Joint Sparsity Model for Multilook Hyperspectral Image Unmixing

J. Bieniarz, E. Aguilera, X. X. Zhu, *Member, IEEE*, R. Müller, and P. Reinartz, *Member, IEEE*

***Abstract*—Recent work on hyperspectral image (HSI) unmixing has addressed the use of overcomplete dictionaries by employing sparse models. In essence, this approach exploits the fact that HSI pixels can be associated with a small number of constituent pure materials. However, unlike traditional least-squares-based methods, sparsity-based techniques do not require a preselection of endmembers and are thus able to simultaneously estimate the underlying active materials along with their respective abundances. In addition, this perspective has been extended so as to exploit the spatial homogeneity of abundance vectors. As a result, these techniques have been reported to provide improved estimation accuracy. In this letter, we present an alternative approach that is able to relax, yet exploit, the assumption of spatial homogeneity by introducing a model that captures both similarities and differences between neighboring abundances. In order to validate this approach, we analyze our model using simulated as well as real hyperspectral data acquired by the HyMap sensor.**

***Index Terms*—Joint sparsity, overcomplete spectral dictionary, spectral unmixing.**

I. INTRODUCTION

HYPERSPECTRAL pixels are usually modeled as a linear combination of spectral signatures typically related to pure materials. Each of these signatures, which are referred to as endmembers, is assumed to be weighted by the abundance of specific pure materials. The basic linear mixing model (LMM) can be formulated as follows:

$$y_j = Ax_j + \varepsilon_j \tag{1}$$

where $y_j \in \mathbb{R}^m$ stands for the measurements of the jth hyperspectal pixel with m spectral channels, $A \in \mathbb{R}^{[m\times n]}$ is the mixing matrix with n endmembers, x_j represents the abundance vector to be estimated for the jth pixel, and ε_j is a residual vector due to noise and modeling errors. The mixing matrix can be either trained from pixels containing pure material or formed by making use of existing spectral libraries containing field or laboratory spectrometer measurements [1].

Once the mixing matrix A has been constructed, the abundance vector x can be estimated from the measurement vector y. Common approaches include the method of least squares (LS), nonnegative LS (NNLS), and fully constrained LS (FCLS) [1]. While NNLS ensures the nonnegativity of the abundances, FCLS additionally enforces that the elements of the abundance vector sum up to one. However, these techniques are restricted to overdetermined mixing models that require a preselection of endmembers from a spectral library based on *a priori* knowledge or retrieving them from the image directly [1]. As a consequence, the amount of allowable endmembers is severely limited.

Recently, sparsity-based spectral unmixing methods have been proposed to estimate abundances in the presence of overcomplete spectral dictionaries [2]–[4]. These works make use of the fact that there are typically only few endmembers inside a pixel (depending on the pixel size) compared with the large endmember spectral library [2], i.e., the abundance vector to be estimated is sparse and has only up to few nonzero elements. This is usually done by solving sparse minimization problems using basis pursuit denoising (BPDN) or equivalent the least absolute shrinkage and selection operator (LASSO) [5], [6].

Iordache *et al.* [7] introduced the idea of collaborative sparse unmixing. In this approach, the unmixing is done by minimizing the residual error for all pixels in the image using the $L_{2,1}$ mixed norm as a regularizer. The mixed $L_{2,1}$ norm promotes sparsity among columns of the abundance matrix while minimizing the energy along the rows, i.e., the algorithm tends to select common endmembers for all pixels in the image. This approach provides good results for homogeneous areas with few active endmembers because the regularization is performed globally. The algorithm presented in [4] in addition to the simultaneous estimation of all abundances enforces the spatial contextual coherence between neighboring pixels. Another concept of sparse unmixing, i.e., sparse unmixing via variable splitting augmented Lagrangian and total variation (SUnSAL-TV), has been introduced in [8], where the L_2/L_1 minimizer from [2] is additionally regularized with a total variation accounting for smooth transitions in the abundances of the same endmember.

