UNSUPERVISED UNMIXING OF HYPERSPECTRAL IMAGES ACCOUNTING FOR ENDMEMBER VARIABILITY

Abderrahim Halimi[(1)]*, Nicolas Dobigeon*[(1)]*, Jean-Yves Tourneret*[(1)] *and Paul Honeine*[(2)]

[(1)] University of Toulouse, IRIT/INP-ENSEEIHT/TéSA, Toulouse, France
[(2)] Institut Charles Delaunay (CNRS), Université de technologie de Troyes, Troyes, France.

ABSTRACT

This paper presents an unsupervised Bayesian algorithm for hyperspectral image unmixing accounting for endmember variability. The pixels are modeled by a linear combination of random endmembers, to take into account their variability in the image, weighted by their corresponding abundances. An additive noise is also considered in the proposed model generalizing the normal compositional model. The proposed Bayesian algorithm exploits the whole image to provide spectral and spatial information. It estimates both the mean and the covariance matrix of each endmember in the image. A spatial segmentation is also obtained based on the estimated abundances. Simulations conducted with synthetic and real data show the potential of the proposed model and unmixing strategy for the analysis of hyperspectral images.

Index Terms— Hyperspectral imagery, endmember variability, image classification, Hamiltonian Monte-Carlo.

1. INTRODUCTION

Spectral unmixing (SU) consists of decomposing the pixel spectrum to recover physical materials, known as *endmembers*, and estimating the corresponding proportions or *abundances* [1]. The variation of endmember spectra along the image has been identified as one of the most profound sources of error in abundance estimation [2, 3]. This variation is called *spectral endmember variability* (SEV) and is knowing growing interest in the hyperspectral community [2, 3]. Many algorithms have been proposed in the literature to describe SEV by considering each endmember as a finite set of spectra or as a random vector with a given statistical distribution. The main contribution of this paper is the consideration of SEV under a statistical Bayesian framework. Two main approaches have been considered in the literature assuming the endmembers are random vectors: the Beta compositional model [4] and the normal compositional model (NCM) [5–7]. This paper proposes a generalization of the NCM model characterized by Gaussian variability for the endmembers (as for the NCM) and an additive Gaussian noise modeling fitting errors (which was not present in the NCM). Moreover, the proposed model considers a different mean and covariance matrix for each endmember to analyze each component separately. These parameters are both estimated to generalize the works of [7] and [6] that estimated only the endmembers means or the covariances, respectively. This paper also considers spatial correlation between pixels using Markov random fields (MRFs) as already done in [8, 9]. The image is segmented into regions sharing similar abundance characteristics which improves the unmixing quality [8, 9].

An unsupervised Bayesian algorithm is proposed to estimate the parameters associated with endmembers and abundances. In addition to the abundance Dirichlet priors, it assumes appropriate prior for the remaining parameters/hyperparameters to satisfy the known physical constraints. However, the classical Bayesian estimators cannot be easily computed from the resulting joint posterior. This problem is alleviated by considering a Markov chain Monte Carlo (MCMC) method that generates sample according to the posterior of interest. The sampling is achieved in this paper using a Gibbs sampler coupled with a Hamiltonian Monte Carlo (HMC) method. HMC is well adapted for problems with a large number of parameters to be estimated [10]. More precisely, this paper considers a constrained-HMC (CHMC) that has been introduced in [10, Chap. 5] and successfully used for hyperspectral SU in [11].

The paper is structured as follows. The proposed mixing model is introduced in Section 2. The different components of the proposed Bayesian model are studied in Section 3. Section 4 introduces the Gibbs sampler and the CHMC method which will be used to generate samples asymptotically distributed according to the joint posterior of the unknown parameters and hyperparameters. Section 5 analyzes the performance of the proposed algorithm when applied to synthetic images. Results on real hyperspectral images are presented in Section 6 whereas conclusions and future works are reported in Section 7.

