.5	1	3.5	4	3	3.5	4	3.5	5	4.5
C	5	3	4	5	4	4	2	4	3	5
ERCs (%)	2.0	1.0	4.6	5.7	15.6	10.0	2.7	5.5	11.8	4.0
ERCc (%)	0.6	1.1	0.4	0.3	0.1	0.3	0.1	0.1	0.3	0.4

3.5. Quantitative results and discussion

The testing phase is performed using all samples from the test dataset. Hence, the performance of the handwritten digit recognition classification is evaluated on an appropriate choice of descriptors using the SVM classifiers and then we evaluate the combination of the SVMs classifiers within DSmT framework.

3.5.1. Comparative analysis of features

The choice of the complementary features is an important step to ensure efficiently the combination. Indeed, the DSmT-based combination allows offering an accurate performance when the selected features are complementary. Hence, we propose in this section the performance of features in order to select the best ones for combining through the DSmT. For this, we evaluate each SVM-OAA implementation using Foreground Features (FF), Background Features (BF), Geometric Features (GF), Uniform Grid Features (UGF), and the descriptors deduced from a concatenation between at least two simple descriptors such as (BF,FF), (BF,FF,GF) and (UGF,BF,FF,GF). Indeed, the experiments have shown that the appropriate choice of both descriptors and concatenation order to represent each digit class in the feature generation step provides an interesting error reduction. In table 3, FF and UGF-based descriptors using SVM classifiers are evaluated. When

concatenating background and foreground (BF,FF)-features, we observe a significant reduction of the MER. Indeed, an error rate reduction of 6.71% is obtained when concatenating BF and FF, respectively. Furthermore, an error rate reduction of 1.5% is obtained when concatenating BF, FF and GF, respectively. This proves that BF, FF and GF are complementary and are more suitable for concatenation. In contrast, when concatenating UGF with BF, FF and GF, the MER is increased to 2.73% comparatively to UGF. This proves that the concatenation does not always allow improving the performance of the classification. Thus, we expect that the UGF and (BF,FF,GF) descriptors are more suitable for combining through the DSmT.

Table 3. Mean error rates of the SVM classifiers using different methods of feature generation

Descriptor	Error Rate (%)
(a) FF	18.87
(b) (BF,FF)	12.16
(c) (BF,FF,GF)	10.66
(d) UGF	6.98
(e) (UGF,BF,FF,GF)	9.71

### 3.5.2. *Performance evaluation of the proposed combination framework*

In these experiments, we evaluate a handwritten digit recognition classification based on a combination of SVM classifiers through DSmT. The proposed combination framework allows exploiting the redundant and complementary nature of the (BF,FF,GF) and UGF-based descriptors and manage the conflict provided from the outputs of SVM classifiers.

