Pairwise Markov fields for segmentation in astronomical hyperspectral images

Jean-Baptiste Courbot[a,b,*], Vincent Mazet[a], Emmanuel Monfrini[c], Christophe Collet[a]

[a] *ICube, Université de Strasbourg - CNRS, 67412 Illkirch, France*
[b] *Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, CRAL UMR5574, F-69230, Saint-Genis-Laval, France*
[c] *SAMOVAR, Département CITI, CNRS, 91011 Évry, France*

Abstract

We consider the problem of segmentation in noisy, blurred astronomical hyperspectral images (HSI). Recent methods based on an hypothesis-testing framework handle the problem, but do not allow to use a prior on the result and often fail in the presence of strong noise. This paper introduces a pairwise Markov field model, allowing the unsupervized Bayesian segmentation of faint sources in astronomical HSI. Results on synthetic images show that the segmentation methods outperform their state-of-the-art counterparts, and allow the detection at very low SNR. Besides, results on real images provide relevant detections with respect to the application.

Keywords: Markov Random Fields, Pairwise Markov Fields, Bayesian Unsupervized Segmentation, Blurred Hyperspectral Image Segmentation

1. Introduction

This paper considers the problem of astronomical Hyperspectral Image (HSI) segmentation, which consists in estimating the classes $\mathbf{X}$ from an observation $\mathbf{Y} = \mathbf{y}$. This kind of images can be provided, *e.g.*, by the Multi-Unit Spectroscopic Explorer (MUSE) instrument [4]. Among others, the MUSE instrument targets the distant Universe to study very faint objects. Considering that these objects are blurred by a spread function and buried in the noise, both originating from the observational setting (instrument and atmosphere), their segmentation is particularly difficult.

The problem of faint source detection in astronomical HSI was recently addressed by dedicated methods within an hypothesis testing framework [8, 3, 2], in which the absence and presence of signal are two competing hypothesis. In [8], we accounted for spatial and spectral priors to perform a local contextual detection, and errors were handled locally with the expected Probability of False Alarm (PFA). In [3, 2], the authors use a matched-filter preprocessing to tackle the spatial structure, and the errors are globally handled through robust False Discovery Rate (FDR) control. Both methods perform locally, and the handling of the Field Spread Function (FSF) relies on contextual information. An alternative can be found in [19], where the authors use a marked point process modeling within a Monte Carlo Markov chain procedure to detect sources in astronomical HSI. This model however relies on the expected shapes of objects (ellipses), which does not matches the general cases of sources detection.

In a more general fashion, the detection can bee seen as a segmentation task. For this purpose, a Markov field modeling is relevant in order to improve the regularity of the segmentation results. Numerous works were published for Bayesian HSI segmentation, using hidden Markov fields (HMF) models [24, 22, 15, 13, 16, 27]. Other works specifically handle astronomical multi-spectral images [23, 26]. However, as far as we know, there is no existing work on the Bayesian segmentation of faint objects in astronomical HSI handling the FSF. Formally, the FSF cannot be modeled within an HMF framework. This is why we turn, in this paper, to the Pairwise Markov Field (PMF) model [21] which is a generalization of HMF models.

Using Markov prior for segmentation of blurred images is not new in image processing, see *e.g.* [1, 28, 25]. However, these methods often take the form of a "hidden" class field that is blurred and corrupted by noise. In this paper, we use the PMF model which allows to overcome the limitations of HMF models by relaxing the Markov distribution assumption on the class field. Indeed, the PMF model considers the joint Markovianity of the (observation, class) pair.

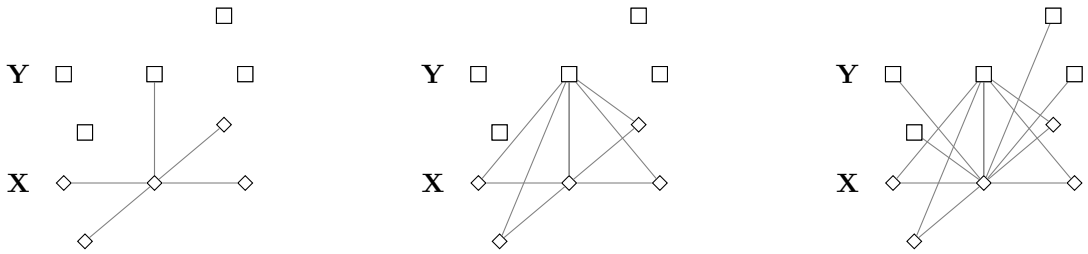
The contributions of this paper are the following. We introduce a PMF model which allows the modeling and the unsupervized Bayesian segmentation of blurred images. This model is applied for the detection of signals in extremely noisy HSI (Section 2 and Section 3). The proposed model is evaluated on synthetic images, and it is shown that it outperforms the state-of-the-art methods from [8, 3, 2] in Section 4. Finally, the segmentation of real MUSE images provides, in Section 5, satisfying results with regards to the target application.

2. Modeling

Along this paper, a random variable (resp. vector) is noted A (resp. $\mathbf{A}$), its realization a (resp. $\mathbf{a}$), and its distribution $p(a)$ (resp. $p(\mathbf{a})$).

Let $\mathbf{Y} = (\mathbf{Y}_s)_{s\in\mathcal{S}}$ and $\mathbf{X} = (X_s)_{s\in\mathcal{S}}$ be respectively the observation process and the class process, which has to be retrieved by

*Corresponding author: jean-baptiste.courbot@uha.fr
JBC is now with Institut de Recherche en Informatique, Mathématiques, Automatique et Signal (IRIMAS EA7499), Université de Haute-Alsace, Mulhouse, France.