2. PROPOSED MIXING MODEL

Let N be the number of pixels of the observed hyperspectral image. Each pixel $\boldsymbol{y}_n$, of size $(L \times 1)$, is observed by L spectral bands. The classical LMM assumes the pixel spectrum $\boldsymbol{y}_n$ is a linear combination of R deterministic endmembers $\boldsymbol{s}_r$, $r \in \{1, \cdots, R\}$, corrupted by an independent additive

noise $\boldsymbol{e}_n$ as follows

$$\boldsymbol{y}_n = \sum_{r=1}^{R} a_{rn}\boldsymbol{s}_r + \boldsymbol{e}_n \tag{1}$$

with $\boldsymbol{e}_n \sim \mathcal{N}\left(\mathbf{0}_L, \psi_n^2 \mathbf{I}_L\right)$, $\mathbf{0}_L$ is an $(L \times 1)$ vector of 0, $\mathbf{I}_L$ is the $(L \times L)$ identity matrix, $\boldsymbol{a}_n = [a_{1n}, \cdots, a_{Rn}]^T$ is the $(R \times 1)$ abundance vector of the nth pixel.

As previously mentioned, the endmembers are generally variable in the observed image [3]. In this paper, we introduce a model taking into account this variability. The proposed model can be seen as a generalization of the NCM model (GNCM) since it introduces an additional residual Gaussian noise $\boldsymbol{e}$ as follows

$$\boldsymbol{y}_n = \sum_{r=1}^{R} a_{rn}\boldsymbol{s}_{rn} + \boldsymbol{e}_n = \boldsymbol{S}_n\boldsymbol{a}_n + \boldsymbol{e}_n \tag{2}$$

where $\boldsymbol{s}_{rn} \sim \mathcal{N}\left(\boldsymbol{m}_r, \text{diag}\left(\boldsymbol{\sigma}_r^2\right)\right)$, $\boldsymbol{S}_n = [\boldsymbol{s}_{1n}, \cdots, \boldsymbol{s}_{Rn}]$, $\boldsymbol{\sigma}_r^2 = \left[\sigma_{r1}^2, \cdots, \sigma_{rL}^2\right]$ is the variance vector of the rth endmember and $\boldsymbol{M} = [\boldsymbol{m}_1, \cdots, \boldsymbol{m}_R]$ is the $(L \times R)$ matrix containing the endmember means of the image. The main difference between model (2) and the LMM is that the endmember matrix $\boldsymbol{S}_n$ depends on each observed pixel in order to introduce the SEV. Each physical element is then represented in a given pixel by an endmember $\boldsymbol{s}_{rn}$ that has its own Gaussian distribution whose variances $\boldsymbol{\sigma}_r^2$ change from one band to another. This allows the GNCM to capture the spectral variations of each physical element with respect to each spectral band. The additional Gaussian noise $\boldsymbol{e}_n$ makes the proposed model more robust with respect to mismodeling. Note finally that the proposed model reduces to the NCM for $\psi_n^2 = 0, \ \forall n$. Thus, it generalizes the model of [6, 7] by considering a non-isotropic covariance matrix for each endmember. Finally, for both LMM and GNCM, the abundance vector $\boldsymbol{a}_n$ represents proportions and should satisfy the physical positivity and sum-to-one (PSTO) constraints $a_{rn} \geq 0, \forall r \in \{1, \ldots, R\}$ and $\sum_{r=1}^{R} a_{rn} = 1$.

3. HIERARCHICAL BAYESIAN MODEL

This section introduces a hierarchical Bayesian model for the proposed GNCM model. The unknown parameters of this model include the endmember mean matrix $\boldsymbol{M}$, the $(R \times L)$ matrix $\boldsymbol{\Sigma}$ gathering the endmember variances (with $\boldsymbol{\Sigma}_{r,l} = \sigma_{rl}^2$), the $R \times N$ abundance matrix $\boldsymbol{A} = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_N]$, the $(1 \times N)$ label vector $\boldsymbol{z}$ and the $(1 \times N)$ vector $\boldsymbol{\Psi}$ containing the noise variances (with $\boldsymbol{\Psi}_n = \psi_n^2$).