Decision making will be only done on the simple classes belonging to the frame of discernment. Hence, we consider in both combination process and calculation of the decision measures the masses associated to all classes representing the partial ignorance $\overline{\theta_i} = \bigcup_{\substack{0 \le j \le n-1 \\ j \ne i}} \theta_j$ and $\overline{\theta_i} \cap \overline{\theta_j}$ such that $i \ne j$. Thus, in order to appreciate the advantage of combining two sources of information through the DSmT-based algorithm, Figure 4 shows values of the distribution of the conflict measured for each test sample between both SVM-OAA implementations using (BF,FF,GF) and UGF-based descriptors for the 10 digit classes $(\theta_i, i = 0,1,\ldots,9)$, respectively. Table 4 reports the minimal and maximal values of the conflict $(K_{ci}, i = 0,1,\ldots,9)$ generated through the supervised model, which represent the mass assigned to the empty set, after combination process. As we can see, the conflict is maximal for the digit 4 while it is minimal for the digit 9.
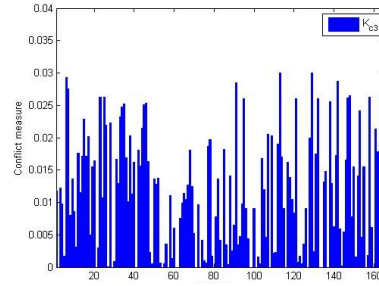
13
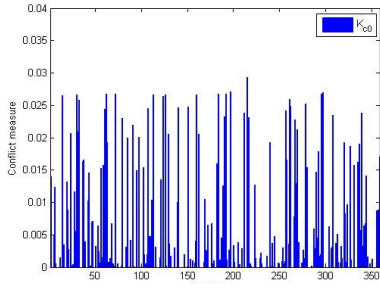
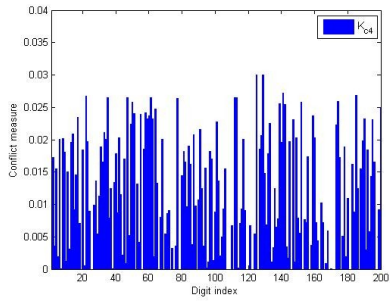(a) Measured conflict for the digits belonging to θ_0
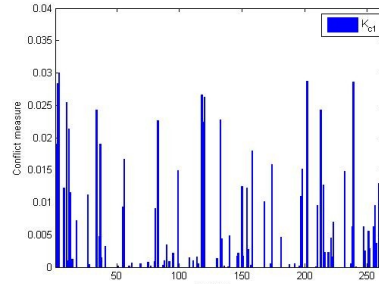
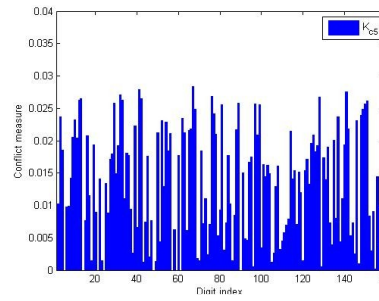(b) Measured conflict for the digits belonging to θ_1
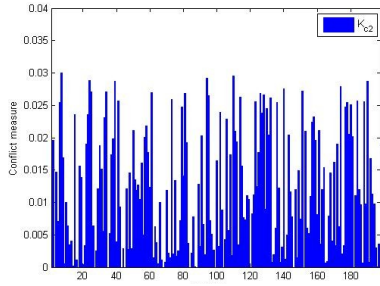
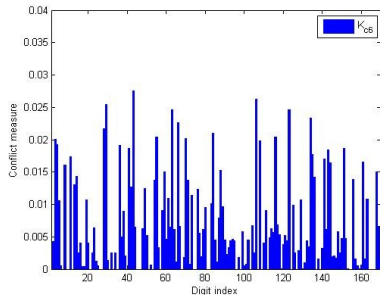(c) Measured conflict for the digits belonging to θ_2

(d) Measured conflict for the digits belonging to θ_3

(e) Measured conflict for the digits belonging to θ_4

(f) Measured conflict for the digits belonging to θ_5

(g) Measured conflict for the digits belonging to θ_6
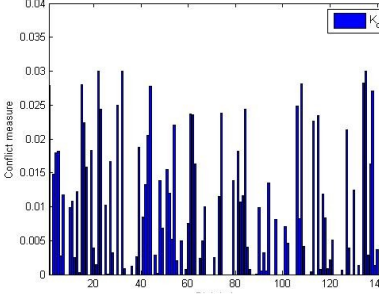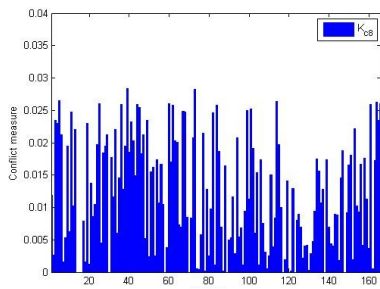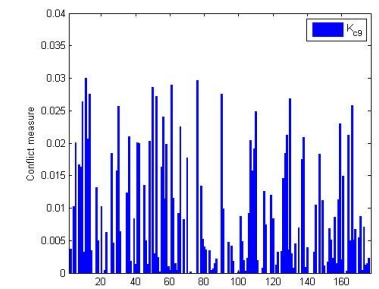
(h) Measured conflict for the digits belonging to θ_7

(i) Measured conflict for the digits belonging to θ_8

(j) Measured conflict for the digits belonging to θ_9

Figure 4. Conflict between both SVMs classifiers using (BF,FF,GF) and UGF-based descriptors for the ten digit classes $\left(\theta_i, i = 0,1,\ldots,9\right)$, respectively.

