

# Compressive imaging spectrometers using coded apertures

D.J. Brady and M.E. Gehm

Duke University Fitzpatrick Center for Photonics, Durham, NC, 27708

## ABSTRACT

A spectral imager provides a 3-D data cube in which the spatial information (2-D) of the image is complemented by spectral information (1-D) about each spatial location. Typically, these systems are operated in a fully-determined (or overdetermined) manner so that the measurements can be computationally inverted into a reliable estimate of the source. We propose a notional system design that is highly underdetermined, yet still computationally invertible. This approach relies on recently-developed concepts in compressive sensing. Because the number of required measurements is greatly reduced from traditional designs, the result is a faster and more economical sensor system.

**Keywords:** Compressive sensing, spectroscopy, multiplex spectroscopy

## 1. INTRODUCTION

There is currently growing interest in a variety of spectral imaging techniques throughout the disparate optical science communities. In this paper, we address *compressive imaging spectroscopy*. Traditional digital imaging techniques produce images with scalar values associated with each spatial pixel location. In imaging spectroscopy, these scalar values are replaced with a vector containing the spectrum spectral information from that spatial location. The resulting *data cube* is therefore three-dimensional (two spatial dimensions and one spectral dimension).

Spectral imaging has traditionally focused on environmental remote sensing and military target recognition tasks.<sup>1,2</sup> In recent years, however, there has been a dramatic growth in biophotonics applications, and with that growth has come an increased interest in spectral imaging for biological applications (especially those with security applications).<sup>3,4</sup>

Straightforward application of traditional spectroscopic techniques to spectral imaging, however, can be problematic. The simplest type of spectral imager combines a tomographic (rotational scanning) or pushbroom (linear scanning) front-end with a traditional slit-based dispersive spectrometer. Unfortunately, the sources tend to be weak and spatially-incoherent. Slit-aperture dispersive spectrometers have extremely poor photon collection efficiency for incoherent sources. When the source is also weak, the absolute number of collected photons can be very small. Further, this small number of photons must be apportioned amongst the large number of “cells” in the data cube. As a result, a given spatio-spectral element tends to contain very few photons and hence has a poor signal-to-noise ratio (SNR).

There have been a number of ingenious solutions to the light collection problem over the years. Two very advanced solutions are the scanning-Michelson Fourier-transform spectrometers, and multiplexed pushbroom designs based on digital micro-mirror (DMM) technology.<sup>5,6</sup> Both approaches have proven highly successful, however they involve expensive components that are not terribly robust.

More robust and inexpensive solutions also exist. The spectral imaging community has developed a number of different *direct-view* designs that maximize the light gathering efficiency of the systems.<sup>7-9</sup> These systems do away with the spectrometer slit altogether and simply view the source through a rotating dispersive element. In this approach, the measurements taken at different rotation angles of the dispersive element are projective measurements through the data cube and can be tomographically reconstructed. While the photon efficiency of this type of approach is quite high, there is a drawback. The geometry of the system necessarily limits the range

---

Further author information: (Send correspondence to D.J. Brady.)  
E-mail: dbrady@duke.edu, Telephone: 1 919 660 5394, www.disp.duke.edu

of angles over which projections are made. As a result of the Fourier-slice theorem, this yields an unsampled region of Fourier space. Consequently, the estimate of the data cube is inexact. In the tomographic community this Fourier undersampling is known as the *missing cone problem*, because the unsampled region is a conical volume in Fourier space. There has been significant work on algorithmic approaches for “filling in” this missing information. The most successful has been the method of projection onto convex sets (POCS).<sup>10</sup>

Additionally, our group has recently developed a coded-aperture based system which is similar to the direct-view methods in that it is inexpensive, mechanically robust, and has a high light-collection efficiency. However, unlike the direct-view methods, it has no missing cone.<sup>11</sup>

