

# ON THE USE OF OVERCOMPLETE DICTIONARIES FOR SPECTRAL UNMIXING

*Jakub Bieniarz, Rupert Müller, Xiaoxiang Zhu, Peter Reinartz*

German Aerospace Center (DLR),  
Remote Sensing Technology Institute (IMF)  
Wessling, 82234, Germany

## ABSTRACT

Hyperspectral unmixing is a sub pixel classification method which aims at recovering fraction and type of materials mixed in a single pixel. This work addresses the unmixing problem from the compressive sensing point of view by using overcomplete dictionaries enabling automatization of the process. However, overcomplete dictionaries of spectra are highly coherent which might confuse the final unmixing result. To deal with this problem we propose the use of differentiated spectra for coherence reduction. In this paper we study the approximation error for the proposed method as well as the correctness of the material detection.

**Index Terms**— Hyperspectral image, unmixing, sparse approximation, derivative, coherence.

## 1. INTRODUCTION

Hyperspectral image data evaluation is one of the remote sensing methods with increasing amount of publications on new applications, methods and algorithms. The main advantage of hyperspectral imagery is the possibility of sampling each pixel with hundreds of contiguous bands spread over a wide spectral range. This relatively new remote sensing technique takes advantage of classical spectroscopy enabling detailed spectral investigation for each pixel including material detection and quantitative analysis of material content for each pixel. However, on the other hand high spectral resolution and huge amount of channels result in reduction of the spatial resolution. Modern spaceborne hyperspectral cameras have ground pixel sizes  $\geq 30$  m. This relatively low spatial resolution leads to mixing of several sources (materials) within one pixel. This makes detailed spectral analysis of the hyperspectral image difficult. In remote sensing the basic sources are pure material spectra which are known in literature as endmembers.

The method for recovery of endmembers and estimation of their amount for each pixel is called spectral unmixing. Recently, spectral unmixing is intensively studied by many researchers resulting in many methods where the endmembers can be retrieved from the image in the manner of geometrical approaches (vertex component analysis - VCA, pixel purity

index - PPI, N-FINDR, iterative error analysis - IEA), statistical approaches (independent component analysis ICA-EA [1]) or manual selection from existing databases of endmembers. The quantitative approximation of the abundances in the HSI data is often regarded as least squares problem. While many existing algorithms are excellent tools for the community and experienced users these methods can however work only with a limited number of endmembers restricted by the dimensionality of the HSI data. If the number of endmembers exceeds the dimensionality of the spectral feature space, fewer equations than unknowns lead to an infinite number of solutions for this problem. To deal with this shortcomings, spectral unmixing can be also addressed as a sparse approximation problem. If we assume that the mixing of endmembers is a linear process, then each measured pixel can be described as a product of a known dictionary and a sparse coefficients vector using  $\ell_1$  norm as minimizer. In this case the dictionary is overcomplete, which means that it consists of more endmembers than the dimensionality of the signal. Dictionaries can be created from existing databases containing hundreds of spectra of different materials, measured using field or laboratory spectrometers. Iordache et al. [2] tested several sparse approximation methods for spectral unmixing with different preselected spectral libraries. In [3] the authors present new algorithms for sparse regression SUNSAL and C-SUNSAL dedicated for solving underdetermined unmixing problems.

In this work we present sparse approximation methods with large overcomplete dictionaries focusing on coherence reduction between the endmembers. We discuss a novel usage of spectral derivatives to increase the detection rate. We test our approach for both, synthetic and real data.

## 2. METHODS

In this section we describe the method used for our experiments with focus on sparse approximation for spectral unmixing and material detection. We also describe spectral derivative as a dictionary coherence reduction method and its application.

