International Journal of Natural Sciences Research
2016 Vol. 4, No. 5, pp. 83-91
ISSN(e): 2311-4746
ISSN(p): 2311-7435
DOI: 10.18488/journal.63/2016.4.5/63.5.83.91

THE EFFECT OF IMPULSE DENOISING ON GEOMETRIC BASED HYPERSPECTRAL UNMIXING

Bilal KOCAKUŞAKLAR[1]+ --- Nihan KAHRAMAN[2]
[1,2]Department of Electronics and Communications Engineering, Yildiz Technical University Istanbul, Turkey

ABSTRACT

Hyperspectral unmixing is a process to find number of spectral component (called endmember), estimation of endmember signatures and their abundance fractions in each pixel on the scene. Geometric based algorithms are developed for hyperspectral unmixing problem in the literature. The distribution of spectra (points in n-dimensional scatterplot) can be used to estimate endmember signatures geometrically. Impulse denoising before unmixing process can help getting better results for endmember extraction. For this reason, General Prior Algorithm (GAP) is used before unmixing process. Experiments using real data demonstrate that this preprocessing step provided better results for endmember estimation.

Keywords: Hyperspectral unmixing, Vertex component analysis, Minimum volume simplex analysis, N-finder, A variable splitting augmented lagrangian approach, General prior algorithm.

Received: 2 November 2016/ **Revised:** 2 December 2016/ **Accepted:** 7 December 2016/ **Published:** 14 December 2016

Contribution/ Originality

This study contributes better estimation of endmember signatures on geometric based unmixing algorithms by applying spatio-spectral correlation for impulse denoising.

1. INTRODUCTION

Hyperspectral imaging is used widely because of the advantage over finding objects, identifying materials or detecting processes with hundreds of spectral bands that range from visible to near-infrared wavelength region (350 nm.-2500 nm.). Hyperspectral image analysis is a popular topic today with applications to military surveillance, mineral exploration, environmental monitoring, investigation of crop fields.

Surface materials have different scattering properties at specific wavelengths. Hyperspectral sensors can provide high spectral resolution. This specialty of hyperspectral sensors provide to make inferences about surface materials in the scene. Hyperspectral images are obtained by sensor mounted on a plane or in a satellite. Every pixel in hyperspectral image coincides large areas. This causes a low-spatial resolution in hyperspectral data cube.

In linear spectral unmixing, spectral components are assumed homogeneously spread out in the pixels. In the nonlinear mixing, components are distributed randomly because of intimate mixture at microscopic scale in the pixels. Hyperspectral unmixing problem aims at firstly finding the number of endmembers, then endmember's signature estimate and lastly abundance fraction that indicate proportion of each endmember in each pixel of the scene.

There are two approaches for unmixing hyperspectral images; statistical and geometric. In this paper, only geometric based unmixing approaches are considered. Under the linear mixing model, some algorithms solve unmixing problem by using geometric feature that samples taken from the scene are in simplex. From convex

+ Corresponding author

geometry, geometric based unmixing deals with finding a data-enclosing simplex [1, 2]. The vertices of a simplex correspond to the endmembers of the image. N-FINDR [2] and vertex component analysis (VCA) [3] are popular geometric based unmixing algorithms used for endmember extraction. They are based on the assumption that for each endmember, at least one pure pixel exist in the image. But, the assumption is not always applicable especially for all endmember classes. Minimum volume simplex analysis (MVSA) [4] algorithm is used for fixing this problem. MVSA solves hyperspectral unmixing problem by fitting a minimum volume simplex to the hyperspectral data. A variable splitting augmented lagrangian approach (SİSAL) [5] is also not based on pure pixel assumption that is used for hyperspectral unmixing.

Hyperspectral images are distorted with impulse noise because of random fluctations in the power supply of sensors or when the response of the sensor is saturated [6, 7]. Reducing impulse noise from the image has been addressed problem [8]. In this paper, reducing impulse noise with General Analysis Prior (GAP) algorithm as a preprocessing step for hyperspectral unmixing is considered.