The goal of this letter is to exploit spatial homogeneity by applying a joint sparsity model (JSM) originally introduced in [9] for unmixing hyperspectral pixels. The JSM model allows for promoting sparse solutions across processed pixels such as in [7] by assuming spatial homogeneity between pixels, i.e., common endmembers, but additionally recovering sparse innovations in the abundance vector, such as additional endmembers and differences in the abundance. The basic inputs to the algorithm are then multiple neighboring pixels (hereinafter referred to as *looks*) as well as a spectral dictionary. Thus, the goal will be to take advantage of the fact that the ensemble

---

Manuscript received July 8, 2014; revised August 8, 2014; accepted August 27, 2014. This work is associated with the project SparsEO funded by Munich Aerospace e.V. Fakultat fr Luft-und Raumfahrt.

J. Bieniarz, X. X. Zhu, R. Müller, and P. Reinartz are with the Remote Sensing Technology Institute (IMF), Earth Observation Center (EOC), German Aerospace Center (DLR), 82234 Wessling, Germany.

E. Aguilera is with the Microwaves and Radar Institute (HR), German Aerospace Center (DLR), 82234 Wessling, Germany.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2014.2358623

of unknown vectors exhibit both common endmembers and abundances. As it has been shown, this approach also allows for capturing their noncommon endmembers. Finally, we validate the applicability of this technique comparing it with state-of-the-art methods (BPDN, NNLS, and SUnSAL-TV) using simulated and real data. Note that the method proposed in [7] is not to be considered, as it is unable to model common abundances locally.

II. METHODOLOGY

A. Sparse Spectral Unmixing

Let us assume an LMM as shown in (1). If the unmixing problem is underdetermined and the vector x_j for the jth pixel is expected to be sparse, i.e., the number of nonzero elements is much smaller than the length of x_j, then x_j can be estimated using L_1-regularized LS. Thus

$$\min_{x_j} \frac{1}{2}\|Ax_j - y_j\|_2^2 + \gamma\|x_j\|_1 \quad \text{subject to} \quad x_j \geq 0. \tag{2}$$

Since the L_1-norm serves as a proxy for sparsity, the penalty parameter γ can be used to trade sparsity of the solution for data mismatch. Moreover, in order to ensure physical validity, the abundance vector is constrained to be nonnegative. The minimization problem (2) is referred to as the nonnegative version of BPDN or equivalent sparse LASSO [6].

Spectral unmixing by means of (2) has been studied in [2] and [3]. These approaches are particularly useful when using spectral dictionaries containing a large number of endmembers. The L_1 term of the minimization (2) imposes a limit on the support of the x vector and, therefore, minimizes the number of active endmembers [1]. The probability of reliable sparse unmixing, however, highly depends on the properties of matrix A and the noise present in the spectrum [1]–[3], [7].

B. Joint Sparsity Model for Multilook Spectral Unmixing

In order to enhance the detection of the mixed endmembers and improve the performance in the presence of noise, we propose using an approach based on the JSM, which uses an ensemble of neighboring pixels in order to exploit spatial correlations between them. This concept is based on the fact that, usually, a pixel will contain a mixture of some spectral signatures from its neighbors, which is indeed a reasonable assumption due to, e.g., oversampling effects in line scanner images [10].

Given the reflectance values of the jth hyperspectral pixel y_j in m hyperspectral bands, the model for joint sparsity unmixing in a window with size of J pixels can be written as

$$Y_j = \tilde{A}X_j + Z_j \tag{3}$$

and $Z_j \in \mathbb{R}^{Jm}$ is an additive residual vector, where $Y_j \in \mathbb{R}_{\geq 0}^{Jm}$ is an ensemble of neighboring hyperspectral pixel, i.e.,

$$Y_j = (y_1, y_2, y_3, \ldots, y_J)^\top \tag{4}$$

and $\tilde{A} \in \mathbb{R}_{\geq 0}^{Jm\times(J+1)n}$ is a JSM dictionary [9] containing endmembers, and $X_j \in \mathbb{R}_{\geq 0}^{(J+1)n}$ is an abundance vector. Both $\tilde{A}$ and X_j are composed as follows:

$$\tilde{A}X_j = \begin{pmatrix} A & A & 0 & 0 & \cdots & 0 \\ A & 0 & A & 0 & \cdots & 0 \\ A & 0 & 0 & A & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ A & 0 & 0 & 0 & \cdots & A \end{pmatrix} \begin{pmatrix} x_c \\ x'_1 \\ x'_2 \\ \vdots \\ x'_J \end{pmatrix}. \tag{5}$$

In particular, the model for the jth pixel can be written as

$$y_j = Ax_j = Ax_c + Ax'_j \tag{6}$$

where x_c is a part of an abundance ensemble X_j containing only common abundance, and x'_j contains abundances present in the jth pixel with $1 \leq j \leq J$, A is the original dictionary containing n endmembers.

