Dynamic Evidential Reasoning for Change Detection in Remote Sensing Images

Zhun-ga Liu, Jean Dezert, Grégoire Mercier, and Quan Pan

***Abstract*—Theories of evidence have already been applied more or less successfully in the fusion of remote sensing images. These attempts were based on the classical evidential reasoning which works under the condition that all sources of evidence and their fusion results are related to the same invariable (static) frame of discernment. When working with multitemporal remote sensing images, some change occurrences are possible between two images obtained at a different period of time, and these changes need to be detected efficiently in particular applications. The classical evidential reasoning is adapted for working with an invariable frame of discernment over time, but it cannot efficiently detect nor represent the occurrence of change from heterogeneous remote sensing images when the frame is possibly changing over time. To overcome this limitation, dynamic evidential reasoning (DER) is proposed for the sequential fusion of multitemporal images. A new state-transition frame is defined in DER, and the change occurrences can be precisely represented by introducing a state-transition operator. Two kinds of dynamical combination rules working in the free model and in the constrained model are proposed in this new framework for dealing with the different cases. Moreover, the prior probability of state transitions is taken into account, and the link between DER and Dezert–Smarandache theory is presented. The belief functions used in DER are defined similarly to those defined in the Dempster–Shafer theory. As shown in the last part of this paper, DER is able to estimate efficiently the correct change detections as a postprocessing technique. Two applications are given to illustrate the interest of DER: The first example is based on a set of two SPOT images acquired before and after a flood, and the second example uses three QuickBird images acquired during an earthquake event.**

***Index Terms*—Change detection, Dezert–Smarandache theory (DSmT), Dempster–Shafer theory (DST), dynamical evidential reasoning, evidence theory, image fusion.**

I. Introduction

INFORMATION fusion resulting from multitemporal and multisource remote sensing images remains an open and important problem [1]. The remote sensing images can be quite different in their modality [2]: Orbits may be ascending and descending, and parameters of acquisitions may differ from one image to the other even when the two acquisitions are derived from the same sensor. In change detection of remote sensing images, the authors focused first on change measure [3]–[7], and then, they developed techniques for classifying changed features [8]–[14]. At this level, the change detection appears as a classification problem. The specific case of very high resolution data suggests the introduction of object recognition so that the temporal behavior of those satellite image objects is being used for change detection applications [15]. In this paper, the change detection is considered as a postclassification procedure that focuses on the transition between classes. Hence, we consider that it is reasonable to deal with uncertain, imprecise, and even conflicting information. Evidence theories, including the Dempster–Shafer theory (DST) [16] and Dezert–Smarandache theory (DSmT) [17]–[19], are valuable for dealing with such information, and they have already been applied for remote sensing applications [1], [20]–[22]. In past works, a particular attention was paid to obtain very specific results for the decision-making support through efficient fusion of the sources of evidence. Thus, many works focused mainly on the redistribution of the conflicting beliefs [23]–[25]. These combination approaches can be called static approaches since they work under the assumption that the frame, on which the decision-making support is based, is temporally invariable in the fusion process. However, in the fusion of the multitemporal remote sensing images, unexpected change occurrences can arise in some parts of the images. The classical combination rules in evidence theories provide specific classification results in the invariable parts of images, but it cannot precisely detect the change occurrences in the variable parts. In fact, the changes can be considered as a conflict between the sources of information in a multitemporal fusion process, but the classical conflict belief produced by the conjunctive combination cannot precisely represent change areas [21].

Therefore, a dynamic evidential reasoning (DER) working under the condition that the frame does not necessarily remain invariable during the fusion is proposed. In this paper, a new frame called "state-transition power set" is defined, and the change occurrences among different hypotheses can be precisely represented by the state-transition operator in this frame. Dynamical combination rules are then proposed to work either in the free model or in the constrained model. The free model is well adapted when no prior knowledge is known on elements of the frame. The constrained model can be used if some integrity constraints between elements of the frame are known.

Manuscript received January 18, 2011; revised May 2, 2011 and June 17, 2011; accepted September 11, 2011. Date of publication October 27, 2011; date of current version April 18, 2012. This work was supported in part by the China Natural Science Foundation under Grant 61075029 and in part by the Ph.D. Thesis Innovation Fund from Northwestern Polytechnical University under Grant cx201015.

Z. Liu is with the School of Automation, Northwestern Polytechnical University, Xi'an 710072, China, and also with the Télécom Bretagne, Technopôle Brest-Iroise, 29238 Brest, France (e-mail: liuzhunga@gmail.com).

J. Dezert is with the Office National d'Etudes et de Recherches Aerospatiales, The French Aerospace Lab, 91761 Palaiseau, France (e-mail: jean.dezert@onera.fr).

G. Mercier is with the Télécom Bretagne, Technopôle Brest-Iroise, 29238 Brest, France (e-mail: Gregoire.Mercier@telecom-bretagne.eu).

Q. Pan is with the School of Automation, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: liuzhunga@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2011.2169075

The conditioning combination rule is also used when the prior probability about the change occurrences is available. There are some links between DER and DST, or DSmT, which will be analyzed in the sequel. The dynamic belief function $\mathrm{Bel}(\cdot)$, plausibility function $\mathrm{Pl}(\cdot)$, and pignistic probability $\mathrm{BetP}(\cdot)$ [26], [27] are defined similarly as in DST. The dynamical approach improves the performances of the classification of areas and also of the estimation of the changes through the fusion of multitemporal sources of evidence.

The next section introduces the DER framework in detail and the dynamic combination that works on the free model (see Section II-B1) as well as the hybrid model (see Section II-B2) and the decision-making support (see Section II-D). Section II-C focuses on the case where some prior knowledge is available to account for some class transition constraint. Section III presents a set of examples that show why the classical belief approach cannot perform accurately for analyzing time series of images. Section IV gives two applications of the DER approach: The first application is based on a set of two SPOT images acquired before and after a flood over Gloucester, U.K., in 2000 (see Section IV-A), and the second one considers a set of three QuickBird satellite images acquired before and after the Boumerdes earthquake in 2003 (see Section IV-B). Section V concludes.

II. DER Approach

A. Space of State Transitions in DER

DSmT has been already applied for the fusion of multitemporal satellite images in [1]. However, the classical frame used is not well adapted for measuring the changes among its elements. The conjunctive elements (intersections) in hyperpower set D^Θ represent either the overlap between hypotheses in free Dezert-Smarandache (DSm) model or the conflict produced by the conjunctive combination in the hybrid DSm model when there is a known integrity constraint between some elements of the frame of discernment. Therefore, the conjunction $A \cap B$ is unable to characterize the transition A changing to B (denoted by $A \to B$) or B changing to A (denoted by $B \to A$). If we need to distinguish two possible state transitions for change detection, we need to define a new operator, and we cannot use the classical conjunctive (intersection) operator as in classical approaches. We propose the state-transition operator "changing to," denoted by $\to$, satisfying the following reasonable conditions:

(C1) Impossible (forward) state transition where $\emptyset$ stands for the null hypothesis

$$A \to \emptyset \triangleq \emptyset$$

(C2) Impossible (backward) state transition

$$\emptyset \to A \triangleq \emptyset$$

(C3) Distributivity of $\cup$ w.r.t. $\to$

$$(A \cup B) \to C = (A \to C) \cup (B \to C)$$

(C4) Distributivity of $\to$ w.r.t. $\cup$

$$A \to (B \cup C) = (A \to B) \cup (A \to C)$$

(C5) Associativity of state transition

$$(A \to B) \to C = A \to (B \to C) = A \to B \to C.$$

For notation convenience, a (state) transition $A \to B$ will be denoted by $t_{A,B}$. It is important to note that the order of indexes does matter because $t_{A,B} \neq t_{B,A}$ in general, but if $A = B$, obviously, $t_{A,A} = A \to A$ represents a particular transition, i.e., the invariable transition actually from state A to state A. A chain of transitions $\theta_1 \to \theta_2 \cdots \to \theta_n$ will be denoted by $t_{1,2,\ldots,n}$, a transition $\theta_i \to (\theta_j \cup \theta_k)$ will be denoted by $t_{i,j\cup k}$, a transition $(\theta_i \cap \theta_j) \to (\theta_k \cup \theta_l)$ will be denoted by $t_{i\cap j,k\cup l}$, etc.

In the theories of belief functions (DST, DSmT, or transferable belief model (TBM) [28]), the result of the fusion of sources of evidence defined on a same frame of discernment Θ is obtained by a given rule of combination relatively to a fusion space G^Θ, where G^Θ can be either the classical power set 2^Θ, a hyperpower set D^Θ, or a superpower set (the power set of the refined frame), depending on the theory used. In this paper, we propose to use another fusion space (the space of transitions), denoted by T^Θ, in order to deal explicitly with all possible state transitions that we want to detect.