Section II of this paper describes linear mixing model. Brief description of geometric based unmixing methods is included in Section III. Section IV describes GAP algorithm for impulse denoising. Section V presents numerical results and the effect of denoising before unmixing. Finally Section VI contains concluding remarks.

2. LINEAR MIXING MODEL

The hypothesis of a linear mixture model is fundamental of geometric based hyperspectral unmixing algorithms. The spectrum of a pixel in the hyperspectral data cube is a linear combination of endmembers that are multiplied with different coefficients.

$$x_l = \sum_{i=1}^{m} a_i \times x_{wi} + \eta \qquad (1)$$

Endmembers are defined by x_{wi}, Pixels of the scene represented by $\{w_1, w_2, w_3, ..., w_m\}$, and η corresponds to noise. $a_1, a_2, a_3, ..., a_m$ coefficients represent proportion of each endmember on examined pixel. m is the number of bands. x_l equals to spectral radiance or reflectance value of the pixel that is the ratio of reflected energy as a function of wavelength. These coefficients have two constraints:

* Sum to one constraint;

$$\sum_{i=1}^{m} a_i = 1 \qquad (2)$$

* The non-negativity constraint;

$$a_1, a_2, a_3, ..., a_m \geq 0 \qquad (3)$$

Estimating the coefficients properly to the Eq. (1) and both constraint is relevant with solving optimization problem examined for example in Heinz and Chang [9].

3. GEOMETRIC BASED UNMIXING ALGORITHMS

A. Vertex Component Analysis (VCA)

VCA is an unsupervised unmixing algorithm. The algorithm provides an geometric solution to the problem. It exploits two facts. First the vertices of a simplex are endmembers. Second, Affine transformation of a simplex is also a simplex particularly by projection operation. It projects the data onto a direction that is orthogonal to the

subspace that are spanned by the endmembers [3]. VCA is based on pure pixel assumption. The algorithm iterates until it reaches previously identified number of endmembers. Vertices of simplex at each iteration is the estimated endmembers.

B. N-Finder

The algorithm aims at finding maximum volume of simplex that are formed by pixels [2]. After finding maximum volume, each vertex of simplex is treated as an endmember. As a dimensional reduction step, principal component analysis (PCA) is applied in this paper.

C. Minimum Volume Simplex Analysis (MVSA)

The aim of MVSA algorithm is to fit a simplex of minimum volume to the dataset. The algorithm is able to unmix hyperspectral data when pure pixel assumption is not provided. That is more realistic scenario. To avoid poor quality local minima, initilization of the algorithm is very important. Initialization of MVSA is obtained by VCA.

D. A Variable Splitting augmented lagrangian approach (SISAL)

SISAL belongs to minimum volume class like MVSA. SISAL is robustness to outliers, noise and poor initialization different from MVSA. The algorithm replaces usual abundance positivity constraint (3), forcing the spectral vectors to belong to the convex hull of the endmember signatures, by soft constraints [5].

4. GAP ALGORITHM

Total variation (TV) regularization is used for image denoising [10] and represented as follows:

$$\|y-x\|_1+\|D_h x\|_1+\|D_v x\|_1 \qquad (4)$$

D_h, D_v are horizontal and vertical finite difference operators. y is the noisy image and x is the image to be recovered. This algorithm takes into account both spectral and spatial correlation for denoising of hyperspectral imaging. GAP algorithm has been compared against recent denoising methods and its superiority in terms of peak signal to noise ratio (PSNR) over another methods has shown in Aggarwal and Majumdar [8].

