

Fuzzy Markov Random Fields versus Chains for Multispectral Image Segmentation

F. Salzenstein ⁺ Ch. Collet [‡]

⁺Laboratoire InESS, UMR CNRS 7163 Université Strasbourg 1 (ULP)

Fabien.Salzenstein@phase.c-strasbourg.fr

[‡] LSIIT UMR CNRS 7005, Université Strasbourg 1 (ULP)

collet@lsiit.u-strasbg.fr

Person to whom correspondence should be sent:

Name: Pr Christophe Collet
Address: ULP-ENSPS-LSIIT Boulevard Sebastien Brant
Pole d'Innovation API, BP 10413 ,
F-67412 ILLKIRCH France.
E-mail: Christophe.Collet@ensps.u-strasbg.fr
Phone: (33)(0) 390 244 490
Fax: (33)(0) 390 244 342

Other authors :

Name: Fabien Salzenstein
Adress: InESS, Institut d'lectronique du Solide et des Systmes
23, rue du Loess - BP 20 CR - F-67037 STRASBOURG Cedex 2 France
E-mail: salzenst@iness.c-strasbourg.fr
Phone: (33)(0) 388 106 558
Fax: (33)(0) 390 244 342

Abstract

This paper deals with a comparison of recent statistical models based on fuzzy Markov random fields and chains for multispectral image segmentation. Fuzzy scheme takes into account discrete and *continuous* classes which models the imprecision of the hidden data. In this framework, we assume the dependence between bands and we express the general model for the covariance matrix. A fuzzy Markov chain model is developed in an unsupervised way. This method is compared with the fuzzy Markovian field model previously proposed by one of the authors. Segmentation task is processed with Bayesian tools, such as the well known MPM (Mode of Posterior Marginals) criterion. Our goal is to compare the robustness and rapidity for both methods (fuzzy Markov fields versus fuzzy Markov chains). Indeed, such fuzzy-based procedures seem to be a good answer, *e.g.*, for astronomical observations when the patterns present diffuse structures. Moreover, these approaches allow us to process missing data in one or several spectral bands which correspond to specific situations in Astronomy [1]. To validate both models, we perform and compare the segmentation on synthetic images and raw multispectral astronomical data.

Index Terms

fuzzy Markov field, fuzzy Markov chain, parameterized joint density, multispectral image segmentation, missing data.

Acknowledgments: The authors thank the French Ministry of Research, ACI IDHA (Images Distribuées Hétérogènes en Astronomie) and ACI MDA (Masses de Données en Astronomie)) for partial financial support of this work.

I. INTRODUCTION

For many years now, the well known markovian models [2], [3] have been powerful statistical tools for classification, denoising or pattern recognition. The segmentation process on multicomponent images requires a non trivial modelization step, in order to allocate a label for different sets of pixels exhibiting similar behaviors. On the one hand, the bands observed are statistically dependent and require an adapted multivariate model. Otherwise, as the observations are intrinsically imprecise, a 'hard' classification which labels the original scene into finite discrete sets artificially eliminates the imprecision. Yet, many observations of the cosmos, such as distant galaxies or dust/gas clouds, present such intrinsic fuzzy properties. The fuzzy segmentation [4]–[6], based on the fuzzy set theory [7], is a possible answer to that requirement: it allows each pixel to belong to several classes simultaneously. Other authors combined the imprecision of the hidden data and the uncertainty of the observed data using probabilistic methods. In this way, a contextual approach has been proposed [8] using a measure of density including components related to the discrete information (Dirac function) and components related to the imprecise information (Lebesgue measure). Some markovian random field models have been introduced successfully in the unsupervised fuzzy classification [9]–[11]. The Markov field model introduced in [9] does not allow any coexistence between the hard and fuzzy classes, where as our fuzzy model does. In [12], authors studied a link between the Dempster-Shafer theory of evidence and the fuzzy classification. The multi-spectral classification scheme based on Markov model has been studied in the discrete case [13]–[15]. Some fuzzy Markov field models were recently extended to multispectral context [16]: under inter-spectral correlated data assumption, we express the covariance matrix related to fuzzy classes. In another way, the fuzzy random Gibbs has been combined with the generalized fuzzy sets [17] in order to provide a multi-class segmentation. Although these results are encouraging, the processing time limits its efficiency: thus it is worth adapting the markovian chain model to a fuzzy multispectral context. The main idea for using the hidden Markov chain model lies in its rapidity and flexibility. Moreover, it offers a specific interest in astronomy when some multispectral data is missing [1]. The fuzzy Markov chains based on the fuzzy set theory have been introduced in [18], [19]. In particular, the hidden fuzzy Markov chain model presented in [18] replaces the probability with a fuzzy measure and uses the fuzzy integral of Choquet for integration over each state. Another

fuzzy Markov chain model, based on a non-parametric density according to a positive measure and using the Lebesgue integration, has been introduced in [20]. It has been also extended to a multivariate context under between-band independence assumption [21], eventually using a P/ICA (Principal or Independent Component Analysis) pre-processing method. Nevertheless, the main drawback of such an approach is the loss of between-spectra information. Moreover, the mixture of weighted spectral bands obtained has no physical significance and remains difficult to analyze. A comparative study between the Markov fields and chains has been performed for radar images in hard classification [22]. In this paper, we propose a comparison between the hidden Markov chains and the hidden Markov fields in the fuzzy context in order to segment astronomical observations using the multispectral driven data model under between-band dependence assumption. The paper is organized as follows. Section II presents the studied models on the hidden variable/image: a fuzzy Markov field model (FMF) and a new stationary fuzzy chain model (FMC) for which the prior joint density is parameterized (this model is denoted by P-FMC). Section III deals with multivariate distribution of the observed data in the case of correlated sensors, whereas section IV describes the segmentation step associated with an MPM strategy (Maximum Posterior Mode [23]). In Section V, we present the estimation of the hyper-parameters (data driven and prior densities) A general overview of the algorithms is given section VI. Finally, in section VII, we compare fuzzy Markov field and fuzzy Markov chain which we developed in the context of multispectral astronomical images. We validate the procedures on synthetic fuzzy images.

NOTATIONS

Variable	Meaning
X_s	Random variable labeling a pixel
N	number of pixels
$X = (X_s)_{1 \leq s \leq N}$	Random variable representing the true field
D	Number of sensors
$\mathbf{Y}_s = (Y_s^{(j)})_{1 \leq j \leq D}$	Random variable representing an observed pixel
$\mathbf{Y} = (\mathbf{Y}_s)_{1 \leq s \leq N}$	Random variable of the observations
Φ_X	Parameters of the a priori distribution
Φ_Y	Parameters of the driven data distribution
Notation	Meaning
FMF	Fuzzy Markov field
FMC	Fuzzy Markov chain
NP-FMC	Fuzzy Markov chain with a non parameterized joint density
P-FMC	Fuzzy Markov chain with a parameterized joint density

II. THE FUZZY RANDOM MARKOV MODELS

A. The fuzzy random Markov field (FMF)

Let $x = (x_s)_{1 \leq s \leq N} = (x_1, \dots, x_N)$ a realization of a random variable $X = (X_s)_{1 \leq s \leq N} = (X_1, \dots, X_N)$, where $s \in 1, \dots, N$ indexes each pixel x_s . The fuzzy segmentation problem consists in estimating the hidden realization $x = (x_s)_{1 \leq s \leq N}$, for a given set of D observations $\mathbf{Y} = \mathbf{y} = \{\mathbf{y}_s \in \mathbb{R}^D\}$, where $x_s = (\varepsilon_1(s), \varepsilon_2(s), \dots, \varepsilon_K(s))$. Each $\varepsilon_i(s)$ represents the contribution of the class ω_i which belongs to a finite discrete set $\Omega = \{\omega_1, \dots, \omega_K\}$ of K hard classes ω_i . The fuzzy belonging of each pixel respect the following normalization condition: $\varepsilon_1(s) + \varepsilon_2(s) + \dots + \varepsilon_K(s) = 1$. In the hard case, all $\varepsilon_i(s)$ are null except one component which equals 1. In the context of astronomical imaging, it worth to consider a fuzzy partition using two discrete classes. Actually, we may have to discriminate a dark cosmic background from emitting structures (galaxy, or a gas/dust cloud, stars ...). A set $\Omega = \{0, 1\}$ yields $x_s \in [0, 1]$. Then, all values $x_s \in [0, 1]$ model the proportion of the class "0" in the pixel related to X_s , whereas $1 - x_s$ corresponds to the proportion of the class "1". The distribution at each random variable X_s is given by a density h_s with respect to a measure ν including discrete components (Dirac functions δ_0, δ_1 on $\{0, 1\}$)

and a continuous component (the Lebesgue measure μ on $]0, 1[$) [8]:

$$\nu = \delta_0 + \delta_1 + \mu$$

thus, h_s verifies the following normalization condition:

$$h_s(0) + h_s(1) + \int_0^1 h_s(\varepsilon) d\varepsilon = 1$$

The discrete components of ν are associated with the hard classes, whereas the continuous component μ is associated with the fuzzy feature. Let us suppose a Markovian random field $X \in [0, 1]^N$. Its distribution h_X with respect to a measure ν^N is given by:

$$h_X(x) = \frac{1}{Z} \cdot e^{-U_f(x)},$$

where the fuzzy energy U_f is a sum of clique functions Φ_C :

$$U_f(x) = \sum_{(x_s, x_t) \in C} \Phi_C(x_s, x_t)$$

where $(x_s, x_t) \in [0, 1]^2$ represents a pairwise of pixels associated with a second order neighborhood (clique C), and the associated functions Φ_C defined on $[0, 1]^2$. We based our study on the model proposed in [10], where the values of Φ_C depend on the nature - i.e. fuzzy or hard - of the neighbored pixels:

1) if $(x_s, x_t) \in \{0, 1\}^2$, i.e., both pixels are hard:

$$\Phi_C(x_s, x_t) = \begin{cases} -\alpha^d & \text{if } x_s = x_t, \\ +\alpha^d & \text{if } x_s \neq x_t. \end{cases} \quad (1)$$

2) if $(x_s, x_t) \in [0, 1]^2 - \{0, 1\}^2$, i.e., at least one fuzzy pixel:

$$\Phi_C(x_s, x_t) = -\alpha^f \cdot (1 - 2|x_s - x_t|).$$

Thus, if we consider four types of neighbors (x_s, x_t) corresponding to the "horizontal", "first diagonal" ("north-east" or "south-west"), "second diagonal" ("north-west" or "south-east") and vertical direction, we need four pairwises $(\alpha^d, \alpha^f) \in (\mathbb{R}^+)^2$ to describe the field X . Thus the prior density is defined by a global parameter $\alpha = \{\alpha_1^d, \alpha_2^d, \alpha_3^d, \alpha_4^d, \alpha_1^f, \alpha_2^f, \alpha_3^f, \alpha_4^f\}$.

B. The fuzzy random Markov chain (FMC)

Scanning a given image is always possible in order to extract the data as a monodimensional vector. In our case, a Hilbert-Peano path which preserves the neighborhood as well as possible [24] is used for images of size $2^n \times 2^m$ pixels with $(n, m) \in \mathbb{Z}^2$. This technique has been successfully applied for the hard Markov chains in the context of the mono-band [25] as well as the multi-band observations [14].

1) *Principe*: Let us consider now a Markov chain $X = (x_s)_{1 \leq s \leq N}$ with continuous statements i.e. $X_s \in [0, 1]$. In that context, at each site s , we have the following condition on the transition density:

$$p(x_s | x_{s-1}, x_{s-2}, \dots, x_1) = p(x_s | x_{s-1})$$

It implies that the fuzzy class distribution at the location s is independent on the previous state conditional on location $s - 1$. To define the distribution $\pi(x)$ of the variable X , we need the density $p(x_1)$ of the initial distribution, and the transition densities $p(x_s | x_{s-1})_{1 \leq s \leq N}$:

$$p(x_1, x_2, \dots, x_N) = p(x_1) \cdot p(x_2 | x_1) \dots p(x_N | x_{N-1})$$

When the chain is stationary, all prior distributions can be deduced from a joint density.

