REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*
15-04-2009

2. REPORT TYPE
Final Technical Report

3. DATES COVERED *(From-to)*
01-May-06 to 30 APR-08

4. TITLE AND SUBTITLE
Compressive Optical Imaging Systems– Theory, Devices and Implementation

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-06-1-0610

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Baraniuk, Richard G.; Kelly, Kevin F.; Brady, David; Willett, Rebecca

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Rice University
6100 Main St. MS 16
Houston TX 77005

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Dennis Healy
DARPA, Microsystems Technology Office
3701 N. Fairfax Dr.
Arlington VA 22203-1714

10. SPONSOR/MONITOR'S ACRONYM(S)
DARPA MTO/SPAWAR

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

In this project we expanded the theory of compressive sensing and explored the connection of CS to advanced optical imaging systems. In regard to theory we incorporated a Hidden Markov Tree signal model into existing CS theory and produced an algorithm for fast signal recover.

We physically implemented CS on a single-pixel imaging testbed and analyzed its performance against existing multiplexing schemes. The insight of CS theory and experience in CS implementation culminated with the creation of our single-shot hyperspectral imager.

15. SUBJECT TERMS
analog-to-information conversion, compressive sampling, sparsity, random filter, random sampling

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE	SAR	34	Richard G. Baraniuk
U	U	U			19b. TELEPONE NUMBER (*Include area code*) 713-348-5132

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI-Std Z39-18

Compressive Optical Imaging Systems – Theory, Devices, and Implementation

Final Technical Report

Richard G. Baraniuk, David Brady, Kevin Kelly, and Rebecca Willett

1 Introduction

In this report we begin by restating the motivation for our work, and review the project objectives. We present our results and follow each research thrust with potential future areas of work. We conclude with a list of publications supported by the grant, and a list of project personnel.

1.1 Review of motivation

Humans are visual creatures, and image sensors that extend our reach cameras have improved dramatically in recent times thanks to the introduction of CCD and CMOS digital imaging technology. Current digital cameras acquire a pixelized image or video sequence by sampling the incident light field using a photon-sensitive array of N pixels. For high-resolution images, N must be large, often $N \gg 10^6$. Reducing the camera storage and communication burden to a manageable level requires compression, such as JPEG [1] or JPEG2000 [2,3]for images and MPEG [4] for video. In a conventional image/video camera system, the pixel array consumes a large chip area, and the compression algorithm, while throwing away well over 90% of the imagers output, consumes a large amount of computational energy. These size, data-inefficiency, and energy issues limit the ubiquity and usefulness of digital cameras, particularly in battery operated scenarios.

Sampling theory is the fundamental basis of sensor systems generally and imaging systems specifically. Conventional conceptions of sampling theory do not distinguish between "measurements" and "samples." In imaging systems, a measurement is a linear projection or inner product $y = \langle x, \phi \rangle$ of the spatial object distribution x.

While the goal of conventional optical design is generally to make ϕ a delta function in the naive sampling space, more complex sampling strategies have a long history. The most obvious example is tomography, which relies on Radon projections of the object distribution. However, more complex coding strategies have long been applied in imaging [5] and spectroscopy [6, 7]. Advanced coding strategies have become increasingly attractive over the past decade as the sophistication of both information processing and electronic sensors has increased. Both dynamic [8, 9] and static [10, 11] coding strategies have recently been demonstrated as mechanisms for improving a variety of sensor system metrics, including system geometry, data load, and sensitivity. The goal of an imaging system is to acquire x. The fundamental advantage of digital systems is that the measured data y and the estimated image $\widehat{x}$ are not the same thing, whereas in direct view and film systems the measurement and the image are identical. In systems constructed to date, the distinction between the electronic signal distribution on a focal plane and the displayed image has mainly been exploited simply as a means of remotely transmitting the image. While a few examples of deliberate distortions of y to improve metrics in $\widehat{x}$ have been demonstrated, many interesting strategies remain unexplored. Strategies applying modern sampling theories to both y and x as discrete objects are particularly under-developed.

1.2 Review of project objectives

This project aimed at exploring the foundations and applications of CS in optical imaging problems. Specifically, we investigated:

- **Advancement of CS theory** In a CS sensor, compression is part of the analog acquisition process. It was our goal to study and design new CS theory and algorithms in the context of optical imaging systems. We developed a theory that incorporates Hidden Markov Tree signal models with CS acquisition and recovery, and produced an algorithm for fast recovery of such signals.

20090618002

- **Implementation of CS practice** We set out to demonstrate Compressive Sensing (CS) in a real-world imaging system. We applied compressive measurement and reconstruction schemes in development of a single-pixel camera.
- **Application of CS principles to optics** Our final objective was to develop a test-bed that uses spectral channels for compact and programmable spatial compressive coding. We used this test-bed to perform single-shot hyper-spectral imaging.

2 Advancement of CS theory

2.1 Summary of results

We developed theory and proposed a new algorithm that enables fast recovery of *piecewise smooth signals*, a large and useful class of signals whose sparse wavelet expansions feature a distinct "connected tree" structure. Our algorithm fused recent results on iterative reweighted ℓ_1-norm minimization with the wavelet Hidden Markov Tree model. The resulting optimization-based solver outperformed the standard compressive recovery algorithms as well as previously proposed wavelet-based recovery algorithms. As a bonus, the algorithm reduced the number of measurements necessary to achieve low-distortion reconstruction.

2.2 Compressive sensing background

Let $x \in \mathbb{R}^N$ be a signal and let the matrix $\Psi := [\psi_1, \psi_2, \ldots, \psi_N]$ be a basis for $\mathbb{R}^N$. We say that x is *K-sparse* if it can be expressed as a linear combination of K vectors from Ψ; that is, $x = \sum_{i=1}^{K} \theta_{n_i} \psi_{n_i}$ with $K \ll N$.

2.2.1 Incoherent measurements

Consider a signal x that is K-sparse in Ψ. Consider also an $M \times N$ *measurement matrix* Φ, $M \ll N$, where the rows of Φ are incoherent with the columns of Ψ. For example, let Φ contain i.i.d. Gaussian entries; with high probability, such a matrix is incoherent with any fixed Ψ (universality). Compute the *measurements* $y = \Phi x$ and note that $y \in \mathbb{R}^M$ with $M \ll N$. The CS theory states that there exists an overmeasuring factor $c > 1$ such that only $M := cK$ incoherent measurements y are required to reconstruct x with high probability [12, 13]. That is, just cK incoherent measurements encode all of the salient information about the K-sparse signal x.

2.2.2 Reconstruction from incoherent projections

The overmeasuring factor c required depends on the (nonlinear) reconstruction algorithm. Under the sparsity assumption, one can search for the sparsest signal that agrees with the obtained measurements; by using the ℓ_0-norm $\|\theta\|_0 = \#\{n : \theta_n \neq 0\}$, the reconstruction algorithm can be expressed as

$$\widehat{\theta} = \arg\min_{\theta} \|\theta\|_0 \quad \text{subject to} \quad \Phi\Psi\theta = y.$$

While this algorithm demands the smallest possible overmeasuring factor ($c = 2$), its computational complexity renders it unfeasible. Most of the existing literature on CS [12–14] has concentrated on optimization-based methods for signal recovery, in particular ℓ_1-norm minimization. The ℓ_1-norm approach seeks a set of sparse coefficients $\widehat{\theta}$ by solving the linear program

$$\widehat{\theta} = \arg\min_{\theta} \|\theta\|_1 \quad \text{subject to} \quad \Phi\Psi\theta = y; \tag{1}$$

the reconstruction of sparse signals via ℓ_1-norm minimization is typically exact, provided that $c = O(\log(N/K))$. Other algorithms, including greedy algorithms [15], have also been proposed for CS reconstruction and require similar oversampling factors. However, all of these algorithms are generic in the sense that, aside from sparsity, they assume no particular structure within the signal coefficients.

2.2.3 Iterative reweighted ℓ_1-norm minimization

When the complexity of the signal is measured using the ℓ_1-norm, individual signal coefficients are penalized according to their magnitude; in contrast, when the ℓ_0-norm is used to measure the signal complexity, the penalty for a nonzero coefficient is independent of its magnitude. The effect of this disparity is reflected in the increase of the oversampling factor between the two algorithms.

A small variation to the ℓ_1-norm penalty function has been suggested to rectify the imbalance between the ℓ_0-norm and ℓ_1-norm penalty functions [16]. The basic goal is to minimize a weighted ℓ_1-norm penalty function $\|W\theta\|_1$, where W is a diagonal "weighting" matrix with entries $W_{n,n}$ approximately proportional to $1/|\theta_n|$. This creates a penalty function that achieves higher magnitude-independence. Since the true values of θ_n are unknown (indeed they are sought), however, an iterative reweighted ℓ_1-norm minimization (IRWL1) algorithm is suggested.

The algorithm starts with the solution to the unweighted ℓ_1-norm minimization algorithm (1), which we name $\widehat{\theta}^{(0)}$. The algorithm then proceeds iteratively: on iteration $i > 0$, it solves the optimization problem

$$\widehat{\theta}^{(i)} = \arg\min_{\theta} \|W^{(i)}\theta\|_1 \quad \text{subject to} \quad \Phi\Psi\theta = y, \tag{2}$$

where $W^{(i)}$ is a diagonal reweighting matrix with entries

$$W^{(i)}_{n,n} = \left(\left|\widehat{\theta}^{(i-1)}_n\right| + \epsilon\right)^{-1},$$

and ϵ is a small regularization constant. The algorithm can be terminated when the change between consecutive solutions is smaller than an established threshold or after a fixed number of iterations. Each iteration of this algorithm can he posed as a linear program, for which there exist efficient solvers.

2.3 CS for wavelet-sparse signals

A widely used sparse representation in signal and image processing is the wavelet transform. Since piecewise polynomial signals have sparse wavelet expansions [17] and since many real-world signals describe punctuated, piecewise smooth phenomena, it follows that many real-world signals have sparse or compressible wavelet expansions. However, the significant wavelet coefficients in general do not occur in arbitrary positions. Instead they exhibit a characteristic signal-dependent structure.

Without loss of generality, we focus on 1D signals, although similar arguments apply for 2D and higher dimensional data in the wavelet or curvelet domains. In a typical 1D wavelet transform, each coefficient at scale $j \in \{1, \ldots, J := \log_2(N)\}$ describes a portion of the signal of size $O(2^{-j})$. With 2^{j-1} such coefficients at each scale, a binary tree provides a natural organization for the coefficients. Each coefficient at scale $j < \log_2(N)$ has 2 *children* at scale $j + 1$, and each coefficient at scale $j > 1$ has one *parent* at scale $j - 1$.

Due to the analysis properties of wavelets, coefficient values tend to persist through scale. A large wavelet coefficient (in magnitude) generally indicates the presence of a singularity inside its support; a small wavelet coefficient generally indicates a smooth region. Thanks to the nesting of child wavelets inside their parents, edges in general manifest themselves in the wavelet domain as chains of large coefficients propagating across scales in the wavelet tree; we call this phenomenon the *persistence property*. Additionally, wavelet coefficients also have *exponentially decaying magnitudes* at finer scales [17]. This causes the significant wavelet coefficients of piecewise smooth signals to concentrate within a connected subtree of the wavelet binary tree.

This connected subtree structure was exploited by previous CS reconstruction algorithms known as Tree Matching Pursuit (TMP) and Tree Orthogonal Matching Pursuit (TOMP) [18, 19]. TMP and TOMP are modifications to the standard greedy algorithms matching pursuit and orthogonal matching pursuit, with the proviso that each selection made by the greedy algorithm build upon a connected tree. Such techniques enable faster algorithms due to the restriction in the greedy search and lower reconstruction distortion due to the inherent regularization.

However, for real-world piecewise smooth signals, the nonzero coefficients generally do not form a perfectly connected subtree. The reasons for this are twofold. First, since wavelets are bandpass functions, wavelet coefficients oscillate between positive and negative values around singularities. Second, due to the linearity of the wavelet transform, two or more singularities in the signal may cause destructive interference among coarse scale wavelet coefficients; that is, the persistence of the wavelets across scale is *weaker at coarser scales*. Either of these factors may cause the wavelet coefficient corresponding to a discontinuity to be small yet have large children, yielding a non-connected set of meaningful wavelet coefficients. TMP and TOMP used heuristic rules to ameliorate the effect of this phenomenon. However, this considerably increases the computational complexity, and the success of such heuristics varies markedly between different signals in the proposed class.

In summary, we have identified several properties of wavelet expansions:

- large/small values of wavelet coefficients generally persist across the scales of the wavelet tree;
- persistence becomes stronger as we move to finer scales; and
- the magnitude of the wavelet coefficients decreases exponentially as we move to finer scales.

We also note that the sparsity of the wavelet transform causes the coefficients to have a peaky, non-Gaussian distribution.

2.4 Hidden Markov Tree models

The properties identified in Section 2.3 induce a joint statistical structure among the wavelet coefficients that is far stronger than simple sparsity or the simple connectivity models used in the TMP and TOMP algorithms. HMT [20] offers one modeling

framework that succinctly and accurately captures this joint structure. HMT modeling has been used successfully to improve performance of denoising, classification, and segmentation algorithms for wavelet-sparse signals.

The HMT models the probability density function of each wavelet as a Gaussian mixture density with a hidden binary state that determines whether the coefficient is large or small. The persistence across scale is captured by a tree-based Markov model that correlates the states of parent and children coefficients. The following properties are captured by the HMT.

Non-Gaussianity: Sparse coefficients can be modeled probabilistically using a mixture of Gaussians: one component features a large variance that models large nonzero coefficients and receives a small weight (to encourage few such coefficients), while a second component features a small variance that models small and zero-valued coefficients and receives a large weight. We distinguish these two components by associating to each wavelet coefficient θ_n an unobserved hidden state $S_n \in \{S, L\}$; the value of S_n determines which of the two components of the mixture model is used to generate θ_n. Thus we have

$$\begin{aligned} f(\theta_n|S_n = S) &= \mathcal{N}(0, \sigma^2_{S,n}), \\ f(\theta_n|S_n = L) &= \mathcal{N}(0, \sigma^2_{L,n}), \end{aligned}$$

with $\sigma^2_{L,n} > \sigma^2_{S,n}$. To generate the mixture, we apply a probability distribution to the available states: $p(S_n = S) = p^S_n$ and $p(S_n = L) = p^L_n$, with $p^S_n + p^L_n = 1$.

Persistence: The perpetuation of large and small coefficients from parent to child is well-modeled by a Markov model that links coefficient states. This induces a Markov tree where the state S_n of a coefficient θ_n is affected only by the state $S_{\mathcal{P}(n)}$ of its parent $\mathcal{P}(n)$. The Markov model is then completely determined by the set of state transition matrices for the different coefficients θ_n at wavelet scales $1 < j \leq J$:

$$A_n = \begin{bmatrix} p^{S\to S}_n & p^{S\to L}_n \\ p^{L\to S}_n & p^{L\to L}_n \end{bmatrix}.$$

The persistence property implies that the values of $p^{L\to L}_n$ and $p^{S\to S}_n$ are significantly larger than their complements. If we are provided the hidden state probabilities for the wavelet coefficient in the coarsest scale p^S_1 and p^L_1, then the probability distribution for any hidden state can be obtain recursively:

$$p(S_n = L) = p^S_{\mathcal{P}(n)} p^{S\to L}_n + p^L_{\mathcal{P}(n)} p^{L\to L}_n.$$

As posed, the HMT parameters include the probabilities for the hidden state $\{p^S_1, p^L_1\}$, the state transition matrices A_n, and Gaussian distribution variances $\{\sigma^2_{L,n}, \sigma^2_{S,n}\}$ for each of the wavelet coefficients θ_n. To simplify the model, the coefficient-dependent parameters are made equal for all coefficients within a scale; that is, the new model has parameters A_j for $1 < j \leq J$ and $\{\sigma^2_{L,j}, \sigma^2_{S,j}\}$ for $1 \leq j \leq J$.