First, the transition frame is given by

$$\begin{aligned}\Theta_{1\to n} &= \Theta_1 \times \Theta_2 \times \cdots \times \Theta_n \\ &= \{t_{X_1,X_2,\ldots,X_n} | X_i \in \Theta_i, \quad i = 1,2,\ldots,n\}\end{aligned}$$

where Θ_i is the frame associated with the ith source in the time series and $\times$ is the Cartesian product operator. The space of transitions will be denoted and defined by

$$T_n^\Theta = G^{\Theta_{1\to n}}.$$

In this paper, we assume to work in a simpler case where $G^\Theta = 2^\Theta$. In other words, we will work with the simpler space denoted by T_n^Θ and defined by

$$T_n^\Theta = 2^{\Theta_{1\to n}} = 2^{\Theta_1 \times \Theta_2 \times \cdots \times \Theta_n}.$$

T_n^Θ can be called state-transition power set, which is composed by all the elements in $\Theta_{1\to n}$ with the union operator $\cup$. We define the union $\cup$ as a componentwise operator in the following way:

$$\forall t_X, t_Y \in T_n^\Theta, \quad t_X \cup t_Y = t_{X\cup Y}. \tag{1}$$

Following conditions C1–C5, we note that, in most cases

$$t_{(X_1,X_2,\ldots,X_n)} \cup t_{(Y_1,Y_2,\ldots,Y_n)} = t_{(X_1,X_2,\ldots,X_n)\cup(Y_1,Y_2,\ldots,Y_n)} \neq t_{X_1\cup Y_1,X_2\cup Y_2,\ldots,X_n\cup Y_n}.$$

If $\forall X_i \neq Y_i$; X_i, Y_i are singletons, $t_{(X_1,X_2,\ldots,X_n)\cup(Y_1,Y_2,\ldots,Y_n)}$ indicates only two kinds of possible transitions, $t_{X=(X_1,X_2,\ldots,X_n)}$ or $t_{Y=(Y_1,Y_2,\ldots,Y_n)}$, whereas the element $t_{X_1\cup Y_1,X_2\cup Y_2,\ldots,X_n\cup Y_n}$ represents $2 \times 2 \times \cdots \times 2 = 2^n$ kinds of possible transitions. It is obvious that they are quite different, and $t_{X_1\cup Y_1,X_2\cup Y_2,\ldots,X_n\cup Y_n}$ is much more imprecise than $t_{(X_1,X_2,\ldots,X_n)\cup(Y_1,Y_2,\ldots,Y_n)}$. For instance, $t_{1,2} \cup t_{2,1} = t_{(1,2)\cup(2,1)} \neq t_{(1\cup 2),(2\cup 1)} = t_{(1,2)\cup(1,1)\cup(2,2)\cup(2,1)}$ following C3 and C4.

As we can see, the important and major difference between the classical approaches (DST, TBM, and DSmT) and DER approach is the choice of the fusion space that we are working with. With DST, TBM, or DSmT, the fusion space is always the same (independent of the number of sources) as soon as the sources are defined with respect to the same frame Θ, whereas with the DER approach, the fusion space is always increasing with the number of sources, even if the sources are all referring to the same frame Θ. This, of course, increases the complexity of DER approach, but this is the "price to pay" to identify and estimate the possible change occurrences in remote sensing image sequences as it will be shown in the last section of this paper. Clearly, DST and DSmT are not well adapted for detecting change occurrences. For example, the transition $t_{A,B,A}$ from state A in source 1 to state B in source 2 and back to state A in source 3 cannot be represented in the fusion space proposed classically with TBM, DST, or DSmT. In such static frames, transitions are stressed in the conflict between sources.

Example 1: Let us consider a simple case $\Theta = \Theta_1 = \Theta_2 = \{\theta_1, \theta_2\}$ with Shafer's model; then, $2^\Theta = \{\emptyset, \theta_1, \theta_2, \theta_1 \cup \theta_2\}$. We first consider at time 1 the initial set of *invariable transitions* defined as follows:

$$\Theta_{1\to 1} = \Theta;$$
$$T_1^\Theta \equiv 2^\Theta = \{t_\emptyset \equiv \emptyset, t_1 \equiv \theta_1, t_2 \equiv \theta_2, t_{1\cup 2} \equiv \theta_1 \cup \theta_2\}.$$

If we want to consider all the possible transitions from time stamp 1 to time stamp 2, one starts with the cross-product frame

$$\Theta_{1\to 2} = \Theta_1 \times \Theta_2 = \{t_{1,1}, t_{1,2}, t_{2,1}, t_{2,2}\}$$

which has $|\Theta_1| \times |\Theta_2| = 2 \times 2 = 4$ distinct elements, and we build its power set $T_2^\Theta = 2^{\Theta_{1\to 2}}$, including its 16 elements, as in the classical way. The empty set element ($\emptyset$) can be interpreted as the following set of impossible state transitions corresponding to $t_{\emptyset,\emptyset}$, $t_{\emptyset,1}$, $t_{\emptyset,2}$, $t_{\emptyset,1\cup 2}$, $t_{1,\emptyset}$, $t_{2,\emptyset}$, and $t_{1\cup 2,\emptyset}$. We recall that the imprecise elements are derived from application of conditions C3 and C4 and not from the componentwise union of n-tuples involved in transition indexes. The cardinality of T_n^Θ increases with the number of the sources n as $|T_n^\Theta| = 2^{|\Theta_1|\times|\Theta_2|\times\cdots\times|\Theta_n|}$.

The power set of transitions that we want to work with for such very simple example will be given by

$$\begin{aligned} T_2^\Theta &= 2^{\Theta_{1\to 2}} \\ &= \Big\{\emptyset, t_{1,1}, t_{1,2}, t_{2,1}, t_{2,2}, \\ &\quad t_{1,1} \cup t_{1,2} = t_{(1,1)\cup(1,2)} = t_{1,1\cup 2}, \\ &\quad t_{1,1} \cup t_{2,1} = t_{(1,1)\cup(2,1)} = t_{1\cup 2,1}, \\ &\quad t_{1,1} \cup t_{2,2} = t_{(1,1)\cup(2,2)}, \\ &\quad t_{1,2} \cup t_{2,1} = t_{(1,2)\cup(2,1)}, \\ &\quad t_{1,2} \cup t_{2,2} = t_{(1,2)\cup(2,2)} = t_{1\cup 2,2}, \\ &\quad t_{2,1} \cup t_{2,2} = t_{(2,1)\cup(2,2)} = t_{2,1\cup 2}, \\ &\quad t_{1,1} \cup t_{1,2} \cup t_{2,1} = t_{(1,1)\cup(1,2)\cup(2,1)}, \\ &\quad t_{1,1} \cup t_{1,2} \cup t_{2,2} = t_{(1,1)\cup(1,2)\cup(2,2)}, \\ &\quad t_{1,1} \cup t_{2,1} \cup t_{2,2} = t_{(1,1)\cup(2,1)\cup(2,2)}, \\ &\quad t_{1,2} \cup t_{2,1} \cup t_{2,2} = t_{(1,2)\cup(2,1)\cup(2,2)}, \\ &\quad t_{1,1} \cup t_{1,2} \cup t_{2,1} \cup t_{2,2} = t_{1\cup 2,1\cup 2}\Big\}. \end{aligned}$$

B. Combination Rules in DER

The classical combination rules usually work under the assumption that all the sources of information refer to the same common frame of discernment. The results of existing combination rules do not deal with unexpected changes in the frame, and they manage the conflicting beliefs without taking into account these possible changes in the frame. Nevertheless, for the detection of changes, the conflicting information is more important than the compatible information.

The fundamental difference between the classical approach and this proposed DER approach is that the frame of the fusion process is considered possibly variable over time, and the fusion process is adapted for the changes' detection and their identification. When the elements of the frame are invariable, they can be considered as a special case of changes corresponding to invariable transition. The dynamical combination rules will be defined by using the state-transition operator to take advantage of the useful information included in the conflict between sources. The combination rules work in the free and constrained models.

Combination Rule in the Free Model of Transitions: In the free model, there is no prior knowledge on state transitions, and all kinds of changes among the elements are considered possible. We start first with the combination of the two temporal sources of evidences. Let $\mathbf{m_1}$ and $\mathbf{m_2}$ be two basic belief assignments (bbas) provided by two temporal sources of evidence, respectively, over the frame of discernment Θ_1 and Θ_2 satisfying Shafer's model. $m_1(A)$, $A \in 2^{\Theta_1}$ is the mass of belief committed to the hypothesis A by source 1, and $m_2(B)$, $B \in 2^{\Theta_2}$ is the mass of belief committed to the hypothesis B by source 2. Sources 1 and 2 are considered independent.

In this paper, we propose to compute the mass of belief of a forward transition $t_{A_k,B_l} = A_k \to B_l$, $l \geq k$ as $m(t_{A_k,B_l}) = m_k(A)m_l(B)$ where k and l are temporal stamps/indexes. Note that, because $m_k(A)m_l(B) = m_l(B)m_k(A)$, this mass of belief also can be associated to the backward transition $t_{B_l\to A_k} = B_l \to A_k$. We assume that we are always working with ordered products corresponding to forward temporal transitions. By convention, the bba $\mathbf{m_i}$ provided by the source i is assumed to be available at time i before the bba $\mathbf{m_j}$ provided by the source j if $i < j$. Indexes of sources correspond actually to temporal stamps. Following this very simple principle, the conjunctive combination rule in the free model of transitions, denoted by DERf, is defined by

$$\forall\, t_{X_1,X_2} \in T_2^\Theta, \quad m_{1\to 2}(t_{X_1,X_2}) = m_1(X_1)m_2(X_2) \qquad (2)$$

where $X_1 \in 2^{\Theta_1}$ and $X_2 \in 2^{\Theta_2}$.

This simple conjunctive rule can be extended easily for computing the mass of belief of a chain of transitions using either a sequential fusion or a joint fusion process as follows.

1) **Sequential fusion of $n > 2$ sources**: $\forall\, t_{X_1,\ldots,X_n} \in T_n^\Theta$

$$m_{1\to n}(t_{X_1,\ldots,X_n}) = m_{1\to n-1}(t_{X_1,\ldots,X_{n-1}})m_n(X_n) \qquad (3)$$

where $t_{X_1,\ldots,X_{n-1}} \in T^{\Theta_{n-1}}$ and $X_n \in 2^{\Theta_n}$.

2) **Direct joint fusion of $n > 2$ sources**: The n sequential sources of evidence can also be directly combined by $\forall\, t_{X_1,X_2,\ldots,X_n} \in T_n^\Theta$

$$m_{1\to n}(t_{X_1,X_2,\ldots,X_n}) = m_1(X_1)m_2(X_2)\cdots m_n(X_n) \quad (4)$$

where $X_i \in 2^{\Theta_i}$, $i = 1,\ldots,n$.