2) *Prior joint density*: The prior joint density $p(x_s, x_{s+1})$ is defined on the pairwise $(x_s, x_{s+1}) \in [0, 1]^2$. Giving such density according to a measure $\nu \otimes \nu$, the normalization condition is:

$$\int_0^1 \int_0^1 p(u, v) d(\nu \otimes \nu)(u, v) = 1 \quad (2)$$

We propose a general model to define it:

$$p(\varepsilon_1, \varepsilon_2) = a \cdot \phi(\varepsilon_1, \varepsilon_2) + b, \text{ with: } \phi(\varepsilon_1, \varepsilon_2) = \phi(\varepsilon_2, \varepsilon_1); \quad (a, b) \in \mathbb{R}^2;$$

The function $\phi(\varepsilon_1, \varepsilon_2)$, is applied when at least one label is fuzzy, i.e. ε_1 or $\varepsilon_2 \in]0, 1[$. If both labels are hard, we note $p(0, 0) = \pi_{00}, p(1, 1) = \pi_{11}, p(0, 1) = \pi_{01}, p(1, 0) = \pi_{10}$. We model a parameterized function as follows:

$$\phi(\varepsilon_1, \varepsilon_2) = (1 - |\varepsilon_1 - \varepsilon_2|)^r \quad r \in \mathbb{R} \quad (3)$$

This expression is analogous to the clique function Φ_C of the fuzzy Markovian field. Thus, the more the values of ε_1 and ε_2 are distant, the lower the corresponding joint probability will be. Moreover, when r increases, the probability of having two similar neighbored pixels increases:

the parameter r has an effect on the image homogeneity. In this paper, we will study different values $1 < r < 50$. The limit conditions yield:

$$p(0, 1) = p(1, 0) = b = \pi_{01} = \pi_{10} \quad (4)$$

Applying equation (2) yields to the general condition:

$$\begin{aligned} & p(0, 0) + p(1, 1) + p(0, 1) + p(1, 0) + 2 \cdot \int_{]0,1[} p(0, u) du + \\ & 2 \cdot \int_{]0,1[} p(1, u) du + \int_{]0,1[} \int_{]0,1[} p(u, v) dudv = 1 \end{aligned} \quad (5)$$

This gives a relationship between all prior parameters $(\pi_{00}, \pi_{11}, \pi_{01}, \pi_{10}, a, b)$. When the prior joint density is defined by the function (3), we compute (5) using a quantization of the interval $[0, 1]$ into M equidistant values: $\{\varepsilon_0 = 0, \varepsilon_1 = \frac{1}{M}, \dots, \varepsilon_i = \frac{i}{M}, \dots, \varepsilon_M = 1\}$. Then, we derive the initial density $p(x_1)$, which corresponds to the marginal distribution (6):

$$p(x_s) = \int_0^1 p(x_s, \varepsilon) d\nu(\varepsilon) = p(x_s, 0) + p(x_s, 1) + \int_{]0,1[} p(x_s, \varepsilon) d\varepsilon \quad (6)$$

In this section we introduced a new fuzzy Markov random chain model. We defined the associated prior joint density, initial and transition probabilities. Let us now describe the law of the observations conditional on X .

III. MULTISPECTRAL NOISE MODEL

The aim of our paper is to process multispectral data. We observe D realizations $(y^{(1)}, y^{(2)}, \dots, y^{(D)})$ of the random vector $\mathbf{Y} = (Y^{(1)}, Y^{(2)}, \dots, Y^{(D)})$. They represent a single scene observed at different wavelengths or from different sensors. Given the set of observations, we wish to estimate one realization $X = x \in [0, 1]^N$. We assume the spatial independence: the variables $\mathbf{Y}_s|_{s \in S}$ are independent conditionally to X , and $P[\mathbf{Y}_s|X] = P[\mathbf{Y}_s|X_s]$. In a spectral correlated context, the distribution of $\mathbf{y}_s = [y_s^{(1)}, y_s^{(2)}, \dots, y_s^{(D)}]^t$ according to $X_s = \varepsilon_s \in [0, 1]$ is Gaussian:

$$f_{x_s}(\mathbf{y}_s) = \frac{1}{2\pi^{D/2}(\det \Gamma_{\varepsilon_s})^{1/2}} \exp\left(\frac{-1}{2}(\mathbf{y}_s - \boldsymbol{\mu}_{\varepsilon_s})^t \boldsymbol{\Gamma}_{\varepsilon_s}^{-1}(\mathbf{y}_s - \boldsymbol{\mu}_{\varepsilon_s})\right) \quad (7)$$

where $\boldsymbol{\mu}_{\varepsilon_s} = [\mu_{\varepsilon_s}^{(1)}, \dots, \mu_{\varepsilon_s}^{(D)}]^t$ and $\boldsymbol{\Gamma}_{\varepsilon_s} \in \mathbb{R}^D \times \mathbb{R}^D$ respectively define a mean vector and variance-covariance matrix, at each fuzzy/hard site. Let be $(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1)$ and $(\boldsymbol{\Gamma}_0, \boldsymbol{\Gamma}_1)$ the mean vectors and

variance-covariance matrix related to the hard classes "0" and "1". For each fuzzy site $X_s = \varepsilon_s$, the related mean vector and matrix $\boldsymbol{\mu}_{\varepsilon_s}$ et $\boldsymbol{\Gamma}_{\varepsilon_s}$ are written:

$$\begin{aligned}\boldsymbol{\mu}_{\varepsilon_s} &= (1 - \varepsilon_s) \cdot \boldsymbol{\mu}_0 + \varepsilon_s \cdot \boldsymbol{\mu}_1 \\ \boldsymbol{\Gamma}_{\varepsilon_s} &= (1 - \varepsilon_s)^2 \cdot \boldsymbol{\Gamma}_0 + \varepsilon_s^2 \cdot \boldsymbol{\Gamma}_1\end{aligned}\quad (8)$$

It is known that the density $f_{X,\mathbf{Y}}(x, \mathbf{y})$ of the distribution of (X, \mathbf{Y}) is Markovian, with respect to a measure $\nu^N \otimes (\mu^N) \otimes (\mu^N) \otimes \dots \otimes (\mu^N)$. In particular, the *a posteriori* field X conditional on \mathbf{Y} is Markovian. Thus, one can process the posterior realizations of the hidden variable X , in the FMF or the FMC context. These properties allow us to perform the segmentation in an unsupervised way with the hyperparameter estimation presented in section V.

IV. SEGMENTATION PROCEDURE

A. MPM criterion adapted to the fuzzy context

Once we estimated the hyperparameters, the second step focuses on a decision/segmentation process: given the set of observations \mathbf{Y} , we wish to estimate one realization $X = x \in [0, 1]^N$. The final decision process is performed with a Bayesian approach as following: let us consider a loss function $L^* : [0, 1]^N \times [0, 1]^N \rightarrow \mathbb{R}^+$. The function $L^*(\hat{x}, x)$ models the severity of attributing the value \hat{x} instead of a true one x . Given an observation $\mathbf{Y} = \mathbf{y}$, the bayesian decision \hat{d} minimizes the expectation $E[L^*(X, \hat{X})]$ with $\hat{d}(\mathbf{Y}) = \hat{X}$. Considering a such loss function on all field X penalizes identically an error on one classified pixel and an error on every classified pixels. However, it is possible to adapt the MPM criterion [23] to the fuzzy context. For a "fuzzy MPM" approach the global loss function L^* is depending on a local loss function $L_s^*(x_s, \hat{x}_s)$ defined at each site s : $L^*(x, \hat{x}) = \sum_{s=1}^{s=N} L_s(x_s, \hat{x}_s)$. A decision \hat{d}_s will then involve minimizing $E[L_s^*(X_s, \hat{X}_s)]$ at each location s , where $\hat{d}_s(\mathbf{Y}) = \hat{X}_s$. In other words, the decision is taking locally at each site s , according to all given information \mathbf{Y} . Although there are numerous possibilities in the choice of the loss function, the 'absolute distance' $L_s(x_s, \hat{x}_s) = |x_s - \hat{x}_s|$, gives efficient results for segmentation tasks. Practically, for one realization $\mathbf{Y} = \mathbf{y}$ we minimize a conditional expectation (9), in order to obtain an optimal value of X_s :

$$\hat{x}_s^{opt} = \arg \min_{\hat{X}_s = \hat{x}_s} E[L_s(X_s, \hat{X}_s) | \mathbf{Y} = \mathbf{y}] \quad (9)$$

For a stationary variable X , the error rate is approximated by:

$$E[L_s(X_s, \widehat{X}_s)] \simeq \frac{1}{N} \sum_{s=1}^{s=N} L_s(x_s, \widehat{x}_s) \quad (10)$$

B. Practical implementation

The calculus of (9) requires the knowledge of the posterior distribution $p_{X_s}^{\mathbf{Y}}$ at each X_s :

$$\begin{aligned} E[L_s(X_s, \widehat{X}_s)|\mathbf{Y} = \mathbf{y}] &= p_{X_s}^{\mathbf{Y}}(0) \cdot L_s(0, \widehat{X}_s) + p_{X_s}^{\mathbf{Y}}(1) L_s(1, \widehat{X}_s) \\ &+ \int_{t \in]0,1[} p_{X_s}^{\mathbf{Y}}(t) L_s(t, \widehat{X}_s) dt \end{aligned} \quad (11)$$

Segmentation is performed by affecting to each pixel a value $\widehat{X}_s \in [0, 1]$ which minimizes (11). For the Markov fields, we estimate numerically $p_{X_s}^{\mathbf{Y}}$ by empirical frequencies computed from several realizations of X according to its posterior distribution. This procedure is time consuming, whereas the Markov chain model provides an efficient tool: one can compute directly the posterior density using the forward/backward procedure [27] in the fuzzy context. The forward and backward densities $\alpha_s(x_s), \beta_s(x_s)$ are defined by:

$$\begin{aligned} \alpha_s(x_s) &= p(x_s, \mathbf{Y}_1, \dots, \mathbf{Y}_s) \\ \beta_s(x_s) &= \frac{p(\mathbf{Y}_{s+1}, \dots, \mathbf{Y}_N | x_s)}{p(\mathbf{Y}_{s+1}, \dots, \mathbf{Y}_N | \mathbf{Y}_1, \dots, \mathbf{Y}_s)} \end{aligned}$$

The recurrence formula providing these quantities, are analogous to the hard segmentation processing:

$$\begin{aligned} \alpha_s(x_s) &= p(\mathbf{Y}_s | x_s) \int_0^1 \alpha_{s-1}(u) \cdot p(x_s | u) d\nu(u) \\ \beta_s(x_s) &= \int_0^1 \beta_{s+1}(u) p(\mathbf{Y}_{s+1} | u) p(u | x_s) d\nu(u) \end{aligned}$$

Practically, we perform the following recursive calculus, using the discretization of the interval $[0, 1]$.

