

# EXPLOITING STRUCTURAL COMPLEXITY FOR ROBUST AND RAPID HYPERSPECTRAL IMAGING

Gregory Ely, Shuchin Aeron and Eric L. Miller

Department of ECE, 161 College Avenue, Tufts University, Medford, MA, 02155

## ABSTRACT

This paper presents several strategies for spectral de-noising of hyperspectral images and hypercube reconstruction from a limited number of tomographic measurements. In particular we show that the non-noisy spectral data, when stacked across the spectral dimension, exhibits low-rank. On the other hand, under the same representation, the spectral noise exhibits a banded structure. Motivated by this we show that the de-noised spectral data and the unknown spectral noise and the respective bands can be simultaneously estimated through the use of a low-rank and simultaneous sparse minimization operation without prior knowledge of the noisy bands. This result is novel for for hyperspectral imaging applications. In addition, we show that imaging for the Computed Tomography Imaging Systems (CTIS) can be improved under limited angle tomography by using low-rank penalization. For both of these cases we exploit the recent results in the theory of low-rank matrix completion using nuclear norm minimization.

**Index Terms**— Hyperspectral imaging, de-noising, Limited angle tomography, low-rank recovery.

## I. INTRODUCTION

This paper addresses two specific image reconstruction challenges encountered in the field of hyperspectral imaging: de-noising in the presence of spectral noise and hypercube reconstruction from a limited set of Radon projections similar to angle limited Computed Tomography Imaging Systems (CTIS).

The first of these two problems is motivated by the desire to remove noise at specific frequency bands from hyperspectral image cubes. This problem frequently arises when using satellites or aircraft to capture hyperspectral images of the earth in which the light reflecting from the surface of the earth must travel through several kilometers of atmosphere to the sensor. The atmosphere even without the presence of clouds has extremely high absorption bands, particularly at 1400 nm and 1900 nm due to water in the atmosphere [1]. This effect leads to numerous bands being discarded for many data classification and analysis algorithms [2] [3].

In order to mitigate the effects of both this spectral and electronic noise several de-noising techniques such as multidimensional Weiner filtering [4] and methods exploiting the use of high order singular value decomposition [5], curvelets [6], and wavelets [2] [7] have been used to de-noise these effects. However, both the intensity dependence the noise and the concentration across a few spectral bands makes the removal of optical noise challenging [8]. Many of these techniques are based on the premise of noise being AWGNG and performance can be poor [9]. Typically a preprocessing (whitening) step is needed to mitigate the effects of the Poisson noise and improve performance [9]. Recently, efforts to de-noise spectral bands have focussed on the use of sparse or

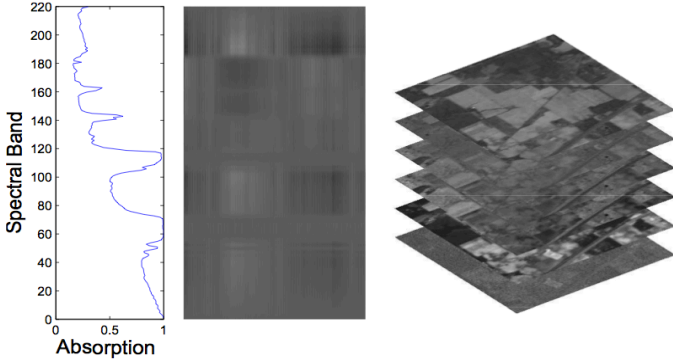
joint penalizations in an appropriate basis such as wavelets [10] and dictionary learning techniques [11].

In this paper we will explore a novel spectral de-noising technique based on a low-rank and simultaneously sparse matrix decomposition. The low-rank sparse matrix decomposition or Robust Principle Component Analysis (RPCA) has been well studied and theoretical limits well characterized in recent years [12] [13]. Furthermore, RPCA has been successfully employed in image and video processing to separate background from the foreground [14] and remove ‘salt and pepper’ noise from imagery [12]. However, little research has been done to explore variations of RPCA such as a low-rank group sparse decomposition and its potential applications. In particular, Tang proposed a feasible solution to solving the group RPCA problem through the method of Augmented Lagrange Multipliers [15] and Ji demonstrated the use of group RPCA to de-noise video data [16]. This paper provides another potential application and extension of RPCA to CTIS systems.