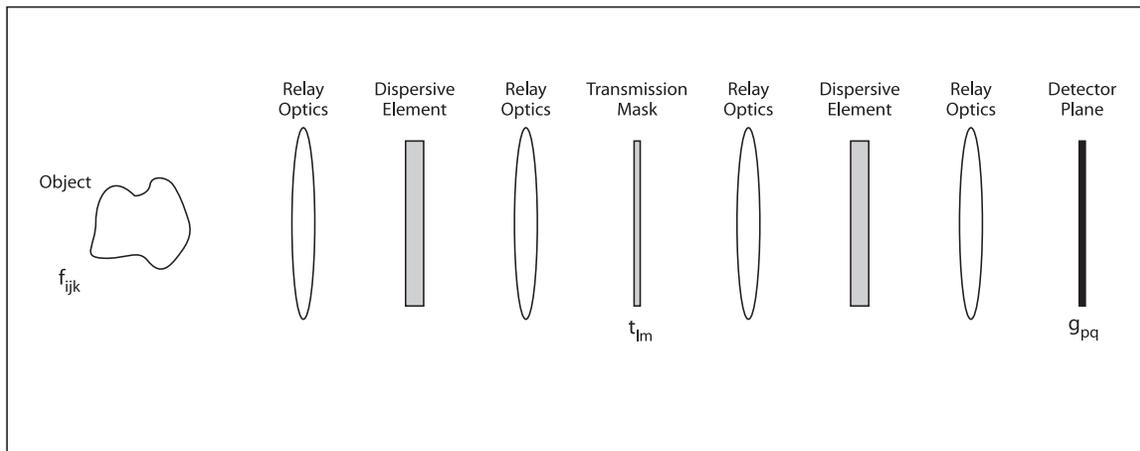
The one characteristic all these methods share, however, is that they generate a number of measurements  $m$  that is equal to, or greater than the number of elements  $n$  in the reconstructed data cube. In this manuscript, we introduce the idea of *compressive spectral imaging*—an approach to spectral imaging that intentionally generates fewer measurements than elements in the reconstruction. This idea is prefigured in the POCS work of Mooney, although here the missing elements are intentional and have a specifically designed structure.

In the past several years, there has been significant work in the area of *compressed sensing*.<sup>12–17</sup> The ability to solve such underdetermined problems relies on the properties of “natural” signals—specifically that they tend to be sparse in some basis other than the naïve Dirac sampling basis.

Previously, we have explored compressive measurement in the context of traditional (non-imaging) spectroscopy.<sup>18</sup> In this manuscript we show how these ideas mesh with our work in coded aperture spectral imagers<sup>11</sup> to yield a spectral imager that is well suited for compressive measurement. The notional design described below allows for the imposition of an arbitrary code along the wavelength dimension of the data cube. Successful reconstruction of the full data cube from the smaller set of measurements then depends on the codesign of the compressive code and the related inference algorithms—work which is still ongoing. The remainder of this manuscript discusses the notional system design, general concepts of coding and inference algorithm design, and presents results from a simple simulation.

## 2. SYSTEM MODEL AND INITIAL RECONSTRUCTION

We begin by considering a simple model of the imaging system. For simplicity, we will assume that all physical scales are matched to the sampling scale of the detector and work solely in a discrete representation. We begin with a source spectral density  $f_{ijk}$ , where the first, second, and third indices indicate the  $x$ ,  $y$ , and  $\lambda$  coordinates, respectively. Light from the source is imaged through a dispersive element onto an intermediate plane containing an aperture mask. This mask modulates the intensity of the light and the resulting distribution is imaged through a second dispersive element onto a two-dimensional detector array which measures the incident intensity profile  $g_{pq}$ . A schematic of this system is shown in Fig. 1.



**Figure 1.** A schematic of the notional imaging system.

To determine the relationship between the detector measurements and the source spectral density, we begin following the propagation of the source through the system. The first dispersive element produces a wavelength-dependent shift of the image on the aperture mask. As a result, the first index of the source (representing the  $x$ -position) becomes a mixture of the spatial and wavelength indices. We can write the spectral density just prior to the aperture mask as  $\mathbf{f}_{(i+k)jk}$ .