Table 4. Ranges of conflict variations measured between both SVM-OAA implementations using (BF,FF,GF) and UGF-based descriptors

Class	Minimal conflict (10^{-5})	Maximal conflict (10^{-2})
0	2.149309	2.9933
1	6.999035	2.9964
2	2.747717	2.9992
3	2.936855	2.9994
4	0.494599	3.0000
5	1.868961	2.9970
6	2.537015	2.9887
7	2.826402	2.9983
8	1.485899	2.9910
9	0.276778	2.9999

For an objective evaluation, Table 5 shows ERC and MER produced from three SVM-OAA implementations using UGF, (BF,FF,GF), the descriptor resulting from a concatenation of both UGF and (BF,FF,GF) (i.e. combination at features level) and finally the PCR6 combination rule (i.e. combination at measure level) performed on (BF,FF,GF) and UGF based descriptors, respectively.

Table 5. Error rates of the proposed framework with PCR6 combination rule using (BF,FF,GF) and UGF descriptors

	Descriptor		Concatenation	Combination rule
ERC (%)	(BF,FF,GF)	UGF	(UGF,BF,FF,GF)	PCR6
0	6.69	**1.95**	9.75	**1.95**
1	4.55	3.79	3.79	**3.03**
2	12.63	8.08	**3.54**	6.06
3	17.47	**10.84**	18.67	**10.84**
4	20.00	11.50	19.50	**9.00**
5	16.87	10.00	10.62	**7.50**
6	**2.94**	5.29	4.71	3.53
7	8.84	8.16	8.84	**4.76**
8	12.05	10.84	10.24	**6.63**
9	10.73	6.21	10.17	**5.65**
MER (%)	**10.66**	**6.98**	**9.71**	**5.43**

Overall, the proposed framework using PCR6 combination rule is more suitable than individual SVM-OAA implementations since it provides a MER of 5.43% comparatively to the concatenation which provides a MER of 9.71%. However, when

inspecting carefully each class, we can note that the PCR6 combination rule allows keeping or reducing in the most cases the ERC except for the samples belonging to classes θ_2 and θ_6. This bad performance is due to the wrong characterization of both UG and (BF,FF,GF)-based descriptors. In other words, the PCR6 combination is not reliable when the complementary information provided from both descriptors is wrongly preserved.

Thus, PCR6 combination rule allows managing correctly the conflict generated from SVM-OAA implementations, even when they provide very small values of the conflict (see Table 4) mainly in the case of samples belonging to θ_8. Thus, the DSmT is more appropriate to solve the problem for handwritten digit recognition. Indeed, the PCR6 combination rule allows an efficient redistribution of the partial conflicting mass only to the elements involved in the partial conflict, contrary to the DST which redistributes the beliefs through a simple normalization by $(1 - K_{cw}, w = 0,1,\ldots,9)$ in the combination process of masses. After redistribution, the combined mass is transformed into the DSm probability and the maximum likelihood (ML) test is used for decision making. Finally, the proposed algorithm in DSmT framework is the most stable across all experiments whereas recognition accuracies pertaining to both individual SVMs classifiers vary significantly.

4. Conclusion and future work

In this paper, we proposed an effective use of the DSmT for multi-class classification using conjointly the SVM-OAA implementation and a supervised model. Exclusive constraints are introduced through a direct estimation technique to compute the belief assignments and reduce the number of focal elements. Therefore, the proposed framework allows reducing drastically the computational complexity of the combination process for the multi-class classification. A case study conducted on the handwritten digit recognition shows that the proposed supervised model with PCR6 rule yields the best performance comparatively to SVM multi-classifications even when they provide uncalibrated outputs. In continuation to the present work, the next objectives consist to adapt the use of one-class classifiers instead of the OAA implementation of SVM in order to obtain a fixed number of focal elements within DSmT combination process. This will allow us to have a feasible computational complexity independently of the number of combined sources and the size of the discernment space.

References

[Duda, 2001] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification. 2nd edition, Wiley Interscience, New York, 2001.

[Cher, 2007] M. Cheriet, N. Kharma, C.-L. Liu, C.Y. Suen, Character Recognition Systems: A Guide for Students and Practitioner, John Wiley & Sons, 2007.

[Rahm, 2003] A. Rahman and M.C. Fairhurst, Multiple classifier decision combination strategies for character recognition: A review, International Journal on Document Analysis and Recognition 5 (4) (2003) 166-194.

