International Journal of Approximate Reasoning 74 (2016) 13–29

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Dempster–Shafer fusion of evidential pairwise Markov fields

Mohamed El Yazid Boudaren [a,*], Lin An [b,c], Wojciech Pieczynski [c]

[a] Ecole Militaire Polytechnique, PO Box 17, Bordj El Bahri, 16111 Algiers, Algeria
[b] National Lab. of Radar Signal Processing, Xidian University, Xi'an 710071, China
[c] SAMOVAR, Télécom SudParis, CNRS, Université Paris-Saclay, 9 rue Charles Fourier, 91000 Evry, France

A R T I C L E I N F O

Article history:
Received 4 May 2015
Received in revised form 23 March 2016
Accepted 24 March 2016
Available online 31 March 2016

Keywords:
Hidden Markov fields
Dempster–Shafer fusion
Theory of evidence
Triplet Markov fields

A B S T R A C T

Hidden Markov fields (HMFs) have been successfully used in many areas to take spatial information into account. In such models, the hidden process of interest X is a Markov field, that is to be estimated from an observable process Y. The possibility of such estimation is due to the fact that the conditional distribution of the hidden process with respect to the observed one remains Markovian. The latter property remains valid when the pairwise process (X, Y) is Markov and such models, called pairwise Markov fields (PMFs), have been shown to offer larger modeling capabilities while exhibiting similar processing cost. Further extensions lead to a family of more general models called triplet Markov fields (TMFs) in which the triplet (U, X, Y) is Markov where U is an underlying process that may have different meanings according to the application. A link has also been established between these models and the theory of evidence, opening new possibilities of achieving Dempster–Shafer fusion in Markov fields context. The aim of this paper is to propose a unifying general formalism allowing all conventional modeling and processing possibilities regarding information imprecision, sensor unreliability and data fusion in Markov fields context. The generality of the proposed formalism is shown theoretically through some illustrative examples dealing with image segmentation, and experimentally on hand-drawn and SAR images.

1. Introduction

Let S be a finite set, with $Card(S) = N$, and let $(Y_s)_{s\in S}$ and $(X_s)_{s\in S}$ be two collections of random variables, which will be called "random fields". We assume that Y is observable with each Y_s taking its values in $\mathbb{R}$ (or $\mathbb{R}^d$) whereas X is hidden with each X_s taking its values from a finite set of "classes" or "labels". Such situation occurs in image segmentation problem, which will be used in this paper as an illustrative frame. Realizations of such random fields will be denoted using lowercase letters. We deal with the problem of the estimation of $X = x$ from $Y = y$. Such estimation subsumes the distribution of (X, Y) to be beforehand defined. One classic way to do so is to define, on one hand, the distribution of X, usually called "prior" distribution, and on the other hand, the distribution of Y conditional on X, usually called "noise" distribution. When the prior distribution is Markov, such models are called "hidden Markov fields" (HMFs). These models are of interest as they allow one to find optimal Bayesian solutions, and are successfully used for about forty years [1,2]. HMFs can be extended to "pairwise Markov fields" (PMFs), in which one directly assumes Markovianity of the pair (X, Y) [3], and PMFs have been extended to "triplet Markov fields" (TMFs), in which a third finite discrete valued random field $(U_s)_{s\in S}$ is introduced and

* Corresponding author.
E-mail addresses: boudaren@gmail.com (M.E.Y. Boudaren), lin.an4579@gmail.com (L. An), wojciech.pieczynski@telecom-sudparis.eu (W. Pieczynski).

http://dx.doi.org/10.1016/j.ijar.2016.03.006
0888-613X/

the triplet (U, X, Y) is assumed Markov [4,5]. Finally, TMFs have been extended to "conditional" TMFs (CTMFs), in which one assumes the Markovianity of (U, X) conditional on Y [6]. Likely to HMFs, Bayesian processing can be performed in PMFs, TMFs, and CTMFs as well.

On the other hand, Dempster–Shafer fusion (DS fusion) performed in the frame of "theory of evidence" (TE) allows one to fuse information of different natures [7–12]. The core point used in the paper is that DS fusion can be seen as an extension of the probabilistic computation of the "*a posteriori*" distribution needed in Bayesian processing mentioned above, and thus this processing can be used in a more general setting. Such ideas have already been applied in some special situations. In particular, it has been shown that using DS fusion and Markov field models simultaneously can be of interest [13–17]. Some extensions of the standard HMFs using the theory of evidence are proposed to segment images in [13]. The problem of data fusion of radar and optical images with cloud cover is considered in [14]. Tupin et al. use DS fusion of several structure detectors for automatic interpretation of SAR images [15]. Notice that theory of evidence has also been used within hidden Markov chains for image modeling-related problems. In [18], hidden evidential Markov chains are applied for nonstationary image segmentation. In [19], DS fusion is used to fuse multisensor data in nonstationary Markovian context. Other applications of evidential Markov models include data fusion and image classification [20], power quality disturbance classification [21], particle filtering [22], prognostics [23] and fault diagnosis [24]. Ramasso and Denoeux use belief functions to introduce partial knowledge about hidden states of an HMM [25]. In [16], authors use evidential reasoning to relax Bayesian decisions given by a Markovian classification. The approach is applied to noisy images classification. In [26], a method is developed to prevent hazardous accidents due to operators' action slip in their use of a Skill-Assist. In [27], a second-order evidential Markov model is introduced. Finally, let us mention that the use of imprecise probabilities [28–30] to extend the above models may also be investigated.

The purpose of this paper is to propose a very general family of models providing an original unifying formalism, allowing different known modeling and processing possibilities regarding information imprecision, sensor unreliability and data fusion in the Markov fields context. More precisely:

(i) the proposed family is closed with respect to DS fusion;
(ii) it contains new "conditional evidential Markov models";
(iii) it contains new nonstationary evidential Markov models.

As will be seen, the first point is the core one as it will allow one to perform DS fusion in a very workable manner, by simply adding the corresponding Markov energies. This is of interest because while trying to perform DS fusion in Markov fields in a classic manner one arrives to a non-tractable sum.

Besides this greater generality and the theoretical interest of the proposed extensions, let us mention a specific advantage of a particular new model with respect to the model proposed in [13]. In the case where the noise is complex and its form is not known, the new model makes it possible to approximate the unknown forms of noise through Gaussian mixtures. This may be of practical interest, as shown through some experiments provided in section 4. Indeed, this is all the more of a practical use since parameters can be estimated with "iterative conditional estimation" (ICE) method [4,5], and thus, segmentation can be achieved in the unsupervised context.

Let us notice that a great deal of papers have been published on HMFs, and the same is true on DS information fusion. However, papers dealing simultaneously with both of these topics are relatively rare. Thus this paper is also intended to readers who are used with one of these theories, and not necessarily with the other one. This is why there are some developments, and numerous examples, which could appear as obvious for some readers but of interest for others. Let us mention that an analogous general formalism has been proposed in the frame of Markov chains in [31].

The remainder of this paper is organized as follows: section 2 recalls different Markov field models, the theory of evidence, and its use within particular Markov models. Section 3 describes the proposed evidential pairwise Markov field and its associated theory. Experimental results obtained on hand-drawn and SAR images are provided in section 4. Concluding remarks and future directions are given in section 5.

2. Theory of evidence and hidden Markov fields

This section contains four paragraphs. In the first one, we briefly recall the basics of the theory of evidence. As it also addresses readers possibly ignoring TE, the theory is presented in a simple classic format. The second paragraph is devoted to illustrate the interest of TE in Bayesian classification. The examples presented are rather simple for TE experts; however, they can be of immediate interest to readers familiar with Bayesian image segmentation, specifying different situations they may be faced with. Classic hidden Markov fields, pairwise Markov fields, and triplet Markov fields are recalled in paragraph 3. Finally, in the last fourth paragraph we recall how a simple HMF can be extended to an evidential Markov field using triplet Markov fields.

2.1. Theory of evidence

Let $\Omega = \{\omega_1, \ldots, \omega_K\}$, and let $P(\Omega) = \{A_1, \ldots, A_q\}$ be its associated powerset, with $q = 2^K$. A function M from $P(\Omega)$ to $[0, 1]$ is called a "basic belief assignment" (*bba*) if $M(\emptyset) = 0$ and $\sum_{A \in P(\Omega)} M(A) = 1$. A *bba* M defines then a "plausibility" function Pl from $P(\Omega)$ to $[0,1]$ by $Pl(A) = \sum_{A\cap B\neq\emptyset} M(B)$, and a "credibility" function Cr from $P(\Omega)$ to $[0,1]$ by $Cr(A) = \sum_{B\subset A} M(B)$. For a given *bba* M, the corresponding plausibility function Pl and credibility function Cr are linked by $Pl(A) + Cr(A^c) = 1$, so that each of them defines the other. Conversely, Pl and Cr can be defined by some axioms, and each of them defines then a unique corresponding *bba* M. More precisely, Cr is a function from $P(\Omega)$ to $[0,1]$ verifying $Cr(\emptyset) = 0$, $Cr(\Omega) = 1$, and $Cr(\bigcup_{j\in J} A_j) \geq \sum_{\emptyset\neq I\subset J} (-1)^{|I|+1} Cr(\bigcap_{j\in I} A_j)$, and Pl is a function from $P(\Omega)$ to $[0,1]$ verifying analogous conditions, with $\leq$ instead of $\geq$ in the third one. A credibility function Cr verifying such conditions also is the credibility function defined by the *bba* $M(A) = \sum_{B\subset A} (-1)^{|A-B|} Cr(B)$. Finally, each of the three functions M, Pl and Cr can be defined in an axiomatic way, and each of them defines the two others. Furthermore, a probability function p can be seen as a particular case in which $Pl = Cr = p$.

When two *bbas* M_1 and M_2 represent two pieces of evidence, we can combine, or fuse, them using the so called "Dempster–Shafer fusion" (DS fusion), which gives $M = M_1 \oplus M_2$ defined by:

$$M(A) = (M_1 \oplus M_2)(A) = \begin{cases} 0 & \text{if } A = \emptyset \\ \frac{1}{1-\kappa} \sum_{B_1\cap B_2=A} M_1(B_1)M_2(B_2) & \text{otherwise} \end{cases} \tag{1}$$

where κ measures the amount of conflict between M_1 and M_2:

$$\kappa = \sum_{B_1\cap B_2=\emptyset} M_1(B_1)M_2(B_2).$$

We will say that a *bba* is "Bayesian" or "probabilistic" when, being null outside singletons, it defines a probability and we will say that it is "evidential" otherwise. One can then see that when either M_1 or M_2 is probabilistic (with $\kappa \neq 1$), the fusion result M is also probabilistic.

2.2. Dempster–Shafer fusion and posterior distribution

In image segmentation context of this paper, the interest of evidential modeling approaches extending the Bayesian frame computations stems from the fact that the posterior distribution can be perceived as the DS combination of two – or more – probabilities. This is a crucial point since it opens ways to wider modeling possibilities by extending one or more among such probabilities to belief functions. This fact remains true when dealing with spatially correlated data, which leads to evidential Markov models. For example, in Markov chains framework, even if the DS fusion result is not necessarily Markovian, it has been shown that it defines a marginal distribution of a Markov model, and hence, all the estimations of interest remain workable [32].

