JOINT SPECTRAL CLASSIFICATION AND UNMIXING USING ADAPTATIVE PIXEL NEIGHBORHOODS

Olivier Eches[1,2], *Jon Atli Benediktsson*[1], *Nicolas Dobigeon*[2] *and Jean-Yves Tourneret*[2]

[1]University of Iceland, Fac. of Electrical and Computer Eng., Hjardarhagi 2-6, 107 Reykjavik, Iceland
[2]University of Toulouse, IRIT/INP-ENSEEIHT, 2 rue Camichel, 31071 Toulouse cedex 7, France

ABSTRACT

A new spatial unmixing algorithm for hyperspectral images is studied. This algorithm is based on the well-known linear mixing model. The spectral signatures (or *endmembers*) are assumed to be known while the mixture coefficients (or *abundances*) are estimated by a Bayesian algorithm. As a pre-processing step, an *area filter* is employed to partition the image into multiple spectrally consistent connected components or *adaptative neighborhoods*. Then, spatial correlations are introduced by assigning to the pixels of a given neighbourhood the same hidden labels. More precisely, these pixels are modeled using a new prior distribution taking into account spectral similarity between the neighbors. Abundances are reparametrized by using logistic coefficients to handle the associated physical constraints. Other parameters and hyperparameters are assigned appropriate prior distributions. After computing the joint posterior distribution, a hybrid Gibbs algorithm is employed to generate samples that are asymptotically distributed according to this posterior distribution. The generated samples are finally used to estimate the unknown model parameters. Simulations on synthetic data illustrate the performance of the proposed method.

1. INTRODUCTION

Spectral unmixing is an important step for hyperspectral image analysis. It aims at estimating the spectral signatures of the pure materials (or *endmembers*) present in any hyperspectral image pixel and its corresponding fractions (or *abundances*) usually assuming that the mixture is linear. More precisely, this linear mixture model (LMM) can be expressed as

$$\boldsymbol{y}_p = \boldsymbol{M}\boldsymbol{a}_p + \boldsymbol{n}_p, \tag{1}$$

where $\boldsymbol{y}_p = [y_{p,1}, \ldots, y_{p,L}]^T$ is the L-spectrum of the observed pixel, $\boldsymbol{M} = [\boldsymbol{m}_1, \ldots, \boldsymbol{m}_R]$ is a known $L \times R$ matrix containing the L-spectra of the endmembers, $\boldsymbol{a}_p$ is the $R \times 1$ abundance vector associated to the pixel p, R is the number of endmembers that are present in the image and $\boldsymbol{n}_p$ is the independent and identically distributed (i.i.d.) zero-mean Gaussian noise sequence with the same variance s^2 for each spectral band. Note that abundances satisfy positivity and sum-to-one constraints since they are proportions. Spectral unmixing begins with the recovery of the endmembers using an endmember extraction algorithm (EEA). EEAs are usually preceded by a dimensionality reduction step using a feature reduction algorithm such as Principal Component Analysis (PCA). In this work, the endmembers are assumed to be known, extracted from a spectral library or recovered by an EEA like the Minimum Volume Simplex Analysis (MVSA) [1] or the well-known N-FINDR [2]. Once the endmembers have been identified, the abundances are estimated in the so-called "inversion" step. Inversion algorithms are based on optimization theory like the FCLS [3], on Bayesian inference in [4] or on a fuzzy membership process [5]. The last mentioned approach employed Markov Random Fields (MRFs) to take into account spatial correlations between image pixels and inspired the hierarchical Bayesian MRF-based model developed in [6] yielding a joint unmixing and segmentation algorithm. However, a major drawback of MRF comes from the resulting high computational complexity.