At each iteration, these quantities are normalized by the factors $\int_0^1 \alpha_s(u) d\nu(u)$ for the backward density, and $\int_0^1 \beta_s(u) d\nu(u)$ for the forward density. The product $\alpha_s(\varepsilon) \cdot \beta_s(\varepsilon)$ is equal to the posterior distribution $p_{X_s}^{\mathbf{Y}}$ which gives immediately the minimization of (11). Moreover, as in the hard context, the posterior realizations of the variable X are simulating using the posterior

transition $p^{\mathbf{Y}^{s+1}}(x_{s+1}|x_s)$ and initial posterior densities $p_{X_s}^{\mathbf{Y}}(x_s)$. For any $(x_s, x_{s+1}) \in [0, 1]^2$:

$$\begin{aligned} p_{X_s}^{\mathbf{Y}}(x_s) &= \alpha_s(x_s) \cdot \beta_s(x_s) \\ p^{\mathbf{Y}^{s+1}}(x_{s+1}|x_s) &= \frac{p(x_{s+1}|x_s) \cdot p(\mathbf{Y}_{s+1}|x_{s+1}) \cdot \beta_{s+1}(x_{s+1})}{\int_0^1 p(x|x_s) p(\mathbf{Y}_{s+1}|x) \cdot \beta_{s+1}(x) d\nu(x)} \end{aligned}$$

C. Segmentation of missing data

Both Markovian approaches allow us to process missing data in one or several spectral bands which corresponds to specific situations in Astronomy. Such segmentations being global methods, it is theoretically possible to extract the information in the presence of missing data [1]. Thus giving a set of observed data \mathbf{Y}_{S_1} where S_1 belongs to the regular lattice S , we are able to extract any information X_{S_2} , where $S_2 \in S$, the set S_1 do not necessarily overlap S_2 . To perform a decision process, we assign to the driven missing-data likelihood to one, i.e., $f_{x_s}(\mathbf{y}_s) = 1$. On this location where the data is missing, only prior distribution takes place in the labelling process. This procedure is named rule 1. Moreover, in the context of the Markov chain model, it is also possible to process only the non missing zone of the image: one has to suppress the missed data in the Hilbert-Peano path. This procedure will be mentioned as rule 2 in the following.

V. HYPER-PARAMETER ESTIMATION

The final segmentation step requires the parameter set $\theta = (\theta_X, \theta_Y)$ where the *prior* parameters θ_X define the fuzzy markovian energy or the prior density of the Markov chain. The parameters $\theta_Y = ((\mu_0, \mu_1); (\mathbf{\Gamma}_0, \mathbf{\Gamma}_1))$ define the distribution of the driven data conditional on X , i.e., the multidimensional Gaussian densities. For the FMF, we use the iterative conditional estimation (ICE) [26] to compute $(\hat{\theta}_X, \hat{\theta}_Y)$ thanks to the conditional expectation $E[\hat{\theta}(X, \mathbf{Y})|\mathbf{Y}]$, where $\hat{\theta}(X, \mathbf{Y}) = (\hat{\theta}_X(X), \hat{\theta}_Y(X, \mathbf{Y}))$ is an estimator of θ for complete data. Although we do not observe any realization of the hidden variable X , it is possible to simulate N_{ICE} realizations of X according to its distribution conditioned on $\mathbf{Y} = \mathbf{y}$, in order to approximate the conditional expectation $E_{\theta^{[p]}}[\hat{\theta}(X, \mathbf{Y})|\mathbf{Y} = \mathbf{y}]$:

- 1) Initialisation : $\theta^{[0]} = \theta_{\text{init}}$
- 2) step p: $\theta^{[p+1]} = (\hat{\theta}_X^{[p+1]}, \hat{\theta}_Y^{[p+1]})$:

$$\hat{\theta}_X^{[p+1]} = \frac{1}{N_{ICE}} (\hat{\theta}_X(x_1^{[p]}) + \dots + \hat{\theta}_X(x_{N_{ICE}}^{[p]})) \quad (12)$$

$$\hat{\theta}_Y^{[p+1]} = \frac{1}{N_{ICE}} (\hat{\theta}_Y(x_1^{[p]}, \mathbf{y}) + \dots + \hat{\theta}_Y(x_{N_{ICE}}^{[p]}, \mathbf{y})) \quad (13)$$

where $x_i^{[p]}$ represents a hidden realization according to the current parameter $\theta^{[p]} = (\hat{\theta}_X^{[p]}, \hat{\theta}_Y^{[p]})$. For each posterior realization of the FMC model, an SEM estimator (empirical frequencies and moments) is used to estimate the hyper-parameters. All values are then averaged by N_{ICE} posterior realizations using (12) and (13). When the sequence $\theta^{[p]}$ approaches steady state - for example 1% of the relative change in the values - we stop the procedure. Let us consider now the problem in estimating θ_X and θ_Y separately.

A. Data driven parameter estimation

Let us suppose now, we observe a realization (x, y) of the pairwise (X, Y) . In a hard classification the empirical moment estimator $\hat{\theta}_Y(X, Y)$ of θ_Y corresponds to the maximum likelihood under conditional Gaussian laws assumption. When we use a fuzzy classification, it is enough to estimate the parameters dealing with hard classes. We generalize the method proposed in [10] applying the empirical moments to the hard pixels. Let be $Q_p = \{s \in S / X_s = p\}$, $p = 0, 1$ the sets of pixels which belong to the hard classes. Our aim is to estimate the set of parameters $\theta_Y = (\mu_0, \Gamma_0; \mu_1, \Gamma_1) = (\mu_0^{(i)}, \Gamma_0^{(i,j)}; \mu_1^{(i)}, \Gamma_1^{(i,j)})$, where $1 < i, j < D$. Applying the empirical moment method on the hard pixels yields:

$$\hat{\mu}_p^{(i)} = \frac{\sum_{s \in Q_p} y_s^{(i)} \cdot \delta(x_s, p)}{\sum_{s \in Q_p} \delta(x_s, p)} \quad (14)$$

$$\hat{\Gamma}_p^{(i,j)} = \frac{\sum_{s \in Q_p} (y_s^{(i)} - \hat{\mu}_p^{(i)}) \cdot (y_s^{(j)} - \hat{\mu}_p^{(j)}) \delta(x_s, p)}{\sum_{s \in Q_p} \delta(x_s, p)} \quad (15)$$

B. Prior parameter estimation

1) *Markovian field model*: The prior parameter is given by the couples $\theta_X = (\alpha^f, \alpha^d)$ associated with the clique functions, according to the different directions. Let us suppose we observe a realization x of the fuzzy field X . The estimator of θ_X is based on a stochastic gradient algorithm [28]:

- 1) Initialization : $\theta_X^{[0]}$
- 2) at the current iteration q :

$$\theta_X^{[q+1]} = \theta_X^{[q]} + \frac{c}{q+1} [U_f'(x^{[q+1]}) - U_f'(x)] \quad (16)$$

The quantities U_f' are the gradients of U_f according to the parameter θ_X . $x^{[q+1]}$ is a realization of X simulated by the fuzzy Gibbs sampler for the current parameter $\theta_X^{[q]}$ ([10], [11]). A descent

step algorithm with $c = 1/N$ gives good results, if we compensate the lost of adaptability (*i.e.*, a decreasing step), by a good initialization. Practically, a value $\theta_X^{[0]}$ corresponding to an homogeneous field gives satisfying results.

2) *Markovian chain model*: Let us suppose we observe a realization x of X . The prior parameter θ_X corresponds to the initial and transition densities. For a stationary chain, they can be deduced from the joint density, as seen in section II-B.2. We consider two hypothesis: i) the joint density depends on a parameterized function ϕ , ii) it is a non parameterized density.

- i) in the context of a parameterized density, we estimate the set of parameters $(\pi_{00}, \pi_{11}, \pi_{01}, \pi_{10}, a, b)$, which is reduced to $(\pi_{00}, \pi_{11}, a, b)$ thanks to condition (4). One parameter, for example a , can be deduced from the others, in order to respect the condition of normalization (5). Finally, we estimate (π_{00}, π_{11}, b) , in the following manner:

$$\pi_{00} = \frac{\sum_{s=1}^{s=N-1} \delta(x_s, 0)\delta(x_{s+1}, 0)}{N} \quad (17)$$

$$\pi_{11} = \frac{\sum_{s=1}^{s=N-1} \delta(x_s, 1)\delta(x_{s+1}, 1)}{N} \quad (18)$$

$$b = \frac{\sum_{s=1}^{s=N-1} \delta(x_s, i)\delta(x_{s+1}, j)}{N}, \quad i, j \in \{0, 1\}, \quad i \neq j \quad (19)$$

- ii) The parameterized density minimizes the amount of prior parameters and allows an homogeneous fuzzy class. However this hypothesis seems too strict, and it is worth to estimate a non parametric density. We compute the empirical $M \times M$ joint probabilities (20) according to different neighborhood configurations:

$$P[X_s \in I_i, X_{s+1} \in I_j] = \int_{I_i} \int_{I_j} p(u, v) d(\nu \otimes \nu)(u, v) \quad (20)$$

We deduce the joint density for different configurations of neighbored pixels. Practically, we process the data for $M = 10$ fuzzy levels. This implies an acceptable compromise between the time processing and the complexity of the data.

VI. GENERAL OVERVIEW OF THE METHODS

We observe a single realization \mathbf{y} of \mathbf{Y} ,

- 1) **Estimation step**: we apply the following estimator $\hat{\theta}(X, \mathbf{Y})$ of $\theta = (\theta_X, \theta_Y)$:

a) *initialization*: $\theta^{[0]} = (\theta_X^{[0]}, \theta_Y^{[0]})$

The parameter $\theta_Y^{[0]}$ is estimated using empirical mean and standard deviation.

i) *FMF model*: $\theta_X^{[0]} = \{\alpha_1^{d,[0]}, \alpha_2^{d,[0]}, \alpha_3^{d,[0]}, \alpha_4^{d,[0]}, \alpha_1^{f,[0]}, \alpha_2^{f,[0]}, \alpha_3^{f,[0]}, \alpha_4^{f,[0]}\}$.

ii) *FMC model*: $\theta_X^{[0]} = \{\pi_{00}^{[0]}, \pi_{11}^{[0]}, \pi_{10}^{[0]}, \pi_{01}^{[0]}\}$

b) *current step*: we compute $\theta^{[p+1]} = (\theta_X^{[p+1]}, \theta_Y^{[p+1]})$ from \mathbf{y} and $\theta^{[p]} = (\theta_X^{[p]}, \theta_Y^{[p]})$ in the following way:

i) simulate N_{ICE} realizations $x^{[p]}$ of X according to the posterior distribution to $\mathbf{Y} = \mathbf{y}$ and the parameter $\theta^{[p]}$.

ii) For each posterior field $x^{[p]}$, estimate the prior parameter $\widehat{\theta}_X(x^{[p]})$:

A) *FMF model*: compute $\widehat{\theta}_X(x^{[p]})$ using (16) and $\theta_X^{[p+1]}$ using (12).

B) *NP-FMC model*: compute $\widehat{\theta}_X(x^{[p]})$ using (17)-(19) and $\theta_X^{[p+1]}$ using (12).

C) *P-FMC model*: compute $\widehat{\theta}_X(x^{[p]})$ using (20) and $\theta_X^{[p+1]}$ using (12).

iii) For each pairwise $(x^{[p]}, \mathbf{y})$, estimate the noise parameter $\widehat{\theta}_Y(x^{[p]}, \mathbf{y})$ using (14)-(15) and compute $\theta_Y^{[p+1]}$ using (13).

c) if the sequence $\theta^{[p]} = (\theta_X^{[p]}, \theta_Y^{[p]})$ approaches steady state, stop.

2) **Segmentation step**: perform the procedure described in the section IV in the following way:

a) *FMF model*: estimate empirically $p_{X_s}^{\mathbf{Y}}$ and compute for each s the value of \widehat{X}_s which minimizes (11).

b) *FMC model*: estimate $p_{X_s}^{\mathbf{Y}}$ using the forward-backward procedure and compute for each s the value of \widehat{X}_s which minimizes (11).

