International Journal of Approximate Reasoning 102 (2018) 41–59

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Unsupervised segmentation of hidden Markov fields corrupted by correlated non-Gaussian noise

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

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

ARTICLE INFO

Article history:
Received 24 December 2017
Received in revised form 3 August 2018
Accepted 3 August 2018
Available online 9 August 2018

Keywords:
Correlated noise
Non-Gaussian noise
Hidden Markov random fields
Pixel labeling
Theory of evidence
Triplet Markov fields

ABSTRACT

Pixel labeling problem stands among the most commonly considered topics in image processing. Many statistical approaches have been developed for this purpose, particularly in the frame of hidden Markov random fields. Such models have been extended in many directions to better fit image data. Our contribution falls under such extensions and consists of introducing two new models allowing one to deal with non-Gaussian correlated noise. The first one is purely probabilistic, whereas the second one calls on Dempster–Shafer theory of evidence, both being particular triplet Markov fields. The interest of the proposed models is assessed in unsupervised segmentation of sampled and real images. While both models exhibit significant improvement with respect to classic models, the evidential model turns out to be of particular interest when the hidden label field presents fine details.

1. Introduction

Thanks to the development of imaging technology, plenty of data covering optical, SAR and magnetic images are available daily for various processing tasks aiming at improving the wellbeing of public. For instance, SAR images are of great importance in disaster monitoring and can be used day and night [1,2]. Satellite images make it possible to generate reliable ground truths to update ground maps [3,4]. Magnetic images are mainly used for medical purposes in diagnoses, or to assist a surgeon [5–8]. Many studies have been carried out to take full use of such data [9], especially in the frame of statistical data analysis covering image classification, image segmentation and image change detection, all of which can be perceived as pixel labeling problems where one has to recover a label "field" from an observable image. In the probabilistic framework of this paper, the latter is considered as a noisy version of the label field. In image classification for instance, one has to assign each pixel to one among a set of predefined set of classes. Image segmentation is a derivative problem where classes are not known in advance. Change detection is another classification problem where each pixel is assigned to "changed" or "unchanged".

Let S be a finite set, with $|S| = N$, let $Y = \{Y_s\}_{s\in S}$ be the random observable field with each Y_s taking its value in $\mathbb{R}$ (or $\mathbb{R}^d$), and let $X = \{X_s\}_{s\in S}$ be the random label field with each X_s taking its values from a finite set $L = \{l_1, ..., l_K\}$

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

https://doi.org/10.1016/j.ijar.2018.08.001
0888-613X/

of "classes" or "labels". Realizations of such random fields will be denoted using lowercase letters. The labeling problem consists of estimating $X = x$ from $Y = y$.

Over the last three decades, there has been a growing research interest in model-based techniques for image labeling [10–13]. 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" or "likelihood" distribution. The most commonly used model-based techniques in image labeling can be divided into two groups [14]: standard Gaussian mixture models (GMMs) [15,16], and hidden Markov field (HMF) models [17–20]. HMFs have received more attention due to their ability to reduce labeling sensitivity to noise. Indeed, the Markovianity of the posterior distribution $p(x|y)$ implies that each pixel can be labeled using the observations on all pixels, whereas in GMMs it is labeled from its only observation. Furthermore, HMFs have been extended in many directions, most of which focus on the prior distribution. In [21,22], authors consider labels nonstationarity using an auxiliary random field modeling the switches of the hidden label field. The resulting model belongs to the general family of "triplet Markov fields" (TMFs) introduced in [4,23], also see [21,22,24]. Another way to handle heterogeneous behavior of the hidden field relies on the use of theory of evidence [25–27]. Let us also mention the recent general family of models including many classic cases, evidential or not [25].

All such extensions focus on improving the prior distribution, whereas the likelihood one is usually considered as simple product of likelihoods defined on pixels. In some Gaussian situations these likelihoods may be suitable; in some other, different densities like Gamma, Raylegh, Fisher, or still Beta can be used [17]. When the noise shape is not known, it can be modeled through a Gaussian mixture, which leads to GMMs. Besides, the use of GMMs can have precise interpretation. In forest monitoring for instance, forest is the only interesting label, but its intensity is just between water and building. Using Gaussian, Gamma, or any other one-peak distribution to model likelihoods can make it harder to distinguish forest from water and building. Of course, one can still classify the data into three classes and then find out the forest, but considering Gaussian mixture likelihood one can classify the image to forest and non-forest directly.

As well known, two assumptions are usually set for the likelihood in HMFs:

- (H1) for a given realization $X = x$, the random variable $(Y_s)_{s\in S}$ are independent on each other; and,
- (H2) the distribution of each Y_s conditional on X is equal to its distribution conditional on X_s.

In practice, such assumptions can turn out to be too restrictive, as one can easily see that there are numerous situations in which neither of (H1) nor (H2) hypotheses is justified. In the example of a satellite image segmentation into three classes "water", "forest" and "desert", one can consider that trees near water have different aspect than those located near desert, so that (H2) is not verified. Similarly, aspect of the forest at a given pixel depends on the neighboring trees aspect, which makes (H1) difficult to justify. Also, it is well known that HMFs can fail when the label images are inhomogeneous. Homogeneity of an area containing two classes can be defined through $p(x_s \neq x_t)$ where s and t are neighbors [28]. Extending such definition to the case of more classes seems more complicated and requires a deeper analysis.

The contribution of our paper is to propose two new models. The first one, called "correlated Gaussian mixture-hidden Markov field" (CGM-HMF), uses correlated mixtures as likelihoods; thus (H1) is no longer assumed. We will also see that (H2) is no longer assumed as well. The second one is an extension of the CGM-HMF to the evidential context. More precisely, the hidden label field in CGM-HMF is extended to an "evidential" label field: the model will be called "correlated Gaussian mixture-hidden evidential Markov field" (CGM-HEMF). The superiority of CGM-HEMF with respect to CGM-HMF is that it is better suited to situations in which there exist very little homogeneous areas in the label field. Indeed, it has been shown in [26] – and confirmed in [25] – that introducing evidential fields can be of interest in such cases. The interest of the proposed models is established via unsupervised segmentations – with parameter estimation using "iterative conditional estimation" (ICE) – on simulated and real images. More precisely, estimated parameters ensure a good balance between homogeneous and inhomogeneous areas in "evidential Markov field" so that variables are more (resp. less) dependent in homogeneous (resp. inhomogeneous) areas. As shown with experiments, this property may increase the segmentation accuracy.

To summarize, we propose two new models: CGM-HMF and CGM-HEMF such that:

(i) CGM-HMF allows one to deal with non Gaussian correlated noise verifying neither (H1) nor (H2);
(ii) CGM-EHMF, which is an extension of the probabilistic Markov label field in CGM-HMF to an evidential one, with as result better efficiency in label images containing homogeneous and inhomogeneous areas case.

Let us point out that these new models are particular cases of the general TMF. Indeed, every Markov field (U, X, Y) where X is searched, Y is observed, and U is finite is a TMF. The model proposed in [25] is thus a particular TMF, and so is the one proposed in this paper. However, the present TMF is different from the one proposed in [25] in two points: (i) the noise considered here is not Gaussian; and (ii), it is correlated.

The remainder of the paper is organized as follows. Section 2 briefly summarizes the hidden Markov fields, Gaussian mixtures, and Dempster–Shafer theory of evidence aspects needed in the paper. The new models are specified in Section 3. Experimental results obtained on synthetic and real images are provided in Section 4. Finally, concluding remarks and future directions are presented in Section 5.

2. Hidden Markov fields, Gaussian mixtures, and theory of evidence

In this section, we briefly recall the basic notions of hidden Markov fields, Gaussian mixture models, and theory of evidence, needed for the sake of this paper.

2.1. Hidden Markov fields

Let us consider the image labeling problem discussed in the introduction. According to hidden Markov fields (HMFs), the random field X is assumed Markovian with respect to a system of cliques C, defined according to some neighborhood system. The model name "hidden Markov field" stands from the very fact that the hidden field X is Markov. According to the Hammersley–Clifford equivalence [29], the random field X such that $p(x) > 0$ for each $x \in N^L$, is a Markov field if and only if its distribution is given by

$$p(x) \propto \exp\left[-\sum_{c\in C} \psi_c(x_c)\right] \tag{1}$$

where for each clique c, $x_c = (x_s)_{s\in c}$ and ψ_c is a function taking its values in $\mathbb{R}$, called "potential function".

On the other hand, in accordance with hypotheses (H1) and (H2) the likelihood distribution $p(y|x)$ is defined by

$$p(y|x) \propto \exp\left[\sum_{s\in S} \log\left(p(y_s|x_s)\right)\right] \tag{2}$$

Then we have

$$\begin{aligned} p(x|y) &\propto p(x,y) = p(x)p(y|x) \\ &\propto \exp\left[-\sum_{c\in C} \psi_c(x_c) + \sum_{s\in S} \log\left(p(y_s|x_s)\right)\right] \end{aligned} \tag{3}$$

It is important to highlight that the posterior distribution (3) remains Markov, which allows one to sample realizations of X. This is at the origin of possibilities of applying Bayesian optimal methods to search realizations of X from $Y = y$.

Remark 2.1. Let us remark that we adopt the qualifier "prior" for the hidden field X distribution as usually done in the literature. However, parameters defining this distribution are not known and are estimated from observations, so that this distribution is partly a "posterior" one. More generally, in TMFs (X, U, Y) we may say that "prior" knowledge is the general form of its distribution, while exact parameters can be seen as a "posterior" one. In addition, neither the distribution of (X, U) nor that of X is Markov in general and cannot be produced, so that its "prior" knowledge directly stems from the knowledge of the whole distribution of (X, U, Y).

2.2. Gaussian mixture models

Mixture models are particularly interesting probabilistic tools. For example, they can handle automatically the possibility of presence of more than one subpopulation within an overall population without requiring the prior identification to which sub-population an individual observation belongs. Another application, used in this paper, is to approximate unknown distributions with Gaussian mixtures. Gaussian mixture model (GMM), which is among the most commonly used mixture models, is defined through the weighted summation of Gaussian distributions with the sum of all weights being 1. More explicitly, it is given by

$$p(y) = \sum_{j=1}^{M} \pi_j P_j(y) \tag{4}$$

where P_j is Gaussian given by some mean μ_j and some variance σ_j^2.

Such mixtures can be inserted in the HMF (1)–(2): for each label $l_k \in L$, the distribution $p(y_s|x_s)$ is a Gaussian mixture [6,14,30]:

$$p(y_s|x_s = l_k) = \sum_{j=1}^{M} \pi_{kj} P_{kj}(y_s) \tag{5}$$

where the likelihood $P_{kj}(y_s)$ is still constructed by a Gaussian distribution.