In this paper, it is proposed to exploit the spatial correlations in a different way. The new joint segmentation and unmixing procedure is based on an *adaptative neighborhood* building process that has been developed for the classification of hyperspectral images [7]. These neighborhoods, regrouping connected pixels that are spectrally consistent, are built by applying a self-complementary area filter [8] on the hyperspectral image. After this pre-processing step, an implicit classification is carried out by assigning class labels to each neighborhood. The classes not only share the same abundance means and covariances but also have similar spectral characteristics. Since the pixels belonging to a given neighborhood must belong to the same class, the spatial dependencies are modeled by assigning to the class labels a prior distribution depending on the spectrum similarity between these neighborhoods. Then, appropriate prior distributions with unknown means and variances depending on the pixel class are chosen for the abundances which are reparametrized in the same way as in [6]. The associated hyperparameters are assigned non-informative prior distributions according to a hierarchical Bayesian model. The joint posterior distribution of the unknown parameters and hyperparameters is finally computed from the likelihood and the prior distributions. However, deriving the Bayesian estimators such as the MMSE and MAP estimators is too difficult from this posterior distribution. Thus, following the strategy in [6], we propose to use Markov Chain Monte Carlo (MCMC) methods to generate samples asymptotically distributed according to this distribution. These samples are then used to estimate the unknown model parameters.

The paper is organized as follows. Section 2 presents the pre-processing step that is used to build the adaptative neighborhoods. In section 3, we present the modified Bayesian model from [6] while the resulting hybrid Gibbs algorithm is given in section 4. Section 5 reports simulation results on synthetic and real data.

2. PRE-PROCESSING STEP

2.1. Adaptative neighborhood

In order to build the adaptative neighborhood on hyperspectral data, the authors in [7] employed a *flattening* procedure stemming from the self-complementarity property [8]. Self-complementarity is an important property in morphological theory and allows the structure

of interest to be preserved independently of their contrasts while removing small meaningless structures (e.g., cars, trees,...) in very high resolution remote sensing images. The algorithm developed by Soille in [8] exploits this property in a two step procedure that divides the image into flat zones, i.e., a region where neighboring pixels have the same values, satisfying any area criterion λ. This procedure is repeated until the desired minimal flat zone size λ is obtained. Note that this self-complementary area filter cannot be directly used on hyperspectral images since the complete ordering property that any morphological operator needs are absent from these data. The proposed strategy in [7] employs a PCA to reduce the dimensionality of the data. Then, the area filtering is computed on the data projected on the first principal component space, i.e., where the corresponding covariance matrix eigenvalue is the greatest. The resulting flat zones regroup pixels that are spectrally consistent and are therefore considered as neighbors.

As stated in the introduction, the main contribution of this paper consists of using the adaptive neighborhood building methods developed in [7] as a pre-processing step for a spatial unmixing algorithm. The neighborhoods resulting from the method derived in [7] are considered for each band of the data. Then, spatial information is extracted from each of the neighborhoods by computing the corresponding vector median value. More precisely, if we denote the number of neighborhoods by S and the sth neighborhood by $\mathbf{\Omega}_s$ ($s = 1, \ldots, S$), then the vector median value for this neighborhood is defined as

$$\mathbf{\Upsilon}_s = \mathrm{med}(\boldsymbol{Y}_{\mathbf{\Omega}_s}), \qquad (2)$$

where $\boldsymbol{Y}_{\mathbf{\Omega}_s}$ represents the matrix of observed pixels belonging to the neighborhood $\mathbf{\Omega}_s$ and $\dim(\mathbf{\Upsilon}_s) = L$ is the number of spectral bands. As explained in [7], the median vector ensures spectral consistency as opposed to the mean vector.

2.2. Image segmentation

As in [6], it is assumed in this paper that the classes contain neighboring pixels that have *a priori* close abundances. This spatial dependency is modelled using the resulting adaptive neighborhoods since they regroup spectrally consistent pixels. Thus, the pixels of a given neighborhood are considered belonging to the same class, since the spectral values are depending on the abundances. In other words, if we let $\mathcal{C}_1, \ldots, \mathcal{C}_K$ defining the classes, a label vector of size $S \times 1$, $S \geq K$ denoted as $\boldsymbol{z} = [z_1, \ldots, z_S]^T$ with $z_s \in \{1, \ldots, K\}$ is introduced to identify the class to which each neighborhood $\mathbf{\Omega}_s$ and its underlying pixels belong, i.e., $z_s = k$ if and only if $\mathbf{\Omega}_s \in \mathcal{C}_k$.