VII. EXPERIMENTS AND RESULTS

A fuzzy Markov field model (FMF) and a fuzzy Markov chain model (FMC) have been presented in the previous sections. We argue that such models are well fitted for segmentation task of diffuse structures in astronomy. This section deals with the comparison of both approaches on synthetic and raw multispectral data sets.

A. Synthetic Images

We applied our methods to fuzzy synthetic images, in order to evaluate their efficiency according to the hyper-parameter estimation, segmentation error, and their ability to preserve the fuzzy information. Moreover, we compare the FMC classification under parameterized and non parameterized density assumption.

1) *Simulation context:* To compare the FMF model with the FMC models, one has to apply both algorithms to noisy Markov fields. A realization of a FMC (resp. a FMF) is shown on Fig.1(a) (resp. Fig.2(a)). The original images contain $M = 10$ fuzzy levels (respectively 24.72 % and 53.96 % of fuzzy rate). We simulated the FMC realization using the parameterized function density described in section II-B.2 with a parameter $r = 30$. The parameter r offers numerous possibilities for simulating such Markov chains with different homogeneous images. The fuzzy rate is computed by dividing the number of fuzzy labels by the number of pixels N . The corresponding model parameters are given in Tab. II and III. For the FMF model, the couples $\{\alpha_i^d, \alpha_i^f\}$ stand for hard and fuzzy potentials of the second order cliques. For the image 1(a) $\pi_{01} = \pi_{10} = b = 0$ and $\pi_{00} = \pi_{11} = 0.4$. Actually, the assumption of no neighborhood configurations containing two different hard classes, is realistic. In the same manner, the FMF 2(a) does not contain such configurations (see Tab. IV), which is compatible with the energy model given in section II-A. With the FMC method, the prior parameters of the Fig.2(a) are computed through an empirical estimation (Tab. IV - first line). Finally, for each field, we simulated three noisy versions by adding gaussian components defined on Tab. I.

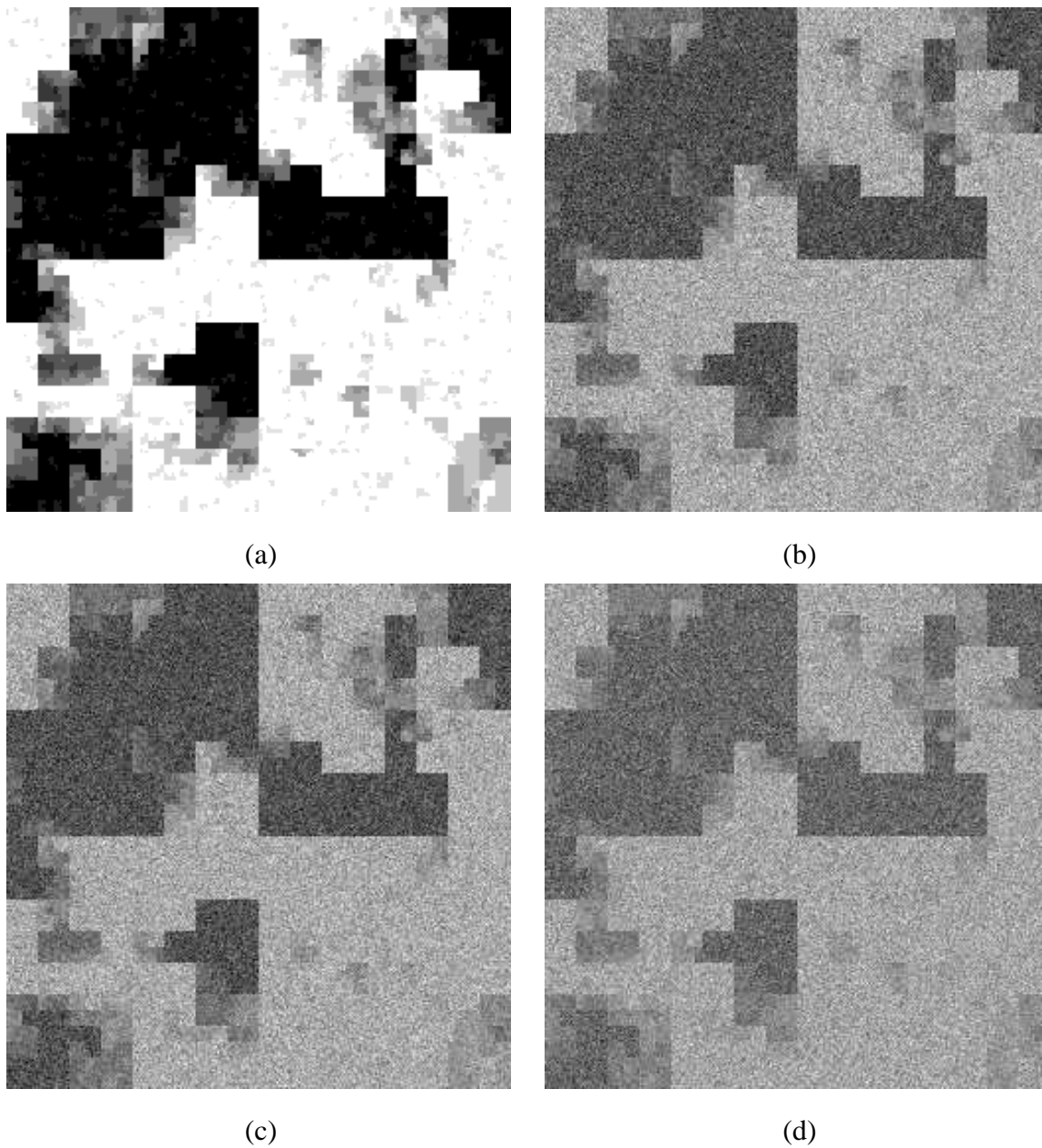


Fig. 1. FMC (a) with prior parameters given in Tab. II, and its noisy versions in 3 bands (b,c,d).

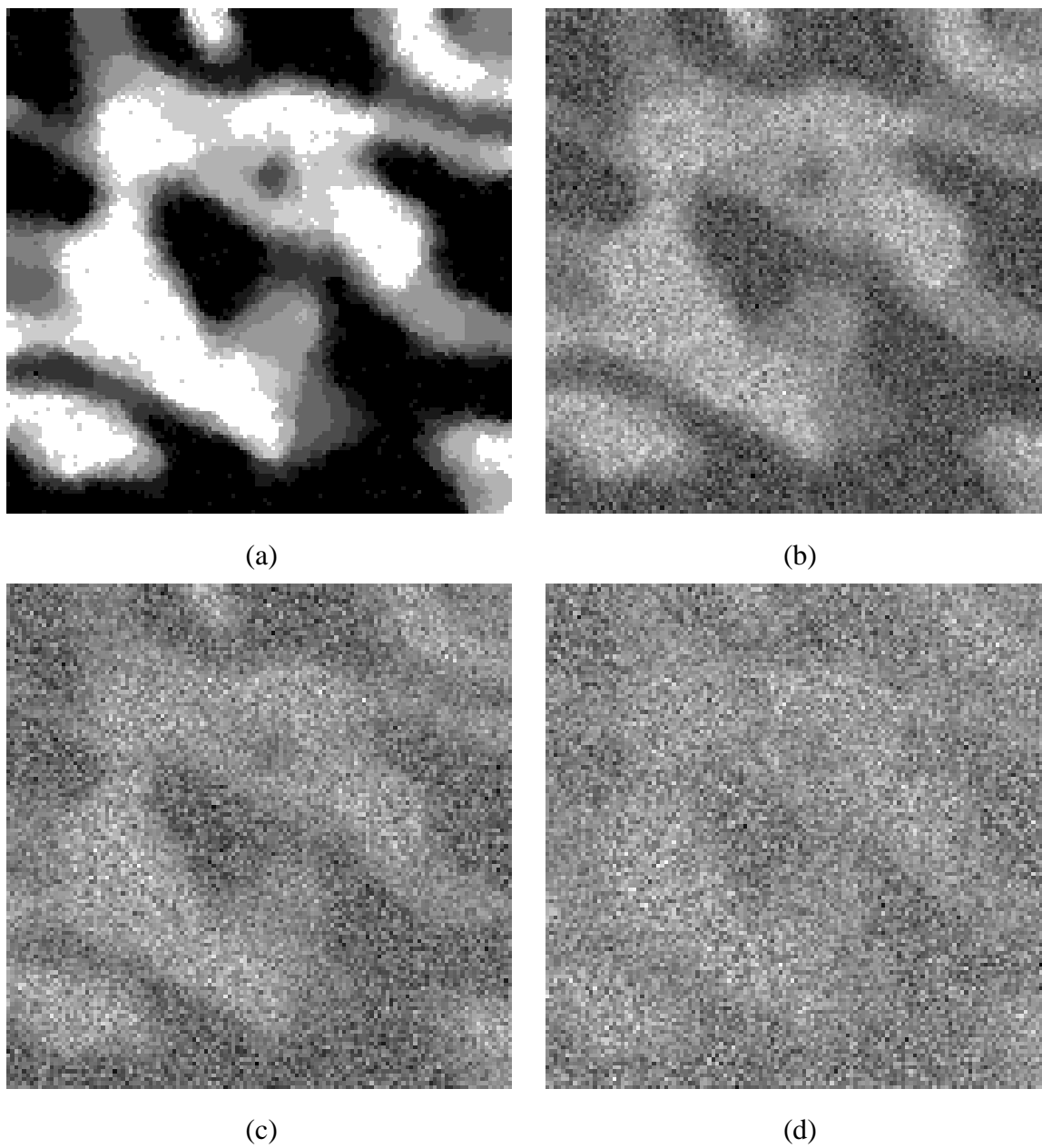


Fig. 2. FMF (a) on 10 fuzzy levels and its noisy versions in 3 bands (b,c,d). The prior parameters are given in Table III

TABLE I

DATA DRIVEN PARAMETERS OF THE IMAGES FIG.1(B)-(D) AND FIG.2(B)-(D)

Image	μ_0	μ_1	σ_0	σ_1
1-2(b)	120	140	4	4
1-2(c)	120	136	4	4
1-2(d)	120	132	4	4

TABLE II

ESTIMATION OF THE PRIOR PARAMETERS OF THE MULTIBAND IMAGES FIG.1 (B)-(D) FOR THE FMC (10 FUZZY LEVELS).

Images	π_{00}	π_{11}	π_{01}	π_{10}	b
real parameters of 1(a)	0.4	0.4	0	0	0
P-FMC $r = 1$	0.36	0.43	0.	0.	0.
P-FMC $r = 50$	0.38	0.42	0.	0.	0.
NP-FMC	0.39	0.42	0.	0.	-

2) *Segmentation results:* One will distinguish parameterized fuzzy Markov chain model (P-FMC) with $1 \leq r \leq 50$, towards non-parameterized FMC models (NP-FMC). For this purpose, one classifies first the multispectral noisy images 1(b,c,d), in order to compare the P-FMC and the NP-FMC. The Fig.3 shows the corresponding segmented images for $r = 1$ (fig. 3 (a)), $r = 50$ (fig. 3(b)) and NP-FMC (fig. 3(c)). Numerical results are given in Tab. II,V,VIII. One observes that both NP-FMC and P-FMC for $r = 50$ give satisfied results of prior distribution estimation

TABLE III

PRIOR PARAMETERS USED FOR FMF SIMULATION AND ESTIMATED PRIOR PARAMETERS OF THE MULTIBAND IMAGES FIG.2 (B)-(D).

	α_1	α_2	α_3	α_4	α_1^f	α_2^f	α_3^f	α_4^f
real parameters of 2(a)	2	2	2	2	2.5	2.5	2.5	2.5
estimated parameters	1.60	1.60	1.58	1.61	2.12	2.07	2.12	2.13

TABLE IV

ESTIMATION OF THE FMC PRIOR PARAMETERS OF THE MULTIBAND IMAGES FIG.2 (B)-(D).

