Single disperser design for coded aperture snapshot spectral imaging

Ashwin Wagadarikar, Renu John, Rebecca Willett, and David Brady*
Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina 27708, USA
*Corresponding author: dbrady@duke.edu

Received 10 September 2007; accepted 28 November 2007; posted 13 December 2007 (Doc. ID 87303); published 8 February 2008

We present a single disperser spectral imager that exploits recent theoretical work in the area of compressed sensing to achieve snapshot spectral imaging. An experimental prototype is used to capture the spatiotemporal information of a scene that consists of two balls illuminated by different light sources. An iterative algorithm is used to reconstruct the data cube. The average spectral resolution is 3.6 nm per spectral channel. The accuracy of the instrument is demonstrated by comparison of the spectra acquired with the proposed system with the spectra acquired by a nonimaging reference spectrometer. © 2008 Optical Society of America

OCIS codes: 300.6190, 110.0110, 120.6200.

1. Introduction

Spectral imaging is a technique that generates a spatial map of spectral variation, making it a useful tool in many applications including environmental remote sensing [1], military target discrimination [2], astrophysics [3], and biomedical optics [4]. When imaging a scene, a spectral imager produces a two-dimensional spatial array of vectors that represents the spectrum at each pixel location. The resulting two-dimensional data set that contains two spatial dimensions and one spectral dimension is known as a data cube.

Many different techniques for spectral imaging have been developed over the years. Whiskbroom [5], pushbroom [6], and tunable filter [7] imagers are all conceptually simple spectral imager designs. These instruments capture a one- or two-dimensional subset of a data cube and thus require temporal scanning of the remaining dimension(s) to obtain a complete data cube. Furthermore, they have poor light collection efficiency for incoherent sources, resulting in a poor signal-to-noise ratio (SNR).

The light collection problem is addressed by multiplex spectral imager designs based on Fourier and Hadamard transforms. Scanning Michelson Fourier transform spectral imagers [8] use a scanning interferometer, thus requiring significant mechanical stability [9]. Multiplex pushbroom designs with Hadamard encoding use expensive spatial light modulators [10]. Although they increase the light throughput, these designs still use some form of scanning, making it difficult to use them for spectral imaging of nonstatic scenes.

Tomographic approaches have also produced major advances. Mooney et al. [11] developed a direct-view design that maximizes the light-gathering efficiency by not requiring any spatial filter, such as a slit. With this design, the source is viewed through a rotating dispersive element. Measurements are taken at different rotation angles. These measurements are projective measurements through the data cube that can be tomographically reconstructed. Although the light-gathering efficiency of such an instrument is high, the geometry of the system limits the range of angles over which projections are made. As a result of the Fourier slice theorem, this results in an unsampled region in Fourier space, a problem known as the missing cone problem because the unsampled region is a conical volume in the Fourier domain representation of the data cube [13]. The computed tomography imaging spectrometer (CTIS) [13] system is a static, snapshot instrument that captures multiple projections of the data cube at once. These capabilities make the CTIS ideal for spectral imaging.
of transient scenes. However, the instrument requires a large focal plane area and also suffers from the missing cone problem [14].

An important characteristic shared by all the designs described above is that the total number of measurements they generate is greater than or equal to the total number of elements in the reconstructed data cube. In contrast, our group has introduced the idea of compressive spectral imaging, an approach to spectral imaging that intentionally generates fewer measurements than elements in the reconstructed data cube [12]. Compressed sensing solves the underdetermined data cube estimation problem by relying on a crucial property of natural scenes, namely, that they tend to be sparse on some multiscale basis. To achieve compressive spectral imaging, our group has developed a new class of imagers dubbed the coded aperture snapshot spectral imager (CASSI). The CASSI utilizes a coded aperture and one or more dispersive elements to modulate the optical field from a scene. A detector captures a two-dimensional, multiplexed projection of the three-dimensional data cube representing the scene. The nature of the multiplexing performed depends on the relative position of the coded aperture and the dispersive element(s) within the instruments.