(a) HMF. (b) CoPMF. (c) MGCoPMF.

Figure 1: Dependency graph corresponding to the CoPMF models introduced in this paper. The graphs represent the distribution of $(\mathbf{X}, \mathbf{Y})$ given the center node $(X_s, \mathbf{Y}_s)$. Roughly speaking, deleting a graph edge is equivalent to not accounting for a variable in one of the conditional distributions ruling the process $(\mathbf{X}, \mathbf{Y})$. The graph corresponds to a local neighborhood, restricted to 4 sites for clarity purpose.

the segmentation task. $\mathcal{S}$ is the lattice of the sites s in the image, which is ruled by a 8-neighborhood system denoted $(N_s)_{s\in\mathcal{S}}$ and associated to two-elements cliques. For all $s \in \mathcal{S}$, X_s takes values in the finite set Ω and $\mathbf{Y}_s \in \mathbb{R}^\Lambda$, Λ being the number of spectral channels of $\mathbf{Y}$.

Within the pairwise Markov field (PMF) modeling [21], the pair $(\mathbf{X}, \mathbf{Y})$ is a Markov field:

$$p(\mathbf{x}, \mathbf{y}) \propto \prod_{s\in\mathcal{S}} p(x_s, \mathbf{y}_s | \mathbf{x}_{N_s}, \mathbf{y}_{N_s}). \tag{1}$$

We assume that $(\mathbf{X}, \mathbf{Y})$ is stationary and that $\mathbf{Y}_s$ and $\mathbf{Y}_{N_s}$ are independent given $(X_s, \mathbf{X}_{N_s})$:

$$p(x_s, \mathbf{y}_s | \mathbf{x}_{N_s}, \mathbf{y}_{N_s}) = p(\mathbf{y}_s | x_s, \mathbf{x}_{N_s}) p(x_s | \mathbf{x}_{N_s} \mathbf{y}_{N_s}) \tag{2}$$

The two resulting distributions are detailed in the following subsections.

2.1. Observation model

We assume the noise to be normally distributed:

$$p(\mathbf{y}_s | x_s, \mathbf{x}_{N_s}) \sim \mathcal{N}(\boldsymbol{\gamma}_s, \boldsymbol{\Sigma}); \tag{3}$$

where $\boldsymbol{\gamma}_s \in \mathbb{R}^\Lambda$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{\Lambda\times\Lambda}$ are the mean and covariance matrix respectively.

The parameter $\boldsymbol{\gamma}_s$ models the convolution by the FSF $\mathbf{f}$[1]:

$$\boldsymbol{\gamma}_s = f_0 \boldsymbol{\mu}_s + \sum_{r\in N_s} f_r \boldsymbol{\mu}_r; \tag{4}$$

where $\boldsymbol{\mu}_s \in \mathbb{R}^\Lambda$ is the spectral mean associated to the class of x_s. We assume in the remaining of this paper that $\mathbf{f}$ is known prior to the segmentation, *e.g.* from instrumental calibration.

In this paper, the covariance matrix is designed to model the MUSE observations (described in Section 5): it is pentadiagonal, with the same parameters for each spectral component [4]:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2 & \rho_1 & \rho_2 & 0 & \cdots & 0 \\ \rho_1 & \ddots & \ddots & \ddots & \ddots & \\ \rho_2 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \rho_2 \\ & \ddots & \ddots & \ddots & \ddots & \rho_1 \\ 0 & & 0 & \rho_2 & \rho_1 & \sigma^2 \end{bmatrix} \tag{5}$$

[1]Note that since since $(\mathbf{X}, \mathbf{Y})$ is stationary, $\mathbf{f}$ does not depend on s.

This configuration can be adapted to other applications.

2.2. Convolutional pairwise Markov fields

In this paper, we investigate two possible choices for the second term of (2):

- in the first case, we assume that X_s and $\mathbf{Y}_{N_s}$ are independent given $\mathbf{X}_{N_s}$, and that $\mathbf{X}$ has a Markov field distribution. We choose to use an Ising potential as it is a widespread prior for image regularization [5]:

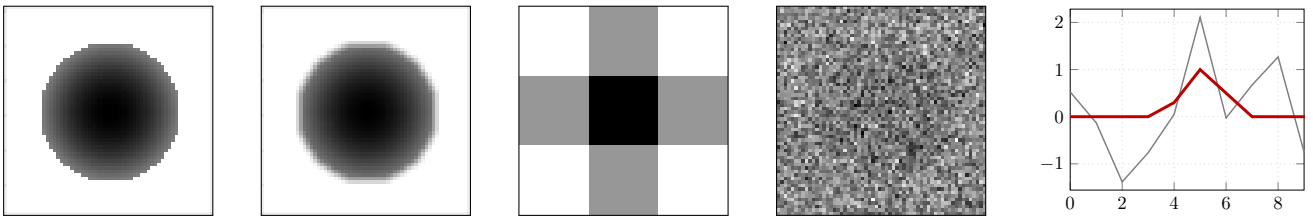
$$p(x_s | \mathbf{x}_{N_s}, \mathbf{y}_{N_s}) = p(x_s | \mathbf{x}_{N_s}) \propto \exp\left(-\sum_{s'\in N_s} \alpha \left(1 - 2\delta_{x_{s'}}(x_s)\right) \right) \tag{6}$$