Images	π_{00}	π_{11}	π_{01}	π_{10}	b
real parameters of 2(a)	0.26	0.16	0.	0.	0.
P-FMC $r = 1$	0.29	0.22	0.	0.	0.
P-FMC $r = 10$	0.17	0.10	0.	0.	-
P-FMC $r = 20$	0.17	0.10	0.	0.	-
P-FMC $r = 50$	0.26	0.15	0.	0.	-
NP-FMC	0.26	0.15	0.0	0.0	-

(Tab. II), as well as data driven densities estimation (Tab. VIII). The final error rate given in Tab. V, is reflecting the competitive behavior of the NP-FMC model, although the simulated data are P-FMC. The P-FMC model with $r = 50$ provides better error rate and also a more homogeneous image than the P-FMC with $r = 1$, as expected. Fig.5 illustrates the segmentation of Fig.2(b,c,d), using the NP-FMC model (Fig.5(a)) and FMF model (Fig.5(b)). Fig. 4(a)-(d) illustrates such segmentation using different P-FMC models for $1 \leq r \leq 50$. On one hand, the final realization obtained by the FMF processing gives satisfied results: it keeps the graduate transition between the levels, *i.e.*, an homogeneous field and gives the best error rate, but tends to lose the fuzzy information (45.91% versus 53.96%, see Tab. VII): the FMF seems to be less robust to the noise, in term of fuzzy rate information. However, the estimated prior parameters (Tab. III) as well as the estimated driven data parameters (Tab. VI) are well fitted with the initial ones. For FMF model, it is highly important to preserve the difference between α^f and α^d more than the exact values. On the other hand, the algorithms based on NP-FMC and P-FMC (for $r \geq 10$) preserve well the fuzzy rate (see Tab. VII) and the hyper parameters (see Tab. IV and Table VII-A.2-VII-A.2). The NP-FMC segmentation provides less homogeneous map than FMF segmentation (Fig.5(a) versus 5(b)), nevertheless FMC error rates are comparable to the FMF error (Tab. VII). However the more the parameter r increases, the more homogeneous is the image (see for example Fig. 4(a) for $r = 1$ vs Fig. 4(d) for $r = 50$). The best result in term of homogeneity, prior parameters and fuzzy rate is obtained with $r = 50$. Let us notice in this

TABLE V

ERROR AND FUZZY RATES FOR THE MULTIBAND IMAGES SEGMENTATION OF FIG. 1 (B)-(D) FOR THE FMC.

Mode	Initial fuzzy rate	Final fuzzy rate	Error rate
P-FMC $r = 1$	24.72 %	18.09 %	3.52 %
P-FMC $r = 50$	24.72 %	23.27 %	1.30 %
NP-FMC	24.72 %	24.52 %	1.37 %

case, the blocked effect due to the Hilbert-Peano path. Again, the parameter $r = 1$ gives the worst result in term of fuzzy rate. FMC Computing time is significantly faster than the FMF one, which takes half an hour on a 128 simulated image versus three minutes. Finally, under three different criteria i.e. error rate, fuzzy rate, processing time, the NP-FMC method and the P-FMC method with $r = 50$ seem to behave more efficiently. If we only take into account such criteria as the processing time and flexibility, the FMF procedure is not competitive.

TABLE VI

ESTIMATED DATA DRIVEN PARAMETERS OF THE IMAGES FIG. 2(B)-(D) FOR A FMF AND NP-FMC

Image	$\hat{\mu}_0$: FMF / NP-FMC	$\hat{\mu}_1$: FMF / NP-FMC	$\hat{\sigma}_0$: FMF / NP-FMC	$\hat{\sigma}_1$: FMF / NP-FMC
2(b)	119.55 / 120.13	140.33 / 139.94	3.98 / 4.11	4.00 / 4.09
2(c)	119.50 / 120.08	136.31 / 135.88	3.92 / 4.06	3.97 / 3.98
2(d)	119.55 / 120.03	132.31 / 132.03	3.91 / 4.02	3.99 / 3.95

captionEstimated data driven parameters of the images Fig. 2(b)-(d) for a P-FMC: $r = 1$ / P-FMC: $r = 10$

Image	$\hat{\mu}_0$	$\hat{\mu}_1$	$\hat{\sigma}_0$	$\hat{\sigma}_1$
2(b)	121.18 / 120.64	137.67 / 138.97	4.42 / 4.15	4.76 / 4.29
2(c)	120.95 / 120.54	134.25 / 135.23	4.24 / 4.10	4.32 / 4.02
2(d)	120.68 / 120.32	130.69 / 131.44	4.03 / 4.04	4.06 / 3.94

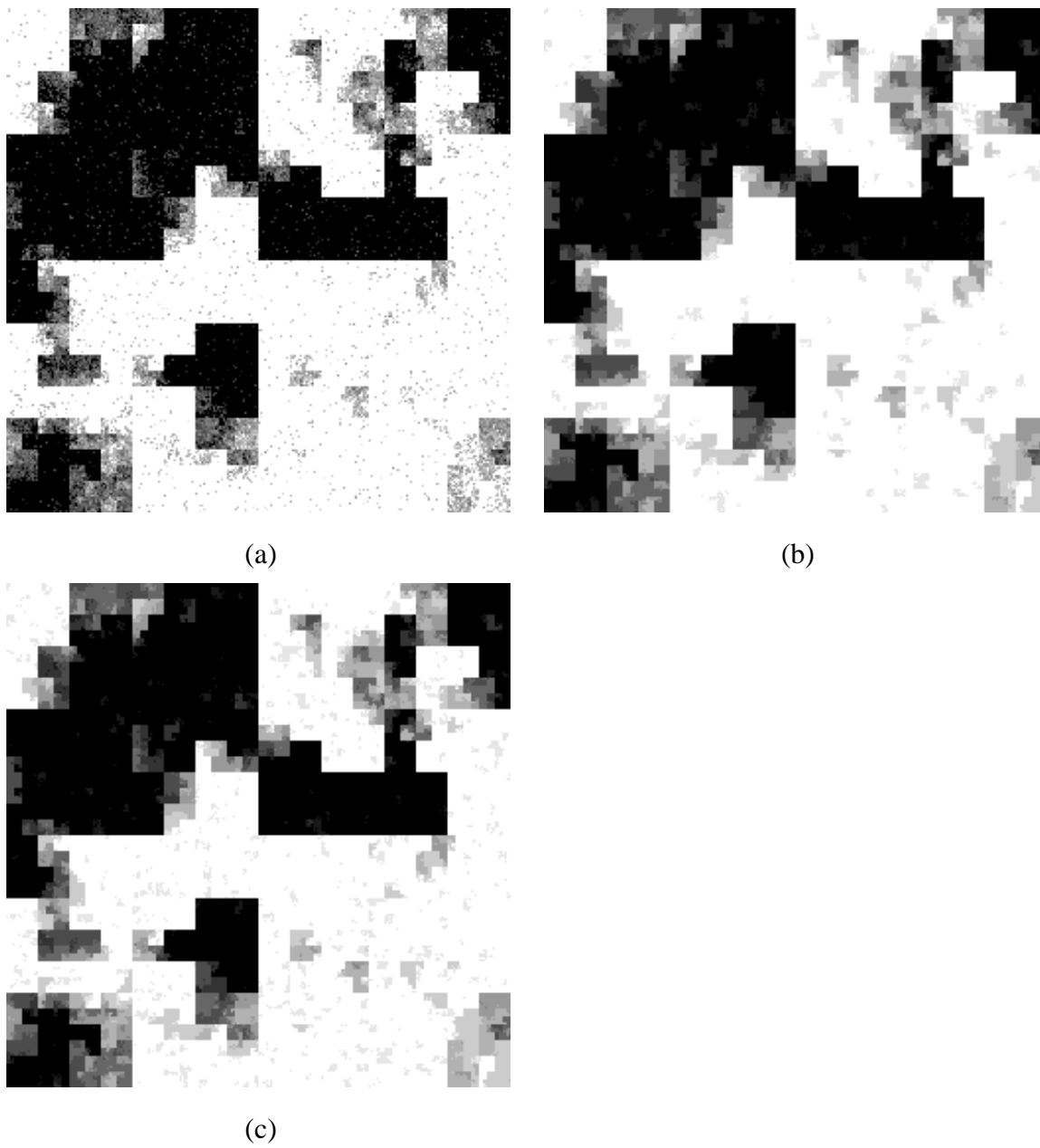


Fig. 3. Segmentation of Fig.1(b)-(d) using P-FMC with $r = 1$ (a), $r = 50$ (b) and NP-FMC (c).

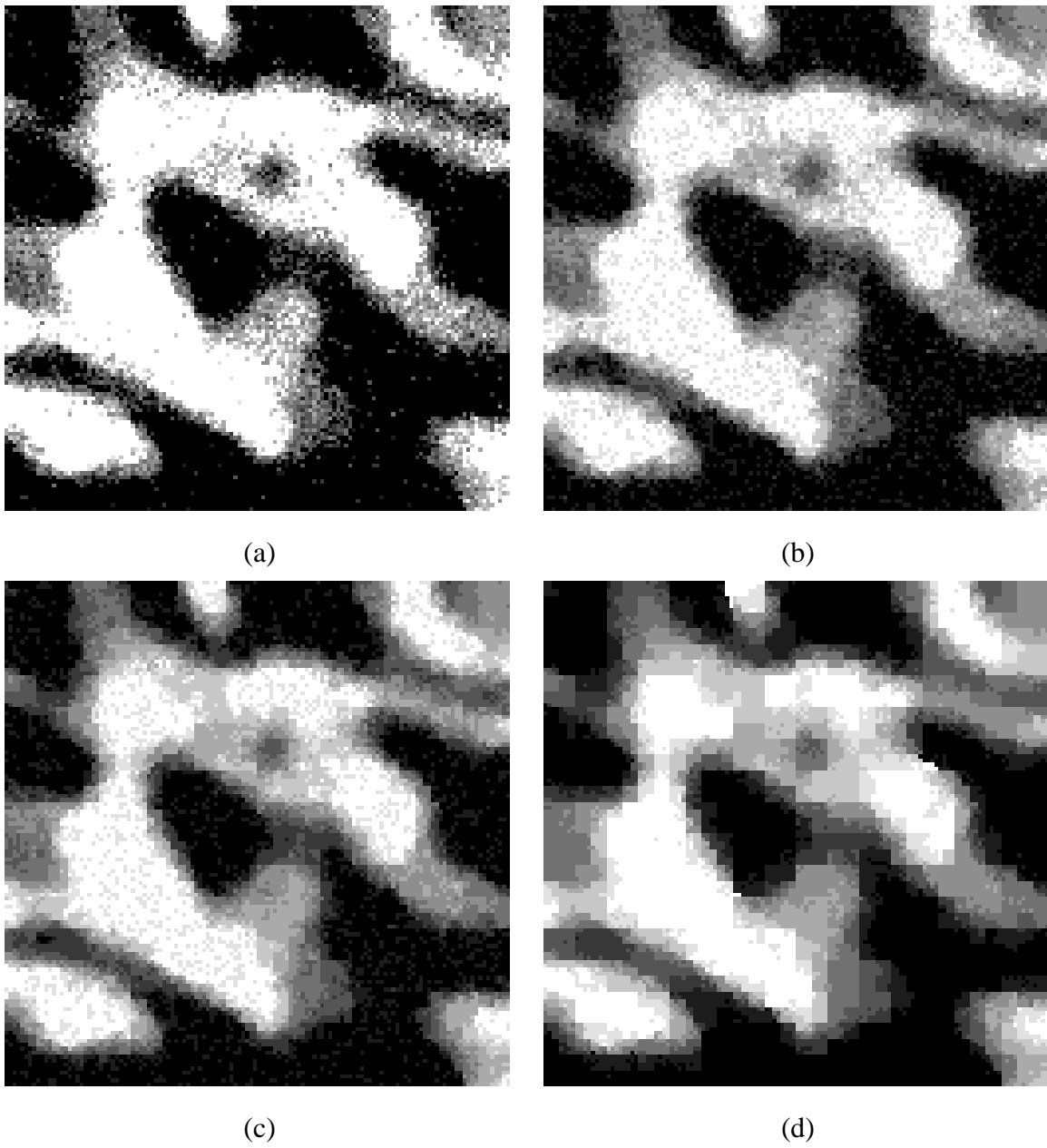


Fig. 4. Segmentation of 2(b)-(d) using P-FMC with (a) $r = 1$, (b) $r = 10$, (c) $r = 20$, (d) $r = 50$. When r reaches large values in the parameterized function (Eq. 3), one observes artefacts (blocky effects) due to the Hilbert-Peano scan

captionEstimated data driven parameters of the images Fig.refImChamp- for a P-FMC: $r = 20$ / P-FMC: $r = 50$