3. BAYESIAN MODEL

In this section, a Bayesian model inspired by [6] is briefly presented.

3.1. Likelihood

The abundances $\boldsymbol{a}_p$ ($p = 1, \ldots, P$) are assumed to depend on logistic coefficients $\boldsymbol{t}_p$ (as in [5],[6]), ensuring positivity and sum-to-one constraints, in such a way that

$$a_{r,p} = \frac{\exp(t_{r,p})}{\sum_{r=1}^{R} \exp(t_{r,p})}, \quad r = 1, \ldots, R. \qquad (3)$$

The noise variances are the same for all the image pixels. Consequently, the unknown parameter vector is $\boldsymbol{\Theta} = \{\boldsymbol{T}, \boldsymbol{z}, s^2\}$ where s^2 is the noise variance, $\boldsymbol{z}$ is the label vector and $\boldsymbol{T} = [\boldsymbol{t}_1, \ldots, \boldsymbol{t}_P]$

with $\boldsymbol{t}_p = [\boldsymbol{t}_{1,p}, \ldots, \boldsymbol{t}_{R,p}]^T$ is the logistic coefficient matrix used for the abundance reparametrization. The additive noise considered as white Gaussian allows us to write for a given pixel[1] p ($p = 1, \ldots, P$) $\boldsymbol{y}_p | \boldsymbol{t}_p, s_p^2 \sim \mathcal{N}\left(\boldsymbol{M} \boldsymbol{a}_p(\boldsymbol{t}_p), s_p^2 \boldsymbol{I}_L\right)$.

The complete expression of the likelihood for a given pixel p and for the whole image is detailed in [6], as for the prior distributions for the logistic coefficients, noise variances and the associated hyperparameters. Only the label prior distribution which is different from that of [6] will be presented in the sequel.

3.2. Label prior

In this paper it is assumed that neighborhoods which have similar spectral median $\mathbf{\Upsilon}_{\mathbf{\Omega}_s}$ should belong to the same class, thus extending the spatial dependency to the whole image and not only the locally connected pixels. More precisely, considering two neighborhoods $\mathbf{\Omega}_s$ and $\mathbf{\Omega}_t$, we introduce the distance $D_{s,t} = \|\mathbf{\Upsilon}_s - \mathbf{\Upsilon}_t\|^2$ ($\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^T \boldsymbol{x}}$ being the standard ℓ^2 norm) and we define $\mathcal{V}_\tau(t)$ the set of index t such that $D_{s,t} < \tau$ where τ is a threshold tuning the degree of spectral similarity between two neighborhoods. The prior distribution for the label z_s of the neighborhood $\mathbf{\Omega}_s$ is

$$P(z_s = k | \boldsymbol{z}_{\text{-}s}) \propto \exp\left[\kappa \sum_{t \in \mathcal{V}_\tau(t)} \delta(z_s - z_t)\right], \qquad (4)$$

where $\propto$ means "proportional to", $\delta(\cdot)$ is the Kronecker function, κ is a parameter tuning the degree of granularity of the neighborhood classes and $\boldsymbol{z}_{\text{-}s} = [z_1, \ldots, z_{s-1}, z_{s+1}, \ldots, z_S]$. The Hammersley-Clifford theorem allows us to demonstrate that the joint prior distribution for the label vector $\boldsymbol{z}$ is well defined and expressed as (see Appendix)

$$P(\boldsymbol{z}) \propto \exp\left[\kappa \sum_{s=1}^{S} \sum_{t \in \mathcal{V}_\tau(t)} \delta(z_s - z_t)\right]. \qquad (5)$$