  where $\delta_{x_{s'}}$ is the Kronecker function of $x_{s'}$ and α is a model parameter controlling the "granularity" of the realization $\mathbf{x}$. This model will be hereafter referred as "Convolutional Pairwise Markov Field" (CoPMF).
- In the second case, we choose to penalize the gaps between the spectral mean at the site s, and the observed spectra in the neighboring sites within N_s. To do so, we add a Gaussian term to (6), yielding:

$$p(x_s | \mathbf{x}_{N_s}, \mathbf{y}_{N_s}) \propto \exp\left[-\sum_{s'\in N_s} \alpha \left(1 - 2\delta_{x_{s'}}(x_s)\right) - \frac{\sqrt{|N_s|}}{2} \left(\boldsymbol{\gamma}_s - \frac{1}{|N_s|} \sum_{s'\in N_s} \mathbf{y}_{s'} \right)^\top \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{\gamma}_s - \frac{1}{|N_s|} \sum_{s'\in N_s} \mathbf{y}_{s'} \right) \right] \tag{7}$$

  where $|\cdot|$ is the cardinal function. This term represents the adequacy of the average of $\mathbf{y}_{s'}$, $s' \in N_s$ with respect to the classes in s seen through convolution. This model will hereafter be referred as "Multivariate Gaussian CoPMF" or MGCoPMF.

Fig. 1 illustrates the dependency graphs yielded by the two models, in comparison with the classical HMF model. The two models will be evaluated in Section 4.



(a) Intensity map. (b) Intensity map after convolution. (c) 3×3 Moffat FSF. (d) Spectral mean of a realization $\mathbf{Y} = \mathbf{y}$. (e) Noiseless (red) and noisy (gray) realizations of $\mathbf{y}_s$.

Figure 2: Synthetic image formation.

3. Unsupervized Bayesian segmentation

3.1. Bayesian segmentation

We compare two estimators for segmentation, namely the Maximum A Posteriori (MAP) [14]:

$$\hat{\mathbf{x}}^{\text{MAP}} = \arg\max_{\omega \in \Omega^{|\mathcal{S}|}} p(\mathbf{X} = \omega | \mathbf{Y} = \mathbf{y}); \tag{8}$$

and the Maximum Posterior Mode (MPM) [17]:

$$\forall s \in \mathcal{S}, \hat{x}_s^{\text{MPM}} = \arg\max_{\omega \in \Omega} p(x_s = \omega | \mathbf{Y} = \mathbf{y}). \tag{9}$$

The two criterion are complimentary, since the MPM minimizes a site-wise error (misclassification) risk while the MAP minimize a global error risk [17].

The exact distributions $p(\mathbf{x}|\mathbf{y})$ is known only up to a constant. However, the segmentation remains possible, using *e.g.* adaptations of Iterated Conditional Modes (ICM) [5] for the MAP and Marroquin's algorithm [17] for the MPM. These methods rely on simulation algorithms, such as the Gibbs sampling [14].

3.2. Parameter estimation

In the following, we focus on a two-class problem, with $\Omega = \{0, 1\}$. We assume $\boldsymbol{\mu}_s = \mathbf{0}$ when $x_s = 0$ and denote $\boldsymbol{\mu}_s = \boldsymbol{m}$ when $x_s = 1$. In an unsupervised context, the unknown parameters $\theta = \{\boldsymbol{m}, \sigma, \rho_1, \rho_2, \alpha\}$ must be estimated prior to segmentation with dedicated methods. It is possible to use, *e.g.*, methods inspired from the Expectation-Maximization (EM) [10, 18], the Stochastic EM (SEM) [7] or the Iterative Conditional Estimator (ICE) [9] algorithms. For robustness reasons [20, 12], we use an adaptation of the SEM method following, at a given step q:

- simulation of $\mathbf{x}^{(q)}$ along $p_{\theta^{(q-1)}}(\mathbf{x}|\mathbf{y})$,
- estimation of $\theta^{(q)}$ from $(\mathbf{x}^{(q)}, \mathbf{y})$ with the estimators described below.

For convenience, we note $\mathbf{w} = (w_s)_{s \in \mathcal{S}} = \mathbf{x} * \mathbf{f}$ the convolution[2] of $\mathbf{x}$ by the FSF $\mathbf{f}$ such that $w_s \boldsymbol{m} = \boldsymbol{\gamma}_s$, $\forall s \in \mathcal{S}$ (see Eq. (4)). For the CoPMF model, the MLE estimators are:

$$\hat{\boldsymbol{m}} = \frac{\sum_{s \in \mathcal{S}} w_s \mathbf{y}_s}{\sum_{s \in \mathcal{S}} w_s^2} \text{ and } \hat{\boldsymbol{\Sigma}} = \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} (\mathbf{y}_s - w_s \hat{\boldsymbol{m}})(\mathbf{y}_s - w_s \hat{\boldsymbol{m}})^\top. \tag{10}$$

[2]This implies a slight abuse of notation by using labels as numerical values.

