Weighted residual NMF with spatial regularization for hyperspectral unmixing

Taner Ince, Nicolas Dobigeon

Weighted Residual NMF with Spatial Regularization for Hyperspectral Unmixing

Taner Ince, *Member, IEEE*, Nicolas Dobigeon, *Senior Member, IEEE*

***Abstract*—This paper proposes a weighted residual nonnegative matrix factorization (NMF) with spatial regularization to unmix hyperspectral data. NMF decomposes a matrix into the product of two nonnegative matrices. However, NMF is known to be generally sensitive to noise, which makes difficult to retrieve the global minimum of the underlying objective function. To overcome this limitation, we include a residual weighting mechanism in the conventional NMF formulation. This strategy treats each row of the residual based on the weighting factor. In this manner, residuals with large values are penalized less and residuals with small values are penalized more to make NMF based unmixing problem more robust. Furthermore, we include a weight term in the form of an ℓ_1 norm regularizer to provide spatial information of the abundance matrix. Experimental results are conducted to validate the effectiveness of the proposed method.**

***Index Terms*—Hyperspectral unmixing, NMF, sparse unmixing, residual weighting.**

I. Introduction

HYPERSPECTRAL sensors capture the light spectrum reflected from objects over hundred narrow wavelengths. The high spectral resolution brought by hyperspectral (HS) sensors allows the properties of imaged materials to be analyzed with an accuracy unattainable by panchromatic and multispectral sensors. Therefore, HS imaging is used in many applicative areas ranging from Earth observation to health science. In the specific context of Earth observation, HS imaging is used mainly to understand the properties of ground materials and to infer a meaningful description of the observed scene. One of the main issues raised by HS imaging is the limited spatial resolution of the images, which leads to the presence of mixed pixels, i.e., the spectrum measured in each spatial position is a combination of the spectral signatures associated with several distinct materials present in the considered. The objective of spectral unmixing (SU) is to find the pure spectral signatures representative of these materials (endmembers) and their related fractions (abundances) in each mixed pixel [1].

SU generally leverages on a prescribed analytic model describing how the endmember spectral signatures contribute to the observed pixel spectra. These models can be linear or nonlinear depending on the light interactions of the materials in the scene. While nonlinear mixing models may account for multiple scattering effects occurring in particular scenarios [2], the linear mixing model (LMM) is generally considered as quite reasonable first approximation of the physical process underlying the measurements. LMM assumes that the measured pixel spectra result from weighted linear combinations of the endmember spectra. Based on the LMM, several classes of unmixing algorithms have been proposed in the literature. Geometrical approaches recover the endmembers by identifying the vertices of the simplex formed by the mixed pixels and abundances are subsequently obtained by conducting an inversion step [3]–[5]. Sparse regression based unmixing aims at estimating the abundances using an *a priori* available spectral library [6]–[10]. Alternatively, nonnegative matrix factorization (NMF) [11] embraces a wide class of algorithms that have encountered a certain success in address the problem of HS unmixing [12]–[14]. This success can be easily explained by the fact that the generic formulation of NMF is particularly well suited to conduct an LMM-based unmixing task. Indeed, NMF decomposes the matrix associated with the measurements into the product of two nonnegative matrices, associated with the endmember signatures and abundances in the context of HS unmixing. However, besides nonnegativity, the canonical instance of NMF does not impose any additional constraint on the sought solution. Due to the nonconvexity of the underlying optimization problem, the retrieved solution is rarely guaranteed to be a global minimum and the algorithm is very sensitive to the initialization. Some strategies have been deployed to stabilize the solution by incorporating additional constraints imposed to the endmembers, e.g., penalizing the volume of the simplex they span [12], their variance [15] or their similarity [16].

Another way to reach a relevant solution consists in regularizing the abundance matrix. Most of the approaches advocated in the literature rely on sparsity-promoting penalizations to reflect the fact that each pixel is composed of a few endmembers. Various sparsity measures have been considered since the seminal work of Jia and Qian [13], including the use of ℓ_1-norm [17] or $\ell_{1/2}$-norm [18]. To extract the structural information in HS data, the $\ell_{1/2}$-NMF (GLNMF) proposed in [19] constructs a graph to extract the local information and $\ell_{1/2}$-norm enforces the sparsity of the abundance matrix. It is also known that pixels within a local region are likely to share the same set of endmembers. Thus spatial group sparsity regularized NMF (SGSNMF) is proposed in [20] where local spatial regions are extracted using superpixel segmentation. Furthermore, using the piecewise smooth property of abundance map, total variation (TV) regularized reweighted sparse

Part of this work has been supported by the ANR-3IA Artificial and Natural Intelligence Toulouse Institute (ANITI) under grant agreement ANITI ANR-19-PI3A-0004 and by the French National Research Agency under grant agreement ANR-21-CE29-0007.

