JOINT TRACE/TV NORM MINIMIZATION: A NEW EFFICIENT APPROACH FOR SPECTRAL COMPRESSIVE IMAGING

Mohammad Golbabaee and Pierre Vandergheynst

Signal Processing Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
E-mail: {mohammad.golbabaei, pierre.vandergheynst}@epfl.ch

ABSTRACT

In this paper we propose a novel and efficient model for compressed sensing of hyperspectral images. A large-size hyperspectral image can be subsampled by retaining only 3% of its original size, yet robustly recovered using the new approach we present here. Our reconstruction approach is based on minimizing a convex functional which penalizes both the trace norm and the TV norm of the data matrix. Thus, the solution tends to have a simultaneous low-rank and piecewise smooth structure: the two important priors explaining the underlying correlation structure of such data. Through simulations we will show our approach significantly enhances the conventional compression rate-distortion tradeoffs. In particular, in the strong undersampling regimes our method outperforms the standard TV denoising image recovery scheme by more than 17dB in the reconstruction MSE.

Index Terms— Hyperspectral images, Compressed sensing, Low rank matrix recovery, TV norm, Trace norm, Convex optimization.

1. INTRODUCTION

Over the last decade compressive sampling (CS) theory has been introduced as an alternative to the Shannon's sampling theorem, wherein the main idea consists of combining the two conventional sampling and compression steps together in order to represent data with sampling rates much lower than the Nyquist rate [1] [2]. Instead of taking n periodic Nyquist samples to discretize data into the vector $x \in \mathbb{R}^n$, $m \ll n$ linear measurements are collected from data in a vector $y \in R^m$ that can be expressed as $y = Ax + z$, where each measurement is the inner product of x and a row of the sampling matrix $A \in \mathbb{R}^{m \times n}$, and $z \in \mathbb{R}^n$ represents the noise vector due to the quantization, transmission, etc. It has been shown that, for signals having compact *sparse* representations in an orthonormal basis (i.e., $x = \Phi\theta$, $\Phi \in \mathbb{R}^{n \times n}$, $\theta \in \mathbb{R}^n$ with $\|\theta\|_{\ell_0} \leq k \ll n$), and for random sampling matrices drawn from certain distributions (e.g., A being an i.i.d subgaussian matrix), a robust reconstruction is achievable by taking $m \gtrsim \mathcal{O}\left(k \log(n/k)\right)$ measurements and solving the following convex problem, called the *Basis Pursuit Denoising* (BPDN):

$$\arg \min_{\theta \in \mathbb{R}^n} \|\theta\|_{\ell_1} \qquad \text{subject to} \qquad \|y - A\Phi\theta\|_{\ell_2} \leq \varepsilon. \qquad (1)$$

Similar results have been obtained regarding compressed sampling of low-rank matrices. Assume a linear mapping $\mathcal{A} : \mathbb{R}^{n_1 \times n_2} \rightarrow$

m applies on a rank $r \ll \min(n_1, n_2)$ data matrix $X \in \mathbb{R}^{n_1 \times n_2}$ to collect $m \ll n_1 n_2$ linear measurements $y \in \mathbb{R}^m$ i.e.,

$$y = \mathcal{A}(X) + z. \qquad (2)$$

Candés and Plan [3] show that, for an $\mathcal{A}$ which is drawn at random (e.g, i.i.d. subgaussian ensemble), a robust recovery of X is achievable from the following convex *trace norm minimization*

$$\underset{X \in \mathbb{R}^{n_1 \times n_2}}{\arg\min} \|X\|_* \qquad \text{subject to} \qquad \|y - \mathcal{A}(X)\|_{\ell_2} \leq \varepsilon, \qquad (3)$$

provided by $m \gtrsim \mathcal{O}(rn_1 + rn_2)$, that is the same order as the degrees of freedom of such low-rank matrix. Note that in (3), $\|.\|_*$ stands for the trace norm (equivalently, the nuclear norm) of a matrix that is the sum of its singular values.