[Xu, 1992] L. Xu, A. Krzyzak, C. Suen, Methods of combining multiple classifiers and their applications to handwriting recognition, IEEE Transactions on Systems, Man, and Cybernetics 22 (3) (1992) 418-435.

[Jain, 2000] A. Jain, R. Duin, J. Mao, Statistical pattern recognition: a review, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22 (1) (2000) 4-37.

[Ruta, 2000] D. Ruta and B. Gabrys, An overview of classifier fusion methods, Computing and Information Systems 7 (1) 2000 1-10.

[Abba, 2012] N. Abbas, Y. Chibani, H. Nemmour, Handwritten Digit Recognition Based On a DSmT-SVM Parallel Combination, in: Proceedings of the 13th International Conference on Frontiers in Handwriting Recognition (ICFHR), 2012, pp. 241-246. <http://www.icfhr2012.uniba.it/paper038.pdf>

[Zade, 1968] L.A. Zadeh, Fuzzy algorithm. Information and Control 12 (1968) 94-102.

[Demp, 1967] A.P. Dmpster, Upper and lower probabilities induced by a multivalued maping, Annals of Mathematical Statistics 38 (2) (1967) 325-339.

[Shaf, 1976] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.

[Dubo, 1988] D. Dubois, H. Prade, Representation and combination of uncertainty with belief functions and possibility measures, Computational Intelligence 4 (1988) 244-264.

[Smar, 2004] F. Smarandache, J. Dezert, Advances and Applications of DSmT for Information Fusion, American Research Press, 2004.

[Smar, 2006] F. Smarandache and J. Dezert, Advances and Applications of DSmT for Information Fusion, American Research Press, 2006.

[Smar, 2009] F. Smarandache and J. Dezert, Advances and Applications of DSmT for Information Fusion, American Research Press, 2009.

[Shaf, 1990] G. Shafer, Perspectives on the Theory and Practice of Belief Functions, International Journal of Approximate Reasoning 4 (1990) 323-362.

[Smet, 1999] Ph. Smets, Practical uses of belief functions, In K.B. Laskey, H. Prade, editors, Fifteenth Conference on Uncertainty in Artifficial Intelligence 99 (1999) 612-621.

[Prov, 1992] G.M. Provan, The validity of Dempster-Shafer Belief Functions, International Journal of Approximate Reasoning 6 (1992) 389-399.

[Dubo, 1992] D. Dubois, H. Prade, Evidence, Knowledge and Belief Functions, International Journal of Approximate Reasoning, 6 (1992) 295-319.

[Rusp, 1992] E.H. Ruspini, D.J. Lowrance, T.M. Start, Understanding Evidential Reasoning, International Journal of Approximate Reasoning 6 (1992) 401-424.

[Corg, 2003] S. Corgne, L. Hubert-Moy, J. Dezert, G. Mercier, Land cover change prediction with a new theory of plausible and paradoxical reasoning, in: Proceedings of the 6th International Conference on Information Fusion, 2003, pp. 1141-1148.

[Maup, 2004] P. Maupin and A.-L. Jousselme, "Vagueness, a multifacet concept - a case study on Ambrosia artemisiifolia predictive cartography," in: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), vol. 1, 2004, pp. 20-24.

[Elha, 2011] A. Elhassouny, S. Idbraim, A. Bekkari, D. Mammass, D. Ducrot, Change Detection by Fusion/Contextual Classification based on a Hybrid DSmT Model and ICM with Constraints, International Journal of Computer Applications 35 (8) (2011) 28-40.

[Zhun, 2012] Z. Liu, J. Dezert, G. Mercier, Q. Pan, Dynamical Evidential Reasoning For Changes Detections In Remote Sensing Images, IEEE Transactions on Geoscience and Remote Sensing 50 (5) (2012) 1955-1967.

[Pann, 2008] B. Pannetier, J. Dezert, E. Pollard, Improvement of Multiple Ground Targets Tracking with GMTI Sensor and Fusion of Identification Attributes, Aerospace Conference, 2008, pp. 1-13.

[Pann, 2009] B. Pannetier, J. Dezert, GMTI and IMINT data fusion for multiple target tracking and classification, in: Proceedings of the 12th International Conference on Information Fusion, 2009, pp. 203-210.