Taner Ince is with the Department of Electrical and Electronics Engineering, Gaziantep University, 27310 Gaziantep, Turkey (e-mail:tanerince@gantep.edu.tr).

Nicolas Dobigeon is with the University of Toulouse, IRIT/INP-ENSEEIHT, 31000 Toulouse, France, and also with the Institut Universitaire de France (IUF), France (e-mail: nicolas.dobigeon@enseeiht.fr).

NMF (TV-RSNMF) is proposed in [21]. Reweighting strategy enhances the sparsity of the abundance matrix and TV captures the spatial information of the abundance map. Spectral-spatial weighted sparse NMF (SSWNMF) [22] includes spatial and spectral information as a weight term in the ℓ_1-NMF framework. Furthermore, robust NMF approaches are developed to make unmixing robust to outliers [23]. Sparsity regularized robust NMF (RNMF) [24] solves a sparsity regularized NMF problem and at the same time eliminates sparse noise by introducing an additional regularizer in the model. Robust collaborative NMF (R-CoNMF) [25] exploits the row-sparse structure of the abundance matrix using a $\ell_{2,1}$-mixed norm regularizer.

In this paper, we capitalize on this latest advances to define a new NMF-based unmixing algorithm. More precisely, a weighting term is included into the data fitting term to make the decomposition more robust in high noise scenarios. Moreover, a spatial weighting term is incorporated as a regularization to exploit the smoothness of the abundance maps. These two simple ingredients are shown to lead to state-of-the art unmixing results in the considered experimental setups.

The remaining of this letter is organized as follows. Section II gives a brief background and explains the proposed method. Experimental results are reported in Section III and Section IV concludes the paper.

II. Weighted Residual NMF (WRNMF)

A. Problem formulation

Let $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_N] \in \mathbb{R}^{L\times N}$ denote the matrix composed of N observed pixels spectra in L spectral bands. LMM states that the observation matrix can be decomposed as $\mathbf{Y} = \mathbf{MA} + \mathbf{E}$ where $\mathbf{M} \in \mathbb{R}^{L\times K}$ is the matrix containing the K endmember spectral signatures, $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_N] \in \mathbb{R}^{K\times N}$ is the fractional abundance matrix and $\mathbf{E} \in \mathbb{R}^{L\times N}$ accounts for noise and modeling error. Because of physical considerations, the entries of these matrices are expected to be nonnegative. Thus, recovering the endmember and abundance matrices $\mathbf{M}$ and $\mathbf{A}$ from the observation matrix $\mathbf{Y}$ can be cast as an NMF problem. This unmixing task consists in solving the following optimization problem

$$\min_{\mathbf{M},\mathbf{A}} \frac{1}{2}\|\mathbf{Y} - \mathbf{MA}\|_{\mathrm{F}}^2 \quad \text{s.t} \quad \mathbf{M} \geq \mathbf{0}, \quad \mathbf{A} \geq \mathbf{0} \tag{1}$$

where $\geq$ stands for an entry-wise inequality.

Several class of algorithms have been proposed to solve this optimization problem. Most of them consist in updating the endmember matrix $\mathbf{M}$ and the abundance matrix $\mathbf{A}$ iteratively and individually, using conventional first order descent methods or multiplicative updates [26]. However, the equation (1) defines a non-convex optimization problem, which is known to raise several issues to yield physically interpretable results, in particular the high sensitivity with respect to noise and to the algorithm initialization. To alleviate, additional constraints and/or regularizations can be included. First, in the specific context of HS unmixing, the abundances are generally accompanied by a sum-to-one constraint to associate these mixing coefficients with proportions. Moreover, one can capitalize on the expected spatial homogeneity of the abundance maps to be estimated. Thus, we propose to complement the optimization problem (1) with a specific form of an ℓ_1-norm regularization promoting this spatially homogeneous abundance maps. Besides, to mitigate the impact of noise, we propose to deploy a reweighting strategy, which consists in adjusting the data fitting term (also referred to as the residual) along the algorithm iterations. To summarize, in this work, the proposed unmixing formulation can be defined as the optimization problem

