

**Military Technical College  
Kobry El-Kobbah,  
Cairo, Egypt**



**11<sup>th</sup> International Conference  
on Electrical Engineering  
ICEENG 2018**

## **Implementation of Sparsity-Based Unsupervised Unmixing of Hyperspectral Imaging Data using Coiflet**

A. Elrewayny\*, Sherif S. Sherif \*\*

### **ABSTRACT**

Unsupervised spectral unmixing is one of the most important processing tasks performed on hyperspectral imaging data. It is a blind source separation problem where every spectral pixel in the hyperspectral data cube is separated into a set of pure spectra, i.e., endmembers, without any prior knowledge about them. This problem could be solved sparsely using the *basis pursuit* optimization problem which is a well-known technique used for sparse source recovery. The *basis pursuit* considers the endmembers sparse in a basis known as dictionary. The contribution of this paper is the fast implementation of the *basis pursuit* unmixing algorithm due to using the Coiflet orthogonal dictionary. Using Coiflet orthogonal dictionary results in sparser spectral pixels which lead to fast computation by reducing the iterations of the *basis pursuit* unmixing algorithm. The *basis pursuit* unmixing algorithm using Coiflet was applied to a synthetic data cube acquired from few materials selected from the given ASTER spectral library.

### **KEY WORDS**

Basis Pursuit (BP), Coiflet, Daubechies, Hyperspectral Imaging (HSI), Spectral Unmixing (SU), Wavelets.

---

\* Egyptian Armed Forces.

\*\* Dept. of Electrical and Computer Engineering, University of Manitoba, Canada.

## 1. Introduction

Hyperspectral imaging (HSI) systems acquire images using hundreds of contiguous wavelengths in potentially different electromagnetic bands [1]. Hyperspectral cameras typically have low spatial resolutions that usually provide the acquired spectral pixel as mixture of spectra of pure materials present in the scene called endmembers. Unsupervised unmixing does this task without any prior knowledge about the nature of the scene nor the present endmembers.

Every spectral pixel in the hyperspectral data cube is separated into a set of endmembers [2]. We model the spectral pixels in the data cube using a linear mixture model such that:

$$Y = SA + E \quad (1)$$

where  $Y = [y_1, \dots, y_N]$  is  $L \times N$  matrix represents the spectral pixels,  $L$  is the number of wavelengths,  $N$  is the number of spectral pixels in the data cube,  $S = [s_1, \dots, s_M]$  is  $L \times M$  matrix represents the endmembers spectra,  $M$  indicates the number of endmembers exist in the scene,  $A = [a_1, \dots, a_N]$  is  $M \times N$  matrix represents all the abundances (proportions) of the endmembers in every pixel such that the abundance vector  $a_i$  for every pixel is subject to the nonnegativity and the sum-to-one constraints and  $E$  represents an additive perturbation (noise). The problem of unsupervised spectral unmixing is to estimate the endmember matrix  $S$  and the abundance matrix  $A$  given the spectral pixels matrix  $Y$ . This is considered as a Blind Source Separation (BSS) problem [3].

Sparse signal restoration algorithms assume that the unknown signal is sparse in an appropriate domain called dictionary. So, signal sparsity could be used as prior information to obtain an estimate of the unknown signal. This problem is termed *basis pursuit* problem [4]. The *basis pursuit* optimization problems could be used in the hyperspectral unmixing problem. In [5], the *basis pursuit* unmixing problem is solved using proximal methods which in the case of  $l_1$ -norm problems will tend to iterative thresholding [6], but they didn't choose any specified dictionary.

The choice of the appropriate dictionary is important for obtaining sparser representation for better unmixing results. Two main approaches of dictionaries for sparse representation; analytic dictionaries approach and learned custom dictionaries approach [7]. In the analytic dictionary approach, the dictionary is described by a transform family. In the learned custom dictionaries approach, the dictionary is derived from the given measurements.