The design of matrix $\tilde{A}$ promotes joint sparsity. The first column of the dictionary $\tilde{A}$ is responsible for recovery of the common accounts for x_c contained in all J pixels. The diagonal part of $\tilde{A}$ reconstructs the endmembers, which are not present in all pixels and are referred to as innovations. As a result, the multilook JSM reconstruction (MLJSR) can be accomplished by means of the previously introduced nonnegative version of BPDN, i.e.,

$$\min_{X_j} \frac{1}{2}\|\tilde{A}X_j - Y_j\|_2^2 + \gamma\|X_j\|_1 \quad \text{subject to} \quad X_j \geq 0 \tag{7}$$

where, again, γ is responsible for trading sparsity for data mismatch. In light of previous considerations, we will provide a numerical evaluation of MLJSR based on (7).

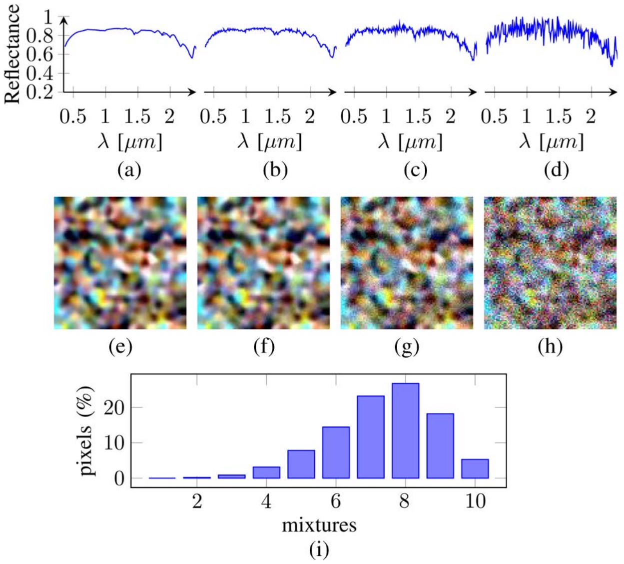
III. EXPERIMENTAL RESULTS

We have tested our algorithm using simulated and real hyperspectral data. The algorithm least angle regression/LASSO used for optimization problems (2) and (7) is reported in [5]. We considered two sliding windows for MLJSR, namely, a square-shaped (with $J = 9$) and a cross-shaped (with $J = 5$) window. For both, the unmixed pixel y_j was located in the center of the window. Additionally, we compared the results with the SUnSAL-TV algorithm using the solver from [11].

A. Simulated Scenario

In this experiment, we used $n = 240$ spectra selected from the United States Geological Survey Digital Spectral Library (splib06) [12] to form the dictionary $A \in \mathbb{R}^{[m\times n]}$ with $m = 224$ spectral channels such as in [8]. The simulated images were created as follows.

1) Form a mixing matrix $M \in \mathbb{R}^{[m\times n']}$ by selecting $n' = 10$ endmembers at random from A.
2) Create an abundance cube $X'(c, r, n')$, where columns c and rows r are $c = r = 100$. For all layers n', randomly set 144 elements to 1.
3) Blur all layers n' using a Gaussian filter with standard deviation $\sigma = 0.025$.



(a) (b) (c) (d)

(e) (f) (g) (h)

(i)

Fig. 1. Simulated data. Subfigures (a)–(d) present a central pixel spectrum from the simulated image with (a) no noise and with SNR of (b) 40 dB, (c) 30 dB, and (d) 20 dB. Similarly, the false color composition of bands 50, 100, 150 contaminated with different noise levels is shown in (e)–(h). The bar plot (i) shows the distribution of pixels according to the number of endmembers mixed in it.