$$\begin{aligned} &\min_{\mathbf{M},\mathbf{A}} \frac{1}{2}\|\mathbf{W}(\mathbf{Y} - \mathbf{MA})\|_F^2 + \lambda\|\mathbf{S} \odot \mathbf{A}\|_1 \\ &\text{s.t} \quad \mathbf{M} \geq \mathbf{0}, \quad \mathbf{A} \geq \mathbf{0}, \quad \mathbf{1}_K^{\mathrm{T}}\mathbf{A} = \mathbf{1}_N^{\mathrm{T}} \end{aligned} \tag{2}$$

where $\mathbf{W} \in \mathbb{R}^{L\times L}$ is a diagonal matrix acting as a weighting operator, $\mathbf{S} \in \mathbb{R}^{K\times N}$ is the spatial weighting matrix and λ is used to control the spatial regularization of $\mathbf{A}$. The definition of the weighting matrices will be discussed in paragraph II-C.

B. Optimization problem

Adopting the strategy followed by most of the algorithms already proposed in the literature, we propose to split the optimization problem (2) into two subproblems to update $\mathbf{M}$ and $\mathbf{A}$ alternately using multiplicative update rules [26]. These two steps are detailed in what follows. Interested readers are invited to consult the supporting document [27] for more details.

Updating $\mathbf{M}$ – The optimization problem related to the update of the endmember matrix $\mathbf{M}$ is given by

$$\min_{\mathbf{M}} \mathcal{J}(\mathbf{M}) \triangleq \frac{1}{2}\|\mathbf{W}(\mathbf{Y} - \mathbf{MA})\|_{\mathrm{F}}^2 \text{ s.t. } \mathbf{M} \geq \mathbf{0}.$$

Following standard derivations similar to those detailed in [26], the updating rules can be written in a matrix form as

$$\mathbf{M} \leftarrow \mathbf{M} \cdot \frac{\mathbf{W}^T\mathbf{WYA}^{\mathrm{T}}}{\mathbf{W}^T\mathbf{WMAA}^{\mathrm{T}}}$$

where the multiplication $\cdot$ and the division should be understood as entrywise (i.e., component-by-component).

Updating $\mathbf{A}$ – Regarding the abundance, the sum-to-one constraint is implicitly included into the optimization problem by resorting to the trick proposed in [28]. It consists in replacing the observation and endmember matrices $\mathbf{Y}$ and $\mathbf{M}$ by the extended ones defined as

$$\tilde{\mathbf{Y}} = \begin{bmatrix} \mathbf{Y} \\ \delta\mathbf{1}_N^T \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{M}} = \begin{bmatrix} \mathbf{M} \\ \delta\mathbf{1}_K^T \end{bmatrix} \tag{3}$$

where δ is used to adjust the importance of the sum-to-one constraint and $\mathbf{1}_d \in \mathbb{R}^d$ denote a vector composed of 1's. Then, the optimization problem becomes

$$\min_{\mathbf{A}} \mathcal{J}(\mathbf{A}) \triangleq \frac{1}{2}\|\tilde{\mathbf{W}}(\tilde{\mathbf{Y}} - \tilde{\mathbf{M}}\mathbf{A})\|_{\mathrm{F}}^2 + \lambda\|\mathbf{S} \odot \mathbf{A}\|_1 \text{ s.t. } \mathbf{A} \geq \mathbf{0}$$

where $\tilde{\mathbf{W}}$ is a diagonal matrix whose elements are defined as

$$\tilde{w}_{ll} = \begin{cases} w_{ll} & \text{if } l \in \{1, \ldots, L\} \\ \beta & \text{if } l = L + 1. \end{cases} \tag{4}$$

TABLE I
SIMULATED DATA: aSAD AND aMSE AVERAGED OVER 10 MONTE CARLO RUNS.