In the second part of this paper we will focus on the problem of estimation of the hyperspectral data cube from limited number of tomographic projections. Here we show how the use of low-rank regularization can be used to improve an existing class of hyperspectral imagers. These hyperspectral imagers [17], [18], [19] sample the hyperspectral image cube by *simultaneously* (i.e. not sequentially) taking a number of Radon type projections of the 3D data cube onto a 2D focal plane array using diffractive optics. Traditionally filtered back-projection methods have been employed to recover the data cube from these tomographic projections. However, these techniques need a large number of projections to ensure accurate results, avoid the so called missing cone problem [20] and often fail in noisy environments. This need for a large number of projections increases the necessary focal plane size beyond what is often feasible. In this context we demonstrate how one can exploit the low-rank regularization to improve the reconstruction under these classes of simultaneous and *compressive measurements*. Note that although some research has focussed on the use of both sparse and low-rank reconstructions of hyperspectral image cubes, these studies use *practically infeasible* sampling techniques such as randomly sampling a small set of pixels within the image cube [21].

## II. STRUCTURAL COMPLEXITY OF HYPERSPECTRAL IMAGES

A hyperspectral image or data cube consists of many images of the same size collected over a number of spectral bands. Mathematically the hyperspectral image can be considered as a three-dimensional matrix  $\mathbf{L} \in \mathbb{R}^{m \times n \times l}$  with spatial dimensions of  $m$  and  $n$  pixels and at  $l$  wavelengths. One can reshape the hyperspectral image as a two-dimensional array with a number of columns equal to the number of spectral bands and where each column is the vectored image at the given band, see Figure 1.



**Fig. 1.** Left: Normalized total counts in the AVIRIS image as a function of band. We see two pronounced absorption bands. Right & Center: This figure shows a 3D and 2D representation of a hyperspectral image. One notices the horizontal bands of spectral noise in the two dimensional image that align with the absorption bands. Much of the structure in the matrix appears to be vertical but the horizontal bands are spectral noise at absorption bands.

With slight abuse of notation we denote the reshaped image by  $\mathbf{L} \in \mathbb{R}^{mn \times l}$ . We now present two observations regarding the structural complexity of the image data which will be exploited for recovery and de-noising.

#### II-A. Low-rank structure of the hyperspectral data cube

Although a hyperspectral image/data may have numerous bands, it has been shown that signal subspace is significantly smaller than the number of bands [22] [23]. In particular the eigenvalues of the reshaped hyperspectral cube  $\mathbf{L}$  obey a power law decay. This means the vector of eigenvalues has a small weak- $\ell_p$  norm [24] which implies that image is compressible under the suitable transformation. This intuition can be physically explained by considering the Singular Value Decomposition (SVD) of the (reshaped) hyperspectral matrix  $\mathbf{L}$ .

$$\mathbf{L} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^* \quad (1)$$

We can think of the right singular vectors as giving the spectra of the common elements in the scene and the left singular values as the concentration map of these spectra. The singular values then give the relative amount each compound in the scene. Low-rank of the image can then be interpreted as presence of a few spectra with a correlated concentration profile across space.

#### II-B. Sparsity structure of hyperspectral noise

In hyperspectral imaging the atmosphere can lead to vastly different absorption rates across the spectrum of interest. In particular as shown in Figure 1, the two water absorption bands are attenuated, roughly at band 60 and 100. In a typical hyperspectral data processing the data from these two bands would be discarded. On the other hand we note that in the noisy reshaped image, the spectral noise exhibits a banded structure which is mathematically equivalent to saying that hyperspectral noise exhibits a simultaneous sparse structure under the given reshaping of the data cube.

Therefore, the noisy reshaped hyperspectral data cube can be represented as  $\mathbf{Y} = \mathbf{L} + \mathbf{S}$  where  $\mathbf{L}$  is the low-rank non-noisy image and  $\mathbf{S}$  is the spectral noise which is simultaneously or group sparse across bands.

### III. ROBUST & RAPID HYPERSPECTRAL IMAGING

Both the spectral de-noising and limited angle reconstruction problems can be viewed through the following framework in which we observe noisy measurement,  $\mathbf{Y}$  of hyperspectral image cube  $\mathbf{L}$  through a measurement system described by the linear operator (matrix)  $\mathcal{A}$ , i.e.

$$\mathbf{Y} = \mathcal{A}(\mathbf{L} + \mathbf{S}) \quad (2)$$

The problem is that given the observation  $\mathbf{Y}$  and the sensing operator  $\mathcal{A}$  (to be defined below for both problems of interest) we want to recover the de-noised image  $\mathbf{L}$  while removing the noise  $\mathbf{S}$ .