To illustrate the interest of such extensions, let us consider the following problem of airborne image segmentation. Let us consider r sensors: $S_1, \ldots, S_r$ providing r observed images $Y^1, \ldots, Y^r$, respectively. In particular, let S_1 be an optical sensor and S_2 be an infrared one. The aim is then to segment the scene $Y = y$ into K classes – or states – by estimating X where each X_s takes its values in the finite set $\Omega = \{\omega_1, \ldots, \omega_K\}$. In the examples bellow, when many sensors are concerned we will take $r = 2$, but the extension to $r > 2$ is straightforward. We will set $K = 4$, thus $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ where $\omega_1, \omega_2, \omega_3$ and ω_4 denote "water", "uninhabited land", "city" and "village" respectively. As a first step, we will limit the frame to a simple context without Markovianity, but we will show in the next section that each of the examples below can be extended to the general Markov context to take into account the spatial information. In all examples bellow, X and Y take their values in Ω and $\mathbb{R}$ (or $\mathbb{R}^2$) respectively (s is removed for the sake of simplicity).

Example 2.1. Let us consider the optical sensor S_1 alone, and let us suppose that our knowledge about the distribution $p(x)$ is $p_1 = p(x = \omega_1) \geq \varepsilon_1, \ldots, p_4 = p(x = \omega_4) \geq \varepsilon_4$ with $\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4 \leq 1$. We see that ε measures the degree of knowledge of $p(x)$ in a "continuous" manner: for $\varepsilon = 1$, the distribution $p(x)$ is perfectly known, and for $\varepsilon = 0$, nothing is known about $p(x)$. Assume that $p(y|x = \omega_1), \ldots, p(y|x = \omega_4)$ are known, and let us consider the distribution $q^y = (q_1^y, q_2^y, q_3^y, q_4^y)$ with

$$q_k^y = \frac{p(y|x = \omega_k)}{\sum_{i=1}^{4} p(y|x = \omega_i)}.$$

Using Bayesian classification to estimate $X = x$ from $Y = y$ requires the knowledge of $p(x|y) \propto p(x)p(y|x)$ which is thus only partly known. How could one use this partial knowledge to perform Bayesian classification? This is made possible by introducing the following *bba* A on $P(\Omega)$: A is null outside $\Delta = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \Omega\}$ and $A[\{\omega_1\}] = \varepsilon_1, \ldots, A[\{\omega_4\}] = \varepsilon_4$, $A[\Omega] = 1 - (\varepsilon_1 + \ldots + \varepsilon_4) = 1 - \varepsilon$. The DS fusion of A with $q^y = (q_1^y, q_2^y, q_3^y, q_4^y)$ gives a probability p^* defined on Ω by

$$p^*(\omega_i) = \frac{(\varepsilon_i + 1 - \varepsilon)q_i^y}{\sum_{j=1}^{4} (\varepsilon_j + 1 - \varepsilon)q_j^y}.$$

Then using p^* to perform the classification allows one to use the partial knowledge of $p(x)$ in a "continuous" manner: perfect knowledge of $p(x)$ corresponds to $\varepsilon = 1$ and indeed, when $\varepsilon = 1$ we have $p^*(x) = p(x|y)$. The case $\varepsilon = 0$ corresponds to the case where $p(x)$ is not known at all and, indeed, this case implies $p^*(x) = q^y(x)$, and the corresponding classification rule is the maximum likelihood classification.

Example 2.2. Let us consider again the optical sensor S_1 with the distribution $p(x)$ on $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ known. Assume that $p(y|x = \omega_1), \ldots, p(y|x = \omega_4)$ are also known; however, there exists an additional class ω_5 corresponding to "clouds" which "hides" the classes of interest forming Ω, and which produces $p(y|x = \omega_5)$. In such situations, one can consider Q^y is null outside $\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \Omega\}$ and

$$Q^y[\{\omega_i\}] = \frac{p(y|x = \omega_i)}{\sum_{j=1}^{5} p(y|x = \omega_j)}, \text{ for } i \in \{1, 2, 3, 4\} \text{ and}$$

$$Q^y(\Omega) = \frac{p(y|x = \omega_5)}{\sum_{j=1}^{5} p(y|x = \omega_j)}.$$

The DS fusion of p with Q^y gives a probability p^* defined on Ω by

$$p^*(\omega_i) = \frac{p(\omega_i)\left[Q^y[\{\omega_i\}] + Q^y(\Omega)\right]}{\sum_{\omega_j \in \Omega} p(\omega_j)\left[Q^y[\{\omega_j\}] + Q^y(\Omega)\right]}$$

This fusion is mathematically similar to that used in *Example 2.1*; however, it models a quite different situation. A Markov extension of such models have been successfully used in cloudy images segmentation in [17].

Example 2.3. Let us consider the infrared sensor S_2 alone and let us assume that the distribution $p(x)$ on $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ is known. As S_2 mainly measures temperature, it can only detect a difference between the urban area (cities and villages) and other classes. Hence, there is a partition of Ω into two subsets $\Omega_1 = \{\omega_1, \omega_2\}$ and $\Omega_2 = \{\omega_3, \omega_4\}$ such that the elements of each subset Ω_i produce a same q_i^y. Then we can consider a *bba* B null outside $\{\Omega_1, \Omega_2\}$ and defined for each Ω_i by

$$B^y(\Omega_i) = \frac{p(y|x \in \Omega_i)}{\sum_{j=1}^{2} p(y|x \in \Omega_j)}$$

The DS fusion of $p(x)$ with B^y gives a probability p^* defined on Ω by

$$p^*(\omega_i) = \frac{p(\omega_i) B^y(\Omega_i \ni \omega_i)}{\sum_{j=1}^{4} p(\omega_j) B^y(\Omega_j \ni \omega_j)}$$

As in *Example 2.1* above, when B^y is the distribution $q^y = (q_1^y, \ldots, q_K^y)$ itself, i. e. each Ω_i is a singleton, p^* is the classic posterior distribution.

Example 2.4. Let us consider the optical sensor S_1 and let us suppose that the distributions $p(x)$ and, $p(y|x = \omega_1), \ldots, p(y|x = \omega_4)$ are known. Let us assume, at a first step, that the knowledge of $p(x)$ is poor. This fact can be taken into account through the *bba* A defined on $\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \Omega\}$ by $A[\{\omega_i\}] = \varepsilon p(\omega_i)$ for $i \in \{1, 2, 3, 4\}$ and $A(\Omega) = 1 - \varepsilon$ as in *Example 2.1*. Considering the distribution q^y defined on Ω by $q_i^y \propto p(y|x = \omega_i)$, the DS fusion of A with q^y gives then a probability p^* which can also be defined on Ω by

$$p^*(\omega_i) \propto \left[A[\omega_i] + A(\Omega)\right] q_i^y.$$

Let us now consider the infrared sensor S_2. The unreliability is then related to the distributions $p(y|x = \omega_i)$ as in *Example* 2.2 and *Example* 2.3. Hence, there is a partition of Ω into two subsets $\Omega_1 = \{\omega_1, \omega_2\}$ and $\Omega_2 = \{\omega_3, \omega_4\}$ such that the elements of each subset Ω_i produce a same q_i^y. Then we can consider a *bba* B null outside $\{\Omega_1, \Omega_2\}$ and defined for each Ω_i by $B^y(\Omega_i) \propto p(y|x \in \Omega_i)$. If $p(x)$ is perfectly known, the DS fusion of $p(x)$ with B^y gives the probability p^* of *Example* 2.3 defined on Ω by $p^*(\omega_i) \propto p(\omega_i) B^y(\Omega_i \ni \omega_i)$. On the other hand, if $p(x)$ is also unreliable, the DS fusion $A \oplus B^y$ is then a *bba* defined on $\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \Omega_1, \Omega_2\}$ by

$$[A \oplus B^y][\{\omega_i\}] \propto A[\{\omega_i\}] B^y(\Omega_j \ni \omega_i) \quad \text{for } i \in \{1, 2, 3, 4\} \text{ and}$$

$$[A \oplus B^y](\Omega_j) \propto A(\Omega) B^y(\Omega_j) \qquad \text{for } j \in \{1, 2\}$$

Example 2.5. Let us now consider both optical sensor S_1 and infrared sensor S_2 and let us assume that they are independent; it is to say, that Y^1 and Y^2 are independent conditional on X. Let us consider that the sensor S_1 provides the observation Y^1 according to the distributions $p(y^1|x = \omega_1), \ldots, p(y^1|x = \omega_4)$ and let us set $q^y = (q_1^y, q_2^y, q_3^y, q_4^y)$ defined on $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ with

$$q_k^y = \frac{p(y^1|x=\omega_k)}{\sum_{i=1}^4 p(y^1|x=\omega_i)}.$$

On the other hand, let us consider that the sensor S_2 produces Y^2 according to the *bba* B^y null outside $\{\Omega_1, \Omega_2\}$ where $\Omega_1 = \{\omega_1, \omega_2\}$ and $\Omega_2 = \{\omega_3, \omega_4\}$ and defined by

$$B^y(\Omega_i) = \frac{p(y^2|x \in \Omega_i)}{\sum_{j=1}^2 p(y^2|x \in \Omega_j)}$$

Assuming all p, q^y and B^y known, the DS fusion $p \oplus q^y \oplus B^y$ gives a probability p^* defined on Ω with

$$p^*(\omega_i) = \frac{p(x=\omega_i)q_i^y B^y(\Omega_i \ni \omega_i)}{\sum_{j=1}^4 p(x=\omega_j)q_j^y B^y(\Omega_j \ni \omega_j)}$$

Example 2.6. Let us consider the same situation of *Example* 2.5 with the knowledge about the distribution $p(x)$ given by $p_1 = p(x=\omega_1) \geq \varepsilon_1, \ldots, p_4 = p(x=\omega_4) \geq \varepsilon_4$ where $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4 = \varepsilon \leq 1$ as in *Example* 2.1. We consider the *bba* A defined on $P(\Omega)$ and null outside $\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \Omega\}$ with $A[\{\omega_1\}] = \varepsilon_1, \ldots, A[\{\omega_4\}] = \varepsilon_4$, $A[\Omega] = 1 - (\varepsilon_1 + \ldots + \varepsilon_4) = 1 - \varepsilon$.

Assuming q^y and B^y defined as in *Example* 2.5 known, the DS fusion $A \oplus q^y \oplus B^y$ gives a probability p^* defined on Ω with

$$p^*(\omega_i) = \frac{\left[A[\{\omega_i\}] + A(\Omega)\right]q_i^y B^y(\Omega_i \ni \omega_i)}{\sum_{j=1}^4 \left[A[\{\omega_i\}] + A(\Omega)\right]q_j^y B^y(\Omega_j \ni \omega_j)}$$

Example 2.7. Let us consider again the situation of *Example* 2.5 with sensor S_2 exhibiting an additional class ω_5 and producing $p(y^2|x=\omega_5)$ as in *Example* 2.2. For this purpose, let us consider Q^y null outside $\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \Omega\}$ with

$$Q^y[\{\omega_i\}] = \frac{p(y^2|x=\omega_i)}{\sum_{j=1}^5 p(y^2|x=\omega_j)}, \text{ for } i \in \{1,2,3,4\} \text{ and}$$

$$Q^y[\Omega] = \frac{p(y^2|x=\omega_5)}{\sum_{j=1}^5 p(y^2|x=\omega_j)}.$$

Assuming B^y, defined as in *Example* 2.5, and p known, the DS fusion $p \oplus Q \oplus B$ gives a probability p^* defined on Ω by

$$p^*(\omega_i) = \frac{p(\omega_i)\left[Q^y[\{\omega_i\}] + Q^y(\Omega)\right]B^y(\Omega_i \ni \omega_i)}{\sum_{\omega_j \in \Omega} p(\omega_j)\left[Q^y[\{\omega_j\}] + Q^y(\Omega)\right]B^y(\Omega_j \ni \omega_j)}$$

Example 2.8. Let us again consider the two sensors of *Example* 2.5 and let us assume now that the knowledge of $p(x)$ is poor as in *Example* 2.4. One can then consider the *bba* A defined on $\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \Omega\}$ by $A(\{\omega_i\}) = \varepsilon p(\omega_i)$ for $i \in \{1,2,3,4\}$ and $A(\Omega) = 1 - \varepsilon$. Considering q^y and B^y known, the DS fusion $A \oplus q^y \oplus B^y$ gives a probability p^* defined on Ω by

$$p^*(\omega_i) \propto \left[A[\{\omega_i\}] + A(\Omega)\right]q_i^y B^y(\Omega_j \ni \omega_i)$$

Remark 2.1. The examples described above correspond to some standard situations that can be further extended. The associated modeling and processing techniques are straightforward. For instance, one may know the exact ratios of "water" and "land" of the considered region, associated to subsets $\{\omega_1\}$ and $\{\omega_2, \omega_2, \omega_3\}$ respectively, whereas the proportions of each of the elements of $\{\omega_2, \omega_3, \omega_4\}$ can be imprecise. This fact is a particular case of the situation of *Example* 2.1. Some of the considered examples can also be blended like the situations of Examples 2.1 and 2.3.