We recently reported on a dual disperser (DD) CASSI design [15], which consists of two sequentially dispersive arms arranged in opposition so that the dispersion in the second arm cancels the dispersion introduced by the first arm. A coded aperture is placed between the two arms. Recovery of the data cube from the detector measurement is performed by use of an expectation-maximization method designed for hyperspectral images. Such a design applies spatially varying, spectral filter functions with narrow features. Through these filters, the detector measures a projective measurement of the data cube in the spectral domain. In essence, the DD CASSI sacrifices spatial information to gain spectral information about the data cube. Spectral information from each spatial location in the scene is multiplexed over a localized region of the detector. A useful property of the design is that the measurement resembles the scene, making it easy to focus the camera on objects in the scene. This also makes it possible to perform local block processing of the detector data to generate smaller data cubes of subsets of the entire scene.

We report a new compressive CASSI instrument dubbed the single disperser (SD) CASSI. Like the DD CASSI, the SD CASSI does not directly measure each voxel in the desired three-dimensional data cube. It collects a small number (relative to the size of the data cube) of coded measurements and a sparse reconstruction method is used to estimate the data cube from the noisy projections. The instrument disperses spectral information from each spatial location in the scene over a large area across the detector. Thus, spatial and spectral information from the scene is multiplexed on the detector, implying that the null space of the sensing operation of the SD CASSI differs from that of the DD CASSI. Also, a raw measurement of a scene on the detector rarely reveals the spatial structure of the scene and makes block processing more challenging.

Since the DD CASSI multiplexes only spectral information in the data cube, it cannot reconstruct the spectrum of a point source object. On the other hand, the SD CASSI can reconstruct the spectrum of a point source, provided that the source spatially maps to an open element on the coded input aperture. This implies that, for reconstructions that demand high spatial resolution with less stringent demands on spectral resolution, the DD CASSI instrument should be the compressive spectral imager of choice. On the other hand, when spectral resolution is more critical than spatial resolution in the data cube, the SD CASSI instrument should be chosen. The SD CASSI has the additional benefit of requiring fewer optical elements, making optical alignment much easier. The SD CASSI also removes the CTIS constraints of measuring multiple projections of the data cube and using a large focal plane array. Essentially, just one spectrally dispersed projection of the data cube that is spatially modulated by the aperture code over all wavelengths is sufficient to reconstruct the entire spectral data cube.

In Section 2 we detail our system model that mathematically describes the sensing operation implemented by the SD CASSI. In Section 3 we describe the reconstruction technique used to generate a three-dimensional data cube from an underdetermined set of measurements collected by the detector of a SD CASSI. In Section 4 we describe the details of an experimental SD CASSI prototype constructed to experimentally verify the snapshot, compressive imaging concept. Finally, in Section 5 we present the experimental results demonstrating the reconstruction of a data cube of a scene.

2. System Model

A schematic of the compressive, snapshot SD CASSI is shown in Fig. 1. A standard imaging lens is used to form an image of a remote scene in the plane of the coded aperture. The coded aperture modulates the spatial information over all wavelengths in the spectral cube with the coded pattern. Imaging the cube
from this plane through the dispersive element results in multiple images of the code-modulated scene at wavelength-dependent locations in the plane of the detector array. The spatial intensity pattern in this plane contains a coded mixture of spatial and spectral information about the scene. We note that this design essentially extends the architecture of our computational static multimodal, multiplex spectrometers [10] to computational spectral imaging, with the addition of an imaging lens placed in front of the input aperture code.

Below we mathematically model the SD CASSI sensing process. The model assists us in developing the system operator matrix for the reconstruction algorithm. The model does not account for optical distortions introduced by the optical elements in the instrument such as blurring and smile distortion [16]. However, it does help us to understand the basic operations implemented by the optical elements.