[Kech, 2009] P. Kechichian, B. Champagne, An improved partial Haar dual adaptive filter for rapid identification of a sparse echo channel, Signal Processing 89 (5) (2009) 710-723.

[Sun, 2010] Y. Sun, L. Bentabet, A particle filtering and DSmT Based Approach for Conflict Resolving in case of Target Tracking with multiple cues, Journal of Mathematical Imaging and Vision 36 (2) (2010) 159-167.

[Deze, 2010] J. Dezert, B. Pannetier, A PCR-BIMM filter for maneuvering target tracking, in: Proceedings of the 13th International Conference on Information Fusion, 2010.

[Pann, 2011] B. Pannetier, J. Dezert, Extended and multiple target tracking: Evaluation of an hybridization solution, in: Proceedings of the 14th International Conference on Information Fusion, 2011.

[Sing, 2008] R. Singh, M. Vatsa, A. Noore, Integrated Multilevel Image Fusion and Match Score Fusion of Visible and Infrared Face Images for Robust Face Recognition, Pattern Recognition - Special Issue on Multimodal Biometrics 41 (3) (2008) 880-893.

[Vats, 2009a] M. Vatsa, R. Singh, A. Noore, Unification of Evidence Theoretic Fusion Algorithms: A Case Study in Level-2 and Level-3 Fingerprint Features, IEEE Transaction on Systems, Man, and Cybernetics - A 29 (1) (2009).

[Vats, 2009b] M. Vatsa, R. Singh, A. Noore, M. Houck, Quality-Augmented Fusion of Level-2 and Level-3 Fingerprint Information using DSm Theory, International Journal of Approximate Reasoning 50 (1) (2009).

[Vats, 2010] M. Vatsa, R. Singh, A. Noore, A. Ross, On the Dynamic Selection in Biometric Fusion Algorithms, IEEE Transaction on Information Forensics and Security 5 (3) (2010) 470-479.

[Garc, 2008] E. Garcia, L. Altamirano, Multiple Cameras Fusion Based on DSmT for Tracking Objects on Ground Plane, in: Proceedings of the 11th International Conference on Information Fusion, 2008.

[Khod, 2010] M. Khodabandeh, A. Mohammad-Shahri, Data Fusion of Cameras' Images for Localization of an Object; DSmT-based Approach, in: Proceedings of the 1st International Symposium on Computing in Science & Engineering, 2010.

[Deze, 2011] J. Dezert, Z. Liu, G. Mercier, Edge Detection in Color Images Based on DSmT, in: Proceedings of the 14th International Conference on Information Fusion, 2011.

[Huan, 2006] X. Huang, X. Li, M. Wang, J. Dezert, A FUSION MACHINE BASED ON DSmT AND PCR5 FOR ROBOT'S MAP RECONSTRUCTION, International Journal of Information Acquisition (IJIA) 3 (3) (2006) 201-211.

[Li, 2006a] X. Li, X. Huang, M. Wang, Robot Map Building from Sonar Sensors and DSmT, International Journal of Information & Security 20 (2006) 104-121.

[Li, 2006b] X. Li, X. Huang, M. Wang, Sonar Grid Map Building of Mobile Robots Based on DSmT, Information Technology Journal 5 (2) (2006) 267-272.

[Li, 2007] X. Li, X. Huang, J. Dezert, L. Duan, M. Wang, A SUCCESSFUL APPLICATION OF DSmT IN SONAR GRID MAP BUILDING AND COMPARISON WITH DST-BASED APPROACH, International Journal of Innovative Computing, Information and Control (ICIC) 3 (3) (2007) 539-549.

[Li, 2008] P. Li, X. Huang, M. Wang, X. Zeng, Multiple Mobile Robots Map Building Based on DSmT, IEEE International Conference on Robotics, Automation and Mechatronics (RAM) (2008) 509-514.

[Huan, 2009] X. Huang, P. Li, M. Wang, Evidence Reasoning Machine based on DSmT for mobile robot mapping in unknown dynamic environment, in: Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO), 2009, pp. 753-758.

