EUSIPCO 2013 1569744489

HYPERSPECTRAL CODED APERTURE (HYCA): A NEW TECHNIQUE FOR HYPERSPECTRAL COMPRESSIVE SENSING

Gabriel Martín[1], José M. Bioucas-Dias[2] and Antonio Plaza[1]

[1]Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications
University of Extremadura, Escuela Politécnica de Cáceres, E-10071, Spain

[2]Instituto de Telecomunicações, Instituto Superior Técnico
Torre Norte, Piso 10, Av. Rovisco Pais, 1049-001 Lisbon, Portugal

ABSTRACT

Hyperspectral imaging is an active research area in remote sensing. Due to the high volume of hyperspectral image data, the exploration of compression strategies has received a lot of attention in recent years. In this paper, we introduce a new compressed sensing methodology, termed Hyperspectral coded aperture (HYCA), which exploits the high correlation existing among the components of remotely sensed hyperspectral data sets to reduce the number of measurements necessary to correctly reconstruct the original data. HYCA relies on two central properties of most hyperspectral images: i) the spectral vectors live on a low dimensional subspace and ii) the spectral bands are piecewise smooth. The former property allows to represent the data vectors using a small number of coordinates, and the latter implies that each coordinate is piecewise smooth and thus compressible on local differences. The reconstruction of the data cube is obtained by minimizing a convex objective function containing a data term associated to the compressed measurements and a total variation spatial regularizer. A series of experiments with simulated and real data show the effectiveness of the newly developed HYCA, indicating that the proposed scheme has a high potential in real-world applications.

Index Terms— Hyperspectral imaging, compressive sensing, coded aperture, signal subspace, total variation, optimization.

1. INTRODUCTION

Hyperspectral imaging spectrometers collect hundreds or thousands of bands (at different wavelength channels) for the same area on the surface of the Earth [1]. For instance, the NASA Jet Propulsion Laboratory's Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS) covers the wavelength region from 0.4 to 2.5 microns using 224 spectral channels, at nominal spectral resolution of 10 nanometers [2]. The resulting multidimensional data cube typically comprises several gigabytes per flight.

Due to the extremely large volumes of data collected by imaging spectrometers, hyperspectral data compression has received considerable interest in recent years [3, 4]. These data are usually collected by a satellite or an airbone instrument and sent to a ground station on Earth for subsequent processing. Usually the bandwidth connection between the satellite/airborne platform and the ground station is reduced, which limits the amount of data that can be transmitted. As a result, there is a clear need for (either lossless or lossy) hyperspectral data compression techniques that can be applied onboard the imaging instrument.

In this paper, we develop a new compressive sensing (CS) framework [5, 6] for hyperspectral images, termed hyperspectral coded aperture (HYCA), which exploits two characteristics of the hyperspectral images: i) the hyperspectral vectors belong to a low dimensional subspace, and ii) the imaged scene components present a high spatial correlation. These two characteristics mean that the hyperspectral data cubes are compressible, *i.e.,* they admit a representation in given base or frame in which most of the coefficients are small, and, thus, the data is well approximated with just a small number of large coefficients.

Compressibility, or sparsity[1], is a necessary condition for success of compressive sensing. In our approach, and having in mind the characteristics i) and ii), we represent the spectral vectors in a basis of the signal subspace and model the spatial correlation by promoting small local differences on the images of coefficients by minimizing their total variation [7]. Under the linear mixing model [8], if the spectral signatures of the endmembers are used to represent the spectral vectors, then the representation coefficients are the abundance fractions of the pure materials. In this way, the proposed approach is strongly connected with unmixing. To be more precise, and assuming we use the spectral signatures of the endmembers to represent the spectral vectors, our methodology implements hyperspectral unmixing in addition to hyperspectral compressive sensing.

[1]A vector is k-sparse is only k of its components are different from zero.

The remainder of the paper is organized as follows. Section 2 describes the proposed methodology. Section 3 describes the experimental results, conducted in this work using both synthetic and real hyperspectral data. Section 4 concludes the paper with some remarks and hints at plausible future research lines.

2. DESCRIPTION OF THE METHOD

Let $\mathbf{X} \in \mathbb{R}^{n_b \times n_p}$ represent, in matrix format, a hyperspectral image with n_b spectral bands and $n_p := n_r n_c$ pixels where n_r and n_c denote, respectively, the number of rows and columns of the hyperspectral image in the spatial domain.