In this free model, the result of the combination is very specific since all kinds of change occurrences are distinguished in the results, but the computation burden is very large because of the large increase of the cardinality of T_n^Θ with n.

Combination Rule in the Constrained Model of Transitions: In the constrained model of transitions, one knows that some kinds of changes among the different elements cannot occur according to our prior knowledge. The set $\emptyset \triangleq \{\emptyset_\mathcal{M}, \emptyset\}$ can be defined in introducing some integrity constraints as done in the hybrid model of DSmT. $\emptyset_\mathcal{M}$ includes all the transitions in T_i^Θ, $i = 1,2,\ldots,n$, which have been forced to be empty because of the chosen integrity constraints in the model $\mathcal{M}$, and $\emptyset$ is the classical empty set. If the sources of evidence share the same reliability in the combination, the conflict among the evidences will be regarded as possible changes or as empty sets depending on the constraints that we have. The mass of the empty sets arising from integrity constraints can be distributed to the other focal elements. The notation $t_{A,B} \overset{\mathcal{M}}{=} t$ means that the transition $t_{A,B}$ is equivalent to the transition t in the underlying model $\mathcal{M}$ given the integrity constraints.

- **$\mathbf{DER_{DS}}$ rule of combination:** If the mass of empty sets is small, it can be proportionally distributed to the other focal elements similarly to Dempster–Shafer's rule, and we denote this rule by $\mathrm{DER_{DS}}$ for short. This distribution of the mass of the conflict can be interpreted as optimistic and is mathematically defined as follows: $\forall t_{n-1} \in T_{n-1}^\Theta$, $\forall X_n \in 2^{\Theta_n}$, $\forall t_n \in T_n^\Theta$, $n \geq 2$

$$m_{1\to n}(t_n) = \frac{\sum_{t_{t_{n-1},X_n} \overset{\mathcal{M}}{=} t_n} m_{1\to n-1}(t_{n-1})m_n(X_n)}{1-K} \quad (5)$$

where K represents the mass of belief committed to the empty sets (i.e., the degree of conflict) which is given by

$$K = \sum_{t_{t_{n-1},X_n} \in \emptyset} m_{1\to n-1}(t_{n-1})m_n(X_n). \quad (6)$$

When considering the direct combination of n sequential sources altogether, one has $\forall t_n \in T_n^\Theta$, and $X_i \in 2^{\Theta_i}$, $i = 1,2,\ldots,n$

$$m_{1\to n}(t_n) = \frac{\sum_{t_{X_1,X_2,\ldots,X_n} \overset{\mathcal{M}}{=} t_n} m_1(X_1)\cdots m_n(X_n)}{1-K} \quad (7)$$

where

$$K = \sum_{t_{X_1,X_2,\ldots,X_n} \in \emptyset} m_1(X_1)\cdots m_n(X_n). \quad (8)$$

- **$\mathbf{DER_Y}$ rule of combination:** If the mass of the empty sets is quite large, it implies that the combination results will not be very credible, and all the mass of empty sets can be prudently allocated to the ignorance transition $t_{\Theta^n} = t_{\Theta_1,\Theta_2,\ldots,\Theta_n}$ similarly to Yager's rule [29]. For short, we denote this rule by $\mathrm{DER_Y}$, and it is mathematically defined by the following: $\forall t_n \in T_n^\Theta$, $t_n \notin \emptyset$, and $X_i \in 2^{\Theta_i}$, $i = 1,2,\ldots,n$

$$m_{1\to n}(t_n) = \sum_{t_{X_1,\ldots,X_n} \overset{\mathcal{M}}{=} t_n \notin \emptyset} m_1(X_1)\cdots m_n(X_n) + \sum_{t_{X_1,\ldots,X_n} \in \emptyset, t_n = t_{\Theta^n}} m_1(X_1)m_2(X_2)\cdots m_n(X_n). \quad (9)$$

Remark: The summation introduced in (5), (7), and (9) allows to take into account the integrity constraints of the model of the space of transitions as shown in the next example.

For simplicity and in our application, we assume that all frames $\Theta_1,\Theta_2,\ldots,\Theta_n$ do not change with time and they are the same and equal to Θ. That is, $\Theta_1 = \Theta_2 = \cdots = \Theta_n = \Theta$.

Example 2: Let us consider the frame $\Theta = \{\theta_1,\theta_2\}$ with the following two integrity constraints representing the impossible state transitions $\emptyset_\mathcal{M} \triangleq \{t_{1,2}, t_{2,2}\}$. Therefore, due to these integrity constraints, the fusion space T_2^Θ, as given in detail in Example 1, reduces to the following simple set:

$$T_2^\Theta = \{\emptyset, t_{1,1}, t_{2,1}, t_{1\cup 2,1}\}.$$

We also consider the following bba inputs:

	θ_1	θ_2	Θ
$\mathbf{m_1}$	0.4	0.6	0
$\mathbf{m_2}$	0.5	0.2	0.3

According to the underlying hybrid model $\mathcal{M}$, only the products $m_1(\theta_1)m_2(\theta_2)$ and $m_1(\theta_2)m_2(\theta_2)$ take part in the conflict. The other products $m_1(\theta_1)m_2(\theta_1)$, $m_1(\theta_1)m_2(\theta_1\cup\theta_2)$, $m_1(\theta_2)m_2(\theta_1)$, and $m_1(\theta_2)m_2(\theta_1\cup\theta_2)$ correspond to some mass of belief of admissible state transitions. The mass of belief committed to the total conflict is therefore given by

$$K = m_1(\theta_1)m_2(\theta_2) + m_1(\theta_2)m_2(\theta_2) = 0.2.$$

The conjunctive masses of belief of possible transitions are given by

$$\begin{aligned} m_\cap(t_{1,1}) &= m_1(\theta_1)m_2(\theta_1) = 0.4\times 0.5 = 0.20 \\ m_\cap(t_{1,1\cup 2}) &= m_1(\theta_1)m_2(\theta_1\cup\theta_2) = 0.4\times 0.3 = 0.12 \\ m_\cap(t_{2,1}) &= m_1(\theta_2)m_2(\theta_1) = 0.6\times 0.5 = 0.30 \\ m_\cap(t_{2,1\cup 2}) &= m_1(\theta_2)m_2(\theta_1\cup\theta_2) = 0.6\times 0.3 = 0.18. \end{aligned}$$

Due to the integrity constraints $t_{1,2} = (\theta_1 \to \theta_2) \overset{\mathcal{M}}{=} \emptyset$ and $t_{2,2} = (\theta_2 \to \theta_2) \overset{\mathcal{M}}{=} \emptyset$ and the condition C4 given in Section II-A, one has

$$\begin{cases} t_{1,1\cup 2} = t_{1,1} \cup t_{1,2} \overset{\mathcal{M}}{=} t_{1,1} \cup \emptyset = t_{1,1} \\ t_{2,1\cup 2} = t_{2,1} \cup t_{2,2} \overset{\mathcal{M}}{=} t_{2,1} \cup \emptyset = t_{2,1}. \end{cases}$$

Therefore, the mass $m_{\cap}(t_{1,1\cup 2})$ must be transferred to $t_{1,1}$, whereas the mass $m_{\cap}(t_{2,1\cup 2})$ must be transferred to $t_{2,1}$ only. Finally, the result given by the DER_{DS} rule is

$$\begin{aligned} m(t_{1,1}) &= \frac{m_{\cap}(t_{1,1}) + m_{\cap}(t_{1,1\cup 2} \equiv t_{1,1})}{1-K} \\ &= \frac{m_1(\theta_1)m_2(\theta_1) + m_1(\theta_1)m_2(\theta_1 \cup \theta_2)}{1-K} \\ &= \frac{0.20 + 0.12}{0.8} = 0.40 \\ m(t_{2,1}) &= \frac{m_{\cap}(t_{2,1}) + m_{\cap}(t_{2,1\cup 2} \equiv t_{2,1})}{1-K} \\ &= \frac{m_1(\theta_2)m_2(\theta_1) + m_1(\theta_2)m_2(\theta_1 \cup \theta_2)}{1-K} \\ &= \frac{0.30 + 0.18}{0.8} = 0.60. \end{aligned}$$

In the hybrid model, if all transitions among the exclusive elements are not allowed, the DER_{DS} combination rule reduces to the classical Dempster–Shafer's rule of combination.

If DER_{Y} is used instead of DER_{DS}, one obtains

$$\begin{aligned} m(t_{1,1}) &= m_{\cap}(t_{1,1}) + m_{\cap}(t_{1,1\cup 2} \equiv t_{1,1}) = 0.32 \\ m(t_{2,1}) &= m_{\cap}(t_{2,1}) + m_{\cap}(t_{2,1\cup 2} \equiv t_{2,1}) = 0.48 \\ m(t_{\Theta,\Theta}) &= m(t_{1\cup 2,1\cup 2}) = m(t_{(1,1)\cup(2,1)\cup(1,2)\cup(2,2)}) \\ &= m(t_{(1,1)\cup(2,1)\cup\emptyset\cup\emptyset}) = m(t_{(1,1)\cup(2,1)}) \\ &= K = 0.2. \end{aligned}$$

For decision-making support, the mass of partial ignorance and/or $m(t_{\Theta,\Theta})$ can be redistributed to the other more specific state-transition hypotheses in many ways using different probabilistic transformations of a bba into a subjective probability measure, like for instance BetP or DSmP transformations [17].

C. Prior Knowledge Consideration

In the previous section, we have only considered the bbas provided by the sources of evidences in the DER fusion rules. In some applications, however, we may also have access to extra prior information on the state transitions. This prior information is supposed to be modeled as a prior Bayesian bba $m_0(\cdot)$ on the space of all transitions or, equivalently, to a given prior probability function $p(\cdot)$. If this prior information is known and fits with the reality and if it is taken properly into account, the results in the estimation of change detections are expected to be more precise.