Due to the physical nature of the spectral pixels which have a small number of peaks corresponding to resonant absorption properties, the wavelet is chosen as an analytic dictionary for the sparsity representation of the spectral pixels. Studying the properties of the Coiflet gives us the motivation of using it as an analytic dictionary for promoting the sparsity of the unknown endmembers in solving the *basis pursuit* unmixing problem.

This paper is arranged as follows; Section 2 illustrates the *basis pursuit* unmixing algorithm. The choice of Coiflet is discussed in section 3. Section 4 shows convergence of the *basis pursuit* unmixing algorithm using three different analytic dictionaries.

## 2. Sparsity-based unmixing algorithm

As a preprocessing step, we can transform the problem into a different domain using the dictionary  $\Phi$  for promoting the sparsity of  $S$  which could be represented in  $\Phi$  as:

$$S = \Phi\alpha \quad (2)$$

where  $\alpha$  is the coefficients of  $S$  in  $\Phi$ . This step is considered as a warm start to the optimization problem to promote the sparsity of  $S$ . The data  $Y$  could be represented in the dictionary  $\Phi$  obtaining the matrix  $\beta = \Phi^T Y$ , the matrix where each of its columns stores the coefficients of each transformed pixel  $y_i$ . The unmixing problem could be formulated as constraint minimization problem by adding the sparsity penalty term:

$$\arg \min_{\alpha, A} \frac{1}{2} \|\beta - \alpha A\|_2^2 + \lambda \|\alpha\|_1 \quad (3)$$

where  $\lambda$  is a regularization parameter that quantifies the tradeoff between the two terms of the equation.

Indeed, as  $N \gg M$ , it turns out that (3) is a multichannel overdetermined linear system. The solution is obtained by an iterative and alternate estimation of  $\alpha$  and  $A$  [6]:

- Update the coefficients: solving  $\min_{\alpha} |A$  when  $A$  is fixed, this is the *basis pursuit* problem.  $\alpha$  is obtained using the proximal method which is ‘hard thresholding’ in this case [8]:

$$\alpha = \text{Thresh}_{\delta}(A^{\dagger}\beta) \quad (4)$$

where  $A^{\dagger}$  is the pseudo-inverse of the current estimate of  $A$ ,  $\text{Thresh}_{\delta}$  is a thresholding operator and the threshold  $\delta$  decreases with increasing iteration count.

- Update the mixing matrix: solving  $\min_A |_{\alpha}$  when  $\alpha$  is fixed.  $A$  is obtained using the least-squares method with the constraints of the abundance matrix in the linear mixture model:

$$A = (\alpha^T \alpha)^{-1} \alpha^T \beta \quad (5)$$

The two stages iterative process leads to the solution of (3). After convergence and obtaining the coefficient matrix  $\alpha$  and the abundance matrix  $A$ , the spectra of the endmembers  $S$  could be restored using the same dictionary  $\Phi$  using  $S = \Phi\alpha$ .

### 3. Coiflet dictionary for spectral unmixing

The wavelet transform is an excellent basis to sparsely represent 1-D smooth signals having a small number of irregular points. The wavelet family was chosen as our potential analytic dictionary because spectral pixels are smooth signals and the physical nature of it indicates that these spectral pixels have a small number of peaks corresponding to resonant absorption peaks.

The vanishing moments of a wavelet function  $\psi$  which is an important property of wavelet:

$$\int x^l \psi(x) dx = 0 \quad \text{for } 0 \leq l \leq m \quad (6)$$

A wavelet function has  $m$  vanishing moments can be represented as the  $m^{\text{th}}$  order derivative of a given function  $\theta$ . Thus the resulting wavelet transform is equivalent to a multiscale differential operator and it kills its peaks [9].

$$\psi(t) = (-1)^m \frac{d^m \theta(t)}{dt^m} \quad (7)$$

So, wavelet with a large number of vanishing moments gives sparser representation for most signals. However, there is a trade-off between the number of vanishing moments of  $\psi$  and the support in the time domain of  $\psi$  [9]. The support of a wavelet in the time domain and its number of vanishing moments are independent. But, for an orthogonal discrete wavelet  $\psi$ , if it has  $m$  vanishing moments, then its support in the time domain will be  $2m - 1$ . The Daubechies wavelet have minimum support in the time domain for a given number of vanishing moments; that's why it's considered optimal.