Let $\mathbf{y} \in \mathbb{R}^m$ denote the CS measurements modeled as

$$\mathbf{y} = A(\mathbf{X}) + \mathbf{w}, \tag{1}$$

where $A : \mathbb{R}^{n_b \times n_p} \to \mathbb{R}^m$ is a linear operator which computes m inner products between known m vectors and the elements of $\mathbf{X}$ and $\mathbf{w}$ models additive perturbation, hereafter termed noise, accounting for, *e.g.*, modeling errors and system noise. Since A is a linear operator, then we have $A(\mathbf{X}) = \mathbf{A}\mathbf{x}$, where $\mathbf{x} := \text{vec}(\mathbf{X})$ is the vectorization of matrix $\mathbf{X}$ by stacking its columns and $\mathbf{A} \in \mathbb{R}^{m \times n}$, with $n := n_b n_p$, is the matrix representation of the linear operator A.

The objective of CS is to recover $\mathbf{x}$ from $\mathbf{y}$ with $m \ll n$, in order to have compression in the acquisition. Without any further information, this recovering is impossible even in the absence of noise because the matrix $\mathbf{A}$ is undetermined. If, however, vector $\mathbf{x}$ admits a sparse representation with respect to a given frame $\mathbf{\Phi}$, *i.e.,* $\mathbf{x} = \mathbf{\Phi}\boldsymbol{\theta}$ with $\boldsymbol{\theta}$ sparse, then the solution of the optimization problem

$$\min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_0 \quad \text{subject to: } \|\mathbf{y} - \mathbf{A}\mathbf{\Phi}\boldsymbol{\theta}\| \leq \delta \tag{2}$$

yields, in given conditions, a good approximation for $\mathbf{x}$. In (2), the notation $\|\mathbf{x}\|_0$, abusively termed the ℓ_0 norm of $\mathbf{x}$, denotes the number on non-null components of $\mathbf{x}$ and $\delta \geq 0$ is a parameter depending on the "size" of the noise level.

For $\delta = 0$, if the system of linear equations $\boldsymbol{y} = \mathbf{A}\mathbf{\Phi}\boldsymbol{\theta}$ has a solution satisfying $2\|\boldsymbol{\theta}\|_0 < \text{spark}(\mathbf{A}\mathbf{\Phi})$, where $\text{spark}(\mathbf{A}) \leq \text{rank}(\mathbf{A}) + 1$ is the smallest number of linearly dependent columns of $\mathbf{A}\mathbf{\Phi}$, it is necessarily the unique solution of (2) [9, 10]. For $\delta > 0$, the concept of uniqueness of the sparsest solution is replaced with that of stability [11, 12, 13]. For example in [12], it is shown that, given a solution vector $\boldsymbol{\theta}$ of the noiseless ($\delta = 0$) problem (2), satisfying the sparsity constraint $\|\boldsymbol{\theta}\|_0 < (1/\mu(\mathbf{A}\mathbf{\Phi}) + 1)/2$, where $\mu(\mathbf{A}\mathbf{\Phi})$ is the mutual coherence of matrix $\mu(\mathbf{A}\mathbf{\Phi})$, then $\boldsymbol{\theta}$ is unique and any solution $\boldsymbol{\theta}^\delta$ of (2) for a given $\delta > 0$ such that $\|\mathbf{y} - \mathbf{A}\mathbf{\Phi}\boldsymbol{\theta}^\delta\| \leq \delta$ satisfies

$$\|\boldsymbol{\theta}^\delta - \boldsymbol{\theta}\|^2 \leq \frac{4\delta^2}{1 - \mu(\mathbf{A}\mathbf{\Phi})(2\|\boldsymbol{\theta}\|_0 - 1)}.$$

The problem (2) is NP-hard [14] and therefore there is little hope in solving it in a straightforward way. A possible strategy is to use convex relaxation via analysis based regularization, which replaces the ℓ_0 norm with the ℓ_1 norm obtaining

$$\min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_1 \quad \text{subject to: } \|\mathbf{y} - \mathbf{A}\mathbf{\Phi}\boldsymbol{\theta}\| \leq \delta. \tag{3}$$