For the MGCoPMF model, the estimators are:

$$\hat{\boldsymbol{m}} = \frac{\sum_{s \in \mathcal{S}} w_s \left(\mathbf{y}_s + \sqrt{|N_s|} \sum_{s' \in N_s} \mathbf{y}_{s'} \right)}{(1 + \sqrt{|N_s|}) \sum_{s \in \mathcal{S}} w_s^2} \tag{11}$$

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{2|\mathcal{S}|} \sum_{s \in \mathcal{S}} \left[(\mathbf{y}_s - w_s \hat{\boldsymbol{m}})(\mathbf{y}_s - w_s \hat{\boldsymbol{m}})^\top + \left(w_s \hat{\boldsymbol{m}} - \sum_{s' \in N_s} \mathbf{y}_{s'} \right) \left(w_s \hat{\boldsymbol{m}} - \sum_{s' \in N_s} \mathbf{y}_{s'} \right)^\top \right]. \tag{12}$$

The proof for the estimators $\hat{\boldsymbol{\mu}}$ are given in Appendix A and the derivations for $\hat{\boldsymbol{\Sigma}}$ follow the same principles. Thanks to the function invariance property of the maximum likelihood estimator [6], the estimators of σ, ρ_1, ρ_2 are obtained by:

$$\hat{\sigma} = \sqrt{\frac{1}{\Lambda} \sum_{i=1}^{\Lambda} \hat{\boldsymbol{\Sigma}}_{i,i}}; \quad \hat{\rho}_1 = \frac{1}{\Lambda - 1} \sum_{i=1}^{\Lambda - 1} \hat{\boldsymbol{\Sigma}}_{i,i+1}; \quad \hat{\rho}_2 = \frac{1}{\Lambda - 2} \sum_{i=1}^{\Lambda - 2} \hat{\boldsymbol{\Sigma}}_{i,i+2}. \tag{13}$$

Finally, α is estimated with the least-squares estimator of [11]. To sum up, the Bayesian unsupervized detection within blurred and noisy HSI is made possible by both CoPMF and MGCoPMF models, using either the MPM or the MAP estimators.

4. Numerical results

4.1. Settings and alternatives

In this section, we evaluate the method for detection of faint sources in hyperspectral images. The numerical experiments are conducted on HSI containing 60×60 spectra and $\Lambda = 10$ spectral channels. The noiseless image is a 2D Gaussian where the faintest intensities are set to zero (cf. Fig. 2a). The intensities are then convolved by a 3×3 Moffat FSF (cf. figures 2c, 2b and also [8]). Finally, $\boldsymbol{m}$ contains 3 non-zero coefficients, as shown in Fig. 2e.

We evaluate the four proposed methods:

1. the CoPMF model with the MAP (8) ;
2. the CoPMF model with the MPM (9);
3. the MGCoPMF model with the MAP (8);

4. the MGCoPMF model with the MPM (9);

For comparison, we also evaluate:

5. the detection method introduced in [8], based on an hypothesis-testing framework. We use the same experimental settings proposed as in [8] concerning the choice of the catalog and the P_{FA}, and use the same FSF as in the synthetic images;

6. the detection method proposed in [3] and extended in [2]. We perform the preprocessing with the same FSF as in the images, and set the FDR value to 0.1.

All methods are evaluated under a varying signal-to-noise ratio (SNR), defined as

$$\mathrm{SNR} = 10\log_{10}\left(\frac{\|\overline{\boldsymbol{m}}\|_2^2}{\Lambda\sigma^2}\right); \tag{14}$$

where $\overline{\boldsymbol{m}}$ the product of $\boldsymbol{m}$ by the average of non-zero intensities in the intensity map. An insight of the results for the six evaluated method is provided in Fig. 3.

4.2. MGCoPMF and CoPMF

In Fig. 4, we report the performances of the CoPMF-based methods in terms of false classification, false positive and false negative rates. These results yield the following findings:

- the four methods yield very close averaged error rates, providing for all of them less than 10% error rate for SNR $>$ -14dB. In most cases, the MGCoPMF with the MAP yields slightly better results than the other methods, and for both MGCoPMF and MPM, the MAP estimator yields smaller error rates than the MPM estimator;
- these two estimators have however distinct behaviors in terms of false positive rates. One one hand, the MAP provides less false negatives that any other methods (less than 10% when RSB $>$ -14 dB), at the cost of having small, but non-zero false positive rates. On the other hand, the MPM provides almost no false alarm (less than 3% at all SNR), but has a non-zero false negative rate. We interpret this observation as an insight of the stability of the MPM criterion with respect to the MAP in practical implementation.
- in average, the MGCoPMF yields better results with either the smaller error rates for the MGCoPMF with the MAP, or the smallest error rate with zero-valued false alarms for the MGCoPMF with the MPM.

4.3. Comparison with state-of-the-art alternatives

Fig. 5 reports the performance comparison of the MGCoPMF model with [8] and [3, 2]. The investigation of these results reveals the following points:

- the CoPMF segmentations yield, at all SNR, the best error rate with respect to the alternatives, with similar error rates for the MAP and MPM estimators;
- the method from [2] yields fairly high error rate when SNR $<$ -8 dB, due to an excess of non-detection required for the FDR control;