We obtain a richer model, that we will call "blind Gaussian mixture-HMF" (BGM-HMF), being able to approximate non-Gaussian noises. However, hypotheses (H1) and (H2) continue to hold. The word "blind" has two meanings here. First, it means that given $X = x$, the distribution of each Y_s only depends on X_s and thus is "blind" of X_t for $t \neq s$, which is (H2). Second, given $X = x$ the distribution of each Y_s is "blind" of Y_t for $t \neq s$, which is (H1).

2.3. Theory of evidence

In the language of theory of evidence (TE, [31]), the label set $L = \{l_1, ..., l_K\}$ is called a "universe of discourse", or "frame of discernment". As above, X_s will take its values from L. Let 2^L be the powerset associated to L. A "basic belief assignment" (*bba*) is a function τ from 2^L to $[0, 1]$ satisfying $\tau(\emptyset) = 0$ and $\sum_{A\in 2^L} \tau(A) = 1$. A *bba* τ then defines a "plausibility" function Pl and a "credibility" function Cr [32], both defined from 2^L to $[0, 1]$ by $Pl(B) = \sum_{A\cap B\neq\emptyset} \tau(A)$ and $Cr(B) = \sum_{A\subseteq B} \tau(A)$ respectively. For a given *bba* τ, related Pl and Cr are linked by $Pl(B) + Cr(B^c) = 1$, where B^c is the complementary set of B with respect to L. Then, a probability p can be perceived as a special case for which τ is null outside of singletons. Indeed, in such a case $Pl = Cr$ and they are a probability, once singletons $\{l_k\}$ are considered as elements l_k. Moreover, if two *bba*s τ_1 and τ_2 represent two pieces of evidence, we can merge, or fuse, them using the so called "Dempster–Shafer fusion" (DS fusion), which defines $\tau = \tau_1 \oplus \tau_2$ given by:

$$\tau(A) = (\tau_1 \oplus \tau_2)(A) \propto \sum_{A_1\cap A_2 = A\neq\emptyset} \tau_1(A_1)\tau_2(A_2) \tag{6}$$

We can see that if either τ_1 or τ_2 is a probability, then $\tau_1 \oplus \tau_2$ is a probability. The interest of TE in Bayesian classification is the following. Let us consider a pixel $s \in S$, a prior probability $p^*(x_s)$ on L, and K noise densities $p(y_s|x_s)$ on $\mathbb{R}$. Then, the posterior probability $p^{y_s}(x_s) = p(x_s|y_s)$ is itself the DS fusion of the probability p^* with the probability q^{y_s} defined on L by $q^{y_s}(x_s) = \frac{p(y_s|x_s)}{\sum_{k=1}^{K} p(y_s|x_s=l_k)}$:

$$p^{y_s} = p^* \oplus q^{y_s} \tag{7}$$

Formula (7) is of interest since when either p^* or q^{y_s} is extended to a *bba* which is not a probability, p^{y_s} remains a probability. Then, it can be seen as an extension of the posterior distribution and thus, can used to perform Bayesian classification. Such extensions can be of interest in image processing using Markov models, as has been described in recent papers [25,33], and also in signal processing using Markov chains models, as studied in [34–36].

3. Probabilistic and evidential hidden Markov fields with correlated non-Gaussian noise

This section contains four subsections. In the first one, we recall the HMF with "blind" Gaussian mixture, called "blind Gaussian mixture-HMF" (BGM-HMF). In subsection 3.2, we propose a new model, which is purely probabilistic (without theory of evidence), called "correlated Gaussian mixture-HMF" (CGM-HMF). The former model is well suited to non Gaussian independent noise, while the latter is likely to improve it in the case of non-Gaussian correlated noise. In subsection 3.3, we recall the recent model tested in [25,37], which will be called "blind Gaussian mixture-hidden evidential Markov field" (BGM-HEMF). So BGM-HEMF is an extension of BGM-HMF obtained by extending the Markov label field to an evidential one. Finally, in subsection 4.2 we introduce the new "correlated Gaussian mixture-HEMF" (CGM-HEMF) model, which is a particularly well-suitable extension of BGM-HMF to deal with non-Gaussian correlated noise. Besides, extending probabilistic BGM-HMF and CGM-HMF to evidential BGM-HEMF and CGM-HEMF allows one to better label images containing fine details. Denoting by G-HMF the Gaussian HMF and by G-HEMF the Gaussian HEMF, the generality order of different extensions is illustrated in Fig. 1.

3.1. Blind Gaussian mixture-hidden Markov field (BGM-HMF)

The expression of the hidden Markov field is given by (1)–(3). One can see that extending the likelihood to Gaussian mixture is equivalent to introducing an auxiliary field $U = (U_s)_{s\in S}$, where $U_s \in \Lambda = \{\lambda_1, ..., \lambda_M\}$ and setting

$$p(x, u, y) \propto \exp\left[-\sum_{c\in C} \psi_c(x_c) - \sum_{s\in S} \eta_s(x_s, u_s) + \sum_{s\in S} \log\left(p(y_s|x_s, u_s)\right)\right] \tag{8}$$

Indeed, since $p(x, y) = \sum_{u\in\Lambda^N} p(x, y, u)$, we have:

$$p(x, y) \propto \exp\left[-\sum_{c\in C} \psi_c(x_c)\right] \prod_{s\in S}\left[\sum_{u_s\in\Lambda} \exp\left[-\eta_s(x_s, u_s)\right] p(y_s|x_s, u_s)\right] \tag{9}$$

and thus, $p(x, y)$ is a HMF with $p(y_s|x_s)$ being mixtures $\sum_{u_s\in\Lambda} \alpha(u_s) p(y_s|x_s, u_s)$, where the mixture coefficients are $\alpha(u_s) = \exp[-\eta_s(x_s, u_s)]$. These models can approximate situations in which $p(y_s|x_s)$ are not Gaussian; however, both (H1) and (H2)

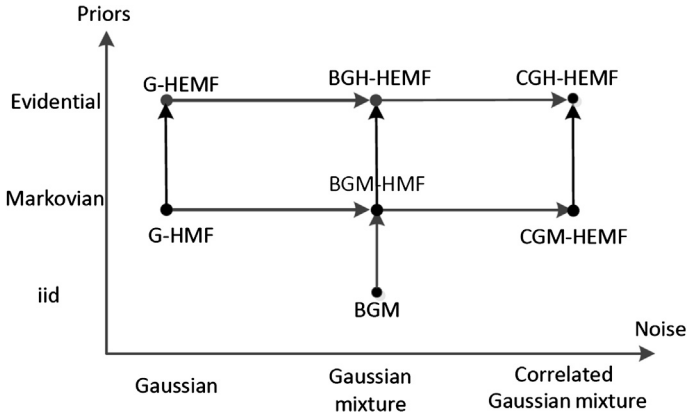


Fig. 1. Growing generality of hidden probabilistic models G-HMF, BGM-HMF, CGM-HMF; and hidden evidential ones BGM-HEMF, CGM-HEMF. G-HMF and G-HEMF are the Gaussian HMF, and the Gaussian evidential HMF, respectively.

hypotheses remain. For this reason, such models will be called "blind Gaussian mixture-HMF" (BGM-HMF) where the word "blind" refers to the fact that the distribution of each Y_s conditional on X_s, being independent of other Y_t and X_t, is "blind" of context.

3.2. Correlated Gaussian mixture-hidden Markov field (CGM-HMF)

Let us consider the following extension of (8):

$$p(x,u,y) \propto \exp\left[-\sum_{c\in C}\psi_c(x_c) - \sum_{c\in C}\varphi_c(u_c) - \sum_{s\in S}\eta_s(x_s,u_s) + \sum_{s\in S}\log\left(p(y_s|x_s,u_s)\right)\right] \tag{10}$$

which will be called "correlated Gaussian mixture-hidden Markov field" (CGM-HMF). The difference is that a term $-\sum_{c\in C}\varphi_c(u_c)$ has been added in the energy. As (10) remains a triplet Markov model, the usual Bayesian segmentation can be performed. Indeed, setting $V=(X,U)$, the field (V,Y) is a classic HMF. However, the distribution of the noise $p(y|x)=\sum_{u\in\Lambda^N}p(u,y|x)$ is complex and does verify neither (H1) nor (H2). Indeed, according to (10) one can see that $p(u,y|x)$ can be seen as a standard HMF, where the field U is hidden Markov whereas the field Y models independent noise (conditionally on U). Then, it is well known that $p(y|x)$, which is the marginal distribution of the HMF $p(y,u|x)$, does verify neither (H1) nor (H2). Let us briefly show on a simple example why the last assertion is true.

Example 3.1. Let us consider just two neighboring pixels s and t. Equation (10) becomes: $p(x_s,x_t,u_s,u_t,y_s,y_t)=p(x_s,x_t)p(u_s,u_t|x_s,x_t)p(y_s|u_s,x_s)p(y_t|u_t,x_t)$. So, $p(y_s,y_t|x_s,x_t)=\sum_{u_s,u_t}p(u_s,u_t|x_s,x_t)p(y_s|u_s,x_s)p(y_t|u_t,x_t)$, $p(y_s|x_s,x_t)=\sum_{u_s,u_t}p(u_s,u_t|x_s,x_t)p(y_s|u_s,x_s)$ and $p(y_t|x_s,x_t)=\sum_{u_s,u_t}p(u_s,u_t|x_s,x_t)p(y_t|u_t,x_t)$. Then, taking $p(u_s,u_t|x_s,x_t)$ such that $p(u_s,u_t|x_s,x_t)\neq p(u_s|x_s,x_t)p(u_t|x_s,x_t)$ and $p(u_s|x_s,x_t)\neq p(u_s|x_s)$, we see that $p(y_s,y_t|x_s,x_t)\neq p(y_s|x_s,x_t)p(y_t|x_s,x_t)$ and $p(y_s|x_s,x_t)\neq p(y_s|x_s)$.

Finally, we note that CGM-HMF is able to take the non-Gaussian and correlated nature of the noise into account simultaneously.