The above formulation is synthesis-based, in the sense that $\mathbf{x} = \mathbf{\Phi}\boldsymbol{\theta}$, *i.e.*, $\mathbf{x}$ is synthesized from the coefficients $\boldsymbol{\theta}$. Here, we adopt an analysis-based approach, *i.e.*, we use an analysis operator $\mathbf{\Psi}$ such that the analysis of the transformed coefficients $\mathbf{\Psi}\mathbf{x}$ are sparse. Then, we obtain

$$\min_{\mathbf{x}} \|\mathbf{\Psi}\mathbf{x}\|_1 \quad \text{subject to: } \|\mathbf{y} - \mathbf{A}\mathbf{x}\| \leq \delta. \tag{4}$$

It should be noted that there is empirical evidence of the superiority of the analysis based approaches [15].

In our approach, we design the measurement matrix $\mathbf{A}$ as block diagonal, *i. e.*,

$$\mathbf{A} = \text{bdiag}(\mathbf{A}_1, \ldots, \mathbf{A}_{n_p})$$

where $\mathbf{A}_i \in \mathbb{R}^{q \times n_b}$ acts on the spectral vector $\mathbf{x}_i$ computing q projections. In this way, we obtain qn_p measurements and thus a compression rate of q/n_b. In our implementation, and for algorithmic complexity reasons, we have $\mathbf{A}_i \in \{\mathbf{H}_1, \ldots, \mathbf{H}_{n_h}\}$, *i.e.*, we will have only n_h different matrices $\mathbf{A}_i$ where the way we assign a matrix to a given pixel is depending on the pixel. Specifically, we partition the image into square windows, an inside each window we assign $\mathbf{H_i}$ to the i-th window pixel. Finally, another important observation is that hyperspectral datasets live systematically in low dimensional subspaces and therefore we can write

$$\mathbf{X} = \mathbf{E}\mathbf{Z}, \tag{5}$$

where $\mathbf{E} \in \mathbb{R}^{n_b \times p}$ is a full column rank matrix, possibly orthogonal, which spans the signal subspace. Usually, we have $p \ll n_b$. In this work, we assume that $\mathbf{E}$ is known (this is a reasonable assumption since the complete dataset is acquired onboard and the bottleneck is in the transmission of the data to a control station on Earth). In other words, the current trend is to compute matrix $\mathbf{E}$ onboard and then do CS acquisition and send only $m = qn_p$ samples with $q < p$. If $\mathbf{E}$ is the mixing matrix and the linear mixing model is valid, then $\mathbf{Z}$ are the fractional abundances.

Given that $\mathbf{x} = \text{vec}(\mathbf{X}) = \text{vec}(\mathbf{E}\mathbf{Z}) = (\mathbf{I} \otimes \mathbf{E})\mathbf{z}$, with $\mathbf{z} := \text{vec}(\mathbf{Z})$, then we have

$$\mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{I} \otimes \mathbf{E})\mathbf{z} \tag{6}$$

$$= \text{bdiag}(\mathbf{A}_1\mathbf{E}, \ldots, \mathbf{A}_{n_p}\mathbf{E})\mathbf{z} \tag{7}$$

$$= \mathbf{K}\mathbf{z}, \tag{8}$$

where

$$\mathbf{K} := \text{bdiag}(\mathbf{A}_1\mathbf{E}, \ldots, \mathbf{A}_{n_p}\mathbf{E}).$$

Similarly to the criteria (4), we propose the following convex optimization problem to recover $\mathbf{z}$

$$\min_{\mathbf{z}} \|\mathbf{\Psi z}\|_1 \quad \text{subject to: } \|\mathbf{y} - \mathbf{Kz}\| \leq \delta. \tag{9}$$

Our regularizer $\|\mathbf{\Psi z}\|_1$ is the sum of total variations of the p images of coefficients $\mathbf{Z}$:

$$\|\mathbf{\Psi z}\|_1 = \sum_{i=1}^{p} \text{TV}(\mathbf{Z}_i) \tag{10}$$

$$:= \text{TV}(\mathbf{z}), \tag{11}$$

where $\mathbf{Z}_i := \mathbf{Z}(i,:)$ (matlab notation) is the ith image of the representation coefficients with respect to matrix $\mathbf{E}$. Therefore, $\mathbf{\Psi}$ is the discrete gradient operating over the images $\mathbf{Z}_i$. By minimizing $\text{TV}(\mathbf{z})$, we are promoting piecewise-smooth images of coefficients. As mentioned before, when $\mathbf{E}$ is the mixing matrix, we can interpret the above results in terms of abundance fractions.