Magnitude decay: To enforce the decay of the coefficient magnitudes across scale, the variances $\sigma^2_{L,j}$ and $\sigma^2_{S,j}$ are modeled so that they decay exponentially as the scale becomes finer [21]:

$$\begin{aligned} \sigma^2_{L,j} &= C_{\sigma_L} 2^{-j\alpha_L}, \\ \sigma^2_{S,j} &= C_{\sigma_S} 2^{-j\alpha_S}. \end{aligned}$$

Since the wavelet coefficients that correspond to signal discontinuities decay slower than those representing smooth regions, the model sets $\alpha_S \geq \alpha_L$.

Scale-dependent persistence: To capture the weaker persistence present in the coarsest scales, the values of the state transition matrices A_j follow a model that strengthens the persistence at finer scales [21]. Additionally, the model must reflect that in general, any large parent generally implies only one large child (that which is aligned with the discontinuity). This implies that the probability that $S_n = L$, given that $S_{p(n)} = L$, should be roughly $1/2$. HMT accounts for both factors by setting

$$\begin{aligned} p^{L\to L}_j &= \frac{1}{2} + C_{LL} 2^{-\gamma_L j}, \quad p^{L\to S}_j = \frac{1}{2} - C_{LL} 2^{-\gamma_L j} \\ p^{S\to S}_j &= 1 - C_{SS} 2^{-\gamma_S j}, \quad \text{and } p^{S\to L}_j = C_{SS} 2^{-\gamma_S j}. \end{aligned}$$

Estimation: We can obtain estimates of all parameters

$$\Theta = \{p^S_1, p^L_1, \alpha_S, \alpha_L, C_{\sigma_L}, C_{\sigma_S}, \gamma_L, \gamma_S, C_{LL}, C_{SS}\}$$

for a set of coefficients θ using maximum likelihood estimation:

$$\Theta_{ML} = \arg\max_{\Theta} f(\theta|\Theta). \tag{3}$$

The expectation-maximization (EM) algorithm in [20] efficiently performs this estimation. Similarly, one can obtain the state probabilities $p(S_n = S|\theta, \Theta)$ using the Viterbi algorithm; the state probabilities for a given coefficient will be dependent on the states and coefficient values of all of its predecessors in the wavelet tree.

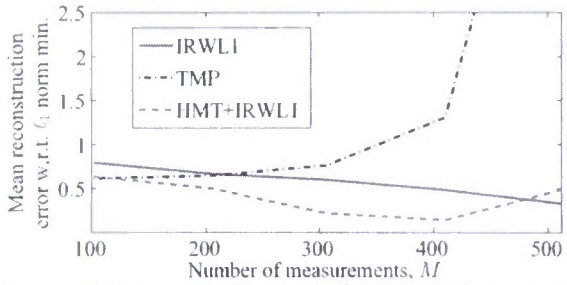


Figure 1: *Performance of IRWL1 algorithm, normalized by the performance of ℓ_1-norm minimization. Since all values are less than 1, IRWL1 and HMT+IRWL1 consistently outperforms ℓ_1-norm minimization.*

2.5 HMT-based weights for IRWL1

The IRWL1 algorithm described in Sec. 2.2.3 provides an opportunity to implement flexible signal penalizations while retaining the favorable computational complexity of ℓ_1-norm minimizations.

We now pose a new weight rule for the IRWL1 algorithm that integrates the HMT model to enforce the wavelet coefficient structure during CS reconstruction. Our weighting scheme, dubbed HMT+IRWL1, employs the following weighting scheme:

$$W_{n,n}^{(i)} = \left(p\left(S_n = L|\widehat{\theta}^{(i-1)}, \Theta\right) + \delta\right)^{-q}.$$

In words, for each wavelet coefficient in the current estimate we obtain the probability that the coefficient's hidden state is large; in the next iteration, we apply to that coefficient a weight that is inversely proportional to that probability. The parameter δ is a regularization parameter for cases where $p(S_n = L|\widehat{\theta}^{(i-1)})$ is very small, and the exponent q is a parameter that regulates the strength of the penalization for small coefficients. The goal of this weighting scheme is to penalize coefficients with large magnitudes that have low likelihood of being generated by a wavelet sparse signal; these coefficients are often the largest contributors to the reconstruction error.

The first step of HMT+IRWL1 consists of an initial training stage in which an EM algorithm solves (3) to estimate the values of the parameters for a representative signal; additionally, the solution $\widehat{\theta}^{(0)}$ for the standard formulation (1) is obtained. Subsequently, we proceed iteratively with two alternating steps: a weight update step in which the Viterbi algorithm for state probability calculations is executed for the previous solution $\widehat{\theta}^{(i-1)}$, and a reconstruction step in which the obtained weights are used in (2) to obtain an updated solution $\widehat{\theta}^{(i)}$. The convergence criterion for this algorithm is the same as for the IRWL1 algortihm.

Other probabilistic models for wavelet-sparse signals can also be used in combination with the IRWL1 algorithm, including generalized Gaussian densities [22], Gaussian scales mixtures [23], and hierarchical Dirichlet processes [24].

2.6 Simulations

We now compare the IRWL1 and HMT+IRWL1 algorithms. We use piecewise-smooth signals of length $N = 1024$, with 5 randomly placed discontinuities and cubic polynomial pieces with random coefficients. Daubechies-4 wavelets are used to sparsify the signals. Measurements are obtained using a matrix with i.i.d. Gaussian entries. For values of M ranging from 102 to 512, we test the ℓ_1-norm minimization and the IRWL1, TMP [18] and HMT+IRWL1 algorithms. We fix the number of iterations for IRWL1 and HMT+IRWL1 to 10. The parameters are set for best performance to $\epsilon = 0.2$, $q = 0.1$, and $\delta = 10^{-10}$. For each M we perform 100 simulations using different randomly generated signals and measurement matrices.

Figure 1 shows the magnitude of the reconstruction error for each of the algorithms, normalized by the error of the unweighted ℓ_1-norm minimization reconstruction, as a function of the iteration count. Figure 2 shows a reconstruction example. TMP performs well for smaller numbers of measurements M. IRWL1 consistently outperforms ℓ_1 minimization. Our proposed HMT+IRWL1 algorithm outperforms IRWL1 for most values of M. For large M near $N/2$, HMT+IRWL1 becomes less efficient than IRWL1; we speculate that at this stage the recovered signal has roughly equal numbers of large and small wavelet coefficients, which begins to violate the HMT model. Figure 2 plots the various reconstructed signals for one realization of the experiment, with $M = 300$.

2.7 Future work

We showed in our work that for piecewise smooth signals, the HMT+IRWL1 scheme reduces the artifacts of the ℓ_1-norm minimization reconstruction over fewer iterations than the standard iterative reweighted ℓ_1 minimization algorithm. Our scheme can be used as well in the extensions of the iterative reweighted ℓ_1-norm minimization scheme, such as reconstruction from noisy measurements. We look to extend this line of work to higher dimensional data. As this work demonstrated that using HMT assumptions aids in compressive sensing and signal recovery, we will broaden our approach to further explore a model-based paradigm for CS.

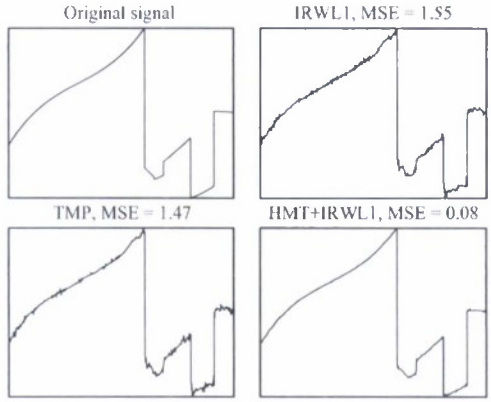


Figure 2: *Example outputs for the reconstruction algorithms.*

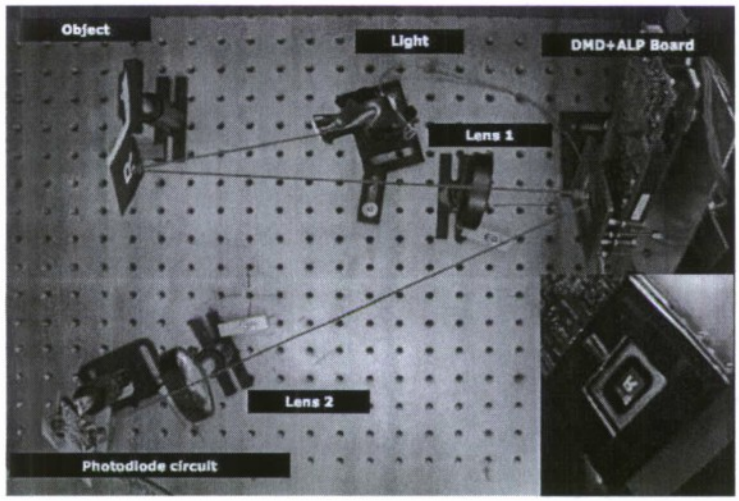


Figure 3: *Aerial view of the single-pixel compressive sampling (CS) camera in the lab [25].*

3 Implementation of CS practice

3.1 Summary of results

The single-pixel camera is a flexible architecture we used to physically implement CS. We analyzed the performance of CS and two other candidate multiplexing methodologies and compared them to the performance of a brute-force array of N pixel sensors. Integral to our analysis is the consideration of Poisson photon counting noise at the detector, which is image-dependent. We conducted two separate analyses to assess the "bang for the buck" of CS. The first was a theoretical analysis that provides general guidance. The second was an experimental study that indicates how the systems typically perform in practice.

3.2 Single-pixel camera overview

The single-pixel camera is an optical computer that sequentially measures the inner products $y[m] = \langle x, \phi_m \rangle$ between an N-pixel sampled version x of the incident light-field from the scene under view and a set of two-dimensional (2D) test functions $\{\phi_m\}$ [25]. As shown in Fig. 3, the light-field is focused by biconvex Lens 1 not onto a CCD or CMOS sampling array but rather onto a DMD consisting of an array of N tiny mirrors.

Each mirror corresponds to a particular pixel in x and ϕ_m and can be independently oriented either towards Lens 2 (corresponding to a 1 at that pixel in ϕ_m) or away from Lens 2 (corresponding to a 0 at that pixel in ϕ_m). The reflected light is then collected by biconvex Lens 2 and focused onto a single photon detector (the single pixel) that integrates the product $x[n]\phi_m[n]$ to compute the measurement $y[m] = \langle x, \phi_m \rangle$ as its output voltage. This voltage is then digitized by an A/D converter. Values of ϕ_m between 0 and 1 can be obtained by dithering the mirrors back and forth during the photodiode

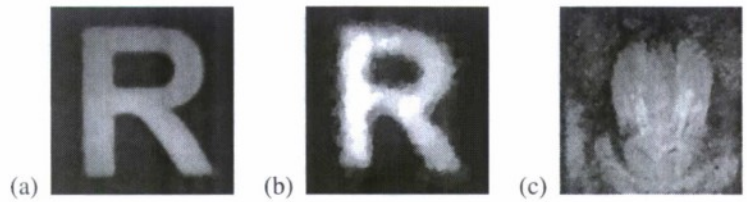


Figure 4: *Single-pixel photo album. (a) 256×256 conventional image of a black-and-white* R. *(b) Single-pixel camera reconstructed image from $M = 1300$ random measurements ($50\times$ sub-Nyquist). (c) 256×256 pixel color reconstruction of a printout of the Mandrill test image imaged in a low-light setting using a single photomultiplier tube sensor, RGB color filters, and $M = 6500$ random measurements.*

integration time. To obtain ϕ_m with both positive and negative values (± 1, for example), we estimate and subtract the mean light intensity from each measurement, which is easily measured by setting all mirrors to the full-on 1 position.

To compute CS randomized measurements $\mathbf{y} = \Phi\mathbf{x}$ as in

$$\mathbf{y} = \Phi\mathbf{x} = \Phi\Psi\alpha,$$

we set the mirror orientations ϕ_m randomly using a pseudo-random number generator, measure $y[m]$, and then repeat the process M times to obtain the measurement vector $\mathbf{y}$. We can set $M = O(K \log(N/K))$ which is $\ll N$ when the scene being imaged is compressible by a compression algorithm like JPEG or JPEG2000. Since the DMD array is programmable, we can also employ test functions ϕ_m drawn randomly from a fast transform such as a Walsh, Hadamard, or Noiselet transform [26, 27].

The single-pixel design reduces the required size, complexity, and cost of the photon detector array down to a single unit, which enables the use of exotic detectors that would be impossible in a conventional digital camera. Example detectors include a photomultiplier tube or an avalanche photodiode for low-light (photon-limited) imaging (more on this below), a sandwich of several photodiodes sensitive to different light wavelengths for multimodal sensing, a spectrometer for hyperspectral imaging, and so on.

In addition to sensing flexibility, the practical advantages of the single-pixel design include the facts that the quantum efficiency of a photodiode is higher than that of the pixel sensors in a typical CCD or CMOS array and that the fill factor of a DMD can reach 90% whereas that of a CCD/CMOS array is only about 50%. An important advantage to highlight is the fact that each CS measurement receives about $N/2$ times more photons than an average pixel sensor, which significantly reduces image distortion from dark noise and read-out noise. Theoretical advantages that the design inherits from the CS theory include its universality, robustness, and progressivity.

The single-pixel design falls into the class of multiplex cameras [28]. The baseline standard for multiplexing is classical raster scanning, where the test functions $\{\phi_m\}$ are a sequence of delta functions $\delta[n - m]$ that turn on each mirror in turn. As we will see below, there are substantial advantages to operating in a CS rather than raster scan mode, including fewer total measurements (M for CS rather than N for raster scan) and significantly reduced dark noise.

3.2.1 Image acquisition examples

Figure 4 (a) and (b) illustrates a target object (a black-and-white printout of an "R") $\mathbf{x}$ and reconstructed image $\widehat{x}$ taken by the single-pixel camera prototype in Fig. 3 using $N = 256 \times 256$ and $M = N/50$ [25]. Fig. 4(c) illustrates an $N = 256 \times 256$ color single-pixel photograph of a printout of the Mandrill test image taken under low-light conditions using RGB color filters and a photomultiplier tube with $M = N/10$. In both cases, the images were reconstructed using Total Variation minimization, which is closely related to wavelet coefficient ℓ_1 minimization [29].