The spectral density that enters the instrument and is relayed to the plane of the coded aperture can be represented as $f_0(x, y; \lambda)$. If we represent the transmission function printed on the coded aperture as $T(x, y)$, the spectral density just after the coded aperture is

$$f_l(x, y; \lambda) = f_0(x, y; \lambda)T(x, y).$$

After propagation through the coded aperture, the relay optics and the dispersive element, the spectral density at the detector plane is

$$f_d(x, y; \lambda) = \int \int \delta(x - [x + \alpha(\lambda - \lambda_c)]) \delta(y - y') f_l(x', y'; \lambda) dx' dy' = \int \int \delta(x' - [x + \alpha(\lambda - \lambda_c)]) \delta(y' - y) f_0(x', y'; \lambda)T(x', y')dx' dy' = f_0(x + \alpha(\lambda - \lambda_c), y; \lambda)T(x + \alpha(\lambda - \lambda_c), y),$$

where the Dirac delta functions describe propagation through unity magnification imaging optics and a dispersive element with linear dispersion $\alpha$ and center wavelength $\lambda_c$. The detector array is insensitive to wavelength and measures the intensity of incident light rather than the spectral density. Thus, the continuous image on the detector array can be represented as

$$g(x, y) = \int f_0(x + \alpha(\lambda - \lambda_c), y; \lambda)T(x + \alpha(\lambda - \lambda_c), y)d\lambda.$$  

This image is a sum over the wavelength dimension of a mask-modulated and later sheared data cube. Note that this is in contrast with the spatially regis-

tered image formed on a DD CASSI detector as a result of summing over the wavelength dimension of a data cube that is first sheared, mask modulated, and finally unsheared.

Since the detector array is spatially pixelated with pixel size $\Delta$, $g(x, y)$ is sampled across both dimensions of the detector. In the presence of noise $w$, the measurements at position $(n, m)$ on the detector can be represented as

$$g_{nm} = \int \int g(x, y)\text{rect}\left(\frac{x}{\Delta - m}, \frac{y}{\Delta - n}\right) dx dy + w_{nm} = \int \int f_0(x + \alpha(\lambda - \lambda_c), y; \lambda)T(x + \alpha(\lambda - \lambda_c), y) \times \text{rect}\left(\frac{x}{\Delta - m}, \frac{y}{\Delta - n}\right) dx dy d\lambda + w_{nm}.$$  

Reconstruction of data cubes from the physical system benefits if the coded aperture feature size is an integer multiple $q$ of the size of detector pixels $\Delta$. This avoids the need for subpixel positioning accuracy of the coded aperture. The aperture pattern $T(x, y)$ can be represented as a spatial array of square pinholes, with each pinhole having a side length of $q\Delta$ and $t_{n'm'}$ representing an open or closed pinhole at position $(n', m')$ in the pinhole array:

$$T(x, y) = \sum_{m'n'} t_{m'n'}\text{rect}\left(\frac{x}{q\Delta - m'}, \frac{y}{q\Delta - n'}\right).$$  

With this representation for the aperture pattern, the detector measurements as represented in Eq. (4) become

$$g_{nm} = \sum_{m'n'} t_{n'm'}\int \int \text{rect}\left(\frac{x + \alpha(\lambda - \lambda_c)}{q\Delta} - m', \frac{y}{q\Delta} - n'\right) \times \text{rect}\left(\frac{x}{\Delta - m}, \frac{y}{\Delta - n}\right) \times f_0(x + \alpha(\lambda - \lambda_c), y; \lambda)dx dy d\lambda + w_{nm}.$$  

Denoting the source spectral density $f_0(x, y; \lambda)$ in discrete form as $f_{ijk}$ and the aperture code pattern $T(x, y)$ as $t_{ij}$, the detector measurements in matrix form can be written as

$$g_{nm} = \sum_k f_{(m+k)n}k t_{(m+k)n} + w_{nm} = (Hf)_{nm} + w_{nm},$$  

where $H$ is a linear operator that represents the system forward model.

