ORIENTED TRIPLET MARKOV FIELDS FOR HYPERSPECTRAL IMAGE SEGMENTATION

Jean-Baptiste Courbot[1,3]*, Emmanuel Monfrini*[2]*, Vincent Mazet*[1]*, Christophe Collet*[1]

[1] ICube, Université de Strasbourg - CNRS, 67412 Illkirch, France
e-mail: jb.courbot@unistra.fr
[2] SAMOVAR, Département CITI, CNRS, 91011 Évry, France
[3] Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, CRAL UMR5574, F-69230, Saint-Genis-Laval, France

ABSTRACT

Hyperspectral image processing benefits greatly from using spatial information. Markov field modeling is a well-known statistical model class for considering spatial relationships between sites of an image. Often, the model restricts to Hidden Markov Field, therefore cannot handle non-stationarities in the images. This paper presents a Triplet Markov Field model for hyperspectral image segmentation, allowing the joint retrieving of image classes and local orientations. Segmentation results on synthetic data validate the methods, and results on real astronomical data are presented.

***Index Terms*—** Triplet Markov Field, Bayesian Segmentation, Orientation Retrieving.

1. INTRODUCTION

Cosmological theory predicts the existence of a "cosmic web", that is weak gas concentration, linking galaxies through elongated structures [1]. These structures are assumed to trace galaxy formation. The Multi-Unit Spectroscopic Explorer (MUSE) instrument is designed to reveal such features in deep field hyperspectral images (HSI) [2]. The purpose of this paper is to study the relevance of Triplet Markov Fields (TMF) [3] for segmenting oriented, faint structures in HSI. Indeed, modeling local orientation of elongated structures can be performed within TMF and could improve the segmentation.

Priors used for HSI segmentation are generally either spectral or spatial. For instance, spectral priors are used in supervised anomaly detection [4]. On the other hand, spatial priors can translate into object-wise processing [5], morphology-based processing [6] or local priors, as in the case of Markov Fields. In this framework, numerous works were published for HSI processing, most of them using Hidden Markov Field (HMF) models [7–12]. Despite its common use, standard HMF modeling has significant lacks: especially, spatial noise correlation and non-stationarity are not modeled. Generalizing HMF, pairwise [13] and triplet [3] Markov field models handle these two aspects.

In addition, spatial priors can enhance directions, meaning that *oriented* object are searched for. To our knowledge, there exist only a few works reporting the search for orientations in HSI. Most of them are related to edge detection techniques, targeting for instance road network detection in remote sensing [14]. Varying local orientations in images can be seen as varying stationarities of the underlying stochastic processes. Hence, it is desirable to provide an accurate modeling in the context of HSI statistical segmentation.

This paper presents a TMF model for segmentation in HSI, recovering jointly classes of interest and local orientations in the image. The main objective is to enhance the segmentation of oriented structures in HSI. The novelties of the approach are both the use of TMF and the estimation of orientations in HSI.

This paper uses the following notations. $\mathcal{S}$ is the set of sites in the image. A set of random variables over the image is noted $\mathbf{A} = (A_s)_{s\in\mathcal{S}}$ and the corresponding realization set is $\boldsymbol{a} = (a_s)_{s\in\mathcal{S}}$. We will note $\mathbf{A}_{N_s}$ the set of random variables over the local neighborhood N_s of the site s, and $\mathbf{A}_{\mathcal{S}\setminus s}$ the random variable set over the image, except in the site s. In the remainder of this paper, Section 2 introduces the proposed model and corresponding methods, and Section 3 presents some experimental validations on HSI.

2. MODEL

2.1. Oriented Triplet Markov Fields

Triplet Markov Fields (TMF) [3] are Markov fields with three constituting processes: $\mathbf{T} = (\mathbf{Y}, \mathbf{X}, \mathbf{V})$. The process $\mathbf{Y}$ represents the observed HSI, containing the spectra $\mathbf{Y}_s \in \mathbb{R}^\Lambda$, with Λ bands per site s. The class process $\mathbf{X}$ takes its value from a finite set $\Omega_x = \{\omega_1, \dots \omega_K\}$. Similarly, the *auxiliary* process $\mathbf{V}$ takes values from $\Omega_v = \{\nu_1, \dots, \nu_L\}$.