## 2.1. Linear spectral unmixing

Usually most HSI pixels are mixed i.e. they consist of signals from more than one source (endmember). The mixing process can be modeled as a transformation  $T$  over the matrix  $M^{[L \times N]}$  containing  $N$  endmembers to describe the pixel  $y$

$$y = T\{M\} + \varepsilon, \quad (1)$$

where  $\varepsilon$  is an additive noise and  $L$  is the dimensionality of the pixel  $y$ . If we assume that the transformation  $T$  is linear i.e. only single scattering occur for each detected photon, then the mixing model can be formulated as

$$y = Ma + \varepsilon, \quad (2)$$

where  $a \geq 0$  is a coefficient vector which defines the fractional abundance of each material [1]. Linear spectral unmixing aims recovery of the endmembers and their abundances contained in a pixel.

## 2.2. Sparse Approximation

### 2.2.1. Spectral unmixing as a compressive sensing problem

In order to automate the unmixing process we could use a large overcomplete mixing matrix containing all available spectra (e. g. measured in field or in laboratory). If each column of the mixing matrix is normalized to  $\ell_2$  unit length it is called dictionary  $D$  with the columns as atoms  $\varphi$ . Then a hyperspectral pixel  $y^{[L \times 1]}$  can be interpreted as the measurement of a sparse signal. Hence, the redundant dictionary  $D^{[L \times N]}$  is in this case a measurement matrix,

$$y = Dx, \quad (3)$$

and  $x^{[N \times 1]}$  is the coefficient vector. Through  $D$  with  $N > L$ , the system of equations is underdetermined and has normally an infinite number of solutions. Therefore, instead of solving  $y = Dx$  we can find such a solution  $x$  that minimizes  $\|Dx - y\|_2$  where  $\|\cdot\|_2$  is the euclidean  $\ell_2$  norm with  $\|x\|_p = \sqrt[p]{\sum |x_i|^p}$ .

Since we expect that each pixel is a mixture of only few endmembers we can look for a sparse approximation of  $x$ . If we denote by  $\|x\|_0 = k$  the pseudo-norm  $\ell_0$  indicating the number of non-zero elements of the vector  $x$ , then the  $x$  vector is  $k$ -sparse.

$$\min \|x\|_0 \text{ s.t. } \|Dx - y\|_2 \leq \xi, \quad (4)$$

where  $\xi > 0$  is the tolerance value. The solution of the above problem is called the sparse solution to equation. 3.

### 2.2.2. Basis Pursuit

Since equation 4 is a non-convex, combinatorial optimization problem, it is difficult to solve. Instead we can use the  $\ell_1$

norm as a minimizer which also promotes the sparsest solution. This minimization problem can be written as

$$\min \|x\|_1 \text{ s.t. } \|Dx - y\|_2 \leq \xi, \quad (5)$$

known as basis pursuit (BP) [4].

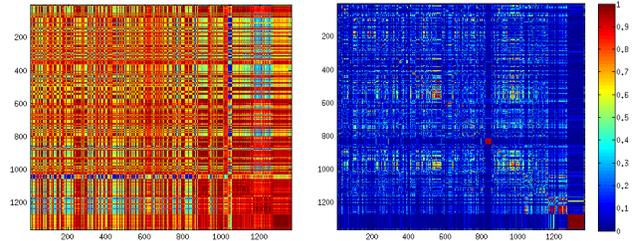
### 2.2.3. Coherence of the dictionary

BP does not recover the solution for problem 4 directly, but under certain condition it is able to recover the sparsest solution for the underdetermined system. One of the most important condition in sparse approximation methods is the mutual coherence. Mutual coherence is a measure of maximum coherence between pairs of atoms  $\varphi$  of the dictionary.

$$\mu = \max_{i \neq j} |\varphi_i^T \cdot \varphi_j|. \quad (6)$$

However, mutual coherence is a worst-case measure, in principle the smaller the coherence between the atoms, the sparser the abundance vector which can be approximated using the  $\ell_1$  minimization [5].