3.3. Joint distribution

We first recall the hyperparameters from [6] that are grouped into the vector $\Phi = \{\boldsymbol{\Psi}, \boldsymbol{\Sigma}, \upsilon^2, \delta\}$ where $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_K]$, $\boldsymbol{\Sigma} = \left\{\mathrm{diag}\left(\sigma_{r,1}^2\right), \ldots, \mathrm{diag}\left(\sigma_{r,K}^2\right)\right\}$ are associated to the logistic coefficient matrix $\boldsymbol{T}$, δ is associated to the noise variance and υ^2 to $\boldsymbol{\Psi}$. Using Bayes theorem, the joint posterior distribution can be obtained by replacing the former prior label distribution by (5). By denoting $n_k = \mathrm{card}(\mathcal{I}_k)$, this leads to

$$f(\boldsymbol{\Theta}, \Phi | \boldsymbol{Y}) \propto \left(\frac{1}{s^2}\right)^{\frac{LP}{2}} \prod_{p=1}^{P} \exp\left[-\frac{\|\boldsymbol{y}_p - \boldsymbol{M} \boldsymbol{a}_p(\boldsymbol{t}_p)\|^2}{2s^2}\right]$$

$$\times \exp\left[\kappa \sum_{t=1}^{S} \sum_{t \in \mathcal{V}_\tau(t)} \delta(z_s - z_t)\right]$$

$$\times \frac{\delta^{\nu-1}}{(s^2)^{\nu+1}} \exp\left(-\frac{\delta}{s^2}\right) \prod_{p=1}^{P} \left(\frac{1}{\upsilon^2}\right)^{\frac{RK}{2}+1}$$

$$\times \prod_{r,k} \frac{1}{\sigma_{r,k}^{n_k+1}} \exp\left[-\left(\frac{\psi_{r,k}^2}{2\upsilon^2} + \frac{2\gamma + \sum_{p \in \mathcal{I}_k}(t_{r,p} - \psi_{r,k})^2}{2\sigma_{r,k}^2}\right)\right] \qquad (6)$$

---

[1] Note that the dependence of the abundance vector $\boldsymbol{a}_p$ on the logistic coefficient vector $\boldsymbol{t}_p$ through (3) has been explicitly mentioned by denoting $\boldsymbol{a}_p = \boldsymbol{a}_p(\boldsymbol{t}_p)$.

This distribution is too complex to easily compute the MMSE and MAP estimators of $\boldsymbol{\Theta}$. Thus we propose to use a Hybrid Gibbs sampler inspired by [6], generating samples that are asymptotically distributed according to (6). These samples are then used to approximate the Bayesian estimators.

4. HYBRID GIBBS ALGORITHM

In a Gibbs sampler, samples are iteratively drawn from the full conditional distributions associated to the posterior of interest. Since the conditional distributions of the parameters and hyperparameters are the same as in [6] except for $\boldsymbol{z}$, we will only derive the $\boldsymbol{z}$ conditional distribution in this section.

Using the same algorithm as in [6], the conditional distribution for each neighborhood label is computed using the Bayes relation. For a given neighborhood $\boldsymbol{\Omega}_s$, we define $\boldsymbol{I}_s$ as the set of pixel indexes belonging to $\boldsymbol{\Omega}_s$. The conditional distribution of z_s is then

$$P\left[z_s = k | \boldsymbol{z}_{\text{-}s}, \boldsymbol{T}_s, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k\right] \propto f(z_s | \boldsymbol{z}_{\text{-}s}) \prod_{i \in \boldsymbol{I}_s} f(\boldsymbol{t}_i | \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k),$$

where $\boldsymbol{\Sigma}_k = \text{diag}\left(\sigma_{r,k}^2\right)$ and $\boldsymbol{T}_s$ is the logistic coefficient matrix reduced to the pixel index included in the neighborhood $\boldsymbol{\Omega}_s$. Since the label of a given neighborhood is the same for all pixels, it makes sense that every corresponding logistic coefficient contributes to the conditional distribution of z_s. The complete expression of the conditional distribution is