3.3. Blind Gaussian mixture-hidden evidential Markov field (BGM-HEMF)

Let us recall the basic evidential hidden Markov field (HEMF), introduced in [26] and shown to be efficient in unsupervised segmentation of SAR images in [37]. The idea is to notice that the posterior distribution (3) in the basic HMF can be seen as the DS fusion (6) of the Markov field (1) and the probability distribution

$$q^y(x)=\prod_{s\in S}\frac{p(y_s|x_s)}{\sum_{l\in L}p(y_s|x_s=l)}. \tag{11}$$

Indeed, recalling $p(x)=\gamma\exp\left[-\sum_{c\in C}\psi_c(x_c)\right]$ and $p(y|x)=\prod_{s\in S}p(y_s|x_s)$, we have

$$\begin{aligned} p(x|y) &= \frac{p(x)p(y_s|x_s)}{p(y)} \\ &= \frac{\exp\left[-\sum_{c\in C}\psi_c(x_c)\right]\prod_{s\in S}p(y_s|x_s)}{p(y)} \end{aligned}$$

$$= \left(\frac{\sum_x \prod_{s\in S} p(y_s|x_s)}{p(y)}\right)\left(\frac{\exp\left[-\sum_{c\in C}\psi_c(x_c)\right]\prod_{s\in S} p(y_s|x_s)}{\sum_x \prod_{s\in S} p(y_s|x_s)}\right)$$
$$= \left(\frac{\sum_x \prod_{s\in S} p(y_s|x_s)}{\gamma\, p(y)}\right)\left(\gamma \exp\left[-\sum_{c\in C}\psi_c(x_c)\right]\prod_{s\in S}\frac{p(y_s|x_s)}{\sum_{x_s} p(y_s|x_s)}\right)$$
$$= \gamma(y)p(x)q^y(x)$$

with $\gamma(y) = \frac{\sum_x \prod_{s\in S} p(y_s|x_s)}{\gamma\, p(y)}$. This means that $p(x|y)$ is DS fusion between Markov $p(x)$ and $q^y(x)$.

Then, the Markov field of (1) is extended to an "evidential" Markov field (EMF) $M = (M_s)_{s\in S}$, where each M_s takes its values in the powerset 2^L instead of L. Thus $p(m) \propto \exp\left[-\sum_{c\in C}\psi_c(m_c)\right]$. As (11) is a probability, the DS fusion of such an EMF with (11) gives a probability $p(x|y)$, which is written $p(x|y) \propto \sum_{x\in m} p(m)q^y(x)$, where the sum $\sum_{x\in m}$ is made over all $m = (m_s)_{s\in S} \in [2^L]^N$ such that $x_s \in m_s$ for each $s \in S$. Then one observes that computing DS fusion is equivalent to computing $p(x|y)$ in the triplet Markov field (X, M, Y), where $(X, M) = (X_s, M_s)_{s\in S}$ takes its values in the subset Λ of $L \times 2^L$ of couples (x_s, m_s) verifying $x_s \in m_s$. Thus DS fusion can be seen as computation of a marginal distribution in a probabilistic triplet Markov field, which is a core point making link between probabilistic Markov fields and theory of evidence.

Finally, an "evidential hidden Markov field" (EHMF) is mathematically equivalent to a TMF (X, M, Y), whose distribution is of the form:

$$p(x, m, y) \propto 1_{x\in m} \exp\left[-\sum_{c\in C}\psi_c(m_c) + \sum_{s\in S}\log(p(y_s|x_s))\right] \tag{12}$$

and is defined on a subset of $[L \times 2^L \times \mathbb{R}]^N$ of (x_s, m_s, y_s) such that $x_s \in m_s$. We see how EHMFs extend HMFs: when this subset is reduced to (x_s, m_s, y_s) such that $m_s = \{x_s\}$, EHMF is reduced to a HMF.

Example 3.2. Let us consider the standard EHMF, which turned out to improve HMF's results in situations where the label field contains fine details [26]. The subset of $[L \times 2^L \times \mathbb{R}]^N$ of (x_s, m_s, y_s) such that $x_s \in m_s$ is reduced to two elements of the form $(x_s, \{x_s\}, y_s)$, (x_s, L, y_s). Thus for K labels, the number of "new labels" when extending HMF to HEMF is only multiplied by two. However, as we will see in section 4, this slight extension will be of importance in finding fine details when segmenting noisy images.

Blind Gaussian mixture-hidden evidential Markov field (BGM-HEMF) is obtained from HEMF in a similar way that BGM-HMF is obtained from HMF, by introducing an auxiliary field U. We obtain:

$$p(x, m, u, y) \propto 1_{x\in m} \exp\left[-\sum_{c\in C}\psi_c(m_c) - \sum_{s\in S}\eta_s(x_s, u_s) + \sum_{s\in S}\log(p(y_s|x_s, u_s))\right] \tag{13}$$

Some examples of application of BGM-HMF can be seen in [25].

3.4. Correlated Gaussian mixture-hidden evidential Markov field (CGM-HEMF)

Finally, the new "correlated Gaussian mixture-HEMF" (CGM-HEMF) model proposed is obtained from BGM-HEMF in a similar manner that CGM-HMF is obtained from BGM-HMF, by considering potential functions φ_c. We get:

$$p(x, m, u, y) \propto 1_{x\in m} \exp\left[-\sum_{c\in C}\psi_c(m_c) - \sum_{c\in C}\varphi_c(u_c) - \sum_{s\in S}\eta_s(x_s, u_s) + \sum_{s\in S}\log(p(y_s|x_s, u_s))\right] \tag{14}$$

As it remains a triplet Markov field, with the third discrete finite field being (M, U), searching X from Y by conventional Bayesian methods, like Iterated Conditional Modes (ICM) [38], maximum posterior marginal (MPM) [39], maximum *a posteriori* (MAP) [40], or graph cut [41], remains feasible.

3.5. Parameter estimation and Bayesian segmentation

All models above are of the form $p(w, y) \propto \exp\left[-\sum_{c\in C}\xi_c(w_c) + \sum_{s\in S}\log(p(y_s|w_s))\right]$. Indeed, $W = (X, U)$ in BGM-HMF and CGM-HMF; and $W = (X, M, U)$ in BGM-HEMF and CGM-HEMF. Thus at the processing level, they can finally be dealt with as in the hidden Markov field context. To perform segmentation we use ICM. We scan the image and compute, at each $s \in S$, the distribution $p(w_s|w_{\mathcal{N}_s}, y_s)$, where $\mathcal{N}_s$ is the neighborhood of s. Then at scan n, the $\hat{w}_s^n$ set at s is the label maximizing $p(w_s|w_{\mathcal{N}_s}, y_s)$. After few scans – about five or six – the estimated image $\hat{w}^n = (\hat{w}_s^n)_{s\in S}$ stabilizes and is considered as segmentation result [42].

For parameter estimation, we use the method specified in [4]. Let us set $\theta = (\alpha, \beta)$ with α being the parameters defining potential functions ξ_c and β being the ones defining Gaussian densities $p(y_s|w_s) = p(y_s|u_s, x_s)$. We use particular Iterative Conditional Estimation (ICE) method producing a sequence $\theta^0, \theta^1, ..., \theta^q, \theta^{q+1}, ...$ according to:

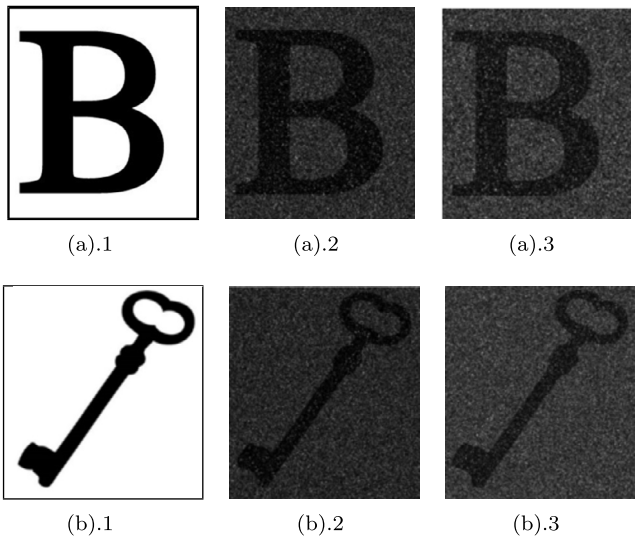


Fig. 2. Synthetic binary class images corrupted by independent and correlated noises. 1. Original class-images; 2. Synthetic image corrupted by independent noise; 3. Synthetic image corrupted by correlated noise; of: (a) Letter B; and, (b) Key.

(i) initialization θ^0 is given by automatic histogram thresholding segmentation;
(ii) to compute θ^{q+1} from θ^q do:
- sample w^q according to $p(w|y,\theta^q)$, using Gibbs sampler;
- use (w^q, y) to estimate β^{q+1} using classic estimators;
- use w^q to estimate α^{q+1} with least squares fit [43];

(iii) stop when an end criterion is satisfied (in this study, when the number of changing pixels is less than 50).

Let us notice that "iterative conditional estimation" (ICE) method is different from the expectation-maximization (EM) one, even if can present some similarities. The main difference is that ICE uses conditional expectation while EM uses the maximum likelihood [44]. In particular the proposed method is different from the EM one proposed in [45].

Remark 3.1. Our paper deals with hidden evidential Markov fields obtained by extension to evidential masses of classic hidden Markov fields. Similar extension of classic hidden Markov chains (HMCs [46,47]) leads to "evidential Markov chains" (HEMCs [48,33]). Some of extensions of the classic hidden Markov chains use imprecise probability theory [49–51], or hidden fuzzy sets [52–54]. The difference with HEMCs is that in these different extensions uncertainty is dealt with at imprecise transitions level, while in HEMCs the class chain searched for is not Markov, and thus transitions are not considered. However, different ideas proposed in imprecise Markov chains could possibly be transposed to Markov fields, and then models so obtained should be compared to different HEMFs considered in this paper.

4. Experiments

In this section, we assess the performance of our proposed models against some conventional models described in section 3. Besides, we also compare the CGM-HEMF with the recent method SVA-B [55]. In the first subsection, we test BGM-HMF, CGM-HMF and SVA-B on simulated and real images, while the second paragraph is devoted to BGM-HEMF, CGM-HEMF and SVA-B, tested on simulated, real images and Berkeley dataset [56]. The comparation of [55] with [57–60] has shown the superiority of [55]. For all experiments, the performance of each model will be evaluated both qualitatively and quantitatively. For this purpose, all approaches will be assessed with respect to the reference map in terms of overall accuracy (OA) and Kappa coefficient (κ) [61]. A classification is better as the OA, resp. κ, is higher. While overall accuracy may appear more intuitive, such metric may introduce a significant bias in the case of unbalanced classes. Hence, the Kappa coefficient (κ) is usually adopted as a universal metric for objective comparison purpose.