2.3. Hidden, pairwise and triplet Markov fields

In hidden Markov fields (HMFs) context, the field $X = (X_s)_{s \in S}$ is assumed Markovian with respect to a neighborhood system $\mathcal{N} = (\mathcal{N})_{s \in S}$. X is then called a Markov random field (MRF), and its distribution verifies

$$p\left(X_s = x_s|(X_t)_{t \in S, t \neq s}\right) = p\left(X_s = x_s|(X_t)_{t \in \mathcal{N}_s}\right) \tag{2}$$

Under some conditions usually assumed in digital imagery, the Hammersley–Clifford theorem [1] establishes the equivalence between an MRF, verifying (2) with respect to the neighborhood system $\mathcal{N}$, and a Gibbs field with potentials associated with $\mathcal{N}$. Such potentials, describing the elementary relationships within the neighborhood, are computed with respect to the system of cliques C, where a clique $c \in C$ is a subset of S which is either a singleton or a set of pixels mutually neighbors with respect to $\mathcal{N}$. Setting $x_c = (x_s)_{s\in c}$, $\phi_c(x_c)$ denotes the potential associated to the clique c.

Finally, the distribution of X is given by

$$p(X = x) = \gamma \exp\left[-\sum_{c\in C} \phi_c(x_c)\right] \tag{3}$$

where γ is a normalizing constant, which is impossible to compute in practice given the very high number of possible configurations K^N. The quantity $E(x) = \sum_{c\in C} \phi_c(x_c)$ is called "energy". Then, the local conditional probability (2) is

$$p\left(X_s = x_s | (X_t)_{t\in S, t\neq s}\right) = \gamma_s \exp\left[-E_s(x_s)\right]$$

where $E_s(x_s) = \sum_{c\ni x_s} \phi_c(x_c)$, and γ_s is a computable normalizing constant. To define the distribution of Y conditional on X, two assumptions are usually set:

(i) the random variables $(Y_s)_{s\in S}$ are independent conditional on X;
(ii) the distribution of each Y_s conditional on X is equal to its distribution conditional on X_s.

When these two assumptions hold, the noise distribution is fully defined through K distributions $(f_i)_{1\leq i\leq K}$ on $\mathbb{R}$ where f_i denotes the density, with respect to the Lebesgue measure on $\mathbb{R}$, of the distribution of Y_s conditional on $X_s = \omega_i$: $p(Y_s = y_s | x_s = \omega_i) = f_i(y_s)$. Then we have

$$p(Y = y | X = x) = \prod_{s\in S} f_{x_s}(y_s) \tag{4}$$

that can equivalently be written as

$$p(Y = y | X = x) = \exp\left[\sum_{s\in S} \log f_{x_s}(y_s)\right] \tag{5}$$

Since $p(x, y) = p(x)p(y|x)$, we obtain

$$p(X = x, Y = y) = \gamma \exp\left[-\sum_{c\in C} \phi_c(x_c) + \sum_{s\in S} \log f_{x_s}(y_s)\right] \tag{6}$$

Hence, according to (6), the couple (X, Y) is a Markov field and the distribution of X conditional on $Y = y$ is also Markovian. This allows to sample a realization of X according to its posterior distribution $p(x|y)$ and hence, to apply Bayesian techniques like marginal posterior mode (MPM) and maximum a posteriori (MAP).

The feasibility of the different estimations of interest in HMFs stems from the possibility of sampling realizations of the hidden process X from $Y = y$ according to the posterior distribution $p(x|y)$, which is possible when this latter distribution is of Markov form. On the other hand, the Markovianity of this latter distribution relies itself on the assumption that the random variables $(Y_s)_{s\in S}$ are independent conditionally on X. However, such an assumption turns out to be too strong, particularly for the texture modeling problem where some texture classes cannot be handled [2]. To overcome this drawback, HMFs have been extended to pairwise Markov fields (PMFs) [3] and triplet Markov fields (TMFs) [4,5], that are briefly recalled and commented in what follows.

Considering that the pair (X, Y) is a pairwise Markov field consists in assuming that the pair $W = (X, Y)$ is a Markov field, which ensures the Markovianity of both $p(x|y)$ and $p(y|x)$. We have

$$p(W = w) = \gamma \exp\left[-\sum_{c\in C} \phi_c(w_c)\right] \tag{7}$$

The HMF given by (6) is then a particular PMF for which X is Markov. Thus compared to HMFs, the Markovianity of $p(y|x)$ allows more modeling possibilities of noise while the Markovianity of $p(x|y)$ allows the same processing properties.

The generalization of PMFs to TMFs consists in introducing a third process $U = (U_s)_{s\in S}$, where each U_s takes its values in a finite set $\Lambda = \{\lambda_1, \ldots, \lambda_J\}$, and considering that $T = (U, X, Y)$ is a Markov field:

$$p(T = t) = \gamma \exp\left[-\sum_{c\in C} \phi_c(t_c)\right] \tag{8}$$

The problem remains to estimate X from $Y = y$. TMFs are more general than PMFs since the distribution of (X, Y) is not necessarily Markov. Still, the conventional processing methods remain workable. As specified in [4], the auxiliary random field U can have different meanings and its estimation, which is also possible, may be of interest. For instance, U can model the fact that the field X may be nonstationary, which seems to present encouraging perspective, particularly in textured image segmentation [5].

2.4. Hidden evidential Markov fields

Let us return to random fields $X = (X_s)_{s\in S}$, $Y = (Y_s)_{s\in S}$ with $Card(S) = N$. We recall in this paragraph some recent "evidential" extensions of the classic HMF. Thus, let us consider the classic HMF of equation (6).

Let

$$p_1(x) = \gamma \exp\left[-\sum_{c\in C}\phi_c(x_c)\right],$$

and let

$$p^y(x) = \prod_{s\in S}\left[\frac{f_{x_s}(y_s)}{\sum_{x'_s\in\Omega} f_{x'_s}(y_s)}\right].$$

Then the posterior distribution $p(x|y)$ associated to the HMF given by (6) can be seen as the DS fusion of p_1 and p^y:

$$p(x|y) = (p_1 \oplus p^y)(x) \tag{9}$$

That is of importance as it opens way to different possibilities of extensions [13]. More precisely, if either p_1 or p^y is extended in $p_1 \oplus p^y$ to an evidential *bba*, the fusion result remains a probability distribution, which can then be seen as an extension of the classic posterior probability $p(x|y)$. Furthermore, if the "evidential" extension of p_1 or p^y is of a similar Markovian form, in spite of the fact that the fusion result is no longer necessarily a Markov field, the computation of posterior margins $p(x_s|y)$ remains feasible. In fact, the core point is to remark that the fusion (1) can be interpreted as the computation of some marginal distribution, which leads, in the Markov fields context we deal with in this paper, to consider a particular triplet Markov fields.

For instance, if p_1 is replaced by a Markov *bba* M (called evidential Markov field and denoted EMF [13]) defined on $[P(\Omega)]^N$ by

$$M(u) = \gamma \exp\left[-\sum_{c\in C}\phi_c(u_c)\right], \tag{10}$$

then, the DS fusion $M \oplus p^y$ is the posterior distribution $p(x|y)$ associated to $p(x, y)$ which is itself a marginal distribution of the distribution $p(u, x, y)$ of a TMF $T = (U, X, Y)$ below and hence X can still be estimated from $Y = y$.

More precisely, we have:

$$\begin{aligned}
(M \oplus p^y)(x) &\propto \sum_{u\ni x}\exp\left[-\sum_{c\in C}\phi_c(u_c)\right]\prod_{s\in S} p(y_s|x_s)\\
&\propto \sum_{u\ni x}\exp\left[-\sum_{c\in C}\phi_c(u_c) + \sum_{s\in S}\log(p(y_s|x_s))\right]\\
&\propto \sum_{u\ni x}\exp\left[-\sum_{c\in C}\varphi_c(u_c, x_c, y_c)\right]
\end{aligned}$$

where $x \in u$ means $x_s \in u_s$ for each $s \in S$.

The interest of such extension has been shown in hidden nonstationary Markov fields segmentation [13]: when the prior distribution p_1 is nonstationary, replacing p_1 with a stationary *bba* M of (10) form yields significantly better results in unsupervised segmentation than replacing it with any other stationary classic HMF. This is crucial in unsupervised context, where all parameters have to be estimated from $Y = y$. Using some estimation method leads, when keeping the classic model given by (6), to a stationary $\hat{p}_1$. When using a stationary extension (10), the model parameters can still be estimated from $Y = y$ and $\hat{M}$-based segmentation provides better results. Such model is called "hidden evidential Markov field" (HEMF).

3. DS-fusion of evidential Markov fields

In this section, we introduce an original model, called "evidential pairwise Markov field" (EPMF), generalizing a wide range of Markov field models including PMFs and EHMFs. We first describe the new model. Then, we show how some examples of section 2 can be extended, using the proposed model, to take the spatial information into account. Let us mention that there have been many attempts in the literature to consider the contextual information in DS fusion of images since the paper [14]. The originality and interest of the proposed model relies in the introduction of an auxiliary random field taking its values in a finite set $\Lambda = \{\lambda_1, \ldots, \lambda_J\}$ which allows one, roughly speaking, to keep the Markovian form of the considered distributions after Dempster–Shafer fusion. In other words, the main idea is to consider each of the *bba*s of interest as a marginal distribution of a Markov field rather than a Markov field.