TMF, while generalizing HMF modeling, still benefit from posterior Markovianity: $(\mathbf{X}, \mathbf{V})$ conditionally to $\mathbf{Y}$ is a

J.-B. Courbot acknowledges support from the ERC advanced grant 339659-MUSICOS. This work was funded in part by the DSIM project under grant ANR-14-CE27-0005.

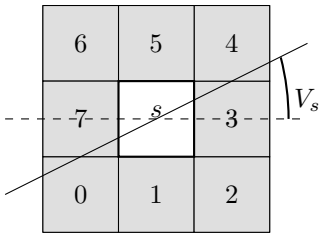


Fig. 1. Binary clique indexing and local orientation measurment within a local neighborhood.

Markov field [3]. This ensures that statistical segmentation is possible, as in HMF. The model verifies:

$$\begin{aligned} p(\mathbf{Y},\mathbf{X},\mathbf{V}) &\propto p(\mathbf{Y}|\mathbf{X},\mathbf{V})p(\mathbf{X},\mathbf{V}) \\ &\propto p(\mathbf{Y}|\mathbf{X},\mathbf{V})p(\mathbf{X}|\mathbf{V})p(\mathbf{V}). \end{aligned}$$

For Maximum Posterior Mode (MPM) segmentation [15], posterior probabilities must be computed on each site. In our model, $\mathbf{V}$ is a Markov field. Besides, $p(\mathbf{X}|\mathbf{V})$ is a Markov field distribution, in which $\mathbf{V}$ rules $\mathbf{X}$ in each site independently.

Then, for each site s on the lattice $\mathcal{S}$:

$$\begin{aligned} &p(\mathrm{X}_s,\mathrm{V}_s|\mathbf{Y},\mathbf{X}_{\mathcal{S}\setminus s},\mathbf{V}_{\mathcal{S}\setminus s}) \\ &\quad\propto p(\mathbf{Y}|\mathbf{X},\mathbf{V})p(\mathrm{X}_s|\mathbf{X}_{N_s},\mathbf{V})p(\mathrm{V}_s|\mathbf{V}_{N_s}) \\ &\quad\propto p(\mathbf{Y}|\mathbf{X},\mathbf{V})p(\mathrm{X}_s|\mathbf{X}_{N_s},\mathrm{V}_s)p(\mathrm{V}_s|\mathbf{V}_{N_s}). \end{aligned} \tag{1}$$

Conditionally to $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{V}$ are assumed independent. Furthermore, we assume that the $\mathbf{Y}_s$ are independent conditionally to $\mathbf{X}$. Finally, each $\mathbf{Y}_s$ depends only on the corresponding $\mathbf{X}_s$:

$$\begin{aligned} p(\mathbf{Y}|\mathbf{X},\mathbf{V}) &= p(\mathbf{Y}|\mathbf{X}) \\ &= \prod_{s\in\mathcal{S}} p(\mathbf{Y}_s|\mathbf{X}) = \prod_{s\in\mathcal{S}} p(\mathbf{Y}_s|\mathrm{X}_s). \end{aligned}$$

$\mathbf{V}$ relies on a Ising/Potts potential (as used in [16]), ruled by a positive, real-valued parameter α^v:

$$\forall s:\ p(\mathrm{V}_s|\mathbf{V}_{N_s}) \propto \exp\left[\alpha^v \sum_{s'\in N_s} 1-2\delta(\mathrm{V}_s,\mathrm{V}_{s'})\right]. \tag{2}$$

We restrict to binary cliques, covering the site s and a neighboring site s' within a 8-neighborhood.