	SNR	WRNMF	SSWNMF	SGSNMF	TV-RSNMF	RSNMF	GLNMF	$L_{1/2}$-NMF	VCA-FCLS
aSAD	**10**	**0.0999**	0.1475	0.5853	0.1081	0.1083	0.1351	0.1360	0.1170
	20	**0.0406**	0.0434	0.1178	0.0429	0.0429	0.0427	0.0432	0.0514
	30	**0.0072**	0.0079	0.0075	0.0079	0.0082	0.0081	0.081	0.0137
	40	**0.0016**	0.0023	0.0028	0.0022	0.0024	0.0024	0.0024	0.0040
aMSE	**10**	**0.3051**	0.3623	0.4565	0.3476	0.3502	0.3561	0.3552	0.3622
	20	**0.1198**	0.1258	0.1377	0.1233	0.1283	0.1239	0.1274	0.1465
	30	**0.0145**	0.0161	0.0146	0.0159	0.0164	0.0163	0.0163	0.0180
	40	**0.0016**	0.0018	0.0024	0.0018	0.0019	0.0019	0.0019	0.0021

where β is a small constant to emphasize the sum-to-one constraint. The resulting updating rule writes

$$\mathbf{A} \leftarrow \mathbf{A} \cdot \frac{\tilde{\mathbf{M}}^T \mathbf{W}\mathbf{W}^T \tilde{\mathbf{Y}}}{\tilde{\mathbf{M}}^{\mathrm{T}} \mathbf{W}\mathbf{W}^T \tilde{\mathbf{M}}\mathbf{A} + \lambda \mathbf{S}}. \tag{5}$$

C. Adjusting the weights

This paragraph describes the strategy proposed to adjust the weighting matrices $\mathbf{W}$ and $\mathbf{S}$. Regarding $\mathbf{W}$, its main objective aims at mitigating the effect of noise whose level may be different from one band to another. By denoting $\mathbf{R}^{(t)} \triangleq \mathbf{Y} - \mathbf{M}^{(t)}\mathbf{A}^{(t)}$ the residual at the tth iteration of the algorithm detailed in the previous paragraph, the diagonal entries of $\mathbf{W}$ are defined as (for $l = 1, \ldots, L$)

$$w_{ll} = \exp\left(-\frac{1}{\mu}\|\mathbf{R}^{(t)}_{l,:}\|_2\right) \tag{6}$$

where $\mathbf{R}_{l,:}$ denotes the lth row of $\mathbf{R}$ and μ is a constant to specify the exponential decay of the residual norm. The terms with high residuals are penalized less and the terms having low residuals are penalized more. This reweighting strategy allows the impact of band with high level of noise to be mitigated during the estimation process.

Besides, at the tth iteration of the algorithm, the element of the spatial weighting matrix $\mathbf{S}$ associated with the abundance a_{kn} of the kth endmember in the nth pixel is adjusted as (for $k = 1, \ldots, K$ and $n = 1, \ldots, N$)

$$s_{kn} = \left[\frac{1}{|\mathcal{V}(n)|}\sum_{p \in \mathcal{V}(n)} a^{(t)}_{kp} + \epsilon\right]^{-1} \tag{7}$$

where $\mathcal{V}(n)$ denotes the set of pixel indices in the spatial neighborhood of the nth pixel and $|\mathcal{V}(n)|$ is its cardinality. The small constant ϵ aims to avoid numerical instabilities. This spatial weighting strategy exploits the spatial correlation of the nearby pixels in a small neighborhood to promote the expected piecewise smoothness of the abundance maps.

III. EXPERIMENTAL RESULTS

Based on experiments conducted on simulated and real data, this section compares the performance of the proposed method with well-known NMF-based unmixing algorithms proposed in the literature. More precisely, the proposed weighted residual NMF method (referred to as WRNMF) has been compared to SSWNMF [22], SGSNMF [20], TV-RSNMF [21], RNMF [24], $\ell_{1/2}$-NMF [18], GLNMF [19] and VCA-FCLS [5]. The unmixing results are assessed with respect to the abilility of recovering the endmember spectral signatures and estimating the abundance maps. The average spectral angle distance (aSAD) is used to evaluate the similarity of the actual and estimated endmember signatures as follows

$$\text{aSAD} = \frac{1}{K}\sum_{k=1}^{K} \arccos\left(\frac{\langle \mathbf{m}_k, \hat{\mathbf{m}}_k\rangle}{\|\mathbf{m}_k\|\|\hat{\mathbf{m}}_k\|}\right) \tag{8}$$