Lately, it has been shown in [4] that, if the sparsity and low-rank assumptions simultaneously hold for data i.e., X being a k-joint-sparse (few k rows/columns of X are nonzero) and low-rank, then one can expect a robust recovery by taking $m \gtrsim \mathcal{O}\left(rk \log(k/n_1) + rn_2\right)$ measurements and solving the following joint trace-$\ell_{2,1}$ norm minimization (for some regularization parameter λ):

$$\underset{X \in \mathbb{R}^{n_1 \times n_2}}{\arg\min} \|X\|_* + \lambda\|X\|_{\ell_{2,1}} \qquad (4)$$
$$\text{subject to} \qquad \|y - \mathcal{A}(X)\|_{\ell_2} \leq \varepsilon$$

The key observation here is that all these methods aim for a robust representation of high dimensional data provided by few non-adaptive linear measurements proportional to the underlying degrees of freedom.

Throughout these years, a massive amount of research has been conducted to benefit from the CS theory in varieties of signal processing applications. In this regard, a few number of image acquisition systems have been proposed based on the CS idea, for example in [5, 6], in order to reduce the number of the photodiodes (some designs such as the Rice' Single-pixel camera have been already implemented and examined). Two standard and popular methods that are extensively used by the community to reconstruct the original image from the CS measurement are i) BPDN using varieties of wavelet basis for Φ due to the well-established capability of the wavelet transforms for compact image representations, and ii) the following *Total Variation norm minimization/Denoising* (TVDN):

$$\arg \min_{x \in \mathbb{R}^n} \|x\|_{TV} \qquad \text{subject to} \qquad \|y - Ax\|_{\ell_2} \leq \varepsilon. \qquad (5)$$

The Total Variation norm $\|.\|_{TV}$ of a 2-dimensional image x has been introduced in [7] as an efficient prior to minimize for the standard image denoising problem.

1.1. Hyperspectral Compressive Imagery

Hyperspectral Images (HSI) are huge collection of images that are acquired simultaneously from a scene in a few hundred narrow adjacent frequency bands. All substances have their own specific *spectral signature* or frequency absorption features and therefore once the frequency bands are sampled with highly enough resolution, Hyperspectral imagery becomes a very powerful tool for characterizing the components of the observed scenes. As a result, this type of images finds a wide variety of applications in remote sensing such as detection and identification of the ground surface as well as atmospheric composition, analysis of soil type, agriculture, mineral exploration and environmental monitoring. The price to pay for such high spatio-spectral resolution is to handle extremely large data size. For example, each instance of the HSI acquired by the NASA's Airborne Visible InfraRed Imaging Spectrometer (AVIRIS) contains 224 spectral bands and roughly more than 140 MBytes of data.

Compressive sampling becomes particularly interesting for hyperspectral imagery mainly because of two reasons that are both related to the enormous amount of data in such application: i) the photodiodes sensitive to the light outside the visible range are severely costly and thus, incorporating a huge number of them into a high-resolution sensor makes hyperspectral imagers extremely expensive devices, ii) handling such huge amount of informations, even at the compression step, brings serious challenges, particularly to embedded systems, such as spacecrafts, where the power consumption, memory storage, computational complexity and bandwidth are posing tight constraints on system implementation. Despite their high dimensionality, hyperspectral images are known to contain massive correlations along both spatial and spectral domains and with such inspiration, there has been many works over the past twenty years on compression of HSI. Considering such enormous redundancy, one major question would naturally arise; Why do we need to make such costly effort to entirely acquire data while only a very small amount of it is kept after yet another computationally-expensive compression procedure?

In order to address this question, a few number of novel designs have been recently proposed based on the new CS sampling paradigm, aiming for spectral image acquisition by very few number of measurements [8] [9]. The main idea is to benefit from the *asymmetric* complexity of CS (i.e., low-complex sampling, high-complex reconstruction) to overcome the aforementioned practical limitations. For data reconstruction, the authors of [9] apply the standard TVDN approach in (5) independently on each spectral band in order to find smooth spatial images with few gradient variations. This approach clearly neglects the existing correlations in the spectral domain. For joint CS reconstruction, Duarte and Baraniuk [10] additionally take into account the piecewise smooth variation of HSI along the spectral domain and reconstruct a sparse representation of HSI in a 3D wavelet basis using the standard BPDN method i.e., Φ in (1) would be an orthonormal basis formed by the *Kronecker* product of a 2D spatial and a 1D spectral wavelet basis. Recently, its has been shown in [11] that incorporating the low-rank structure of the HSI can better handle the underlying spatial correlations, and together with the assumption of spatial sparsity in a 2D wavelet basis, one can achieve a significant enhancement in HSI reconstruction for severely under-sampled regimes by using the joint trace/$\ell_{2,1}$ minimization approach in (4).