In this section, we propose a solution to improve the DER approach by including the prior probability about the change from the hypothesis X to Y. For this purpose, one assumes that probabilities $p(t_{X,Y})$ for all X, $Y \in 2^{\Theta}$ are known and given. The mass of belief of this change occurrence is then calculated as the weighted conjunctive product $m_{1\to 2}(t_{X,Y}) = m_1(X)p(t_{X,Y})m_2(Y)$. In the fusion of n sequential sources of evidence, the fusion results including prior information, using a first-order time-homogeneous Markov chain of transitions for $X_i \in 2^{\Theta}$, $i = 1, 2, \ldots, n$, will be given by

$$m_{1\to n}(t_{X_1,X_2,\ldots,X_n}) = m_1(X_1)\prod_{i=2}^{n} p(t_{X_{i-1},X_i})m_i(X_i). \quad (10)$$

The DER_{DS} taking into account these weighted (discounted) bbas is then given by the following: $\forall t_n \in T_n^{\Theta}$ and $X_i \in 2^{\Theta}$ for $i = 1, 2, \ldots, n$

$$m_{1\to n}(t_n) = \frac{\sum_{t_{X_1,\cdots,X_n} \overset{\mathcal{M}}{=} t_n} m_1(X_1)\prod_{i=2}^{n} p(t_{X_{i-1},X_i})m_i(X_i)}{\sum_{t_{X_1,\cdots,X_n} \neq \emptyset} m_1(X_1)\prod_{i=2}^{n} p(t_{X_{i-1},X_i})m_i(X_i)}. \quad (11)$$

If the prior probability of a given transition is 0, i.e., $p(t_{X,Y}) = 0$, then $t_{X,Y}$ will be forced to be equal to the empty set as in a constrained model. The combination rule in the constrained model can be seen as a special case of this rule where the forced empty sets are transitions assigned with zero probability $p\{\emptyset_{\mathcal{M}}\} = 0$, and the other transitions are equiprobable. The prior and transition probabilities must be acquired by experience, or from a training database.

Remark: Because the order in the computation of the product does matter in DER to identify the transitions, all the dynamical combination rules developed in the free model or constrained models (with or without prior information on the probability of transitions) are not commutative since the transitions are oriented. Therefore, the fusion of the sources must be done properly following a (forward) temporal sequence. DSmT can be seen as a generalization of DST, the power set 2^{Θ} is a subset of hyperpower set D^{Θ}, and we have

$$m_{\cap}(A) = \sum_{\substack{X_1 \cap X_2 \cdots \cap X_n = A \\ A \in D^{\Theta}, X_i \in 2^{\Theta}}} m_{1\to n}(t_{X_1,X_2,\ldots,X_n}) \quad (12)$$

where $m_{\cap}(\cdot)$ is obtained by the combination of n sources of evidence in DSm free model and $m_{1\to n}(\cdot)$ is the combination results of these n sources of evidence by DERf. This relation indicates that the combination results in DER are much more specific than that in DSmT and DST, and that is why DER is well adapted to estimate the belief in state transitions using these fusion rules.

D. Decision-Making Support With DER

For decision-making support, the element of the frame having either the maximum of belief function $\text{Bel}(\cdot)$, the maximum of plausibility function $\text{Pl}(\cdot)$, or the maximum of pignistic probability $\text{BetP}(\cdot)$ [26], [27] is often chosen.[1] These functions can be also used in DER approach as well. Indeed, all the elements in $\Theta^{1\to n}$ composed by the state transitions through the singleton elements have a specific and unique meaning, and they are considered as the singleton elements in DER. All the focal elements in T_n^{Θ} can be decomposed in the disjunctive canonical form using these singleton elements with the operator $\cup$, and we call them the canonical focal elements. For

[1]The $\text{DSmP}(\cdot)$ transformation proposed in [18] which provides better probabilistic informational content than $\text{BetP}(\cdot)$ can also be chosen instead. However, $\text{DSmP}(\cdot)$ is more complicated to implement than $\text{BetP}(\cdot)$, and it has not been tested in our application for now.

example, $m(t_{(\theta_1\cup\theta_2),\theta_3}) = m(t_{\theta_1,\theta_3} \cup t_{\theta_2,\theta_3})$ because of the condition (C3). The belief, plausibility functions, and the pignistic transformation are defined in DER similarly as in DST; that is

$$\text{Bel}(A) = \sum_{A,B\in T_n^\Theta; B\subseteq A} m(B) \tag{13}$$

$$\text{Pl}(A) = \sum_{A,B\in T_n^\Theta; A\cap B\neq\emptyset} m(B). \tag{14}$$

The interval $[\text{Bel}(A), \text{Pl}(A)]$ is then interpreted as the lower and upper bounds of imprecise probability for decision-making support [16], and the pignistic probability $\text{BetP}(A)$ commonly used to approximate the unknown probability in $[\text{Bel}(A), \text{Pl}(A)]$ is calculated by

$$\text{BetP}(A) = \sum_{A,B\in T_n^\Theta, A\subseteq B} \frac{|A\cap B|}{|B|} m(B) \tag{15}$$

where $|X|$ stands for the cardinality of the element X. In DER, the cardinality of $A \in T_n^\Theta$ is the number of the singleton elements that it contains in its canonical form. For example, $|t_{(\theta_1\cup\theta_2),\theta_3}| = |t_{\theta_1,\theta_3} \cup t_{\theta_2,\theta_3}| = 2$.

III. Numerical Examples

We show here how dynamical evidential reasoning works. Its results are compared to those based on the classical fusion rules. Results of DER for change detections in remote sensing images are given in the next section.

Example 3: Let us consider two sources of information in the same frame of discernment $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, and the bbas are given by

	θ_1	θ_2	θ_3	Θ
$\mathbf{m_1}$	0.4	0	0.3	0.3
$\mathbf{m_2}$	0	0.3	0.2	0.5

- **With DSm free model of** Θ: The results obtained are the following:
  1) With DSm classic/conjunctive rule (DSmC)

$$m(\theta_1\cap\theta_2) = 0.12$$
$$m(\theta_1\cap\theta_3) = 0.08$$
$$m(\theta_1\cap\theta_4) = 0.20$$
$$m(\theta_2\cap\theta_3) = 0.09$$
$$m(\theta_2\cap\theta_4) = 0.09$$
$$m(\theta_3\cap\theta_4) = \mathbf{0.21}$$
$$m(\theta_3) = 0.06$$
$$m(\theta_4) = 0.15.$$

  2) With DERf rule: We identify each possible transition, and we get

$$m(t_{1,2}) = m_1(\theta_1)m_2(\theta_2) = 0.12$$
$$m(t_{1,3}) = m_1(\theta_1)m_2(\theta_3) = 0.08$$
$$m(t_{1,4}) = m_1(\theta_1)m_2(\theta_4) = 0.20$$
$$m(t_{3,2}) = m_1(\theta_3)m_2(\theta_2) = 0.09$$
$$m(t_{3,3}) = m_1(\theta_3)m_2(\theta_3) = 0.06$$
$$m(t_{3,4}) = m_1(\theta_3)m_2(\theta_4) = \mathbf{0.15}$$
$$m(t_{4,2}) = m_1(\theta_4)m_2(\theta_2) = 0.09$$
$$m(t_{4,3}) = m_1(\theta_4)m_2(\theta_3) = \mathbf{0.06}$$
$$m(t_{4,4}) = m_1(\theta_4)m_2(\theta_4) = 0.15.$$

  Note that the main difference between DSmC and DERf lies in the ability of DERf to refine the partial conflicts into several distinct transitions states. More precisely, the mass 0.21 of the partial conflict $\theta_3\cap\theta_4$ computed as $m(\theta_3\cap\theta_4) = m_1(\theta_3)m_2(\theta_4) + m_1(\theta_4)m_2(\theta_3) = 0.21$ with classical DSm free rule becomes clearly split into two transitions $t_{3,4} \triangleq \theta_3 \to \theta_4$ and $t_{4,3} \triangleq \theta_4 \to \theta_3$ with associated masses 0.15 and 0.06.

  One sees that, if the DSmC rule is used to detect the change occurrence reflecting conflicting information, $\theta_3\cap\theta_4$ gets the biggest mass which indicates that the transition $\theta_3 \to \theta_4$ or $\theta_4 \to \theta_3$ is most likely to happen. Nevertheless, the most mass (or credibility) is actually committed to θ_1 by source 1 and to θ_4 by source 2. Therefore, the change occurrence from θ_1 to θ_4 is intuitively most likely to occur, which is consistent with the result provided by the DERf rule.

- **With a hybrid model for** Θ: We assume now the following set $\emptyset_\mathcal{M} \triangleq \{t_{3,3}, t_{1,4}, t_{4,1}, t_{2,4}, t_{4,2}\}$ of impossible transitions in the model. The corresponding DSm hybrid model is therefore defined by the following corresponding set of integrity constraints:

$$\emptyset_\mathcal{M} \triangleq \{\theta_3, \theta_1\cap\theta_4, \theta_2\cap\theta_4\}.$$

  1) With the DSm hybrid rule of combination (DSmH) [19] proposed in the DSmT framework as an extension of Dubois–Prade's rule, one gets

$$m(\theta_1\cap\theta_2) = 0.16$$
$$m(\theta_1) = 0.12$$
$$m(\theta_2) = 0.12$$
$$m(\theta_4) = 0.27$$
$$m(\theta_1\cup\theta_4) = 0.12$$
$$m(\theta_2\cup\theta_4) = 0.12$$
$$m(\Theta) = 0.09.$$

2) With the DER_{DS} rule of combination, one gets

$$m(t_{1,2}) = 0.64$$
$$m(t_{4,4}) = 0.36.$$

In this hybrid model, DER_{DS} gets the most specific results about the state transition because of the use of the constraints. DSmH yields to more uncertain results since it works under the assumption that the element for the decision making is invariable in the fusion process, and it tries to overestimate the invariable hypothesis by the redistribution of conflicting beliefs to partial ignorance. Therefore, DSmH appears to be not well adapted for the change detection.