3.1. Likelihood

Using the observation model (2), the Gaussian properties of both the noise sequence $\boldsymbol{e}_n$ and the endmembers, and exploiting independence between the observations in different spectral bands, yield

$$f(\boldsymbol{y}_n|\boldsymbol{A}, \boldsymbol{M}, \boldsymbol{\Sigma}, \boldsymbol{z}, \boldsymbol{\Psi}) \propto \left(\frac{1}{\prod_{l=1}^{L} \boldsymbol{\Omega}_{ln}}\right)^{\frac{1}{2}} \times \exp\left\{-\frac{1}{2}\boldsymbol{\Lambda}_{:n}^T \left[(\boldsymbol{y}_n - \boldsymbol{M}\boldsymbol{a}_n) \odot (\boldsymbol{y}_n - \boldsymbol{M}\boldsymbol{a}_n)\right]\right\} \tag{3}$$

where $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^T (\boldsymbol{A} \odot \boldsymbol{A}) + \boldsymbol{K}$ is an $(L \times N)$ matrix, $\boldsymbol{K} = \mathbf{1}_L \otimes \boldsymbol{\Psi}$ is an $(L \times N)$ matrix whose rows are equal to $\boldsymbol{\Psi}$, $\mathbf{1}_L$ is an $(L \times 1)$ vector of 1, $\boldsymbol{\Lambda}$ is an $(L \times N)$ matrix with $\boldsymbol{\Lambda}_{ln} = \frac{1}{\boldsymbol{\Omega}_{ln}}$, $\odot$ denotes the Hadamard (termwise) product and $\otimes$ denotes the Kronecker product. Moreover, contrary to the LMM, Eq. (3) shows that the elements[1] of $\boldsymbol{\Omega}$ depend jointly on the pixel abundances and on the pixel index $\#n$. This property was also satisfied by the NCM model as previously shown in [6, 7]. Note finally that by assuming independence between the observed pixels, the joint likelihood of the observation matrix $f(\boldsymbol{Y}|\boldsymbol{T}, \boldsymbol{M}, \boldsymbol{\Sigma}, \boldsymbol{z}, \boldsymbol{\Psi})$ can be expressed as the product of the N likelihoods shown in (3).

3.2. Parameter/hyperparameter priors

This section introduces the prior distributions that we have chosen for the parameters of interest $\boldsymbol{z}$, $\boldsymbol{A}$, $\boldsymbol{M}$, $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$.

3.2.1. Classification prior modeling

This paper considers the spatial correlation between the image pixels by dividing the observed image into K classes sharing the same abundance properties [9]. Each pixel is assigned to a specific class by using a latent label variable z_n. A Markov random field prior is then assigned for z_n as follows $f\left(z_n|\boldsymbol{z}_{\backslash n}\right) = f\left(z_n|\boldsymbol{z}_{\nu(n)}\right)$, where $\nu(n)$ denotes the pixel neighborhood (a four neighborhood structure will be considered in this paper), $\boldsymbol{z}_{\nu(n)} = \{z_i, i \in \nu(n)\}$ and $\boldsymbol{z}_{\backslash n} = \{z_i, i \neq n\}$. As in [8, 9, 12], this paper considers a Potts-Markov model which is appropriate for hyperspectral image segmentation. The prior of $\boldsymbol{z}$ is obtained using the Hammersley-Clifford theorem, $f(\boldsymbol{z}) = \frac{1}{G(\beta)} \exp\left[\beta \sum_{n=1}^{N} \sum_{n' \in \nu(n)} \delta(z_n - z_{n'})\right]$, where $\beta > 0$ is the granularity coefficient, $G(\beta)$ is a normalizing (or partition) constant and $\delta(.)$ is the Dirac delta function. The parameter β controls the degree of homogeneity of each region in the image. It is assumed known a priori in this paper. However, it could be also included within the Bayesian model and estimated using the strategy described in [13].