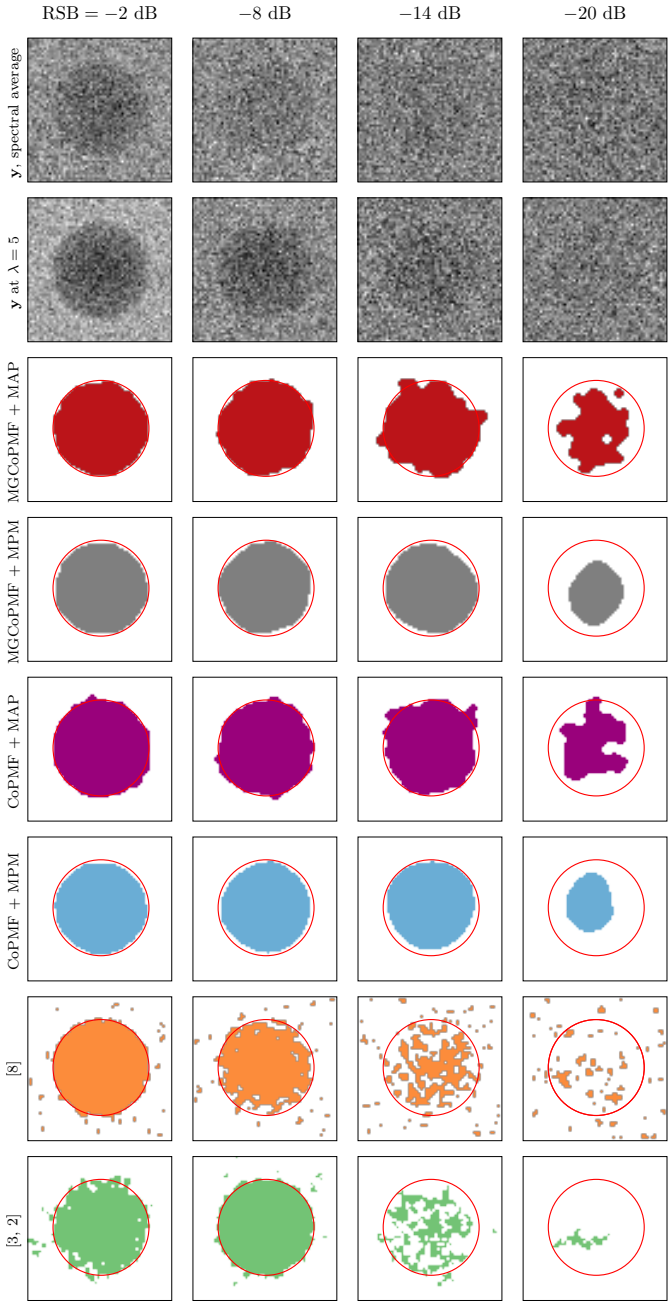


Figure 3: Results insights under low SNR for synthetic images. Red circles represent the ground truth (cf. Fig. 2a).

- the method from [8] provides no clear advantage over the other detection methods. This is partly caused by the use of the 3×3 pixels FSF from the simulations.

5. Real MUSE images

Detection in astronomical hyperspectral images may reveal spatially extended, spectrally-located light emission called Lyman-alpha, tracing the circum-galactic medium (CGM). The CGM may be occurring in the vicinity of known galaxies, making possible to extract from a MUSE observation (see [4]) a smaller HSI centered spatially and spectrally on the emission

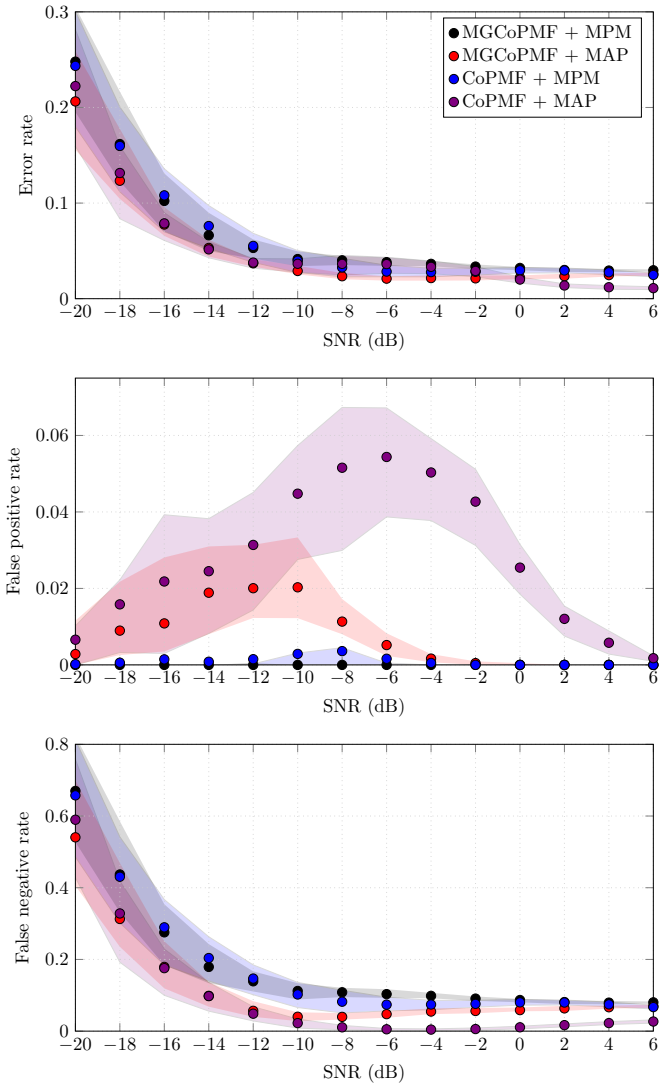


Figure 4: Average performances of the MGCoPMF and CoPMF models using the MAP (8) and the MPM (9). Each point is an average results, and the shaded regions are between the first and fourth quartiles. The results were obtained with 100 realizations $\mathbf{Y} = \mathbf{y}$.