4.1. Unsupervised segmentation using BGM-HMF and CGM-HMF

We present two series of experiments: simulated and real images. The first series deals with two sampled class-images corrupted by non-Gaussian correlated noise. Let us consider the two label-images "Letter B" and "Key" presented in Fig. 2. Let X be the label field with each X_s taking its values from $L = \{l_1, l_2\}$ where l_1 and l_2 corresponds to black pixels and white ones respectively. The non-Gaussian noise considered here is the Gamma one. In independent case, the two densities

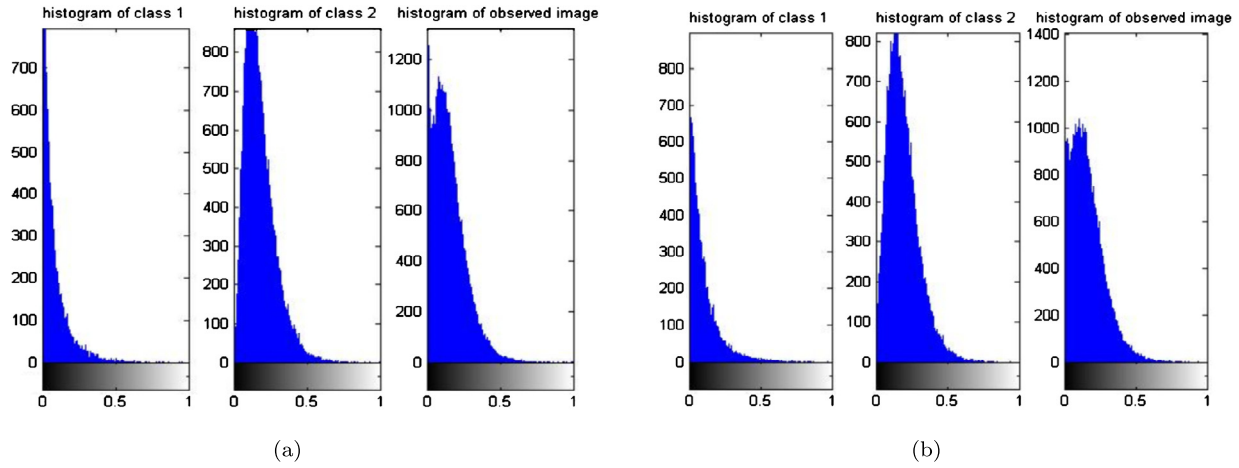


Fig. 3. Histograms of Letter B image in the: (a) Independent noise case; and, (b) Correlated noise case. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

related to the two classes are Gamma $\mathcal{G}_1(0.5, 2)$ and $\mathcal{G}_2(3, 1)$ given by the densities: $f_1(y) = \frac{y^{0.5-1}e^{-\frac{y}{2}}}{2^{0.5}\Gamma(0.5)}$ and $f_1(y) = \frac{y^{3-1}e^{-\frac{y}{1}}}{1^3\Gamma(3)}$ respectively. Histograms in Fig. 3(a) show that they are quite different from Gaussian densities.

On the other hand, the correlated noise is obtained as follows:

$$y_{i,j} = \left[\frac{(y^k_{i,j} - \mu^k) + (y^k_{i-1,j} - \mu^k) + (y^k_{i,j-1} - \mu^k)}{\sqrt{3}} + \mu^k \right] \tag{15}$$

where i, j is the location of the pixel; y^1 and y^2 are two independent noises with the densities being $\mathcal{G}_1(0.5, 2)$ and $\mathcal{G}_2(3, 1)$ respectively; $\mu_1 = 1$ and $\mu_2 = 3$ are the means; l_k is the label of the pixel with $k \in \{1, 2\}$. This produces a sampled image exhibiting a complicated correlated noise. More explicitly, one obtains a correlation coefficient of 0.23 (evaluated from the resulted image). We show the histograms of noises and the related observed images in Fig. 2 and Fig. 3.

The presented results and other similar results we obtained allow drawing the following conclusions:

- for the independent noise SVA-B performs a little better than BGM-HMF and CGM-HMF, but for correlated noise CGM-HMF performs much better than SVA-B;
- in the case of independent noise, CGM-HMF yields comparable results to BGM-HMF. This could appear as evident; however, this is not in the unsupervised case considered here. In fact, this means that the parameter estimation method used in more complex CGM-HMF is able to find again the simpler BGM-HMF;
- the most interesting conclusion is that the noise correlation is taken into account by CGM-HMF which results in significant improvement of unsupervised segmentation. Indeed, CGM-HMF performs as well on correlated noise context as on independent noise one. (See Figs. 4 and 5.)

In the second series, we consider two cases of SAR images. In the first one, shown in Fig. 6(a), the ground truth considered is hand-drawn. The second image “Amazonia”, shown in Fig. 6(b), has a genuine ground truth. Visual results are given in Fig. 6, while ratios of recovered classes and Kappa coefficient are specified in Table 1.

According to the presented results we can state that CGM-HMF is able to take the noise correlation into account, which results in improvement of well classified pixels rate and Kappa coefficient as well.

4.2. Unsupervised segmentation using BGM-HEMF and CGM-HEMF

4.2.1. Hand written images segmentation

Let us first apply BGM-HEMF and CGM-HEMF to the same synthetic images as in Section 4.1. The results obtained are provided in Fig. 7 and 8. A comparative performance evaluation is also provided through quantitative metrics in Table 2.

We can see that results are comparable when overall accuracy or Kappa coefficient are considered. However, errors are not made in a similar manner and results are visually somewhat different. In particular, edges seem to be better found by the evidential models. These differences will give a striking advantage to evidential models when label images contain fine details we are going to consider now.

Let us consider three images: “Nazca bird”, “Fingerprint”, and “Tree”, presented in Fig. 9. Such images are noised using the same independent and correlated noise as in the Section 4.1.

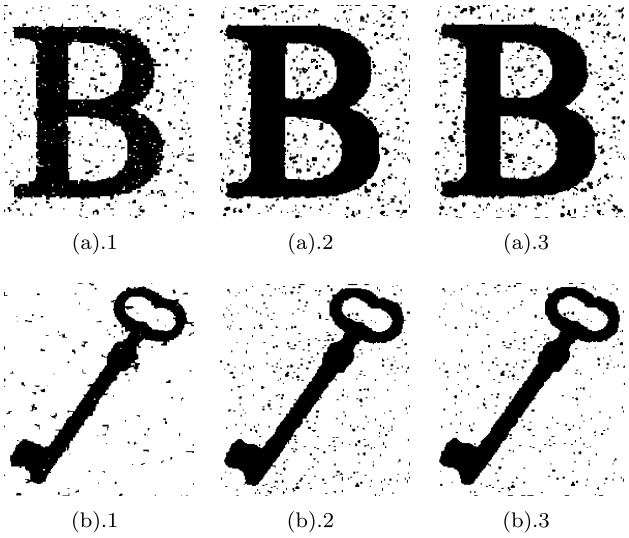


Fig. 4. Unsupervised segmentation results of synthetic image corrupted by independent noise: (a) Letter B; and, (b) Key; using 1. SVA-B; 2. BGM-HMF; and, 3. CGM-HMF.

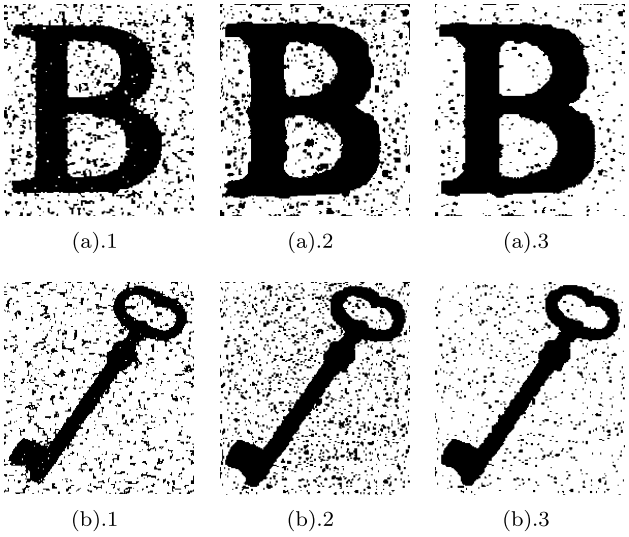


Fig. 5. Unsupervised segmentation results of synthetic image corrupted by correlated noise: (a) Letter B; and, (b) Key; using 1. SVA-B; 2. BGM-HMF; and, 3. CGM-HMF.

We segment the observed images by SVA-B, BGM-HMF, CGM-HMF, BGM-HEMF and CGM-HEMF respectively. Results corresponding to independent noise are given in Fig. 10, while results corresponding to correlated noise are given in Fig. 11. Overall, accuracy and Kappa coefficients are given in Table 3.

According to these different results, we can state the following:

(i) SVA-B incorporates spatial information and Potts model to establish a tradeoff between fine details and noises. However, CGM-HEMF performs still better, especially in the case of correlated noise;
(ii) BGM-HEMF and CGM-HEMF seem to benefit from the efficiency of BGML in parts of images presenting fine details, and from the efficiency of BGM-HMF and CGM-HMF in homogeneous parts. A theoretical partial explanation of this phenomenon is given in Remark 4.1 below;
(iii) when the noise is independent both BGM-HEMF and CGM-HEMF give similar results. This means that the parameters estimation method applied to CGM-HEMF is robust enough to find that the original data is not correlated;
(iv) the most important ascertainment is to notice that when the noise is correlated, unsupervised segmentation based on the new CGM-HEMF can strongly improve the results obtained with BGM-HEMF.