3.3 Comparative scanning methodologies

In our comparative analysis, the four imaging methodologies we consider are:

- **Pixel array** (PA): a separate sensor for each of the N pixels receives light throughout the total capture time T. This is actually not a multiplexing system, but we use it as the gold standard for comparison.
- **Raster scan** (RS): a single sensor takes N light measurements sequentially from each of the N pixels over the capture time. This corresponds to test functions $\{\phi_m\}$ that are delta functions and thus $\Phi = I$. The measurements y thus directly provide the acquired image $\widehat{x}$.
- **Basis scan** (BS): a single sensor takes N light measurements sequentially from different combinations of the N pixels as determined by test functions $\{\phi_m\}$ that are not delta functions but from some more general basis [30]. In our analysis,

Table 1: *Comparison of the four camera scanning methodologies.*

	Pixel Array	Raster Scan	Basis Scan	Compressive Sampling
Number of measurements	N	N	N	$M \leq N$
Dynamic range	D	D	$\frac{ND}{2}$	$\frac{ND}{2}$
Quantization (total bits)	NB	NB	$N(B + \log_2 N)$	$M(B + \log_2 N + \log_2 C_N + 1)$
Photon counting MSE	$\frac{P}{T}$	$N\frac{P}{T}$	$(3N-2)\frac{P}{T}$	$< 3C_N^2 M \frac{P}{T}$

we assume a Walsh basis modified to take the values 0/1 rather than ± 1; thus $\Phi = W$, where W is the 0/1 Walsh matrix. The acquired image is obtained from the measurements y by $\widehat{x} = \Phi^{-1}y = W^{-1}y$.

- **Compressive sampling** (CS): a single sensor takes $M \leq N$ light measurements sequentially from different combinations of the N pixels as determined by random 0/1 test functions $\{\phi_m\}$. Typically, we set $M = O(K \log(N/K))$ which is $\ll N$ when the image is compressible. In our analysis, we assume that the M rows of the matrix Φ consist of randomly drawn rows from a 0/1 Walsh matrix that are then randomly permuted (we ignore the first row consisting of all 1's). The acquired image is obtained from the measurements y via a sparse reconstruction algorithm such as BPIC (see Sidebar: Compressive Sampling in a Nutshell).

3.4 Theoretical analysis

We conduct a theoretical performance analysis of the above four scanning methodologies in terms of the required dynamic range of the photodetector, the required bit depth of the A/D converter, and the amount of Poisson photon counting noise. Our results are pessimistic in general; we show in the next section that the average performance in practice can be considerably better. Our results are summarized in Table 1. An alternative analysis of CS imaging for piecewise smooth images in Gaussian noise has been reported in [31].

Dynamic range: We first consider the photodetector dynamic range required to match the performance of the baseline PA. If each detector in the PA has a linear dynamic range of 0 to D, then it is easy to see that single-pixel RS detector need only have that same dynamic range. In contrast, each Walsh basis test function contains $N/2$ 1's, and so directs $N/2$ times more light to the detector. Thus, BS and CS each require a larger linear dynamic range of 0 to $ND/2$. On the positive side, since BS and CS collect considerably more light per measurement than the PA and RS, they benefit from reduced detector nonidealities like dark currents.

3.4.1 Quantization error

We now consider the number of A/D bits required within the required dynamic range to match the performance of the baseline PA in terms of worst-case quantization distortion. Define the mean-squared error (MSE) between the true image x and its acquired version $\widehat{x}$ as MSE $= \frac{1}{N}\|\mathbf{x} - \widehat{\mathbf{x}}\|_2^2$. Assuming that each measurement in the PA and RS is quantized to B bits, the worst-case mean-squared quantization error for the quantized PA and RS images is MSE $= \sqrt{N}D2^{-B-1}$ [32]. Due to its larger dynamic range, BS requires $B + \log_2 N$ bits per measurement to reach the same MSE distortion level. Since the distortion of CS reconstruction is up to C_N times larger than the distortion in the measurement vector, we require up to an additional $\log_2 C_N$ bits per measurement. One empirical study has found roughly that C_N lies between 8 and 10 for a range of different random measurement configurations [33]. Thus, BS and CS require a higher-resolution A/D converter than PA and RS to acquire an image with the same worst-case quantization distortion.

3.4.2 Photon counting noise

In addition to quantization error from the A/D converter, each camera will also be affected by image-dependent Poisson noise due to photon counting [34]. We compare the MSE due to photon counting for each of the scanning methodologies. The details are summarized in Table 1. We see that the MSE of BS is about three times that of RS. Moreover, when $M < \frac{N}{3C_N^2}$, the MSE of CS is lower than that of RS. We emphasize that in the CS case, we have only a fairly loose upper bound and that there exist alternative CS reconstruction algorithms with better performance guarantees, such as the Dantzig Selector [12].

3.4.3 Summary

From Table 1, we see that the advantages of a single-pixel camera over a PA come at the cost of more stringent demands on the sensor dynamic range and A/D quantization and larger MSE due to photon counting effects. Additionally, the linear de-

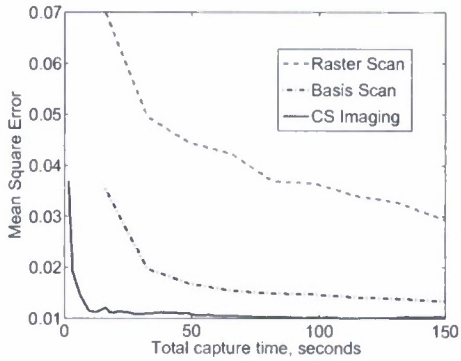


Figure 5: *Average MSE for Raster Scan (RS), Basis Scan (BS), and Compressive Sampling (CS) single-pixel images as a function of the total image capture time T (real data).*

pendence of the MSE on the number of image pixels N for BS and RS is a potential deal-breaker for high-resolution imaging. The promising aspect of CS is the logarithmic dependence of its MSE on N through the relationship $M = O(K \log(N/K))$.

3.5 Experimental results

Since CS acquisition/reconstruction methods often perform much better in practice than the above theoretical bounds suggest, in this section we conduct a simple experiment using real data from the CS imaging testbed depicted in Fig. 3. Thanks to the programmability of the testbed, we acquired RS, BS, and CS measurements from the same hardware. We fixed the number of A/D converter bits across all methodologies. Figure 5 shows the pixel-wise MSE for the capture of a $N = 128 \times 128$ pixel "R" test image as a function of the total capture time T. Here the MSE combines both quantization and photon counting effects. For CS we took $M = N/10$ total measurements per capture and used a Daubechies-4 wavelet basis for the sparse reconstruction.

We make two observations. First, the performance gain of BS over RS contradicts the above worst-case theoretical calculations. We speculate that the contribution of the sensor's dark current, which is not accounted for in our analysis, severely degrades RS's performance. Second, the performance gain of CS over both RS and BS is clear: images can either be acquired in much less time for the same MSE or with much lower MSE in the same amount of time.

3.6 Future work

We presented an overview of a simple yet flexible single-pixel architecture for CS based on a DMD spatial light modulator. While there are promising potential applications where current digital cameras have difficulty imaging, there are clear tradeoffs and challenges in the single-pixel design. Our current and planned work involves better understanding and addressing these tradeoffs and challenges. Other potential avenues for research include extending the single-pixel concept to wavelengths where the DMD fails as a modulator, such as THz and X-rays.

4 Application of CS principles to optics

4.1 Summary of results

We developed two single-shot spectral imaging approaches based on the concept of compressive sensing. The primary features of the system designs were either one or two dispersive elements in conjunction with a binary-valued aperture code. In contrast to thin-film approaches to spectral filtering, this structure resulted in easily-controllable, spatially-varying, spectral filter functions with narrow features. Measurement of the input scene through these filters was equivalent to projective measurement, and hence was treated with compressive sensing frameworks. We created two alternative reconstruction frameworks and demonstrated their application to experimental data.

4.2 Compressive spectral imaging

Spectral imaging is an emerging tool for a variety of scientific and engineering applications because of the additional information it provides about the nature of the materials being imaged. Traditional imagers produce two-dimensional spatial arrays of scalar values representing the intensity of a scene. A spectral imager, in contrast, produces a two-dimensional spatial array of

vectors which contain the spectral information for the respective spatial locations. The resulting data is known as the *spectral data cube* because of its three-dimensional nature. The addition of spectral information can provide valuable information in a variety of contexts ranging from environmental monitoring [35, 36] to astrophysics [37] and from biochemistry [38, 39] to security applications [40].

Adoption of spectral imaging has been slow because of a fundamental tradeoff between spatial resolution, spectral resolution, light collection, and measurement acquisition time. Standard spectral imaging designs can simultaneously optimize only two of the four quantities—resulting in relatively poor overall performance. The origin of these tradeoffs can be readily understood. Traditional, non-multiplexed, spectrometers already exhibit a tradeoff between light collection and spectral resolution. Any system that attempts to use one of these systems as the spectrograph in a spectral imager inherits this limitation. Further, the fact that the spectral image data cube is three-dimensional, while available detector arrays are two-dimensional results in either a need for scanning (which increases the overall acquisition time) or in the tiling of the detector array with multiple two-dimensional slices of the cube (which, if the field of view is held fixed, limits the spatial resolution). In the past decade, however, there have been a number of ingenious designs that allow independent control of three of these quantities simultaneously [41–43], largely through the introduction of various multiplex measurement techniques.

The work described in this report attempts to eliminate the remaining tradeoffs by presenting two novel spectral imaging systems and associated reconstruction frameworks that fully decouple the four operational quantities. Both imagers are completely static, single-shot designs, resulting in mechanically robust and inexpensive systems. In these respects the system architectures are a descendant of the coded-aperture spectroscopy architecture previously developed by several of the current authors [44–48]. In our first extension to spectral imaging [43], however, the implementation required a sequence of exposures in order to measure the contents of the spectral data cube. The system architecture described in this report has been developed precisely to avoid that shortcoming. In this system, the imager does not directly measure each voxel in the data cube. Instead, it collects a small number (relative to the size of the data cube) of coded measurements, and then a novel reconstruction method is used to estimate the spectral image from the noisy projections. The two systems described in this report are called the *dual disperser coded aperture snapshot spectral imager* (DD-CASSI) and *single disperser coded aperture snapshot spectral imager* (SD-CASSI), depending on whether one or two dispersive elements are present in the design.

This approach draws heavily on ideas in the emerging field of compressed sensing [13, 29, 49]. In compressed sensing, certain design strategies are incorporated into measurement systems in a way that can dramatically improve the system's ability to produce high-quality reconstructions from a limited number of measurements. The basic idea of this theory is that when the signal of interest is very sparse (i.e., zero-valued at most locations) or highly compressible in some basis, relatively few *incoherent* observations are necessary to reconstruct the most significant non-zero signal components. In the remainder of this manuscript we demonstrate the practical application of these ideas to spectral imaging. We describe a particular system design, present a multiscale reconstruction algorithm, and demonstrate that accurate spectroscopic images can be estimated from an under-determined set of noisy projections.

4.3 DD-CASSI

This section describes a single-shot spectral imaging approach based on the concept of compressive sensing called the *dual disperser coded aperture snapshot spectral imager* (DD-CASSI). The primary features of the system design are *two dispersive elements*, arranged in opposition and surrounding a binary-valued aperture coded. In contrast to thin-film approaches to spectral filtering, this structure results in easily-controllable, spatially-varying, spectral filter functions with narrow features. Measurement of the input scene through these filters is equivalent to projective measurement in the spectral domain, and hence can be treated with the compressive sensing frameworks recently developed by a number of groups. We present a reconstruction framework and demonstrate its application to experimental data.

4.3.1 DD-CASSI system design

The system is comprised of two sequential dispersive arms of the 4-f type commonly used (singly) as a traditional dispersive spectrometer. The two arms are arranged in opposition so that the second arm exactly cancels the dispersion introduced by the first arm. A coding aperture occupies the plane separating the two arms. A schematic of the system is shown in Fig. 6.

The operational characteristics of the system can be easily understood on a conceptual level. A standard imaging relay (not shown) is used to form an image of a remote scene in the plane of the input aperture. The input aperture is then imaged through the first arm onto the plane containing the coding aperture. However, because the arm contains a dispersive element, multiple images are formed at wavelength-dependent locations. At this point, the spatial structure in the plane of the coding aperture contains a mixture of spatial and spectral information about the source. Passing through the coding aperture modulates this information with the applied pattern. The second arm then undoes the spatial-spectral mixing introduced by the first arm as it forms an image of the source on the detector. In the process of undoing the effects of arm 1, the spatial modulation introduced

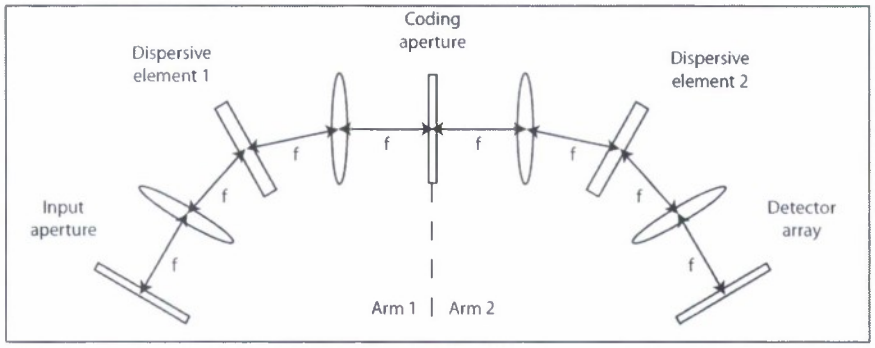


Figure 6: Schematic of the spectral imager.

by the coding aperture is transformed into a spatial and spectral modulation. The details of the process are described in the following section.

4.3.2 System model

We denote the spectral density entering the instrument as $S_0(x, y; \lambda)$. The spectral density just prior to the code aperture is then

$$\begin{aligned} S_1(x, y; \lambda) &= \iint dx'\, dy'\, \delta(x' - [x + \alpha(\lambda - \lambda_c)])\, \delta(y' - y)\, S_0(x', y'; \lambda) \\ &= S_0(x + \alpha(\lambda - \lambda_c), y; \lambda). \end{aligned} \tag{4}$$

Here the Dirac delta functions describe propagation through unity-magnification imaging optics and a dispersive element with linear dispersion α and center wavelength λ_c. (Note that this model assumes linear dispersion—which is approximately true only over limited wavelength ranges. The system can still be operated in nonlinear regions, as the processing algorithm is calibration based. The linear model is only used to provide guidance about the aperture code.)

Immediately after the coding aperture the spectral density is given by

$$S_2(x, y; \lambda) = T(x, y)\, S_1(x, y; \lambda) = T(x, y)\, S_0(x + \alpha(\lambda - \lambda_c), y; \lambda),$$

with $T(x, y)$ the spatial transmission pattern imposed by the coding aperture.

Propagation through the second set of imaging optics and the second dispersive element results in a spectral density in the detector plane of

$$\begin{aligned} S_3(x, y; \lambda) &= \iint dx'\, dy'\, \delta(x' - [x - \alpha(\lambda - \lambda_c)])\, \delta(y' - y)\, S_2(x', y'; \lambda) \\ &= T(x - \alpha(\lambda - \lambda_c), y)\, S_0(x, y; \lambda) \\ &= F(x, y; \lambda)\, S_0(x, y; \lambda). \end{aligned} \tag{5}$$

Again, the Dirac delta functions describe propagation through the imaging optics and disperser (note the internal sign change representing the reversed orientation of the disperser). Here we have defined the spectral density filter function $F(x, y; \lambda) = T(x - \alpha(\lambda - \lambda_c), y)$. This formulation makes explicit the fact that the two-dimensional coding pattern introduces a three-dimensional filter function that acts on the source spectral density.