3. Data Cube Reconstruction Method
The SD CASSI measures a two-dimensional, spatiotemporal multiplexed projection of the three-
dimensional data cube representing the scene. Reconstructing a data cube from this compressed measurement relies on the assumption that the sources in the scene have piecewise smooth spatial structure, making the data cube highly compressible in the wavelet basis. Here we describe the reconstruction method used to estimate the data cube from the detector measurements. The data cube can be represented as

\[ f = W\theta, \]  

where \( \theta \) is a vector composed of the two-dimensional wavelet transform coefficients for each spectral band concatenated to form one vector, and \( W \) denotes the inverse two-dimensional wavelet transform applied to each spectral band to form data cube \( f \). The SD CASSI detector measurement can be represented as

\[ g_{nm} = (HW\theta)_{nm} + w_{nm}, \]  

where \( H \) is a representation of the system forward model described in Section 2 and \( w_{nm} \) is noise.

If data cube \( f \) consists of \( (p \times p) \) spatial channels with \( q \) spectral channels, it can be represented as a cube of size \( (p \times p \times q) \). The corresponding detector measurements \( g \) can be represented as a matrix of size \( (p \times (p + q - 1)) \). The number of columns in this matrix reflects the fact that the detector measurement is a sum of coded images of the scene at each spectral channel, with each spectral image displaced by a column of pixels from the adjacent image. If we represent \( f \) and \( g \) as column vectors, the linear operator matrix \( H \) can be represented as a matrix of size \( [p(p + q - 1)] \times (p^2q) \). The vector of wavelet coefficients of data cube \( \theta \) is of size \( [p^2q(3 \log_2(p) + 1)] \times 1 \). The size of this vector reflects the fact that the wavelet decomposition of the data cube is performed as a two-dimensional undecimated transform on each of the \( q \) spectral channels. The undecimated transform is used to ensure that the resulting method is translation invariant.

An estimate \( \hat{f} \) for the data cube can be found by solving the problem

\[ \hat{f} = W\left[\arg\min_{\theta'}\{\|g - HW\theta'\|_2^2 + \tau\|\theta'\|_1\}\right]. \]  

The solution of this nonlinear optimization problem has received significant attention recently [17–19]. We use the gradient projection for sparse reconstruction (GPSR) method developed by Figueiredo et al. [20]. This approach is based on a variant of the Barzilai–Borwein gradient projection method [21] and has code available online [22] . The reconstruction method searches for a data cube estimate with a sparse representation in the wavelet basis; i.e., a \( \theta \) that contains mostly zeros and a relatively small number of large coefficients. The first term in this optimization equation minimizes the \( \ell_2 \) error between the measurements modeled from the estimate and the true measurement. The second term is a penalty term that encourages sparsity of the reconstruction in the wavelet domain and controls the extent to which piecewise smooth estimates are favored. In this formulation, \( \tau \) is a tuning parameter for the penalty term and higher values of \( \tau \) yield sparser estimates of \( \theta \).

4. Experimental Setup

To verify the SD CASSI spectral imaging concept experimentally, we constructed a proof-of-concept prototype as shown in Fig. 2. The prototype consisted of (i) a coded aperture, lithographically patterned on a chrome-on-quartz mask, (ii) three lenses from Schneider Optics (Hauppauge, New York) with an f/# of 1.4 and a focal length of 22.5 mm, (iii) an equilateral prism from Edmund Optics (Barrington, New Jersey) as a dispersive element, (iv) a monochrome charge-coupled device (CCD) detector from Photometrics (Tucson, Arizona) with 1040 × 1392 pixels that are each 4.65 \( \mu \)m square, and (v) a 500–620 nm bandpass filter that was placed in front of the imaging lens to remove the impact of stray light on the experimental measurements. MATLAB routines were written to control and capture data on the CCD.