We denote the aperture pattern as  $\mathbf{t}_{lm}$ . The coding plane is then imaged through another dispersive element (with a dispersion equal and opposite to the first dispersive element) onto a two-dimensional detector plane. This undoes the index mixing in  $\mathbf{f}$  but introduces a similar index mixing in  $\mathbf{t}$ . The intensity pattern recorded on the detector can be written as

$$\begin{aligned} \mathbf{g}_{pq} &= \sum_{klmpq} \mathbf{f}_{ijk} \mathbf{t}_{(l-k)m} \delta_{li} \delta_{mj} \delta_{pi} \delta_{qj} \\ &= \sum_k \mathbf{f}_{pqk} \mathbf{t}_{(p-k)q} . \end{aligned} \quad (1)$$

The Kronecker-deltas enforce the imaging properties of the optical system, while the sum in  $k$  represents that fact that the detector is wavelength insensitive. This measurement can be viewed as the imposition of the  $m$ th row-code along the wavelength direction of  $\mathbf{f}$ . However, we have the added complication that the code is circularly shifted by an amount that depends on the value of the  $p$ -index.

This single measurement from the detector plane, however, does not provide enough information to reconstruct  $\mathbf{f}$ . In our previous spectral imaging work,<sup>11</sup> we solve this problem by making a series of measurements where the aperture code is shifted relative to the source. Considering a set of shifts, indicated by the index  $\Delta$ , we can write the full set of measurements as

$$\mathbf{g}_{pq\Delta} = \sum_k \mathbf{f}_{pqk} \mathbf{t}_{(p-k)(q-\Delta)} . \quad (2)$$

Previously, we considered only code patterns  $\mathbf{t}$  which were *complete*. Here we intentionally consider *incomplete* code patterns where  $\mathbf{t}$  is row-deficient, such that the computational problem remains significantly underdetermined.

We rewrite Eq. 2 in operator form

$$\mathbf{g}_{pq\Delta} = \sum_k \mathbf{H}_{k\Delta} \mathbf{f}_{pqk} , \quad (3)$$

and solve this linear equation in a least-squares sense for our initial reconstruction

$$\widetilde{\mathbf{f}}_{pqk} = \arg \min \left( \left| \mathbf{g}_{pq\Delta} - \sum_k \mathbf{H}_{k\Delta} \mathbf{f}_{pqk} \right|^2 \right) . \quad (4)$$

### 3. NULL-SPACE SMOOTHING AND FINAL RECONSTRUCTION

The initial reconstruction  $\widetilde{\mathbf{f}}_{pqk}$  is almost certainly going to be poor because of the underdetermined nature of the problem. To improve on the estimate, we must draw on prior knowledge about the source. This prior knowledge, represented in the form of an *inference algorithm*, may take many forms, and may draw upon many types of prior knowledge, both structural and statistical.

Recently, we have been considering the relatively simple idea that “natural” spectra tend to be smooth (or contain a small number of jump discontinuities). However, as we attempt to smooth our spectra to improve the estimates, we are subject to a very important constraint—the estimate, when inserted in Eq. 2, must produce measurements identical to those we observed.

We call this approach *null-space smoothing*, as we can obey the constraint by modifying the estimate only in the space spanned by  $\mathbf{N}_{k\Delta}$ , the null space of the operator  $\mathbf{H}_{k\Delta}$ . Full details on this technique can be found in another publication by our group.<sup>18</sup> Operating in this space, the smoothness criterion is implemented as a