As the detector is wavelength-insensitive, the ultimate quantity measured is not the spectral density in the detector plane, but the intensity

$$I(x, y) = \int d\lambda\, F(x, y; \lambda)\, S_0(x, y; \lambda).$$

Further, the detector array is spatially pixelated. If we take the pixel size as Δ, then the detector measurements become

$$I_{nm} = \iiint dx\, dy\, d\lambda\, \mathrm{rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n\right) F(x, y; \lambda)\, S_0(x, y; \lambda). \tag{6}$$

It then becomes natural to consider a coding aperture where the transmission function T is pixelated with features that are the same size as the detector pixels (Note that for implementation reasons, it is actually preferable to consider codes where the

features sizes are integer multiples of the detector pixel size. Extending this formal treatment to that case is straightforward.)

$$T(x,y) = \sum_{m',n'} T_{n'm'} \text{ rect}\left(\frac{x}{\Delta} - m', \frac{y}{\Delta} - n'\right). \tag{7}$$

With this, (6) becomes

$$I_{nm} = \sum_{m'n'} \iiint dx\, dy\, d\lambda \text{ rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n\right) \text{rect}\left(\frac{x - \alpha(\lambda - \lambda_c)}{\Delta} - m', \frac{y}{\Delta} - n'\right) \times T_{n'm'}\, S_0(x,y;\lambda). \tag{8}$$

To gain an understanding of how the coding pattern influences the measured intensity distribution, it is instructive to consider two highly constrained cases. First we consider a monochromatic source with $S_0(x,y,\lambda) = I_0(x,y)\,\delta(\lambda - \lambda_c)$, where $I_0(x,y)$ is the intensity distribution of the monochromatic scene. In this case, (8) simplifies significantly

$$\begin{aligned} I_{nm}|_{\lambda=\lambda_c} &= \sum_{m'n'} \iiint dx\, dy\, d\lambda \text{ rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n\right) \text{rect}\left(\frac{x}{\Delta} - m', \frac{y}{\Delta} - n'\right) \\ &\qquad \times T_{n'm'}\, I_0(x,y)\,\delta(\lambda - \lambda_c) \\ &= \sum_{m'n'} \delta_{mm'}\,\delta_{nn'}\, T_{n'm'}\, I_{0,nm} \\ &= T_{nm}\, I_{nm}, \end{aligned} \tag{9}$$

where $I_{0,nm}$ is a spatially-pixelated version of the monochromatic source intensity $I_0(x,y)$. Next, we consider the response to a monochromatic source at $\lambda = \lambda_c + \Delta\lambda$ with $\Delta\lambda = \Delta/\alpha$, with α the linear dispersion of the dispersive elements. In this case, we find

$$\begin{aligned} I_{nm}|_{\lambda=\lambda_c+\Delta\lambda} &= \sum_{m'n'} \iiint dx\, dy\, d\lambda \text{ rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n\right) \text{rect}\left(\frac{x}{\Delta} - (m'+1), \frac{y}{\Delta} - n'\right) \\ &\qquad \times T_{n'm'}\, I_0(x,y)\,\delta(\lambda - (\lambda_c + \Delta\lambda)) \\ &= \sum_{m'n'} \delta_{mm'}\,\delta_{nn'}\, T_{n'(m'+1)}\, I_{0,nm} \\ &= T_{n(m-1)}\, I_{0,nm}. \end{aligned} \tag{10}$$

Thus we see that the contribution from a particular wavelength is the pixelated source weighted by a version of the aperture code that shifts in the x direction as the wavelength changes. At the center wavelength, the weighting pattern is aligned with the detector pixels. For wavelength shifts that are integer multiples of $\Delta\lambda$, the pattern is again registered with the detector pixels. However, for intermediate wavelengths, the elements in the coding pattern straddle multiple detector pixels.

Finally, we define the filter function

$$\begin{aligned} h_{nmp} = \sum_{m'n'} \iiint dx\, dy\, d\lambda \text{ rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n, \frac{\lambda - \lambda_c}{\Delta\lambda} - p\right) \\ \times \text{rect}\left(\frac{x - \alpha(\lambda - \lambda_c)}{\Delta} - m', \frac{y}{\Delta} - n'\right) T_{n'm'}, \end{aligned} \tag{11}$$

and the fully pixelated source spectral density as

$$s_{nmp} = \frac{1}{\Delta^2 \Delta\lambda} \iiint dx\, dy\, d\lambda \text{ rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n, \frac{\lambda - \lambda_c}{\Delta\lambda} - p\right) S_0(x,y;\lambda).$$

If the source spectrum is slowly varying on the scale $\Delta\lambda$, then we can approximate (6) as

$$I_{nm} = \sum_p h_{nmp}\, s_{nmp}. \tag{12}$$

4.3.3 Code design

The analysis in the previous section determined the nature of the spatio-spectral filter that is applied and how it is related to the code pattern placed in the system. It is most convenient to view the filter function w_{nmp} as a spatially varying collection of spectral filters. In this framework, the analysis above reveals the following facts:

1. The spectral filter for a given spatial location may contain structure on a scale as narrow as $\Delta\lambda = \Delta/\alpha$. Typically, this is significantly narrower than the features in a traditional thin-film interference filter.

2. The spectral filters in different rows of the detector (different n) are unconstrained. There are no correlations imposed by the physical structure.

3. The spectral filters within a row (same n) are cyclic permutations of each other. This arises from the spatial shift in the m-direction that arises from spectral shifts. (e.g., In a hypothetical instrument with 4 spectral channels, the spectral filter applied to a particular pixel had values $\{a, b, c, d\}$, then the filters on its left/right neighbors are $\{b, c, d, a\}$ and $\{d, a, b, c\}$, respectively).

4. The measurement at a given detector pixel is the inner product of the spectral filter for that pixel and the pixelated source spectrum for that spatial location (see (12)).

To create a system with M spectral channels requires a 1-D code of at least length M. For the remainder of this section, we assume a code of length M for simplicity. The physical nature of the system produces spectral filters that are all M possible cyclic shifts of the fundamental code. Thus we are led to consider well-conditioned codes that consist of cyclic permutations of a single master codeword. The canonical example of these types of codes are those based on the *cyclic S-matrices* [7]. For our initial system (and earliest simulations), we drew upon the order-15 cyclic S-matrix with the fundamental codeword "100100011110101". This code provides several advantages, including M *unique* cyclic shifts and an overall transmission efficiency of $8/15$.

If the coding plane were directly tiled with this pattern, the various filter functions would be implemented on the detector plane in the manner depicted in Fig. 7. In this image, the number k denotes the kth spectral filter function, which corresponds to the fundamental codeword circularly shifted by $k-1$ bits; thus 1 refers to the fundamental codeword "100100011110101", 2 refers to the shifted codeword "001000111101011", etc. The difficulty with this arrangement is that there are no compact regions that contain all of the filter functions.

15	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1
15	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1
15	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1

Figure 7: Distribution of filter functions that arises from simple tiling of the fundamental codeword. No compact region contains all 15 filters.

If we instead tile the aperture plane with a unit cell of the form

100100011110101
001111010110010
101011001000111

(three copies of the fundamental codeword, but with circular shifts of five elements between rows) then the filter functions are arrayed on the detector plane as displayed in Fig. 8. In this scheme, any 3×5 region on the detector contains all 15 spectral filters.

10	11	12	13	14	15	1
15	1	2	3	4	5	6
5	6	7	8	9	10	11
10	11	12	13	14	15	1
15	1	2	3	4	5	6

Figure 8: Distribution of filter functions that arises from the more complicated unit cell. Any 3×5 region contains all 15 filters.

4.3.4 Multiscale reconstruction method

In this section we describe the multiscale reconstruction process employed to get the spatial and spectral information from the mask-modulated intensity information represented in (12). First we formulate a model for the noisy observations we measure at the detector array, and then describe the estimation process. We compute an optimal solution to this underdetermined problem using an expectation maximization algorithm combined with a multiscale Poisson denoising technique. The following subsections give a systematic description of our approach.

Let the spectral image of interest be denoted $\boldsymbol{s} = \{s_{nmp}\}$, and the intensity of the observations as $\boldsymbol{I} = \{I_{nm}\}$, so that equation (12) can be written in matrix-vector notation as

$$\boldsymbol{I} = \boldsymbol{H}\boldsymbol{s},$$

where the matrix $\boldsymbol{H}$ performs the discretized filtering described in Sec. 4.3.2. In addition, we model our observed data as

$$\boldsymbol{d} \sim \text{Poisson}(\boldsymbol{I}) = \text{Poisson}(\boldsymbol{H}\boldsymbol{s}),$$

so that the likelihood of observing $\boldsymbol{d}$ given spectral image $\boldsymbol{s}$ is

$$p(\boldsymbol{d}|\boldsymbol{H}\boldsymbol{s}) = \prod_{n=1}^{N}\prod_{m=1}^{M} \frac{e^{-\sum_p h_{nmp}s_{nmp}}\left(\sum_p h_{nmp}s_{nmp}\right)^{d_{nm}}}{d_{nm}!}.$$

Note that $\boldsymbol{H}$ has many more columns than rows (by a factor of M), making this a very underdetermined problem.

To solve this challenging inverse problem, we seek a solution ($\widehat{\boldsymbol{s}}$) which is both a good match to the data ($\boldsymbol{d}$) and *sparse*. In particular, we solve the following optimization problem:

$$\widehat{\boldsymbol{s}} = \arg\min_{\widetilde{\boldsymbol{s}}\in\mathcal{S}} \left\{-\log p(\boldsymbol{d}|\boldsymbol{H}\widetilde{\boldsymbol{s}}) + \text{pen}(\widetilde{\boldsymbol{s}})\right\}, \tag{13}$$

where $\mathcal{S}$ is a collection of estimators to be described below, and $\text{pen}(\widetilde{\boldsymbol{s}})$ is a penalty is proportional to the sparsity of $\widetilde{\boldsymbol{s}}$ in a multiscale spatio-spectral representation. This will be described in detail below.

The use of sparsity to solve challenging and ill-posed inverse problems has received widespread attention recently [13, 14, 29, 33, 49]. The objective function proposed in (13) in particular is similar to those proposed in [14, 33], in that we seek a solution accurately represented by a small number of cells in a *recursive dyadic partition* (*i.e.* sparse), as described below. We compute the solution to this problem using an Expectation-Maximization algorithm, which in this case is a regularized version of the Richardson-Lucy algorithm [50–52]. The method consists of two alternating steps:

Step 1:

$$\boldsymbol{y}^{(t)} = \widehat{\boldsymbol{s}}^{(t)}.\times \boldsymbol{H}^T(\boldsymbol{d}./\boldsymbol{H}\widehat{\boldsymbol{s}}^{(t)}),$$

where $.\times$ and $./$ denote element-wise multiplication and division, respectively.

Step 2: Compute $\widehat{\boldsymbol{s}}^{(t+1)}$ by denoising $\boldsymbol{y}^{(t)}$:

$$\widehat{\boldsymbol{s}}^{(t+1)} = \arg\min_{\widetilde{\boldsymbol{s}}} \left\{-\log p(\boldsymbol{y}^{(t)}|\widetilde{\boldsymbol{s}}) + \text{pen}(\widetilde{\boldsymbol{s}})\right\}.$$

The penalized likelihood denoising method employed in Step 2 takes advantage of correlations in the data between both wavelengths and spatial locations. The proposed method entails performing hereditary Haar intensity estimation via tree pruning in the spatial dimensions, with each leaf of the resulting unbalanced quad-tree decomposition corresponding to a region of spatially homogeneous spectra.

In particular, we determine the ideal partition of the spatial domain of observations and use maximum likelihood estimation to fit a single spectrum to each square in the optimal spatial partition. The space of possible partitions is a nested hierarchy defined through a recursive dyadic partition (RDP) of the datacube domain. The optimal partition is selected by merging neighboring squares of (*i.e.* pruning) a quad-tree representation of the observed data to form a data-adaptive RDP $\mathcal{P}$. Each of the terminal squares in the pruned spatial RDP could correspond to a region of intensity which is spatially homogeneous or smoothly varying (regardless of the regularity or irregularity between the spectral bands). This gives our estimators the capability of spatially varying the resolution to automatically increase the smoothing in very regular regions of the intensity and to preserve detailed structure in less regular regions.

Given a partition $\mathcal{P}$, $\widetilde{s}(\mathcal{P})$ can be calculated by finding the "best" spectrum fit to the observations over each cell in $\mathcal{P}$. This can be accomplished by simply computing the mean observed spectrum in each cell of $\mathcal{P}$. The final spatio-spectral estimate is then calculated by finding the partition which minimizes the total penalized likelihood function:

$$\begin{aligned} \widehat{\mathcal{P}} &\equiv \arg\min_{\mathcal{P}} \left\{ -\log p(\boldsymbol{y}^{(t)} | \widetilde{s}(\mathcal{P})) + \mathrm{pen}(\mathcal{P}) \right\} \\ \widehat{s} &\equiv \widetilde{s}(\widehat{\mathcal{P}}), \end{aligned} \tag{14}$$

where $\mathrm{pen}(\mathcal{P})$ is a penalty proportional to the number of cells in the RDP, encouraging a sparse solution (in terms of the size of the RDP). This method is described in detail in [53].

This approach is similar to the image estimation method described in [54, 55], with the key distinction that the proposed method *forces the spatial RDP to be the same at every spectral band.* A sample such partition is displayed in Fig. 9. This constraint makes it impossible for the method to perform spatial smoothing at some spectral bands but not others. In other words, when a tree branch is pruned in the proposed framework, it means partition cells are merged in every spectral band simultaneously at the corresponding spatial location. This approach is effective because an outlier observation in one spatio-spectral voxel may not be recognized as such when spectral bands are considered independently, but may be correctly pruned when the corresponding spectrum is very similar to spatially nearby spectra.

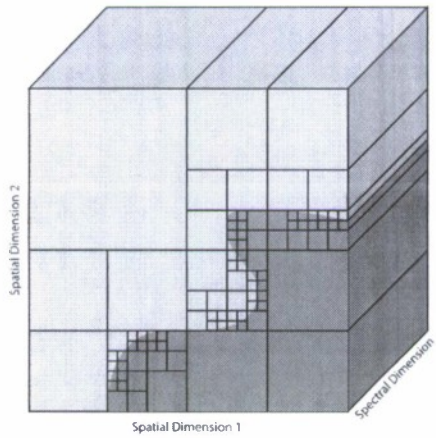


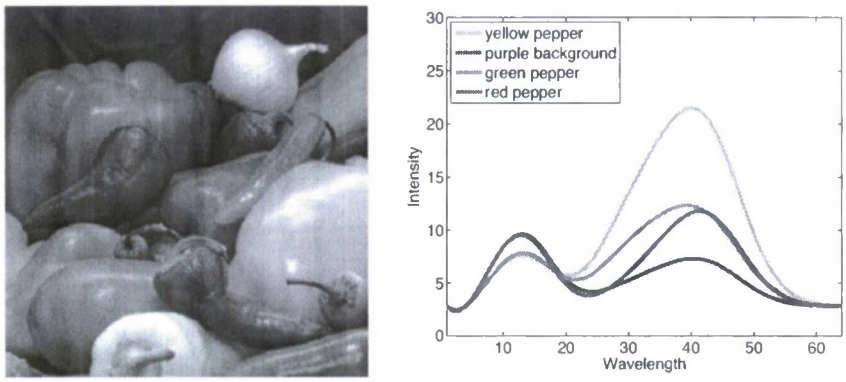
Figure 9: Sample partition of a spatio-spectral data cube. The spatial partition is the same at each spectral band, making it impossible for the estimation method to perform spatial smoothing at some spectral bands but not others.