[Abba, 2012a] N. Abbas, Y. Chibani, SVM-DSmT combination for the simultaneous verification of off-line and on-line handwritten signatures, International Journal of Computational Intelligence and Applications (IJCIA) 11 (3) 2012.

[Abba, 2012b] N. Abbas, Y. Chibani, SVM-DSmT Combination for Off-Line Signature Verification, in: Proceedings of the IEEE International Conference on Computer, Information and Telecommunication Systems (CITS), 2012, pp. 1-5.

[Abba, 2012c] N. Abbas, Y. Chibani, H. Nemmour, Handwritten Digit Recognition Based On a DSmT-SVM Parallel Combination, in: Proceedings of the 13th International Conference on Frontiers in Handwriting Recognition (ICFHR), 2012, pp. 241-246.

[Dede, 1897] R. Dedekink, Über Zerlegungen von Zahlen durch ihre grössten gemeinsammen Teiler, in: Gesammelte Werke, Bd. 1, 1897, pp. 103-148.

[Comt, 1974] L. Comtet, Sperner Systems, sec. 7.2 in Advanced Combinatorics: The Art of Finite and Infinite Expansion, D. Reidel Publ. Co. (1974) 271-273.

[Tomb, 2001] M. Tombak, A. Isotamm, T. Tamme, On logical method for counting Dedekind numbers, Lecture Notes in Computer Science, vol. 2138, Springer-Verlag, 2001, pp. 424-427. <www.cs.ut.ee/people/m_tombak/publ.html>

[Deze, 2004a] J. Dezert, F. Smarandache, The generation of the hyper-power sets, Chap. 2, in: Advances and Application of DSmT for Information Fusion, American Research Press, 2004, pp. 37-48.

[Deze, 2004b] J. Dezert, F. Smarandache, Partial ordering on hyper-power sets, Chap. 3, in: Advances and Application of DSmT for Information Fusion, American Research Press, 2004, pp. 49-60.

[Kenn, 1992] R. Kennes, Computational Aspect of the Möbius Transformation of Graphs, IEEE Transaction on Systems, Man, and Cybernetics – Part A: Systems and Humans 22 (2) (1992) 201-223.

[Smet, 2002] Ph. Smets, The application of matrix calculus for belief functions, International Journal of Approximate Reasoning 31 (2002) 1-30.

[Deno, 2001] T. Denoeux, Inner and outer approximation of belief structures using a hierarchical clustering approach, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 9 (4) (2001) 437-460.

[Djik, 2004] P. Djiknavorian, D. Grenier, Reducing DSmT hybrid rule complexity through optimization of the calculation algorithm, Chap. 15, in: Advances and Application of DSmT for Information Fusion, American Research Press, 2006, pp. 345-440.

[Mart, 2006] A. Martin, C. Osswald, A new generalization of the proportional conflict redistribution rule stable in terms of decision, Chap. 2, in: Advances and Application of DSmT for Information Fusion, American Research Press, 2006, pp. 69-88.

[Mart, 2009] A. Martin, Implementing general belief function framework with a practical codification for low complexity, Chap. 7, in: Advances and Application of DSmT for Information Fusion, American Research Press, 2009, pp. 217-273.

[Deze, 2004c] J. Dezert, F. Smarandache, Combination of beliefs on hybrid DSm models, Chap. 4, in: Advances and Application of DSmT for Information Fusion, American Research Press, 2004, pp. 61-103.

[Li, 2011] X. Li, J. Dezert, F. Smarandache, X. Huang, Evidence Supporting Measure of Similarity for Reducing the Complexity in Information Fusion, Information sciences 181 (10) (2011) 1818–1835.

[Deng, 1999] P.S. Deng, H.Y.M. Liao, C.-W. Ho, H.-R. Tyan, Wavelet based off-line handwritten signature verification, Computer Vision and Image Understanding 76 (3) (1999) 173–190.

[Fang, 2003] B. Fang, C.H. Leung, Y.Y. Tang, K.W. Tse, P.C.K. Kwok, Y.K. Wong, Off-line signature verification by the tracking of feature and stroke positions, Pattern Recognition 36 (2003) 91–101.

[Guo, 2001] J.K. Guo, D. Doermann, A. Rosenfeld, Forgery detection by local correspondence, International Journal of Pattern Recognition and Artificial Intelligence 15 (4) (2001) 579–641.