Remark 3.1. Let us specify the main idea of the paper within the simple framework of section 2.1, without Markov context. As the set $P(\Omega)$ is finite and small enough to rapidly compute the sum of any positive function $f : B \longrightarrow f(B)$, having a *bba* M exactly or knowing that it is proportional to f is the same information. Indeed, normalizing f to obtain a *bba* is a fast operation. Let us now imagine that f is of the form $f(B) = \sum_{\lambda \in \Lambda} g(B, \lambda)$, where Λ is finite and small enough to make rapidly computable this sum for any B. Similarly, we can say that having M or having g (which defines f) is equivalent. Indeed, a given g allows a fast computation of the associated M, and a given M can be considered as a g with Λ reduced to a singleton. The idea is to consider the DS fusion at the level of functions g instead of *bbas* M, and its interest is the following. Let M_1, M_2 be two *bbas* given by g_1 (with Λ_1) and g_2 (with Λ_2), respectively. Then their DS fusion $M_1 \oplus M_2$, which at each $A \in P(\Omega)$ is proportional to $\sum_{B_1 \cap B_2 = A} M_1(B_1) M_2(B_2)$, is also proportional to

$$\sum_{B_1 \cap B_2 = A} \Big[\sum_{\lambda_1 \in \Lambda_1} g_1(B_1, \lambda_1) \Big] \Big[\sum_{\lambda_2 \in \Lambda_2} g_2(B_2, \lambda_2) \Big] = \sum_{(\lambda_1, \lambda_2, B_1, B_2) \in \Lambda_1 \times \Lambda_2 \times [P(\Omega)]^2} g_1(B_1, \lambda_1) g_2(B_2, \lambda_2) 1_{[B_1 \cap B_2 = B]}$$

Thus it is given by g defined on $P(\Omega) \times \Lambda$, with $\Lambda = \Lambda_1 \times \Lambda_2 \times [P(\Omega)^2]$, by

$$g(B, \lambda) = g(B, (\lambda_1, \lambda_2, B_1, B_2)) = g_1(B_1, \lambda_1) g_2(B_2, \lambda_2) 1_{[B_1 \cap B_2 = B]} \tag{11}$$

This allows one to define the DS fusion without a sum. This is not of importance in the simple framework considered here, but it will be in Markov context considered below because of the fact that the related sums cannot be computed. This possibility will be the core point exploited in the general family of Markov models proposed below. More precisely, if g_1 and g_2 are of Markov forms, we will see that g also is of Markov form. Then classic Markovian methods can be used to sample realizations (z, u) according to *bba* z defined by g, and to estimate $g_s(z_s, u_s)$ on each $s \in S$, which is the core point. To finish, the sum of $g_s(z_s, u_s)$ over u_s is made on each $s \in S$, which is feasible. Then the result can be used to perform segmentation. In other words, segmentation can be performed thanks to the fact that Markov form is saved (thanks to the fact that sums are not performed) through successive fusions of different informations provided by different Markov models.

As in the previous section, we will adopt the language of image processing; however, the proposed model could be of interest in any other application area involving hidden discrete Markov fields. Let us return to the classic hidden Markov field (HMF) model (6). For S a set of pixels, which will be seen in this paper as a rectangular table on which are defined images, we consider two sets of random variables $(X_s)_{s \in S}$ and $(Y_s)_{s \in S}$ where each X_s taking its values in $\Omega = \{\omega_1, \ldots, \omega_K\}$ and each Y_s taking its values in $\mathbb{R}^d$. We have seen in the previous section that in the classic HMF the posterior distribution $p(x|y)$ could be interpreted as the DS fusion of the prior distribution $p(x) = \gamma \exp[-\sum_{c \in C} \phi_c(x_c)]$ and the "likelihood provided" (LP) distribution $q^y(x) = \prod_{s \in S} [\frac{p(y_s|x_s)}{\sum_{\omega \in \Omega} p(y_s|x_s = \omega)}]$. The latter will be called "blind" LP distribution where "blind" refers to the fact that the variables corresponding to such a product are independent, and thus are "blind" of the context. Thus in the classic HMF, we can say that the "blind" priors have been extended to a Markov distribution, while the "blind" LP distribution has been kept.

We propose the following general Markov model, which incorporates different known evidential Markov fields, hidden evidential Markov fields, conditional random Markov fields, or still, triplet Markov fields.

Definition 3.1. Let us consider:

(i) $\Omega = \{\omega_1, \ldots, \omega_K\}$ a set of classes, and $P(\Omega) = \{A_1, \ldots, A_q\}$ its associated powerset, with $q = 2^K$;
(ii) $\Lambda = \{\lambda_1, \ldots, \lambda_J\}$ a finite set;
(iii) S a set of pixels with $N = Card(S)$, and $V = (V_s)_{s \in S} = (Z, U) = (Z_s, U_s)_{s \in S}$ a random field, each (Z_s, U_s) taking its values in $P(\Omega) \times \Lambda$;
(iv) $I_{V_s} \subset P(\Omega) \times \Lambda$ is the image set common for all (Z_s, U_s), $s \in S$.

Then $V = (Z, U) = (Z_s, U_s)_{s \in S}$ is called "evidential pairwise Markov field" (EPMF) if its distribution is given on $I_V = [I_{V_s}]^N$ by

$$p(v) = \gamma \exp\left[-\sum_{c \in C} \phi_c(v_c) \right] \tag{12}$$

where C is the set of cliques related to some neighborhood system.

Definition 3.2. Let us consider the context of Definition 3.1. Let $(Y_s)_{s \in S}$ be a random field, each Y_s taking its values in $\mathbb{R}^d$. $V = (Z, U) = (Z_s, U_s)_{s \in S}$ is called "conditional evidential pairwise Markov field" (CEPMF) if its distribution conditional on $Y = y$ is an EPMF:

$$p(v|y) = \gamma(y) \exp\left[-\sum_{c \in C} \phi_c^y(v_c) \right] \tag{13}$$

Remark 3.2. It is of importance to notice that the support $I_{V_s} \subset P(\Omega) \times \Lambda$ of the law of $V_s = (Z_s, U_s)$ is not, in general, the whole $P(\Omega) \times \Lambda$, but a part of it. It will be convenient to consider I_{V_s} as being defined by:

(i) the set $I_{Z_s} \subset P(\Omega)$ such that there exists $u_s \in \Lambda$ for which $(z_s, u_s) \in I_{V_s}$;
(ii) and by the function which associates to each $z_s \in I_{Z_s}$ the set $\Lambda(z_s)$ of elements u_s in Λ such that $(z_s, u_s) \in I_{V_s}$.

Definition 3.3. Let $V = (Z, U)$ be an EPMF (a CEPMF, respectively). The distribution of Z, which simply is the marginal distribution, will be said "stemming from EPMF $V = (Z, U)$" (from CEPMF, respectively).

Remark 3.3. The following two points make the interest of EMPFs and CEMPFs:

(i) Let $V = (Z, U)$ be an EPMF (or a CEPMF). For each $s \in S$, the distribution $p(z_s)$ can be estimated in a similar way as the classical posterior distribution $p(x_s|y)$ is estimated in the very classic hidden Markov fields (6). Thus, in a CEMPF, $p(z_s|y)$ is easily estimable. Indeed, realizations of $V = (Z, U)$ can be sampled with some classic method like Gibbs sampler, which makes possible the estimation of $p(z_s, u_s|y)$. Then $p(z_s|y)$ is easily computable with $p(z_s|y) = \sum_{u_s \in \Lambda(z_s)} p(z_s, u_s|y)$. Then there are two possibilities: $p(z_s|y)$ is a probability or a *bba*. In the first case, the classic Bayesian MPM method is used to estimate the hidden class. In the second case, one can compute the plausibility $Pl(\omega_i) = \sum_{z_s|\omega_i \in z_s} p(z_s|y)$, and estimate the hidden class by the "maximum of plausibility". Finally, important is that in the general CEPMF the hidden classes can be searched by a method extending the classic MPM method, which is merely as simple as the latter used in classic HMFs;
(ii) Let M_1 and M_2 be two *bba*s stemming from EPMFs (or CEPMFs) $V^1 = (Z^1, U^1)$ and $V^2 = (Z^2, U^2)$, respectively. As we will see in Proposition 3.1 below, the DS fusion $M_1 \oplus M_2$ also stems from an EPMF (or CEPMF), $V = (Z, U)$, with the distribution easily computable from the distributions of $V^1 = (Z^1, U^1)$ and $V^2 = (Z^2, U^2)$. Thus the set of "*bba*s stemming from EPMFs (or CEPMF)" is closed with respect to the DS fusion, and this fusion is easy to compute. Thus using this set rather than the set of hidden evidential Markov fields, which is not closed with respect to DS fusion, is more interesting. Of course, such models are of practical interest, as discussed in Remark 3.4 below.

Remark 3.4. Considering the particular purely probabilistic case of CEPMF with blind LP distribution, in which the distribution of $(Z_s)_{s\in S}$ is probabilistic, we find again the so-called "triplet Markov fields" (TMFs) in which the set $\Lambda = \{\lambda_1, \ldots, \lambda_J\}$ can be seen as modeling the nonstationarity. Introduced in [33] and developed in [4,5], TMFs have recently given different extensions particularly useful in SAR images processing [34–39,6,40,41].

Proposition 3.1. *Let M_1 and M_2 be two* bba*s stemming from EPMFs (or CEPMFs) $V^1 = (Z^1, U^1)$ and $V^2 = (Z^2, U^2)$ respectively, both Markovian with respect to the same neighborhood. Let $p(z^1, u^1) = \gamma_1 \exp\left[-\sum_{c\in C} \phi_c^1(z_c^1, u_c^1)\right]$ and $p(z^2, u^2) = \gamma_2 \exp\left[-\sum_{c\in C} \phi_c^2(z_c^2, u_c^2)\right]$. For each $s \in S$, both Z_s^1 and Z_s^2 take their values from $I_{Z_s^1}$ and $I_{Z_s^2}$ respectively. For each $s \in S$, $z_s^1 \in I_{Z_s^1}$ and $z_s^1 \in I_{Z_s^2}$, U_s^1 and U_s^2 take their values from $\Lambda^1(z_s^1)$ and $\Lambda^2(z_s^2)$, respectively. This defines the image sets $I_{V_s^1}$ and $I_{V_s^2}$. Thus $V^1 = (Z^1, U^1)$ and $V^2 = (Z^2, U^2)$ take their values from $I_{V^1} = [I_{V_s^1}]^N$ and $I_{V^2} = [I_{V_s^2}]^N$, respectively. Then DS fusion $M = M_1 \oplus M_2$ stems from an EPMF (or a CEPMF) $V = (Z, U) = (Z_s, U_s)_{s\in S}$ whose distribution $p(z, u) = \gamma \exp\left[-\sum_{c\in C} \phi_c(z_c, u_c)\right]$ is defined as follows:*

(i) *I_{Z_s} is defined with $[z_s \in I_{Z_s}] \Leftrightarrow [\exists z_s^1 \in I_{Z_s^1}, \exists z_s^2 \in I_{Z_s^2} | z_s = z_s^1 \cap z_s^2]$ (and thus $z = (z_s)_{s\in S} \in I_Z = [I_{Z_s}]^N$ iff the previous equivalence holds for each $s \in S$);*
(ii) *For each $z_s \in I_{Z_s}$, let us define $\Lambda_s(z_s) \subset I_{Z_s^1} \times I_{Z_s^2} \times \Lambda^1(z_s^1) \times \Lambda^2(z_s^2)$ by $[u_s = (z_s^1, z_s^2, u_s^1, u_s^2) \in \Lambda_s(z_s)] \Leftrightarrow [z_s = z_s^1 \cap z_s^2, u_s^1 \in \Lambda^1(z_s^1), u_s^2 \in \Lambda^2(z_s^2)]$. Thus if this equivalence holds for each $s \in S$, we have $u = (u_s)_{s\in S} \in \Lambda(z) = \prod_{s\in S} \Lambda_s(z_s)$ (the product being the Cartesian product of sets);*
(iii) *points (i) and (ii) above define $I_{V_s} = \{(z_s, u_s) | z_s \in I_{Z_s}$ and $u_s \in \Lambda_s(z_s)\}$, and $I_V = [I_{V_s}]^N$, on which is defined $p(z, u)$;*
(iv) *for each clique $c \in C$, $\phi_c(z_c, u_c)$ is given by*

$$\phi_c(z_c, u_c) = \phi_c(z_c, (z_c^1, z_c^2, u_c^1, u_c^2)) = [\phi_c^1(z_c^1, u_c^1) + \phi_c^2(z_c^2, u_c^2)], \tag{14}$$

with $z_c \in I_{Z_c}$ and $(z_c^1, z_c^2, u_c^1, u_c^2) \in \Lambda_c(z_c)$.