3.2.2. Abundance matrix $\boldsymbol{A}$

A Dirichlet prior is assigned to the abundances to satisfy the physical PSTO constraints [14]. Each spatial class k is assigned a different Dirichlet parameters $\boldsymbol{c}_k = (c_{1k}, \cdots, c_{Rk})^T$

[1] The matrix $\boldsymbol{\Omega}$ gathers the noise and endmember variances.

as follows

$$\boldsymbol{a}_n|z_n = k, \boldsymbol{c}_k \sim \text{Dir}(\boldsymbol{c}_k), \text{for } n \in \mathcal{I}_k \tag{4}$$

where Dir(.) denotes the Dirichlet distribution, and $n \in \mathcal{I}_k$ means that $\boldsymbol{y}_n$ belongs to the kth class (which is also equivalent to $z_n = k$). This prior allows the data to be located in several different clusters inside the simplex [14]. Moreover, to simplify the sampling procedure, we use the reparametrization from $\boldsymbol{a}_n$ to the $(R-1) \times 1$ vector $\boldsymbol{t}_n$ introduced in [15]. This reparametrization expresses the physical PSTO constraints by only using nonnegativity constraint $\boldsymbol{t}_n \in [0,1]^{R-1}$, which is easily handled by the sampling procedure (see [11, 15] for more details about this reparametrization). Note finally that assigning a Dirichlet prior for $\boldsymbol{a}_n$ corresponds to a beta distribution prior for the independent coefficients t_{rn}, $r \in \{1, \cdots, R-1\}$ and that $f(\boldsymbol{t}_n|z_n = k, \boldsymbol{c}_k) = \prod_{r=1}^{R-1} f(t_{rn}|z_n = k, \boldsymbol{c}_k)$.

3.2.3. Endmember means

The endmember mean matrix $\boldsymbol{M}$ contains reflectances that should satisfy the following constraints $0 < \boldsymbol{m}_{rl} < 1, \forall r \in \{1, \cdots, R\}, \forall l \in \{1, \cdots, L\}$ [11]. Therefore, we choose a truncated Gaussian prior for each endmember [7, 11] as follows $\boldsymbol{m}_r \sim \mathcal{N}_{[0,1]^L}\left(\widetilde{\boldsymbol{m}}_r, \epsilon^2 \mathbb{I}_l\right)$, where $\widetilde{\boldsymbol{m}}_r$ denotes an estimated endmember (resulting from an endmember extraction algorithm such as VCA [16]) and ϵ^2 is a variance term defining the confidence that we have on this estimated endmember $\widetilde{\boldsymbol{m}}_r$.

3.2.4. Endmember variances

The absence of knowledge about the endmember variances can be considered by choosing a Jeffreys distribution for the parameters σ_{rl}^2, i.e., $f(\boldsymbol{\Sigma}_{:l}) \propto \prod_{r=1}^{R} \frac{1}{\sigma_{rl}^2} \mathbf{1}_{\mathbb{R}+}\left(\sigma_{rl}^2\right)$,where we have assumed prior independence between the endmember variances.

3.2.5. Noise variance prior

To ensure sparsity, the noise variance is assigned an exponential prior $f\left(\psi_n^2|\lambda\right) = \lambda \exp\left(-\lambda \psi_n^2\right) \mathbf{1}_{\mathbb{R}+}\left(\psi_n^2\right)$, where λ is a large coefficient imposing sparsity for ψ_n ($\lambda = 10^7$ in our simulations). We furthermore assume prior independence between the random variables $\psi_n^2, \forall n \in \{1, \cdots, N\}$.

3.2.6. Dirichlet parameters

The Dirichlet parameters $\boldsymbol{c}_k$ are assigned the conjugate prior described in [17]. The parameter of this prior have been chosen to ensure a non-informative prior (flat distribution) [18].