4) Normalize $\forall\, c, r, n' : X'(c, r, n') = X'(c, r, n') / \sum_n X'(c, r, n')$.
5) Create a hyperspectral data cube $Y'(c, r, m)$ by means of LMM $\forall\, c, r : Y'(c, r, m) = MX'(c, r, n') + \varepsilon'(c, r, m)$, where $\varepsilon'(c, r, m)$ is Gaussian noise with a signal-to-noise ratio $\mathrm{SNR[dB]} = 10 \log_{10} \|Y'(c, r, m)\|_2^2 / \|\varepsilon'(c, r, m)\|_2^2$.

In our simulations, we considered three test cases with SNR values of 20, 30, and 40 dB [see Fig. 1(a)–(h)]. The simulated images contained mixtures of from two to ten endmembers distributed as shown in Fig. 1(i). In order to evaluate the abundance recovery performance, we considered two measures. First, the signal-to-reconstruction error (SRE) is defined as

$$\mathrm{SRE[dB]} = 10 \log_{10} \left(\frac{1}{M} \sum_{j=1}^{M} \frac{\left\| x_j^o \right\|_2^2}{\left\| x_j^o - x_j \right\|_2^2} \right) \tag{8}$$

where x_j^o is the ground truth abundance of the jth x_j pixel, and M is the total number of pixels. Second, the mean absolute error (MAE) is given as

$$\mathrm{MAE}_j = \frac{1}{n} \left\| x_j^o - x_j \right\|_1 \tag{9}$$

where n is the number of endmembers.

For each test case, we calculated solutions using NNLS, BPDN, SUnSAL-TV, and MLJSR with 60 uniformly distributed γ parameter settings such that $\gamma = e^i,\ i \in \langle -1, 11 \rangle$. Additionally, since SUnSAL-TV requires two regularization parameters, we have tested the second parameter λ_{TV} [7] with values 0.0001, 0.0005, 0.001, 0.005, 0.01, and 0.05 resulting in 360 different parameter settings for each test case.

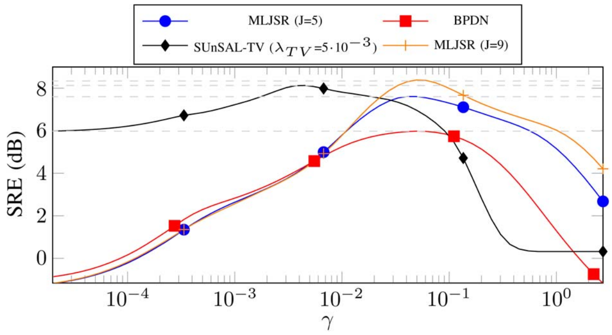


Fig. 2. Plot of the SRE values as a function of the γ regularization parameter. Results for SNR = 30 dB using BPDN, SUnSAL-TV, and MLJSR methods.