Image	$\hat{\mu}_0$	$\hat{\mu}_1$	$\hat{\sigma}_0$	$\hat{\sigma}_1$
2(b)	120.28 / 120.11	139.52 / 139.88	4.13 / 4.13	4.18 / 4.19
2(c)	120.25 / 120.09	135.66 / 135.84	4.07 / 4.09	3.99 / 4.02
2(d)	120.17 / 120.05	131.79 / 131.98	4.09 / 4.03	3.95 / 3.94

TABLE VII

ERROR AND FUZZY RATES FOR THE MULTIBAND IMAGES SEGMENTATION OF FIG.2(B)-(D) FOR THE FMC AND FMF

Segmentation mode	Initial fuzzy rate	Final fuzzy rate	Error rate
P-FMC $r = 1$	53.96 %	32.77 %	7.17 %
P-FMC $r = 10$	53.96 %	55.24 %	4.78 %
P-FMC $r = 20$	53.96 %	59.34 %	4.63 %
P-FMC $r = 50$	53.96 %	58.37 %	4.95 %
NP-FMC	53.96 %	55.76 %	4.39 %
FMF	53.96 %	45.91 %	2.94 %

B. Some perspectives on multispectral astronomical data

We present in Fig.6 multiband astronomical images in infrared domain at 60, 100 and 170 nanometers. We wish to measure the fluctuations due to the presence of dust cloud (cirrus) and galaxies. This problem, which is of great importance for the astrophysic community, did not find yet significant solutions. These fluctuations, produced by interstellar dust, bother the observations of galaxies behind them. Specific dusty area of the sky owns a threading-based structure, we want to recover. Moreover, they exhibit a fuzzy shape which leads us to apply our processing schemes. We performed the previous algorithms (FMF, NP-FMC, P-FMC for $r = 50$) and show

TABLE VIII

ESTIMATED PARAMETERS OF THE IMAGES FIG.1(B)-(D) FOR THE P-FMC: $r = 1$ / P-FMC: $r = 50$ / NP-FMC

Image	$\hat{\mu}_0$	$\hat{\mu}_1$	$\hat{\sigma}_0$	$\hat{\sigma}_1$
1(b)	120.34 / 120.00 / 120.03	139.73 / 140.05 / 140.12	4.14 / 4.03 / 4.04	4.07 / 3.99 / 4.00
1(c)	120.26 / 120.00 / 120.01	135.76 / 136.01 / 136.08	4.03 / 4.00 / 3.99	4.00 / 3.99 / 3.98
1(d)	120.22 / 120.01 / 120.04	131.85 / 132.03 / 132.08	4.04 / 4.04 / 4.03	4.01 / 4.02 / 4.03

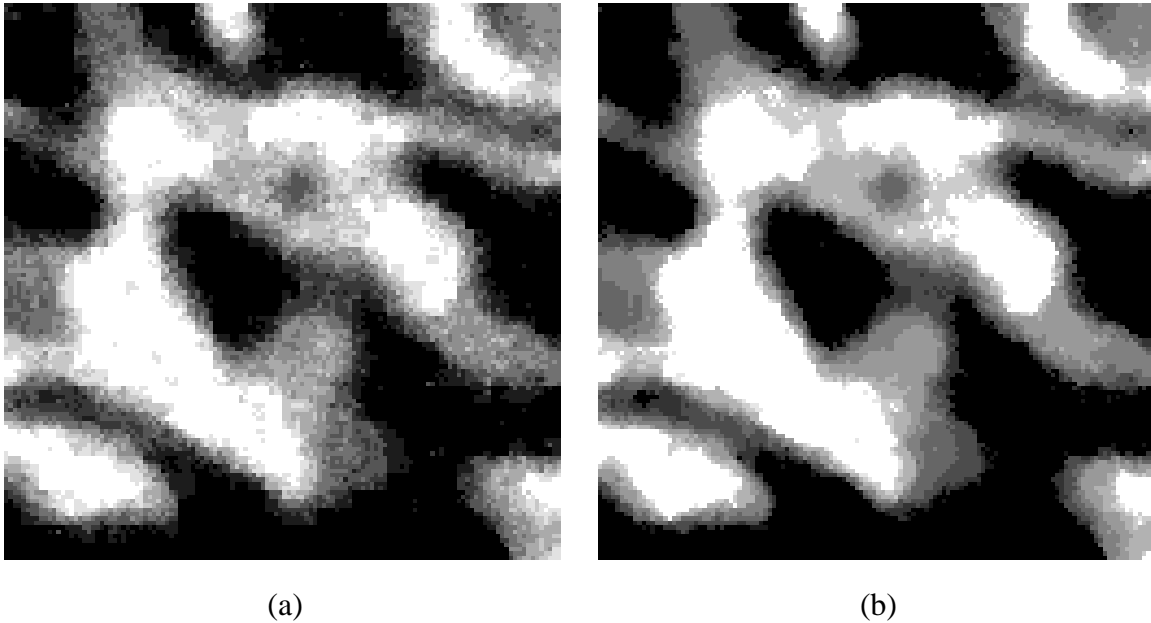


Fig. 5. Segmentation of 2(b)-(d) using NP-FMC (a) and FMF (b)

the classification maps in Fig.7(a,b,c), as well as the numerical results in Tab. IX, X, XI, XII. In Tab. X, the NP-FMC gives comparable results as P-FMC with $r = 50$ as mentioned in the previous results. The related segmented images contain comparable fuzzy information, similar structures at the bottom (Fig.7 and data driven parameters (Tab. XI). Another set of observation, Fig. 8 with missing data (black areas), are then investigated. We apply the rules described in section IV for FMC model. According to rule 1 we obtain FMC/FMF-maps on Fig.9(a,b,d), whereas the FMC-maps on Fig.9(c) have been processed with rule 2. Similar segmentation maps are observed between Fig. 9(a), 9(b), 9(c), 9(d). Nevertheless, rule 2 is interesting because in this case the Peano path does not take into account missing data locations. However, the P-FMC model does not converge when using the rule 2 and $r > 5$. Finally for a complete unsupervised algorithm, it worths to estimate the optimal value of r in presence of real data. The FMF, NP-FMC and P-FMC ($r = 50$) models give comparable fuzzy rates (see Tab.XVI).

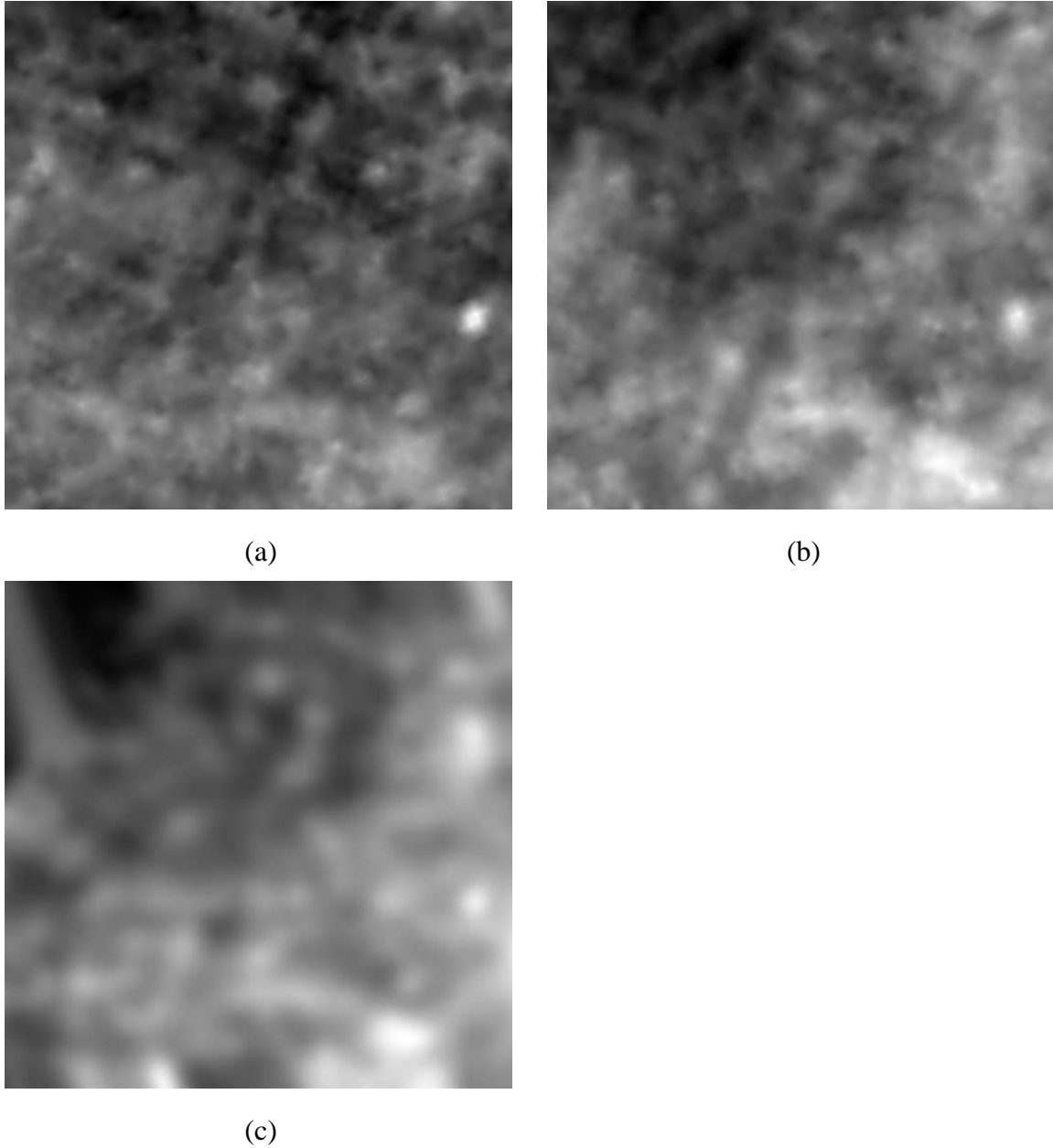


Fig. 6. Three multiband data in our galaxy, showing the fluctuations produced by an interstellar dust.

TABLE IX

ESTIMATION OF THE FMF PRIOR PARAMETERS, OF THE MULTIBAND IMAGES FIG.6(A)-(C).