In this work we consider two criteria:

C-HYCA criterion:

$$\min_{\mathbf{z}} \text{TV}(\mathbf{z}) \quad \text{subject to: } \|\mathbf{y} - \mathbf{Kz}\| \leq \delta, \quad \mathbf{z} \geq \mathbf{0}. \tag{12}$$

HYCA criterion:

$$\min_{\mathbf{z}} (1/2)\|\mathbf{y} - \mathbf{Kz}\|^2 + \lambda \text{TV}(\mathbf{z}) \tag{13}$$

$$\text{subject to: } \mathbf{z} \geq \mathbf{0}. \tag{14}$$

The constraint $\mathbf{z} \geq 0$ is enforced only if $\mathbf{E}$ is the mixing mitrix. Otherwise, it is disregarded. The optimization problems (12) and (13) are convex, although very complex from the computational point of view due to its dimension and due to the presence of non-smooth terms in the objective functions. We tackle these difficulties by using instances of the alternate direction method of multipliers described in [16] which convert a very difficult problem into a sequence of simpler problems.

3. EXPERIMENTAL RESULTS

In this section, we conduct a series of experiments using different versions of HYCA and C-HYCA algorithms on real and simulated data. Specifically, we have considered three cases, first we have considered the problem without the non-negativity constraint, in this case we have considered that $\mathbf{E}$ is an orthonormal matrix and the sampling matrix $\mathbf{H}_i$ is a random, Gaussian iid matrix. In the second case we have considered that $\mathbf{H}_i = \mathbf{G}_i\mathbf{E}^\sharp$, where $\mathbf{G}_i \in \mathbb{R}^{q\times p}$ is a random, Gaussian iid matrix, and $\mathbf{E}^\sharp$ is the pseudo-inverse of the matrix $\mathbf{E}$. In the third case we have considered the non-negativity constraint and we assume that $\mathbf{E}$ is the mixing matrix which contain the endmembers and that $\mathbf{H}_i$ is a random, Gaussian iid matrix.

Table 1. Average NMSE between the original and the reconstructed dataset for $q = 3$ and different SNR values, after 10 Monte-Carlo runs.

Version	SNR=30db	SNR=50db	SNR=70db	SNR=∞
HYCA *	$8.27 \cdot 10^{-4}$	$0.52 \cdot 10^{-4}$	$0.52 \cdot 10^{-4}$	$0.30 \cdot 10^{-4}$
HYCA †	$0.44 \cdot 10^{-4}$	$0.45 \cdot 10^{-4}$	$0.40 \cdot 10^{-4}$	$0.37 \cdot 10^{-4}$
HYCA §	$21.08 \cdot 10^{-4}$	$0.68 \cdot 10^{-4}$	$0.33 \cdot 10^{-4}$	$0.20 \cdot 10^{-4}$
C-HYCA *	$22.7 \cdot 10^{-4}$	$3.35 \cdot 10^{-4}$	$3.24 \cdot 10^{-4}$	$3.36 \cdot 10^{-4}$
C-HYCA †	$2.73 \cdot 10^{-4}$	$2.00 \cdot 10^{-4}$	$1.31 \cdot 10^{-4}$	$1.59 \cdot 10^{-4}$
C-HYCA §	$7.26 \cdot 10^{-4}$	$0.51 \cdot 10^{-4}$	$0.29 \cdot 10^{-4}$	$0.28 \cdot 10^{-4}$

* First case without non-negativity and $\mathbf{E}^* := orth(\mathbf{E})$. † Second case with $\mathbf{H}_i = \mathbf{G}_i\mathbf{E}^\sharp$. § Third case using the non-negativity constraint.

3.1. Experiments with Synthetic data

The synthetic dataset used in this experiments were generated from spectral signatures randomly selected from the United States Geological Survey (USGS) [2]. The simulated images consist of a set of 5×5 squares of 10×10 pixels each one, for a total size of 110×110 pixels. The first row of squares contains the endmembers, the second row contains mixtures of two endmembers, the third row contains mixtures of three endmembers, and so on. Zero-mean Gaussian noise was added to the synthetic scenes in with signal-to-noise ratios (SNRs) defined as $\text{SNR} = 10 \cdot \log_{\mathbf{10}} \frac{\mathbb{E}\|\mathbf{EZ}\|_\mathbf{F}^\mathbf{2}}{\mathbb{E}\|\mathbf{w}\|_\mathbf{F}^\mathbf{2}}$, where $\mathbb{E}$ denotes mean value, to simulate contributions from ambient and instrumental noise sources. Fig. 1 displays the ground-truth abundance maps used for generating the simulated imagery.

