

Hyperspectral unmixing with material variability using social sparsity

Travis Meyer, Lucas Drumetz, Jocelyn Chanussot, Andrea Bertozzi, Christian Jutten

▶ **To cite this version:**

Travis Meyer, Lucas Drumetz, Jocelyn Chanussot, Andrea Bertozzi, Christian Jutten. Hyperspectral unmixing with material variability using social sparsity. ICIP 2016 - 23rd IEEE International Conference on Image Processing, IEEE, Sep 2016, Phoenix, United States. 10.1109/ICIP.2016.7532746 . hal-01364247

HAL Id: hal-01364247
https://hal.science/hal-01364247

Submitted on 12 Sep 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HYPERSPECTRAL UNMIXING WITH MATERIAL VARIABILITY USING SOCIAL SPARSITY

Travis R. Meyer, Lucas Drumetz†, Jocelyn Chanussot†, Andrea L Bertozzi*, Christian Jutten†*

* University of California, Los Angeles - Department of Mathematics. Los Angeles, California, USA
† GIPSA-Lab, Grenoble Alpes University, France

ABSTRACT

We apply social ℓ-norms for the first time to the problem of hyperspectral unmixing while modeling spectral variability. These norms are built with inter-group penalties which are combined in a global intra-group penalization that can enforce selection of entire endmember bundles; this results in the selection of a few representative materials even in the presence of large endmembers bundles capturing each material's variability. We demonstrate improvements quantitatively on synthetic data and qualitatively on real data for three cases of social norms: group, elitist, and a fractional social norm, respectively. We find that the greatest improvements arise from using either the group or fractional flavor.

***Index Terms*—** Remote sensing, hyperspectral imaging, endmember variability, unmixing, endmember bundles

1. INTRODUCTION

Spectral variability [1, 2] arises during hyperspectral unmixing due to each material having potentially multiple associated spectral signatures. As an example, lighting variation can cause the spectral signature to be scaled by a factor. To compensate for this, it is necessary to develop methods that allow for material endmember variability.

Vertex component analysis (VCA) [3], a technique for hyperspectral endmember extraction, can be used to capture spectral variability found in hyperspectal imagery (HSI) [4]. By extracting the endmembers as learned with VCA, from random subsets of pixels in the image, one produces endmembers that capture variability in the imaged materials. The selection of pixels is related to material variability – sampling small subsets can capture more variability while sampling large percentages of the data captures only slight variation. VCA must have pure pixels to discover endmembers, an assumption which some methods do not assume [5]. This means too small of a sampling subset risks missing desired pure endmembers.

With extracted endmembers, each ideally representing a pure signature resulting from one state of one material, a standard constrained fit to these endmembers for each pixel produces an unmixing result. What remains is to identify which endmembers correspond to the same material by clustering the endmembers into material groups. The final abundance map for each material is then the sum of contributions from each endmember in that material's cluster.

This is made concrete in the following way: Let H be the hyperspectral data matrix which is s-by-n, for s the number of spectral bands and n the number of image pixels. Each of k times, for some input k, we use VCA to extract m endmembers from a random sampling of $x\%$ of the columns of H for some x. This produces vectors which form the columns of a matrix M, the s-by-km endmember matrix.

Next, a clustering technique is applied to the columns of M to produce m clusters resulting in a partition $G_1, G_2, ..., G_m$ of the endmembers such that $\{M_{:,j} \ \forall \ j \in G_i\}$ forms signatures corresponding to the i^{th} material. In the work of [4] endmembers are clustered with k-means, however in this work we use instead the more robust technique of spectral clustering [1] [6]. We consider the case of scaling first studied by [7] which can be a large source of variability e.g. due to illumination differences caused by terrain slope. The spectral angle measure naturally identifies signatures with this type of variability and is often applied to signature similarity [8, 9]. Let $||\cdot||_F$ denote the Frobenius matrix norm. The second part of the problem assumes a linear mixture model and solves the optimization problem

$$\min_X ||H - MX||_F^2 \tag{1}$$

subject to X non-negative with unit-sum columns, a technique known as fully constrained linear spectral unmixing (FCLSU) [10]. This provides the abundance matrix X that contains the abundance of each endmember for each material in each pixel. Summing the rows of X corresponding to like-partitioned elements of G results in the abundance of the materials at each pixel. The matrix M has km columns, which means even for small k there are considerably more columns than materials. The presence of many representatives for each material means the fit (1) frequently produces

[1] By spectral clustering, it is meant the grouping of the extracted endmembers into several subsets that are spectrally homogeneous, assuming that a given subset will correspond to all the different instances of the extracted endmembeers corresponding to one given material.

a mixture of nearly every material in each pixel. A natural way to prevent this is for example using sparsity [11, 12]. We generalize the technique of [13] to enforce this sparsity across bundles with social norms.