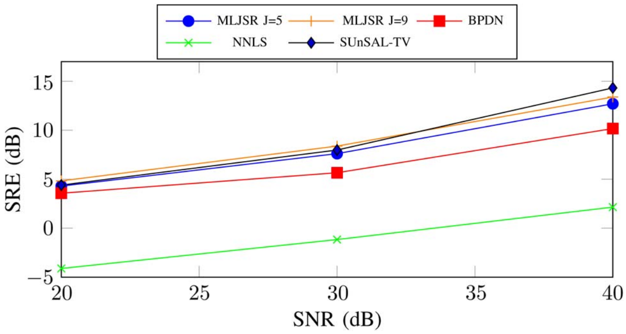
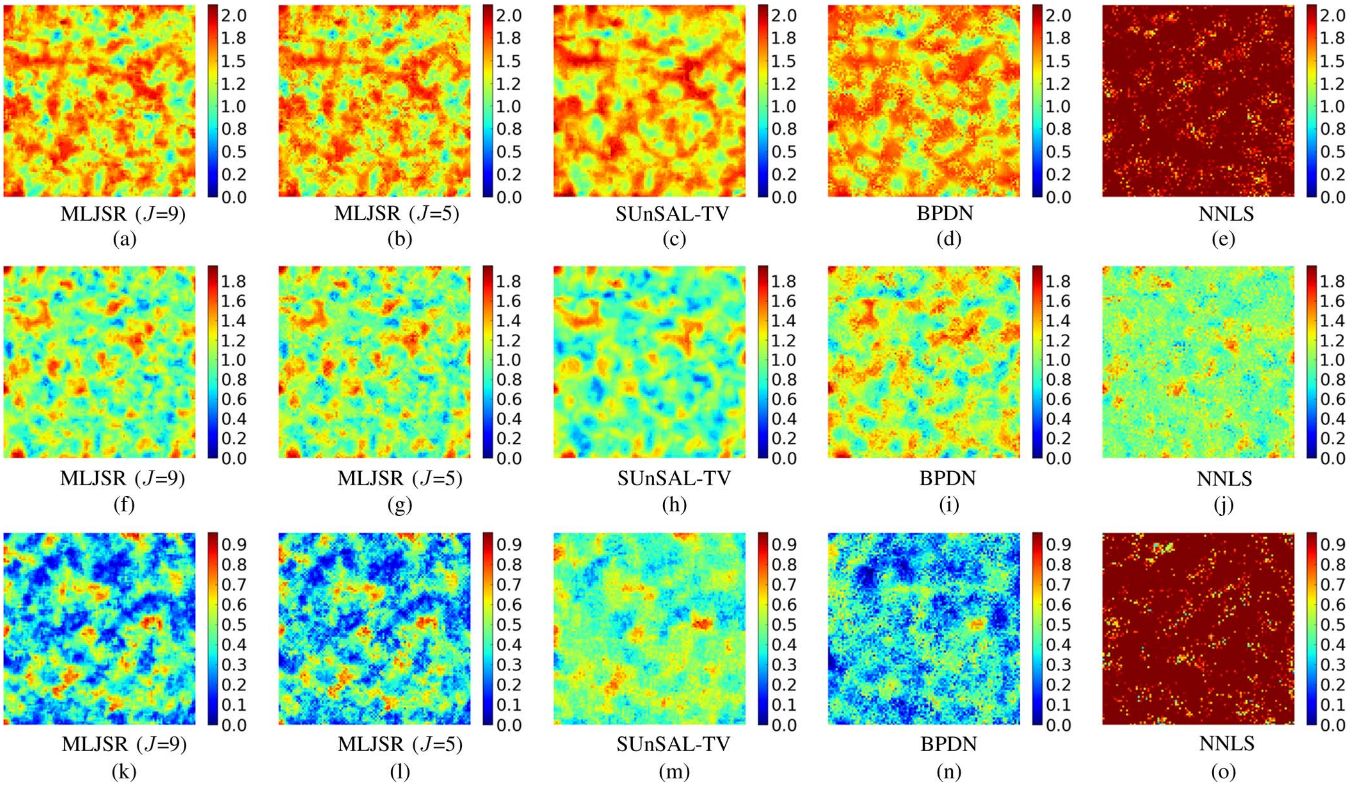


Fig. 3. Comparison of NNLS, BPDN, SUnSAL-TV, and MLJSR algorithms using SRE measure for signal with different SNR values.

Fig. 2 shows the SRE for different γ settings using the MLJSR, BPDN, and SUnSAL-TV methods. In this particular case, MLJSR exhibits better performance than BPDN and SUnSAL-TV. In Fig. 3, the mean SRE is shown as a function of SNR. For all SNR values, the γ parameter has been chosen from the set of solutions so as to favor each particular algorithm (see, for example, Fig. 2). In terms of SRE, both settings of the MLJSR algorithm outperformed NNLS and BPDN at all noise levels. MLJSR with $J = 9$ window is reported to have higher SRE than SUnSAL-TV for low SNR values of 20 and 30 dB.