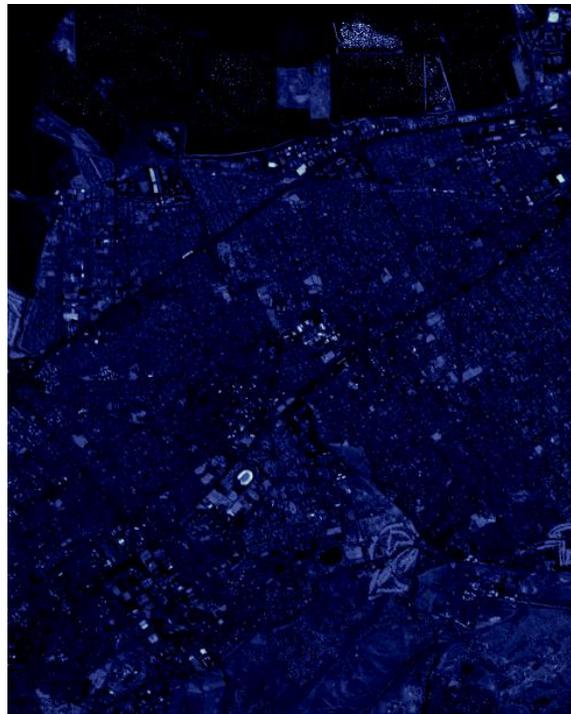
*least gradient* problem, which itself can be written in least squares form. The output of the null-space smoothing algorithm  $\widehat{\mathbf{f}}_{pqk}$  forms our final reconstruction.

Ultimately, null-space smoothing will be only a part of any successful inference approach. However, to explore the performance of this important component, the following section provides results of an initial simulation of these ideas.

#### 4. SIMULATION

We simulated the notional system using freely available spectral imaging data of the Moffett Field region of the San Francisco Bay area. This data was generated by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) instrument.<sup>19</sup> This dataset is divided into four roughly equal spatial regions and contains 224 spectral channels. For computational regions, we limited our source to one of the spatial regions ( $614 \times 495$  pixels and the first 128 spectral channels). An intensity image of the source (sum of all spectral channels) is shown in Fig. 2.

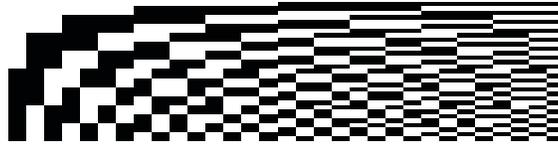
Intensity Image of Source



**Figure 2.** Intensity image of source (sum of all spectral channels).

The coding matrix used was derived from the order-128 Hadamard matrix. Because the system operates on the optical intensity, code patterns are restricted to the range  $[0 \ 1]$ . To comply with this restriction, all  $-1$  elements in the Hadamard matrix were mapped to zero. For the purposes of this manuscript, we kept only the 32 lowest frequency rows of the Hadamard matrix—equivalent to a  $4\times$  downsampling, but with better performance in the presence of noise. The coding matrix used in the simulation is shown in Fig. 3. (As stated above, ultimate performance of the system will depend on the codesign of the coding pattern and the inference algorithms. This work is ongoing, and the final coding pattern is unlikely to be of the form shown here.)

The spectral information at each spatial location was then encoded with a properly shifted version of the coding matrix. The measurements were then inverted into an initial estimate using a least-squares solver. The initial estimate was then used as the starting solution in the null-space solver. The output from this routine was

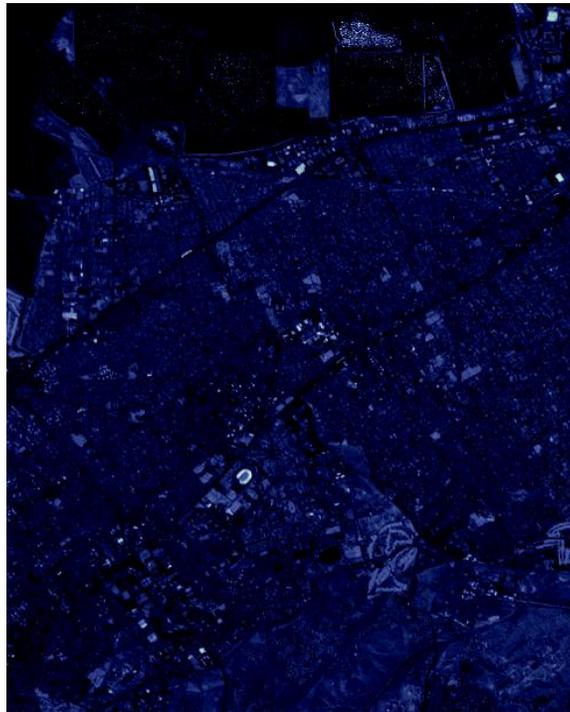


**Figure 3.** Aperture code used in the simulation. White indicates transparent regions of the mask, while black indicates opaque regions.