Similarly, the distribution $p(\mathrm{X}_s|\mathbf{X}_{N_s},\mathrm{V}_s)$, parametrized by $\alpha^{x|v}\in\mathbb{R}^+$, is expressed as:

$$\begin{aligned} &\forall s:\ p(\mathrm{X}_s|\mathbf{X}_{N_s},\mathrm{V}_s) \\ &\quad\propto \exp\left[\alpha^{x|v}\sum_{s'\in N_s}\phi^{s'}(\mathrm{V}_s)\,(1-2\delta(\mathrm{X}_s,\mathrm{X}_{s'}))\right]; \end{aligned} \tag{3}$$

where $\phi^{s'}$ is an *orientation function*, defining the clique potential weights used for $\mathbf{X}$ in the case of Oriented TMF (OTMF)

Algorithm 1 SEM algorithm for OTMF

Input: Observed realization $\boldsymbol{y}$, stop criterion
Output: Parameter estimation Θ^{SEM}
Initialization: $\Theta^{\mathrm{noise}\ (0)}$, $\Theta^{\mathrm{field}\ (0)}$
while stop criterion not reached – iteration q **do**
Simulate $(\boldsymbol{x}^q,\boldsymbol{v}^q)$ from $p_{\Theta^{q-1}}(\mathbf{X},\mathbf{V}|\boldsymbol{y})$.
Estimate $\Theta^{\mathrm{noise}\ (q)}$ through MLE on $(\boldsymbol{x}^q,\boldsymbol{y})$.
Estimate $\Theta^{\mathrm{field}\ (q)}$ with DE, with $(\boldsymbol{x}^q,\boldsymbol{v}^q)$.
end while

accounting for the favored direction described in V_s. The orientation function is:

$$\phi^{s'}(\mathrm{V}_s) = \begin{cases} |\cos(\mathrm{V}_s)| & \text{if } s'\in\{3,7\} \\ |\cos(\mathrm{V}_s-\pi/4)| & \text{if } s'\in\{4,0\} \\ |\cos(\mathrm{V}_s-\pi/2)| & \text{if } s'\in\{5,1\} \\ |\cos(\mathrm{V}_s-3\pi/4)| & \text{if } s'\in\{6,2\} \end{cases};$$

where s' denotes the index within the neighborhood N_s (see Fig. 1).

To summarize, the distributions expressed in Eqs (2) and (3), together with the knowledge of the noise behavior, enable the Bayesian segmentation according to the MPM criterion [15]. $\forall s\in\mathcal{S}$:

$$(\hat{x}_s,\hat{v}_s)^{\mathrm{MPM}} = \underset{(\omega,\nu)\in\Omega_x\times\Omega_v}{\arg\max}\, p(X_s=\omega,V_s=v|\boldsymbol{y}). \tag{4}$$

There is no exact analytical expression for these estimations, however simulations are possible through Gibbs sampling [17]. Then, MPM estimates amount to choose for each site the most frequent Gibbs sampler realization.

2.2. Parameter Estimation

When performing unsupervised segmentation, the model parameters must be estimated. The parameter set Θ splits into field parameters, $\Theta^{\mathrm{fields}}=\{\alpha^v,\alpha^{x|v}\}$ and noise parameters, Θ^{noise}. From a realization $(\boldsymbol{x},\boldsymbol{v})$, the parameter set Θ^{fields} is estimated by the method from Derin and Elliott (DE) [16]. From complete data $(\boldsymbol{y},\boldsymbol{x},\boldsymbol{v})$, Θ^{noise} is obtained through the Maximum Likelihood Estimator (MLE). Complete data are however unavailable: this is why we use the Stochastic Expectation-Maximization (SEM) algorithm [18], whose principle is to mimic complete realization with iterative simulations. Let us note that it has been shown in [19][20] that SEM provides in practice slightly better results than the Expectation-Maximization (EM) [21] or the Iterative Conditional Estimator (ICE) [22]. The SEM algorithm is reported in Alg. 1.

The stop criterion is a similarity measure between consecutive estimations, to ensure that convergence is reached. Initial parameters are estimated through simple estimates: K-means clustering [23] provides an initial $\mathbf{X}$, and the initial $\mathbf{V}$ is obtained through image gradient measure on the initial $\mathbf{X}$.