$$P\left[z_s = k | \boldsymbol{z}_{\text{-}s}, \boldsymbol{T}_s, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k\right] \propto \exp\left[\kappa \sum_{t \in \mathcal{V}_\tau(t)} \delta(z_s - z_t)\right]$$
$$\times \prod_{i \in \boldsymbol{I}_s} f(\boldsymbol{t}_i | \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k), \quad (7)$$

where $f(\boldsymbol{t}_i | \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k)$, is given in [6]. Note that sampling from this conditional distribution can be achieved using the same method as in [6], i.e., by drawing a discrete value in the finite set $\{1, ..., K\}$ with the normalized probabilities (7). As mentioned above, the conditional distributions for the other parameters can be found in [6] and will not be detailed in this article for conciseness. After defining the neighborhoods and extracting the median vector, the proposed hybrid Gibbs sampler iteratively generates N_{MC} samples asymptotically distributed according to the conditional distributions. The first generated samples N_{bi} belonging to the so-called *burn-in* period are ignored whereas the last samples are used to estimate the unknown model parameters and hyperparameters. The estimates of the class labels are obtained using the MAP estimator approximated by retaining the samples that maximizes the posterior conditional probabilities of $\boldsymbol{z}$. Then, the abundances are estimated conditionnally on these MAP estimates using the MMSE estimator, approximated by averaging over the $N_{\text{MC}} - N_{\text{bi}}$.

5. SIMULATION RESULTS

We consider a 25×25 synthetic image generated with $K = 3$ different classes. The image contains $R = 3$ mixed components whose spectra ($L = 413$ spectral bands) have been extracted from the spectral libraries distributed with the ENVI package [9] and represent construction concrete, green grass and micaceous loam. A label map shown in Fig.1 (left) has been generated using a Markov random field with a granularity coefficient $\beta = 2$. Then, abundances have been generated using different means and variances for each class (see Table 1 for the actual values). Fig. 2 (top) depicts the generated abundance maps. Note that a black (resp. white) pixel indicates a

weak (resp. strong) value of the abundance coefficient. The noise variance has been fixed to $s^2 = 0.001$ leading to a signal-to-noise ratio of 20dB. The neighborhoods have been built with a minimal flat zone size parameter $\lambda = 5$. The threshold for the norm of the difference between the medians $D_{s,t}$ is equal to $\tau = 5 \times 10^{-3}$ and the granularity parameter has been set to $\kappa = 1$.
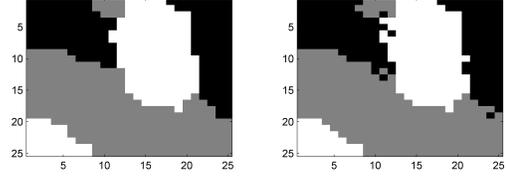

Fig. 1. Left: original labels. Right: labels estimated by the proposed hybrid Gibbs sampler.
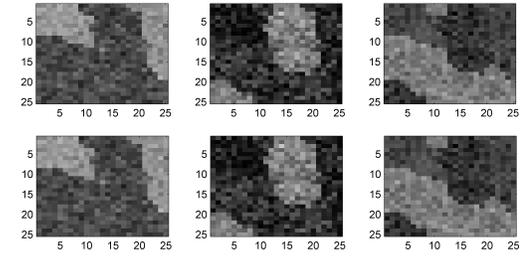

Fig. 2. Top: abundance maps of the 3 pure materials for LMM. Bottom: abundance maps of the 3 pure materials estimated by the proposed Gibbs sampler (from left to right: construction concrete, green grass, micaceous loam).

The classification map obtained from the MAP estimates of the label samples is depicted in Fig. 1 and is in good agreement with the actual one. Conditionnally to these MAP estimates, the MMSE estimates of the abundances are given in Fig. 2 and the estimated values of the abundance means and variances are reported in Table 1. The estimated classes, abundance coefficients and abundance mean vectors estimated by our algorithm are clearly in accordance with the actual values of these parameters.