3.3. Posterior distribution

The parameters of the proposed Bayesian model are included in the vector $\boldsymbol{\theta} = \{\boldsymbol{\theta}_p, \boldsymbol{\theta}_h\}$ where $\boldsymbol{\theta}_p = \{\boldsymbol{A}, \boldsymbol{M}, \boldsymbol{\Sigma}, \boldsymbol{z}, \boldsymbol{\Psi}\}$ (parameters) and $\boldsymbol{\theta}_h = \{\boldsymbol{C}\}$ (hyperparameters).

The joint posterior distribution of the unknown parameter/hyperparameter vector $\boldsymbol{\theta}$ can be computed from the following hierarchical structure

$$f\left(\boldsymbol{\theta}_p, \boldsymbol{\theta}_h|\boldsymbol{Y}\right) \propto f\left(\boldsymbol{Y}|\boldsymbol{\theta}_p, \boldsymbol{\theta}_h\right) f\left(\boldsymbol{\theta}_p, \boldsymbol{\theta}_h\right) \tag{5}$$

where $f\left(\boldsymbol{Y}|\boldsymbol{\theta}_p, \boldsymbol{\theta}_h\right) = f\left(\boldsymbol{Y}|\boldsymbol{\theta}_p\right)$ and $f\left(\boldsymbol{\theta}_p, \boldsymbol{\theta}_h\right) = f\left(\boldsymbol{T}|\boldsymbol{C}\right)$ $f\left(\boldsymbol{M}\right) f\left(\boldsymbol{\Sigma}\right) f\left(\boldsymbol{z}\right) f\left(\boldsymbol{\Psi}\right) f\left(\boldsymbol{C}\right)$, where we have assumed prior independence between the parameters.

Unfortunately, it is difficult to obtain closed form expressions for the standard Bayesian estimators associated with (5). In this paper, we propose to use MCMC methods to generate samples asymptotically distributed according to (5) and to build estimators of $\boldsymbol{\theta}$ from these generated samples. Due to the large number of parameters to be sampled, we use a CHMC algorithm which improves the mixing properties of the sampler [10]. The parameters are finally estimated using the minimum mean square error (MMSE) estimator for $\{\boldsymbol{A}, \boldsymbol{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi}, \boldsymbol{C}\}$ and the maximum a posteriori (MAP) estimator for the labels $\boldsymbol{z}$. The next section defines the proposed sampling procedure based on a hybrid Gibbs sampler including a CHMC method.

4. HYBRID GIBBS ALGORITHM

The principle of the Gibbs sampler is to generate samples according to the conditional distributions of the target distribution $f\left(\boldsymbol{\theta}_p, \boldsymbol{\theta}_h|\boldsymbol{Y}\right)$ [19]. In this paper, we propose to sample sequentially the abundance matrix $\boldsymbol{A}$, the mean endmember matrix $\boldsymbol{M}$, the variance of endmembers $\boldsymbol{\Sigma}$, the labels $\boldsymbol{z}$, the noise variances $\boldsymbol{\Psi}$ and the Dirichlet parameters $\boldsymbol{C}$.

The abundance matrix $\boldsymbol{A}$ or equivalently $\boldsymbol{T}$ It can be shown that the N vectors $\boldsymbol{t}_n, n \in \{1, \cdots, N\}$ are a posteriori independent in each spatial class k. Moreover, using the likelihood (3) and the prior distribution $f(\boldsymbol{t}_n|z_n = k, \boldsymbol{c}_k)$ provides the conditional distribution for $\boldsymbol{t}_n$. The CHMC framework is used for sampling the independent vectors $\boldsymbol{t}_n, n \in \{1, \cdots, N\}$ in an effective parallel procedure that reduces the computational cost (see [18] for more details about the CHMC procedure).

The endmember mean $\boldsymbol{M}$ It can be shown from (5) that the $\boldsymbol{M}$ rows ($\boldsymbol{M}_{\ell:}$) are a posteriori independent. Moreover, straightforward computations shows that $f\left(\boldsymbol{M}_{l:}|\boldsymbol{Y}_{l:}, \boldsymbol{T}, \boldsymbol{\Sigma}_{:l}\right)$ is a truncated Gaussian distribution that is easily sampled using a parallel CHMC procedure (see [18] for more details).