taken as the final reconstructed spectrum for the spatial location and was inserted in the appropriate location in the reconstruction datacube.

Initial results are quite promising. An intensity image of the reconstructed data (sum over all spectral channels) is shown in Fig. 4. This should be compared to the source intensity image of Fig. 2.

Intensity Image of Reconstruction



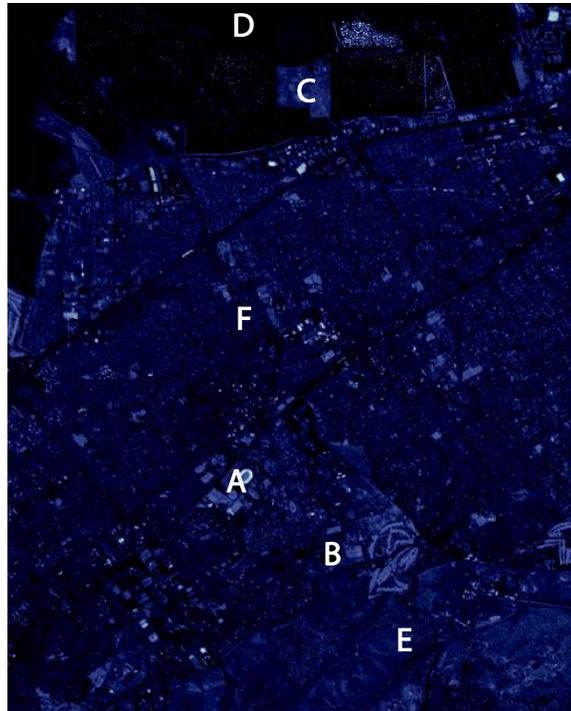
**Figure 4.** Intensity image of reconstruction (sum of all spectral channels).

Of course, the overall performance of a spectral imaging system is hard to judge from an intensity image, since all spectral information is lost. Further comparisons will have to examine spectra at specific spatial locations or images in specific spectral channels. From the intensity images, we can identify a set of seven spatial locations for further examination. These locations are indicated in Fig. 5, and (to our untutored eyes) appear to be

- A. Clearly a manmade structure. From its oval shape, perhaps a stadium of some type.
- B. A region which appears relatively dark in all spectral bands. Composition unknown.
- C. A smooth, grassy region.
- D. A stand of trees or agricultural region.

- E. A grassy hillside.
- F. The center of an urban area.

Intensity Image of Source



**Figure 5.** Spatial locations for subsequent spectral comparisons.

Spectral comparisons between the source and the reconstructed spectra at these six spatial locations are shown in Fig. 6. In all cases, the original spectrum is plotted as a solid line, while the reconstructed values are plotted as a sequence of points.