## 2.3. Derivative of spectrum



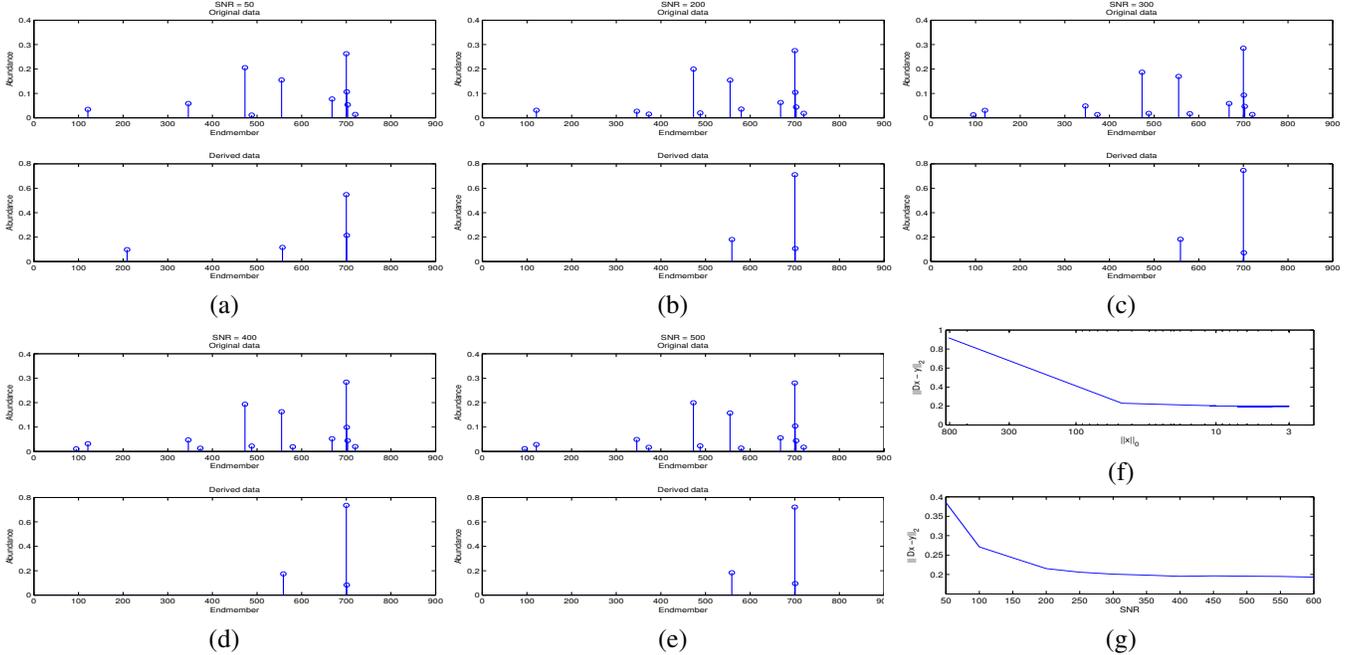
(a) Original dictionary. (b) Transformed dictionary.

**Fig. 1.** Coherence matrices for the dictionaries: a) containing original spectra and b) containing derivatives of original spectra.

The differentiation of the function estimates the slope over the changing independent variable. In our case the independent variable is the band number. We calculate the derivative of a spectrum in the following way

$$\frac{\partial y}{\partial b_i} = \frac{y(b_i) - y(b_j)}{\Delta b}, \quad (7)$$

where  $b$  is a hyperspectral band,  $\Delta b = b_i - b_j$  and  $b_i > b_j$ . The differentiation of spectra does not result in more information that is contained in the original bands but it is possible to decrease background reflectance and can therefore considerably improve the detection of convoluted weaker absorption features. We calculate the derivatives for each atom in the library. The comparison of the coherence matrices for both the original and the transformed dictionary using the derivative method can be seen on fig. 1. The derivative methods are further described in [6, 7].