Coiflet was developed by I. Daubechies at the request of R. Coifman. In addition to its wavelet function  $\psi$  with order  $n$  that has  $2n$  vanishing moments, their scaling function  $\varphi$ , also has  $2n - 1$  vanishing moments. Both wavelet function,  $\psi$ , and scaling function,  $\varphi$ , have the minimum possible support in the discrete time domain of  $6n - 1$ . Therefore, among all wavelet families with a given support in the discrete time domain, Coiflet would has the largest number of vanishing moments for both  $\psi$  and  $\varphi$ . This makes the representation of a regular function with a small number of singularities or peaks as the spectral pixels using Coiflet dictionary atoms more sparse than any other wavelet dictionary.

The computational cost of implementing the *basis pursuit* unmixing algorithm using an orthogonal dictionary will reduce to be [6]:

$$O(N_{iter}M^2L) + O(NL) + O(ML) \quad (8)$$

where  $N_{iter}$  is the algorithm number of iterations. Using Coiflet as an orthogonal dictionary will give rise to an initial point of the iterations that is close to the optimal sparse solution for the endmembers which means reducing the number of iterations, so the computational cost given in (15) will be more reduced.

#### 4. Results

This section presents a comparison of using three different analytic dictionaries in solving the *basis pursuit* unmixing algorithm that was discussed in section 2. The three dictionaries are the Coiflet dictionary, the Daubechies dictionary and the Discrete Cosine Transform (DCT) dictionary [10].

We used in our implementation a synthetic cube made from few selected materials from ASTER spectral library [11]. A mixed cube of 3000 spectral pixels was synthesized from known three materials spectra; grass, concrete and asphalt; picked up from the ASTER spectral library [11]. The cube was established using random abundance matrix  $A$  subject to the constraints of the linear mixture model. Additive gaussian noise was added to the synthetic cube spectra with  $SNR = 30$  db . Here in this case, the number of endmembers  $M = 3$  , the number of spectral pixels  $N = 3000$  and the number of wavelengths  $L = 491$ .

The convergence of the *basis pursuit* unmixing algorithm is investigated to assure that using Coiflet dictionary will result in faster convergence. This convergence is obtained by calculating the correlation coefficient between the exact endmember spectrum and the estimated one versus the number of iterations for each endmember. As the correlation coefficient increases to 1, the optimum solution is got faster.

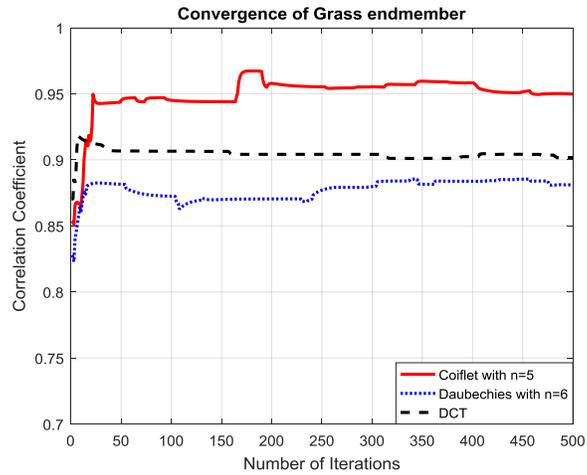
The *basis pursuit* unmixing problem will converge faster when using the Coiflet dictionary as it needs less than 50 iterations to reach around 0.95 correlation values which are too close to the optimum value '1' compared to that of using the other dictionaries as shown in Fig. 1. Decreasing the number of iteration will lead to reducing the computation complexity as shown in equation (8) meaning that the algorithm becomes faster.