2. PROPOSED METHOD

We propose for the first time incorporating the material clustering information into (1) by using social norms described in [14]. Because G is known prior to solving (1), the assumption of each pixel being a mixture of a few materials can be enforced through penalty. Let $x \in \mathbf{R}^s$ be some vector and partition the set $\{1, 2, ..., km\}$ into m groups G_i for $i = 1, 2, ..., m$, as was done in the previous section using spectral clustering with the spectral angle measure. The vector ℓ_p norm of x is given by

$$||x||_p = \left(\sum_{i=1}^{s} |x_i|^p \right)^{\frac{1}{p}}$$

with the cases of $p = 1$ and $p = 2$ very common, and limiting behaviors of $p \to 0$ and $p \to \infty$ approaching the number of non-zero entries of x and the maximum absolute value of all entries, respectively. The group pq-norm given the partition G is given by

$$||x||_{G,p,q} = \left(\sum_{i=1}^{m} ||x_{G_i}||_p^q \right)^{\frac{1}{q}}. \tag{2}$$

This can be generalized to matrices by summing the application of (2) to each column which, for a penalty parameter λ, results in the proposed model

$$\min_X \frac{1}{2}||H - MX||_F^2 + \lambda \sum_{i=1}^{n} ||X_{:,i}||_{G,p,q}. \tag{3}$$

The cases we consider are the group lasso $(p, q) = (2, 1)$, elitist lasso $(p, q) = (1, 2)$, and a fractional case $(p, q) = (1, \frac{9}{10})$. Group lasso tends to select a few groups, in our case materials, and within groups it prefers a dense mixture of members. The elitist lasso selects a dense mixture over the groups and within each group selects a few representative "elites". The final fractional lasso selects a few groups, similar to the group lasso, but it does so without preferring a dense mixture over groups. Because (3) is subject to unit-sum columns, the addition of these penalties can appear contradictory. For example, if each material has a single endmember the group lasso has no influence. This constraint also makes the traditional sparsity-enforcing lasso penalty nonsensical; a fraction of $9/10$ can enforce sparsity even with constrained abundance at the expense of non-convexity.

3. ALGORITHM

In this section we solve (3) using the alternating direction method of multipliers, or ADMM [15], that allows the constraints and complex penalty term to be split into distinct and easily calculable stages. To solve (3), first write the problem in a slightly different way. Consider the optimization problem

$$\min_{X,Y,Z} \frac{1}{2}||H - MX||_F^2 + \lambda||Y||_{G,p,q} \tag{4}$$

subject to the constraints that $X = Z$, $X = Y$, X has unit-sum columns and Z is non-negative. This equivalent problem is a ready form for ADMM with multiplier variable matrices α and β that produce the augmented Lagrangian formulation

$$\begin{aligned} \min_{X,Y,Z} \frac{1}{2}||H - MX||_F^2 + \lambda||Y||_{G,p,q} \\ + \langle \alpha, X - Z \rangle + \frac{\rho}{2}||X - Z||_F^2 \\ + \langle \beta, X - Y \rangle + \frac{\rho}{2}||X - Y||_F^2 \end{aligned} \tag{5}$$

subject to the constraints of X with unit-sum columns and Z non-negative.

Z Subproblem – Minimization with respect to Z is trivial as the problem is separable for each coordinate of the matrix, hence Z is updated by

$$Z \leftarrow \left(X + \frac{\alpha}{\rho} \right)_+$$

with $(\cdot)_+$ indicating the coordinate-wise positive part.

X Subproblem – The update for X is slightly more difficult due to the unit-sum constraint, however this only requires adding n multipliers μ_i

$$\begin{aligned} \min_{X,Y,Z} \frac{1}{2}||H - MX||_F^2 + \lambda||Y||_{G,p,q} \\ + \langle \alpha, X - Z \rangle + \frac{\rho}{2}||X - Z||_F^2 \\ + \langle \beta, X - Y \rangle + \frac{\rho}{2}||X - Y||_F^2 \\ + \sum_j \mu_j \left(\left(\sum_i X_{i,j} \right) - 1 \right) \end{aligned} \tag{6}$$

which results in a linear system of equations given by

$$(M^T M + 2\rho I) \begin{pmatrix} X \\ \mu \end{pmatrix} = \begin{pmatrix} -M^T H + \alpha + \beta - \rho(Z + Y) \\ 1 \end{pmatrix} \tag{7}$$

where μ indicates the row vector of entries μ_i and 1 indicates a matrix of ones with the same shape as μ. This system is easily invertible and, conveniently, the system matrix is state-independent.