The accuracy of these estimates can be augmented by a process called cycle-spinning, or averaging over shifts, resulting in translation-invariant (TI) estimates [56]. Cycle-spinning was derived in the context of undecimated wavelet coefficient thresholding in the presense of Gaussian noise, but is difficult to implement efficiently in the case of tree-pruning methods. The above multiscale tree-pruning methods can be modified to produce the same effect by averaging over shifts, but the increase in quality comes at a high computational cost. Novel computational methods [55], however, can be used to yield TI-Haar tree pruning estimates in $O\left(N^2 M \log N\right)$ time for an $N \times N \times M$ data cube.

4.3.5 Denoising simulation results

As described above, the proposed method exploits spatial homogeneities while simultaneously taking advantage of correlations between neighboring spectral bands. This significantly reduces over-smoothing across spatial boundaries and edges while improving the reconstruction of each spectrum. To test the effectiveness of the proposed method, we develop a phantom using a color image of peppers, as displayed in Figure 10. For each pixel in this image, we define a spectrum in our phantom data cube which corresponds to the color in the image. This results in a $256 \times 256 \times 64$ data cube, with a mean intensity (photon count) of 7.04 per voxel.

The spatially-varying intensities for three representative spectral bands (one from the red portion of the simulated spectrum (16th spectral band), one green (24th spectral band), and one blue (53rd spectral band)) are displayed in Figure 11(a), (e), and (i). Note the faint contrast between some of the features. Noisy observations in the same three spectral bands are displayed in Figure 11(b), (f), and (j). Several faint or low-contrast features are not easily discernible by the eye due to the low photon counts. In this simulation study, the penalty term in (13) was multiplied by a scalar weight to improve empirical performance. The weight was not selected to minimize any particular error metric, but rather to yield generally accurate reconstructions.



(a) (b)

Figure 10: Test data set. (a) Image used to generate $256 \times 256 \times 16$ spatio-spectral data cube. (b) Representative selection of generated spectra in different image regions. Wavelength units are normalized so each spectral band corresponds to one wavelength unit.

If we were to reconstruct this data cube by performing hereditary TI-Haar *image* estimation [55] on each spectral band, we would achieve the results displayed in Figure 11(c), (g), and (k); the mean squared error associated with this imaging-based data cube estimate, denoted $\widehat{s}_I$, is $\|s - \widehat{s}_I\|^2/\|s\|^2 = 5.22 \cdot 10^{-3}$. (While the images may appear oversmoothed, decreasing the weight on the penalty term results in strong noisy artifacts and increases the MSE significantly.) Spatio-spectral analysis offers dramatic advantages over processing individual spectral bands independently, as shown in Figure 11(d), (h), and (l). The mean squared error associated with this data cube estimate, denoted $\widehat{s}_S$, is $\|s - \widehat{s}_S\|^2/\|s\|^2 = 4.02 \cdot 10^{-3}$. Finally, Figure 12 shows four representative spectra and their estimates, further demonstrating the strength of the proposed method.

4.3.6 DD-CASSI reconstruction simulation results

This method is used to reconstruct a spatio-spectral data cube from simulated measurements modeled after the DD-CASSI system. For this experiment, we generated a phantom spatio-spectral data cube similar to the one used in Section 4.3.5, but with only 15 (instead of 64) spectral bands. The intensities in several spectral bands, the color projection of the data cube, and several representative spectra for this phantom are displayed in the first row of Figure 14. The 256×256 noisy simulated observations are displayed in Figure 13; in this experiment, the mean photon count per pixel in x is $1,534$.

The result of the proposed reconstruction method is displayed in the second row of Figure 14. The reconstructed spectral image computed using the multiscale denoising method described in Section 4.3.5 has an MSE of $3.42 \cdot 10^{-2}$, while the reconstructed spectral image computed without any regularization (*i.e.* the conventional Richardson-Lucy reconstruction) has an MSE of $4.64 \cdot 10^{-2}$ and significantly more artifacts.

4.3.7 Experimental results

To test the DD-CASSI architecture, we constructed a proof-of-concept spectral imaging system as described in Sec. 4.3.1. A photograph of the experimental prototype is shown in Fig. 16. As designed and built, the prototype has an operational range of 520–590 nm. It is an interesting consequence of the system approach that shifting the wavelength by one spectral band should produce the same physical shift on the coding array as a spatial shift by one spatial resolution element. As a result, the system requires a low amount of dispersion. For this prototype we achieved this with a prism-based approach (diffractive approaches are possible, but more challenging as low dispersion tends introduce overlap from higher diffraction orders).

Before the system may be used for spectral imaging, it requires calibration. This calibration is performed by sequentially measuring the response of the system to spatially-uniform monochromatic light at 2 nm steps in the range 520–590 nm. To minimize the noise in the calibration, the system response at a given wavelength is taken to be the average of 10 exposures. Further, the monochromator used to generate the light has a non-uniform output as a function of wavelength. To control for this, a small portion of the light output is fed into a photodiode with a known wavelength response. At each wavelength step, the output of this photodiode is used to adjust the exposure time to keep the total energy propagating through the spectral imager constant. This process is automated via MATLAB. When the resulting system responses are stacked according to wavelength, the result is the system filter function h_{nmp} as defined in (11).

As a first test of the system we consider simple targets illuminated with monochromatic light. We begin with a pingpong ball and illuminations at 532 nm and 543 nm. A representative detector image (taken with 532 nm illumination) is shown in Fig. 17(a). Note the modulation introduced by the coding aperture. The pattern is particularly clear because of the monochro-

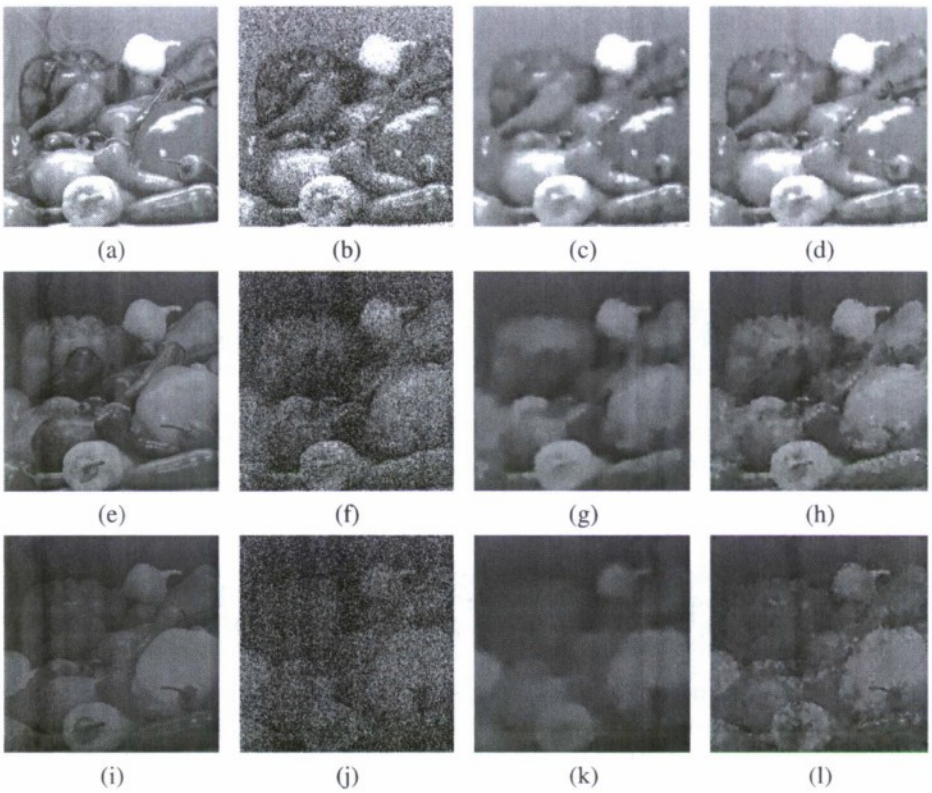


Figure 11: Spatio-spectral denoising results. (a) Spatial variation of intensity in 16th spectral band. (b) Noisy observations, with an average of 7.04 photon count per voxel. (c) Result after using tree-pruning to denoise the image in each spectral band independently; MSE $= 5.22 \cdot 10^{-3}$. (d) Result after proposed reconstruction; MSE $= 4.02 \cdot 10^{-3}$. (e) Spatial variation of intensity in 24th spectral band. (f) Noisy observations. (g) Result after using tree-pruning to denoise the image in each spectral band independently. (h) Result after proposed reconstruction. (i) Spatial variation of intensity in 53rd spectral band. (j) Noisy observations. (k) Result after using tree-pruning to denoise the image in each spectral band independently. (l) Result after proposed reconstruction.

matic source. An estimated intensity image of the scene, generated by summing the spectral channels in the reconstructed datacube is shown in Fig. 17(b). The reconstruction algorithm has largely succeeded in eliminating the modulation introduced by the screen. Figures 17(b) and (c) are the spectral estimates for a particular spatial location in the cases with 532 nm and 543 nm illumination, respectively. The reconstruction accurately locates the spectral peaks.

We next consider the same simple target but now with slightly broader illumination—broadband light filtered by 10 nm FWHM bandpass filters centered at 560 nm and 580 nm, respectively. A representative detector image take with the 560 nm filter is shown in Fig. 18(a). Note that the sharpness of the modulation pattern introduced by the coding aperture has decreased as a result of the increased spectral width. An estimated intensity image of the scene is shown in Fig. 18(b). Figures 18(c) and (d) show the spectral reconstruction at a particular spatial location for the 560 nm and 580 nm bandpass filters, respectively. Again, the reconstruction accurately locates the spectral peaks. However, we note that, for the 580 nm bandpass, a spurious peak now appears near 520 nm. For this bandpass, there is significant source illumination at wavelengths longer than 590 nm. The system is designed for a spectral range of 520–590 nm. Any wavelengths outside this band are aliased back into this range (the illumination produces a coding pattern that is indistinguishable from that generated by a wavelength inside the operational band). The peak at 520 nm, then, is an aliased version of source energy above 590 nm. Any final system would avoid this difficulty by incorporating a bandpass filter matched to the spectral range of the instrument.

Finally, we image a slightly less constrained target under more normal illumination. For this portion of the experiment, we use a collection of citrus fruit under broadband (white) light illumination. The resulting reflection spectrum contains significant energy in the spectral band of interest. Figure 19(a) shows the detector image captured during this experiment. The top fruit is green, the bottom right is yellow-green, and the bottom left is yellow-orange. Note the the broad spectral ranges have made the spectral patterns very indistinct, and that the patterns are subtly different in the three regions as a result of the different spectral components. A reconstructed intensity image is shown in Fig. 19(b). The spectral reconstructions for the three different regions are shown in Fig. 19(c) along with measurements made by a conventional spectrometer. There is good qualitative agreement between the reconstructions and the conventional measurements, especially in the central spectral

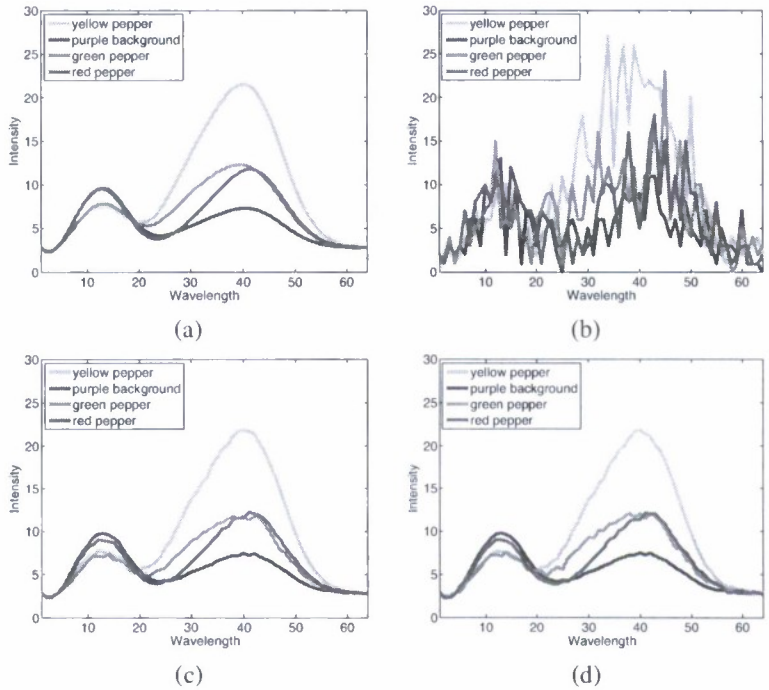


Figure 12: Spatio-spectral denoising results. (a) True spectra at five different spatial locations. (b) Observed spectra at same locations. (c) Estimated spectra at same locations after performing intensity estimation on each spectral band independently, resulting in $\widehat{s}_I$. (d) Estimated spectral at same locations after performing proposed reconstruction, resulting in $\widehat{s}_S$.

regions. We believe that the measurements from the conventional instrument are not fully accurate because of the practical difficulties in coupling the reflected spectra to the probe. However, at least a portion of the deviation arises from inaccurate reconstructions, as we again see some spurious peaks near 520 nm that are the result of spectral aliasing.

Figure 20 shows slices of the reconstructed datacube for 8 specific channels across the spectral range.

4.4 SD-CASSI system

This section reports a new compressive CASSI instrument dubbed the *single disperser coded aperture snapshot spectral imager* (SD-CASSI). Like the DD-CASSI, the SD-CASSI does not directly measure each voxel in the desired three-dimensional datacube. It collects a small number (relative to the size of the datacube) of coded measurements and a sparse reconstruction method is used to estimate the datacube 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 is different from that of the DD-CASSI. Also, a raw measurement of a scene on the detector rarely reveals spatial structure of the scene and makes block processing more challenging.

Since the DD-CASSI only multiplexes spectral information in the datacube, 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 demanding 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, where spectral resolution is more critical than spatial resolution in the datacube, 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 datacube and using a large focal plane array. Essentially, just one spectrally-dispersed projection of the datacube that is spatially modulated by the aperture code over all wavelengths is sufficient to reconstruct the entire spectral datacube.

Figure 13: Simulated noisy observations of peppers spatio-spectral data cube collected by proposed DD-CASSI.

4.4.1 SD-CASSI system model

A schematic of the compressive, snapshot SD-CASSI is shown in Fig. 21. 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 [44] 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 [57]. However, it does help us understand the basic operations implemented by the optical elements.