The endmember variance $\boldsymbol{\Sigma}$ The conditional distribution of $\boldsymbol{\Sigma}_{:\ell}$ is proportional to the product of the joint likelihood $f(\boldsymbol{Y}|\boldsymbol{T}, \boldsymbol{M}, \boldsymbol{\Sigma}, \boldsymbol{z}, \boldsymbol{\Psi})$ with the Jeffreys distribution described in section 3.2.4. The resulting distribution has a complex form and is sampled using a parallel HMC algorithm.

The labels $\boldsymbol{z}$ Sampling z_n from its conditional distribution can be performed by drawing a discrete value in the finite set $\{1, \cdots, K\}$ with known probabilities.

The noise variance $\boldsymbol{\Psi}$ The conditional distribution of ψ_n is proportional to the product of the likelihood (3) with the

exponential distribution described in section 3.2.5. The resulting distribution has a complex form and is sampled using a parallel HMC algorithm.

The Dirichlet coefficients $\boldsymbol{c}_k$ The conditional distribution of $\boldsymbol{c}_k|\boldsymbol{T}, \boldsymbol{z}_{n\in\mathcal{I}_k}$ has a complex form and is also sampled using an HMC procedure.

5. SIMULATION RESULTS ON SYNTHETIC DATA

This section considers a 50×50 synthetic image generated according to (2) with $R = 3$ endmember means (construction concrete, green grass and micaceous loam) that have been extracted from the ENVI software library [20]. Each endmember has its own distribution whose variance changes with respect to spectral bands. This image contains $K = 3$ classes whose label maps have been generated using the Potts-Markov prior with $\beta = 1.5$. The abundances of each class share the same Dirichlet parameters as previously explained. Note that the generated abundances have been truncated ($a_r < 0.9, \forall r$) to avoid the presence of pure pixels in the image. Finally, we have considered a noise variance equal to 10^{-7}. The proposed unsupervised GNCM-based algorithm, denoted by UsGNCM, has been run using $N_{\text{bi}} = 11000$ burn-in iterations and $N_{\text{MC}} = 12000$ iterations. Our algorithm is compared with state of the art algorithms: (i) VCA+FCLS: the endmembers are extracted from the whole image using VCA [16] and the abundances are estimated using the FCLS algorithm [21], (ii) UsLMM: the unsupervised Bayesian algorithm of [22] is used to jointly estimate the endmembers and abundances, (iii) AEB: this is the automated endmember bundles algorithm proposed in [23] (used with 10% image subset and the VCA algorithm), and (iv) UsNCM: the proposed unmixing strategy with $\psi_n = 0$ (i.e., the additive noise $\boldsymbol{e}_n$ of (2) is removed). In this case, the resulting algorithm reduces to the NCM model.

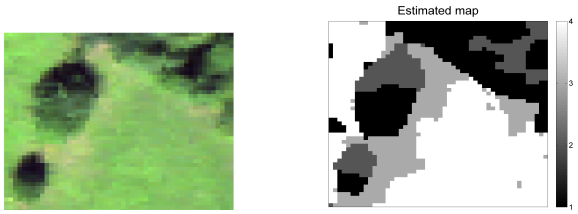
Table 1 reports the quality of the estimated abundances and endmembers when considering the averaged root mean square error (aRMSE) and the averaged spectral angle mapper (aSAM) criteria [18]. This table shows bad performance for VCA+FCLS and AEB algorithms which is mainly due to the absence of pure pixels in the considered images. The UsLMM provides good results, however, it appears to be sensitive to the variation of endmember/noise variances with respect to the spectral band and to the spatial correlations between adjacent pixels. Indeed, the UsLMM did not consider spatial correlation which leads to a performance reduction. Note also that the UsLMM algorithm provides one estimate for each endmember and does not take into account the spatial variability of endmembers in the processed image. The best performance is obtained by the proposed UsNCM and UsGNCM strategies that provide almost similar results. However, the UsGNCM shows better results thanks to the consideration of additive noise. These results confirm the superiority of the proposed approach in presence of SEV, spatial correlation between pixels and in absence of pure pixels in the observed scene.