1.2. Our Main Contributions

This paper aims to present a novel CS reconstruction method based on the joint trace/TV norm minimization and with applications in spectral compressive imaging. Our approach brings two important priors of HSI into the consideration; (i) sparse gradient variations along the spatial domain, and (ii) the low-rank structure due to the high correlations. The recovery problem is relaxed into a convex minimization that is implemented using the proximal splitting method [12] with the capability of a parallel realization. We show through simulations that the global minima of the proposed approach significantly enhances the HSI reconstruction, compared to the standard TVDN-based scheme. Notably and in contrast with the conventional decorrelation-based HSI compression techniques such as [13], the sensor side does not need to compute or transmit any correlation matrix and its complexity reduces to taking few number of linear measurements and forward them to the decoder.

2. COMPRESSIVE HYPERSPECTRAL IMAGERS

In order to represent hyperspectral images we define a matrix $X \in \mathbb{R}^{n_1 \times n_2}$ whose jth column, say X_j, corresponds to a 2D spatial image (reshaped in a vector) in the corresponding spectral band j. Here n_2 denotes the number of spectral bands and n_1 is the resolution of the spatial images in each band. In a compressive acquisition setup, sensors are collecting $m \ll n_1 n_2$ linear measurements from the HSI matrix X in a vector $y \in \mathbb{R}^m$ following the same sampling model as in (2). Note that $\mathcal{A}$ can be explicitly expressed by a matrix $A \in \mathbb{R}^{m \times n_1 n_2}$ through the equivalent expression

$$y = AX_{vec} + z,$$

where the columns of X are stacked into the vector X_{vec}.

Several camera designs have been so far proposed for the single-channel image compressive acquisition. A common point among those is the use of a *random pattern* to modulate the light prior to collecting the measurements. As an example, the random convolution measurement scheme have been proposed by [6], convolves the image light-field with a random pattern using few optical blocks, and finally a few number of random pixels have been acquired from the resulting modulation. All those setups can be easily extended to hyperspectral imaging, by repeating the same acquisition scheme for all spectral bands, however by using an *independent random pattern per channel*. In this case the corresponding measurement matrix A, would be a block diagonal matrix of the form

$$A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_J \end{bmatrix}, \tag{6}$$

where $A_j \in \mathbb{R}^{\widehat{m} \times N}$ is the random measurement matrix that applies on channel j independently from the other spectral bands, and $\widehat{m}$ denotes the number of measurements collected per channel i.e., $m = \widehat{m}J$. In contrast with the Single-pixel hyperspectral imager in [8] which uses a unique random pattern for all spectral bands (i.e., $A_1 = A_2 = \dots = A_J$), using independent blocks as in (6) leads the measurements to benefit more efficiently from the existing information *diversity* across multiple spectral channels.

For more real-time acquisitions, another camera design has been implemented by Wagadarikar *et al*. The Coded Aperture Snapshot Spectral Imager (CASSI) captures a few thousands of CS measurements in snapshot. There, by using optical light modulators the whole spectral information are encoded into a single 2D spatial image, and thus the sampling matrix has a different form than in (6) (for more details see [9]).