Proof. Points (i), (ii) and (iii) stem directly from the Dempster–Shafer fusion principle. To show (iv), let $M = M_1 \oplus M_2$. We have

$$M(z) = [M_1 \oplus M_2](z) \propto \sum_{z^1 \cap z^2 = z} M_1(z^1) M_2(z^2)$$

$$= \sum_{z^1 \cap z^2 = z} \left[\sum_{u^1 \in \Lambda^1(z^1)} \gamma_1 \exp\left[- \sum_{c \in C} \phi_c^1(z_c^1, u_c^1) \right] \right] \left[\sum_{u^2 \in \Lambda^2(z^2)} \gamma_2 \exp\left[- \sum_{c \in C} \phi_c^2(z_c^2, u_c^2) \right] \right]$$

$$= \gamma_1 \gamma_2 \sum_{z^1 \cap z^2 = z} \left[\sum_{(u^1, u^2) \in \Lambda^1(z^1) \times \Lambda^2(z^2)} \exp\left[- \sum_{c \in C} \left[\phi_c^1(z_c^1, u_c^1) + \phi_c^2(z_c^2, u_c^2) \right] \right] \right]$$

$$= \gamma_1 \gamma_2 \sum_{(z^1, z^2, u^1, u^2) \in \Lambda(z)} \exp\left[- \sum_{c \in C} \left[\phi_c^1(z_c^1, u_c^1) + \phi_c^2(z_c^2, u_c^2) \right] \right]$$

This shows that $M = M_1 \oplus M_2$ stems from the EPMF defined on $I_V = [I_{V_s}]^N$ (point (iii)) by $p(z, u) = \gamma \exp\left[- \sum_{c \in C} (\phi_c^1(z_c^1, u_c^1) + \phi_c^2(z_c^2, u_c^2)) \right]$, (recall that for each $s \in S$, $u_s = (z_s^1, z_s^2, u_s^1, u_s^2)$ is such that $z_s = z_s^1 \cap z_s^2, u_s^1 \in \Lambda^1(z_s^1), u_s^2 \in \Lambda^2(z_s^2)$), which ends the proof. □

Let us specify how to proceed in practice. First, one has to define I_{Z_s} from $I_{Z_s^1}$ and $I_{Z_s^2}$. That is done by considering all A_s such that there exist $A_s^1 \in I_{Z_s^1}$ and $A_s^2 \in I_{Z_s^2}$ for which $A_s = A_s^1 \cap A_s^2$. Then for each clique c, each $A_c \in I_{Z_c}$, each $\lambda_c^1 \in \Lambda_1(z_c^1)$ and each $\lambda_c^2 \in \Lambda_2(z_c^2)$, one considers $\lambda_c = 1_{[z_c^1 \cap z_c^2 = A_c]}(z_c^1, z_c^2, \lambda_c^1, \lambda_c^2)$. Finally, $\phi_c(z_c = A_c, u_c = \lambda_c) = \phi_c^1(z_c^1, \lambda_c^1) + \phi_c^2(z_c^2, \lambda_c^2)$.

Let us resume different possibilities of new models. In the context considered in section 2, which can be considered as "pixel by pixel processing" context, we considered four cases: prior $p(x_s)$ and LP $q^y(x_s)$ distributions Bayesian (the classic probabilistic case); prior *bba* $p(x_s)$ general and LP *bba* $q^y(x_s)$ Bayesian, prior *bba* $p(x_s)$ Bayesian and LP *bba* $q^y(x_s)$ general, both of them general. In the context of random fields considered here, in each of these cases prior $p(x_s)$ and LP $q^y(x_s)$ can either be extended to a Markov *bba*, or used to define a "blind" bba. This provides twelve possibilities of new models, all of which being CEPMFs.

For example, in [17] we considered a probabilistic Markov field for the prior distribution, a blind probabilistic LP provided by a radar image, and a blind evidential LP provided by an optical image. Evidential LP was used to model existence of clouds in the scene as in Example 2.2. This is a very simple CEPMF, and even somewhat "degenerate" one. Indeed, contextual information are considered via Bayesian Markov field, and in such a case DS fusion is performed "pixel by pixel", which does not need the use of the general formula (14). Contrariwise, the model introduced in [13] to model priors is a "real" evidential Markov model. It is a particular CEPMF in that LP is Bayesian and blind. Thus general CEPMF extends these LP in two directions: they may be evidential, and they may be Markov. Let us also specify that the general ideas considered here are somewhat similar to those considered in the Markov chains framework [18,19,42,20]. However, in spite of this similarity, precise models and solutions are very different. Indeed, in Markov chains context posterior distributions used in Bayesian processing can be computed, while they must be estimated using some "Monte Carlo Markov Chain" (MCMC) method in Markov fields context considered here.

Remark 3.5. Let (Z, U) be an EPMF (or CEPMF) and Λ the set of possible values of each U_s. According to Proposition 3.1, Λ models DS fusion, which is the core point of the paper. However, Λ can also (simultaneously) classically model the nonstationarity. Thus U can be of the form $U = (U', U'')$ and take its values in a product $\Lambda = \Lambda' \times \Lambda''$, where Λ' possibly models the DS fusion result as specified in (12), and Λ'' models the nonstationarity of $(Z, U'') = ($ or (Z, U'', Y) in the case of CEPMF).

Remark 3.6. Let us briefly mention the problem of estimating the mass elements, crucial for the practical use. The problem of identifying the focal elements, it is to say the sets with non-null bba, is often solved by the physical nature of the problem, as it is the case in different examples given in sub-section 2.1. If not, some methods of automated identification do exist [43–48], and the problem is hard in the general case. If the focal elements are known, there is no theoretical difficulty in applying "iterative conditional estimation" (ICE) to estimate all the model's parameters, once the form of potentials are chosen. Indeed, the result of different fusions (before summing) is a classic triplet Markov field, and ICE, successfully used in [4,5,38–41], can be applied (notice that the hidden class field $X = (X_s)_{s \in S}$ in these references is here $Z = (Z_s)_{s \in S}$, while the observed field – in CEPMF case – is $Y = (Y_s)_{s \in S}$, and the third field is $U = (U_s)_{s \in S}$, likely as in the references). However, we see according to Proposition 3.1 and Remark 3.1 that in the DS fusion case the number of elements in Λ can increase very quickly, and thus in practical applications ICE (or other methods, like those of EM kind) could turn out to be not so easy to use.

Let us specify how Examples 2.1 and 2.2 given in the previous section can be extended to CEPMFs. Let us mention that the other examples can be extended in an analogous manner.

Example 3.1. Let us extend the situation of *Example* 2.1 using the proposed formalism. Let us first consider the extension at the "prior" *bba* level: the bba A defined on $\Delta = \{\{\omega_1\}, \ldots, \{\omega_K\}, \Omega\}$ is extended to a Markov field $Z^1 = (Z_s^1)_{s \in S}$, each Z_s^1

taking its values in Δ: $p(z^1) = \gamma \exp[-\sum_{c\in C} \phi_c(z_c^1)]$. Thus, using the probability $q^{y_s}(x_s = \omega_k) = \frac{p(y_s|x_s=\omega_k)}{\sum_{1\le i\le K} p(y_s|x_s=\omega_k)}$ defined on Ω to define the blind *bba* $Q^y(z^2) = \prod_{s\in S} Q^{y_s}(z_s^2)$, with $Q^{y_s}(\{\omega\}) = q^{y_s}(\omega)$, we obtain

$$p(z) = \gamma \sum_{z^1\cap z^2=z} \exp\left[-\sum_{c\in C} \phi_c(z_c^1) + \sum_{s\in S} \log(Q^{y_s}(z_s^2))\right] \tag{15}$$

Such extension has been introduced and studied in [13], with applications to nonstationary hidden field. Notice that if we set $\sum_{c\in C} \phi_c(z_c^1) = \sum_{s\in S} \phi_s(z_s^1)$, the Markovianity disappears and we find again *Example* 2.1. This shows how more general is this model with respect to the one considered in *Example* 2.1.

Another possible extension would be to extend the probabilities $(q^{y_s})_{s\in S}$ defined on Ω by introducing Markovianity: $q^y(z^2) = \gamma^y \exp[-\sum_{c\in C} \varphi_c^y(z_c^2)]$. Hence, using the *bbas* $(A_s)_{s\in S}$ defined on Δ to define a blind prior *bba* $A(z^1) = \prod_{s\in S} A(z_s^1)$, we obtain the following $p(z)$, stemming from a CEPMF:

$$p(z) = \gamma^y \sum_{z^1\cap z^2=z} \exp\left[\sum_{s\in S} \log(A_s(z_s^1)) - \sum_{c\in C} \varphi_c^y(z_c^2)\right] \tag{16}$$

Finally, we can consider both extensions simultaneously. More explicitly, A and $(q^{y_s})_{s\in S}$ are extended to Markov distributions $p(z^1) = \gamma \exp[-\sum_{c\in C} \phi_c(z_c^1)]$ and $q^y(z^2) = \gamma^y \exp[-\sum_{c\in C} \varphi_c^y(z_c^2)]$, respectively. Then $I_{Z_s^1} = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \Omega\}$, $I_{Z_s^2} = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\}$, and thus $I_{Z_s} = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\}$. Hence, for $z_s = \{\omega_i\} \in I_{Z_s}$ we have $\Lambda(z_s) = \{(\{\omega_i\}, \{\omega_i\}), (\Omega, \{\omega_i\})\}$. This gives I_{V_s}, which thus contains eight elements. Finally, for a clique c, v_c is the set of $(z_s, u_s)_{s\in c}$ such that $u_s \in \Lambda(z_s)$. Then the CEPMF's distribution is $p(v) \propto \exp[-\sum_{c\in C} \psi_c^y(v_c)]$, with $\psi_c^y(v_c)$ defined for each $v_c = (v_s)_{s\in c} = (z_s, u_s)_{s\in c}$ by $\psi_c^y(v_c) = \phi_c(z_c^1) + \varphi_c^y(z_c^2)$. Accordingly, $p(z)$ is given by

$$p(z) \propto \sum_{z^1\cap z^2=z} \exp\left[-\sum_{c\in C} \phi_c(z_c^1) + \varphi_c^y(z_c^2)\right] \tag{17}$$

Example 3.2. Let us consider *Example* 2.2, with a probability p on $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and a *bba* Q^{y_s} (with $y_s \in \mathbb{R}$) on $\Delta = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \Omega\}$. As in *Example* 3.1, there are three possibilities of considering CEPMFs: p is extended to a Markov probabilistic and Q^{y_s} is kept as blind *bba*; p is kept as blind *bba* and Q^{y_s} is extended to a CEMF; or both of them are extended. Let us explicit all of them. The first case has been introduced and studied in [17], while the two other ones are original.