- **With Shafer's model of Θ**: In Shafer's model, all the transitions between different elements are considered impossible to occur.
  1) With Dempster–Shafer's rule, one gets

$$m_{DS}(\theta_3) = 0.3$$
$$m_{DS}(\theta_4) = 0.7.$$

  2) DER_{DS} will provide the same results as

$$m(t_{3,3}) = 0.3$$
$$m(t_{4,4}) = 0.7.$$

  3) With the DER_{Y} rule, one gets

$$m(t_{3,3}) = 0.06$$
$$m(t_{4,4}) = 0.15$$
$$m(t_{\Theta,\Theta}) = m\left(t_{(1,1)\cup(2,2)\cup(3,3)\cup(4,4)}\right) = K = 0.79.$$

This indicates that Shafer's model does not fit well for change detection, and the classical DST approach cannot be used efficiently to detect the change occurrences.

Example 4: Let us see how DER works in a high conflicting situation by analyzing the well-known Zadeh's example. Therefore, we consider the frame $\Theta = \{\theta_1, \theta_2, \theta_3\}$ satisfying Shafer's model with the following Bayesian bbas:

	θ_1	θ_2	θ_3
$\mathbf{m_1}$	0.9	0.1	0
$\mathbf{m_2}$	0	0.1	0.9

- **With Shafer's model, one gets**
  1) DS rule: $m(\theta_2) = 1$.
  2) DER_{DS}: $m(t_{2,2}) = 1$.
  3) DSmH: $m(\theta_2) = 0.01$, $m(\theta_1 \cup \theta_2) = 0.09$, $m(\theta_1 \cup \theta_3) = 0.81$, and $m(\theta_2 \cup \theta_3) = 0.09$
  4) DER_{Y}: $m(t_{2,2}) = 0.01$, and $m(t_{\Theta,\Theta}) = m(t_{(1,1)\cup(2,2)\cup(3,3)}) = 0.99$.

Dempster's rule provides counterintuitive results because of the high conflict in this case. DSmH provides a reasonable result which lies in its proper distribution of the conflicting beliefs, but it cannot well identify and estimate the change occurrences. DER_{DS} gets the absolute state transition because of the constraints. Nevertheless, this absolute result is very risky. DER_{Y} commits all the conflicting mass of belief to the ignorance transition, which yields, of course, to a very imprecise result which can make sense in such high conflicting case. It implies that the prudent distribution of the mass of empty sets by DER_{Y} can be less risky than the optimistic distribution done by DER_{DS} in the potentially high conflicting conditions.

As we can see, the combination results in the hybrid model[2] mainly depend on the given integrity constraints and how the redistribution of the mass corresponding to partial (or total) conflicts is done. Since it is rather difficult to clearly emphasize the differences between DER and DSm in the hybrid model, we propose to analyze and compare DERf with DSmC only in the free model in the following examples.

Example 5: Consider another pair of bbas over the frame $\Theta = \{\theta_1, \theta_2\}$ which is given by

	θ_1	θ_2	$\theta_1 \cup \theta_2$
$\mathbf{m_1}$	0.45	0.2	0.35
$\mathbf{m_2}$	0	1	0

- **With DSm free model**:
  1) DSmC: $m(\theta_1 \cap \theta_2) = 0.45$ and $m(\theta_2) = 0.55$.
  2) DERf: $m(t_{1,2}) = 0.45$, $m(t_{2,2}) = 0.2$, and $m(t_{1\cup 2,2}) = 0.35$.

Source 1 commits most of its belief to θ_1, whereas source 2 is absolutely confident in θ_2. It is quite possible that θ_1 changes to θ_2, and the results produced by DERf reflect the intuitive reasoning. Nevertheless, the results of DSmC show that θ_2 is the most credible since two specific transition states in DERf are linked together to a focal element in Dezert-Smarandache free model as follows: $m(\theta_2) = m(t_{2,2}) + m(t_{1\cup 2,2}) = 0.55$. For the singleton elements, one gets

	Bel$(\cdot)$	BetP$(\cdot)$	Pl$(\cdot)$
$t_{1,2}$	0.45	0.625	0.8
$t_{2,2}$	0.2	0.375	0.55

From Bel$(\cdot)$, Pl$(\cdot)$, or BetP$(\cdot)$ values, we can draw the conclusion for the decision making that the most relevant (credible, plausible, or possible) hypothesis is $t_{1,2}$.

Example 6: Let us consider a pair of bbas over $\Theta = \{\theta_1, \theta_2\}$ given by

	θ_1	θ_2	$\theta_1 \cup \theta_2$
$\mathbf{m_1}$	0.6	0	0.4
$\mathbf{m_2}$	0	1	0
$\mathbf{m_3}$	0	0.5	0.5

- **With DSm free model**
  1) DSmC: $m(\theta_1 \cap \theta_2) = 0.6$ and $m(\theta_2) = 0.4$.
  2) DERf: $m(t_{1,2,2}) = 0.3$, $m(t_{1,2,1\cup 2}) = 0.3$, $m(t_{1\cup 2,2,1\cup 2}) = 0.2$, and $m(t_{1\cup 2,2,2}) = 0.2$.

For the singleton elements, based on DERf, one gets

	Bel$(\cdot)$	BetP$(\cdot)$	Pl$(\cdot)$
$t_{1,2.2}$	0.3	0.60	1
$t_{1,2,1}$	0	0.20	0.5
$t_{2,2,1}$	0	0.05	0.2
$t_{2,2,2}$	0	0.15	0.4

[2]Shafer's model is just a special case of hybrid model [17]–[19].

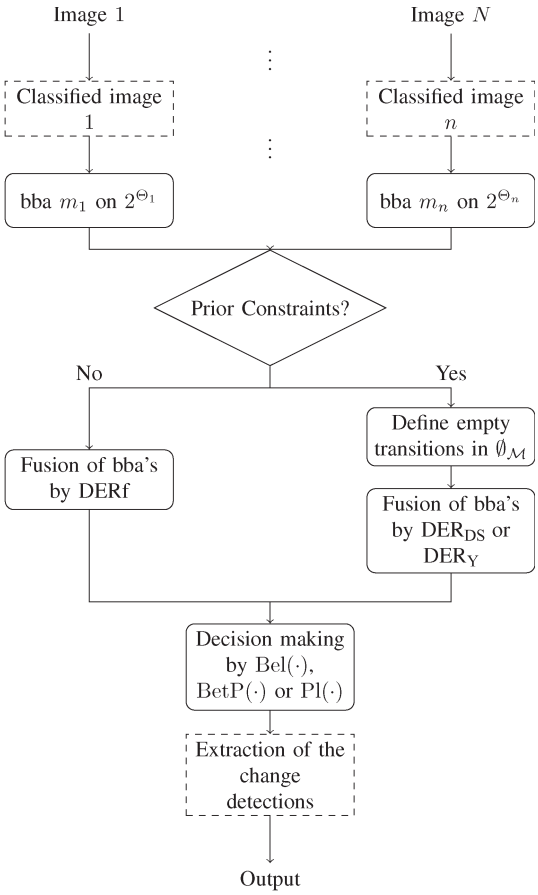


Fig. 1. DER flowchart for change detection of remote sensing images.

The results of DERf precisely represent the belief of change occurrences, whereas the elements in DSmC cannot reflect the process of state transition because of its invariable frame. In the decision making, it indicates that the hypothesis $t_{1,2,2}$ is most likely to happen according to Bel$(\cdot)$, Pl$(\cdot)$, or BetP$(\cdot)$, which means that θ_1 in source 1 (at time 1) changes to θ_2 in source 2 and to θ_2 in source 3.

IV. Application on Real Remote Sensing Images

The application of DER, as a postclassification-based change detection technique from remote sensing images, is presented with the flow chart of Fig. 1.

The process can be summarized as follows: from a time series of N images (for which geometric correction, coregistration, and radiometric correction may be required).

1) Perform an appropriate classification (supervised or unsupervised) on each image.
2) Apply an appropriate mass assignment from the classification outputs. It may be of Bayesian type, distance type, and membership value and has to be transformed into mass type $m_k(X_l)$, including uncertainty (i.e., Θ).
3) Define a (possibly hybrid) frame of discernment T_n^{Θ}. At this step, the class index has to be carefully identified to well define transitions $t_{X_1,\ldots,X_n}$.
4) Apply the DER fusion rule [(7) or (9)], and use Bel, Pl, or BetP criteria for decision making.

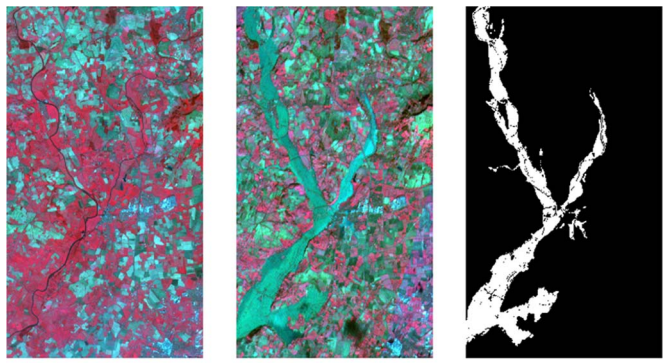


Fig. 2. SPOT images of Gloucester. (a) Before image, acquired on 11/16/1999. (b) After image, acquired on 10/21/2000. (c) Ground truth.