**Fig. 2.** Results for the test with simulated hyperspectral pixel. Sub-figures (a-e) show approximation of the abundance vector  $x$ . Upper plot is for original data, lower for differentiated data. Plots (a-e) are composed of two different equally distributed spectra (no. 559 - 50% and 700 - 50%) SNR: 50, 200, 300, 400, 500. (f) is a plot of ratio between sparsity and approximation error. (g) shows approximation error for different SNR.

### 3. EXPERIMENTAL RESULTS AND DISCUSSION

We have tested our method with simulated as well as real AVIRIS data [8]. The algorithm has been assessed for the detection rate, the  $\ell_2$  approximation error for different signal to noise ratios (SNR), and also how the  $\ell_2$  error changes with respect to the sparsification of the abundance vector.

For the experiment we use the USGS spectral library [9]. The library consists of 1365 spectra, but due to missing readings of many spectra, and incorrect values we have reduced the size of the library atoms to 813. Each atom was re-sampled to 224 bands according to the AVIRIS airborne sensor. After band selection the dictionary size is  $D^{[154 \times 813]}$ .

#### 3.1. Simulated data

Simulated pixels have been created using spectra from the dictionary consisting of equally distributed two endmembers, randomly selected from the dictionary. To test a further parameter different SNR have been added. We have calculated abundances for the test pixels for both, original spectra and derivative of the spectrum. Results for this test are shown in fig. 2 (a-e). Since the selected spectra (Kaoline - no. 559 and Plastic PETE - no. 700) have high coherence (0.90) to each other and to neighborhood it causes confusion in the approximation of the coefficients. Coherence between the both derivatives of the spectra is 0.25. Consequently we test the

simulated pixels for the  $\ell_2$  approximation error for different SNR see fig. 2 (g). The error for SNR = 50 is 0.39 and it consistently decreases with reduction of the noise to 0.18 for SNR = 600. The figure 2 (f) shows the relationship between the sparsity and the  $\ell_2$  approximation error. This results have been calculated for different  $\xi$  tolerance values to indicate the best ratio between the sparsity and the approximation error. Using derivatives the sparsest solution is  $\|x\|_0 = 3$  and the residual  $\|Dx - y\|_2$  is 0.18, while for the original data  $\|x\|_0 = 13$  and the  $\ell_2$  approximation error is equal 0.02. Our method indicated only one wrong endmember (no. 701) for almost all test cases, which has very high coherence of 0.99 to the endmember no. 700. Only for very low SNR = 50 (fig. 2 (a)) our method indicates two wrong endmembers (no. 211 and no. 701), which is not possible to eliminate with the  $\ell_1$  penalty due to relatively high coefficients. However, the use of derivatives increases the detection reliability. But it has to be noted that differentiating the signal enhances not only spectral features but also amplifies the noise.

#### 3.2. AVIRIS data

For this test we have used freely available AVIRIS hyperspectral image data [8]. From this image we have selected a region of  $612 \times 512$  pixels with 154 out of 224 spectral bands. The abundances for each differentiated pixel have been approximated using BP technique. The elements of the dictionary