[Fang, 2001] B. Fang, Y.Y. Wang, C.H. Leung, K.W. Tse, Y.Y. Tang, P.C.K. Kwok, Y. K. Wong, Offline signature verification by the analysis of cursive strokes, International Journal of Pattern Recognition and Artificial Intelligence 15 (4) (2001) 659–673.

[Sabo, 1997] R. Sabourin, G. Genest, F. Prêteux, Off-line signature verification by local granulometric size distributions, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (9) (1997) 976–988.

[Just, 2005] E.J.R. Justino, F. Bortolozzi, R. Sabourin, A comparison of SVM and HMM classifiers in the off-line signature verification, Pattern Recognition Letters 26 (2005) 1377-1385.

[Just, 2001] E.J.R. Justino, F. Bortolozzi, R. Sabourin, Off-line Signature Verification Using HMM for Random, Simple and Skilled Forgeries, in: Proceedings of the International Conference on Document Analysis and Recognition (ICDAR), vol. 1, 2001, pp. 105-110.

[Coet, 2004] J. Coetzer, B.M. Herbst, J.A. de Preez, Offline Signature Verification Using the Discrete Radon Transform and a Hidden Markov Model, EURASIP Journal on Applied Signal Processing 4 (2004) 559-571.

[Kaew, 1999] T. Kaewkongka, K. Chamnongthai, B. Thipakorn, Offline signature recognition using parameterized hough transform, in: Proceedings of the 5th International Symposium on Signal Processing and Its Applications, 1999, pp. 451–454.

[Quek, 2002] C. Quek, R.W. Zhou, Antiforgery: a novel pseudo-outer product based fuzzy neural network driven signature verification system, Pattern Recognition Letters 23 (14) (2002) 1795–1816.

[Weston, 1998] J. Weston, C. Watkins, Support Vector Machines for Multi-Class Pattern Recognition Machines, in: CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London, 1998.

[West, 1998] J. Weston, C. Watknis, Multi-class support vector machines, Tech. Rep. CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London, Egham, TW20 0EX, UK, 1998.

[Guer, 1999] Y. Guermeur, A. Elisseeff, H. Paugam-Moisy, Estimating the sample complexity of a multi-class discriminant model, In: Proceedings of the Industrial Conference on Artificiel Neural Networks, 1999, pp. 310-315.

[Hsu, 2001] C. Hsu, C. Lin, A comparaison of methods for muli-class support vector machines, Tech. Rep, Department of Computer Science and Information Engineering, National Taiwan University, 2001.

[Huan, 2002] H.P. Huang, Y.H. Liu, Fuzzy support vector machines for pattern recognition and data mining, International Journal of Fuzzy Systems 4 (3) (2002) 826-835.

[Martin, 2007] A. Martin, C. Osswald, Toward a combination rule to deal with partial conflict and specificity in belief functions theory, in: Proceedings of the International Conference on Information Fusion, 2007.

[Aregui, 2007] A. Aregui, T. Denoeux, Fusion of one-class classifier in the belief function framework, in: Proceedings of the International Conference on Information Fusion, 2007.

[Quost, 2007a] B. Quost, T. Denoeux, M. Masson, Pairwise classifier combination using belief functions, Pattern Recognition Letters 28 (5) (2007) 644–653.

[Quost, 2005] B. Quost, T. Denoeux, M. Masson, Pairwise classifier combination in the framework of belief functions, in: Proceedings of the International Conference on Information Fusion, 2005.

[Quost, 2006] B. Quost, T. Denoeux, M. Masson, One-against-all combination in the framework of belief functions, in: Proceedings of the IPMU , vol. 1, 2006, pp. 356-363.

[Quost, 2007b] B. Quost, T. Denoeux, M. Masson, Combinaison crédibiliste de classifieurs binaires, Traitement du Signal 24 (2) (2007) 83-101.

[Hu, 2005] Z. Hu, Y. Li, Y. Cai, X. Xu, Method of combining multi-class SVMs using Dempster-Shafer theory and its application, in: Proceedings of the American Control Conference, 2005.

[Martin, 2008] A. Martin, I. Quidu, Decision support with belief functions theory for seabed characterization, in: Proceedings of the International Conference on Information Fusion, 2008.