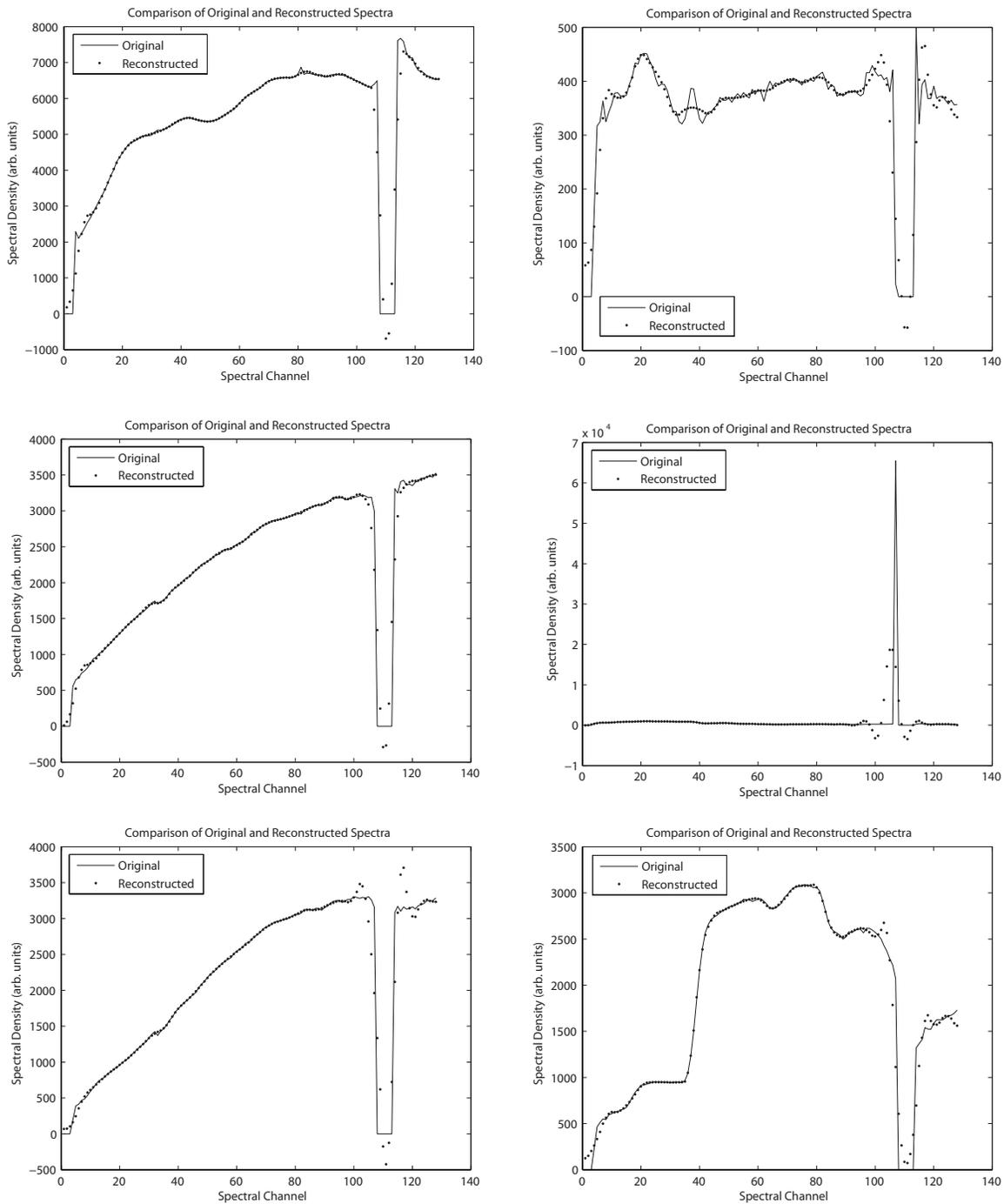
There are a number of interesting features in these performance results. First, we note that the source contains a region of zero value in spectral bands  $\approx 108 - 115$ . This produces sharp discontinuities in the signal which, unsurprisingly, cause problems for our approach which assumes a degree of signal smoothness. In addition to rounding the sharp edges, the reconstruction shows “ringing” in nearby spectral bands. Clearly this is a shortcoming which will be corrected by inference algorithms which go beyond solely null-space smoothing.

Second, we note that the performance on the regions of the spectra which are smooth is almost invariably excellent. The one notable exception is region B, where the signal level is approximately an order of magnitude smaller than the other regions, and contains a number of more narrow features.