(i) p is extended to a Markov probabilistic field: $p(z^1) = \gamma \exp[-\sum_{c\in C} \phi_c(z_c^1)]$, where each $z_s^1 \in \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\}$, and $(Q^{y_s})_{s\in S}$ are used to define a "blind" distribution by $Q^y(z^2) = \prod_{s\in S} Q^{y_s}(z_s^2)$. Then the corresponding $p(z)$, stemming from a CEPMF, is given by:

$$p(z) = \gamma \sum_{z^1\cap z^2=z} \exp\left[-\sum_{c\in C} \phi_c(z_c^1) + \sum_{s\in S} \log(Q^{y_s}(z_s^2))\right] \tag{18}$$

(ii) p is used to define a blind *bba* $P(z^1) = \prod_{s\in S} p(z_s^1)$, and the *bbas* $(Q^{y_s})_{s\in S}$ are extended to a CEMF $Q^y(z^2) = \gamma^y \exp[-\sum_{c\in C} \varphi_c^y(z_c^2)]$. Then the corresponding $p(z)$, stemming from a CEPMF, is given by:

$$p(z) = \gamma^y \sum_{z^1\cap z^2=z} \exp\left[\sum_{s\in S} \log(p(z_s^1)) - \sum_{c\in C} \varphi_c^y(z_c^2)\right] \tag{19}$$

(iii) both p and $(Q^{y_s})_{s\in S}$ are extended to Markov distributions $p(z^1) = \gamma \exp[-\sum_{c\in C} \phi_c(z_c^1)]$ and $Q^y(z) = \gamma^y \exp[-\sum_{c\in C} \varphi_c^y(z_c^2)]$, respectively. Then $I_{Z_s^1} = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\}$, $I_{Z_s^2} = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \Omega\}$, and thus $I_{Z_s} = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\}$. As a consequence, for $z_s = \{\omega_i\} \in I_{Z_s}$ we have $\Lambda(z_s) = \{(\{\omega_i\}, \{\omega_i\}), (\{\omega_i\}, \Omega)\}$. This gives I_{V_s}, which thus contains eight elements. Finally, for a clique c, v_c is the set of $(z_s, u_s)_{s\in c}$ such that $u_s \in \Lambda(z_s)$. Then the CEPMF's distribution is $p(v) = \alpha \exp[-\sum_{c\in C} \psi_c^y(v_c)]$, with $\psi_c^y(v_c)$ defined for each $v_c = (v_s)_{s\in c} = (z_s, u_s)_{s\in c}$ by $\psi_c^y(v_c) = \phi_c(z_c^1) + \varphi_c^y(z_c^2)$. Accordingly, $p(z)$ is given by

$$p(z) \propto \sum_{z^1\cap z^2=z} \exp\left[-\sum_{c\in C} \phi_c(z_c^1) + \varphi_c^y(z_c^2)\right] \tag{20}$$

4. Experiments

4.1. Experiments context

As stated in the previous section, the originality of the proposed extension lies in the introduction of an additional field $(U_s)_{s\in S}$, each U_s taking its values in a finite set $\Lambda = \{\lambda_1, \dots, \lambda_J\}$. In the very classic probabilistic hidden Markov field (X, Y), such introduction leads to triplet Markov fields and the additional field can have at least two interpretations: (i) it can model different stationarities of (X, Y); or (ii) it can model the fact that the noise distributions $p(y_s|x_s)$ are not known exactly and they are approximated with a Gaussian mixture [4, page 483]. In this section we will consider the evidential model proposed in [13], which will be called Model 1, and we will show how the introduction of $(U_s)_{s\in S}$ can improve the segmentation results in the case (ii): the noise distributions are not Gaussian and they are not known.

Let $\Omega = \{\omega_1, \omega_2\}$ be a frame of discernment and let $P(\Omega)$ be its associated powerset. Let us consider $X = (X_s)_{s\in S}$, $Z = (Z_s)_{s\in S}$ and $Y = (Y_s)_{s\in S}$ with X_s, Z_s and Y_s taking their values in Ω, $P(\Omega)$ and $\mathbb{R}$ respectively.

The distribution of Model 1 is given by

$$p(z, x, y) \propto \mathbf{1}_{x\in z} \exp\left[-\sum_{c\in C}\phi_c(z_c) + \sum_{s\in S}\log(p(y_s|x_s))\right] \tag{21}$$

and can also be written through formula (15) of *Example 3.1*.

We will consider a more general EPMF with the priors being evidential and the noise being blind, called Model 2 and extending Model 1 through the introduction of an auxiliary field $U = (U_s)_{s\in S}$ where $U_s \in \Lambda = \{\lambda_1, \lambda_2\}$. Model 2 is given by:

$$p(z, x, u, y) \propto \mathbf{1}_{x\in z} \exp\left[-\sum_{c\in C}\phi_c(z_c) - \sum_{s\in S}\varphi_s(x_s, u_s) + \sum_{s\in S}\log(p(y_s|x_s, u_s))\right] \tag{22}$$

Since $p(z, x, y) = \sum_{u\in\Lambda^N} p(z, x, u, y)$, the noise distribution is the Gaussian mixture one.

We will consider two kinds of experiments: hand-written images extracted from [13], and a real SAR image.

Remark 4.1. Let us specify, according to the context of section 3 and its related notations, how Model 2 is derived from the DS-fusion of two *bbas* M_1 and M_2 stemming from two particular EPMFs: $V^1 = Z^1$ and $V^2 = (Z^2, U^2)$, respectively. For this purpose, let us consider the EMF $Z^1 = (Z_s^1)_{s\in S}$ where each Z_s^1 takes its values in $P(\Omega)$. Z^1, which can be seen as an EPMF V^1 without U^1, is governed by: $p(z^1) \propto \exp[-\sum_{c\in C}\phi_c(z_c^1)]$. Also, let us define the *bba* $Q^y(v^2) = \gamma' \exp[-\sum_{c\in C}\psi_c(v_c^2)] \propto \exp[-\sum_{s\in S}\varphi_s(v_s^2) + \sum_{s\in S}\log(Q^{y_s}(v_s^2))]$, with $Q^{y_s}((\{\omega_i\}, \lambda_j)) \propto p(y_s|x_s = \omega_i, u_s = \lambda_j)$. Then, we have $I_{Z_s^1} = \{\{\omega_1\}, \{\omega_2\}, \Omega\}$, $I_{Z_s^2} = \{\{\omega_1\}, \{\omega_2\}\}$, and thus $I_{Z_s} = \{\{\omega_1\}, \{\omega_2\}\}$. Hence, for $z_s = \{\omega_i\} \in I_{Z_s}$ we have $\Lambda(z_s) = \{(\{\omega_i\}, \{\omega_i\}, \lambda_1), (\{\omega_i\}, \{\omega_i\}, \lambda_2), (\Omega, \{\omega_i\}, \lambda_1), (\Omega, \{\omega_i\}, \lambda_2)\}$. This gives I_{V_s}, which thus contains eight elements. Finally, for a clique c, v_c is the set of $(z_s, u_s)_{s\in c}$ such that $u_s \in \Lambda(z_s)$. Then the CEPMF's distribution is $p(v) \propto \exp[-\sum_{c\in C}\psi_c^y(v_c)]$, with $\psi_c^y(v_c)$ defined for each $v_c = (v_s)_{s\in c} = (z_s, u_s)_{s\in c}$ by $\psi_c^y(v_c) = \phi_c(z_c^1) + \sum_{s\in c}[\varphi_s^y(v_s^2) - \log(Q^{y_s}(v_s^2))]$ and we find formula (22) that implies:

$$p(z) \propto \sum_{z^1\cap z^2 = z, u^2\in\Lambda} \exp\left[-\sum_{c\in C}\phi_c(z_c^1) - \sum_{s\in S}\left[\varphi_s^y(v_s^2) - \log(Q^{y_s}(v_s^2))\right]\right] \tag{23}$$

Notice that model 1 is obtained simply through the DS fusion of M_1 and M_2 stemming from Z^1 and Z^2 respectively. This shows how formula (22) is obtained from formula (21) by introducing an auxiliary field, in accordance with the idea of the paper.

4.2. Unsupervised segmentation of nonstationary hand-drawn images corrupted by general noise densities

The hidden evidential Markov field (21) was proposed in [13] to segment images corrupted by Gaussian noise, and the experimental results confirm its effectiveness. In this experiment, we assess the performance of our proposed method on the same hand-drawn class-image used in [13], shown in Fig. 1, changing the Gaussian noises into Gamma ones, as illustrated in Fig. 2. More explicitly, two observed images are sampled from the class-image through Gamma noise densities. As depicted in Fig. 2, the aspect of one Gamma density used to sample the first image is close to Gaussian form. Both Gamma densities used to produce the second image share the same means with those used for the first image, their aspect is however quite different from any Gaussian one. The noisy images are segmented with three methods based on the following models: Model 1 with true noise densities (Model 1&T), Model 1 with Gaussian noises approximating the true densities (Model 1&G), and Model 2. Except Model 1&T noise parameters, which are the genuine ones, all other parameters are estimated with the general "Iterative Conditional Estimation" (ICE) [4], and maximization is achieved via Iterative Conditional Mode (ICM) [49]. The segmentation results provided by Model 1&T are then considered as a reference for both other models. The segmentation results obtained are shown in Fig. 3, and confusion matrices are provided in Tables 1 and 2. Overall, Model 2

Fig. 1. Nonstationary hand-drawn class-image.