Y Subproblem – The last closed-form subproblem is that of Y which requires the use of a group shrinkage operation $\mathcal{S}_{G,p,q}$, described in [14] with approximate fractional *p-shrinkage* as used in [16], denoted by

$$Y = \mathcal{S}_{G,p,q}\left(X + \frac{1}{\rho}\beta, \frac{\lambda}{\rho}\right). \tag{8}$$

3.1. Final Iterative Scheme

We have described the solution to each subproblem. The full iterative scheme, along with dual updates, is given by:

- Initialize $\alpha \leftarrow 0, \beta \leftarrow 0$.
- Randomly initialize X then set $Y \leftarrow X, Z \leftarrow X$.
- Repeat to convergence:
  - $Z \leftarrow \left(X + \frac{\alpha}{\rho}\right)_+$.
  - $X \leftarrow$ the solution of (7).
  - $Y \leftarrow \mathcal{S}_{G,p,q}\left(X + \frac{1}{\rho}\beta, \frac{\lambda}{\rho}\right)$.
  - $\alpha \leftarrow \alpha + \rho(X - Z)$.
  - $\beta \leftarrow \beta + \rho(X - Y)$.

For our data sets we use $\rho = 10$ with the iteration requiring on the order of 1000 iterations. The scheme is relatively quick, taking only minutes to unmix data of size 100-by-100-by-56. The major challenge is the sum constraint which conflicts with the sparsifying penalties resulting in slower convergence.

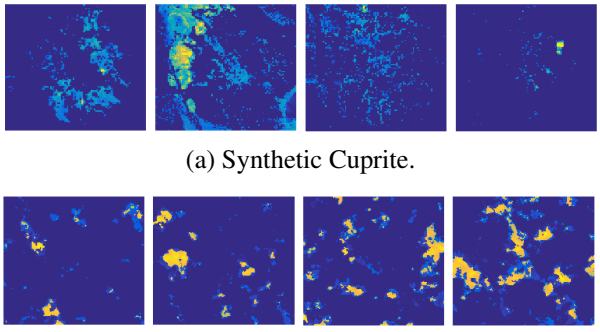
4. EXPERIMENTAL STUDY

4.1. Data Sets

Set	Group	Elitist	Fractional (0.9)	Batchless
Cuprite	88.2%	96.4%	93.2%	203.5%
Islands	94.9%	*94.4%	97.8%	565.0%

Table 1. Mean pixel errors for each of three types of sparsity enforcement algorithms and FCLSU without bundles applied to two synthetic data sets. Shown is the error as a percent of the unpenalized bundle model. (*) performance achieved with a negative value of λ.

We applied our methods to three data sets, two synthetic with known abundances and one real data set with unknown abundances. In figure 1 we show example true abundance maps. The first set is "cuprite" that was generated using the AVIRIS cuprite data set via the method proposed in [17] with artificial spectral variability as described in [18]. This technique uses real spectral signatures and abundance maps based on real data to generate a hyperspectral image with synthetic variability. For both synthetic data sets there are approximately three materials in each image pixel, and the data cube has been reduced by a factor of two due to large performance costs in optimizing λ for the experiments. The third data set, shown in figure 3, is an aerial image taken of a stadium with surrounding parking lots, a parking structure, and trees among other features. For these data sets we use $k = 5$ and learn m endmembers from k randomly selected subsets of 80% of the data. For each subset VCA is run for 100 iterations.



(a) Synthetic Cuprite.

(b) Synthetic Islands.

Fig. 1. Synthetic abundance maps: the (a) cuprite synthetic data has a 100-by-100-by-47 data cube and the (b) islands synthetic data has a 100-by-100-by-56 hyperspectral cube. Shown are the known exact abundance maps for four materials out of a respective total of 14 and 15 simulated materials.



Fig. 2. Real data set: the third data we evaluate the proposed models with is a 105-by-128-by-144 hyperspectral cube taken of the since demolished Robertson stadium at the University of Houston. Shown is the approximate color image using three bands near red, green, and blue.