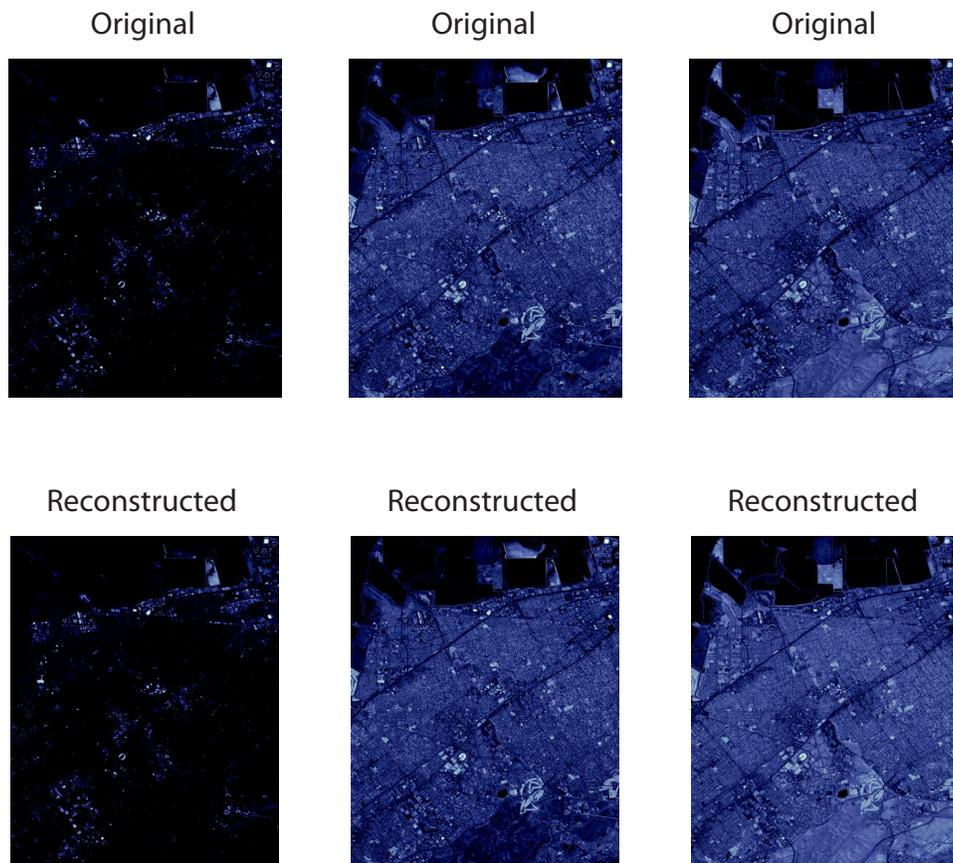
So, in general, the spectral performance of the reconstruction is quite good. To compare the spatial performance, we look at images at particular spectral channels. Based on the spectral plots above, we chose spectral channels 14, 42, and 94. We avoid the regions of spectral ringing as we already know the system performance is substandard there. These comparisons are shown in Fig. 7. To an untrained eye, the reconstructions are indistinguishable from the original source.

## 5. CONCLUSION

We have proposed a straightforward method for combining our previous work on coded-aperture spectral imaging systems with our approach to compressive sensing. The result is an inexpensive, mechanically-robust spectral



**Figure 6.** Spectral comparison at spatial locations A (top left), B (top right), C (middle left), D (middle right), E (bottom left), and F (bottom right).



**Figure 7.** Source and reconstructed spectral images at channels 14 (left) 42 (center) and 94 (right).

imaging system that can impose an arbitrary compressive code on the wavelength dimension of the data cube. As we design more sophisticated coding schemes and inference algorithms, they can be easily be incorporated into the system by changing the only the coding mask.

Regardless of the final form of the inference algorithms, null-space smoothing (or a related concept) is likely to play an important role. We have performed initial simulations of null-space smoothing in the spectral imaging problem, and have shown that it functions well in this domain. Primary difficulties arise near spectral discontinuities, as would be expected. This should be resolved with more sophisticated inference algorithms that incorporate statistical knowledge of the signal as well as structural.