where $\mathbf{m}_k$ and $\hat{\mathbf{m}}_k$ denote the kth actual and estimated endmember signatures, respectively. The actual and estimated abundance maps are estimated with respect to the average mean-square error (aMSE)

$$\text{aMSE} = \frac{1}{N}\sum_{n=1}^{N} \|\mathbf{a}_n - \hat{\mathbf{a}}_n\|^2 \tag{9}$$

where $\mathbf{a}_n$ and $\hat{\mathbf{a}}_n$ are the actual and estimated abundance vectors associated with the nth pixel, respectively. Note that complementary results are reported in the supporting document [27].

Simulated data – To generate synthetic HS images, $K = 9$ endmembers are mixed according to the LMM. More precisely, these endmembers are randomly selected from the digital spectral library (splib06) [29] provided by the U.S. Geological Survey (USGS). This library contains spectrum of 498 materials in $L = 224$ bands ranging from 0.4 to 2.5μm. Abundance maps are randomly generated with smooth and piecewise smooth regions to mimic the expected spatial diversity of the HS scenes. The resulting 100×100 HS images ($N = 10^4$) are corrupted by an additive centered and white Gaussian noise whose variance is adjusted to reach SNR levels in the set $\{10\text{dB}, 20\text{dB}, 30\text{dB}, 40\text{dB}\}$. For the proposed WRNMF algorithm, the parameters in (4) and (6) have been empirically set to $\beta = 0.5$ and $\mu = 20$ to reach good results. A spatial neighborhood $\mathcal{V}(\cdot)$ in (7) of size 3×3 has been chosen. For all algorithms adopting the strategy in (3) to enforce the sum-to-one constraint, in particular FCLS and WRNMF, we set $\delta = 15$.

Table I report the aSAD and aMSE of all algorithms. These results have been averaged over 10 Monte Carlo runs for each noise level and obtained by adjusting the regularization parameters of each algorithm to reach the best performance. We can see that WRNMF performs better

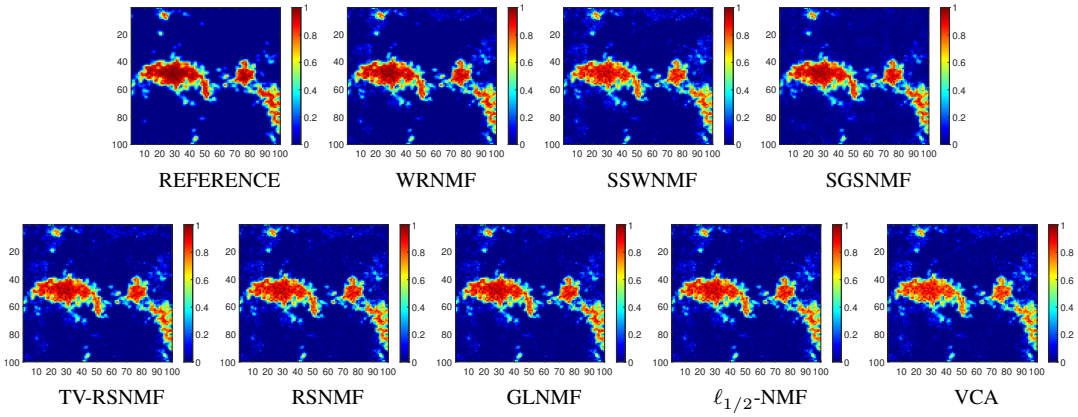

Fig. 1. Simulated data (SNR= 20dB): estimated abundance maps for endmember #9.

TABLE II
REAL URBAN DATA SET: SAD AVERAGED OVER 10 MONTE CARLO RUNS.