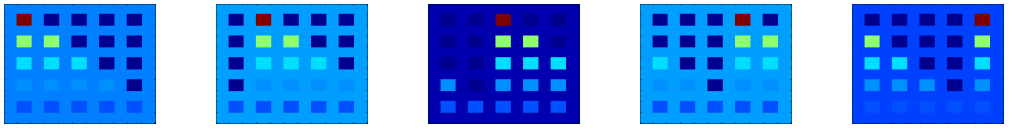
In order to evaluate the performance of the HYCA and C-HYCA, we use as performance indicator the normalized mean squared error (NMSE) of the reconstruction given by

$$\text{NMSE} = ||\mathbf{E}(\widehat{\mathbf{Z}} - \mathbf{Z})||_F^2 / ||\mathbf{EZ}||_F^2, \tag{15}$$

where $\mathbf{Z}$ and $\widehat{\mathbf{Z}}$ denote the original and reconstructed abundance fraction maps, respectively. In our experiments, we set the window size $ws = 2$. In the first case we disabled the non-negativity constraint and we assumed that the mixing matrix $\mathbf{E}^*$ is an orthogonal matrix computed as $\mathbf{E}^* := orth(\mathbf{E})$, where $\mathbf{E}$ denotes the original mixing matrix used to generate the dataset and $\mathbf{E}^*$ denotes the mixing matrix used in the reconstruction algorithm. In the second case we used for sampling $\mathbf{H}_i = \mathbf{G}_i\mathbf{E}^\sharp$ without the non-negativity constraint. Finally, in the third case we used the non-negativity constraint considering the original mixing matrix used for the dataset generation. In this experiment, we set $q = 3$. Since the original data set has $nb = 224$ bands, the compression ratio is $nb/q = 74.67$ and the number of measurements per pixel band is $q/nb = 0.0134$.

Table 1 shows the NMSE obtained for both versions (HYCA and C-HYCA) in different cases. We performed 10 Monte-Carlo runs, sampling not only the noise but also the elements of the linear operator $\mathbf{A}$. The regularization parameter λ in (13) was hand-tuned for optimal performance in

[2] http://speclab.cr.usgs.gov/spectral-lib.html



(a) Endmember #1 (b) Endmember #2 (c) Endmember #3 (d) Endmember #4 (e) Endmember #5

Fig. 1. True abundance maps of endmembers in the synthetic hyperspectral data.

.

the case of the HYCA algorithm. Having in mind the linear model (5), the parameter δ in (12) is set to $\delta = \|\mathbf{A}(\mathbf{w})\|_F$.

As we can see in Table 1, by disabling the non-negativity constraint the algorithm is more robust to noise. However, the implementation with the non-negativity constraint provides better results when there is no noise. This is expected, as noise introduces outliers in the model which lead to errors in the reconstruction when we use the non-negativity constraint. If we compare the HYCA and C-HYCA criteria on Table 1, we can see that both versions provide very good results with low reconstruction errors. We can also observe that HYCA outperforms C-HYCA, which is probably due to the parameter optimization conducted in the case of HYCA.

3.2. Experiment with Real Data

In this experiment, we use the well-known AVIRIS Cuprite data set, available online in reflectance units after atmospheric correction. This scene has been widely used to validate the performance of endmember extraction algorithms. The portion used in experiments corresponds to a 250 by 190 pixels subset of the sector labeled as f970619t01p02_r02_sc03a.rfl in the online data. The scene comprises 224 spectral bands between 0.4 and 2.5 μm, with full width at half maximum of 10 nm and spatial resolution of 20 meters per pixel. Prior to the analysis, several bands were removed due to water absorption and low SNR in those bands, leaving a total of 188 reflectance channels to be used in the experiments. We used a window size of $ws = 2$, so that $m = 4$. Here, we estimated the number of endmembers with Hysime algorithm [17].