3. CS RECONSTRUCTION BASED ON THE JOINT TRACE/TV NORM MINIMIZATION

Perhaps the best way to characterize the compressible priors that will be useful for an efficient CS reconstruction is by looking at the extensive literature of the compression schemes for hyperspectral images. One of the most efficient approaches [13] consists in 2D wavelet coding for the spatial domain, as the natural images can be typically represented by few sparse wavelet coefficients. In addition, a Karhunen-Loéve transform (KLT) is applied to compress data into few principal components along the spectral dimension. The KLT-based approach is data-dependent and in practice costs heavy computations and transmission (to the decoder) of the correlation matrix, however its high efficiency reveals an important point about the data structure and that is, *HSI are typically piecewise smooth along the spatial domain and in addition, they contain very few principal components implying their low-rank structure.*

We propose the following convex optimization for reconstruction of the original HSI from its CS measurements:

$$\underset{X \in \mathbb{R}^{n_1 \times n_2}}{\operatorname{argmin}} \ \|X\|_* + \lambda \sum_{j=1}^{n_2} \|X_j\|_{TV} \qquad (7)$$
$$\text{subject to} \qquad \|y - \mathcal{A}(X)\|_{\ell_2} \leq \varepsilon.$$

The TV norm of a 2D image is defined as the sum of the magnitude of the gradient over all image pixels (u, v) i.e., for each spectral band we have

$$\|X_j\|_{TV} = \sum_{u,v}^{n_1} \left| (\nabla X_j)_{u,v} \right|.$$

As previously mentioned, penalizing the trace norm *or* the TV norm has been widely used in the literature to impose the reconstructed data to have a low-rank or piecewise smooth structure, respectively. However, by penalizing both terms with a proper regularization factor λ in (7), we tend to impose the solutions satisfy *simultaneously* both properties (penalizing the sum of the TV norms implies a piecewise smooth structure along the spatial domain for every spectral band). As a result, applying this approach can efficiently take advantage of the limited degrees of freedom of such data structure in order to economize the number of measurements required for the HSI reconstruction.

In order to solve (7) we use the Parallel Proximal Algorithm (PPXA) proposed by [12]. PPXA is an iterative method for solving convex minimizations with the capability of parallel implementation and in the case of solving (7), it briefly consists of three main steps to proceed with the current solution $X^{[t]}$ at iteration t: i) the singular values soft thresholding ii) multiple TVDN minimizations, each applies to a certain spectral band (they can be performed in parallel) and, iii) projection onto the convex set of matrices satisfying the fidelity constraint $\|y - \mathcal{A}(X)\|_{\ell_2} \leq \varepsilon$. These three steps can run in parallel and at the end of each iteration their results are averaged together for updating the solution. The whole procedure repeats up to a convergence point (for more details see [11, 12]). Note that the orthogonal projection onto the ℓ_2 ball (the fidelity constraint) in the last step can be computed iteratively as suggested by [14] (and within a single iteration, once $\mathcal{A}$ is a tight frame). Finally, for the regularization parameters in (7), we set $\varepsilon = \|z\|_{\ell_2}$ that is the estimated noise power, and we suggest $\lambda \sim \sqrt{r/kn_2}$ which requires a rough estimation of the rank r of the HSI matrix and the sparsity-level k of the spatial image gradient per spectral band.
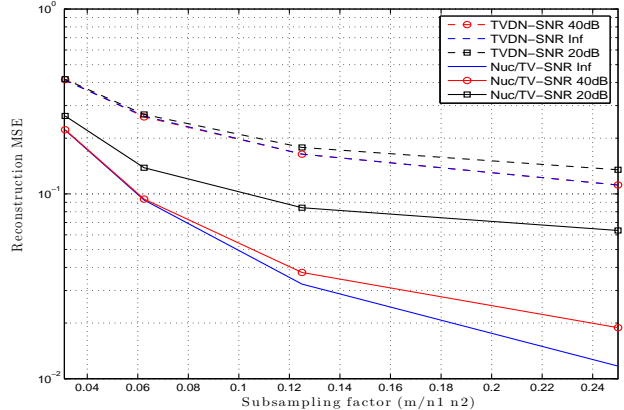

Fig. 1: Reconstruction MSE vs. Subsampling factor, using the joint Trace (Nuclear)/TV norm minimization (Nuc/TV) and TV denoising (TVDN) recovery schemes, and for sampling SNR= ∞, 40dB and 20dB.