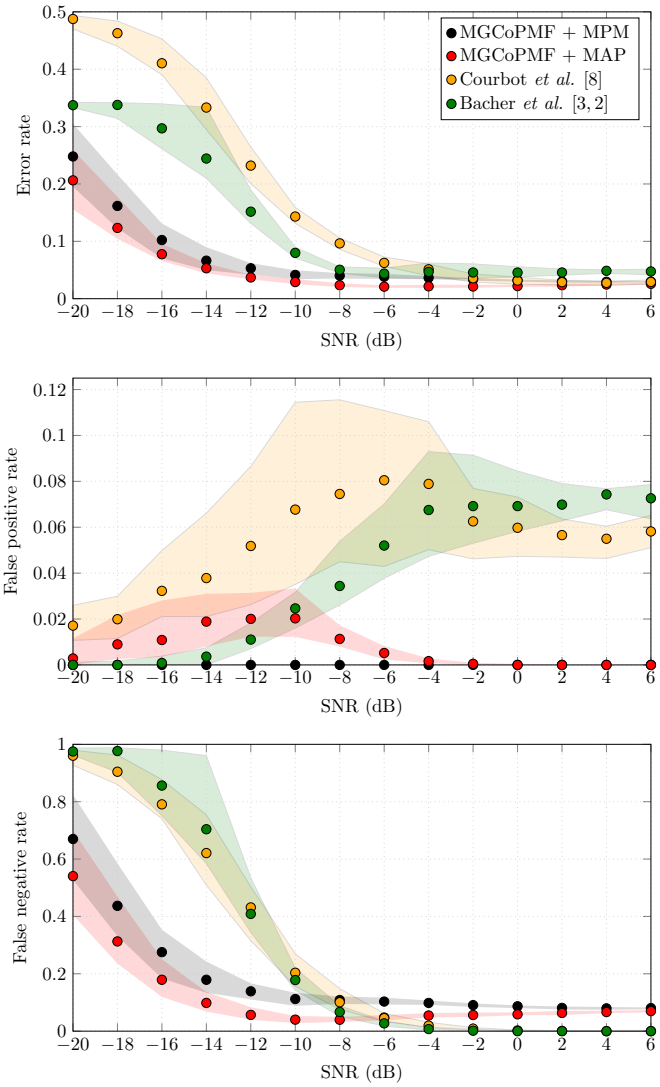


Figure 5: Results for the methods from [8], from [3, 2] and with the MAP (8) and the MPM (9) within the MGCoPMF introduced in this paper (the red and black curves are the same as in Fig. 4). The legend is the same as in Fig. 4.

of interest. Such HSI may be contaminated by brightest objects presenting a "continuum" emission. These contamination are efficiently removed by a spectral median filtering prior to the segmentation.

Fig. 6 illustrates the unsupervized segmentation of 4 MUSE HSI. In the absence of ground truth, the qualification of the results must be conducted thoroughly by experts. Nevertheless, two points are worth mentioning:

- in most cases, the MAP-based segmentation yields a larger detected region than its MPM counterpart. This is consistent with the results obtained on synthetic images;
- the inspection of the average spectra reveals that their shape is consistent with a possible Lyman-alpha emission. Besides, the spatial shapes make likely the presence of an extended emission wider than a FSF blur on a point source, bolstering the hypothesis of the detection of an extended Lyman-alpha emission.

6. Conclusion

In this paper, we introduced the CoPMF and the MGCoPMF model, allowing to handle the convolution in the Markovian modeling of images. This model was specified for the application to detection in astronomical hyperspectral images, and the unsupervized segmentation in this context was presented. Then, comparative results with state-of-the art method indicate that the introduced models outperform their state-of-the-art alternatives, with in addition complimentary results provided by the MAP and MPM estimators. We finally presented results on real MUSE images, showing that the both segmentation method are robust to a real-world use.

Acknowledgement

J.-B. Courbot acknowledges support from the ERC advanced grant 339659-MUSICOS. This work was funded in part by the

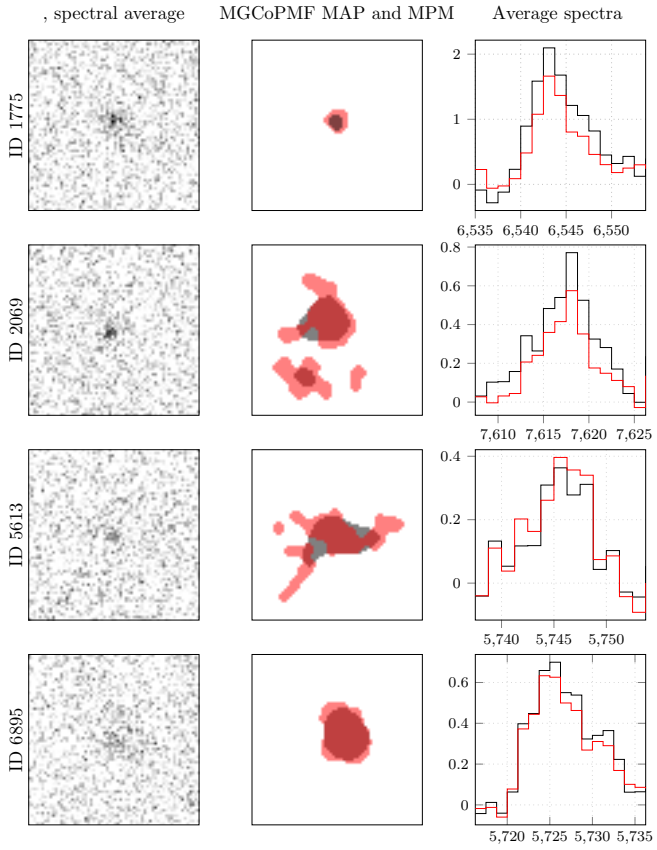


Figure 6: Results on 4 real MUSE images. Left: spectral average of $\mathbf{Y} = \mathbf{y}$ (inverse video) after the spectral median subtraction has removed the brightest objects. Center: detection results with the MGCoPMF model, using the MPM (gray) and MAP (red) criteria. Right: average spectra in the two detected regions, with intensities in 10^{-20}erg s^{-1}cm^{-2}Å^{-1} as a function of the wavelength (in Ångström).