TABLE I
Class Description of the Gloucester SPOT Images

Class	Color in Fig. 2	description
θ_1	Red area	covered field
θ_2	Dark-red area	wooded area
θ_3	Green area	bare soil
θ_4	Dark-green area	senescence of vegetation
Θ		Ignorance

It has been analyzed in theory (in Section II) and shown by the numerical examples (in Section III) that DSmT does not work well for change detection, particularly when the number of sources is larger than 2. Therefore, only DER will be applied in the next two experiments. In those experiments, an unsupervised clustering method has been used. We choose the evidential c-means (ECM) approach, detailed in [30], which uses the radiometric information only. ECM is adapted to the classification of uncertain data, and the imprecise classes can be acquired by ECM. This classification technique has been used here because it is unsupervised, and the results can be directly used as the mass functions (bbas). Nevertheless, any kind of classifiers (including supervised ones) may be used at this level. Moreover, all the other available features which may be based on radiometric [31], morphological, or textural features, as well as decision models based on Bayesian [32], support vector machine [10], [33], [34], or object-based technique [11], [35], may of course be applied in this image classification level. More features are expected to lead to the derivation of better classification results with ECM, but it will increase the computation complexity. In this paper, we want to focus on the way to apply DER after the image classification.

A. Change Detection With a Pair of SPOT Images

In this experiment, a pair of SPOT images is considered, corresponding to a flood over Gloucester, U.K., from October to November 2000 shown in Fig. 2.

The feature space for classification is defined by the three spectral bands of the SPOT image. In addition, the texture information has been taken into account by using the variance of the near infrared (NIR) through a 3×3 sliding window. The number of the clusters is automatically determined by minimizing the validity index of a credal partition as the average normalized specificity proposed in [30], and both the SPOT images are clustered into $k = 4$ groups using ECM as defined in Table I.

TABLE II
TRANSITION DESCRIPTION OF THE GLOUCESTER SPOT IMAGES

Class of image 1	Class of image 2
θ_1	θ_1
θ_2	θ_2
θ_3	θ_3
θ_4	θ_4
$T^{\Theta}_{n=2} \triangleq \{t_{1,1}, t_{2,2}, t_{3,3}, t_{4,4}, t_{1,3}\}$	

Remark: The number of the clusters can be automatically determined by the method in [30] or directly given by the user according to the prior experience. If the number of the clusters is too high, the same object (content) whose pixel values may be only slightly different in one image could be in a different cluster, which will generate problems for the definitions of the state transitions, including the change occurrences. Moreover, the high number of clusters will induce an increase of the algorithm complexity.

The other tuning parameters to be defined in ECM are the following: $\alpha = 1$, $\beta = 2$, $\sigma^2 = 20$, termination threshold $\epsilon = 1$, and the maximum number of iterations $T = 10$. The normalized membership is used as the mass functions (bbas).

According to our prior knowledge that a flood happened between the two images, the change occurrences mainly focus on the flood, which corresponds to the transitions from $\theta_1 = \{\text{Red area}\}$ to $\theta_3 = \{\text{Green area}\}$ in the images as $t_{1,3}$. Therefore, the constrained model can be applied here. The available transitions are given by the set $T^{\Theta}_{n=2} = \{t_{1,1}, t_{2,2}, t_{3,3}, t_{4,4}, t_{1,3}\}$, and all the other transitions are considered as empty transitions in $\emptyset_{\mathcal{M}}$ as shown in Table II.

Remark: The labels of each class in each image are possibly different when using the unsupervised classification methods (i.e., ECM). Nevertheless, once the corresponding relationship between the label and the class is given, the state transitions including the change occurrences can be defined (tuned) accordingly. Therefore, the labels do not affect the change detection process at all. In this paper, we denoted the same class in each image with the same label for the convenience of the notations and representations.

The maximum of pignistic probability has been used in the decision making. The consequence is that DER_{DS} and DER_Y yield the same decision. Their fused image with all the allowed transitions of $T^{\Theta}_{n=2}$ is presented in Fig. 3(a). In this example, we consider that the flood area is characterized by the transition $t_{1,3}$. Fig. 3(b) focuses on this transition and, by using the ground truth of Fig. 2(c), shows in false color the correct, missed, and false detections.

Correct detections, missed detections, and false alarms have been described by the Kappa index, which is often used to demonstrate the quality of classifications for change detection [10], [11], [36]. With the DER approach, we get a Kappa index of 0.82, which indicates that the output is mainly consistent with the ground truth, so the total results are satisfying for change detection applications.

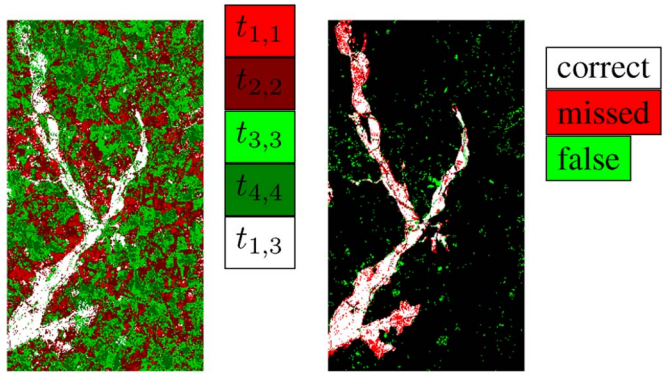


Fig. 3. (a) Fusion results of the pair of SPOT images. (b) Fusion results with all transitions. The comparison of the change detections with ground truth (Kappa $= 0.82$).

One can see that the results of the fusion of images not only accurately detect the change occurrences due to the flood but also produce accurate classifications in the unchanged areas. However, there remain some missed detections and false alarms. As we can see, some regions along the river were displayed in green color before the flood, and they are still in green color after the flood. Therefore, they have not been considered as change occurrences, and this mainly leads to the missed detections, when the detection is focused on the radiometry only. The false alarms are mainly due to the change of land cover in some small areas which are similar to the changes in the flooded areas in the second image. If better prior information about the location of the river is known, the prior probability about the transitions can be given accordingly, and the false alarms could be reduced.

B. Change Detection in a Three-Image Time Series

This section focuses on the event of the 2003 earthquake in the region of Boumerdes, Algeria. Three QuickBird images have been acquired before and after the earthquake. They have been corrected geometrically by using Shuttle Radar Topography Mission elevation data and resampled by a P+Xs pan-sharpening technique to yield 60-cm resolution color images (see Fig. 4) for operational use during this event on behalf of the "Space and Major Disaster" charter [37].

This case study is used to illustrate the ability of DER in dealing with more than two images which are classified with different numbers of classes. The automatic determination of the number of clusters in [30] is still applied here, and the number of clusters in the first image has been evaluated to $C = 4$ groups (see details in Table III); we still take into account the texture information defined by the variance of the NIR by a 3×3 window and the pixel values as the features for classifications in this application. The classified image is shown in Fig. 4(a). The *second* and *third* images are clustered in $C = 5$ groups as defined in Table IV and shown in Fig. 4(b) and (c). The other tuning parameters of the ECM are defined by the following: the maximum number of iterations $T = 20$, the weighting exponent for cardinality $\alpha = 2$, the weighting exponent $\beta = 2$, and the termination threshold $\varepsilon = 1$. The maximum of BetP is used for decision making.

(a)

(b)

(c)

Fig. 4. Multitemporal QuickBird satellite images acquired before and after an earthquake. (a) Before image: 04/22/2002. (b) After image: 05/13/2003. (c) Latter acquisition: 06/13/2003. Data set Boumerdes Copyright SERTIT, 2009, distribution CNES.

TABLE III
CLASS DESCRIPTIONOF THE QUICKBIRD 04/22/2002 IMAGE

Class	Color in Fig. 4	Description
θ_1	Dark area	green plant or shadow
θ_2	Maroon area	farmland
θ_3	Gray area	normal building or road
θ_4	Bright area	bright bared earth
Θ		Ignorance

The classification labels are not very specific since the classification is based on the radiometry only. Different objects in the same image may share similar pixel values. For example, in Fig. 4(a), buildings and roads are confused in θ_3 since the radiometry is similar. On the contrary, a class may appear differently along the time series: For instance, roads around the new building areas, in different images of Fig. 4(a) and (b), appear with a different hue which causes a change in the image classification.

TABLE IV
CLASS DESCRIPTION OF THE QUICKBIRD 05/13/2003 AND 06/13/2003 IMAGES

Class	Color in Fig. 4	Description
θ_1	Dark area	green plant or shadow
θ_2	Maroon area	farmland
θ_3	Gray area	normal building, road or gray soil
θ_4	Bright yellow area	bright bared earth
θ_5	White area	destroyed building or temporary building
Θ		Ignorance

TABLE V
TRANSITION DESCRIPTIONOF THE BOUMERDES QUICKBIRD IMAGES

Class of image 1	Class of image 2	Class of image 3
θ_1	θ_1	θ_1 ($t_{1,1,1}$)
θ_2	θ_2	θ_2 ($t_{2,2,2}$)
θ_3	θ_3	θ_3 ($t_{3,3,3}$)
		θ_4 ($t_{4,4,4}$)
θ_4	θ_4	
		θ_5 ($t_{2,5,5}$)
		($t_{3,5,5}$)

$T^{\Theta}_{n=3} \triangleq \{t_{1,1,1}, t_{2,2,2}, t_{3,3,3}, t_{4,4,4}, t_{2,5,5}, t_{3,5,5}, t_{3,3,5}\}$

The time series analysis is considered for abrupt change detection (ignoring shade in the evolution) in order to focus on the affected area. Then, a constrained model of DER is applied. The state transitions considered from the image in Fig. 4(a) to (b) include the change occurrences from the farmland to temporary building as $t_{2,5}$ (for the notation $t_{i,j} = t_{\theta_i,\theta_j}$) and from normal buildings to the destroyed buildings as $t_{3,5}$ and all unchanged transitions as $t_{1,1}$, $t_{2,2}$, $t_{3,3}$, and $t_{4,4}$. The change occurrences from normal buildings to destroyed buildings as $t_{3,5}$ and the unchanged transitions as $t_{1,1}$, $t_{2,2}$, $t_{3,3}$, $t_{4,4}$, and $t_{5,5}$ occur from the images in Fig. 4(b) to (c). Therefore, the constrained available transitions in the three images are given by $t_{1,1,1}$, $t_{2,2,2}$, $t_{3,3,3}$, $t_{4,4,4}$, $t_{2,5,5}$, $t_{3,5,5}$, and $t_{3,3,5}$ as detailed in Table V.