We rewrite formulation in (4);

$$\arg\min\|Y-X\|_1+\lambda\|D_h XD\|_1+\lambda\|D_v XD\|_1 \qquad (5)$$

$$\text{subject to } \begin{aligned} P &= Y-X \\ Q &= D_h XD \\ R &= D_v XD \end{aligned}$$

$$\text{Let; } \begin{aligned} &y_1 = vec(P-Y-B_1^k), y_2 = vec(Q-B_2^k), \\ &y_3 = vec(R-B_3^k), W_h = D^T \otimes D_h, W_v = D^T \otimes D_v \end{aligned}$$

Bregman variables are updated as:

$$\begin{aligned} B_1^{k+1} &= B_1^k + Y - X - P \\ B_2^{k+1} &= B_2^k + D_h XD - Q \\ B_3^{k+1} &= B_3^k + D_v XD - R \end{aligned}$$

Using Split-Bregman approach the problem can be divided into four separate problems as follows:

$$\left[\mu_1 I + \mu_2 W\right]x = \mu_2\left(W_h^T y_2 + W_v^T y_3\right) - \mu_1 y_1 \qquad (6)$$

$$\begin{aligned} &\min_P \|P\|_1 + \mu_1 \left\|Y - X - P + B_1^k\right\|_F^2 \\ &\min_Q \|Q\|_1 + \mu_2 \left\|Q - D_h XD - B_2^k\right\|_F^2 \\ &\min_R \|R\|_1 + \mu_2 \left\|R - D_v XD - B_3^k\right\|_F^2 \end{aligned} \qquad (7)$$

LSQR [11] is used for solving first problem (6). Another three minimization problem is solved by soft-thresholding. After getting X, P, Q, R values, $B_1^{k+1}, B_2^{k+1}, B_3^{k+1}$ are updated at each iteration.

5. EXPERIMENTS AND RESULTS

A. Hyperspectral Data

Urban HYDICE hyperspectral data is one of the most widely used data for unmixing problems. There are 307x307 pixels in the image and each pixel corresponds to 2x2 m^2 area (Fig.1). There exist 210 wavelength between 400-2500 nm. The spectral resolution of data cube is 10 nm. After removing water absorption and low SNR bands (1-4, 76, 87, 101-111, 136-153, 198-210), 162 usable bands remained.

Fig-1. RGB Urban image
Source: Retrieved from http://www.escience.cn/people/feiyunZHU/Dataset_GT.html

There versions of ground truth (GT) information with 4, 5 and 6 endmembers are obtained in the same way as Jia and Qian [12].

B. Performance Criteria

Spectral Angle Mapper (SAM) is used for measurement of spectral similarity between reference spectra and obtained spectra from algorithms. SAM determines spectral similarity of two spectral by calculating the angle between them. Scaling is not essential because the algorithm uses only vector direction, not vector length. It is the main advantage of SAM. SAM can be computed as follows:

$$a = \arccos \frac{\sum_{t=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^{\,2}} \sqrt{\sum_{i=1}^{n} y_i^{\,2}}} \qquad (8)$$

Here n is the number of bands, x and y are the spectral vector. The other metric root-mean-square-error (rmse) is used to calculate the similarity of true (S) versus estimated ($\hat{S}$) abundances [13].

$$rmse_k = \left(\frac{1}{K} \sum \left(S_{kj} - \hat{S}_{kj} \right)^2 \right)^{\frac{1}{2}} \qquad (9)$$

C. Experimental Results

VCA, N-finder, MVSA and SISAL algorithms are used for endmember extraction on URBAN data. We need hyperspectral data cube and number of endmembers for these algorithms. SAM and rmse calculations for each algorithm are evaluated for 4, 5, 6 endmembers respectively. GAP algorithm exploiting spatio-spectral correlation is used for impulse denoising in image. The image and denoised image at band 80 depict the scene displayed in Figure 2. We have empirically chosen the parameters $(\lambda, \mu_1, \mu_2, \max iter) = (2,2,2,50)$. Now that MATLAB gives out of memory error over 10 bands, first 1-10 bands have been taken for denoising process with GAP algorithm. Then the process repeated for the other bands and finally denoised parts are combined. But we know that increasing the number of bands improve denoising results. Because the algorithm uses both spatial and spectral information. VCA, MVSA, and SISAL algorithm contain random initial conditions that give different results each time they run. The algorithms have run 10 times and the average values of the results are exhibited.