DSIM project under grant ANR-14-CE27-0005. The authors would like to thank R. Bacon (CNRS – CRAL) for providing the MUSE data, and R. Bacher (CNRS – Gipsa-lab) for providing the source code for [3, 2] and for the fruitful discussions on the detection problem. They would also like to thank the anonymous reviewers for their helpful comments on the paper.

Appendix A. Maximum likelihood estimators for $\boldsymbol{m}$

In this appendix, we derive the maximum likelihood estimators (MLE) of $\boldsymbol{m}$, stated in (10) for the CoPMF and the MGCoPMF models introduced in this paper.

1) MLE for the CoPMF model.

Using Equations (1–3) and (6), and recalling that $\boldsymbol{\gamma}_s = w_s\boldsymbol{m}$, we get the log-likelihood:

$$\log \ell(\mathbf{x}, \mathbf{y}; \boldsymbol{m}, \boldsymbol{\Sigma}) = c - \frac{1}{2}\sum_{s\in\mathcal{S}}(\mathbf{y}_s - w_s\boldsymbol{m})^\top\boldsymbol{\Sigma}^{-1}(\mathbf{y}_s - w_s\boldsymbol{m}). \tag{A.1}$$

where c contains the terms that are not related to $\boldsymbol{m}$. The MLE of $\boldsymbol{m}$, noted $\hat{\boldsymbol{m}}$, verifies:

$$\left.\frac{\partial \log \ell(\mathbf{x}, \mathbf{y}, \boldsymbol{m}, \boldsymbol{\Sigma})}{\partial \boldsymbol{m}}\right|_{\boldsymbol{m}=\hat{\boldsymbol{m}}} = 0 \tag{A.2}$$

which can be rewritten using (A.1) as:

$$0 = -\frac{1}{2}\frac{\partial}{\partial \boldsymbol{m}}\left(\sum_{s\in\mathcal{S}} \mathbf{y}_s^\top\boldsymbol{\Sigma}^{-1}\mathbf{y}_s - 2w_s\boldsymbol{m}^\top\boldsymbol{\Sigma}^{-1}\mathbf{y}_s + w_s^2\boldsymbol{m}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{m}\right)\Bigg|_{\boldsymbol{m}=\hat{\boldsymbol{m}}}$$

$$0 = -\frac{1}{2}\sum_{s\in\mathcal{S}} -2w_s\boldsymbol{\Sigma}^{-1}\mathbf{y}_s + 2w_s^2\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{m}}. \tag{A.3}$$

Assuming that $\sum_{s\in\mathcal{S}} w_s^2 \neq 0$, a left-hand product by $\boldsymbol{\Sigma}$ yields:

$$\hat{\boldsymbol{m}} = \frac{\sum_{s\in\mathcal{S}} w_s\mathbf{y}_s}{\sum_{s\in\mathcal{S}} w_s^2}. \tag{A.4}$$

□

2) MLE for the MGCoPMF model.

The derivation for the MGCoPMF follows the same principle as for the CoPMF model. Because an additional term involving $\boldsymbol{m}$ is introduced in 7, and reusing Equations (1–3), we obtain the log-likelihood:

$$\log \ell(\mathbf{x}, \mathbf{y}; \boldsymbol{m}, \boldsymbol{\Sigma}) = c - \frac{1}{2}\sum_{s\in\mathcal{S}}(\mathbf{y}_s - w_s\boldsymbol{m})^\top\boldsymbol{\Sigma}^{-1}(\mathbf{y}_s - w_s\boldsymbol{m})$$
$$-\sum_{s\in\mathcal{S}}\frac{\sqrt{|N_s|}}{2}\left(w_s\boldsymbol{m} - \frac{1}{|N_s|}\sum_{s'\in N_s}\mathbf{y}_{s'}\right)^\top\boldsymbol{\Sigma}^{-1}\left(w_s\boldsymbol{m} - \frac{1}{|N_s|}\sum_{s'\in N_s}\mathbf{y}_{s'}\right) \tag{A.5}$$

Similarly to (A.3), setting the derived log-likelihood to 0 yields:

$$0 = -\frac{1}{2}\sum_{s\in\mathcal{S}} -2w_s\boldsymbol{\Sigma}^{-1}\mathbf{y}_s + 2w_s^2\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{m}}$$
$$-\sum_{s\in\mathcal{S}}\frac{\sqrt{|N_s|}}{2}\left(-\frac{2w_s}{|N_s|}\boldsymbol{\Sigma}^{-1}\sum_{s'\in N_s}\mathbf{y}_{s'} + 2w_s^2\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{m}}\right) \tag{A.6}$$

Considering $|N_s|$ as a constant and assuming that $\sum_{s\in\mathcal{S}} w_s^2 \neq 0$, a left-hand product by $\boldsymbol{\Sigma}$ yields:

$$\hat{\boldsymbol{m}} = \frac{\sum_{s\in\mathcal{S}} w_s\left(\mathbf{y}_s + \sqrt{|N_s|}\sum_{s'\in N_s}\mathbf{y}_{s'}\right)}{(1+\sqrt{|N_s|})\sum_{s\in\mathcal{S}} w_s^2}. \tag{A.7}$$

□

References

[1] Ayasso, H., Mohammad-Djafari, A., 2010. Joint NDT image restoration and segmentation using Gauss–Markov–Potts prior models and variational Bayesian computation. IEEE Transactions on Image Processing 19, 2265–2277.