Material	WRNMF	SSWNMF	SGSNMF	TV-RSNMF	RSNMF	GLNMF	$L_{1/2}$-NMF	VCA-FCLS
Asphalt	**0.0803**	0.0821	0.0415	0.0790	0.1520	0.1515	0.1510	0.0943
Grass	**0.0689**	0.1002	0.1621	0.0972	0.1037	0.1032	0.1029	0.0989
Tree	0.0885	0.0769	0.0629	0.0615	0.0582	**0.0578**	0.0581	0.0751
Roof	0.0552	0.0792	0.0445	**0.0321**	0.0634	0.0637	0.0660	0.0439
Dirt	**0.0226**	0.0577	0.0289	0.0597	0.0615	0.0616	0.0626	0.1507
Mean SAD	**0.0631**	0.0792	0.0680	0.0659	0.0878	0.0876	0.0881	0.0926

than other algorithms for all noise levels. Moreover, even if the hyperspectral data is highly corrupted (SNR= 10dB), WRNMF is shown to provide better results than the compared algorithms in terms of aSAD and aMSE. Furthermore, the generated abundance maps estimated by all algorithms for SNR= 20dB are depicted in Fig. 1. We can observe that the abundance map obtained by WRNMF is more similar to reference map compared to other algorithms.

Real data – To illustrate, the real Urban dataset[1] has been unmixed by the compared algorithms. This image initially contains 210 bands covering the wavelength range 0.4 to 2.5μm with a spatial size of 307 $\times$ 307 pixels. However, only $L = 162$ bands have been kept by removing bands which are highly corrupted. We remove them before the analysis. We consider five types of signatures referred to as "Asphalt", "Grass", "Tree", "Roof" and "Dirt" and the reference spectral signatures can be obtained online[1]. We report the SAD scores associated to this image in Table II. It can be observed that WRNMF provides the lowest overall spectral distorsion over all compared algorithms. Fig. 3 shows the library signatures and the estimated endmember spectra obtained by the WRNMF. It can be observed that endmember signatures are close to library signatures. Moreover, qualitative results for abundance maps obtained by WRNMF are shown in Fig. 2 for materials identified as asphalt, grass, tree, roof and dirt. We can see that abundance maps obtained by WRNMF are in good agreement with the reference abundance maps.

[1]Available online at http://www.tec.army.mil/hypercube.

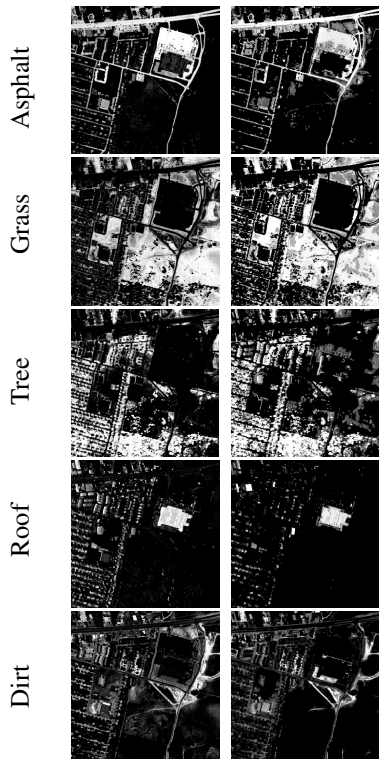

Fig. 2. Real Urban data set: reference (left) and abundance maps estimated by WRNMF (right).

IV. CONCLUSION

In this letter, we proposed an efficient yet simple unmixing method named WRNMF, which employed two weighting

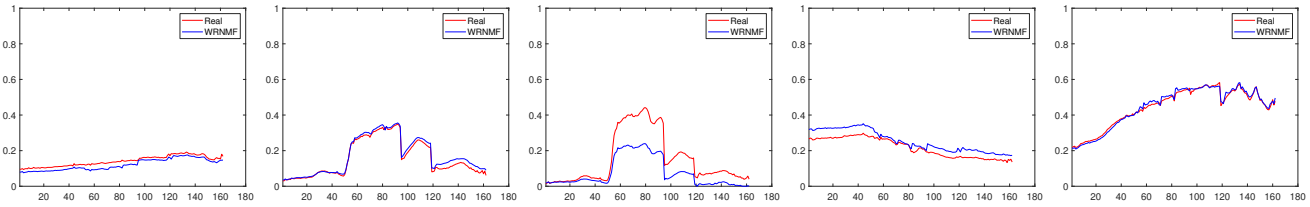

Fig. 3. Real Urban data set: library signatures and endmember estimates obtained by WRNMF. (a) Asphalt road. (b) Grass. (c) Tree. (d) Roof. (f) Dirt.

strategies. One of them applied to the residual term to assign different weights to the data fitting term across bands, which may account for different noise levels over the spectral range. The weighting strategy elaborated on a sparse regularization to adjust the spatial regularization of the abundance maps with respect to the spatial neighborhood. The resulting objective function is minimized by iteratively updating the endmember and abundance matrices using multiplicative rules, yielding a computationally efficient algorithm. The accuracy of the proposed method was assessed by comparing its performance with those of state-of-the-art algorithms when unmixing simulated and real hyperspectral images. WRNMF was observed to outperform recently NMF-based approaches even in the case where the hyperspectral data to be unmixed was highly corrupted.