$\hat{\alpha}_1^d$	$\hat{\alpha}_2^d$	$\hat{\alpha}_3^d$	$\hat{\alpha}_4^d$	$\hat{\alpha}_1^f$	$\hat{\alpha}_2^f$	$\hat{\alpha}_3^f$	$\hat{\alpha}_4^f$
0.69	0.76	0.63	0.64	3.50	3.54	3.41	3.36

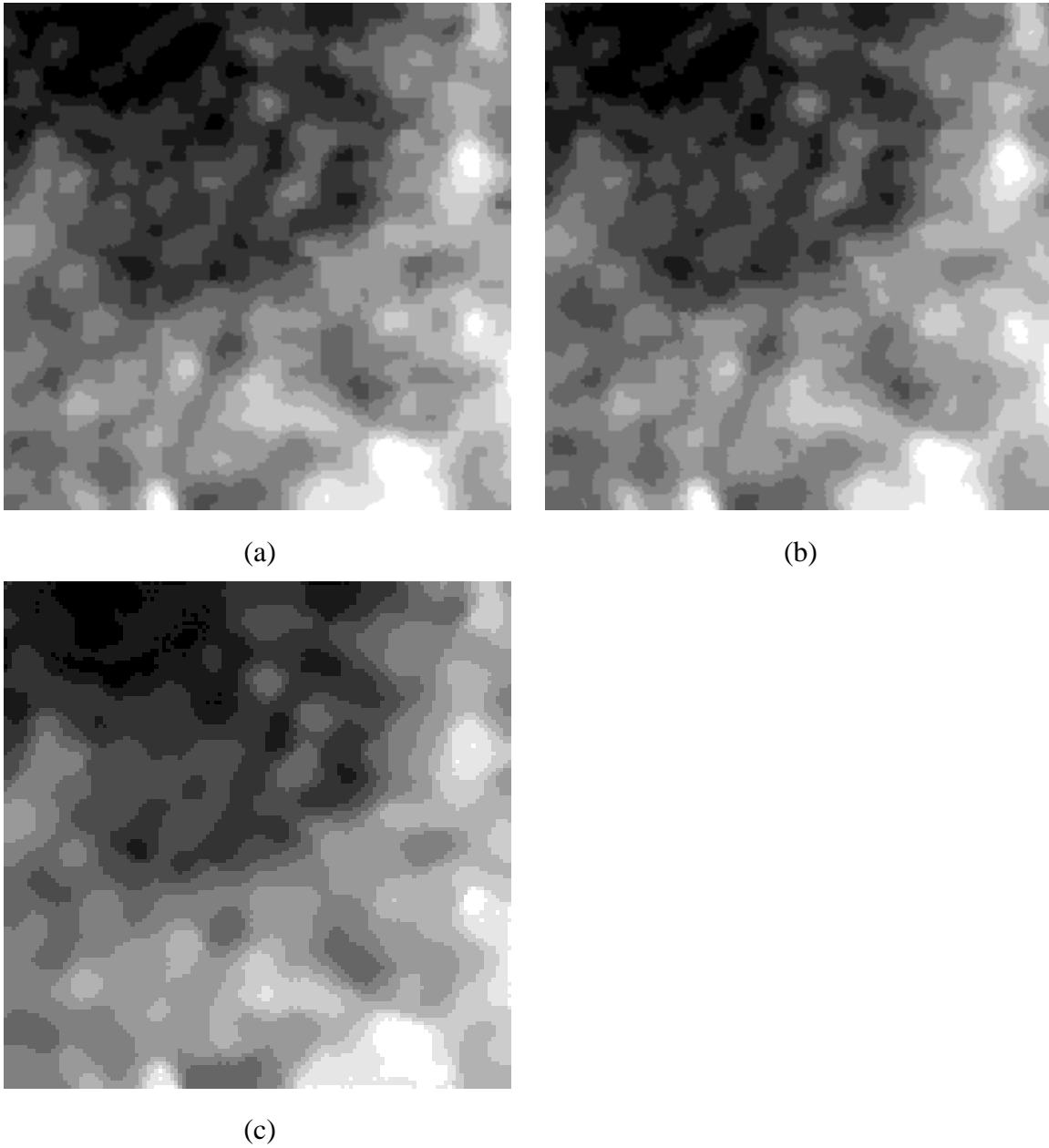


Fig. 7. Segmentation of Fig. 6(a)-(c) using P-FMC $r = 50$ (a), NP-FMC (b), FMF (c)

TABLE X

ESTIMATION OF THE FMC PRIOR PARAMETERS OF THE MULTIBAND IMAGES FIG.6(A)-(C)

	$\hat{\pi}_{00}$	$\hat{\pi}_{11}$	$\hat{\pi}_{01}$	$\hat{\pi}_{10}$	\hat{b}
P-FMC $r = 50$ (rule 1)	0.05	0.03	$1.5 \cdot 10^{-5}$	$1.5 \cdot 10^{-5}$	$1.5 \cdot 10^{-5}$
NP-FMC (rule 1)	0.05	0.02	0.	0.	-

TABLE XI

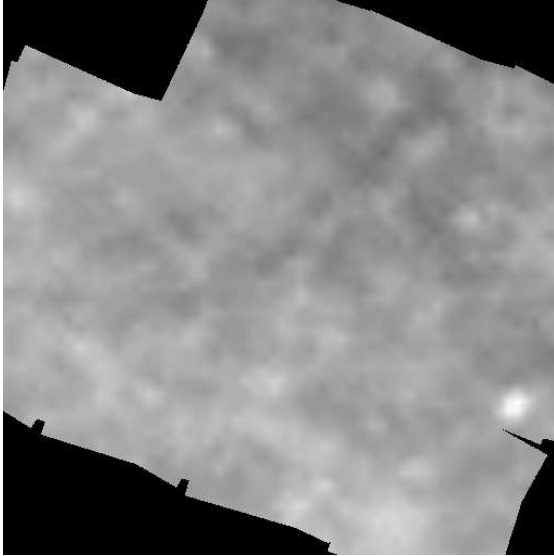
ESTIMATED DATA DRIVEN PARAMETERS OF THE IMAGES FIG.6(A)-(C) FOR A FMF / NP-FMC / P-FMC: $r = 50$.

Image	$\hat{\mu}_0$	$\hat{\mu}_1$	$\hat{\sigma}_0$	$\hat{\sigma}_1$
6(a)	0.37 / 0.37 / 0.37	0.52 / 0.51 / 0.52	0.02 / 0.02 / 0.03	0.04 / 0.04 / 0.04
6(b)	0.55 / 0.55 / 0.55	1.00 / 0.98 / 0.99	0.03 / 0.02 / 0.03	0.06 / 0.09 / 0.07
6(c)	1.47 / 1.49 / 1.50	2.07 / 2.11 / 2.10	0.07 / 0.08 / 0.08	0.07 / 0.04 / 0.05

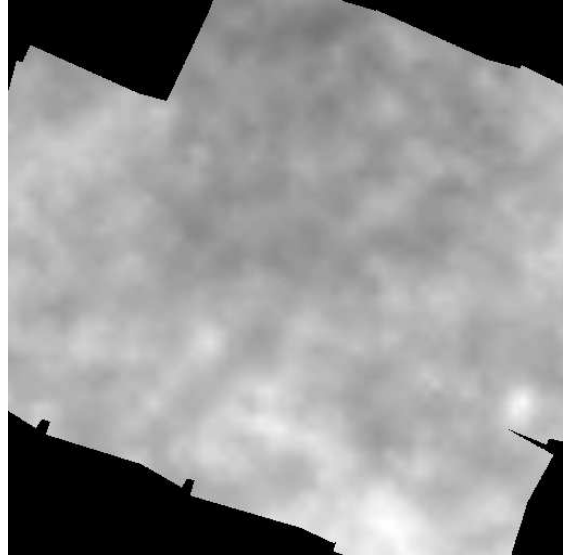
TABLE XII

FUZZY RATES FOR THE MULTIBAND IMAGES SEGMENTATION OF FIG.6 (A)-(C) FOR A FMF AND FMC.

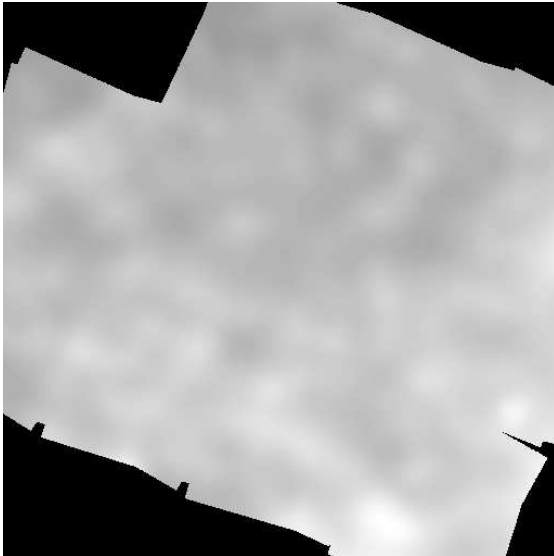
Segmentation mode	Final fuzzy rate
FMF	94.28 %
P-FMC $r = 50$	93.34 %
NP-FMC	93.58 %



(a)



(b)



(c)

Fig. 8. Three multiband astronomical image. One observes missing areas in black. This leads us to apply the methods (rule 1 and rule 2) described in section IV.C

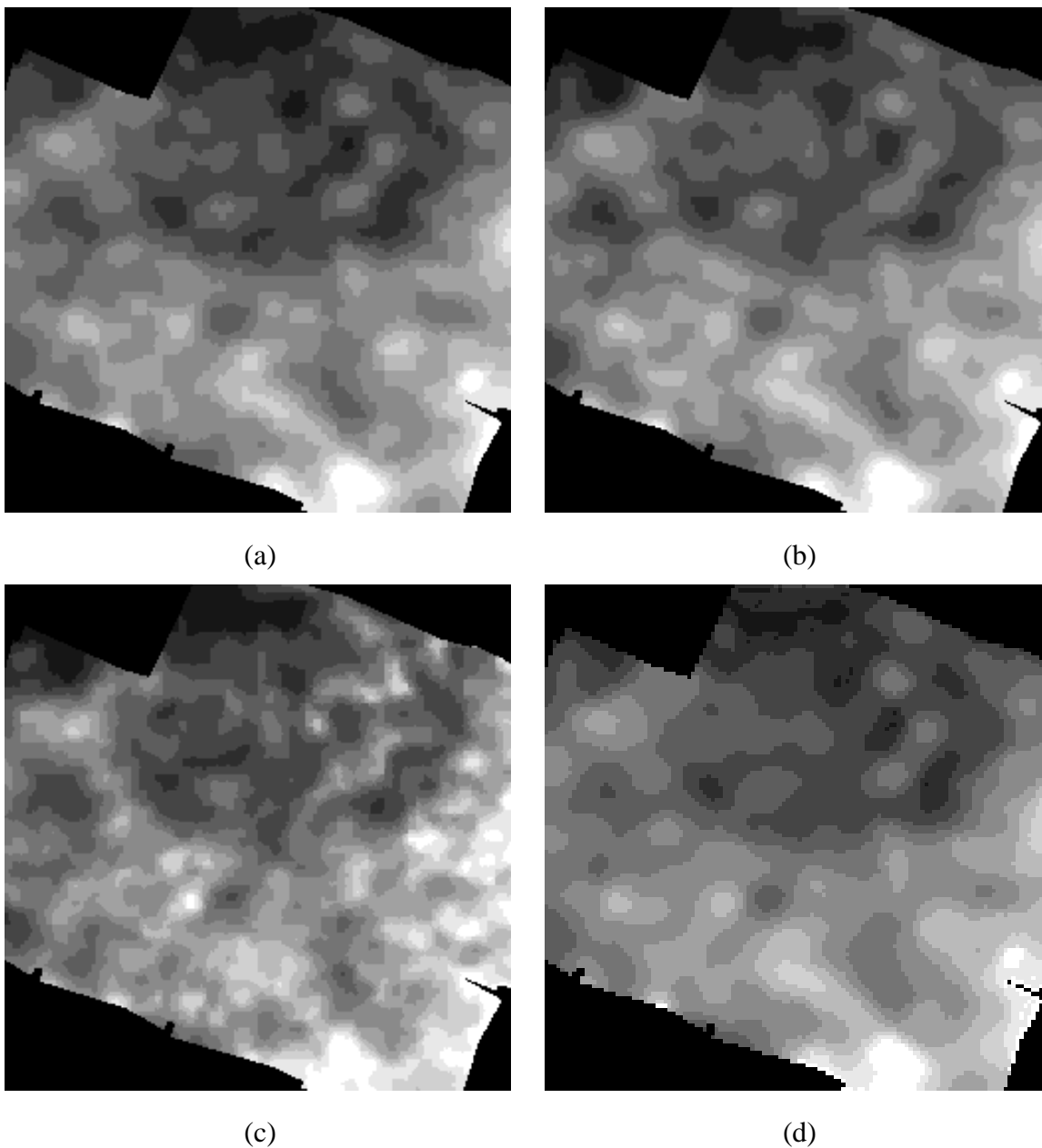


Fig. 9. Segmentation of 8(a)-(c) using P-FMC $r = 50$ using rule 1 (a), NP-FMC using rule 1 (b), rule 2 (c) and FMF using rule 1 (d). The rules are described in section IV.C

TABLE XIII

ESTIMATION OF THE FMF PRIOR PARAMETERS, OF THE MULTIBAND IMAGES FIG.8 (A)-(C).