The spectral density entering the instrument and being relayed to the plane of the coded aperture can be represented as $S_0(x, y; \lambda)$. If we represent the transmission function printed on the coded aperture as $T(x, y)$, the spectral density just after the coded aperture is:

$$S_1(x, y; \lambda) = S_0(x, y; \lambda)\, T(x, y),$$

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

$$\begin{aligned} S_2(x, y; \lambda) &= \iint \delta(x' - [x + \alpha(\lambda - \lambda_c)])\, \delta(y' - y)\, S_1(x', y'; \lambda)\, dx'\, dy' \\ &= \iint \delta(x' - [x + \alpha(\lambda - \lambda_c)])\, \delta(y' - y)\, S_0(x', y'; \lambda)\, T(x', y')\, dx'\, dy' \\ &= S_0(x + \alpha(\lambda - \lambda_c), y; \lambda)\, T(x + \alpha(\lambda - \lambda_c), y), \end{aligned}$$

where the Dirac delta functions describe propagation through unity-magnification imaging optics and a dispersive element with linear dispersion α and center wavelength λ_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:

$$I(x, y) = \int S_0(x + \alpha(\lambda - \lambda_c), y; \lambda)\, T(x + \alpha(\lambda - \lambda_c), y)\, d\lambda.$$

This image is a sum over the wavelength dimension of a mask-modulated and later sheared datacube. Note that this is in contrast to the spatially registered image formed on a DD-CASSI detector as a result of summing over the wavelength dimension of a datacube that is first sheared, mask-modulated and finally unsheared.

Since the detector array is spatially pixelated with pixel size Δ, $I(x, y)$ is sampled across both dimensions on the detector.

The measurements at position (n, m) on the detector can be represented as:

$$\begin{aligned} I_{nm} &= \iint I(x,y)\ \mathrm{rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n\right) dx\, dy \\ &= \iiint S_0(x + \alpha(\lambda - \lambda_c), y; \lambda)\, T(x + \alpha(\lambda - \lambda_c), y) \\ &\quad \times \mathrm{rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n\right) dx\, dy\, d\lambda. \end{aligned} \tag{15}$$

Reconstructions of datacubes from the physical system benefit if the coded aperture feature size is an integer multiple, q, of the size of the detector pixels, Δ. This avoids the need for sub-pixel 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 $q\Delta$ and $T_{n',m'}$ representing an open or closed pinhole at position (n', m') in the pinhole array:

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

With this representation for the aperture pattern, the detector measurements as represented in (15) become:

$$\begin{aligned} I_{nm} &= \sum_{m',n'} T_{n'm'} \iiint \mathrm{rect}\left(\frac{x + \alpha(\lambda - \lambda_c)}{q\Delta} - m', \frac{y}{q\Delta} - n'\right) \mathrm{rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n\right) \\ &\quad \times S_0(x + \alpha(\lambda - \lambda_c), y; \lambda)\, dx\, dy\, d\lambda. \end{aligned}$$

Denoting the source spectral density $S_0(x, y; \lambda)$ in discrete form as s_{ijk} and the aperture code pattern $T(x, y)$ as T_{ij}, the detector measurements in matrix form can be written as:

$$\begin{aligned} I_{nm} &= \sum_k s_{(m+k)nk} T_{(m+k)n} + w_{nm} \\ &= (\boldsymbol{H}\boldsymbol{s})_{nm}, \end{aligned} \tag{16}$$

where $\boldsymbol{H}$ is a linear operator that represents the system forward model.

4.4.2 SD-CASSI reconstruction method

As mentioned previously, the SD-CASSI measures a two-dimensional, spatio-spectral multiplexed projection of the three-dimensional datacube representing the scene. Reconstructing a datacube from this compressed measurement relies on the assumption that the sources in the scene have piecewise smooth spatial structure, making the datacube highly compressible in the wavelet basis. In this section, we describe the reconstruction method used to estimate the datacube from the detector measurements. The datacube can be represented as

$$\boldsymbol{s} = \boldsymbol{W}\,\boldsymbol{\theta},$$

where $\boldsymbol{\theta}$ is a vector composed of the two-dimensional wavelet transform coefficients for each spectral band concatenated to form one vector, and $\boldsymbol{W}$ denotes the inverse two-dimensional wavelet transform applied to each spectral band to form the datacube $\boldsymbol{s}$. In the presense of noise, w, the SD-CASSI detector measurement can be represented as:

$$\begin{aligned} d_{nm} &= I_{nm} + w_{nm} \\ &= (\boldsymbol{H}\,\boldsymbol{W}\,\boldsymbol{\theta})_{nm} + w_{nm}, \end{aligned}$$

where $\boldsymbol{H}$ is a representation of the system forward model described in Section 4.4.1 and w_{nm} is noise.

Because of particularly ill-posed nature of $\boldsymbol{H}$ in the SD-CASSI system, the EM reconstruction method developed for the DD-CASSI system for Poisson observations is ineffective and results in very strong artifacts when applied to the SD-CASSI system. As a result, we apply an alternative reconstruction method based on the assumption of white Gaussian noise to reconstructing the SD-CASSI spectral images. This approach, as demonstrated below, is effective when the number of photons collected is large and the Gaussian assumption is accurate, but is significantly less effective in low light conditions.

If the datacube, $\boldsymbol{s}$, 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, $\boldsymbol{d}$, 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 $\boldsymbol{s}$ and $\boldsymbol{d}$

as column vectors, the linear operator matrix, $\boldsymbol{H}$, can be represented as a matrix of size $\{[p\,(p+q-1)] \times (p^2\,q)\}$. The vector of wavelet coefficients of the datacube, $\boldsymbol{\theta}$, is of size $\{[p^2 q\,(3\log_2(p)+1)] \times 1\}$. The size of this vector reflects the fact that the wavelet decomposition of the datacube 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, $\widehat{\boldsymbol{s}}$, for the datacube can be found by solving the problem:

$$\widehat{\boldsymbol{s}} = \boldsymbol{W}\left[\underset{\boldsymbol{\theta}'}{\operatorname{argmin}}\left\{||\boldsymbol{d} - \boldsymbol{H}\,\boldsymbol{W}\,\boldsymbol{\theta}'||_2^2 + \tau||\boldsymbol{\theta}'||_1\right\}\right].$$

The solution of this nonlinear optimization problem has received significant attention recently [49,58,59]. We use the Gradient Projection for Sparse Reconstruction (GPSR) method developed by Figueiredo et al. [60]. This approach is based upon a variant of the Barzilai-Borwein gradient projection method [61], and has code available online at http://www.lx.it.pt/~mtf/GPSR/. The reconstruction method searches for a datacube estimate with a sparse representation in the wavelet basis; i.e. a $\boldsymbol{\theta}$ which contains mostly zeros and a relatively small number of large coefficients. The first term in this optimization equation minimizes the ℓ_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, τ is a tuning parameter for the penalty term and higher values of τ yield sparser estimates of $\boldsymbol{\theta}$.

4.4.3 Experimental results

To experimentally verify the SD-CASSI spectral imaging concept, we constructed a proof-of-concept prototype as shown in Fig. 22. The prototype consisted of (i) a coded aperture, lithographically patterned on a chrome-on-quartz mask, (ii) three lenses from Schneider Optics Inc. with an f/# of 1.4 and a focal length of 22.5 mm, (iii) an equilateral prism from Edmund Optics Inc. as a dispersive element, (iv) a monochrome charge-coupled device (CCD) detector from Photometrics with 1040×1392 pixels that are $4.65\,\mu$m square each, 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 [7], 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 report. Although we have developed the theory for and conducted experimental studies of optimal aperture codes for coded aperture spectroscopy [44, 62], 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.

An equilateral prism was used instead of a grating because the grating produces overlapping diffractive orders while the prism only refracts the wavelengths into one order. Prisms also have large transmission efficiencies [63]. 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 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.

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 displaying the ability of the SD-CASSI to reconstruct relatively simple spatio-spectral 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 datacube representing the scene, the reconstruction algorithm requires two inputs – the aperture code used by the instrument to encode the datacube and the detector measurement of the datacube. 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 23 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 datacube than the code printed on the aperture mask, and thus produces a more accurate reconstruction of the datacube.

We constructed a scene consisting of two ping pong balls as shown in Fig. 24. One ping pong ball was illuminated with a 543 nm laser and a white light source filtered by a green 560 nm narrow band filter. The other was a red ping pong ball illuminated with a white light source. Figure 25 shows a CCD measurement of the scene by the SD-CASSI. Given the low linear dispersion of the prism, there is spatio-spectral 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 datacube corresponding 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 two in the row and column dimensions prior to performing a reconstruction of the datacube. This downsampling helped reduce the time needed by the reconstruction algorithm to generate an estimate of the datacube. The resulting $\{128 \times 128 \times 28\}$ datacube spanned a spectral range of 540 nm to 640 nm. One hundred iterations of the reconstruction algorithm required about fourteen minutes of runtime on a desktop machine. The GPSR method was run with $\tau = 0.05$. This value was determined via trial and error. Future work is needed to study how to choose the optimal value of τ.

Figure 26 shows the spatial content of each of 28 wavelength channels between 540 nm to 640 nm. Note that the mask modulation on the spatial structure visible in Fig. 25 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 spectrometer. Figure 27 (a) shows the SD-CASSI spectrum at a point on the green ping pong ball, while Fig. 27 (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 due to the non-linear dispersion of the prism across the detector. This calibration was performed by tracking the position of a point source while varying its wavelength. Figures 27 (a) and (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 datacube 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 datacube could be improved if the reconstruction algorithm utilizes 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, then the spectral resolution of the SD-CASSI is determined by the width of the smallest code feature. Consider the case where the smallest code feature maps to two detector pixels. Then imaging two adjacent, monochromatic point sources of close but distinct wavelengths on to the smallest code feature can potentially result in the spatio-spectral 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 nm) across two detector pixels. Since the dispersion of a prism is non-linear, 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 nm 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 per pixel. Since the width of the smallest code feature was 4 detector pixels, the average spectral resolution of the instrument was 3.6 nm/spectral channel.

4.5 Conclusion and future work

A proof-of-concept prototype was constructed and tested on both highly-constrained and real-world sources. The reconstructions accurately captured the spectral features of the source (barring a spectral aliasing that can be eliminated through the incorporation of a bandpass filter on future systems). The spatial structure of the reconstructions was also accurate. The modulation introduced by the coding aperture was successfully removed by the reconstruction (especially for the broadband, real-world scene).

The result was a single-shot spectral imager that, for the first time, mitigated the trade-offs between spatial resolution, spectral resolution, light collection, and measurement acquisition time. While the performance of the system was quite acceptable for a first proof-of-concept, we feel that the results were limited by the stock optics used to create the present prototype. We are currently constructing a new version, with dramatically enhanced spatial and spectral resolution.

5 Publications supported by this grant

R. G. Baraniuk and M. B. Wakin, "Random projections of smooth manifolds," *Foundations of Computational Mathematics*, To appear.

P. Boufounos and R. G. Baraniuk, "1-bit compressive sensing.," in *CISS*, pp. 16–21, IEEE, 2008.

P. Boufounos and R. Baraniuk, "Quantization of sparse representations," *Data Compression Conference*, vol. 0, p. 378, 2007.

P. Boufounos, M. F. Duarte, and R. G. Baraniuk, "Sparse signal reconstruction from noisy compressive measurements using cross validation," in *Proc. IEEE Statistical Signal Processing Workshop (SSP)*, 2007.

V. Cevher, M. F. Duarte, and R. G. Baraniuk, "Distributed target localization via spatial sparsity," in *European Signal Processing Conference (EUSIPCO)*, 2008.

M. A. Davenport, M. F. Duarte, M. B. Wakin, J. N. Laska, D. Takhar, K. F. Kelly, and R. G. Baraniuk, "The smashed filter for compressive classification and target recognition," in *Proc. IS&T/SPIE Symposium on Electronic Imaging: Computational Imaging*, vol. 6498, (San Jose, CA), Jan. 2007.

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 Proc. Mag.*, vol. 25, pp. 83–91, March 2008.

M. F. Duarte, M. B. Wakin, and R. G. Baraniuk, "Wavelet-domain compressive signal reconstruction using a hidden markov tree model," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2008.

R. F. Marcia, C. Kim, C. Eldeniz, J. Kim, D. J. Brady, and R. M. Willett, "Superimposed video disambiguation for increased field of view," *Opt. Express*, vol. 16, no. 21, pp. 16352–16363, 2008.

R. Marcia and R. Willett, "Compressive coded aperture video reconstruction," in *European Signal Processing Conference (EUSIPCO)*, 2008.

R. Marcia and R. Willett, "Compressive coded aperture superresolution image reconstruction," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2008.

S. Sarvotham, D. Baron, and R. Baraniuk, "Measurements vs. bits: Compressed sensing meets information theory," in *Allerton Conference on Communication, Control, and Computing*, 2006.

R. M. Willet, M. E. Gehm, and D. J. Brady, "Multiscale reconstruction for computational spectral imaging," in *Proc. SPIE*, vol. 6498, 2007.

6 Professional personnel

Principle Investigators
Richard G. Baraniuk, David Brady, Kevin Kelly, and Rebecca Willett

Graduate Student Research Assistants
Mark Davenport, Jason Laska, Matthew Moravec, Shriram Sarvotham, Mona Sheikh, Ting Sun, and Dharmpal Takhar

References

[1] W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Data Compression Standard.* Kluwer Academic Publishers, 1992.

[2] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The jpeg 2000 still image compression standard," *IEEE Signal Processing Magazine*, vol. 18, pp. 36–58, 2001.

[3] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice.* Boston: Kluwer, 2002.

[4] A. Puri and A. Eleftheriadis, "Mpeg-4: an object-based multimedia coding standard supporting mobile applications," *Mob. Netw. Appl.*, vol. 3, no. 1, pp. 5–32, 1998.

[5] S. R. Gottesman and E. E. Fenimore, "New family of binary arrays for coded aperture imaging," *Appl. Opt.*, vol. 28, 1989.

[6] M. J. Golay, "Static multislit spectrometry and its application to the panoramic display of infrared spectra," *Journal of the Optical Society of America*, vol. 41, no. 7, pp. 468–472, 1951.

[7] M. Harwit and N. J. A. Sloan, *Hadamard transform optics*. New York: Academic, 1979.

[8] E.P. Wagner II, B. W. Smith, S. Madden, J. Winefordner, and M. Mignardi, "Construction and evaluation of a visible spectrometer using digital micromirror spatial light modulation," *Applied Spectroscopy*, vol. 49, p. 1715, 1995.

[9] R. DeVerse, R. Hammaker, and W. Fateley, "Realization of the hadamard multiplex advantage using a programmable optical mask in a dispersive flat-field near-infrared spectrometer," *Applied Spectroscopy*, vol. 54, no. 12, pp. 1751–1758, 2000.

[10] D. J. Brady, M. Feldman, N. Pitsianis, J. P. Guo, A. Portnoy, and M. Fiddy, "Compressive optical montage photography," in *SPIE: Photonic Devices and Algorithms for Computing VII*, vol. 5907, p. 590708, 2005.