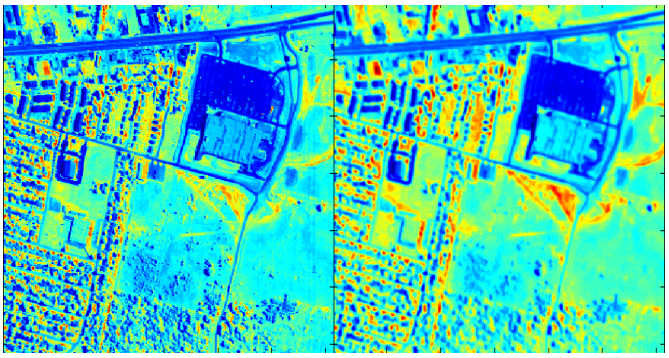

Fig-2. Original image and denoised image (band 80)
Source: MATLAB implementation

Table I quantifies the obtained results for 4 endmembers. The numbers in bold represent the best performance. Among all algorithms, best performance is achieved in denoised images. The algorithms that assume pure pixel assumption (VCA,N-finder) showed better development on SAM scores. Asphalt road is the hardest endmember to extract for geometric based algorithms.

Table-I. Sam Similarity Between Endmembers Extracted from Urban Dataset by Vca, N-Finder, Mvsa, Sisal and the Ground Truth (4 Endmember)

Method	Endmember				
	Asphalt Road	**Grass**	**Tree**	**Roof**	**Average**
VCA (original)	0,8198	0,5842	0,2784	0,8228	0,6263
VCA (denoised)	0,5838	**0,2164**	0,0827	0,438	0,3302
N-finder (original)	1,1535	0,5937	0,0744	0,2174	0,5097

Method	Endmember				
	Asphalt Road	Grass	Tree	Roof	Average
N-finder (denoised)	**0,4902**	0,3876	**0,0678**	**0,1043**	**0,2625**
MVSA (original)	2,6568	0,5547	0,4216	0,1647	0,9495
MVSA (denoised)	2,8098	0,5629	0,1637	0,1196	0,914
SISAL (original)	2,5779	0,5496	0,3164	0,1982	0,9105
SISAL (denoised)	2,6313	0,555	0,2803	0,1478	0,9036

Source: MATLAB implementation

Table II demonstrates 5 endmember SAM scores. Dirt is added as a fifth endmember. Estimation of asphalt road signature is again not close to reference spectrum of it. The reference signatures of grass and tree are close to each other. N-finder uses PCA as a first step of algorithm. This provides removing of noisy bands. Upon this we apply impulse denoising algorithm. It also causes smoothing. Then, N-finder can not distinguish tree and grass signature and thus evaluate them as a single endmember. Average SAM scores of the algorithms improved after denoising.

Table-2. Sam Similarity between Endmembers Extracted From Urban Dataset by Vca, N-Finder, Mvsa, Sisal and the Ground Truth (5 Endmember)

Method	Endmember					
	Asphalt Road	Grass	Tree	Roof	Dirt	Average
VCA (original)	0,8088	0,7233	0,3442	0,8283	0,326	0,6061
VCA (denoised)	0,5644	0,2791	0,1377	0,2324	0,2869	0,3001
N-finder (original)	1,141	**0,116**	**0,1227**	0,174	0,2223	0,3552
N-finder (denoised)	**0,4821**	0,6369	0,1268	**0,1009**	**0,0626**	**0,2819**
MVSA (original)	2,621	0,3461	0,5595	0,3826	0,3289	0,8476
MVSA (denoised)	2,5073	0,1347	0,2528	0,1993	0,2955	0,6779
SISAL (original)	2,5633	0,1778	0,4097	0,3336	0,3161	0,7601
SISAL (denoised)	2,5533	0,1522	0,2579	0,1853	0,2481	0,6794

Source: MATLAB implementation

Table III shows SAM similarity scores for 6 endmembers. The SAM scores of pure pixel assumption based algorithms again get much better after denoising step. While the algorithms are unable to find metal signature with original image, they can get close signature to the original one with denoised image with VCA and N-finder.