#### III-A. Complexity penalized recovery algorithms

To de-noise and recover the hyperspectral data, one can exploit the low-rank and sparse structure of the data and noise and solve the following optimization,

$$\min_{\mathbf{L}, \mathbf{S}} \|\mathcal{A}(\mathbf{L} + \mathbf{S}) - \mathbf{Y}\|_2^2 + \lambda_L \text{rank}(\mathbf{L}) + \lambda_S \|\mathbf{S}\|_{0,2} \quad (3)$$

where  $\lambda_L$  &  $\lambda_S$  control the relative strength of the sparsity and low-rank penalization and  $\|\mathbf{S}\|_{p,q}$  is the  $p$ -norm of the vector formed by taking the  $q$ -norm along the rows of  $\mathbf{S}$  or otherwise also known as  $\ell_{p,q}$  norm. This optimization problem is known to be NP-hard. However, the rank and support penalties can be relaxed to the nuclear norm and  $\ell_{1,2}$  norm, respectively, which makes the optimization tractable while still encouraging the desired structure for  $\mathbf{L}$  and  $\mathbf{S}$  [12]. Therefore we relax the above combinatorial optimization problem to the following convex optimization problem and consider three cases.

$$\min_{\mathbf{L}, \mathbf{S}} \|\mathcal{A}(\mathbf{L} + \mathbf{S}) - \mathbf{Y}\|_2^2 + \lambda_L \|\mathbf{L}\|_* + \lambda_S \|\mathbf{S}\|_{1,2} \quad (4)$$

**Case I - Hyperspectral de-noising with raster scan data** - In this case  $\mathcal{A}$  is an identity operator and therefore the optimization problem becomes,

$$\min_{\mathbf{L}, \mathbf{S}} \|\mathbf{L} + \mathbf{S} - \mathbf{Y}\|_2^2 + \lambda_L \|\mathbf{L}\|_* + \lambda_S \|\mathbf{S}\|_{1,2} \quad (5)$$

In Section IV we will demonstrate the performance of this algorithm on real hyperspectral data and give experimental results that motivate why the sparse component is necessary for the de-noising.

#### Case II. - Image recovery from limited angle tomography:

**No spectral noise** - As pointed out in the introduction the CTIS systems are limited by the size of focal plane array which limits the number of tomographic projections that can be obtained. In this case traditional reconstructions suffer from the missing cone problem [20]. These methods however do not exploit the low complexity of the underlying data cube. Assuming no spectral noise, given the limited number of Radon projections we propose the following algorithm for estimation of the hyperspectral image which exploits the low-rank structure.

$$\min_{\mathbf{X}} \|\mathcal{A}(\mathbf{X}) - \mathbf{Y}\|_2 + \lambda \|\mathbf{X}\|_* \quad (6)$$

#### Case III. - Image recovery from limited angle tomography:

**Noisy case** - Here we consider the most general case where the spectral data is corrupted by banded spectral noise and the data is acquired through a CTIS system with limited number of Radon projections. In this case simultaneous spectral cube recovery and spectral de-noising is affected by solving for the optimization

problem given by equation 4. In the next section we will present detailed experimental results of the proposed algorithms on real data sets.

#### IV. EXPERIMENTAL EVALUATION

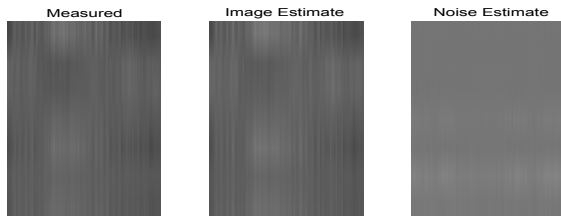
In this section we will use a real hyperspectral image taken from Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) [website- <http://aviris.jpl.nasa.gov/html/aviris.spectrum.html>], of an aerial scene with a spatial dimension of 128 by 128 pixels. The imager uses 220 bands which cover the spectrum from IR to visible range. The two water absorption bands centered at 1400 and 1900 nm corrupt the image. **NOTE:** All the optimization problems below are implemented using TFOCS [25].