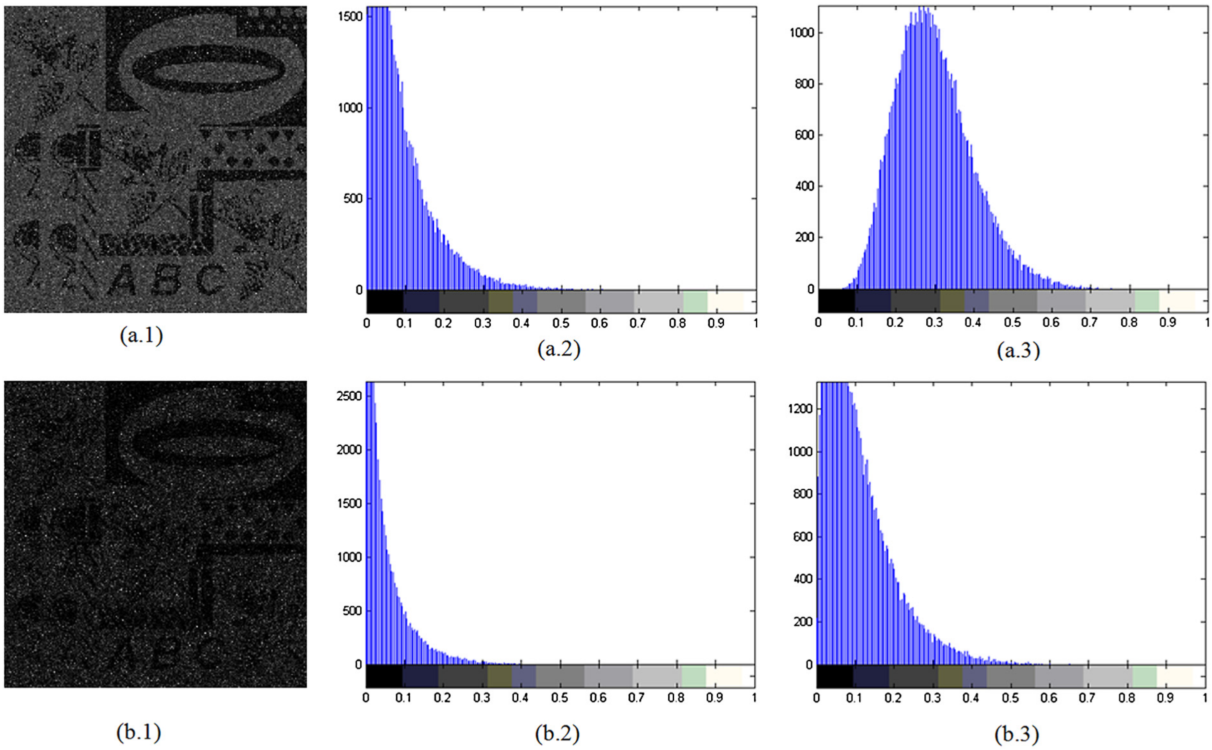


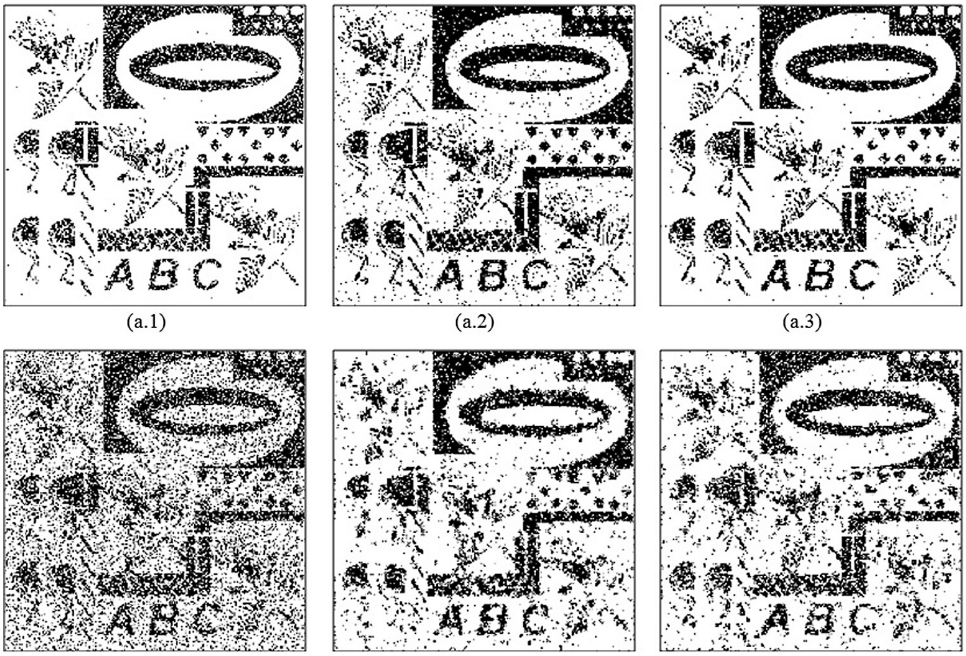
Fig. 2. Nonstationary hand-drawn class-image corrupted by Gamma noise. (a.1) Observed image 1 corrupted by Gamma noise $\Gamma(1,1)$ and $\Gamma(9,1/3)$. (a.2) The histogram of Gamma noise $\Gamma(1,1)$. (a.3) The histogram of Gamma noise $\Gamma(9,1/3)$. (b.1) Observed image 2 corrupted by Gamma noise $\Gamma(0.5,2)$ and $\Gamma(1.5,2)$. (b.2) The histogram of Gamma noise $\Gamma(0.5,2)$. (b.3) The histogram of Gamma noise $\Gamma(1.5,2)$.

yields better results than Model 1&G, especially for the second observed image. One interesting fact is that the overall accuracies of Model 2 on both observed images are very close to the reference ones provided by Model 1&T. Besides, we see that when the noise densities are far from being Gaussian, Model 1&G can perform quite bad in segmentation.

4.3. Unsupervised segmentation of nonstationary SAR images

In this experiment, we consider a real 256×256 SAR image taken by Jet Propulsion Laboratory on L band (see Fig. 4 (a)). Its associated ground truth is shown in Fig. 4(b). According to Fig. 5, we can remark that for such data the noise forms for both classes are not Gaussian. Segmentation is carried out through both Model 1&G and Model 2 (no possibility to consider Model 1&T here). The results are shown in Fig. 4(c) and Fig. 4(d), respectively. The performances of both models are also assessed quantitatively through their confusion matrices, provided in Table 3.

We see that Model 2 outperforms Model 1&G. Indeed, the difference in terms of overall accuracy is high: 96.56% for Model 2 against 85.11% for Model 1. To better understand why Model 2 outperforms Model 1&G, we illustrate in Fig. 5 for each class: the histogram of the actual image intensity, the estimated Gaussian distribution, and the estimated Gaussian mixture distribution. Visually, the Gaussian mixture distribution is better-suited to fit the actual noise density, especially for class ω_2. Let us point out that the different weights of the Gaussian mixture are computed by $\varpi_{ij} = \frac{\exp(\alpha_{ij})}{\sum_{j'=1}^{2} \exp(\alpha_{ij'})}$ where



(a.1) (a.2) (a.3)

(b.1) (b.2) (b.3)

Fig. 3. Segmentation results on nonstationary hand-drawn images corrupted by Gamma noise obtained by: (a.1–b.1) Model 1&G, (a.2–b.2) Mode 1&T, and (a.3–c.3) Model 2. (a.1–a.3) are tested on image 1, and (b.1–b.3) are tested on image 2.

Table 1
Performance evaluation of different models on image 1.

Model 1&G	Predicted ω_1	Predicted ω_2	Accuracy (%)
Actual ω_1	12379	4512	73.29
Actual ω_2	76	41597	99.82
Overall accuracy			92.16
Model 1&T	**Predicted ω_1**	**Predicted ω_2**	**Accuracy (%)**
Actual ω_1	14565	2326	86.23
Actual ω_2	1369	40304	96.71
Overall accuracy			93.69
Model 2	**Predicted ω_1**	**Predicted ω_2**	**Accuracy (%)**
Actual ω_1	13501	3390	79.98
Actual ω_2	314	41359	99.28
Overall accuracy			93.67

Table 2
Performance evaluation of different models on image 2.

Model 1&G	Predicted ω_1	Predicted ω_2	Accuracy (%)
Actual ω_1	12626	4265	74.75
Actual ω_2	7303	34370	82.48
Overall accuracy			80.25
Model 1&T	**Predicted ω_1**	**Predicted ω_2**	**Accuracy (%)**
Actual ω_1	12002	4889	71.06
Actual ω_2	2617	39056	93.72
Overall accuracy			87.18
Model 2	**Predicted ω_1**	**Predicted ω_2**	**Accuracy (%)**
Actual ω_1	11751	5140	69.57
Actual ω_2	2674	38999	93.58
Overall accuracy			86.65

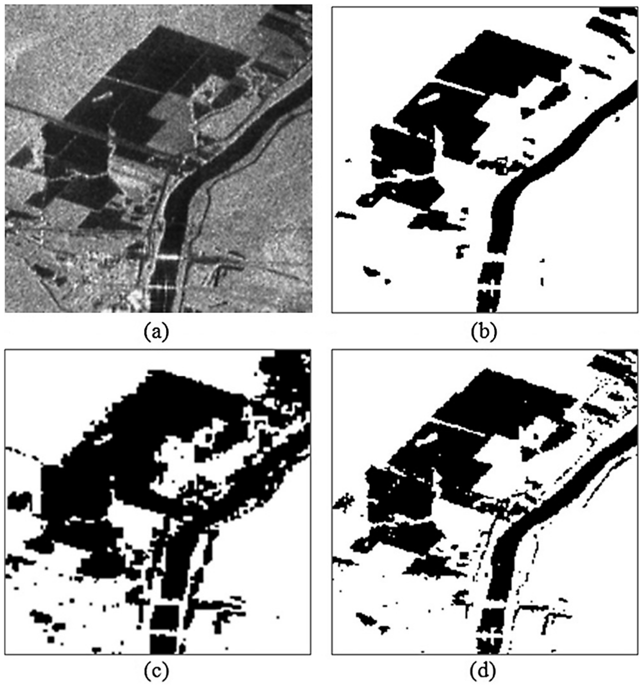


Fig. 4. Unsupervised segmentation of a real SAR image. (a) Real image. (b) Ground truth. (c) Segmentation result based on Model 1. (d) Segmentation result based on Model 2.

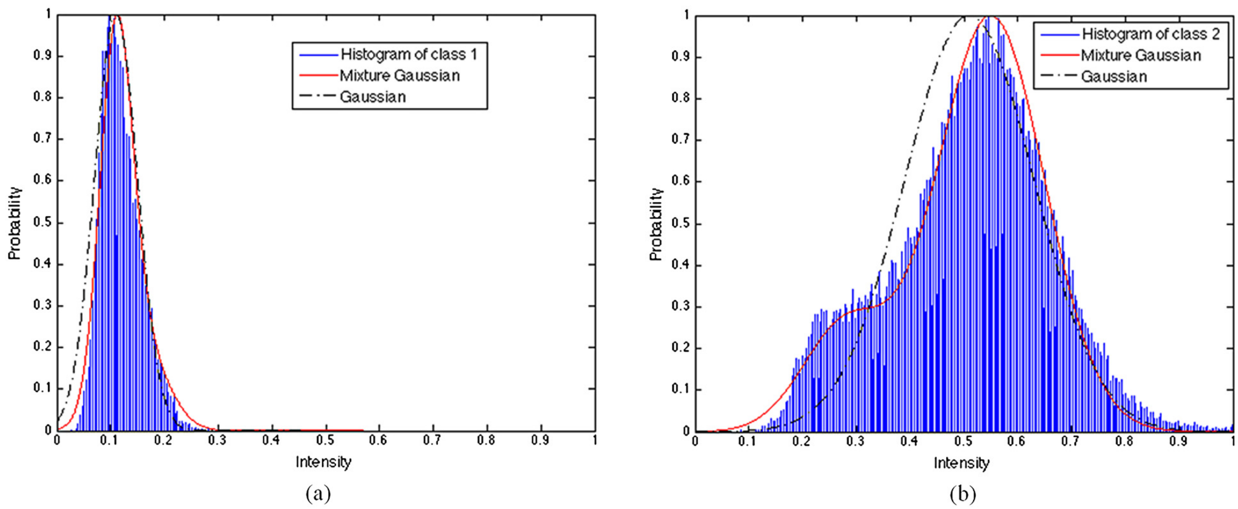


Fig. 5. The actual intensity histogram, the estimated Gaussian distribution and mixture Gaussian distribution of: (a) class 1, and (b) class 2.

$\alpha_{ij} = \varphi_s(x_s = \omega_i, u_s = \lambda_j)$, which agrees with (22) and guarantees that the sum of the weights in one mixture Gaussian distribution is 1.

To conclude, in the particular CEPMF considered in this section the non-stationarity is modeled by the evidential aspect of the model, and the non-Gaussian unknown noise densities are modeled by the additional field $U = (U_s)_{s \in S}$. Then the general ICE method is used to estimate all model's parameters from the sole observation $Y = (Y_s)_{s \in S}$. When the noise is unknown and is not Gaussian, ICM unsupervised segmentation method so obtained can significantly improve the results obtained by the method proposed in [13], which can be seen as a particular case of the considered CEPMF (without $U = (U_s)_{s \in S}$), in both hand-written and real SAR images context.

Finally, let us notice that theoretically Gaussian mixture can fit any types of noises as long as there are enough components. Here we take just two components to assess our model against the one proposed in [13], and it greatly enhances the segmentation results. Of course, our model is open to accept more components, which is likely to save its efficiency in more complex situations.

Table 3
Performance evaluation of different models on SAR image.

Model 1&G	Predicted ω_1	Predicted ω_2	Accuracy (%)
Actual ω_1	15562	0	100
Actual ω_2	9607	39347	80.38
Overall accuracy			85.11
Model 2	**Predicted ω_1**	**Predicted ω_2**	**Accuracy (%)**
Actual ω_1	15485	77	99.51
Actual ω_2	2142	46812	95.62
Overall accuracy			96.56

5. Conclusion

In this paper, we have introduced a unifying framework generalizing a wide family of Markov fields proposed in the literature so far. The proposed formalism allows one on one hand, to consider standard problems regarding information imprecision or unreliability in Markov fields context; and on the other hand, to fuse such information when different sources are concerned. The interest of the proposed evidential pairwise Markov field with respect to the classic models has been shown through some illustrative examples dealing with image segmentation and on real images segmentation. An interesting future direction would be to investigate the possible integration of some sophisticated developments in belief function theory, such as using some recent combination rules that have been proposed instead – or in complement – to the Dempster–Shafer fusion rule considered in this paper [50–52]. Another important future work would be to further investigate the different possibilities of applying the proposed formalism to "SAR image processing" where application results of other TMFs are promising [6,34–36,38–41,53].