The aperture code used in all the experiments was based on an order 192 S-matrix code [23], with features that were four CCD pixels wide and four CCD pixels tall, and two completely closed rows of CCD pixels added between the code rows. The columns of the original S-matrix code were shuffled in a random but repeatable way. The code was originally designed for a coded-aperture spectrometer and was not optimized for the spectral imaging application demonstrated in this paper. Although we have developed the theory for and conducted experimental studies of optimal aperture codes for coded aperture spectroscopy [9,24], it is not applicable to the spectral imaging application demonstrated here. Work on optimal aperture code(s) for the CASSI instruments will be reported in a future manuscript.

![Fig. 2. (Color online) Experimental prototype of the SD CASSI.](Image)
An equilateral prism was used instead of a grating because the grating produces overlapping diffractive orders whereas the prism only refracts the wavelengths into one order. Prisms also have large transmission efficiencies [25]. Given the system geometry and the low dispersion of the equilateral prism, the number of CCD columns illuminated when white light was allowed to pass through the system was less than half of the width of the CCD array. Thus, the spectral range of the instrument was essentially limited by the quantum efficiency of the CCD image sensor.

5. Experimental Results and Discussion

We note from the outset that given the proof-of-concept nature of the prototype system, it was simply put together with off-the-shelf parts and not optimized in optical ray tracing software. Consequently, the results presented in this section are limited to a display of the ability of the SD CASSI to reconstruct relatively simple spatiospectral scenes. We are in the process of constructing a second generation instrument that is sturdier and has much higher measurement quality.

To generate an estimate of the data cube representing the scene, the reconstruction algorithm requires two inputs: the aperture code used by the instrument to encode the data cube and the detector measurement of the data cube. Instead of using the code pattern printed on the aperture mask as the code pattern used for the reconstruction, the instrument was uniformly illuminated with a single wavelength source in the form of a 543 nm laser. Figure 3 shows the detector measurement of the aperture code pattern after propagation through the optics. This is a more accurate representation of the spatial coding scheme applied to the data cube than the code printed on the aperture mask and thus produces a more accurate reconstruction of the data cube.

We constructed a scene consisting of two Ping-Pong balls as shown in Fig. 4. One Ping-Pong ball was illuminated by a 543 nm laser and a white light source filtered by a 560 nm narrowband filter. The other was a red Ping-Pong ball illuminated by a white light source. Figure 5 shows a CCD measurement of the scene obtained with the SD CASSI. Given the low linear dispersion of the prism, there is spatiospectral overlap of the aperture code-modulated images of each ball. The 500–620 nm bandpass filter placed in front of the imaging lens ensures that only the subset of the data cube that corresponds to this band of wavelengths is measured by the instrument.

Detector measurements of both the code pattern and the scene were digitally downsampled by a factor of 2 in the row and column dimensions prior to performing a reconstruction of the data cube. This downsampling helped reduce the time needed by the reconstruction algorithm to generate an estimate of the data cube. The resulting \(128 \times 128 \times 28\) data cube spanned a spectral range from 540 to 640 nm. One hundred iterations of the reconstruction algorithm required approximately 14 min of run time on a desktop machine. The GPSR method was run with \(\tau = 0.05\). This value was determined by trial and error. Future work is needed to study how to choose the optimal value of \(\tau\).

Figure 6 shows the spatial content of each of 28 wavelength channels between 540 and 640 nm. Note
that the mask modulation on the spatial structure visible in Fig. 5 has been removed in all the wavelength channels and that the two balls are spatially separated. To validate the ability of the SD CASSI to reconstruct the spectral signature of objects in the scene, we measured the spectral signatures of each Ping-Pong ball using an Ocean Optics (Dunedin, Florida) spectrometer. Figure 7(a) shows the SD CASSI spectrum at a point on the green Ping-Pong ball, and Fig. 7(b) shows the SD CASSI spectrum at a point on the red Ping-Pong ball. The wavelength axis in the plots had to be calibrated because of the nonlinear dispersion of the prism across the detector. This calibration was performed by tracking the position of a point source while varying its wavelength. Figures 7(a) and 7(b) also show the Ocean Optics spectrum from each ball for comparison. The two reconstructed SD CASSI spectra closely match those generated by the Ocean Optics spectrometer.