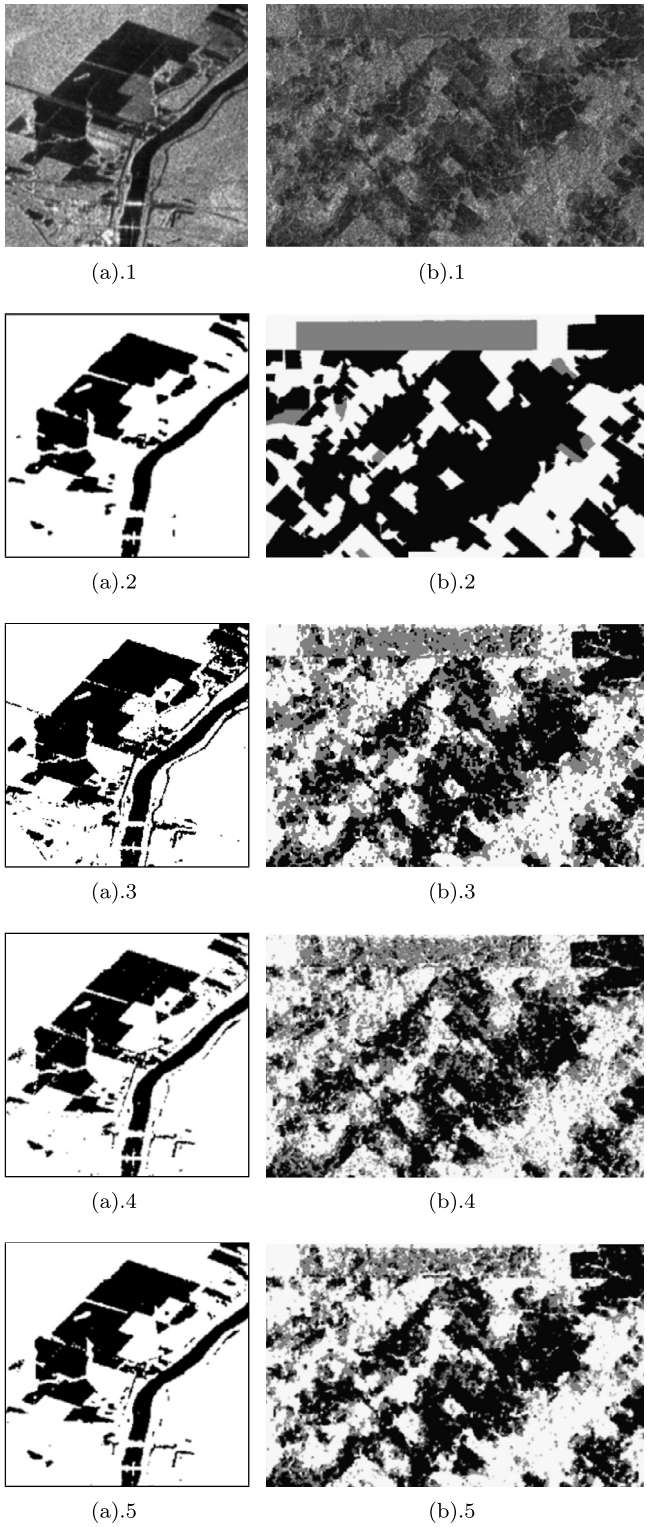


Fig. 6. Unsupervised segmentation of real SAR images. 1. Original image; 2. Ground truth; 3. Segmentation result using SVA-B; 4. Segmentation result using BGM-HMF; and, 5. Segmentation result using CGM-HMF; of (a) Jet propulsion; and, (b) Amazonia.

Table 1
Performance evaluation of different approaches on real SAR images.

Image	SVA-B	BGM-HMF	CGM-HMF
Overall accuracy (%)			
Jet propulsion	91.70	96.66	**97.55**
Amazonia	63.23	61.48	**66.64**
Kappa			
Jet propulsion	0.800	0.912	**0.935**
Amazonia	0.454	0.411	**0.485**

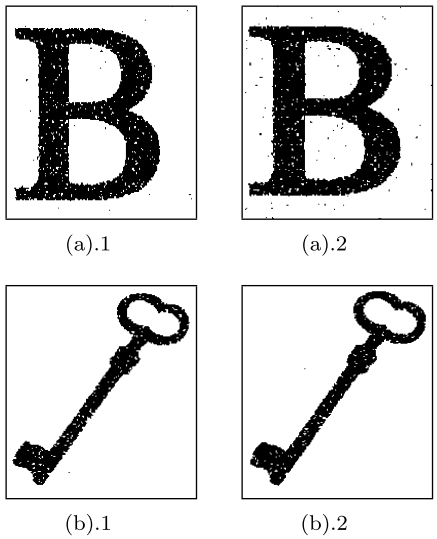


Fig. 7. Unsupervised segmentation results of synthetic image corrupted by independent noise: (a) Letter B; and, (b) Key; using 1. BGM-HEMF; and, 2. CGM-HEMF.

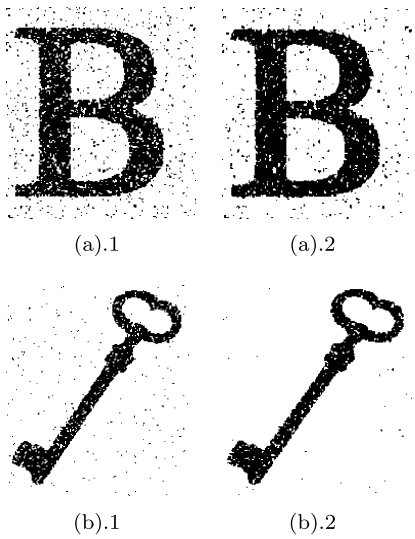


Fig. 8. Unsupervised segmentation results of synthetic image corrupted by correlated noise: (a) Letter B; and, (b) Key; using 1. BGM-HEMF; and, 2. CGM-HEMF.

Comparing results Fig. 11(a).1–11(c).1 with Fig. 11(a).2–11(c).2 is of interest as it probably explains the good results obtained using the proposed CGM-HEMF. We can see that results of Fig. 11(a).1–11(c).1 are bad as the noise is strong and spatial information is not taken into account. However, fine details are not lost, even if they are not found efficiently. On the contrary, results of Fig. 11(a).2–11(c).2 are good outside fine details, but these details are lost. Introducing evidential

Table 2
Performance evaluation of different approaches on synthetic images.

	Image	SVA-B	BGM-HMF	CGM-HMF	BGM-HEMF	CGM-HEMF
				Overall accuracy (%)		
A	Letter B	94.92	93.82	93.29	**95.13**	94.75
	Key	97.67	96.79	97.29	97.78	**98.01**
B	Letter B	89.36	87.81	**92.88**	88.33	92.07
	Key	90.27	91.88	96.18	93.15	**97.17**
				Kappa		
A	Letter B	0.887	0.866	0.851	**0.886**	0.879
	Key	0.904	0.872	0.890	0.897	**0.909**
B	Letter B	0.775	0.747	**0.846**	0.733	0.819
	Key	0.674	0.718	0.850	0.777	**0.870**

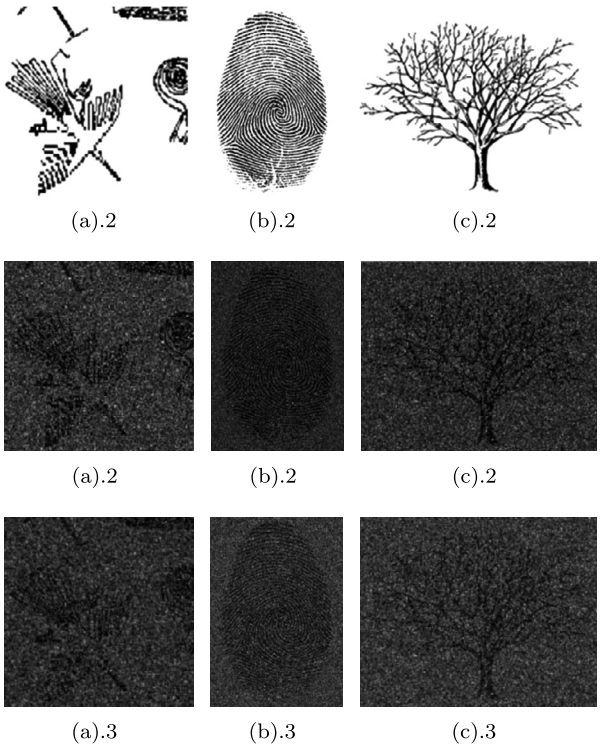


Fig. 9. Synthetic images. 1. Original class-images; 2. Synthetic image corrupted by independent noise; 3. Synthetic image corrupted by correlated noise; of: (a) Nazca bird; (b) Fingerprint; and, (c) Tree.

variables allows one to benefit from these two advantages simultaneously. To better understand why, let us consider a case of two neighboring pixels s and t and two labels $L = \{l_1, l_2\}$. In the probabilistic case, we have $p(x_s, x_t, y_s, y_t) = p(x_s, x_t)p(y_s|x_s)p(y_s|x_s)$. In the evidential case, we have additional variables M_s and M_t such that (M_s, M_t) takes its values in $\{\{l_1\}, \{l_2\}, L\}^2$, and $p(x_s, m_s, y_s, x_t, m_t, y_t) \propto 1_{(x_s,x_t)\in(m_s,m_t)}p(m_s, m_t)p(y_s|x_s)p(y_s|x_s)$. We see that when $p(m_s, m_t)$ mainly charges $\{\{l_1\}, \{l_2\}\}^2$, we are close to the probabilistic case, and we find again this case when $p(m_s, m_t)$ is null outside $\{\{l_1\}, \{l_2\}\}^2$. Similarly, when $p(m_s, m_t)$ mainly charges (L, L) we are close to the independent case $p(x_s, m_s, y_s, x_t, m_t, y_t) \propto p(y_s|x_s)p(y_s|x_s)$, and we find again this case when $p(m_s, m_t)$ is null outside $\{(L, L)\}$. This remark extended to the whole set of pixels in the Markov field context illustrates how the proximity to the first or to the second case can be managed according to the place of a pixel in the image. Such interpretation is confirmed by experiments presented in Fig. 12, where we give estimation of $M = (M_s)_{s\in S}$. Indeed, we find that $\hat{m}_s = L$ in spots with fine details. The interesting point is that the segmentation is unsupervised, and thus the right balance between these two aspects of the model seems to be automatically found, for each pixel, in the parameter estimation stage. Let us remark that similar remarks have been made in the context of hidden Markov chains in [33], also see [62].

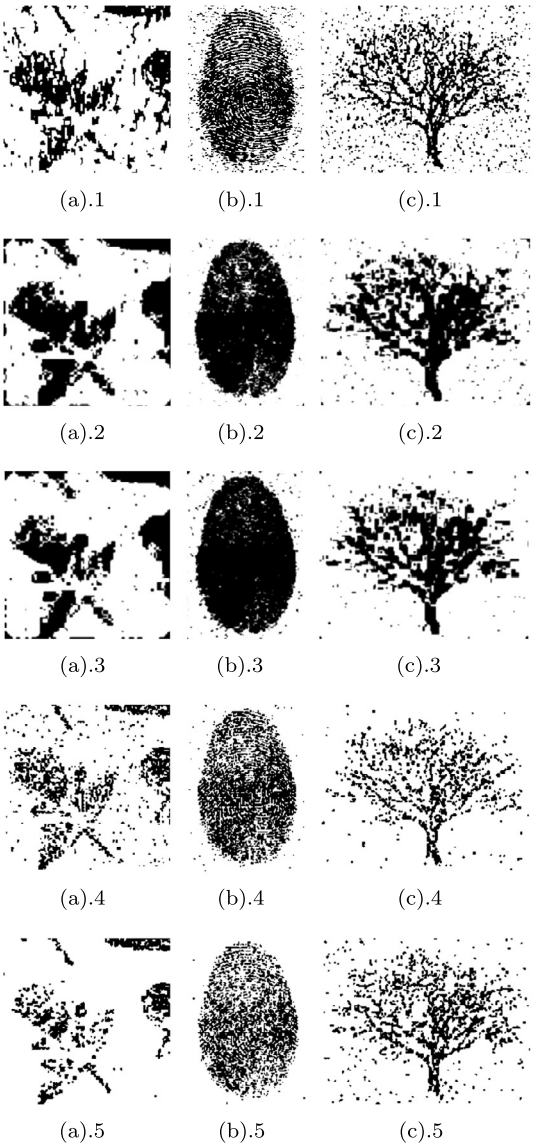


Fig. 10. Unsupervised segmentation of synthetic images corrupted by independent noise: (a) Nazca bird; (b) Fingerprint; and, (c) Tree, by: 1. SVA-B; 2. BGM-HMF; 3. CGM-HMF; 4. BGM-HEMF; and, 5. CGM-HEMF.