4. NUMERICAL EXPERIMENTS

We evaluate the performance of our approach on the standard Moffett field hyperspectral dataset[1]. We crop the corresponding HSI to have the spatial resolution $n_1 = 256 \times 256$, and in addition we select $n_2 = 180$ spectral bands after discarding water absorption bands. The ordered singular values of the HSI matrix have a very fast decay in their magnitudes (though it is not demonstrated in this note), and the whole HSI can be constructed with 2×10^{-4} normalized MSE by considering the first seven principal components. Therefore, such HSI can be well approximated by a low-rank matrix.

For compressive acquisition, a block diagonal random matrix similar to (6) has been applied to collect m linear random measurements from data as in (2). An independent random convolution measurement matrix [6] has been used for each spectral band. In this case, the resulting sampling operator/matrix is a tight frame. Note that, in our experiments the CS measurements are corrupted by the additive white Gaussian noise whose power corresponds to the "sampling SNR".

For the CS reconstruction, we apply two methods namely, the TVDN approach in (5) separately applied on each spectral band and, our trace/TV norm minimization in (7) applied on the whole data for a low-rank and spatially piecewise smooth HSI joint recovery. Figure 1 compares the performance of both methods in terms of normalized reconstruction MSE (i.e., $\|\widehat{X}_{vec} - X_{vec}\|_{\ell_2} / \|X_{vec}\|_{\ell_2}$) for different subsampling factors ($\frac{m}{n_1 n_2}$) and different sampling SNRs (i.e., the energy of noise z). There is an evident gap between the performances of these two methods: the trace/TV norm minimization always outperforms TVDN. In particular, for severely undersampled regimes (when the number of CS measurements are about 3% of the whole HSI cube size) and for moderate sampling SNRs, our recovery approach indicates more than 17dB improvement comparing to the standard TVDN approach. Figure 2 illustrates the reconstruction quality suggested by both methods. As we can see, the joint trace/TV norm minimization is able to robustly perform the HSI reconstruction for severely under-sampling rates for which the other standard CS recovery methods such as TVDN totally fail.

---

[1]Dataset available at http://aviris.jpl.nasa.gov/html/aviris.freedata.html
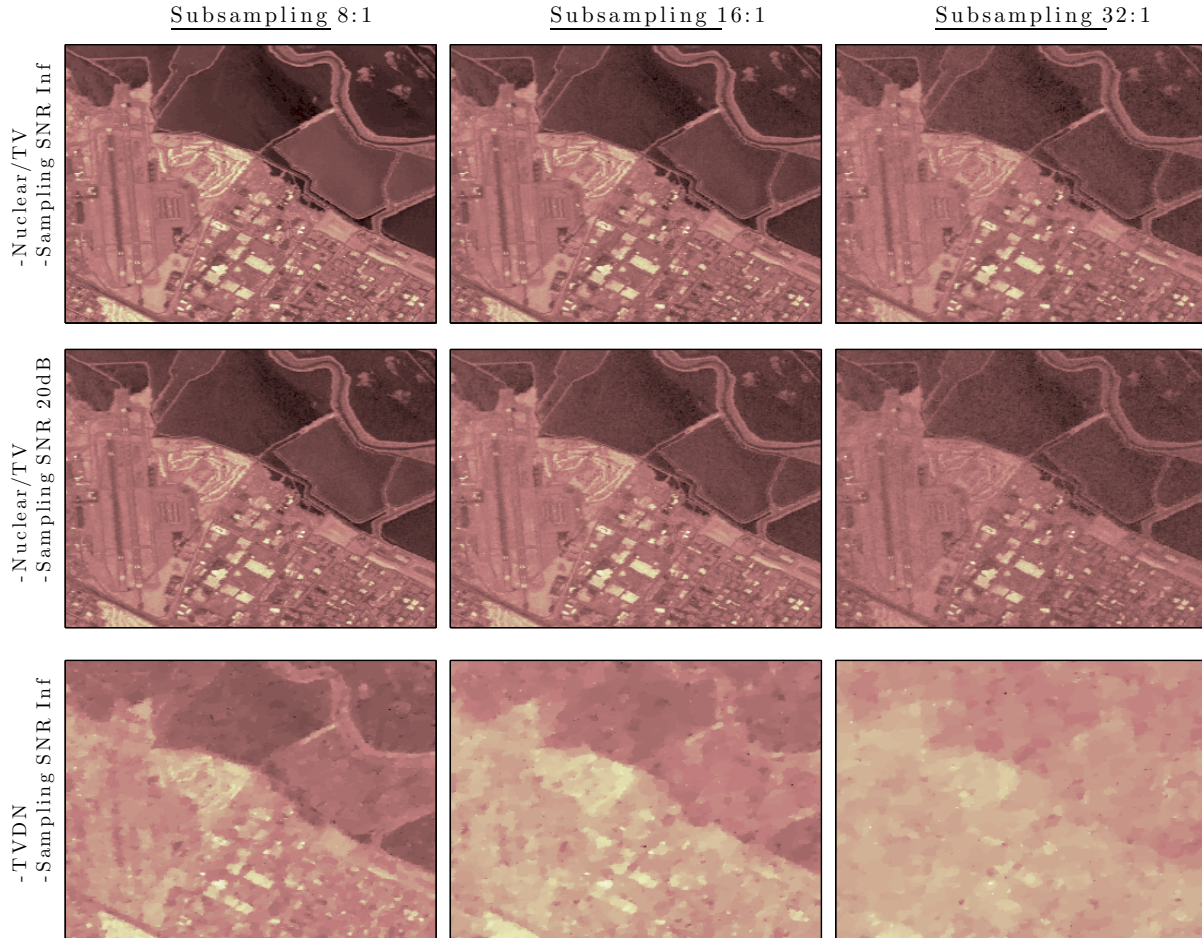
Fig. 2: Reconstruction of the Moffett field HSI using the joint Trace (Nuclear)/TV norm minimization (Nuclear/TV) and TVDN, under various subsampling ratios and sampling SNRs. Results are demonstrated for the spectral band j=50.