**Case I. - Hyperspectral de-noising** - In some of the less noisy bands the structure of the image is still somewhat visible (figure 3). In order to improve the de-noising of the AVIRIS data we first take and record the Frobenius norm of each frame to construct a  $N_\lambda \times 1$  vector  $W$ . We then use this recorded vector to normalize each image at given wavelength such that the signal energy in each band is 1. Because we expect the noise in our experiment to be due to low photon counts in bands of high absorption, we can use the vector  $W$  to weight the minimization operation. In particular we want to encourage row sparsity along the bands with low counts. In order to do so we modify equation 5 to include the weighting factor  $W$ , that makes it more expensive for the intense bands to be decomposed into the sparse matrix.

$$\min_{L, S} \| (L + S) - Y \|_2^2 + \lambda_L \| L \|_* + \lambda_S \| WS \|_{1,2} \quad (7)$$

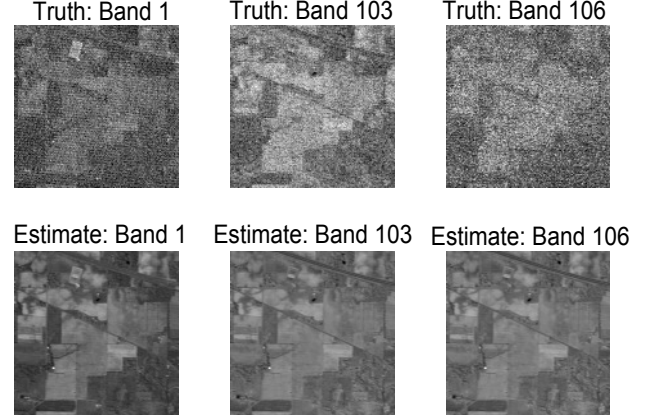
This weighting factor allows our algorithm to be more robust to choices of  $\lambda_S$  and  $\lambda_L$  as it effectively decreases the coherence between the  $\ell_{1,2}$  norm and the nuclear norm.

The minimization operation in Equation (7) is then applied to the hyperspectral image with  $\lambda_S$  of .06 and  $\lambda_L$  of 0.1. The proposed algorithm was successful in de-noising and was able to remove the spectral noise. Figure 3 shows the results of the algorithm applied to a few very noisy bands and figure 2 shows the results applied to all bands of the hyperspectral image. Like in the synthetic example, we can now see structure in the bands that were previously noisy.



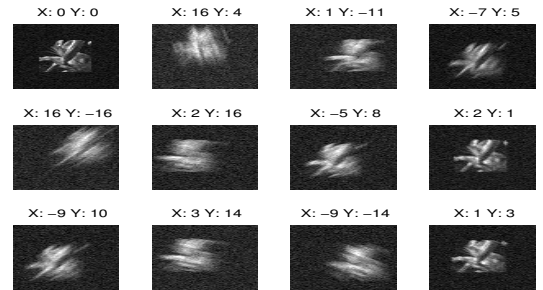
**Fig. 2.** This figure shows 2D hyperspectral cube with noise and low-rank reconstruction.

**Case II. - Hyperspectral imaging from limited Radon projections with no spectral noise** - In the following example we attempt to reconstruct the 32 by 32 image of the hyperspectral flower [26] using a limited number of projections. The projections at various angles for a typical single-shot CTIS system [17], [27] are shown in Figure 4. Gaussian noise was then added to the measured projections, such that the resulting SNR of the projections was 4.5 dB. This projection operation can be represented through the



**Fig. 3.** This figure shows images from AVIRIS data at various bands before de-noising and after de-noising.

underdetermined matrix  $\mathcal{A}$ . In this case we solve the optimization problem (6) for recovery. We compare the performance of this method to the standard Tikhonov regularization approach with  $\ell_2$  norm penalty instead of nuclear norm penalty. The choice of  $\lambda$  for both cases were determined using the Kolmogorov-Smirnov (KS)-test method described in Section IV-A. As expected the low-rank minimization resulted in a better reconstruction of the hyperspectral image cube with normalized mean square error of .23 versus .35 for the Tikhonov reconstruction. The resulting reconstruction for the 12<sup>th</sup> band is shown in Figure 5.