4.2. Results

To quantify performance we compute, for each method, a mean pixel error

$$\mathrm{E}_{\mathrm{Model}} = \frac{1}{\#\ \mathrm{Pixels}} \sum_{\mathrm{Pixels}\ i} \sqrt{\frac{1}{\#\ \mathrm{Materials}} ||a_i - \tilde{a}_i||_2^2}$$

with a_i and $\tilde{a}_i$ the actual and approximate abundance maps for each pixel i. For each method we use the error ratio

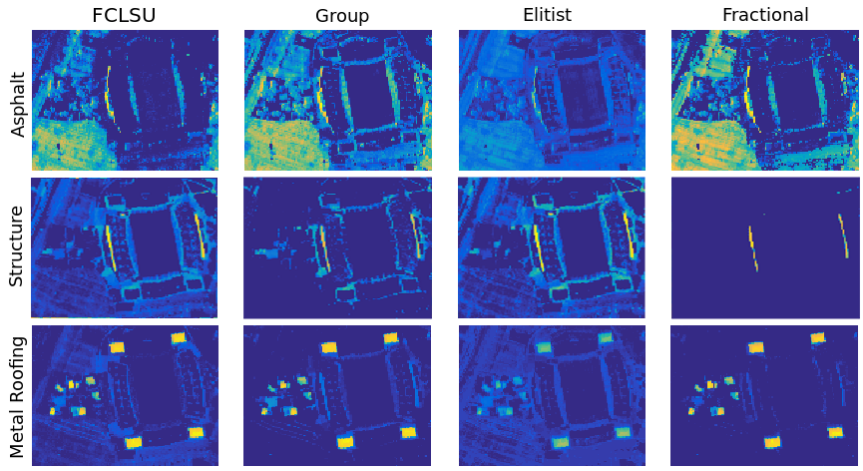


Fig. 3. Material abundance maps corresponding to (top row) asphalt, (middle row) a painted structure, and (bottom row) metal roofing. Shown is the result for each of unpenalized FCLSU, group lasso, elitist lasso, and fractional lasso.

E_{Model}/E_{FCLSU} to quantify improvement over the unpenalized bundle method (1). Table 1 shows this improvement optimized for each method over λ. As a control we include the "batchless" model – the model produced by extracting a single endmember for each material using VCA on the entire data set. Quite evidently, the group lasso performs consistently the best, followed by fractional and lastly elitist, with all new models outperforming batch FCLSU and all models outperforming the batchless control. In the case of the elitist lasso, the beneficial value for λ was negative for the islands data. The effect of a negative λ acts similarly to the fractional lasso by preferring abundance concentrated in a few material groups, as opposed to the dense mixture of the elitist lasso. In the synthetic cases the error is quite low with a value on the order of 10^{-3} with most improvements resulting from intra-material confusion caused by noise.

Qualitatively, results on the stadium are promising. We selected values of λ for each method which appear to improve the unmixing result without heavy distortion. Here we demonstrate the differences for three materials: paint on two buildings, metal roofing, and asphalt as shown in figure 3.

For the two painted buildings and the metal roofing the group and fractional penalties produce clearly distinct regions of the image that are consistent with local structure, a promising result given that no spatial information is used by the model. The elitist lasso does not follow this trend as it appears to prefer spreading information across material bundles. The asphalt is a less clear result, though we note that noise present in the grass field is reduced and removed by both the group and fractional penalties.

5. CONCLUSIONS AND PERSPECTIVES

We have shown that using the material clustering information in the unmixing step is improved by using social norms. We presented an efficient method for accomplishing this that is based on ADMM which breaks the optimization task into fast and approachable subproblems. In all experiments the use of group, elitist, or fractional lassos was able to improve on the baseline model with no penalty term which indicates that these reflect the desired assumption of material selection when using bundles to capture spectral variability. The group and fractional lasso had the best performance while the elitist lasso showed the relatively smallest improvement and unclear results. Questions remain as to the relative effectiveness of the different penalties, for example the use of a negative elitist lasso or other fractional lassos than $p = 9/10$. Furthermore, combinations of these norms would potentially improve the model further by more delicately capturing the intra-bundle abundance statistics at the cost of an additional model parameter. We note, however, that the study of these many variations becomes dependent on the validity of labeled data, which is itself a challenge to come by. A further limitation of these methods is endmember selection based on VCA applied to random subsets of pixels. Because material variability is spatially correlated, using a spatially correlated technique for subset selection such as [19, 20] could improve endmember quality.