Table 1. Actual and estimated abundance mean and variance ($\times 10^{-3}$) in each class.

		Actual values	Estimated values
Class 1	$E[\boldsymbol{a}_{p, \, p \in \mathcal{I}_1}]$	$[0.6, 0.3, 0.1]^T$	$[0.58, 0.29, 0.13]^T$
	$\text{Var}[a_{p,r, \, p \in \mathcal{I}_1}]$	$[5, 5, 5]^T$	$[6.2, 3.9, 7.9]^T$
Class 2	$E[\boldsymbol{a}_{p, \, p \in \mathcal{I}_2}]$	$[0.3, 0.5, 0.2]^T$	$[0.31, 0.49, 0.2]^T$
	$\text{Var}[a_{p,r, \, p \in \mathcal{I}_2}]$	$[5, 5, 5]^T$	$[5.8, 5.2, 8.9]^T$
Class 3	$E[\boldsymbol{a}_{p, \, p \in \mathcal{I}_3}]$	$[0.3, 0.2, 0.5]^T$	$[0.29, 0.2, 0.5]^T$
	$\text{Var}[a_{p,r, \, p \in \mathcal{I}_3}]$	$[5, 5, 5]^T$	$[5.0, 4.7, 9.0]^T$

The proposed spatial hybrid Gibbs sampler is compared respectively with its MRF counterpart developed in [6] and with the non-spatial Bayesian algorithm developed in [4]. The synthetic image

built with SNR = 20dB has been analyzed by these two algorithms with the same number of iterations N_{MC}. As a performance criterion, the global mean square errors (MSEs) of the rth estimated abundances have been computed. This global MSE is defined as

$$\mathrm{MSE}_r^2 = \frac{1}{P} \sum_{p=1}^{P} (\hat{a}_{r,p} - a_{r,p})^2 \qquad (8)$$

where $\hat{a}_{r,p}$ denotes the MMSE estimate of the abundance $a_{r,p}$. Table 2 reports the obtained results with the time of simulation showing that the algorithm developed in this paper (referred to as "Neighborhoods") performs similarly or better than the two other algorithms (referred to as "MRF" and "Bayesian") in terms of estimation performance. However the proposed algorithm shows the lowest computational time which is a very interesting property.

Table 2. Global MSEs of each abundance component and time of simulation for the three unmixing algorithms.

	Bayesian	MRF	Neighborhoods
MSE_1^2	5.8×10^{-3}	3.4×10^{-4}	3.2×10^{-4}
MSE_2^2	5.9×10^{-3}	9.5×10^{-5}	8.3×10^{-5}
MSE_3^2	2.3×10^{-4}	2.3×10^{-4}	2.5×10^{-4}
Time (sec.)	4.6×10^3	2×10^3	1.6×10^3

6. CONCLUSIONS

A new spatial unmixing algorithm based on adaptative neighborhoods was proposed for hyperspectral images. The main novelty of this algorithm is that spatial dependencies were taken into account by assigning hidden discrete variables (labels) to pixels belonging to appropriate neighborhoods partionning the image into multiple classes. The classes were defined by homogeneous abundances with a common mean vector and a common covariance matrix. A new Bayesian model based on the neighborhood structure and on ideas developed in a previous study was derived. The complexity of this Bayesian model was alleviated by implementing an hybrid Gibbs sampler generating data asymptotically distributed according to the posterior distribution of interest. Simulations conducted on synthetic data showed that the proposed algorithm achieved similar classification and unmixing performance than its Markov random field counterpart but with a reduced computational time. Future works will investigate the generalization of the self-complementary area filter to multiple bands by using a method inspired by [10].