We are continuing our codesign of the aperture codes and the inference algorithms and hope to begin constructing a prototype system in the near future.

## ACKNOWLEDGMENTS

The authors would like to thanks N.P. Pitsianis and X. Sun for stimulating discussions on this topic. This work was supported by the Air Force Office of Scientific Research through grant #F49620-02-1-0335.

## REFERENCES

1. W. Smith, D. Zhou, F. Harrison, H. Revercomb, A. Larar, A. Huang, and B. Huang, "Hyperspectral remote sensing of atmospheric profiles from satellites and aircraft," *Proc. SPIE* **4151**, pp. 94–102, 2001.
2. C. Stellan, F. Olchowski, and J. Michalowicz, "WAR HORSE (wide-area reconnaissance: hyperspectral overhead real-time surveillance experiment)," *Proc. SPIE* **4379**, pp. 339–346, 2001.

3. T. Pham, F. Bevilacqua, T. Spott, J. Dam, B. Tromberg, and S. Andersson-Engles, "Quantifying the absorption and reduced scattering coefficients of tissue-like turbid media over a broad spectral range with noncontact Fourier-transform hyperspectral imaging," *Appl. Opt.* **39**, pp. 6487–6497, 2000.
4. R. Schultz, T. Nielsen, J. Zavaleta, R. Ruch, R. Wyatt, and H. Garner, "Hyperspectral imaging: A novel approach for microscopic analysis," *Cytometry* **43**, pp. 239–247, 2001.
5. C. Snively, G. Katzenberger, and J. Lauterbach, "Fourier-transform infrared imaging using a rapid-scan spectrometer," *Opt. Lett.* **24**, pp. 1841–1843, 1999.
6. A. Wuttig and R. Riesenberger, "Sensitive Hadamard transform imaging spectrometer with a simple MEMS," *Proc. SPIE* **4881**, pp. 167–178, 2003.
7. M. Descour and E. Dereniak, "Computed-tomography imaging spectrometer: experimental calibration and reconstruction results," *Appl. Opt.* **34**, pp. 4817–4826, 1995.
8. P. Bernhardt, "Direct reconstruction methods for hyperspectral imaging with rotational spectrotomography," *J. Opt. Soc. Am. A* **12**, pp. 1884–1901, 1995.
9. J. Mooney, V. Vickers, M. An, and A. Brodzik, "High-throughput hyperspectral infrared camera," *J. Opt. Soc. Am. A* **14**, pp. 2951–2961, 1997.
10. A. Brodzik and J. Mooney, "Convex projections algorithm for resotation of limited-angle chromotomographic images," *J. Opt. Soc. Am. A* **16**, pp. 246–257, 1999.
11. M. Gehm, S. McCain, N. Pitsianis, D. Brady, P. Potuluri, and M. Sullivan, "Static 2D aperture coding for multimodal multiplex spectroscopy," *Appl. Opt.* , May 2006. To appear.
12. M. A. Neifeld and P. Shankar, "Feature-specific imaging," *Applied Optics* **42**(17), pp. 3379–3389, 2003.
13. H. S. Pal and M. A. Neifeld, "Multispectral principal component imaging," *Optics Express* **11**(18), pp. 2118–2125, 2003.
14. E. Candes and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?," *Unpublished manuscript.* , 2004.
15. D. Donoho, "Compressed sensing," *Tech. Rep., Stanford University* , 2004.
16. Y. Tsaig and D. Donoho, "Extensions of compressed sensing," *Tech. Rep., Stanford University* , 2004.
17. N. Pitsianis, D. Brady, and X. Sun, "Sensor-layer image compression based on the quantized cosing transform," *Proc. SPIE* **5817**, pp. 250–257, 2005.
18. D. Brady, M. Gehm, N. Pitsianis, X. Sun, and P. Potuluri, "Compressive sampling strategies for integrated microspectrometers," *Proc. SPIE* , 2006.
19. See <http://aviris.jpl.nasa.gov/>.