[Burger, 2006] T. Burger, O. Aran, Modeling hesitation and conflict: a belief-based approach for multi-class problems, in: Proceedings of the 5th International Conference on Machine Learning and Applications-ICMLA, 2006, pp. 95-100.

[Burger, 2008] T. Burger, A. Urankar, O. Aran, L. Akarun, A. Caplier, A Dempster-Shafer theory based combination of classifiers for hand gesture recognition, Computer Vision and Cumputer Graphics: Theory and Applications 21 (2008) 137-150.

[Laanay, 2010] H. Laanaya, A. Martin, D. Aboutajdine, A. Khenchaf, Support vector regression of membership functions and belief functions-Application for pattern recognition, Information Fusion 11 (4) (2010) 338–350.

[Bott, 1994] L. Bottou, C. Cortes, J. Drucker, I. Guyon, L. LeCunn, U. Muller, E. Sackinger, P. Simard, V. Vapnik, Comparaison of Classifier methods: a case study in handwriting digit recognition, In: Proceedings of the International Conference on Pattern Recognition, 1994, pp. 77-87.

[Deno, 1997] T. Denoeux, Analysis of evidence theoretic decision rules for pattern classification, Pattern recognition 30 (7) (1997) 1095-1107.

[Smet, 1994] Ph. Smets, R. Kennes, The transferable belief model, Artificial Intelligence 66 (1994) 191-234.

[Dubo, 1994] D. Dubois, H. Prade, Fuzzy Sets-A Convenient Fiction for Modeling Vagueness and Possibility, IEEE Transactions on Fuzzy Systems 2 (1) (1994).

[Appr, 1999] A. Appriou, Multisensor signal processing in the framework of the theory of evidence, NATO/RTO, Application of Mathematical Signal Processing Techniques to Mission Systems, 1999.

[Plat, 1999] J.C. Platt, Probabilities for SV Machines, Advances in Large Margin Classifiers, MIT Press, 1999, pp. 61-74.

[Bloc, 2003] I. Bloch. Fusion d'informations en traitement du signal et des images. IC2. Hermès Science, 2003.

[Fren, 1985] S. French, Group Consensus Probability Distributions: A Critical Survey, In J. Bernardo et al., editor, Bayesian Statistics, Elsevier, 1985, pp. 183-201.

[Cook, 1988] R. Cooke, Uncertainty in Risk Assessment: A Probabilist's Manifesto, Reliability in Engineering and Systems Safety 23 (1988) 277-283.

[Cook, 1991] R. Cooke. Experts in Uncertainty. Oxford University Press, 1991.

[Otsu, 1979] N. Otsu, A Threshold Selection Method from Gray-Level Histograms, IEEE Transactions on Systems, Man, and Cybernetics 9 (1979) 62-66.

[Brit, 2004] A. Britto, R. Sabourin, F. Bortolozzi, C. Suen, Foreground and Background Information in an HMM-based Method for Recognition of Isolated Characters and Numeral Strings, in: Proceedings of the International Workshop on Frontiers in Handwriting Recognition, 2004, pp. 371-376.

[Cava, 2006] P.R. Cavalin, A. Britto, F. Bortolozzi, R. Sabourin, L. Oliveira, An implicit segmentation-based method for recognition of handwritten strings of characters, ACM symposium on Applied computing, 2006, pp. 836-840.

[Fava, 1996] J. Favata, G. Srikantan, A Multiple Feature/Resolution Approach To Handprinted Digit and Character Recognition, Proceedings of the International journal of imaging systems and technology 7 (4) (1996) 304–311.

[Abba, 2011] N. Abbas, Y. Chibani, Combination of Off-Line and On-Line Signature Verification Systems Based on SVM and DST, in: The 11th International Conference on Intelligent Systems Design and Applications, 2011, pp. 855 – 860.

[Vapn, 1995] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer, 1995.

[Cort, 1995] D. Cortes, V. Vapnik, Support vector networks, Machine Learning 20 (1995) 273-297.

[Lowr, 1991] J.D. Lowrance, T.M. Strat, L.P. Wesley, T.D. Garvey, E.H. Ruspini, D.E. Wilkins, The Theory, Implementation and Practice of Evidential Reasoning, SRI project 5701 final report, SRI, Palo Alto, 1991.