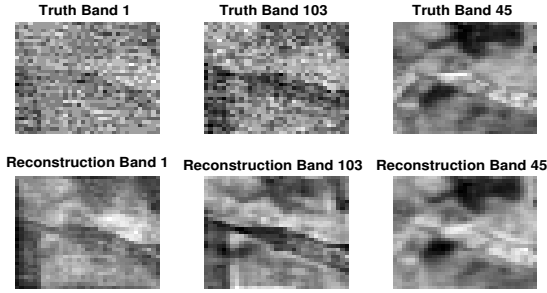
**Fig. 4.** This figure shows the 12 noisy radon projections of the hyperspectral image cube. With 12 projections the system is underdetermined.

**Case III. - Simultaneous tomographic reconstruction and de-noising** - Here we attempted to remove spectral noise from a hyperspectral data cube and reconstruct the cube from a limited number of Radon projections. In order to do so we use a 64x64 section of the original AVIRIS image as used above and observe cube through the same Radon matrix as in the flower example. The simultaneous reconstruction and spectral noise was recovered by solving the optimization problem given in (4) the results are shown in Figure 6. A  $\lambda_S$  of ?? and  $\lambda_L$  of ?? were used. These regularization parameters were chosen using the KS-surface method described in the following section. We show good



**Fig. 5.** This figure shows an example of the true image, low-rank reconstruction, and least square reconstruction, from the hyperspectral flower at band 12.

reconstruction outside of the noisy bands and significant reduction of noise within the spectrally corrupted bands. Although the noise was somewhat reduced in the corrupted bands, as is to be expected in this limited data case, the images still remained noisier than the case when the image hypercube was directly observed with the identity operator.



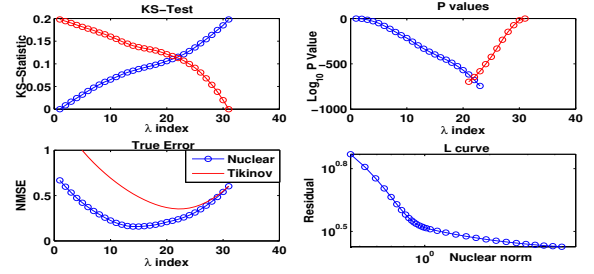
**Fig. 6.** This figure shows the reconstructed and original hypercube at two noisy bands 1 & 103 and at the clean band 45. The reconstruction at the noiseless bands highly resemble the original image. Although somewhat de-noised, the the images at the corrupted bands remain somewhat blurry and the presence of noise is still visible.

#### IV-A. Selection of parameters $\lambda_L$ and $\lambda_S$

In all of the above algorithms the issue of selection of complexity regularization parameters  $\lambda_S$  is of practical importance. In the following experiments that we carried out on real and synthetic data sets the choice of both  $\lambda_S$  and  $\lambda_L$  was determined using a one and two dimension variation of the Komolgrov-Smirnov test method proposed in [28]. The method essentially computes the KS test statistics of errors for a particular value of regularization parameter with respect to error residuals at extreme values of regularization parameter(s) and generates two curves. The operating point is then picked at the intersection of these two curves.

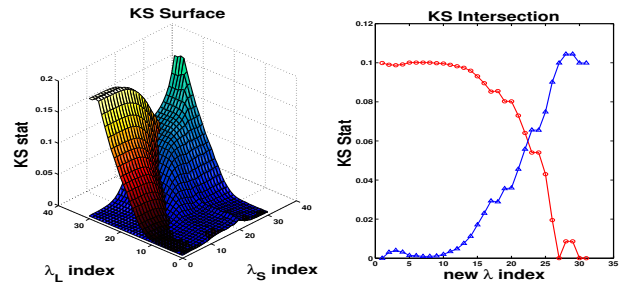
**Selection of regularization parameter for limited angle tomography: no spectral noise** - The KS plot was generated with logarithmic spaced choice of  $\lambda$  from  $10^{-1}$  to  $10^2$ . The KS statistic values and the associated p-values are shown in the top of the Figure 7 and their intersection yield a slight suboptimal  $\lambda$  of 12.6. For reference the KS-test was preformed for Tikhonov regularization and the optimal, see Figure 7-bottom left plot. From this plot we can see that for all feasible values of  $\lambda$  Tikhonov

regularization results in a poorer reconstruction than the nuclear norm reconstruction. In addition to performing the KS-test for selection of regularization parameter the L-curve method [29], commonly used for the selection of the  $\lambda$  was also generated, see Figure 7 bottom right. The L-curve method results in a shallow curve without the presence of sharp knee typical of L-curve plots. The lack of the knee makes it very difficult to select a  $\lambda$  and introduces an opportunity for user bias in the selection of the regularization parameter.