We have used the MAE to evaluate the accuracy of abundance estimation at the pixel level for whole abundance vectors [see Fig. 4(a)–(e)], only abundance of present endmembers [see Fig. 4(f)–(j)], and false recovery of nonpresent endmembers, i.e., endmembers present in recovered abundance but not present in the ground truth abundance [see Fig. 4(k)–(o)]. For the 20-dB test case, all L_1-regularized algorithm exhibit significantly lower overall MAE than NNLS (mean MAE for the whole test image was equal to 1.43 for MLJSR with $J = 9$, 1.44 for MLJSR with $J = 5$, 1.47 for SUnSAL-TV, 1.49 for BPDN, and 3.42 for NNLS). Both settings of MLJSR (mean MAE $= 0.38$) along with BPDN (mean MAE $= 0.34$) have lower MAE for nonpresent endmembers when compared with SUnSAL-TV (mean MAE $= 0.48$) and NNLS (mean MAE $= 2.36$). All algorithms were able to recover abundances of present endmembers with relatively similar MAE.

We have compared the computation time for all algorithms implemented in MATLAB using Intel(R) Core(TM)2 Quad CPU Q9550 at 2.83 GHz with 8 GB of RAM. In this experiment, MLJSR and BPDN are solved with [5], SUnSAL-TV using [11], and NNLS with MATLAB built-in function *lsqnonneg*. The computation time for the 30-dB test case resulting in highest SRE was, respectively, 24.97 s using MLJSR ($J = 9$, $\gamma = 0.05$), 39.04 s using MLJSR ($J = 5$, $\gamma = 0.049$), 382.17 s using SUnSAL-TV ($\gamma = 0.0045$, $\lambda_{\mathrm{TV}} = 0.005$), 4.12 s using BPDN ($\gamma = 0.05$), and 101.5 s using NNLS.



MLJSR (J=9) (a) MLJSR (J=5) (b) SUnSAL-TV (c) BPDN (d) NNLS (e)

MLJSR (J=9) (f) MLJSR (J=5) (g) SUnSAL-TV (h) BPDN (i) NNLS (j)

MLJSR (J=9) (k) MLJSR (J=5) (l) SUnSAL-TV (m) BPDN (n) NNLS (o)

Fig. 4. Comparison of NNLS, BPDN, SUnSAL-TV, and MLJSR algorithms using the SNR $= 20$ dB test case and the MAE measure for (a)–(e) all endmembers, (f)–(j) endmembers present in the image only, and (k)–(o) endmembers not present in the image (false abundance recovery). The color bar range, for all subfigures in a row, has been set at maximum MAE of one of the L_1-regularized methods.

B. Real HyMap Image

For the experiment with real data, we used a hyperspectral scene acquired over Munich, Nymphenburg district, with the airborne sensor HyMap [13]. The scene has a size of 100×100 pixels with originally 126 spectral channels reduced to 96 after discarding noisy and absorption bands. The image has been atmospherically [14] and geometrically [10] corrected with a ground resolution of 4 m. To form a spectral library, we used 90 man-made materials, vegetation, and soil spectra selected from the splib06 [12], merged with 39 endmembers from a roof-material dictionary [15]. The dictionary had finally $m = 129$ endmembers.