Table-3. Sam Similarity Between Endmembers Extracted From Urban Dataset By Vca, N-Finder, Mvsa, Sisal And The Ground Truth (6 Endmember)

Method	Endmember						
	Asphalt Road	Grass	Tree	Roof	Metal	Dirt	Average
VCA (original)	0,5928	0,4428	0,3542	0,8129	0,685	0,3787	0,5444
VCA (denoised)	0,4812	0,4345	0,1998	0,1952	0,3473	0,1379	0,2993
N-finder (original)	**0,4754**	**0,116**	0,1227	0,174	1,1624	0,2771	0,3879
N-finder (denoised)	0,5638	0,6423	**0,0868**	**0,1009**	**0,2076**	**0,096**	**0,2829**
MVSA (original)	1,6643	1,0515	0,4603	0,5098	1,2594	0,3242	0,8783
MVSA (denoised)	1,8943	0,8288	0,2626	0,2089	1,1455	0,2388	0,7632
SISAL (original)	1,9077	0,7402	0,363	0,3189	1,8632	0,2866	0,9133
SISAL (denoised)	1,8508	1,6237	0,2346	0,1854	1,0219	0,1972	0,8523

Source: MATLAB implementation

Fully constrained least squares (FCLS) was developed [14] as an abundance estimator of the endmembers present in an image pixel. The method can ensure both sum-to-one (2) and non-negativity (3) constraint. FCLS demands signature of endmembers. They are obtained with VCA, N-finder, MVSA, SISAL and their denoised

versions. Table IV shows average rmse scores between GT-FCLS and geometric based unmixing algorithms-FCLS. Denoised-VCA demonstrated better results than original-VCA. When rmse values are considered separately for each endmember, poor results of denoised image in other algorithms result from asphalt road and grass estimates. It shows that better estimation of endmember signatures as an average doesn't mean better estimate for abundance estimate also. Each endmember signature has to be estimated close to the actual signature for a good abundance estimate. Only one incorrect signature can disrupt total results. Because according to (2-3) they are dependent on each other. In Fig. 3, abundance estimate is obtained with applying FCLS to GT. Figure (4-5) displays best abundance estimates among algorithms of original data and denoised data.

Table-4. Rmse Similarity Between Abundance Estimates From Urban Dataset By Vca, N-Finder, Mvsa, Sisal And The Ground Truth (Fcls Is Used For Abundance Estimate)

Method	4 Endmember	5 Endmember	6 Endmember
VCA (original)	0,3418	0,371	0,3023
VCA (denoised)	**0,2552**	**0,1845**	0,2004
N-finder (original)	0,3566	0,2178	**0,1949**
N-finder (denoised)	0,3971	0,2533	0,2273
MVSA (original)	0,3032	0,2223	0,2135
MVSA (denoised)	0,3226	0,2123	0,2181
SISAL (original)	0,3103	0,2053	0,2161
SISAL (denoised)	0,3333	0,2122	0,2118