In this experiment, we have also considered the same three cases as in the experiment with simulated data. But in this case the mixing matrix was estimated from the original data using the vertex component algorithm (VCA) [18], so that in the first case we use $\mathbf{E}^* := orth(\widehat{\mathbf{E}})$ where $\widehat{\mathbf{E}}$ is the mixing matrix estimated by VCA algorithm. In the second case $\mathbf{H}_i = \mathbf{G}_i\widehat{\mathbf{E}}^\sharp$ and $\mathbf{E}^* := \widehat{\mathbf{E}}$. Finally in the third case, due to the non-linear mixtures and outliers present in the real images, the non-negativity constraint may be violated. In order to ensure that the mixing matrix encloses the whole data set and the non-negativity constraint is satisfied, we open the cone defined in the mixing matrix $\mathbf{E}^*$ as follows:

$$\mathbf{E}^* := \widehat{\mathbf{E}} + \Delta \cdot (\widehat{\mathbf{E}} - \overline{\mathbf{E}}) \tag{16}$$

where Δ is a scalar which defines how much the cone is opened and $\overline{\mathbf{E}}$ is a matrix containing the mean spectrum of the endmembers. By choosing a value of Δ large enough, then all observed spectral vectors are inside the cone implying that $\mathbf{z} \geq 0$. In the current data set, $\Delta = 6$ ensures this constraint. In order to evaluate the performance of HYCA and C-HYCA with the real dataset, we perform experiments with the compression ratios $188/q$ with $q = 5, 9, 13, 17$. In all cases we used a window size of $ws = 2$.

Table 2. Average NMSE between the AVIRIS Cuprite data and its reconstructed version (after 10 Monte-Carlo runs) for different compression ratios.

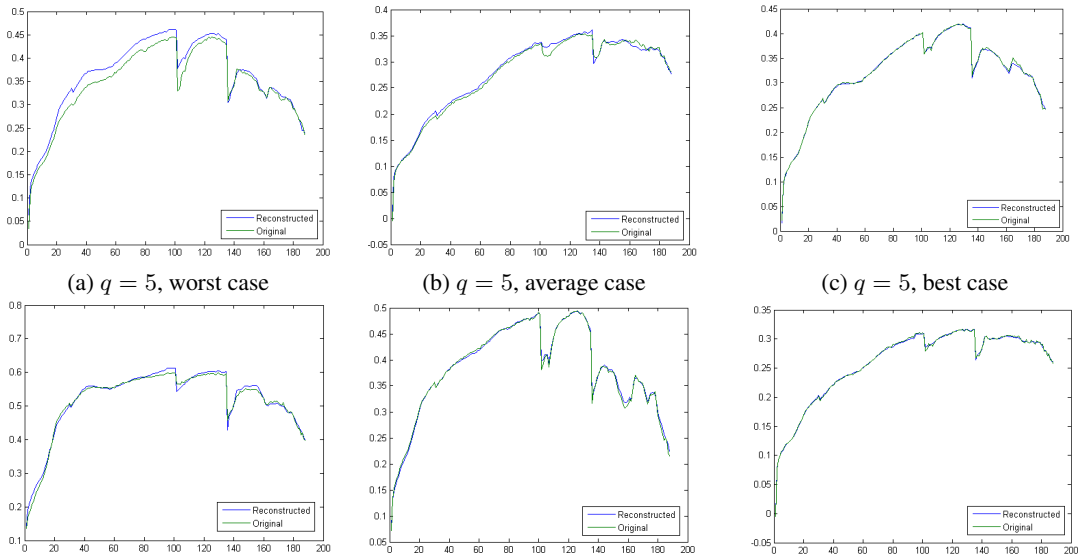
Version	$q = 5$	$q = 9$	$q = 13$	$q = 17$
HYCA *	$98.13 \cdot 10^{-4}$	$5.21 \cdot 10^{-4}$	$2.59 \cdot 10^{-4}$	$4.54 \cdot 10^{-4}$
HYCA †	$376.83 \cdot 10^{-4}$	$200.39 \cdot 10^{-4}$	$108.80 \cdot 10^{-4}$	$34.56 \cdot 10^{-4}$
HYCA §	$4.66 \cdot 10^{-4}$	$2.50 \cdot 10^{-4}$	$1.28 \cdot 10^{-4}$	$1.05 \cdot 10^{-4}$
C-HYCA *	$24.05 \cdot 10^{-4}$	$5.77 \cdot 10^{-4}$	$3.54 \cdot 10^{-4}$	$2.63 \cdot 10^{-4}$
C-HYCA †	$590.82 \cdot 10^{-4}$	$351.57 \cdot 10^{-4}$	$268.54 \cdot 10^{-4}$	$225.88 \cdot 10^{-4}$
C-HYCA §	$3.84 \cdot 10^{-4}$	$2.24 \cdot 10^{-4}$	$1.74 \cdot 10^{-4}$	$1.31 \cdot 10^{-4}$