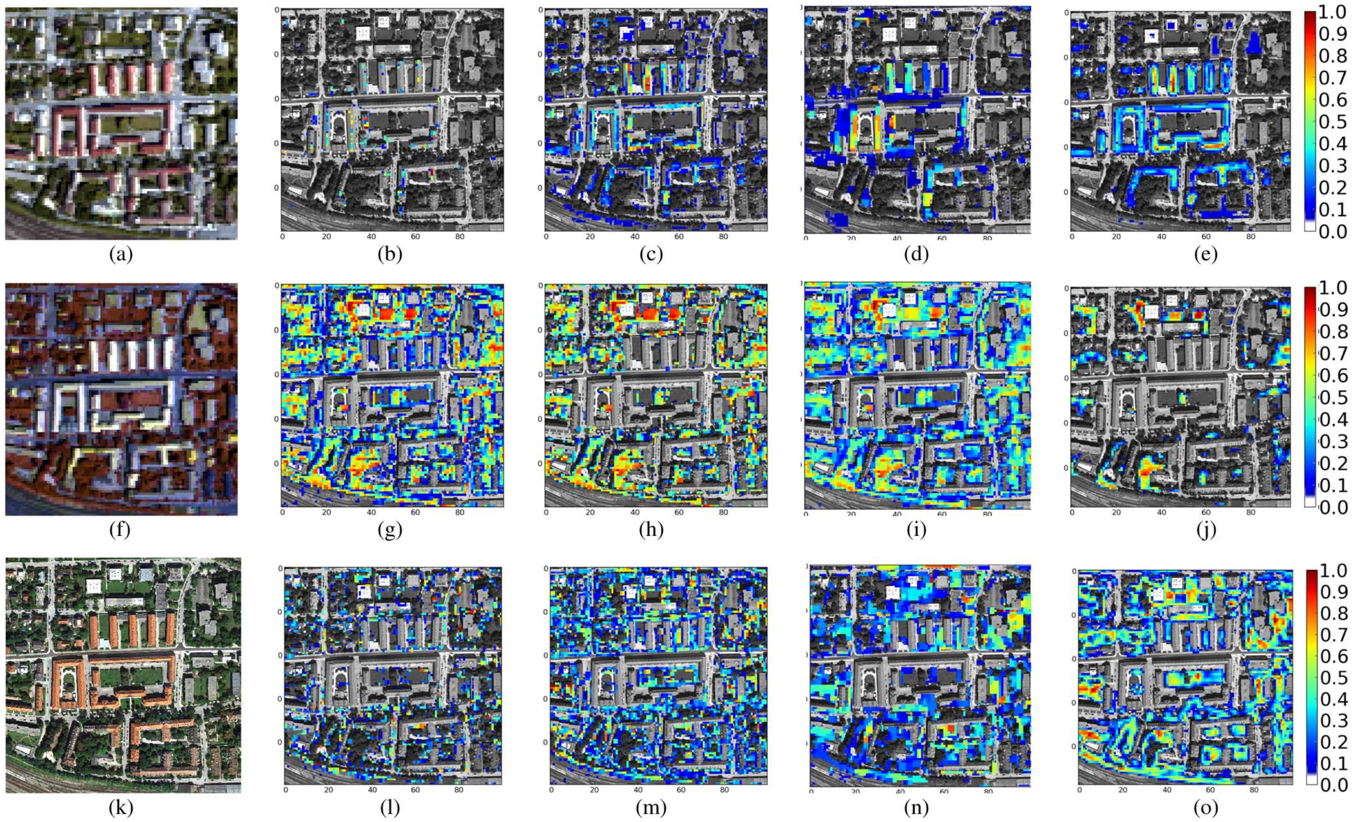
We have performed the reconstruction of abundances by means of NNLS; BPDN with $\gamma = 0.5$; SUnSAL-TV with $\gamma = e^{-3}$ and $\lambda_{\mathrm{TV}} = e^{-3}$ [8]; and MLJSR with $\gamma = 1.5$ and $J = 9$. Again, we have chosen the regularization parameters so as to favor each particular algorithm and, thus, allow for a fair comparison. All unmixing results have been overlaid on a high-resolution image (see Fig. 5), henceforth employed as a pseudo-ground truth. NNLS, BPDN, SUnSAL-TV, and MLJSR have detected, respectively, 107, 22, 98, and 38 active materials. Fig. 5(a)–(k) shows the estimated abundances for red roof shielding, lawn grass, and trees, respectively.

When compared with BPDN and NNLS, the MLJSR algorithm leads to smoother transitions from one material to another, which is consistent with the assumption of localized spatial homogeneity. For example, Fig. 5(d) shows a gradual decrease in the roof abundance from the middle of the roof toward the edges. Noticeably, this is not the case for SUnSAL-TV [see Fig. 5(c)], which tends to exhibit very similar neighboring abundances. Also, whereas BPDN [see Fig. 5(b)] is unable to detect the roof edges, NNLS misses a great deal of roof endmembers [see Fig. 5(a)]. Moreover, NNLS seems to provide reliable results only for abundances associated with tree [see Fig. 5(e)]. Finally, note that, while NNLS [see Fig. 5(i)] and BPDN [see Fig. 5(j)] fail for grass detection, MLJSR provides satisfactory results [see Fig. 5(k)].

IV. CONCLUSION

In this letter, we have proposed a multilook joint sparse reconstruction for sparse spectral unmixing. The proposed model takes advantage of the neighboring information by means of joint approximation of abundances in a sliding window and particularly a joint approximation of common endmembers.