Table 1. Results on synthetic data.

	Criteria ($\times 10^{-2}$)		
	aRMSE ($\boldsymbol{A}$)	aRMSE ($\boldsymbol{M}$)	aSAM ($\boldsymbol{M}$)
VCA+FCLS	3.71	2.68	6.74
UsLMM	0.76	0.49	0.94
AEB	9.46	4.20	8.72
UsNCM	0.56	0.19	0.43
UsGNCM	**0.48**	**0.16**	**0.41**

6. SIMULATION RESULTS ON REAL DATA

This section illustrates the performance of the proposed UsGNCM algorithm when applied to a real hyperspectral data set. The considered real image was acquired in 2010 by the Hyspex hyperspectral scanner over Villelongue, France. The dataset contains $L = 160$ spectral bands, 50×50 pixels and $R = 4$ components: tree, grass, soil and shadow (see Fig. 1 (left)).



(a) Madonna image. (b) Classification map.

Fig. 1. Real Madonna image and the estimated classification map using UsGNCM.

The estimated abundances using the UsGNCM algorithm are in good agreement with the FCLS and UsLMM results. These results are not presented here for brevity (see [18] for more details). In addition to unmixing, UsGNCM also provides a spatial segmentation of the considered scenes as shown in Fig. 1 (right). These classifications clearly highlight the area of each physical element in the scene. UsGNCM estimates both the mean and variance of each physical element in the scene which provides a measure the SEV in the considered image. Fig. 2 shows the estimated endmember distributions as blue level areas for each endmember. These distributions are in good agreement with the point estimates of VCA and UsLMM algorithms.

7. CONCLUSIONS

This paper introduced a Bayesian model for unsupervised unmixing of hyperspectral images accounting for SEV. The proposed algorithm was based on a generalization of the NCM

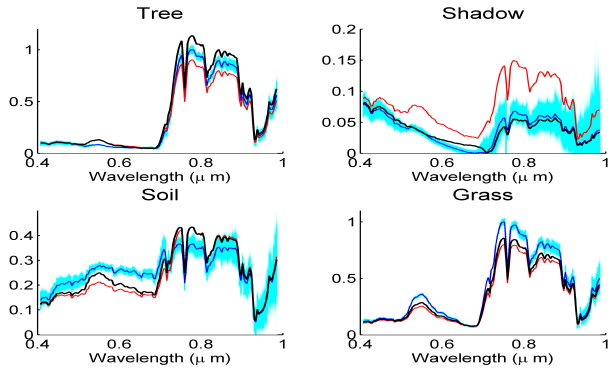


Fig. 2. The $R = 4$ endmembers estimated by VCA (continuous red lines), UsLMM (continuous black lines), UsGNCM (continuous blue lines) and the estimated endmember distribution (blue level areas) for the Madonna image.

and includes an additive Gaussian noise for modeling errors. This algorithm estimated the endmembers of the scene, their variabilities provided by their variances and the corresponding abundances. The observed image was also spatially segmented into regions sharing homogeneous abundance characteristics. The physical constraints of the abundances were ensured by choosing a Dirichlet distribution for each spatial class of the image. Due to the complexity of the resulting joint posterior distribution, an MCMC procedure based on a Gibbs algorithm was used to sample the posterior of interest and to approximate the Bayesian estimators of the unknown parameters using the generated samples. The sampling was achieved using an HMC method which is well suited for problems with a large number of parameters. The proposed algorithm showed good performance when processing data presenting SEV, spatial correlation between pixels and in absence of pure pixels in the observed scene. Future work includes the study of SEV with nonlinear mixing models. This point is an interesting issue that is currently under investigation.