**Fig. 7.** Top plots - KS test plot for recovery under limited Radon projections for the case considered. Bottom plots: (Left) - MMSE computed using the true image for various values of  $\lambda$  for Tikhonov and RPCA methods; (Right) - L-curve for the RPCA method.

**Selection of regularization parameter for limited angle tomography with spectral noise** - For this we extend the one dimensional KS test method in [28] to a two dimension variation by generating a KS test surface. In order to generate the KS-surface many KS-tests were run with a fixed  $\lambda_S$  and the  $\lambda_L$  was varied from  $10^{-3}$  to  $10^{-1}$ . This process was then repeated for a range of  $\lambda_S$  from  $10^{-3}$  to  $10^{-1}$ , effectively generating a KS-plot for the selection of  $\lambda_L$  for a given value of  $\lambda_S$ . We can then view these multiple KS-tests as two surfaces of KS statistics as shown in Figure 8 (left plot), where the intersection of the two surfaces represents best choice of  $\lambda_L$  as a function of  $\lambda_S$ . From this KS-surface we can then interpret the line defining the intersection of the two surfaces line of optimal regularization pairs. We can then take the pair corresponding to the smallest value and largest value of  $\lambda_L$  and generate a conventional one dimensional KS-test along the intersection of  $\lambda$  pairs as shown in Figure 8 right. In this was the one dimensional KS-test selects the best pair set among the intersection of pair sets.



**Fig. 8.** KS Surface for selecting regularization parameters for simultaneous data cube recovery and hyperspectral noise elimination.