[2] Bacher, R., Chatelain, F., Michel, O., 2017a. Global error control procedure for spatially structured targets, in: Signal Processing Conference (EUSIPCO), 2017 25th European, IEEE. pp. 206–210.

[3] Bacher, R., Meillier, C., Chatelain, F., Michel, O., 2017b. Robust Control of Varying Weak Hyperspectral Target Detection With Sparse Nonnegative Representation. IEEE Transactions on Signal Processing 65, 3538–3550.
[4] Bacon, R., Conseil, S., Mary, D., Brinchmann, J., Shepherd, M., Akhlaghi, M., Weilbacher, P., Piqueras, L., Wisotzki, L., Lagattuta, D., et al., 2017. The MUSE Hubble Ultra Deep Field Survey: I. Survey description, data reduction and source detection. Astronomy & Astrophysics .
[5] Besag, J., 1986. On the statistical analysis of dirty pictures. Journal of the Royal Statistical Society. Series B (Methodological) , 259–302.
[6] Casella, G., Berger, R.L., 2002. Statistical inference. volume 2. Duxbury Pacific Grove, CA.
[7] Celeux, G., Diebolt, J., 1992. A stochastic approximation type EM algorithm for the mixture problem. Stochastics: An International Journal of Probability and Stochastic Processes 41, 119–134.
[8] Courbot, J.B., Mazet, V., Monfrini, E., Collet, C., 2017. Extended faint source detection in astronomical hyperspectral images. Signal Processing 135, 274–283.
[9] Delignon, Y., Marzouki, A., Pieczynski, W., 1997. Estimation of generalized mixtures and its application in image segmentation. Image Processing, IEEE Transactions on 6, 1364–1375.
[10] Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the royal statistical society. Series B (methodological) , 1–38.
[11] Derin, H., Elliott, H., 1987. Modeling and segmentation of noisy and textured images using Gibbs random fields. IEEE Transactions on Pattern Analysis and Machine Intelligence , 39–55.
[12] Dias, J.G., Wedel, M., 2004. An empirical comparison of EM, SEM and MCMC performance for problematic Gaussian mixture likelihoods. Statistics and Computing 14, 323–332.
[13] Eches, O., Benediktsson, J.A., Dobigeon, N., Tourneret, J.Y., 2013. Adaptive Markov random fields for joint unmixing and segmentation of hyperspectral images. Image Processing, IEEE Transactions on 22, 5–16.
[14] Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. Pattern Analysis and Machine Intelligence, IEEE Transactions on , 721–741.
[15] Li, J., Bioucas-Dias, J.M., Plaza, A., 2012. Spectral–spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields. Geoscience and Remote Sensing, IEEE Transactions on 50, 809–823.
[16] Li, W., Prasad, S., Fowler, J.E., 2014. Hyperspectral image classification using Gaussian mixture models and Markov random fields. Geoscience and Remote Sensing Letters, IEEE 11, 153–157.
[17] Marroquin, J., Mitter, S., Poggio, T., 1987. Probabilistic solution of ill-posed problems in computational vision. Journal of the American Statistical association 82, 76–89.
[18] McLachlan, G., Krishnan, T., 2007. The EM algorithm and extensions. volume 382. John Wiley & Sons.
[19] Meillier, C., Chatelain, F., Michel, O., Ayasso, H., 2015. Nonparametric Bayesian extraction of object configurations in massive data. Signal Processing, IEEE Transactions on 63, 1911–1924.
[20] Monfrini, E., Pieczynski, W., 2005. Estimation de mélanges généralisés dans les arbres de Markov cachés, application à la segmentation des images de cartons d'orgue de barbarie. Traitement du Signal 22.
[21] Pieczynski, W., Tebbache, A.N., 2000. Pairwise Markov random fields and segmentation of textured images. Machine graphics and vision 9, 705–718.
[22] Rellier, G., Descombes, X., Falzon, F., Zerubia, J., 2004. Texture feature analysis using a Gauss-Markov model in hyperspectral image classification. Geoscience and Remote Sensing, IEEE Transactions on 42, 1543–1551.
[23] Salzenstein, F., Collet, C., 2006. Fuzzy Markov random fields versus chains for multispectral image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 28, 1753–1767.
[24] Schweizer, S.M., Moura, J.M., 2000. Hyperspectral imagery: Clutter adaptation in anomaly detection. Information Theory, IEEE Transactions on 46, 1855–1871.
[25] Vacar, C., Giovannelli, J.F., 2019. Unsupervised joint deconvolution and segmentation method for textured images: A Bayesian approach and an advanced sampling algorithm. URL: `http://giovannelli.free.fr/Papers/TextSegDCV.pdf`.
[26] Vollmer, B., Perret, B., Petremand, M., Lavigne, F., Collet, C., Van Driel, W., Bonnarel, F., Louys, M., Sabatini, S., MacArthur, L., 2013. Simultaneous multi-band detection of low surface brightness galaxies with Markovian modeling. The Astronomical Journal 145, 36.
[27] Xia, J., Chanussot, J., Du, P., He, X., 2015. Spectral–spatial classification for hyperspectral data using rotation forests with local feature extraction and Markov random fields. Geoscience and Remote Sensing, IEEE Transactions on 53, 2532–2546.
[28] Zhao, N., Basarab, A., Kouamé, D., Tourneret, J.Y., 2016. Joint segmentation and deconvolution of ultrasound images using a hierarchical bayesian model based on generalized gaussian priors. IEEE transactions on Image Processing 25, 3736–3750.