3. EXPERIMENTS

3.1. Observation model

Detection aims at segmenting significant signal versus noise only. Therefore, Ω_x contains two labels $\omega_0 = 0$ and $\omega_1 = 1$, denoting signal absence and presence, respectively. Furthermore, the spectral noise behavior is assumed identical with respect to the considered band, and correlated over three adjacent bands. The spectral covariance matrix is written as:

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma^2 & \rho_1 & \rho_2 & 0 & \cdots & 0 \\ \rho_1 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \rho_2 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \rho_2 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \rho_1 \\ 0 & \cdots & 0 & \rho_2 & \rho_1 & \sigma^2 \end{bmatrix}$$

Assuming a zero-mean, additive Gaussian spectral noise implies that $\Theta^{\text{noise}} = \{\boldsymbol{\mu}, \sigma, \rho_1, \rho_2\}$. $\boldsymbol{\mu}$ is the spectrum to detect, σ stands for the noise standard deviation, and ρ_1, ρ_2 are the spectral noise correlation parameters. In this paper, we use two orthogonal directions for the auxiliary process: $\Omega_v = \{\pi/4, 3\pi/4\}$; to ensure a proper discrimination between orientations. Note that we restrict to a simple instance, but that generalizing to more subdivisions is straightforward.

3.2. Images

Three images are used for experiments:

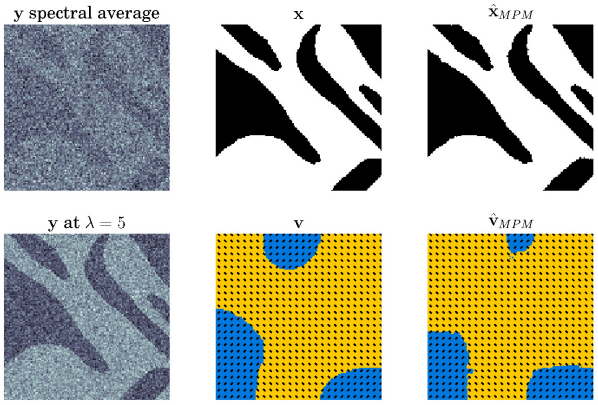
- an image generated through OTMF;
- a synthetic image generated conditionally to a fixed auxiliary process $\mathbf{V}$, which is not a Markov field realization;
- a real HSI from the MUSE instrument [2], obviously without $\mathbf{V}$ nor $\mathbf{X}$ ground truths for such data.

The synthetic images, of size 128×128 px, are presented in Figs. 2a and 2b. We use a weak average signal spectrum, presenting a few non-zero coefficient (see Fig. 2c) over $\Lambda = 20$ spectral bands. Let us remark that we consider that the images are *hyperspectral* in the sense that they correspond to very fine and close bandwidths. Two cases are investigated: $\text{SNR} = -5$ dB and $\text{SNR} = -10$ dB. The SNR is defined as:

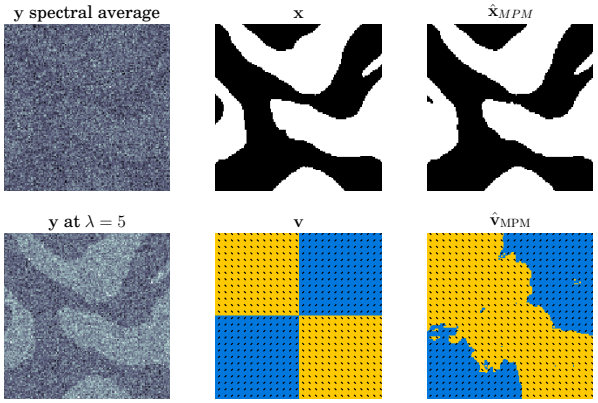
$$\text{SNR} = 10\log_{10}\left(\|\boldsymbol{\mu}\|_2^2/\text{Tr}(\mathbf{\Sigma})\right) = 10\log_{10}\left(\|\boldsymbol{\mu}\|_2^2/\Lambda\sigma^2\right).$$

Finally, the correlation coefficients were set to $\rho_1 = 0.5\sigma^2$ and $\rho_2 = 0.25\sigma^2$.