8. REFERENCES

[1] J.M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Qian Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 5, no. 2, pp. 354–379, April 2012.

[2] B. Somers, G. P. Asner, L. Tits, and P. Coppin, "Endmember variability in spectral mixture analysis: A review," *Remote Sensing of Environment*, vol. 115, no. 7, pp. 1603 – 1616, 2011.

[3] A. Zare and K.C. Ho, "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 95–104, Jan 2014.

[4] A. Zare, P. Gader, D. Drashnikov, and T. Glenn, "Beta compositional model for hyperspectral unmixing," in *Proc. IEEE GRSS Workshop on Hyperspectral Image and SIgnal Processing: Evolution in Remote Sensing (WHISPERS)*, Gainesville, USA, Jun. 2013.

[5] D. Stein, "Application of the normal compositional model to the analysis of hyperspectral imagery," in *Advances in Techniques for Analysis of Remotely Sensed Data, 2003 IEEE Workshop on*, Oct 2003, pp. 44–51.

[6] O. Eches, N. Dobigeon, C. Mailhes, and J.-Y. Tourneret, "Bayesian estimation of linear mixtures using the normal compositional model. Application to hyperspectral imagery," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1403–1413, June 2010.

[7] A. Zare, P. Gader, and G. Casella, "Sampling piecewise convex unmixing and endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 3, pp. 1655–1665, March 2013.

[8] N. Bali and A. Mohammad-Djafari, "Bayesian approach with hidden markov modeling and mean field approximation for hyperspectral data analysis," *IEEE Trans. Image Process.*, vol. 17, no. 2, pp. 217–225, Feb. 2008.

[9] O. Eches, N. Dobigeon, and J.-Y. Tourneret, "Enhancing hyperspectral image unmixing with spatial correlations," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, Nov 2011.

[10] S. Brooks, A. Gelman, G. L . Jones, and X.-L. Meng, *Handbook of Markov chain Monte Carlo*, ser. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. Taylor & Francis, 2011.

[11] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, "Unsupervised post-nonlinear unmixing of hyperspectral images using a Hamiltonian Monte Carlo algorithm," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2663–2675, June 2014.

[12] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, "Residual component analysis of hyperspectral images: Application to joint nonlinear unmixing and nonlinearity detection," *IEEE Trans. Image Process.*, vol. 23, no. 5, pp. 2148–2158, May 2014.

[13] M. Pereyra, N. Dobigeon, H. Batatia, and J. Tourneret, "Estimating the granularity coefficient of a Potts-Markov random field within a Markov Chain Monte Carlo algorithm," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2385–2397, June 2013.

[14] J.M.P. Nascimento and J.M. Bioucas-Dias, "Hyperspectral unmixing based on mixtures of Dirichlet components," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 863–878, March 2012.

[15] M. J. Betancourt, "Cruising the simplex: Hamiltonian Monte Carlo and the Dirichlet distribution," in *ArXiv e-prints*, Oct. 2013.

[16] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, April 2005.

[17] Zhanyu Ma, "Bayesian estimation of the Dirichlet distribution with expectation propagation," in *EUSIPCO-12*, Bucharest, Romania, Aug. 2012.

[18] A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Unsupervised unmixing of hyperspectral images accounting for endmember variability," in *ArXiv e-prints*, Jun. 2014.

[19] C. P. Robert, *The Bayesian Choice: from Decision-Theoretic Motivations to Computational Implementation*, Springer Texts in Statistics. Springer-Verlag, New York, 2 edition, 2007.

[20] RSI (Research Systems Inc.), *ENVI User's guide Version 4.0*, Boulder, CO 80301 USA, Sept. 2003.

[21] D. C. Heinz and C. -I Chang, "Fully constrained least-squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 29, no. 3, pp. 529–545, March 2001.

[22] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009.

[23] B. Somers, M. Zortea, A. Plaza, and G.P. Asner, "Automated extraction of image-based endmember bundles for improved spectral unmixing," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 5, no. 2, pp. 396–408, April 2012.