References

[1] J. Besag, Spatial interaction and the statistical analysis of lattice systems, J. R. Stat. Soc. B 6 (1974) 192–236.
[2] H. Derin, H. Elliot, Modelling and segmentation of noisy and textured images using Gibbs random fields, IEEE Trans. Pattern Anal. Mach. Intell. 9 (1) (1987) 39–55.
[3] W. Pieczynski, A.-N. Tebbache, Pairwise Markov random fields and segmentation of textured images, Mach. Graph. Vis. 9 (4) (2000) 705–718.
[4] D. Benboudjema, W. Pieczynski, Unsupervised image segmentation using triplet Markov fields, Comput. Vis. Image Underst. 99 (3) (2005) 476–498.
[5] D. Benboudjema, W. Pieczynski, Unsupervised statistical segmentation of nonstationary images using triplet Markov fields, IEEE Trans. Pattern Anal. Mach. Intell. 29 (8) (2007) 1367–1378.
[6] X. Lian, Y. Wu, W. Zhao, F. Wang, Q. Zhang, M. Li, Unsupervised SAR image segmentation based on conditional triplet Markov fields, IEEE Geosci. Remote Sens. Lett. 11 (7) (2014) 1185–1189.
[7] G. Shafer, A Mathematical Theory of Evidence, vol. 1, Princeton University Press, Princeton, 1976.
[8] P. Smets, R. Kennes, The transferable belief model, Artif. Intell. 66 (2) (1994) 191–234.
[9] C. Romer, A. Kandel, Applicability analysis of fuzzy inference by means of generalized Dempster–Shafer theory, IEEE Trans. Fuzzy Syst. 3 (4) (1995) 448–453.
[10] F. Janez, A. Appriou, Theory of evidence and non-exhaustive frames of discernment: plausibilities correction methods, Int. J. Approx. Reason. 18 (1) (1998) 1–19.
[11] C. Lucas, B.N. Araabi, Generalization of the Dempster–Shafer theory: a fuzzy-valued measure, IEEE Trans. Fuzzy Syst. 7 (3) (1999) 255–270.
[12] R.R. Yager, Set measure directed multi-source information fusion, IEEE Trans. Fuzzy Syst. 19 (6) (2011) 1031–1039.
[13] W. Pieczynski, D. Benboudjema, Multisensor triplet Markov fields and theory of evidence, Image Vis. Comput. 24 (1) (2006) 61–69.
[14] S. Le Hegarat-Mascle, I. Bloch, D. Vidal-Madjar, Introduction of neighborhood information in evidence theory and application to data fusion of radar and optical images with partial cloud cover, Pattern Recognit. 31 (11) (1998) 1811–1823.
[15] F. Tupin, I. Bloch, H. Maître, A first step toward automatic interpretation of SAR images using evidential fusion of several structure detectors, IEEE Trans. Geosci. Remote Sens. 37 (3) (1999) 1327–1343.
[16] S. Foucher, M. Germain, J.-M. Boucher, G.B. Benie, Multisource classification using ICM and Dempster–Shafer theory, instrumentation and measurement, IEEE Transactions on 51 (2) (2002) 277–281.
[17] A. Bendjebbour, Y. Delignon, L. Fouque, V. Samson, W. Pieczynski, Multisensor image segmentation using Dempster—Shafer fusion in Markov fields context, IEEE Trans. Geosci. Remote Sens. 39 (8) (2001) 1789–1798.
[18] P. Lanchantin, W. Pieczynski, Unsupervised restoration of hidden nonstationary Markov chains using evidential priors, IEEE Trans. Signal Process. 53 (8) (2005) 3091–3098.
[19] M.E.Y. Boudaren, E. Monfrini, W. Pieczynski, A. Aïssani, Dempster–Shafer fusion of multisensor signals in nonstationary Markovian context, EURASIP J. Adv. Signal Process. 2012 (2012) 134.
[20] L. Fouque, A. Appriou, W. Pieczynski, An evidential Markovian model for data fusion and unsupervised image classification, in: Proceedings of the Third International Conference on Information Fusion, 2000, vol. 1, FUSION 2000, IEEE, Paris, France, 2000, pp. TUB4–25.
[21] H. Dehghani, B. Vahidi, R. Naghizadeh, S. Hosseinian, Power quality disturbance classification using a statistical and wavelet-based hidden Markov model with Dempster–Shafer algorithm, Int. J. Electr. Power Energy Syst. 47 (2013) 368–377.
[22] T. Reineking, Particle filtering in the Dempster–Shafer theory, Int. J. Approx. Reason. 52 (8) (2011) 1124–1135.
[23] E. Ramasso, R. Gouriveau, Prognostics in switching systems: evidential Markovian classification of real-time neuro-fuzzy predictions, in: Prognostics and Health Management Conference, 2010, PHM'10, IEEE, Portland, Oregon, 2010, pp. 1–10.
[24] E. Ramasso, Contribution of belief functions to hidden Markov models with an application to fault diagnosis, in: Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing, Grenoble, France, 2009, pp. 1–6.
[25] E. Ramasso, T. Denoeux, Making use of partial knowledge about hidden states in HMMs: an approach based on belief functions, IEEE Trans. Fuzzy Syst. 22 (2) (2014) 395–405.

[26] Y. Yoji, M. Tetsuya, U. Yoji, S. Yukitaka, A method for preventing accidents due to human action slip utilizing HMM-based Dempster–Shafer theory, in: IEEE International Conference on Robotics and Automation, 2003. Proceedings, vol. 1, ICRA'03, IEEE, Taipei, Taiwan, 2003, pp. 1490–1496.
[27] J. Park, M. Chebbah, S. Jendoubi, A. Martin, Second-order belief hidden Markov models, in: Belief Functions: Theory and Applications, Springer, Oxford, UK, 2014, pp. 284–293.
[28] L.A. Zadeh, Toward a perception-based theory of probabilistic reasoning with imprecise probabilities, J. Stat. Plan. Inference 105 (1) (2002) 233–264.
[29] M.C. Troffaes, Decision making under uncertainty using imprecise probabilities, Int. J. Approx. Reason. 45 (1) (2007) 17–29.
[30] P. Walley, Towards a unified theory of imprecise probability, Int. J. Approx. Reason. 24 (2) (2000) 125–148.
[31] M.E.Y. Boudaren, W. Pieczynski, Dempster–Shafer fusion of evidential pairwise Markov chains, IEEE Trans. Fuzzy Syst. (2016), http://dx.doi.org/10.1109/TFUZZ.2016.2543750.
[32] W. Pieczynski, Multisensor triplet Markov chains and theory of evidence, Int. J. Approx. Reason. 45 (1) (2007) 1–16.
[33] W. Pieczynski, D. Benboudjema, P. Lanchantin, Statistical image segmentation using triplet Markov fields, in: International Symposium on Remote Sensing, SPIEs, Crete, Greece, 2002, pp. 22–27.
[34] P. Zhang, M. Li, Y. Wu, L. Gan, M. Liu, F. Wang, G. Liu, Unsupervised multi-class segmentation of SAR images using fuzzy triplet Markov fields model, Pattern Recognit. 45 (11) (2012) 4018–4033.
[35] P. Zhang, M. Li, Y. Wu, M. Liu, F. Wang, L. Gan, SAR image multiclass segmentation using a multiscale TMF model in wavelet domain, IEEE Geosci. Remote Sens. Lett. 9 (6) (2012) 1099–1103.
[36] F. Wang, Y. Wu, Q. Zhang, P. Zhang, M. Li, Y. Lu, Unsupervised change detection on SAR images using triplet Markov field model, IEEE Geosci. Remote Sens. Lett. 10 (4) (2013) 697–701.
[37] D. Benboudjema, F. Tupin, Markovian modelling and Fisher distribution for unsupervised classification of radar images, Int. J. Remote Sens. 34 (22) (2013) 8252–8266.
[38] F. Wang, Y. Wu, Synthetic aperture radar image segmentation using fuzzy label field-based triplet Markov fields model, IET Image Process. 8 (12) (2014) 856–886.
[39] F. Wang, Y. Wu, Q. Zhang, W. Zhao, M. Li, G. Liao, Unsupervised SAR image segmentation using higher order neighborhood-based triplet Markov fields model, IEEE Trans. Geosci. Remote Sens. 52 (8) (2014) 5193–5205.
[40] G. Liu, M. Li, Y. Wu, P. Zhang, L. Jia, H. Liu, PolSAR image classification based on Wishart TMF with specific auxiliary field, IEEE Geosci. Remote Sens. Lett. 11 (7) (2014) 1230–1234.
[41] L. Gan, Y. Wu, F. Wang, P. Zhang, Q. Zhang, Unsupervised SAR image segmentation based on triplet Markov fields with graph cuts, IEEE Geosci. Remote Sens. Lett. 11 (4) (2014) 853–857.
[42] M.E.Y. Boudaren, W. Pieczynski, Unified representation of sets of heterogeneous Markov transition matrices, IEEE Trans. Fuzzy Syst. (2015), http://dx.doi.org/10.1109/TFUZZ.2015.2460740.
[43] Y.M. Zhu, L. Bentabet, O. Dupuis, D. Babot, M. Rombaut, et al., Automatic determination of mass functions in Dempster–Shafer theory using fuzzy c-means and spatial neighborhood information for image segmentation, Opt. Eng. 41 (4) (2002) 760–770.
[44] W. Jiang, Y. Deng, J. Peng, A new method to determine BPA in evidence theory, J. Comput. 6 (6) (2011) 1162–1167.
[45] T. Denoeux, A k-nearest neighbor classification rule based on Dempster–Shafer theory, IEEE Trans. Syst. Man Cybern. 25 (5) (1995) 804–813.
[46] T. Denoeux, A neural network classifier based on Dempster–Shafer theory, systems, man and cybernetics, Part A: Systems and humans, IEEE Transactions on 30 (2) (2000) 131–150.
[47] R.R. Yager, An intuitionistic view of the Dempster–Shafer belief structure, Soft Comput. 18 (11) (2014) 2091–2099.
[48] P. Xu, Y. Deng, X. Su, S. Mahadevan, A new method to determine basic probability assignment from training data, Knowl.-Based Syst. 46 (2013) 69–80.
[49] J. Besag, R. Kempton, Statistical analysis of field experiments using neighbouring plots, Biometrics (1986) 231–251.
[50] T. Denoeux, The cautious rule of combination for belief functions and some extensions, in: 2006 9th International Conference on Information Fusion, Florence, Italy, IEEE, 2006, pp. 1–8.
[51] T. Denoeux, Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence, Artif. Intell. 172 (2) (2008) 234–264.
[52] A. Kallel, S. Le Hégarat-Mascle, Combination of partially non-distinct beliefs: the cautious-adaptive rule, Int. J. Approx. Reason. 50 (7) (2009) 1000–1021.
[53] L. Gan, Y. Wu, M. Liu, P. Zhang, H. Ji, F. Wang, Triplet Markov fields with edge location for fast unsupervised multi-class segmentation of synthetic aperture radar images, IET Image Process. 6 (7) (2012) 831–838.