5. CONCLUSIONS

In this paper we have proposed a novel convex optimization formulation for recovering the spectral images from very few CS measurements. Our approach penalizes both the trace norm and the TV norm of the data matrix in order to reconstruct simultaneously the low-rank and spatially piecewise smooth structure of the data. An algorithm to solve this convex optimization has been proposed based on proximal splitting methods. Through simulations we have shown that, our approach is robust against noise and the number of measurements required for the HSI reconstruction is significantly reduced comparing to the standard TVDN method.

6. REFERENCES

[1] D.L. Donoho, "Compressed sensing," *IEEE Trans. on Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] E. J. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements.," *Pure Appl. Math.*, vol. 59, pp. 1207–1223, 2005.

[3] E. J . Candes and Y. Plan, "Tight oracle bounds for low-rank matrix recovery from a minimal number of random measurements.," *IEEE Trans. on Inf. Theory*, 2009.

[4] M. Golbabaee and P. Vandergheynst, "Guaranteed recovery of a low-rank and joint-sparse matrix from incomplete and noisy measurement," in *Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS11)*, 2011.

[5] M.F. Duarte, M.A. Davenport, D. Takhar, J.N. Laska, T. Sun, K.F. Kelly, and R.G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 83–91, 2008.

[6] J. Romberg, "Compressive sensing by random convolution," *SIAM J. Imaging Sciences*, 2009.

[7] 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.

[8] T. Sun and K.F. Kelly, "Compressive sensing hyperspectral imager," *Comp. Optical Sensing and Imaging (COSI), San Jose, CA, Oct. 2009.*

[9] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Applied Optics*, vol. 47, pp. B44–B51, 2008.

[10] M.F. Duarte and R.G. Baraniuk, "Kronecker Compressive Sensing," *to appear in the IEEE Trans. on Image Processing*, 2009.

[11] M. Golbabaee and P. Vandergheynst, "Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2012.

[12] P. L. Combettes and J. C. Pesquet, "Proximal splitting methods in signal processing," *in: Fixed-Point Algorithms for Inverse Problems in Science and Engineering, Springer-Verlag*, vol. 49, pp. 185–212, 2011.

[13] Q. Du and J. E. Fowler, "Hyperspectral image compression using jpeg2000 and principal component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 201–205, April 2007.

[14] Mohamed-Jalal Fadili and Jean-Luc Starck, "Monotone operator splitting for optimization problems in sparse recovery," in *ICIP*, 2009, pp. 1461–1464.