Compared with the standard sparse BPDN or NNLS method, experiments with simulated data demonstrate that the proposed method restores abundances for pixels sharing common endmembers more accurately in terms of the SRE measure and overall MAE. Compared with the SUnSAL-TV method, our approach exhibited higher SRE for SNR values equal to 20 and 30 dB. The fact that our method requires only one regularization parameter, it is easier to use compared with SUnSAL-TV. Additionally, MLJSR in both settings performs faster than SUnSAL-TV or NNLS. The outcome of MAE comparison shows that the MLJSR approach recovers endmembers with similar accuracy as state-of-the-art algorithms



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

(k) (l) (m) (n) (o)

Fig. 5. Unmixing results using NNLS [(b), (g), (l)], BPDN [(c), (h), (m)], SUnSAL-TV [(d), (i), (m)], and MLJSR [(e), (j), (o)], respectively. The first row of images presents the abundance map for red roof shielding. In the second row are the abundances for the tree spectrum, and in third row are the abundances for lawn grass. (a) RGB composition (channels 7, 5, 3) from the HyMAP image, (f) false RGB composition (channels 96, 60, 16), (k) high-resolution Google Earth Image of the corresponding area.

with low detection of nonpresent endmembers. The qualitative analysis of the results on the HyMAP image shows the MLJSR approach to be competitive with respect to BPDN and NNLS or SUnSAL-TV methods and visually provides more consistent results.

For the future work, we address tests on real hyperspectral data including quantitative analysis, different ensemble sizes and shapes, and dependence from the spatial resolution. The JSM can be extended to a more complex design, e.g., by incorporating shared endmembers between singular pixels in the window by weighting of the central pixel or by different weighting factors for the pixels according to the point spread function of the hyperspectral sensor.

ACKNOWLEDGMENT

The authors would like to thank Dr. U. Heiden for providing the roof spectral library.

REFERENCES

[1] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[2] 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, Jun. 2011.

[3] J. Bieniarz, R. Müller, X. X. Zhu, and P. Reinartz, "On the use of overcomplete dictionaries for spectral unmixing," presented at the 4th WHISPERS, Jun. 2012.

[4] W. Tang, Z. Shi, and Y. Wu, "Regularized simultaneous forward–backward greedy algorithm for sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5271–5288, Sep. 2014.

[5] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *J. Mach. Learn. Res.*, vol. 11, pp. 19–60, Mar. 2010.

[6] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Statist.*, vol. 32, no. 2, pp. 407–499, 2004.

[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] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, Nov. 2012.

[9] D. Baron, M. F. Duarte, S. Sarvotham, M. B. Wakin, and R. G. Baraniuk, "An information theoretic approach to distributed compressed sensing," in *Proc. Allerton Conf. Commun., Control, Comput.*, 2005, pp. 1–12.

[10] R. Müller *et al.*, "A program for direct georeferencing of airborne and spaceborne line scanner images," in *Proc. ISPRS Commission I Symp.; Integrating Remote Sens. Global, Regional Local Scale*, Nov. 2002, pp. 148–153.

[11] J. M. Bioucas-Dias, 2012. [Online]. Available: http://www.lx.it.pt/~bioucas/code/demo~sparse~tv.rar

[12] R. N. Clark *et al.*, "Usgs digital spectral library splib06a," in *Digital Data Series 231*. Reston, VA, USA: USGS, 2007.

[13] T. Cocks, R. Jenssen, A. Stewart, I. Wilson, and T. Shields, "The HyMap (TM) airborne hyperspectral sensor: The system, calibration, and performance," presented at the 1st EARSeL Workshop Imag. Spectroscopy, 1998.

[14] R. Richter and D. Schlpfer, "Atmospheric/Topographic correction for airborne imagery," DLR, Cologne, Germany, DLR-IB 565-02/13, 2011.

[15] U. Heiden, K. Segl, S. Roessner, and H. Kaufmann, "Determination and verification of robust spectral features for an automated classification of sealed urban surfaces," in *Proc. 4th EARSeL Workshop Imag. Spectroscopy Warsaw*, 2005, pp. 151–159.