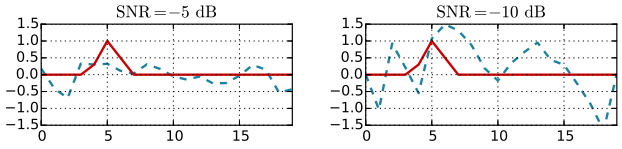
The MUSE HSI has 90×55 px over $\Lambda = 50$ spectral bands, and is illustrated in Fig. 3. Segmentation on this image should reveal faint, spectrally-sparse and spatially wide light emission (called the Lyman-alpha line) located in the vicinity of distant galaxies. In such deep field observation, these galaxies occupy no more than 5 to 10 pixels. Detection of an oriented structure in this image may be a clue for a filamentary inter-galactic structure.



(a) First synthetic image, at $\text{SNR} = -5$ dB.



(b) Second synthetic image, at $\text{SNR} = -10$ dB.



(c) Instances of noisy synthetic spectra (dashed) and $\boldsymbol{\mu}$ (plain).

Fig. 2. Synthetic images, spectra and the correponding segmentation results.

3.3. Results

Results on parameter estimation and segmentation error rate are reported in Table 1. Segmentation results are also shown in Fig. 2. The segmentation results are compared with a classical HMF modeling, using also SEM for parameter estimation and MPM for segmentation.

The estimated values for noise are very close to real parameters, with less than 1% difference in most cases. On the other hand, field parameters $\alpha^{x|v}$ and α^v are poorly estimated, with a strong variance (not shown). One can notice that this misestimation does not affect significantly segmentation result: in this model, their low influence on the posterior distribution from Eq. (1) makes it difficult to provide precise estimations (see *e.g* [24] for an instance of similar conclusions).

Table 1. Real parameters, SEM estimations and segmentation error rates.

			$\alpha^{x\|v}$	α^{v}	σ^2	ρ_1	ρ_2	$\boldsymbol{\mu}$ (MSE)	$E^{\mathrm{HMF}}_{\hat{\boldsymbol{x}}}$	$E^{\mathrm{OTMF}}_{\hat{\boldsymbol{x}}}$	$E^{\mathrm{OTMF}}_{\hat{\boldsymbol{v}}}$
Im1	-5 dB	Real	5.0	20.0	0.2119	0.1059	0.0530	2.10^{-4}	0.29%	**0.27%**	10.50%
		Est.	0.8194	0.8097	0.2131	0.1065	0.0532				
	-10 dB	Real	5.0	20.0	0.6700	0.3350	0.1675	9.10^{-4}	3.20%	**0.78%**	14.09%
		Est.	0.6543	0.9304	0.6683	0.3335	0.1657				
Im2	-5 dB	Real	8.0	–	0.2119	0.1059	0.0530	5.10^{-4}	0.40%	**0.27%**	17.41%
		Est.	0.9357	0.8641	0.2114	0.1058	0.0525				
	-10 dB	Real	8.0	–	0.6700	0.3350	0.1675	10.10^{-4}	3.67%	**0.95%**	22.31%
		Est.	0.7493	0.8843	0.6699	0.3351	0.1672				
Im3	–	Est.	1.2527	0.5878	0.5856	0.1227	0.0154	–	–	–	–

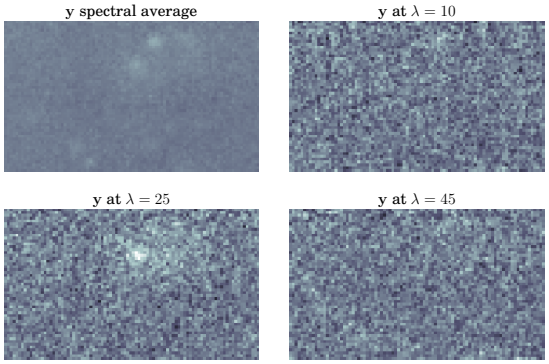
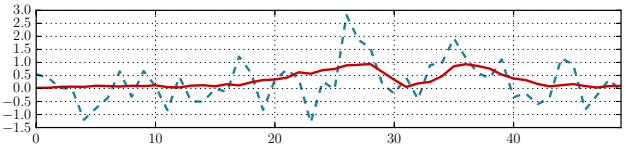


Fig. 3. Astronomical HSI from MUSE (Im3).