Source: MATLAB implementation

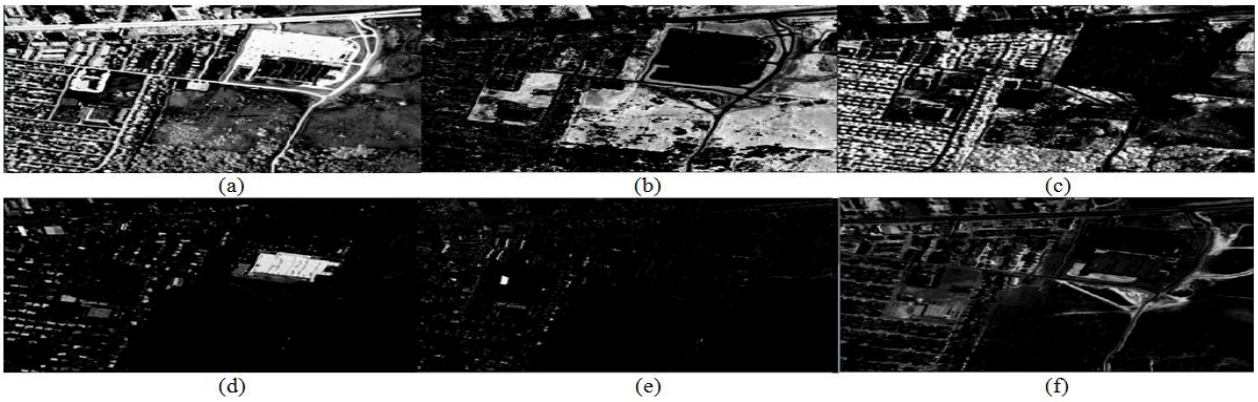

Fig-3. Abundance map estimate (GT is used for endmembers) (a) asphalt road, (b) grass, (c) tree, (d) roof, (e) metal, (f) dirt
Source: MATLAB implementation

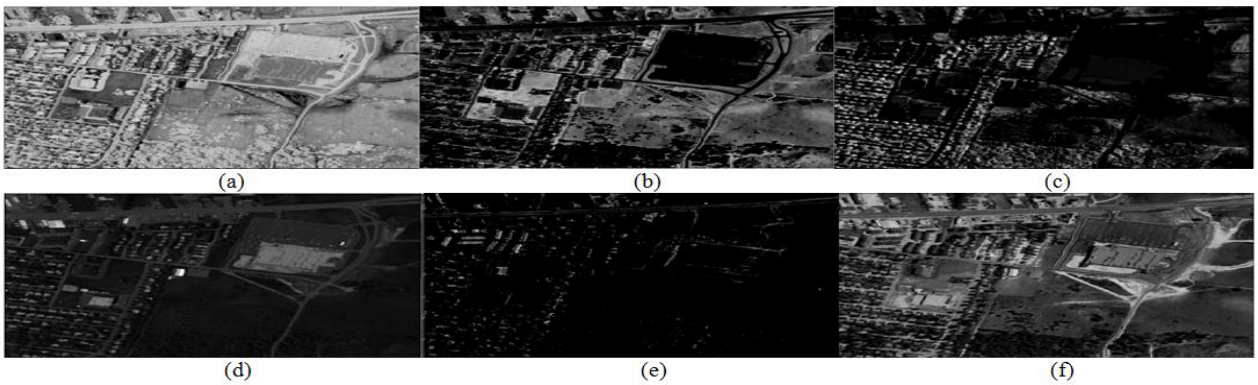

Fig-4. Abundance map estimate (endmembers extracted by algorithms with original data) (a) asphalt road, (b) grass, (c) tree, (d) roof, (e) metal, (f) dirt
Source: MATLAB implementation