4.2.2. Unsupervised segmentation of real images

This part is devoted to the segmentation of real images. More explicitly we consider one SAR image (Fig. 13(a)), and two optical images: a gray-level image "Landscape 1" (Fig. 14(a)); and a color image "Landscape 2" (Fig. 15(a)). The first two images are to be segmented into three classes whereas the last one is to be segmented into five classes.

As well acknowledged, the histogram of SAR image is not Gaussian distribution, and SAR image is heavily interfered by speckle noise. Thus it is challenging to process SAR data. Besides, according to Fig. 13(a), the data used in our paper includes plenty of small edges, which makes it even more difficult to enhance the robustness against noise and keep the edges simultaneously. We segment the data by above five approaches and show the result in Fig. 13. According to the results obtained, SVA-B, BGM-HMF and CGM-HMF are robust against noise but cannot segment the fine lines. BGM-HEMF and CGM-HEMF are robust against noise and also keep the edges. Besides, CGM-HEMF performs better than BGM-HEMF in some spots; in particular around the rich-edge area. We can see from the red panels that the segmentation obtained by CGM-HEMF includes more details with respect to the one obtained through BGM-HEMF.

For the optical image "Landscape 1", it is reasonable to assume the histogram of one specific class as Gaussian, but when there are some subsets within one class the distribution will be Gaussian mixture. Thus except the SAR data, we also test the proposed method on complex optical image as shown in Fig. 14(a). This optical image includes a rich-edge area (the tree) and a stationary area (the field). Besides, according to its histogram there are some subsets in this image. According

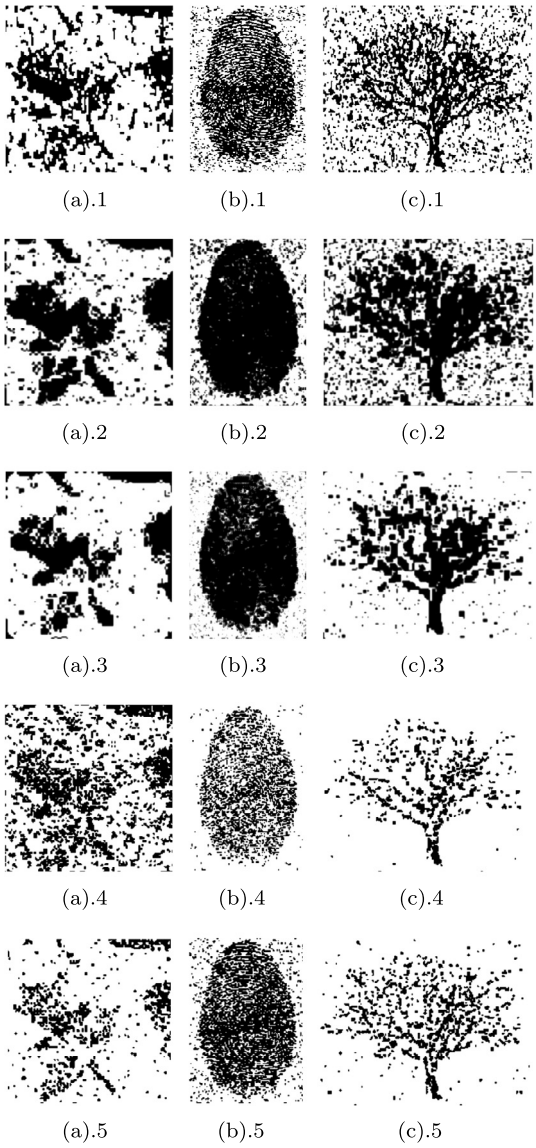
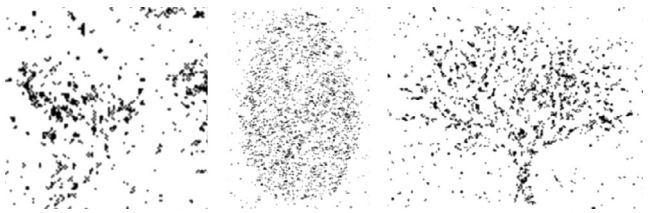


Fig. 11. Unsupervised segmentation of synthetic images corrupted by correlated noise: (a) Nazca bird; (b) Fingerprint; and, (c) Tree, by: 1. SVA-B; 2. BGM-HMF; 3. CGM-HMF; 4. BGM-HEMF; and, 5. CGM-HEMF.

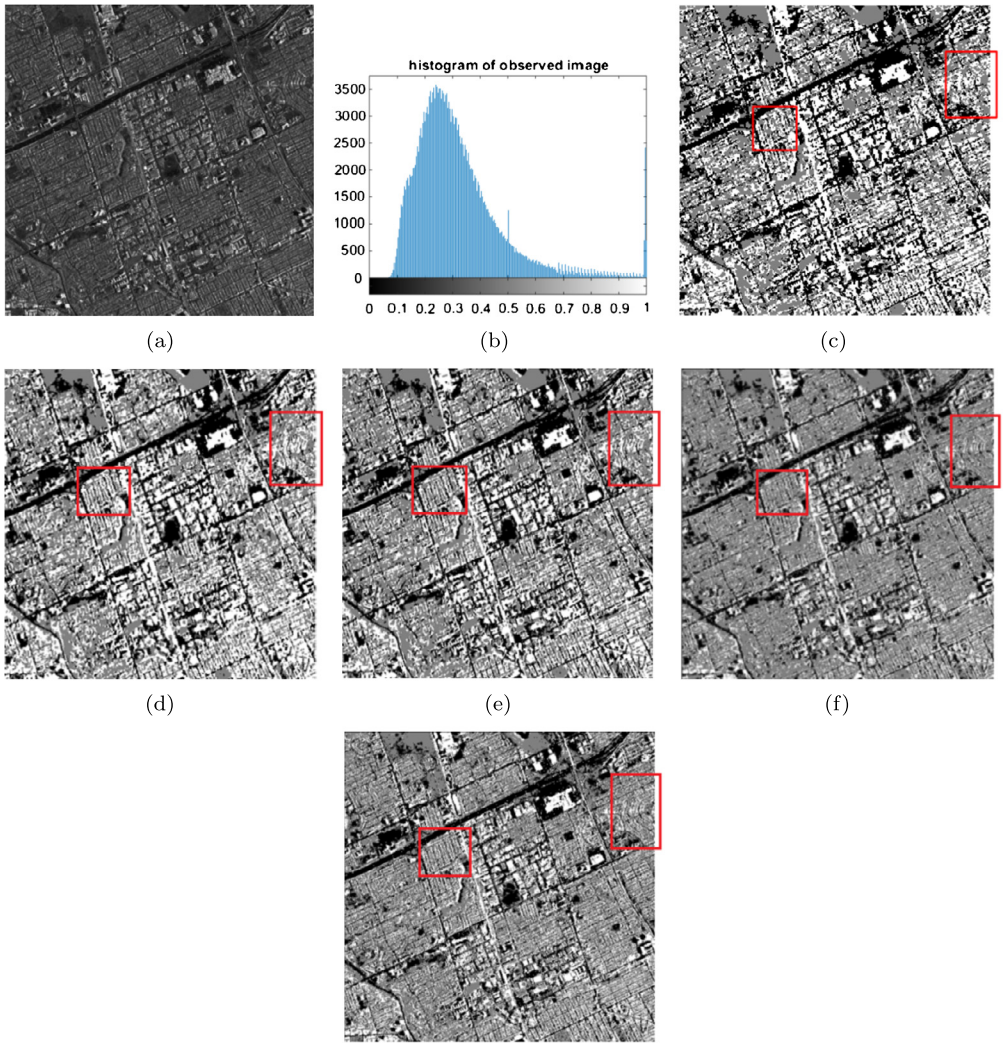
Table 3
Performance evaluation of different approaches on synthetic images.

	Image	SVA-B	BGM-HMF	CGM-HMF	BGM-HEMF	CGM-HEMF
				Overall accuracy (%)		
A	Nazca	86.15	43.02	44.33	90.47	**90.85**
	Fingerprint	82.62	83.11	81.51	87.50	**87.88**
	Tree	85.69	81.42	84.30	**90.17**	89.40
B	Nazca	73.86	50.56	54.33	85.69	**88.41**
	Fingerprint	76.81	73.74	75.41	83.50	**85.39**
	Tree	76.62	63.23	80.02	87.34	**90.70**
				Kappa		
A	Nazca	0.5633	0.1261	0.1315	0.6166	**0.6193**
	Fingerprint	0.6135	0.6203	0.5593	0.6811	**0.6963**
	Tree	0.5235	0.4761	0.4971	**0.5760**	0.5570
B	Nazca	0.4223	0.1696	0.2824	0.5468	**0.5976**
	Fingerprint	0.4817	0.4701	0.4998	0.5484	**0.5981**
	Tree	0.3790	0.2365	0.4070	0.5541	**0.6005**



(a) (b) (c)

Fig. 12. Estimates of $\mathcal{L}$ extracted from the estimated evidential field $\hat{M} = (\hat{M}_s)_{s \in S}$, in the CGM-HEMF case. Pixels are labeled black if $\hat{M}_s = L$, and white if either $\hat{M}_s = \{l_1\}$, or $\hat{M}_s = \{l_2\}$. (a) Nazca bird; (b) Fingerprint; and, (c) Tree.



(a) (b) (c)

(d) (e) (f)

(g)

Fig. 13. Unsupervised segmentation of a real SAR image. (a) Original image; (b) Image histogram; (c) SVA-B segmentation; (d) BGM-HMF segmentation; (e) CGM-HMF segmentation; (f) BGM-HEMF segmentation; and, (g) CGM-HEMF segmentation.

to Fig. 14, the segmentation results obtained by correlated-likelihood based methods experience less noises than the ones obtained by independent-likelihood based methods. BGM-HMF and CGM-HMF perform well in enhancing the robustness against noise, but they cannot segment the branches of the tree well. BGM-HEMF and CGM-HEMF can still segment the image better.