$\hat{\alpha}_1^d$	$\hat{\alpha}_2^d$	$\hat{\alpha}_3^d$	$\hat{\alpha}_4^d$	$\hat{\alpha}_1^f$	$\hat{\alpha}_2^f$	$\hat{\alpha}_3^f$	$\hat{\alpha}_4^f$
0.74	0.78	0.7	0.73	3.48	3.5	3.43	3.4

TABLE XIV

ESTIMATION OF THE FMC PRIOR PARAMETERS OF THE MULTIBAND IMAGES FIG.8(A)-(C)

	$\hat{\pi}_{00}$	$\hat{\pi}_{11}$	$\hat{\pi}_{01}$	$\hat{\pi}_{10}$	\hat{b}
P-FMC $r = 50$ (rule 1)	0.04	0.02	9.10^{-5}	9.10^{-5}	9.10^{-5}
NP-FMC (rule 1)	0.04	0.03	0.	0.	-
NP-FMC (rule 2)	0.03	0.03	0.	0.	-

TABLE XV

ESTIMATED DATA DRIVEN PARAMETERS OF THE IMAGES FIG.8(A)-(C) FOR A FMF / NP-FMC / P-FMC: $r = 50$.

Image	$\hat{\mu}_0$	$\hat{\mu}_1$	$\hat{\sigma}_0$	$\hat{\sigma}_1$
8(a)	0.35 / 0.37 / 0.35	0.52 / 0.50 / 0.52	0.02 / 0.03 / 0.02	0.05 / 0.06 / 0.05
8(b)	0.63 / 0.66 / 0.64	1.1 / 1.02 / 1.06	0.04 / 0.05 / 0.04	0.08 / 0.11 / 0.09
8(c)	1.2 / 1.21 / 1.22	1.8 / 1.76 / 1.77	0.05 / 0.03 / 0.05	0.06 / 0.05 / 0.05

TABLE XVI

FUZZY RATES FOR THE MULTIBAND IMAGES SEGMENTATION OF FIG.8 (A)-(C) FOR FMF AND FMC.

Segmentation mode	Final fuzzy rate
FMF	96.84%
P-FMC (rule 1) $r = 50$	95.88 %
NP-FMC (rule 1)	95.37 %
NP-FMC (rule 2)	94.45 %

VIII. CONCLUSION

In this paper we used two main fuzzy hidden Markov models (FMF and FMC models) in order to segment multispectral images. The respective discrete and continuous nature of the hard and fuzzy intermediate classes are formalized by a measure including two Dirac functions and a Lebesgue measure on the set $]0, 1[$. One model was the Fuzzy Markov Field (FMF) described by an energy computed on clique functions. The other model we propose is based on fuzzy Markov chains (FMC) and is described by a joint density computed on neighbored pixels. We proposed to characterize this density using a parametric (P-FMC) and a non parametric (NP-FMC) approach. The parameterized density models the transition between the hard classes. The driven data distributions are multivariate gaussian densities under dependent sensor assumption: the variance-covariance matrix corresponding to the fuzzy classes depends linearly on the variance-covariance matrix related to the hard classes. We performed the segmentation in an unsupervised context using the MPM strategy with an absolute distance loss function. We compared these approaches according to three criteria: fuzzy rate, error rate and computing time. The results were performed on synthetic images, and some perspectives are drawn in the specific case of astronomical images, where fuzzy model seems to be of great interest. Finally, the NP-FMC and P-FMC model with a higher power, behave efficiently on synthetic and astronomical images. They preserve well the fuzzy rate. In particular, the P-FMC model gives good results in term of fuzzy gradation and homogeneity for higher parameters. Moreover, it is a fast method, because one has to estimate a small amount of hyper-parameters. These NP-FMC and P-FMC seem us to be a good compromise regarding to the other methods. We believe an extension of the fuzzy Markov models to the quadtree [13], [29] could be a powerful tool to process multiscale data.

REFERENCES

- [1] Louys M., Bot C., Oberto A., Collet C., "Multiwavelength image analysis of the Small Magellanic Cloud using hierarchical Markovian segmentation", *Third workshop on Physics in Signal and Image Processing*, pp 29-31, Grenoble, France, Jan. 2003.
- [2] L.E. Baum, T. Petrie, G. Soules, and N. Weiss, "Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains", *Ann. of Math. Stat.*, vol. 41, no. 1, pp. 164-171, 1970.
- [3] Geman S., Geman D., "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images", *IEEE, Trans. On Pattern Analysis and Machine Intelligence*, Vol. 6, pp 721-741, 1984.
- [4] Bezdek J.C., "Pattern Recognition and Fuzzy Objective Function Algorithm", *Plenum Press*, New York, 1981.
- [5] Gath I., Geva A.B., "Unsupervised Optimal Fuzzy Clustering", *IEEE, Trans. On Pattern Analysis and Machine Intelligence*, Vol. 11, n°7, pp 773-781, 1989.
- [6] Pedrycz W., "Fuzzy Sets in Pattern Recognition: Methodology and Methods", *Pattern Recognition*, Vol. 23, n°1/2, pp 121-146, 1990.
- [7] Zadeh L.A., "Fuzzy Sets", *Information and Control*, Vol 8, pp 338-352, 1965.
- [8] Caillol H., Hillion A., Pieczynski W., "Fuzy Random fields and Unsupervised Image Segmentation", *IEEE Trans. Geosci. Remote Sensing*, GE-31, n°4, pp 801-810, 1993.
- [9] J.T. Kent, K.V. Mardia, "Spatial Classification Using Fuzzy Membership", *IEEE, Trans. On Pattern Analysis and Machine Intelligence*, Vol. 10, n°5, pp 659-671, 1988.
- [10] Salzenstein F., Pieczynski W., "Parameter Estimation in Hidden Fuzzy Markovian Fields and Image Segmentation", *Graph. Models Process.*, Vol 59, n°4, pp 205-220, 1997.
- [11] Ruan S., Moretti B., Fadilli J., Bloyet D., "Fuzzy Markovien Segmentation in Application of Magnetic Resonance Images", *Computer Vision and Image Understanding*, Vol. 85, pp 54-69, 2002.
- [12] M. Germain, M. Voorons, J.M. Boucher and G.B. Bni, "Fuzzy statistical classification for multiband image fusion", *Proceedings of the Fifth International Conference on Information Fusion*, Vol. 1, pp. 178-184, Annapolis, USA, July 2002.
- [13] Laferte J.-M., Perez P., Heitz F. "Discrete Markov Image Modeling and Inference on the Quad-tree", *IEEE, Trans. On Image Processing*, Vol. 9, n° 3, pp. 390-404, 2000.
- [14] N. Giordana, W. Pieczynski, "Estimation of Generalized Multisensor Hidden Markov Chains and Unsupervised Image Segmentation", *IEEE Transactions on PAMI*, Vol. 19, n°5, pp 465-475, 1997.
- [15] J.-N Provost, Ch. Collet, P. Rostaing, P. Perez, P. Boutheymy, "Segmentation of Multispectral Spot Images for the Reconstruction of Water Depth Maps", *Computer Vision and Image Understanding*, Vol 93, Issue 2, pp. 155-174, January 2004.
- [16] Salzenstein F., Collet Ch., Petremand M., "Champs de Markov flous pour imagerie multispectrale - Fuzzy Markov random fields for multispectral images", *Traitement du Signal*, Vol. 21, n°1, pp 37-55, 2004.
- [17] Y. Lin, W. Chen and F.H. Chan, "Multi-class segmentation based on generalized fuzzy Gibbs random fields", *Proceedings of the Int. Conference on Image Processing (ICIP)*, Barcelona, Spain, September 2003.
- [18] M.A. Mohammed and P. Gader, "Generalized hidden Markov models-Part I: Theoretical frameworks", *IEEE Trans. Fuzzy Syst.*, Vol. 8, n°1, pp 67-81, 2000.
- [19] K. Avrachenkov and E. Sanchez, "Fuzzy Chains and Decision-Making", *Fuzzy Optimization and Decision Making*, Vol. 1, n°2, pp 143-159, 2002.

- [20] C. Carincotte, S. Derrode, G. Sicot and JM Boucher, "Unsupervised image segmentation based on a new fuzzy hidden Markov chain model", *Proceedings of the IEEE Int. Conference on Acoustic, Speech, Signal Processing (ICASSP)*, Montreal, Canada, May 2004.
- [21] C. Carincotte, S. Derrode and S. Bourennane, "Multivariate fuzzy hiddenMarkov chains model applied to unsupervised multiscale SAR image segmentation", *IEEE Int. Conference on Fuzzy Systems*, Reno, USA, 2005.
- [22] R. Fjørtoft, Y. Delignon, W. Pieczynski, M. Sigelle and F. Tupin, "Unsupervised segmentation of radar images using hidden Markov chains and hidden Markov random fields", *IEEE Transactions on GRS*, Vol. 41, n°3, pp 675-686, 2003.
- [23] Maroquin J.,Mitte S., Poggio T., " Probabilistic Solution of Ill-posed Problems in Computational Vision ", *Journal of The American Statistical Association*, Vol. 82, pp 76-89, 1987.
- [24] Hilbert D., "Über die stetige abbildung einer linie auf ein flächenstück", *Math. Ann.*, vol.38, pp.459-460, 1891
- [25] B. Benmiloud, W. Pieczynski, "Estimation des paramtres dans les chanes de Markov caches et segmentation d'images", *Traitement du Signal*, Vol. 12, n°5, pp 433-454, 1995.
- [26] Pieczynski W., "Statistical Image Segmentation", *Machine Graphics Vision*, Vol. 1, n°1, pp 261-268, 1992.
- [27] P. Devijver, "Baum's forward backward algorithm revisited", *Pattern Recognition Letters*, vol. 3, pp 369-373, 1985.
- [28] Younes L., "Parametric Inference for Imperfectly Observed Gibbsian Fields", *Probability Theory and Related Fields*, Vol. 82, pp 625-645, 1989.
- [29] Crouse M.S., Novak R.D., and Baraniuk R. G. - "Wavelet-Based Signal Processing using Hidden Markov Models", *IEEE Transactions on Signal Processing*, Vol.46, n°4, pp 886-902, 1998.