All the other possible transitions which are considered not applicable are defined as empty sets through constrained model in DER.

Remark: Some specific transitions cannot exist in the temporal behavior of some classes. For instance, building roof, road, and destroyed or temporary building may be considered as similar from the radiometric point of view. Even if the transitions $t_{2,5}$ and $t_{3,5}$ may involve many kinds of possible change occurrences, some of them are impossible according to our prior knowledge: Farmland cannot become destroyed building in $t_{2,5}$, and only the specific transition from farmland to temporary building is possible. In $t_{3,5}$, only the change occurrence from normal building to destroyed building may occur, and the other combinations are impossible according to the prior knowledge that one has on the problem. Therefore, it yields the temporal interpretation of Table VI.

The combination results are shown in Fig. 5. Change occurrences are considered more important than the invariable

TABLE VI
MEANING OF THE TRANSITION

Transition in $T^{\Theta}_{n=2}$	interpretation
$t_{2,5}$	farmland → building
$t_{3,5}$	normal building → destroyed building
$t_{2,5,5}$	farmland → building → building
$t_{3,5,5}$	normal building → destroyed building → destroyed building
$t_{3,3,5}$	normal building → normal building → destroyed building

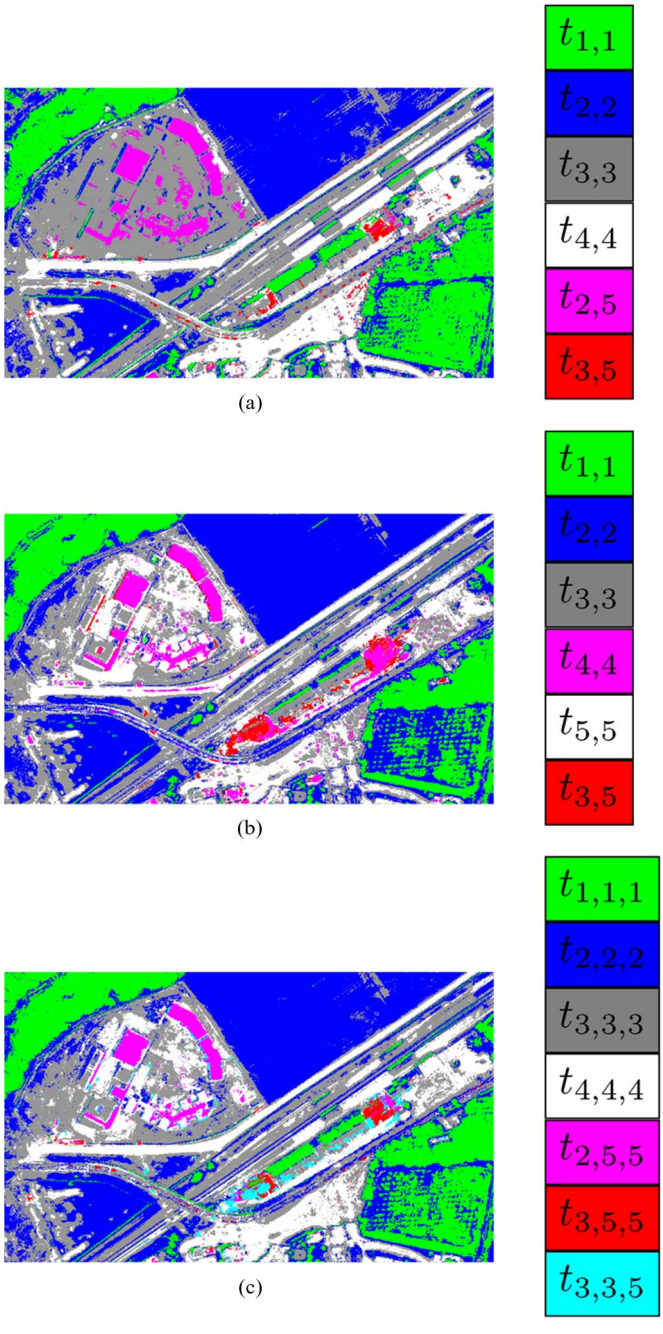


Fig. 5. Fusion results of classified images by DER_{DS} (same results are obtained by DER_Y). (a) Fusion of the first and the second image. (b) Fusion of the second and the third image. (c) Fusion of the first, second, and third images.

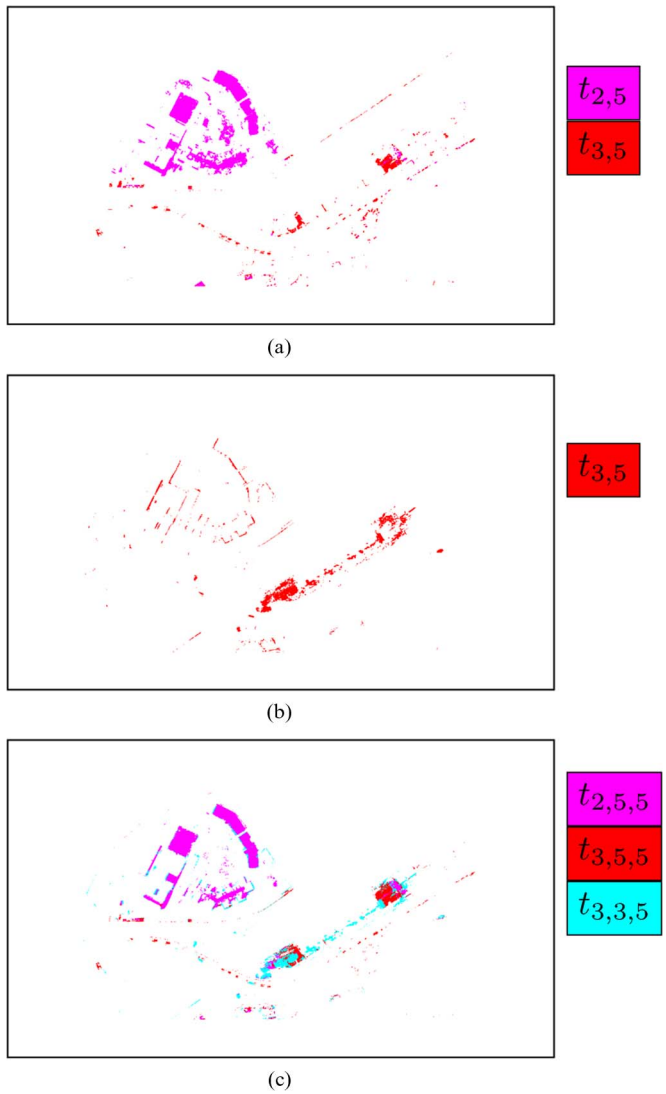


Fig. 6. Significant damage map extracted from DER decision of Fig. 5. (a) Changes from the first to the second classified image. (b) Changes from the second to the third classified image. (c) Changes through the first, second, and third classified images.

transitions for the evaluation of the disaster, and they have been extracted in Fig. 6. The transitions $t_{2,5}$ and $t_{3,5}$ between the first and the second images correspond to actual changes, and they are linked to damage mapping; see Fig. 5(a). Fig. 6(a) focuses on those changes. As we can see, part of the farmland has been converted to temporary building which may be a refuge after the earthquake or new constructions represented by purple color. On the contrary, in the middle of the image, a normal building became a destroyed building represented by the red area in Fig. 6(a). What is interesting in the focus on the $t_{3,5}$ transition is the ability to show a misalignment in the building in the center of the image. It highlights the fact that the buildings (at least one of them) have been affected by the earthquake (up to a grade 3 or 4 in the masonry building damage classification of EMS-98 [38]) but did not collapse. In the ground truth available from Centre National d'Etudes Spaciales and Service Régional de Traitement d'Image et de Télédétection, the level of destruction is not given, but we know that those buildings have been affected.

The transitions $t_{3,5}$ from the second to the third image in Fig. 5(b) correspond to changes that occurred after the earthquake. A building collapsed later after the earthquake while the first collapsed building had been excavated. Fig. 6(b) focuses on

this transition only. It shows that there are some change occurrences in front of those buildings which is evident because there was no path along these buildings before the third image was taken, but the path emerged on the third image. The emergence of the path could not be distinguished from the change of the building to the destroyed building since the path had the same white pixel values as the destroyed building. Fig. 6(c) shows the sequential transitions of the three images. These fusion results will be helpful for the disaster evaluation in real applications.

In the fusion process of DER, some change occurrences are possibly caused by the differences in the geometry of the acquisitions. These changes associated with the geometry errors (i.e., from building to shadow) can be defined with prior experience in a similar way as in [39], and they will be considered as noisy changes in the fusion results. These noisy changes would be removed to reduce the influence of geometric errors. Of course, if some ancillary information about the sensor geometry and elevation models is available for postprocessing, the noisy changes can be significantly reduced.

V. Conclusion

A DER approach has been proposed in this paper for state-transition estimation and has been applied for change detection purpose on real multiple temporal remote sensing images. The DER approach is a postclassification-based change detection technique that starts with the sequential construction of the power set of admissible state transitions taking into account, if necessary, some integrity constraints representing some known unacceptable (impossible) transitions. When the prior knowledge of the constraints of the space of transitions is unknown, the free dynamical rule DERf can be used to compute the mass of belief of the state transitions with a greater computation burden. If some change occurrence constraints are known, the corresponding constrained model must be chosen for better fusion results with less computational complexity. Based on a particular rule-based algebra, the mass of belief of the change occurrences can be acquired using two different rules of combinations in red, with the constrained model depending on the degree of conflict between sources: the $\mathrm{DER}_{\mathrm{DS}}$ rule or the $\mathrm{DER}_{\mathrm{Y}}$ rule, which are direct extensions of the Dempster–Shafer's or Yager's rule in the particular context of state-transition estimation. We have also shown in this paper how DER can be improved for taking into account the prior probability of the state transitions. Several simple numerical examples were given to show how to use DER and to show its difference from the classical fusion approaches. Finally, two experiments on the fusion of multitemporal satellite images illustrate the suitability and the efficiency of DER for change detection and estimation. The DER fusion of images can detect accurately the change occurrences and characterize the unchanged areas. The possible extension of this work will concern the redefinition of DER for working with heterogeneous data.