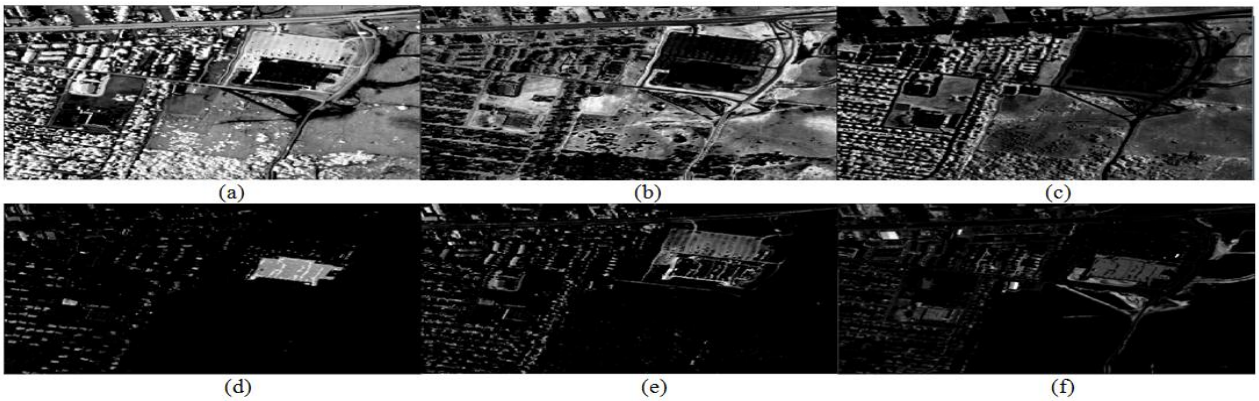

Fig-5. Abundance map estimate (endmembers extracted by algorithms with denoised data) (a) asphalt road, (b) grass, (c) tree, (d) roof, (e) metal, (f) dirt
Source: MATLAB implementation

6. CONCLUSION AND FUTURE WORK

Removing impulse noise in hyperspectral image is an important pre-processing step. This step provided better results for endmember extraction on geometric based unmixing algorithms. But, impulse denoising in scene was unable to get better estimates of abundance fractions excluding VCA algorithm. Among the algorithms, pure-pixel assumption based like VCA, N-finder get closer result to GT information in endmember extraction step.

As future work, Denoising of other noise types like Gaussian and its effect on hyperspectral unmixing are needed for further research. In this paper, geometric based unmixing algorithms are considered. Therefore statistical based unmixing algorithms should also be studied to see denoising effect better.

Funding: This study received no specific financial support.
Competing Interests: The authors declare that they have no competing interests.
Contributors/Acknowledgement: All authors contributed equally to the conception and design of the study.

REFERENCES

[1] J. W. Boardman, "Automating spectral unmixing of aviris data using convex geometry concepts," in *Proc. Summ. 4th Annu. JPL Airborne Geosci. Workshop, R. O. Green, Ed*, 1994, pp. 11-14.

[2] A. Plaza and C. I. Chang, "An improved n-findr algoritm in implementation," in *Algorithm and Technology for Multispectral, Hyperspectral and Ultraspectral Imagery XI*, n.d.

[3] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 898-910, 2005.

[4] J. Li and J. Bioucas-Dias, "Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data," presented at the IEEE International Geoscience and Remote Sensing Symposium, 2008.

[5] J. Bioucas-Dias, "A variable splitting augmented lagrangian approach to linear spectral unmixing," presented at the IEEE GRSS Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, Grenoble, France, 2009.

[6] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising employing a spectral-spatial adaptive total variation model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, pp. 3660-3677, 2012.

[7] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising with a spatial-spectral view fusion strategy," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, pp. 2314-2325, 2014.

[8] H. K. Aggarwal and A. Majumdar, "Exploiting spatio-spectral correlation for impulse denoising in hyperspectral images," *SPIE Journal of Electronic Imaging*, vol. 24, pp. 013027-013027, 2015.

[9] D. Heinz and C. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, pp. 529-545, 2001a.

[10] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, pp. 259-268, 1992.

[11] M. A. Saunders, "Solution of sparse rectangular systems using LSQR and CRAIG," *BIT Numerical Mathematics*, vol. 35, pp. 588-604, 1995.

[12] S. Jia and Y. Qian, "Spectral and spatial complexity-based hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 3867-3879, 2007.

[13] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitive and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, pp. 650-663, 2004.

[14] D. Heinz and C. I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, pp. 529-545, 2001b.