REFERENCES

[1] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, April 2012.

[2] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 82–94, 2014.

[3] M. E. Winter, "N-findr: An algorithm for fast autonomous spectral endmember determination in hyperspectral data," in *Imaging Spectrometry V*, vol. 3753. International Society for Optics and Photonics, 1999, pp. 266–275.

[4] J. W. Boardman, F. A. Kruse, and R. O. Green, "Mapping target signatures via partial unmixing of AVIRIS data," in *Fifth JPL Airborne Earth Science Workshop*, vol. 95. JPL Publication, 1995, pp. 23–26.

[5] J. M. P. Nascimento and J. M. B. 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.

[6] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. 2nd Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, June 2010, pp. 1–4.

[7] M. D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 341–354, Jan 2014.

[8] T. Ince, "Superpixel-based graph Laplacian regularization for sparse hyperspectral unmixing," *IEEE Geosci. Remote Sens. Lett.*, pp. 1–5, 2020.

[9] ——, "Double spatial graph Laplacian regularization for sparse unmixing," *IEEE Geosci. Remote Sens. Lett.*, pp. 1–5, 2021.

[10] N. Dobigeon, J. Y. Tourneret, and C. I. Chang, "Semi-supervised linear spectral unmixing using a hierarchical bayesian model for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2684–2695, July 2008.

[11] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[12] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, March 2007.

[13] S. Jia and Y. Qian, "Constrained nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 161–173, Jan 2009.

[14] R. Huang, X. Li, and L. Zhao, "Spectral–spatial robust nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 8235–8254, 2019.

[15] A. Huck, M. Guillaume, and J. Blanc-Talon, "Minimum dispersion constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2590–2602, 2010.

[16] N. Wang, B. Du, and L. Zhang, "An endmember dissimilarity constrained non-negative matrix factorization method for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 554–569, 2013.

[17] M. D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, June 2011.

[18] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via $l_{1/2}$ sparsity-constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4282–4297, 2011.

[19] X. Lu, H. Wu, Y. Yuan, P. Yan, and X. Li, "Manifold regularized sparse nmf for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2815–2826, 2013.

[20] X. Wang, Y. Zhong, L. Zhang, and Y. Xu, "Spatial group sparsity regularized nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6287–6304, 2017.

[21] W. He, H. Zhang, and L. Zhang, "Total variation regularized reweighted sparse nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3909–3921, 2017.

[22] S. Zhang, G. Zhang, F. Li, C. Deng, S. Wang, A. Plaza, and J. Li, "Spectral-spatial hyperspectral unmixing using nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, pp. 1–13, 2021.

[23] C. Févotte and N. Dobigeon, "Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization," *IEEE Trans. Image Processing*, vol. 24, no. 12, pp. 4810–4819, Dec. 2015.

[24] W. He, H. Zhang, and L. Zhang, "Sparsity-regularized robust non-negative matrix factorization for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4267–4279, Sept 2016.

[25] J. Li, J. M. Bioucas-Dias, A. Plaza, and L. Liu, "Robust collaborative nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6076–6090, Oct 2016.

[26] C. Févotte and J. Idier, "Algorithms for nonnegative matrix factorization with the β-divergence," *Neural computation*, vol. 23, no. 9, pp. 2421–2456, 2011.

[27] T. Ince and N. Dobigeon, "Weighted residual NMF with spatial regularization for hyperspectral unmixing – Complementary results and supplementary materials," France, Tech. Rep., May 2022. [Online]. Available: http://dobigeon.perso.enseeiht.fr/papers/Ince_TechReport_2022.pdf

[28] D. C. Heinz and Chein-I-Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar 2001.

[29] R. N. Clark *et al.*, *USGS digital spectral library splib06a*. U.S. Geological Survey Denver, CO, 2007.