Let us now consider the color optical image “Landscape 2”. Segmentation is applied on the gray levels at each of the three primitive colors and joint probability is computed simply as the product of the three posterior probabilities (which subsumes

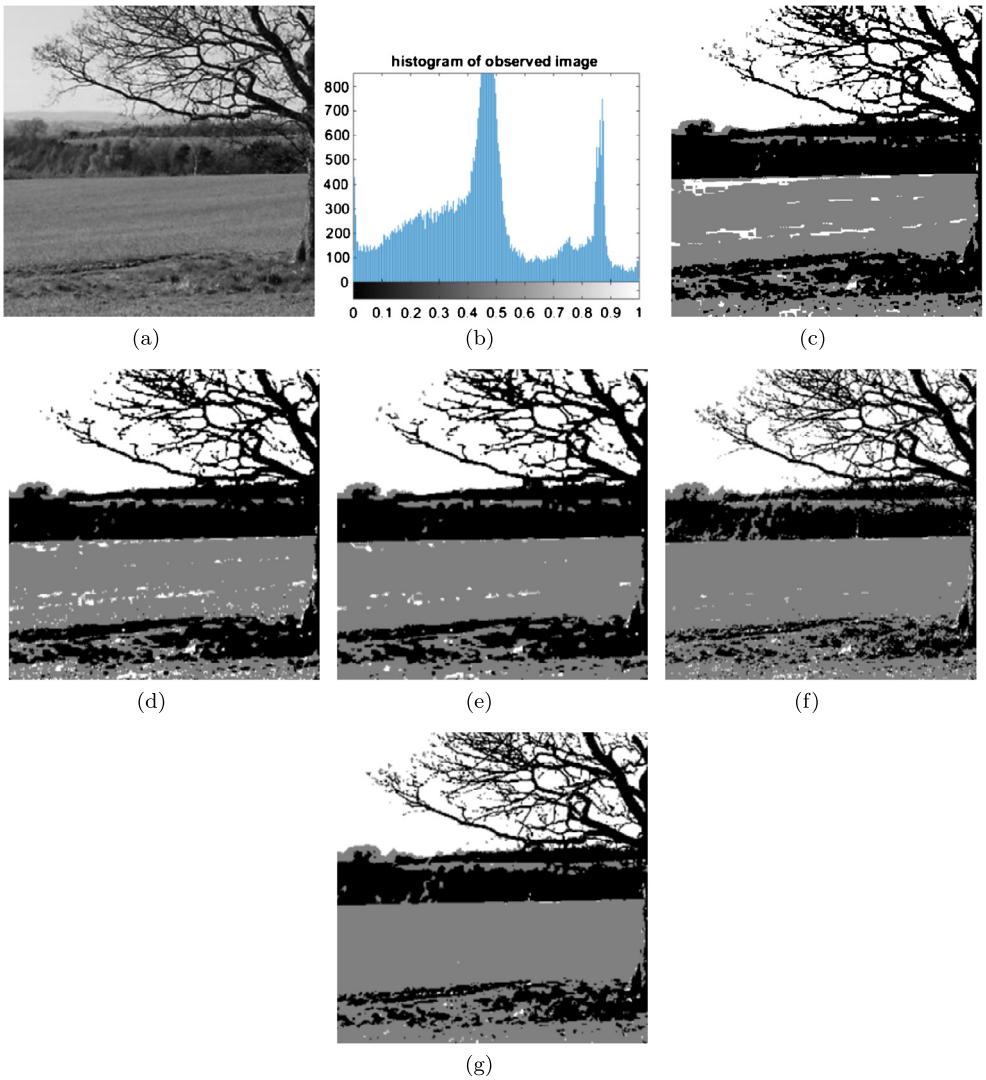


Fig. 14. Unsupervised segmentation of an optical image "Landscape". (a) Original image; (b) Image histogram; (c) SVA-B segmentation; (d) BGM-HMF segmentation; (e) CGM-HMF segmentation; (f) BGM-HEMF segmentation; and, (g) CGM-HEMF segmentation.

independence of the three features). One can see from the results provided in Fig. 15 that HEMF-based segmentation allows one to find the road which is a rich-edge region while taking advantages of Markov regularization. One can conclude that HEMF-based segmentation establishes an adequate tradeoff between regularization ability and edge preserving. This confirms the interest of introducing theory of evidence.

Finally, estimates of some parameters of CGM-HEMF related to some data considered in this study are provided in Table 4.

Results in Table 4 are related to BGM-HEMF (14), with cliques of second order. Thus c in $\varphi_c(m_c)$ and $\phi_c(m_c)$ have two neighbors (s,t), which can be horizontal or vertical. For horizontal neighbors (s,t) we set $\begin{cases} -\alpha_n^h & \text{for} \quad m_s = m_t \\ \alpha_n^h & \text{for} \quad m_s \neq = m_t \end{cases}$, and similarly for vertical neighbors with α_m^v instead of α_m^h. Similarly for $\phi_{(s,t)} = (u_s, u_t)$, which thus are defined with α_u^v and α_u^h. We can notice that ideal α_u^v and α_u^h should be null for non-correlated noise, which is not case and thus parameter estimation method seems little robust with respect to these parameters. However, this is possibly due to slow variation of the whole model with them, and thus their little influence on segmentation result.

4.2.3. Segmentation of Berkeley dataset

To make strong conclusions, we also consider the Berkeley segmentation dataset to evaluate the proposed methods along with the existing ones. To this end, we choose 50 images exhibiting fine details, that we will refer to as fine details

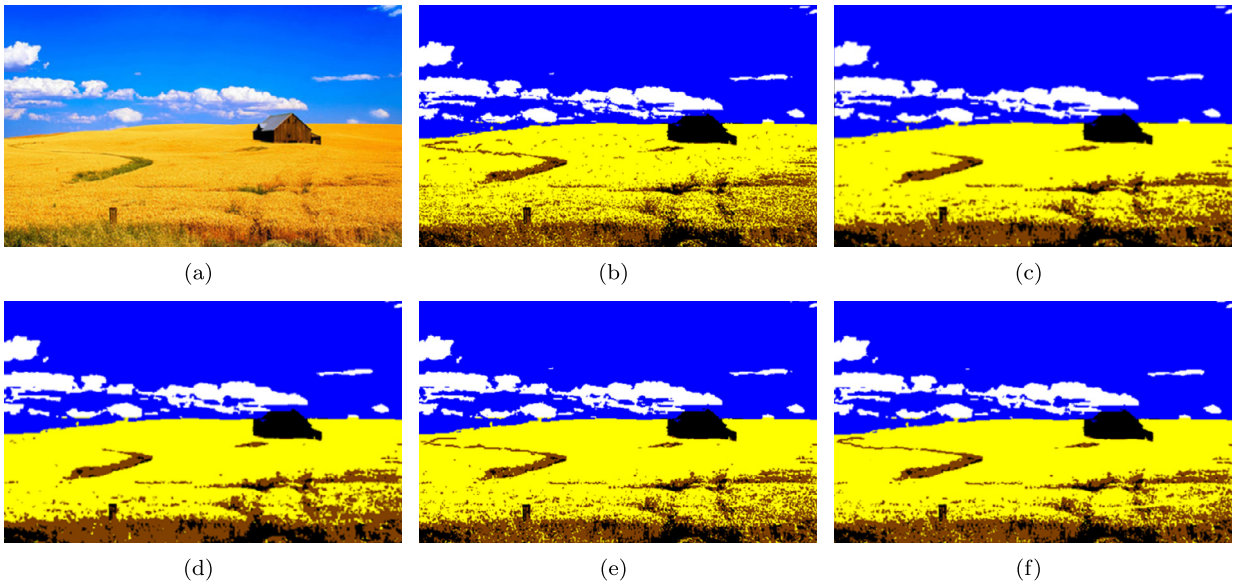


Fig. 15. Unsupervised segmentation of a color optical image "Landscape 2". (a) Original image; (b) SVA-B segmentation; (c) BGM-HMF segmentation; (d) CGM-HMF segmentation; (e) BGM-HEMF segmentation; and, (f) CGM-HEMF segmentation.

Table 4
Estimated parameter according to CGM-HEMF model.

Parameter	Indep. Letter B	Corr. Letter B	Indep. Nazca bird	Corr. Nazca bird	Jet propulsion	Landscape
α_m^h	0.8546	0.8722	0.3233	0.3848	0.7655	0.4553
α_m^v	0.6565	0.7014	0.4012	0.3992	0.6980	0.5334
α_u^h	0.1725	0.3022	0.1056	0.2020	0.2432	0.2873
α_u^v	0.1346	0.3123	0.1043	0.2143	0.2544	0.2489

Table 5
Performance evaluation of different approaches on Berkeley dataset.

Image	SVA-B	BGM-HMF	CGM-HMF	BGM-HEMF	CGM-HEMF
			Overall accuracy (%)		
FDI	83.45	76.31	80.25	87.11	**91.02**
NFDI	94.05	91.31	**94.56**	90.23	94.34
			Kappa		
FDI	0.6071	0.5002	0.5229	0.6237	**0.7043**
NFDI	0.8006	0.7933	**0.8490**	0.7423	0.8181

images (FDI); and 50 images without fine details, that we will refer to as non-fine details images (NFDI). The average overall accuracy and Kappa coefficient are illustrated in Table 5.

For NFDI, all the CGM-HMF, CGM-HEMF and SVA-B perform well and yield comparable overall accuracy and Kappa coefficient. On the other hand for FDI, BGM-HEMF and CGM-HEMF perform much better than SVA-B, BGM-HMF and CGM-HMF. Indeed, CGM-HEMF benefits from the following three aspects: (i) Markov regularization which allows to take into account the spatial contextual information; (ii) theory of evidence which establishes a good tradeoff between noise and prior information; and (iii) considering correlated noise which allows to model the actual noise more efficiently.

Remark 4.1. The ability of detecting fine details and large-size classes observed in Markov fields with correlated noise context would probably be kept in mono-dimensional problems modeled with Markov chains. Thus extending the evidential Markov chain proposed in [48,62] probably is an interesting perspective.

5. Conclusion

In this paper, we have proposed two original Markov field models aiming to segment images corrupted with non Gaussian and correlated noise. Both models have been assessed in unsupervised segmentation. The first one, probabilistic, efficiently improves results when the noise is correlated. The second one, which extends former "evidential" models [25],

turns out to be particularly interesting when images simultaneously contain homogeneous parts (regions of a same label of large size) and inhomogeneous parts (regions of a same label of small size). An interesting future direction would be to consider by considering other compound hypotheses than the frame of discernment L in the case of multiclass data classification.