(a) Estimated average spectrum $\boldsymbol{\mu}$ (plain) and detected spectrum instance (dashed).

(b) MPM estimations for the $\mathbf{X}$ and $\mathbf{V}$ processes.

Fig. 4. Estimation results on the MUSE HSI.

The segmentation results on $\mathbf{X}$ (error rate: $E^{\mathrm{OTMF}}_{\hat{\boldsymbol{x}}}$) show the improvements of the considered model with respect to HMF modeling ($E^{\mathrm{HMF}}_{\hat{\boldsymbol{x}}}$), when processing oriented, thin structures. Indeed, the isotropic potential used in HMF may make the corresponding segmentation fail in such cases. On the other hand, the MPM estimation on $\mathbf{V}$ (error rate: $E^{\mathrm{OTMF}}_{\hat{\boldsymbol{v}}}$) is not as efficient as on $\mathbf{X}$. This is mainly because some cases (*e.g.* wide, homogeneous region) make the direction estimation hard to perform. Therefore, there is an intrinsic imprecision on the segmentation result. Let us also remark that the segmentation results are better on the first synthetic image than on the second, because the fixed $\mathbf{V}$ in the second case is not a Markov field realization.

Estimated parameters for the real MUSE HSI are also reported in Table 1, and the estimated signal spectrum is displayed in Fig. 4a. The SNR, estimated at -5.04 dB, is close to the SNR used in synthetic cases, which induces to believe that parameter estimation are of similar quality. One can especially notice that the correlation coefficients are significantly far from zero. This is due to MUSE data pre-processing, and hence such correlation must be considered when handling MUSE observations.

The segmentation results are reported in Fig. 4b. Ground truths are naturally not available. One can notice, however, that the detected region $\hat{\boldsymbol{x}}^{\mathrm{MPM}}$ is significantly larger than what is expected from a single galaxy, and that the wide detected emission cannot be due to the instrument point spread function alone, because of its asymmetry. Concerning the estimated orientations $\hat{\boldsymbol{v}}^{\mathrm{MPM}}$, the result is difficult to interpret: we notice, however, that the border of the detected region and the corresponding orientations are coherent. Let us finally remark that spatial noise correlation could induce mis-estimations, because the proposed model assigns such correlation on the hidden processes $\mathbf{X}$ and $\mathbf{V}$.

4. CONCLUSION

This paper presented a model for joint Bayesian segmentation of classes and orientations in an HSI. The method was validated in the case of signal detection on synthetic images, outperforming HMF, and applied to real astronomical HSI. Further work will include models and algorithms refinement, to handle more complex cases.

5. ACKNOWLEDGEMENT

The authors would like to thank Roland Bacon (France, CNRS–CRAL) for providing the MUSE data.

References

[1] Rien van de Weygaert and Willem Schaap, "The cosmic web: geometric analysis," in *Data Analysis in Cosmology*, pp. 291–413. Springer, 2008.

[2] Roland Bacon, Jarle Brinchmann, Johan Richard, Thierry Contini, Alyssa Drake, Marijn Franx, Sandro Tacchella, Joel Vernet, Lutz Wisotzki, Jeremy Blaizot, et al., "The MUSE 3D view of the Hubble Deep Field South," *Astronomy & Astrophysics*, vol. 575, pp. A75, 2015.

[3] Dalila Benboudjema and Wojciech Pieczynski, "Unsupervised image segmentation using triplet Markov fields," *Computer Vision and Image Understanding*, vol. 99, no. 3, pp. 476–498, 2005.