We note that the reconstruction algorithm used to generate the data cube assumed that the system response when the aperture code was fully illuminated with any wavelength was identical to the system response at 543 nm. Thus, it did not account for an anamorphic horizontal stretch of the image that is wavelength dependent. The quality of the reconstructed data cube could be improved if the reconstruction algorithm were to utilize a system response that captures a fully illuminated aperture code pattern at all the wavelengths in the spectral range of the instrument.

An important characteristic of any spectrometer or spectral imager is its spectral resolution. If we ignore the optical distortions such as the blurring and the smile distortion, the spectral resolution of the SD CASSI would then be determined by the width of the smallest code feature. Consider the case in which the smallest code feature maps to two detector pixels. Then imaging two adjacent, monochromatic point

---

**Fig. 6.** (Color online) Spatial content of the scene in each of 28 spectral channels between 540 and 640 nm. The green ball can be seen in channels 3, 4, 5, 6, 7, and 8; the red ball can be seen in channels 23, 24, and 25.

**Fig. 7.** (Color online) (a) Spectral intensity through a point on the Ping-Pong ball illuminated by a 543 nm green laser and a white light source filtered by a 560 nm narrowband filter. (b) Spectral intensity through a point on the red Ping-Pong ball illuminated by a white light source. Spectra from an Ocean Optics nonimaging reference spectrometer are shown for comparison.
sources of close but distinct wavelengths to the smallest code feature can potentially result in spatiotemporal mapping of both point sources to the same pixel on the detector. Thus, the spectral resolution for the SD CASSI, i.e., the separation between the spectral channels, is determined by the amount of dispersion (in nanometers) across two detector pixels. Since the dispersion of a prism is nonlinear, the spectral resolution of this instrument varies with wavelength.

Accordingly, an average spectral resolution can be computed for the SD CASSI experimental prototype. This was determined by removing the bandpass filter in front of the imaging lens and illuminating the instrument with a 543 and 632 nm laser. The separation between the aperture images resulting from these wavelengths was approximately 100 pixels, corresponding to an average dispersion of 0.9 nm/pixel. Since the width of the smallest code feature was four detector pixels, the average spectral resolution of the instrument was 3.6 nm/spectral channel.

6. Summary
We have developed a new class of spectral imagers known as the coded aperture snapshot spectral imager (CASSI). The dual disperser CASSI measures a two-dimensional coded spectral projection of the three-dimensional data cube. We have reported on the single disperser CASSI, which measures a two-dimensional coded spatiotemporal projection of the three-dimensional data cube that represents the scene. The differences between these instruments are listed in table format in Fig. 8.

The GPSR method was used to estimate the data cube from the SD CASSI measurements by assuming spatial sparsity of the scene in the wavelet basis. A first generation prototype was used to provide experimental verification of the compressive spectral imaging concept. The spatial resolution was enough to allow the identification of two distinct objects in the constructed scene. The spectral signatures of the scene as measured by the experimental prototype compared well with a reference nonimaging spectrometer and the average spectral resolution was 3.6 nm/spectral channel. Future work will focus on the development of a second generation SD CASSI with better image quality, improving the quality of reconstructed data cubes by utilizing the complete wavelength-dependent system response in the reconstruction process, and determining the optimal aperture code for the SD CASSI instrument. Finally, we note that the DD CASSI and SD CASSI represent two extreme forms of compressive spectral imagers. Modifying the second dispersive arm of the DD CASSI so that its dispersion does not completely cancel the dispersion caused by the first arm will result in a new CASSI instrument that will measure something between a spatially well registered image and a spatiotemporal projection of the data cube.

This research was supported by the Defense Advanced Research Projects Agency’s Microsystems Technology Office through a collaborative project with Rice University (ONR grant N 00014-06-1-0610).

References