Acknowledgements

This work was supported by the Natural Science Foundation of China No. 61271297, No. 61272281, and No. 61301284, and the Specialized Research Fund for the Doctoral Program of Higher Education No. 20130203120006.

References

[1] M. Gong, L. Su, M. Jia, W. Chen, Fuzzy clustering with a modified MRF energy function for change detection in synthetic aperture radar images, IEEE Trans. Fuzzy Syst. 22 (1) (2014) 98–109.
[2] L. An, M. Li, P. Zhang, Y. Wu, L. Jia, W. Song, Discriminative random fields based on maximum entropy principle for semisupervised SAR image change detection, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (8) (2016) 3395–3404.
[3] A. Baraldi, T. Wassenaar, S. Kay, Operational performance of an automatic preliminary spectral rule-based decision-tree classifier of spaceborne very high resolution optical images, IEEE Trans. Geosci. Remote Sens. 48 (9) (2010) 3482–3502.
[4] 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.
[5] D. Tomazevic, B. Likar, T. Slivnik, F. Pernus, 3-D/2-D registration of CT and MR to X-ray images, IEEE Trans. Med. Imaging 22 (11) (2003) 1407–1416.
[6] M. Chegini, H. Ghassemian, Spatial spectral Gaussian mixture model approach for automatic segmentation of multispectral MR brain images, in: 19th Iranian Conference on Electrical Engineering, IEEE, 2011, pp. 1–6.
[7] H.S. Choi, D.R. Haynor, Y. Kim, Partial volume tissue classification of multichannel magnetic resonance images—a mixel model, IEEE Trans. Med. Imaging 10 (3) (1991) 395–407.
[8] Y. Zhang, M. Brady, S. Smith, Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm, IEEE Trans. Med. Imaging 20 (1) (2001) 45–57.
[9] 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.
[10] J. Liu, Y.-H. Yang, Multiresolution color image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 16 (7) (1994) 689–700.
[11] M. Hayat, M. Bennamoun, S. An, Deep reconstruction models for image set classification, IEEE Trans. Pattern Anal. Mach. Intell. 37 (4) (2015) 713–727.
[12] 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.
[13] V.A. Krylov, G. Moser, S.B. Serpico, J. Zerubia, Supervised high-resolution dual-polarization SAR image classification by finite mixtures and copulas, IEEE J. Sel. Top. Signal Process. 5 (3) (2011) 554–566.
[14] T.M. Nguyen, Q.J. Wu, Gaussian-mixture-model-based spatial neighborhood relationships for pixel labeling problem, IEEE Trans. Syst. Man Cybern., Part B, Cybern. 42 (1) (2012) 193–202.
[15] R.A. Redner, H.F. Walker, Mixture densities, maximum likelihood and the EM algorithm, SIAM Rev. 26 (2) (1984) 195–239.
[16] A.R. Webb, Statistical Pattern Recognition, John Wiley & Sons, 2003.
[17] D. Benboudjema, F. Tupin, Markovian modelling and Fisher distribution for unsupervised classification of radar images, Int. J. Remote Sens. 34 (22) (2013) 8252–8266.
[18] A. Diplaros, N. Vlassis, T. Gevers, A spatially constrained generative model and an EM algorithm for image segmentation, IEEE Trans. Neural Netw. 18 (3) (2007) 798–808.
[19] A. Blake, P. Kohli, C. Rother, Markov Random Fields for Vision and Image Processing, MIT Press, 2011.
[20] J. Li, R.M. Gray, Image Segmentation and Compression Using Hidden Markov Models, vol. 571, Springer Science & Business Media, 2012.
[21] F. Wang, Y. Wu, J. Fan, X. Zhang, Q. Zhang, M. Li, Synthetic aperture radar image segmentation using fuzzy label field-based triplet Markov fields model, IET Image Process. 8 (12) (2014) 856–865.
[22] 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.
[23] W. Pieczynski, D. Benboudjema, P. Lanchantin, Statistical image segmentation using triplet Markov fields, in: International Symposium on Remote Sensing, International Society for Optics and Photonics, 2003, pp. 92–101.
[24] J.B. Courbot, E. Monfrini, V. Mazet, C. Collet, Oriented triplet Markov fields, Pattern Recognit. Lett. 103 (7–8) (2018) 457–460.
[25] M.E.Y. Boudaren, L. An, W. Pieczynski, Dempster–Shafer fusion of evidential pairwise Markov fields, Int. J. Approx. Reason. 74 (2016) 13–29.
[26] W. Pieczynski, D. Benboudjema, Multisensor triplet Markov fields and theory of evidence, Image Vis. Comput. 24 (1) (2006) 61–69.
[27] S. Foucher, M. Germain, J.-M. Boucher, G.B. Benie, Multisource classification using ICM and Dempster–Shafer theory, IEEE Trans. Instrum. Meas. 51 (2) (2002) 277–281.
[28] F. Salzenstein, W. Pieczynski, Sur le choix de méthode de segmentation statistique d'images, Trait. Signal 15 (2) (1998) 119–127.
[29] J.M. Hammersley, P. Clifford, Markov fields on finite graphs and lattices, 1971.
[30] K. Kurisu, N. Suematsu, K. Iwata, A. Hayashi, Image segmentation using a spatially correlated mixture model with Gaussian process priors, in: 2013 2nd IAPR Asian Conference on Pattern Recognition, IEEE, 2013, pp. 59–63.
[31] G. Shafer, A Mathematical Theory of Evidence, vol. 1, Princeton University Press, Princeton, 1976.
[32] P. Smets, R. Kennes, The transferable belief model, Artif. Intell. 66 (2) (1994) 191–234.
[33] M.E.Y. Boudaren, W. Pieczynski, Dempster–Shafer fusion of evidential pairwise Markov chains, IEEE Trans. Fuzzy Syst. 26 (1) (2016) 338–352.
[34] W. Pieczynski, Multisensor triplet Markov chains and theory of evidence, Int. J. Approx. Reason. 45 (1) (2007) 1–16.
[35] 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) (2013) 395–405.
[36] E. Ramasso, Inference and learning in evidential discrete latent Markov models, IEEE Trans. Fuzzy Syst. 25 (5) (2017) 1102–1114.
[37] M.E.Y. Boudaren, L. An, W. Pieczynski, Unsupervised segmentation of SAR images using Gaussian mixture-hidden evidential Markov fields, IEEE Geosci. Remote Sens. Lett. 13 (12) (2016) 1865–1869.
[38] J. Besag, On the statistical analysis of dirty pictures, J. R. Stat. Soc., Ser. B, Methodol. 48 (1986) 259–302.
[39] J. Marroquin, S. Mitter, T. Poggio, Probabilistic solution of ill-posed problems in computational vision, J. Am. Stat. Assoc. 82 (397) (1987) 76–89.

[40] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE Trans. Pattern Anal. Mach. Intell. 6 (1984) 721–741.
[41] Y. Boykov, O. Veksler, R. Zabih, Fast approximate energy minimization via graph cuts, IEEE Trans. Pattern Anal. Mach. Intell. 23 (11) (2001) 1222–1239.
[42] W. Pieczynski, Sur la convergence de l'estimation conditionnelle itérative, C. R. Math. 346 (7–8) (2008) 457–460.
[43] S.Z. Li, Markov Random Field Modeling in Image Analysis, Springer Science & Business Media, 2009.
[44] J.P. Delmas, An equivalence of the EM and ICE algorithm for exponential family, IEEE Trans. Signal Process. 45 (10) (1997) 2613–2615.
[45] T. Denœux, Maximum likelihood estimation from fuzzy data using the EM algorithm, Fuzzy Sets Syst. 183 (1) (2011) 72–91.
[46] O. Cappé, E. Moulines, T. Rydén, Inference in Hidden Markov Models, Springer, New York, 2005.
[47] Y. Ephraim, N. Merhav, Hidden Markov processes, IEEE Trans. Inf. Theory 48 (6) (2002) 1518–1569.
[48] H. Soubaras, On Evidential Markov Chains, Springer, Berlin, Heidelberg, 2010.
[49] T. Krak, J.D. Bock, A. Siebes, Imprecise continuous-time Markov chains, Int. J. Approx. Reason. 88 (2017) 178–200.
[50] M.L.P. Bueno, A. Hommersom, P.J.F. Lucas, A. Linard, Asymmetric hidden Markov models, Int. J. Approx. Reason. 88 (2017) 169–191.
[51] G.D. Cooman, J.D. Bock, S. Lopatatzidis, Imprecise stochastic processes in discrete time: global models, imprecise Markov chains, and ergodic theorems, Int. J. Approx. Reason. 76 (2016) 18–46.
[52] S. Fabien, C. Christophe, H. Mathieu, Non stationary fuzzy Markov chains, Pattern Recognit. Lett. 28 (2007) 2201–2208.
[53] C. Carincotte, S. Derrode, S. Bourennane, Unsupervised change detection on SAR images using fuzzy hidden Markov chains, IEEE Trans. Geosci. Remote Sens. 44 (2) (2006) 432–441.
[54] S. LeCam, F. Salzenstein, C. Collet, Fuzzy pairwise Markov chain to segment correlated noisy data, Signal Process. 88 (10) (2008) 2526–2541.
[55] M. Pereyra, S. McLaughlin, Fast unsupervised Bayesian image segmentation with adaptive spatial regularisation, IEEE Trans. Image Process. 26 (6) (2017) 2577–2587.
[56] P. Arbeláez, M. Maire, C. Fowlkes, J. Malik, Contour detection and hierarchical image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 33 (5) (2011) 898–916.
[57] Bresson, Esedoglu, Vandergheynst, J.P. Thiran, Osher, Fast global minimization of the active contour/snake model, J. Math. Imaging Vis. 28 (2) (2007) 151–167.
[58] X. Cai, R. Chan, T. Zeng, A two-stage image segmentation method using a convex variant of the Mumford–Shah model and thresholding, SIAM J. Imaging Sci. 6 (1) (2013) 368–390.
[59] M. Pereyra, H. Batatia, S. Mclaughlin, Exploiting information geometry to improve the convergence of nonparametric active contours, in: IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2013, pp. 165–168.
[60] S. Bagon, Matlab wrapper for graph cut, http://www.wisdom.weizmann.ac.il/~bagon, December 2006.
[61] G. Poggi, G. Scarpa, J.B. Zerubia, Supervised segmentation of remote sensing images based on a tree-structured MRF model, IEEE Trans. Geosci. Remote Sens. 43 (8) (2005) 1901–1911.
[62] P. Lanchantin, W. Pieczynski, Unsupervised restoration of hidden nonstationary Markov chains using evidential priors, IEEE Trans. Signal Process. 53 (8) (2005) 3091–3098.