## V. REFERENCES

- [1] "AVIRIS-spectra," [aviris.jpl.nasa.gov/html/aviris.spectrum.html](http://aviris.jpl.nasa.gov/html/aviris.spectrum.html).
- [2] S. Kaewpjit, J. Le Moigne, and T. El-Ghazawi, *A wavelet-based PCA reduction for hyperspectral imagery*, Ieee, New York, 2002, WOS:000179116800839.
- [3] C. A. Shah, P. Watanachaturaporn, P. K. Varshney, and M. K. Arora, *Some recent results on hyperspectral image classification*, Ieee, New York, 2004, WOS:000222142800050.
- [4] Damien Letexier and Salah Bourennane, "Multidimensional wiener filtering using fourth order statistics of hyperspectral images," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing, Vols 1-12*, pp. 917–920. Ieee, New York, 2008, WOS:000257456700230.
- [5] Guangyi Chen and Shen-En Qian, "Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage," *Ieee Transactions on Geoscience and Remote Sensing*, vol. 49, no. 3, pp. 973–980, Mar. 2011, WOS:000287658000007.
- [6] Sun Lei and Luo Jian-shu, "Junk band recovery for hyperspectral image based on curvelet transform," *Journal of Central South University of Technology*, vol. 18, no. 3, pp. 816–822, June 2011, WOS:000291185400032.
- [7] P. Scheunders and J. Driesen, "Least-squares interband denoising of color and multispectral images," in *Icip: 2004 International Conference on Image Processing, Vols 1- 5*, pp. 985–988. Ieee, New York, 2004, WOS:000228043501069.
- [8] Nicola Acito, Marco Diani, and Giovanni Corsini, "Signal-dependent noise modeling and model parameter estimation in hyperspectral images," *Ieee Transactions on Geoscience and Remote Sensing*, vol. 49, no. 8, pp. 2957–2971, Aug. 2011.
- [9] Xuefeng Liu, Salah Bourennane, and Caroline Fossati, "Non-white noise reduction in hyperspectral images," *Ieee Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 368–372, May 2012, WOS:000301236900010.
- [10] Adam C. Zelinski and Vivek K. Goyal, *Denoising Hyperspectral Imagery and Recovering Junk Bands using Wavelets and Sparse Approximation*, Ieee, New York, 2006, WOS:000260989400100.
- [11] Zhengming Xing, Mingyuan Zhou, Alexey Castrodad, Guillermo Sapiro, and Lawrence Carin, "Dictionary learning for noisy and incomplete hyperspectral images," *Siam Journal on Imaging Sciences*, vol. 5, no. 1, pp. 33–56, 2012, WOS:000302220800002.
- [12] Emmanuel J. Candes, Xiaodong Li, Yi Ma, and John Wright, "Robust principal component analysis?," *Journal of the Acm*, vol. 58, no. 3, May 2011, WOS:000291246000003.
- [13] Daniel Hsu, Sham M. Kakade, and Tong Zhang, "Robust matrix decomposition with sparse corruptions," *Ieee Transactions on Information Theory*, vol. 57, no. 11, pp. 7221–7234, Nov. 2011, WOS:000297046100001.
- [14] G. Pope, M. Baumann, C. Studer, and G. Durisi, "Real-time principal component pursuit," in *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2011, pp. 1433–1437.
- [15] Gongguo Tang and A. Nehorai, "Robust principal component analysis based on low-rank and block-sparse matrix decomposition," in *2011 45th Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2011, pp. 1–5.
- [16] Hui Ji, Sibin Huang, Zuowei Shen, and Yuhong Xu, "Robust video restoration by joint sparse and low rank matrix approximation," *Siam Journal on Imaging Sciences*, vol. 4, no. 4, pp. 1122–1142, 2011.
- [17] M. R. Descour, C. E. Volin, D. S. Sabatke, E. L. Dereniak, K. J. Thome, A. B. Schumacher, D. W. Wilson, and P. D. Maker, "Demonstration of a high speed non-scanning imaging spectrometer," in *Infrared Detectors and Focal Plane Arrays V*, E. L. Dereniak and R. E. Sampson, Eds., vol. 3379, pp. 250–258. Spie-Int Soc Optical Engineering, Bellingham, 1998, WOS:000075913800024.
- [18] Ashwin A. Wagadarikar, Nikos P. Pitsianis, Xiaobai Sun, and David J. Brady, "Spectral image estimation for coded aperture snapshot spectral imagers," *Proceedings of SPIE*, vol. 7076, no. 1, pp. 707602–707602–15, Aug. 2008.
- [19] J. M. Mooney, V. E. Vickers, M. An, and A. K. Brodzik, "High-throughput hyperspectral infrared camera," *Journal of the Optical Society of America a-Optics Image Science and Vision*, vol. 14, no. 11, pp. 2951–2961, Nov. 1997, WOS:A1997YD31900009.
- [20] Ping Feng, Sze Fong Yau, and Y. Bresler, "A multicoset sampling approach to the missing cone problem in computer-aided tomography," in *Circuits and Systems, 1996. ISCAS '96, Connecting the World., 1996 IEEE International Symposium on*, May 1996, vol. 2, pp. 734–737 vol.2.
- [21] A.E. Waters, A.C. Sankaranarayanan, and R.G. Baraniuk, "SpaRCS: recovering low-rank and sparse matrices from compressive measurements," *Neural Information Processing Systems (NIPS)*, p. 10891097, 2011.
- [22] J.C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 4, pp. 779–785, July 1994.
- [23] D. Landgrebe, "Hyperspectral image data analysis," *Signal Processing Magazine, IEEE*, vol. 19, no. 1, pp. 17–28, Jan. 2002.
- [24] David L. Donoho, "Unconditional bases are optimal bases for data compression and for statistical estimation," *Applied and Computational Harmonic Analysis*, vol. 1, no. 1, pp. 100–115, 1993.
- [25] S. Becker, E. J. Candes, and M. Grant, "Templates for convex cone problems with applications to sparse signal recovery," *Mathematical Programming Computation*, vol. 3, 2011.
- [26] D. H. Foster, S. M. C. Nascimento, and K. Amano, "Information limits on neural identification of colored surfaces in natural scenes," *Visual Neuroscience*, vol. 21, no. 3, pp. 331–336, June 2004.
- [27] William R. Johnson, Daniel W. Wilson, and Greg Bearman, "Spatial-spectral modulating snapshot hyperspectral imager," *Applied Optics*, vol. 45, no. 9, pp. 1898–1908, Mar. 2006.
- [28] Shuchin Aeron, Sandip Bose, Henri-Pierre Valero, and Venkatesh Saligrama, "Broadband dispersion extraction using simultaneous sparse penalization," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4821–4837, 2011.
- [29] Per Christian Hansen and Dianne Prost O'Leary, "The use of the l-curve in the regularization of discrete ill-posed problems," *SIAM J. Sci. Comput.*, vol. 14, no. 6, pp. 1487–1503, Nov. 1993.