6. ACKNOWLEDGEMENTS

This work was partially funded by the European Research Council (ERC) under grant ERC AdG-2012-320684 CHESS, NSF grant DMS-1118971, NSF grant DMS-1417674, ONR grant N000141210838, UC Lab Fees Research grant 12-LR-236660, and by DGA under grant 2015 60 0012 00.470.75.01.

7. REFERENCES

[1] Alina Zare and KC Ho, "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *Signal Processing Magazine, IEEE*, vol. 31, no. 1, pp. 95–104, 2014.

[2] Abderrahim Halimi, Nicolas Dobigeon, and Jean-Yves Tourneret, "Unsupervised unmixing of hyperspectral images accounting for endmember variability," *arXiv preprint arXiv:1406.5071*, 2014.

[3] José MP Nascimento and José M Bioucas Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 4, pp. 898–910, 2005.

[4] Ben Somers, Maciel Zortea, Antonio Plaza, and Gregory 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, 2012.

[5] Javier Plaza, Eligius MT Hendrix, Inmaculada García, Gabriel Martín, and Antonio Plaza, "On endmember identification in hyperspectral images without pure pixels: A comparison of algorithms," *Journal of Mathematical Imaging and Vision*, vol. 42, no. 2-3, pp. 163–175, 2012.

[6] Andrew Y Ng, Michael I Jordan, Yair Weiss, et al., "On spectral clustering: Analysis and an algorithm," *Advances in neural information processing systems*, vol. 2, pp. 849–856, 2002.

[7] José MP Nascimento and José M Bioucas Dias, "Does independent component analysis play a role in unmixing hyperspectral data?," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 1, pp. 175–187, 2005.

[8] Nirmal Keshava and John F Mustard, "Spectral unmixing," *Signal Processing Magazine, IEEE*, vol. 19, no. 1, pp. 44–57, 2002.

[9] José M Bioucas-Dias, Antonio Plaza, Nicolas Dobigeon, Mario Parente, Qian Du, Paul Gader, and Jocelyn 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, 2012.

[10] Chein-I Chang, *Hyperspectral imaging: techniques for spectral detection and classification*, vol. 1, Springer Science & Business Media, 2003.

[11] Feiyun Zhu, Ying Wang, Bin Fan, Shiming Xiang, Geofeng Meng, and Chunhong Pan, "Spectral unmixing via data-guided sparsity," *Image Processing, IEEE Transactions on*, vol. 23, no. 12, pp. 5412–5427, 2014.

[12] Chengbo Li, Ting Sun, Kevin F Kelly, and Yin Zhang, "A compressive sensing and unmixing scheme for hyperspectral data processing," *Image Processing, IEEE Transactions on*, vol. 21, no. 3, pp. 1200–1210, 2012.

[13] Yuntao Qian, Sen Jia, Jun Zhou, and Antonio Robles-Kelly, "Hyperspectral unmixing via sparsity-constrained nonnegative matrix factorization," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 11, pp. 4282–4297, 2011.

[14] Matthieu Kowalski, Kai Siedenburg, and Monika Dorfler, "Social sparsity! neighborhood systems enrich structured shrinkage operators," *Signal Processing, IEEE Transactions on*, vol. 61, no. 10, pp. 2498–2511, 2013.

[15] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[16] Rick Chartrand, "Fast algorithms for nonconvex compressive sensing: Mri reconstruction from very few data," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*. IEEE, 2009, pp. 262–265.

[17] Zhipeng Hao, Mark Berman, Yi Guo, Glenn Stone, and Iain Johnstone, "Semi-realistic simulations of natural hyperspectral scenes," in *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*. IEEE, 2015, pp. 1004–1007.

[18] Lucas Drumetz, Simon Henrot, Miguel Angel Veganzones, Jocelyn Chanussot, and Christian Jutten, "Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability," in *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS 2015)*, 2015.

[19] Miguel A Veganzones, Guillaume Tochon, Mauro Dalla-Mura, Antonio J Plaza, and Jocelyn Chanussot, "Hyperspectral image segmentation using a new spectral unmixing-based binary partition tree representation," *Image Processing, IEEE Transactions on*, vol. 23, no. 8, pp. 3574–3589, 2014.

[20] Yoann Altmann, Marcelo Pereyra, and Jose Bioucas-Dias, "Collaborative sparse regression using spatially correlated supports-application to hyperspectral unmixing," *arXiv preprint arXiv:1409.8129*, 2014.