Acknowledgment

The authors would like to thank the anonymous reviewers for their remarks which helped improve the quality of this paper.

References

[1] A. Bouakache, A. Belhadj-Aissa, and G. Mercier, "Satellite image fusion using Dezert–Smarandache theory," in *Advances and Applications of DSmT for Information Fusion*, F. Smarandache and J. Dezert, Eds. Rehoboth, DE: Amer. Res. Press, 2009, ch. 22.

[2] G. Mercier, G. Moser, and S. Serpico, "Conditional copula for change detection on heterogeneous SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1428–1441, May 2008.

[3] J. Inglada and G. Mercier, "A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1432–1446, May 2007.

[4] F. Bovolo and L. Bruzzone, "A theoretical framework for unsupervised change detection based on change vector analysis in polar domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 218–236, Jan. 2007.

[5] G. Mercier, S. Derrode, E. Trouvé, and L. Bombrun, "Change detection in remote sensing observation," in *Multivariate Image Processing: Methods and Applications*, J. Chanussot, K. Chehdi, and C. Collet, Eds. Paris, France: Hermes, 2009.

[6] G. Mercier, "Progressive change detection in time series of SAR images," in *Proc IEEE IGARSS*, 2010, pp. 3086–3089.

[7] A. Robin, G. Mercier, G. Moser, and S. Serpico, "An a-contrario approach for unsupervised change detection in radar images," in *Proc. IEEE IGARSS*, 2009, pp. IV-240–IV-243, Invited paper in the session "Time-series analyses for change detection".

[8] G. Moser, S. B. Serpico, and G. Vernazza, "Unsupervised change detection from multichannel SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 2, pp. 278–282, Apr. 2007.

[9] F. Pacifici, F. Del Frate, C. Solimini, and W. Emery, "An innovative neural-net method to detect temporal changes in high-resolution optical satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 9, pp. 2940–2952, Sep. 2007.

[10] F. Bovolo, L. Bruzzone, and M. Marconcini, "A novel approach to unsupervised change detection based on a semisupervised SVM and a similarity measure," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2070–2082, Jul. 2008.

[11] F. Bovolo, "A multilevel parcel-based approach to change detection in very high resolution multitemporal images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 1, pp. 33–37, Jan. 2009.

[12] G. Moser and S. Serpico, "Unsupervised change detection from multichannel SAR data by Markovian data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2114–2128, Jul. 2009.

[13] Y. Bazi, F. Melgani, and H. Al-Sharari, "Unsupervised change detection in multispectral remotely sensed imagery with level set methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 8, pp. 3178–3187, Aug. 2010.

[14] T. Celik, "Change detection in satellite images using a genetic algorithm approach," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 386–390, Apr. 2010.

[15] C. Huo, Z. Zhou, H. Lu, C. Pan, and K. Chen, "Fast object-level change detection for VHR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 118–122, Jan. 2010.

[16] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton Univ. Press, 1976.

[17] F. Smarandache and J. Dezert, *Advances and Applications of DSmT for Information Fusion (Collected works)*. Rehoboth, DE: Amer. Res. Press, 2009. [Online]. Available: http://www.gallup.unm.edu/~smarandache/DSmT-book3.pdf

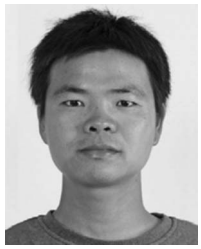
[18] F. Smarandache and J. Dezert, *Advances and Applications of DSmT for Information Fusion (Collected works)*. Rehoboth, DE: Amer. Res. Press, 2006. [Online]. Available: http://www.gallup.unm.edu/~smarandache/DSmT-book2.pdf

[19] F. Smarandache and J. Dezert, *Advances and Applications of DSmT for Information Fusion*. Rehoboth, DE: Amer. Res. Press, 2004. [Online]. Available: http://www.gallup.unm.edu/~smarandache/DSmT-book1.pdf

[20] S. Corgne, L. Hubert-Moy, J. Dezert, and G. Mercier, "Land cover change prediction with a new theory of plausible and a pradoxical reasoning," in *Advances and Applications of DSmT for Information Fusion*, F. Smarandache and J. Dezert, Eds. Rehoboth, DE: Amer. Res. Press, Jun. 2004.

[21] S. L. Hegarat-Mascle, I. Bloch, and D. Vidal-Madjar, "Application of Dempster–Shafer evidence theory to unsupervised classification in multisource remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 4, pp. 1018–1031, Jul. 1997.

[22] S. Hachicha and F. Chaabane, "Application of DSM theory for SAR image change detection," in *Proc. 16th IEEE ICIP*, Nov. 2009, pp. 3733–3736.
[23] W. Liu, "Analyzing the degree of conflict among belief functions," *Artif. Intell.*, vol. 170, no. 11, pp. 909–924, Aug. 2006.
[24] A. Martin, A. L. Jousselme, and C. Osswald, "Conflict measure for the discounting operation on belief functions," in *Proc. Fusion*, Cologne, Germany, 2008, pp. 1–8.
[25] C. K. Murphy, "Combining belief functions when evidence conflicts," *Decision Support Syst.*, vol. 29, no. 1, pp. 1–9, Jul. 2000.
[26] P. Smets, "Decision making in the TBM: The necessity of the pignistic transformation," *Int. J. Approx. Reason.*, vol. 38, no. 2, pp. 133–147, Feb. 2005.
[27] P. Smets, "The combination of evidence in the transferable belief model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 5, pp. 447–458, May 1990.
[28] P. Smets and R. Kennes, "The transferable belief model," *Artif. Intell.*, vol. 66, pp. 191–234, 1994.
[29] R. R. Yager, "On the relationships of methods of aggregation of evidence in expert systems," *Cybern. Syst.*, vol. 16, pp. 1–21, 1985.
[30] M. H. Masson and T. Denoeux, "ECM: An evidential version of the fuzzy c-means algorithm," *Pattern Recognit.*, vol. 41, no. 4, pp. 1384–1397, 2008.
[31] J. Chen, X. Chen, X. Cui, and J. Chen, "Change vector analysis in posterior probability space: A new method for land cover change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 2, pp. 317–321, Mar. 2011.
[32] C. Benedek and T. Sziranyi, "Change detection in optical aerial images by a multilayer conditional mixed Markov model," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 10, pp. 3416–3430, Oct. 2009.
[33] F. Pacifici and F. Del Frate, "Automatic change detection in very high resolution images with pulse-coupled neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 58–62, Jan. 2010.
[34] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Rojo-Alvarez, and M. Martinez-Ramon, "Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1822–1835, Jun. 2008.
[35] C. Huo, Z. Zhou, H. Lu, C. Pan, and K. Chen, "Fast object-level change detection for VHR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 118–122, Jan. 2010.
[36] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, Apr. 1960.
[37] R. Andréoli, H. Yésou, N. Tholey, B. Allenbach, and S. Clandillon, Exploitation en Crise et Post Crise de Données Satellites Haute et Trés Haute Résolution Pour la Cartographie de Dégâts de Séismes, Montpellier, France, Mar. 10–11, 2005, cas de Bam, Boumerdés et Al Hoceima.
[38] G. Gruenthal, Ed., *European Macroseismic Scale 1998 (EMS-98)*. Luxembourg: Cahier du centre européen de Geodynamique et de Séismologie 15, 1998.
[39] S. Marchesi, F. Bovolo, and L. Bruzzone, "A context-sensitive technique robust to registration noise for change detection in VHR multispectral images," *IEEE Trans. Image Process.*, vol. 19, no. 7, pp. 1877–1889, Jul. 2010.



Zhun-ga Liu was born in China in 1984. He received the B.S. and M.S. degrees from Northwestern Polytechnical University, Xi'an, China, in 2007 and 2010, respectively, where he is currently working toward the Ph.D. degree.

He has been studying in Telecom Bretagne, Brest, France, since August 2010, for one year. His research work focuses on belief function theory and its application in remote sensing image processing.

Jean Dezert was born in France in 1962. He received the electrical engineering degree from Ecole Française de Radioélectricité Electronique and Informatique, Paris, France, in 1985, the D.E.A. degree from the University Paris VII, Paris, in 1986, and the Ph.D. degree from the University Paris XI, Orsay, France, in 1990.

Since 1993, he has been a Senior Research Scientist with the Information Modeling and Processing Department, ONERA, Palaiseau, France. His current research interest focuses on belief function theory, particularly for DSmT which has been developed by him and Prof. Smarandache.

Grégoire Mercier was born in France in 1971. He received the Engineering degree from Institut National des Télécommunications, Évry, France, in 1993, and the Ph.D. degree from the University of Rennes, Rennes, France, in 1999.

Since 1999, he has been with Telecom Bretagne, Brest, France, where he is currently a Professor in the Image and Information Processing Department. His research interests include remote sensing image processing and information fusion.

Dr. Mercier is an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS and the President of the French Chapter of the IEEE Geoscience and Remote Sensing Society.

Quan Pan was born in China in 1961. He received the B.S. degree from Huazhong University of Science and Technology, Wuhan, China, and the M.S. and Ph.D. degrees from Northwestern Polytechnical University (NWPU), Xi'an, China, in 1991 and 1997, respectively.

Since 1998, he has been a Professor with NWPU. His current research interests include image processing and information fusion.