* First case without non-negativity and $\mathbf{E}^* := orth(\mathbf{E})$. † Second case with $\mathbf{H}_i = \mathbf{G}_i\mathbf{E}^\sharp$. § Third case using the non-negativity constraint.

Table 2 shows the value of NMSE over 10 Monte-Carlo runs for several versions of the proposed method with different compression ratios over the Cuprite dataset. In this experiment we can see that the version with the non-negativity constraint provide better results than the other versions. Here, we observe a slightly worse performance in the second case which is probably due to the fact that we are estimating the endmembers from the original data. For illustrative purposes, Fig. 1 shows the reconstructed and the original spectral signatures with highest, average and lowest error for the C-HYCA algorithm (with the non-negativity constraint) for different compression ratios. In this plot, we can see that, even in the worst case, the reconstructed pixel preserves the shape of the original pixel, which means that the features of the signature are well-preserved. In the other two cases the reconstructed and the original pixels are extremely similar.

4. CONCLUSION

In this paper we discuss a compressive sensing algorithm called hyperspectral coded aperture (HYCA) and its constrained version (C-HYCA). HYCA framework takes advantage of two main properties of hyperspectral data, namely the high spatial correlation of abundance fractions and the low number of endmembers to explain the observed data. The



(a) $q = 5$, worst case (b) $q = 5$, average case (c) $q = 5$, best case

(d) $q = 15$, wort case (e) $q = 15$, average case (f) $q = 15$, best case

Fig. 2. Worst (a,d), average (b,e), and best (c,f) reconstructed pixels in the AVIRIS Cuprite scene for values of $q = 5$ and $q = 15$, respectively.

.

former property is exploited by minimizing the total variation (TV) of the reconstructed images of abundance fractions and the latter property is exploited by formulating the reconstruction problem with respect to abundance fractions, which have much lower dimension than the original data. While HYCA depends on the tuning of a regularization parameter controlling the relative weight between the TV regularizer and the data term, C-HYCA does not depend on any regularization parameter. As demonstrated by our experiments with synthetic and real hyperspectral data, both approaches provide good results in the task of compressing remotely sensed hyperspectral data sets and are strongly related with the concept of spectral unmixing that has been widely used to interpret hyperspectral data.

5. REFERENCES

[1] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectrometry for Earth remote sensing," *Science*, vol. 228, pp. 1147–1153, 1985.

[2] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. Solis *et al.*, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sensing of Environment*, vol. 65, no. 3, pp. 227–248, 1998.

[3] G. Motta, F. Rizzo, and J. A. Storer, *Hyperspectral data compression*. Berlin: Springer, 2006.

[4] B. Huang, *Satellite data compression*. Berlin: Springer, 2011.

[5] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, pp. 1289–1306, 2006.

[6] J. E. Candès, T. Romberg, and Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, p. 1207, 2006.

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

[8] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing: geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2012.

[9] D. Donoho and M. Elad, "Optimal sparse representation in general (non-orthogonal) dictionaries via l1 minimization," *Proceedings of the National Academy of Sciences*, vol. 100, pp. 2197–2202, 2003.

[10] I. Gorodnitsky and B. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 600–616, 1997.

[11] S. Foucart and M. Lai, "Sparsest solutions of underdetermined linear systems via lq-minimization for 0," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 395–407, 2009.

[12] D. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 6–18, 2006.

[13] J. Tropp, "Just relax: convex programming methods for subset selection and sparse approximation," *ICES report*, pp. 04–04, 2004.

[14] B. Natarajan, "Sparse approximate solutions to linear systems," *SIAM journal on computing*, vol. 24, no. 2, pp. 227–234, 1995.

[15] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, "The cosparse analysis model and algorithms," *Applied and Computational Harmonic Analysis*, 2012.

[16] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 681 –695, 2011.

[17] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, 2008.

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