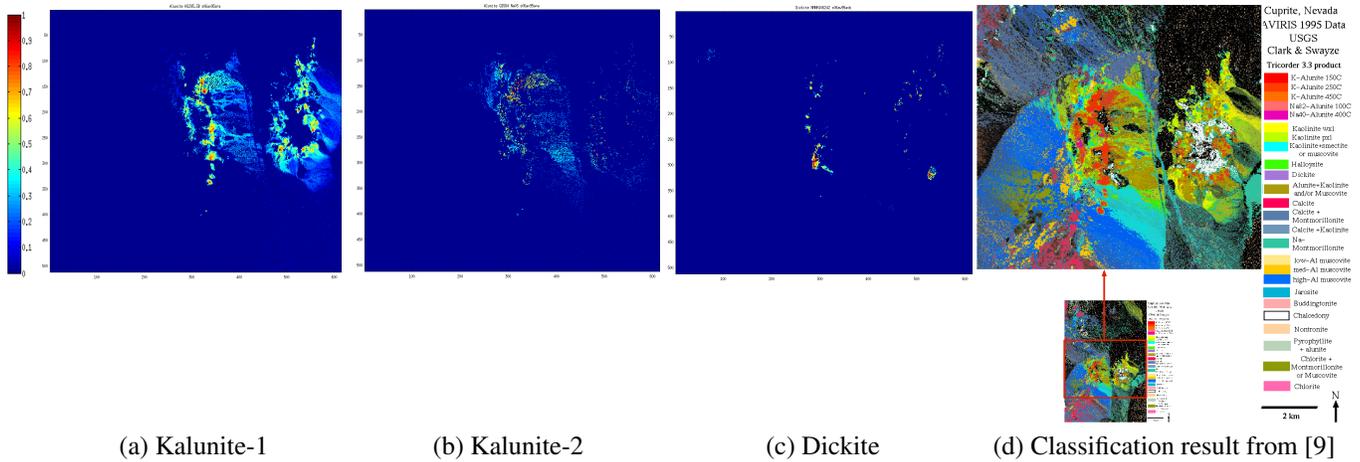


Fig. 3. Abundance maps created using the proposed method.

$D^{[154 \times 813]}$  have been differentiated, too. Mean  $\|x\|_0$  for all pixels in the image is about 3.06. In this study case we concentrated on material detection and approximation of abundance values. Here we imply that all materials on ground are available in the dictionary. Selected abundance maps are shown in fig. 3. This experiment shows feasibility of use of the proposed algorithm with the real data. Further quantitative analysis is addressed as future work.

#### 4. CONCLUSIONS

In this work we propose to use sparse approximation methods with large overcomplete dictionaries for spectral unmixing and material detection. To deal with detection confusion caused by high coherence of the dictionary, we introduce derivatives of spectra. This transformation of the signal significantly decreases overall coherence of the dictionary in comparison to the original spectra.

Experiments on simulated data show that the reduced coherence of differentiated dictionary enables sparser approximation and thus smaller detection confusion. However, use of the derivative method increases the approximation error. Hence, the improvement of the trade off between error and detection of actual endmembers is an important issue for future work.

We also show the applicability of the proposed method using real hyperspectral image data producing reasonable sparse results.

#### 5. REFERENCES

[1] J.M. Bioucas-Dias and A. Plaza, "An overview on hyperspectral unmixing: Geometrical, statistical, and sparse regression based approaches," in *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, July 2011, pp. 1135–1138.

[2] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 6, pp. 2014–2039, 2011.

[3] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," *ArXiv e-prints*, Feb. 2010.

[4] Scott Shaobing Chen, David L. Donoho, Michael, and A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, pp. 33–61, 1998.

[5] Jean-Luc Starck, Fionn Murtagh, and Jalal M. Fadili, *Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity*, New York : Springer, 2010.

[6] Gerhard Talsky, *Derivative Spectrophotometry of First and Higher Orders*, VCH Verlagsgesellschaft mbH, Weinheim (Federal Republic of Germany) VCH Publishers, Inc., New York, NY (USA), 1994.

[7] Fuan Tsai and William Philpot, "Derivative analysis of hyperspectral data," *Remote Sensing of Environment*, vol. 66, pp. 4151, October 1998.

[8] JPL, "<http://aviris.jpl.nasa.gov/html/aviris.freedata.html>," 1996.

[9] Roger N. Clark, Gregg A. Swayze, K. Eric Livo, Raymond F. Kokaly, Steve J. Sutley, J. Brad Dalton, Robert R. McDougal, and Carol A. Gent., "Imaging spectroscopy: Earth and planetary remote sensing with the usgs tetra-corder and expert systems," *Journal of Geophysical Research*, vol. 108, pp. 5131, 2003.