[4] Stefania Matteoli, Marco Diani, and Giovanni Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 25, no. 7, pp. 5–28, 2010.

[5] Céline Meillier, Florent Chatelain, Olivier Michel, and Hacheme Ayasso, "Nonparametric Bayesian extraction of object configurations in massive data," *Signal Processing, IEEE Transactions on*, vol. 63, no. 8, pp. 1911–1924, 2015.

[6] Mathieu Fauvel, Yuliya Tarabalka, Jon Atli Benediktsson, Jocelyn Chanussot, and James C Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652–675, 2013.

[7] Susan M Schweizer and José MF Moura, "Hyperspectral imagery: Clutter adaptation in anomaly detection," *Information Theory, IEEE Transactions on*, vol. 46, no. 5, pp. 1855–1871, 2000.

[8] Guillaume Rellier, Xavier Descombes, Frederic Falzon, and Josiane Zerubia, "Texture feature analysis using a Gauss-Markov model in hyperspectral image classification," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, no. 7, pp. 1543–1551, 2004.

[9] Jun Li, José M Bioucas-Dias, and Antonio Plaza, "Spectral–spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 50, no. 3, pp. 809–823, 2012.

[10] Olivier Eches, Jon Atli Benediktsson, Nicolas Dobigeon, and Jean-Yves Tourneret, "Adaptive Markov random fields for joint unmixing and segmentation of hyperspectral images," *Image Processing, IEEE Transactions on*, vol. 22, no. 1, pp. 5–16, 2013.

[11] Wei Li, Santasriya Prasad, and James E Fowler, "Hyperspectral image classification using Gaussian mixture models and Markov random fields," *Geoscience and Remote Sensing Letters, IEEE*, vol. 11, no. 1, pp. 153–157, 2014.

[12] Junshi Xia, Jocelyn Chanussot, Peijun Du, and Xiyan He, "Spectral–spatial classification for hyperspectral data using rotation forests with local feature extraction and Markov random fields," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 53, no. 5, pp. 2532–2546, 2015.

[13] Wojciech Pieczynski and Abdel-Nasser Tebbache, "Pairwise Markov random fields and segmentation of textured images," *Machine graphics and vision*, vol. 9, no. 3, pp. 705–718, 2000.

[14] Mehdi Kamandar and Hassan Ghassemian, "Linear feature extraction for hyperspectral images based on information theoretic learning," *Geoscience and Remote Sensing Letters, IEEE*, vol. 10, no. 4, pp. 702–706, 2013.

[15] Jose Marroquin, Sanjoy Mitter, and Tomaso Poggio, "Probabilistic solution of ill-posed problems in computational vision," *Journal of the American Statistical association*, vol. 82, no. 397, pp. 76–89, 1987.

[16] Haluk Derin and Howard Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, , no. 1, pp. 39–55, 1987.

[17] Stuart Geman and Donald Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, , no. 6, pp. 721–741, 1984.

[18] Gilles Celeux and Jean Diebolt, "A stochastic approximation type EM algorithm for the mixture problem," *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 41, no. 1-2, pp. 119–134, 1992.

[19] José G Dias and Michel Wedel, "An empirical comparison of EM, SEM and MCMC performance for problematic Gaussian mixture likelihoods," *Statistics and Computing*, vol. 14, no. 4, pp. 323–332, 2004.

[20] Emmanuel Monfrini and Wojciech Pieczynski, "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*, vol. 22, no. 2, 2005.

[21] Arthur P Dempster, Nan M Laird, and Donald B Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38, 1977.

[22] Yves Delignon, Abdelwaheb Marzouki, and Wojciech Pieczynski, "Estimation of generalized mixtures and its application in image segmentation," *Image Processing, IEEE Transactions on*, vol. 6, no. 10, pp. 1364–1375, 1997.

[23] James MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, 1967, vol. 1, pp. 281–297.

[24] Fabien Salzenstein and Wojciech Pieczynski, "Parameter estimation in hidden fuzzy Markov random fields and image segmentation," *Graphical models and image processing*, vol. 59, no. 4, pp. 205–220, 1997.