[11] Y. W. N. P. P. M. E. S. S. T. McCain, M. E. Gehm and D. J. Brady, "Multimodal, multiplex, raman spectroscopy of alcohol in diffuse, fluorescent media," in *SPIE: Novel Optical Instrumentation for Biomedical Applications II*, vol. 5864, p. 58640L, 2005.

[12] E. J. Candès, "Compressive sampling," in *Proc. International Congress of Mathematicians*, vol. 3, (Madrid, Spain), pp. 1433–1452, 2006.

[13] D. Donoho, "Compressed sensing," *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[14] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Trans. on Information Theory*, vol. 52, no. 9, pp. 4036–4048, 2006.

[15] J. Tropp and A. C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," *IEEE Trans. Info. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[16] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by weighted ℓ_1 minimization," tech. rep., California Institute of Technology, Pasadena, CA, Oct. 2007.

[17] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA, USA: Academic Press, 1999.

[18] M. F. Duarte, M. B. Wakin, and R. G. Baraniuk, "Fast reconstruction of piecewise smooth signals from random projections," in *Proc. SPARS05*, (Rennes, France), Nov. 2005.

[19] C. La and M. N. Do, "Signal reconstruction using sparse tree representation," in *Proc. Wavelets XI at SPIE Optics and Photonics*, (San Diego, CA), Aug. 2005.

[20] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using Hidden Markov Models," *IEEE Trans. Signal Processing*, vol. 46, pp. 886–902, Apr. 1998.

[21] J. K. Romberg, H. Choi, and R. G. Baraniuk, "Bayesian tree-structured image modeling using wavelet-domain Hidden Markov Models," *IEEE Trans. Image Processing*, vol. 10, pp. 1056–1068, July 2001.

[22] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Trans. Image Processing*, vol. 11, no. 2, pp. 146–158, 2002.

[23] M. J. Wainwright and E. P. Simoncelli, "Scale mixtures of Gaussians and the statistics of natural images," in *Neural Information Processing Systems (NIPS)* (S. A. Solla, T. K. Leen, and K.-R. Müller, eds.), vol. 12, (Cambridge, MA), pp. 855–861, MIT Press, Dec. 2000.

[24] J. Kivinen, E. Sudderth, and M. Jordan, "Image denoising with nonparametric Hidden Markov Trees," in *IEEE International Conference on Image Processing (ICIP)*, (San Antonio, TX), Sept. 2007.

[25] R. G. Baraniuk, D. H. Johnson, K. K. Kelly, and Y. Massoud, "A new camera architecture based on optical-domain compression," *preprint*, 2005.

[26] R. Coifman, F. Geshwind, and Y. Meyer, "Noiselets," *Appl. Comp. Harmonic Analysis*, vol. 10, pp. 27–44, 2001.

[27] E. J. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, pp. 969–985, 2007.

[28] D. Brady, "Multiplex sensors and the constant radiance theorem," *Optics Letters*, vol. 27, no. 1, pp. 16–18, 2002.

[29] E. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Info. Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[30] R. A. DeVerse, R. R. Coifman, A. C. Coppi, W. G. Fateley, F. Geshwind, R. M. Hammaker, S. Valenti, and F. J. Warner, "Application of spatial light modulators for new modalities in spectrometry and imaging," in *Proc. SPIE*, vol. 4959, pp. 12–22, 2003. http://www.math.yale.edu/ mmm82/hyperspectral.html.

[31] J. Haupt and R. Nowak, "Compressive sampling versus conventional imaging," in *IEEE Int. Conf. Image Proc.*, (Atlanta, GA), pp. 1269–1272, Oct. 2006.

[32] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2325–2383, Oct. 1998.

[33] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, pp. 1207–1223, Aug. 2006.

[34] R. Constantini and S. Susstrunk, "Virtual sensor design," in *Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications V*, vol. 5301, pp. 408–419, Jun. 2004.

[35] P. A. Townsend, J. R. Foster, J. R. A. Chastain, and W. S. Currie, "Application of imaging spectroscopy to mapping canopy nitrogen in the forests of the central appalachian mountains using hyperion and aviris," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 41, no. 6, pp. 1347–1354, 2003.

[36] 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," in *Proc. SPIE*, vol. 4151, pp. 94 – 102, 2001.

[37] R. P. Lin, B. R. Dennis, and A. O. B. (Eds.), *The Reuven Ramaty High-Energy Solar Spectrscopic Imager (RHESSI) - Mission Description and Early Results*. Dordrecht: Kluwer Academic Publishers, 2003.

[38] W. Denk, J. Strickler, and W. Webb, "Two-photon laser scanning fluorescence microscopy," *Science*, vol. 248, no. 4951, pp. 73–76, 1990.

[39] R. Schultz, T. Nielsen, J. Zavaleta, R. Ruch, R. Wyatt, and H. Garner, "Hyperspectral imaging: A novel approach for microscopic analysis," *Cytometry*, vol. 43, pp. 239 – 247, 2001.

[40] M. Hinnrichs, J. Jensen, and G. McAnally, "Handheld hyperspectral imager for standoff detection of chemical and biological aerosols," *Proc. SPIE*, vol. 5268, pp. 67–78, 2003.

[41] J. Mooney, V. Vickers, M. An, and A. Brodzik, "High-throughput hyperspectral infrared camera," *J. Opt. Soc. Am. A*, vol. 14, no. 11, pp. 2951–2961, 1997.

[42] M. Descour, C. Volin, E. Dereniak, T. Gleeson, M. Hopkins, D. Wilson, and P. Maker, "Demonstration of a computed-tomography imaging spectrometer using a computer-generated hologram disperser," *image*, vol. 1501, p. 1, 1980.

[43] M. E. Gehm and D. J. Brady, "High-throughput hyperspectral microscopy," *Proc. SPIE*, vol. 6090, pp. 13–21, 2006.

[44] M. Gehm, S. McCain, N. Pitsianis, D. Brady, P. Potuluri, and M. Sullivan, "Static two-dimensional aperture coding for multimodal, multiplex spectroscopy," *Appl. Opt.*, vol. 45, pp. 2965–2974, 2006.

[45] S. McCain, M. Gehm, Y. Wang, N. Pitsianis, and D. Brady, "Coded aperture raman spectroscopy for quantitative measurements of ethanol in a tissue phantom," *Appl. Spectrosc.*, vol. 60, pp. 663–671, 2006.

[46] E. Cull, M. Gehm, D. Brady, C. Hsieh, O. Momtahan, and A. Adibi, "Dispersion multiplexing with broadband filtering for miniature spectrometers," *Appl. Opt.*, vol. 46, pp. 365–374, 2007.

[47] C. Fernandez, B. Guenther, M. Gehm, D. Brady, and M. Sullivan, "Longwave infrared (lwir) coded aperture dispersive spectrometer," *Opt. Express*, vol. 15, pp. 5742–5753, 2007.

[48] M. E. Gehm, D. J. Brady, N. Pitsianis, and X. Sun, "Compressive sampling strategies for integrated microspectrometers," in *Proc. SPIE*, vol. 6232, 2006.

[49] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489 – 509, 2006.

[50] W. Richardson, "Bayesian-based iterative method of image restoration," *J. Opt. Soc. of Am.*, vol. 62, pp. 55–59, 1972.

[51] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *Astron. J.*, vol. 79, pp. 745–754, 1974.

[52] R. Nowak and E. Kolaczyk, "A multiscale statistical framework for Poisson inverse problems," *IEEE Trans. Info. Theory*, vol. 46, pp. 1811–1825, 2000.

[53] R. Willett, "Multiscale intensity estimation for marked poisson processes," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2007.

[54] E. Kolaczyk and R. Nowak, "Multiscale likelihood analysis and complexity penalized estimation," *Annals of Stat.*, vol. 32, pp. 500–527, 2004.

[55] R. Willett and R. Nowak, "Fast multiresolution photon-limited image reconstruction," in *Proc. IEEE Int. Sym. Biomedical Imaging — ISBI '04*, (15-18 April, Arlington, VA, USA), 2004.

[56] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, "Noise reduction using an undecimated discrete wavelet transform," *IEEE Signal Processing Letters*, vol. 3, no. 1, pp. 10–12, 1996.

[57] D. Schroeder, *Astronomical Optics*. Academic Press, 1987.

[58] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61 (electronic), 1998.

[59] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.

[60] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems." To appear in *IEEE Journal of Selected Topics in Signal Processing: Special Issue on Convex Optimization Methods for Signal Processing*, 2007.

[61] Y.-H. Dai and R. Fletcher, "Projected barzilai-borwein methods for large-scale box-constrained quadratic programming," *Numer. Math.*, vol. 100, pp. 21–47, 2005.

[62] A. A. Wagadarikar, M. E. Gehm, and D. J. Brady, "Performance comparison of aperture codes for multimodal, multiplex spectroscopy," *Applied Optics*, vol. 46, no. 22, pp. 4932–4942, 2007.

[63] J. M. Lerner, "Imaging spectrometer fundamentals for researchers in the biosciences - a tutorial," *Cytometry Part A*, vol. 69, no. 8, pp. 712–734, 2006.

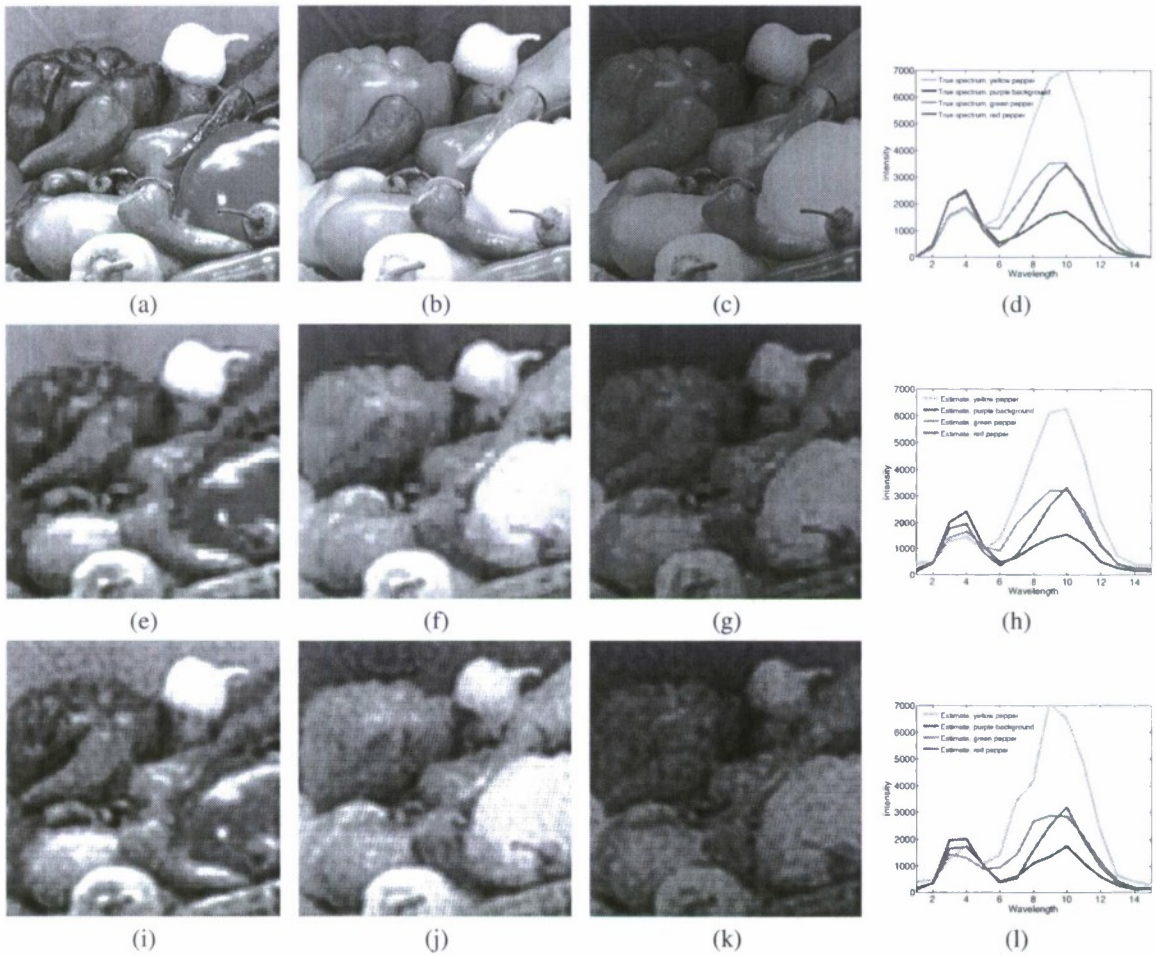


Figure 14: DD-CASSI reconstruction simulation results. (a) True intensity in 4th spectral band. (b) True intensity in 8th spectral band. (c) True intensity in 12th spectral band. (d) Representative spectra from true spatio-spectral data cube. (e) Estimated intensity in 4th spectral band, computed using reconstruction method described above with multiscale spatio-spectral regularization; $MSE = 3.42 \cdot 10^{-2}$. (f) Estimated intensity in 8th spectral band, computed using reconstruction method described above with multiscale spatio-spectral regularization. (g) Estimated intensity in 12th spectral band, computed using reconstruction method described above with multiscale spatio-spectral regularization. (h) Representative spectra from estimated spatio-spectral data cube, computed using reconstruction method described above with multiscale spatio-spectral regularization. (i) Estimated intensity in 4th spectral band, computed using reconstruction method described above with no regularization; $MSE = 4.64 \cdot 10^{e-2}$. (j) Estimated intensity in 8th spectral band, computed using reconstruction method described above with no regularization. (k) Estimated intensity in 12th spectral band, computed using reconstruction method described above with no regularization. (l) Representative spectra from estimated spatio-spectral data cube, computed using reconstruction method described above with no regularization.

(a) (b) (c)

Figure 15: DD-CASSI reconstruction simulation results, color projections. (a) Color projection of true spatio-spectral data cube. (b) Color projection of estimated spatio-spectral data cube, computed using reconstruction method described above with multiscale spatio-spectral regularization. (c) Color projection of estimated spatio-spectral data cube, computed using reconstruction method described above with no regularization.

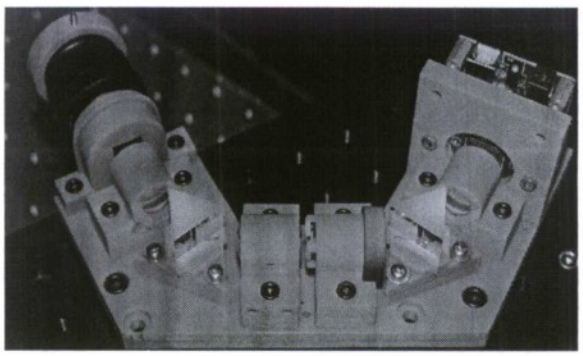


Figure 16: The DD-CASSI experimental prototype.