#### 5. Conclusion

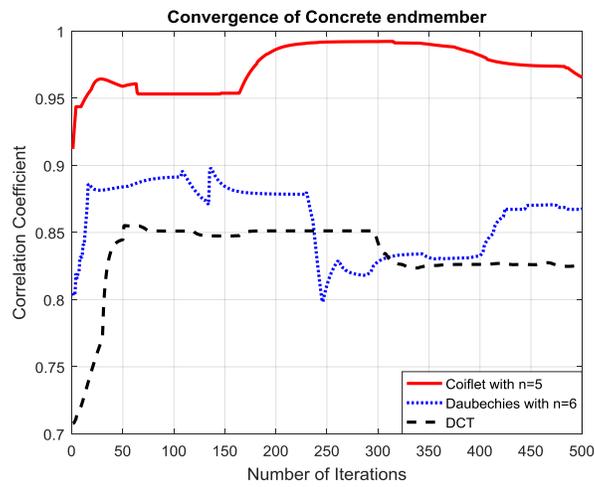
Coiflet dictionary was chosen as an optimal analytic dictionary for sparse representation of the hyperspectral spectral pixels due to its largest number of vanishing moments. Using Coiflet dictionary causes faster and simpler computation of solving the *basis pursuit* unmixing problem.

#### 6. REFERENCES

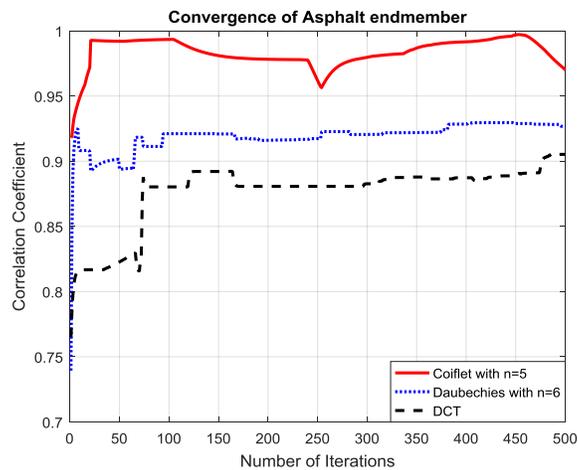
- [1] D. Manolakis, D. Marden and G. A. Shaw, "Hyperspectral image processing for automatic target detection applications," Lincoln Laboratory Journal, vol. 14, (1), pp. 79-116, 2003.
- [2] J. M. Bioucas-Dias et al, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 5, (2), pp. 354-379, 2012.
- [3] P. Comon, C. Jutten and J. Herault, "Blind separation of sources, Part II: Problems statement," Signal Process, vol. 24, (1), pp. 11-20, 1991.
- [4] S. S. Chen, D. L. Donoho and M. A. Saunders, "Atomic decomposition by basis pursuit," SIAM Rev, vol. 43, (1), pp. 129-159, 2001.
- [5] Y. Moudden and J. Bobin, "Hyperspectral BSS using GMCA with spatio-spectral sparsity constraints," IEEE Trans. Image Process., vol. 20, (3), pp. 872-879, 2011.
- [6] J. Starck, F. Murtagh and J. M. Fadili, Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity. Cambridge university press, 2010.
- [7] R. Rubinstein, A. M. Bruckstein and M. Elad, "Dictionaries for sparse representation modeling," Proc IEEE, vol. 98, (6), pp. 1045-1057, 2010.
- [8] I. Rish and G. Grabarnik, Sparse Modeling: Theory, Algorithms, and Applications. CRC Press, 2014.
- [9] S. Mallat, A Wavelet Tour of Signal Processing. Academic press, 1999.
- [10] K. R. Rao and P. Yip, Discrete Cosine Transform: Algorithms, Advantages, Applications. Academic press, 2014.
- [11] ASTER Spectral Library, <http://speclib.jpl.nasa.gov>.



(a)



(b)



(c)

**Fig.1.** The correlation coefficient between the exact and estimated endmembers using three different dictionaries: (a) for the grass endmember, (b) for the Concrete endmember and (c) for the Asphalt endmember.