7. ACKNOWLEDGEMENTS

8. APPENDIX: LABEL PRIOR DISTRIBUTION

The Hammersley-Clifford theorem [11, p.231] allows us to write

$$P(\boldsymbol{z}) \propto \prod_{s=1}^{S} \frac{P(z_s | z_1^*, \ldots, z_{s-1}^*, z_{s+1}, \ldots, z_S)}{P(z_s^* | z_1^*, \ldots, z_{s-1}^*, z_{s+1}, \ldots, z_S)},$$

where the labels marked with a star z_s^* are arbitrary auxiliary variables.

Therefore,

$$P(\boldsymbol{z}) \quad \propto \quad \exp \Bigg\{ \kappa \sum_{s=1}^{S} \Bigg[\Bigg(\sum_{t \in \mathcal{V}_\tau(t), t<s} \delta(z_s - z_t^*) \\ + \sum_{t \in \mathcal{V}_\tau(t), t>s} \delta(z_s - z_t) \Bigg) \\ - \Bigg(\sum_{t \in \mathcal{V}_\tau(t), t<s} \delta(z_s^* - z_t^*) + \sum_{t \in \mathcal{V}_\tau(t), t>s} \delta(z_s^* - z_t) \Bigg) \Bigg] \Bigg\}.$$

By interverting the indexes s and t, we have

$$\sum_{s=1}^{S} \sum_{t \in \mathcal{V}_\tau(t), t<s} \delta(z_s - z_t^*) = \sum_{s=1}^{S} \sum_{t \in \mathcal{V}_\tau(t), t>s} \delta(z_s^* - z_t). \qquad (9)$$

This allows us to write the final expression

$$P(\boldsymbol{z}) \propto \exp \Bigg[\kappa \sum_{s=1}^{S} \sum_{t \in \mathcal{V}_\tau(t)} \delta(z_s - z_t) \Bigg]. \qquad (10)$$

9. REFERENCES

[1] J. Li and J. M. Bioucas-Dias, "Minimum volume simplex analysis: a fast algorithm to unmix hyperspectral data," in *Proc. IEEE Int. Conf. Geosci. Remote Sens. (IGARSS)*, vol. 3, Boston, MA, Jul. 2008, pp. 250–253.

[2] M. E. Winter, "Fast autonomous spectral endmember determination in hyperspectral data," in *Proc. 13th Int. Conf. on Applied Geologic Remote Sensing*, vol. 2, Vancouver, April 1999, pp. 337–344.

[3] 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. and Remote Sensing*, vol. 39, no. 3, pp. 529–545, March 2001.

[4] N. Dobigeon, J.-Y. Tourneret, and C.-I Chang, "Semi-supervised linear spectral using a hierarchical Bayesian model for hyperspectral imagery," *IEEE Trans. Signal Processing*, vol. 56, no. 7, pp. 2684–2696, July 2008.

[5] J. T. Kent and K. V. Mardia, "Spatial classification using fuzzy membership models," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 10, no. 5, pp. 659–671, 1988.

[6] O. Eches, N. Dobigeon, and J.-Y. Tourneret, "Enhancing hyperspectral image unmixing with spatial correlations," *IEEE Trans. Geosci. and Remote Sensing*, 2011, submitted. [Online]. Available: http://www3.hi.is/~eches

[7] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Adaptive pixel neighborhood definition for the classification of hyperspectral images with support vector machines and composite kernel," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, San Diego, CA, USA, Dec. 2008, pp. 1884–1887.

[8] P. Soille, "Beyond self-duality in morphological image analysis," *Image and Vision Computing*, vol. 23, no. 2, pp. 249–257, June 2005.

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

[10] S. Velasco-Forero and J. Angulo, "Spatial structuresdetection inhyperspectral images using mathematical morphology," in *Proc. IEEE GRSS Workshop on Hyperspectral Image and SIgnal Processing: Evolution in Remote Sensing (WHISPERS)*, Reykjavik, Iceland, June 2010, pp. 1–4.

[11] J.-M. Marin and C. P. Robert, *Bayesian core: a practical approach to computational Bayesian statistics*. New-York: Springer, 2007.