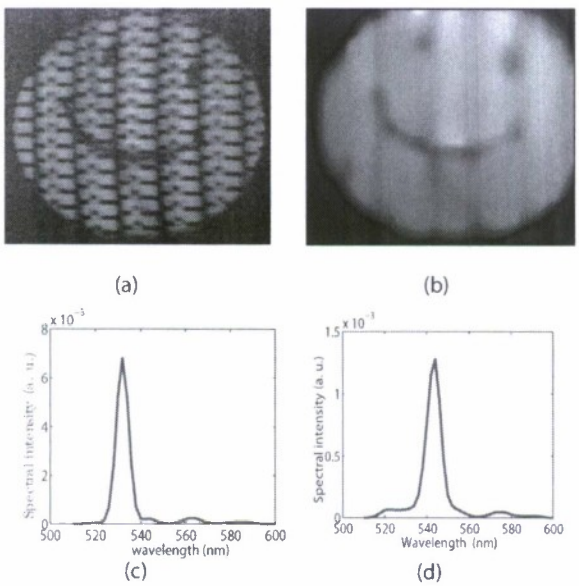


Figure 17: DD-CASSI experimental results from simple targets with monochromatic illumination. (a) Detector image recorded for 532 nm illumination. (b) Intensity image generated by summing the spectral information in the reconstruction for 532 nm illumination. (c) Spectral reconstruction at a particular spatial location for 532 nm illumination. (d) Spectral reconstruction at a particular spatial location for 543 nm illumination.

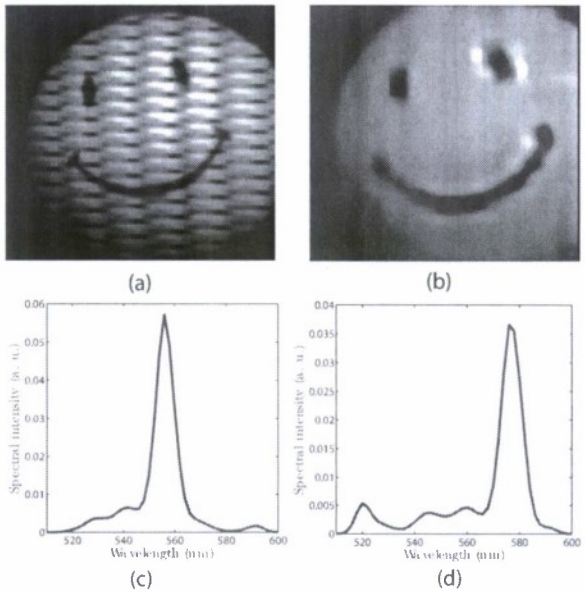


Figure 18: DD-CASSI experimental results from simple targets with narrow-band illumination. (a) Detector image recorded for illumination with a 10 nm FWHM bandpass centered at 560 nm. (b) Intensity image generated by summing the spectral information in the reconstruction for the 560 nm bandpass. (c) Spectral reconstruction at a particular spatial location for the 560 nm bandpass. (d) Spectral reconstruction at a particular spatial location for the 580 nm bandpass. The origin of the small peak near 520 nm is explained in the text.

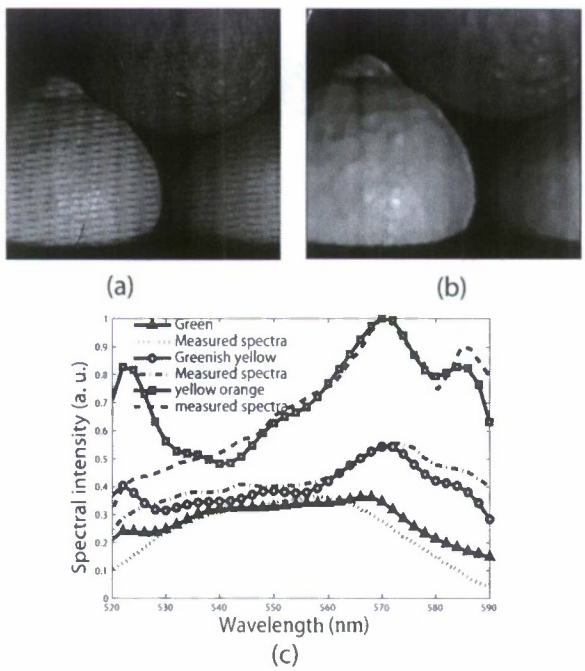


Figure 19: DD-CASSI experimental results from real-world objects under broadband (white) illumination. (a) Detector image recorded by the system. (b) Reconstructed intensity image of the scene. (c) Spectral reconstructions for spatial locations in the three regions.

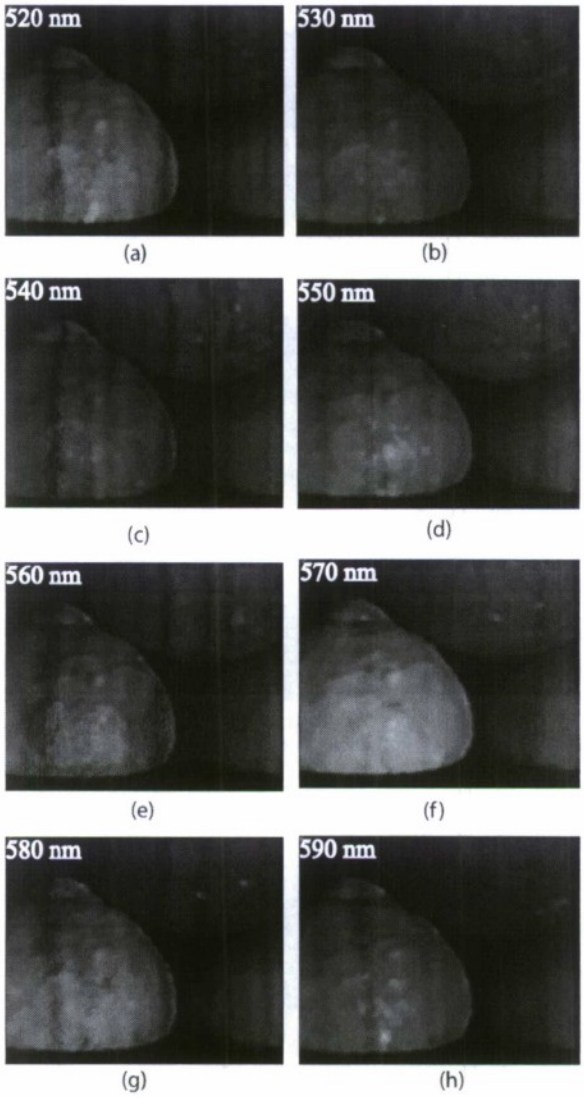


Figure 20: Slices through the reconstructed datacube at 8 particular spectral channels.

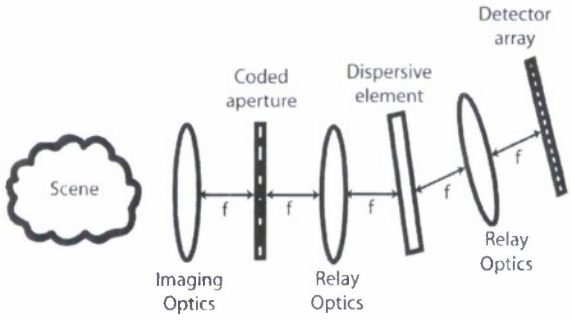


Figure 21: Schematic of an SD-CASSI. The imaging optics image the scene on to the coded aperture. The relay optics relay the image from the plane of the coded aperture to the detector through the dispersive element.

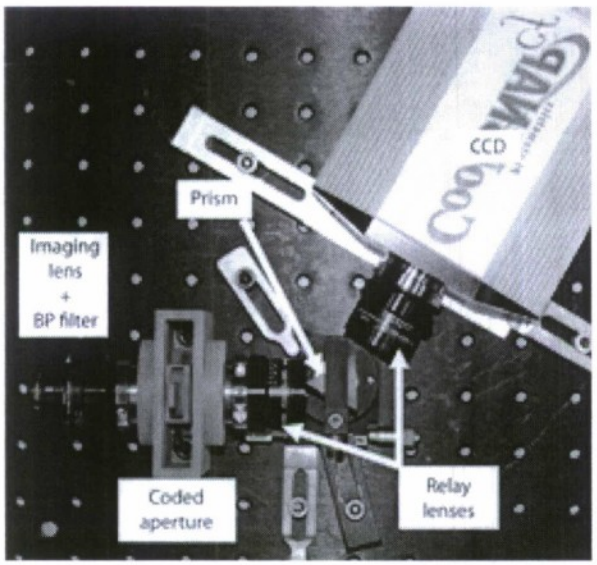


Figure 22: Experimental prototype of the SD-CASSI. The prototype consists of (i) a coded aperture, lithographically patterned on a chrome-on-quartz mask, (ii) three lenses from Schneider Optics Inc. with an f/# of 1.4 and a focal length of 22.5 mm, (iii) an equilateral prism from Edmund Optics Inc. as a dispersive element, (iv) a monochrome charge-coupled device (CCD) detector from Photometrics with 1040×1392 pixels that are $4.65\ \mu$m square each, and (v) a $500 - 620$ nm bandpass filter that is placed in front of the imaging lens to remove the impact of stray light on the experimental measurements.

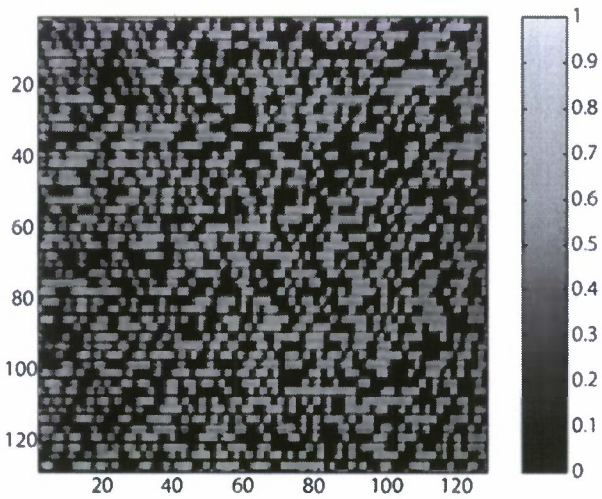


Figure 23: Aperture code pattern used by the SD-CASSI reconstruction algorithm to generate an estimate of the datacube.

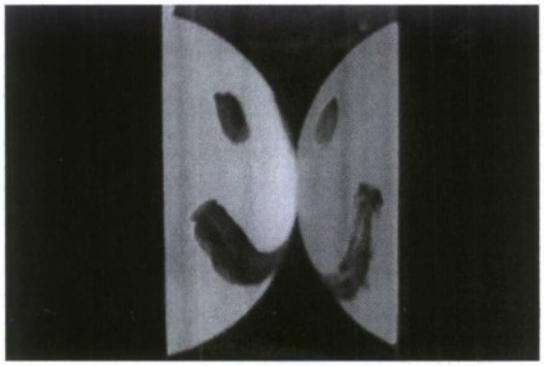


Figure 24: A scene consisting of a ping pong ball illuminated with a 543 nm green laser and a white light source filtered by a 560 nm narrow band filter (left), and a red ping pong ball illuminated with a white light source (right).

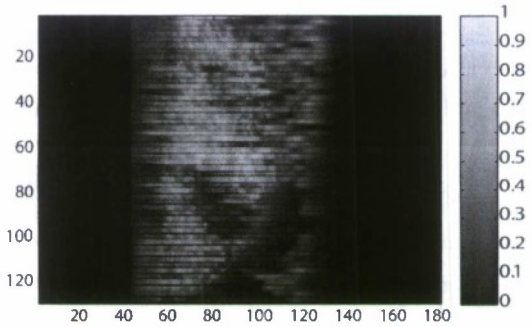


Figure 25: SD-CASSI detector measurement of the scene consisting of the two ping pong balls. Given the low linear dispersion of the prism, there is spatio-spectral overlap of the aperture code-modulated images of each ball.

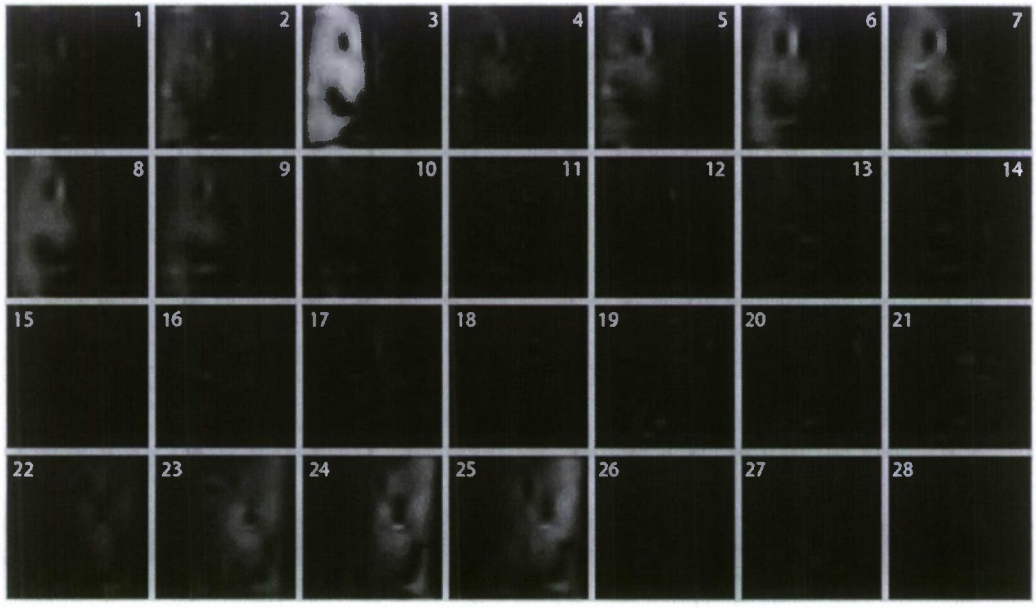


Figure 26: SD-CASSI estimated spatial content of the scene in each of 28 spectral channels between 540 nm and 640 nm. The green ball can be seen in channels 3, 4, 5, 6, 7 and 8, while the red ball can be seen in channels 23, 24 and 25.

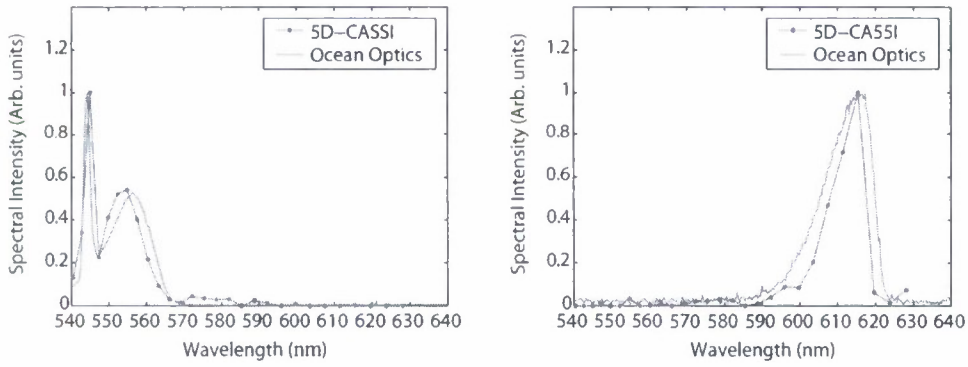


Figure 27: (a) SD-CASSI spectral intensity through a point on the ping pong ball illuminated with a 543 nm green laser and a white light source filtered by a 560 nm narrow band filter. (b) SD-CASSI spectral intensity through a point on the red ping pong ball illuminated with a white light source. Spectra from an Ocean Optics non-imaging reference spectrometer are shown for comparison.