

# Video rate spectral imaging using a coded aperture snapshot spectral imager

Ashwin A. Wagadarikar<sup>a</sup>, Nikos P. Pitsianis<sup>abc</sup>, Xiaobai Sun<sup>b</sup>, David J. Brady<sup>a</sup>

<sup>a</sup>Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708

<sup>b</sup>Department of Computer Science, Duke University, Durham, NC 27708, USA

<sup>c</sup>Department of Electrical and Computer Engineering, Aristotle University, Thessaloniki, 54124, Greece

[dbrady@duke.edu](mailto:dbrady@duke.edu)

<http://www DISP.duke.edu>

**Abstract:** We have previously reported on coded aperture snapshot spectral imagers (CASSI) that can capture a full frame spectral image in a snapshot. Here we describe the use of CASSI for spectral imaging of a dynamic scene at video rate. We describe significant advances in the design of the optical system, system calibration procedures and reconstruction method. The new optical system uses a double Amici prism to achieve an in-line, direct view configuration, resulting in a substantial improvement in image quality. We describe NeAREst, an algorithm for estimating the instantaneous three-dimensional spatio-spectral data cube from CASSI's two-dimensional array of encoded and compressed measurements. We utilize CASSI's snapshot ability to demonstrate a spectral image video of multi-colored candles with live flames captured at 30 frames per second.

© 2009 Optical Society of America

**OCIS codes:** (300.6190) Spectrometers, (110.0110) Imaging systems, (120.6200) Spectrometers and spectroscopic instrumentation, (999.9999) Coded aperture spectral imaging

---

## References and links

1. J. Mooney, V. Vickers, and A. Brodzik, "High throughput hyperspectral infrared camera," *J. Opt. Soc. Am. A* **14**, 2951–2961 (1997).
2. C. Volin, B. Ford, M. Descour, J. Garcia, D. Wilson, P. Maker, and G. Bearman, "High-speed spectral imager for imaging transient fluorescent phenomena," *Applied Optics* **37**(34), 8112–8119 (1998).
3. K. Hege, D. O'Connell, W. Johnson, S. Basty, and E. Dereniak, "Hyperspectral imaging for astronomy and space surveillance," vol. 5159 of *Proc. SPIE*, pp. 380–391 (2003).
4. W. Johnson, D. Wilson, W. Fink, M. Humayun, and G. Bearman, "Snapshot hyperspectral imaging in ophthalmology," *Journal of Biomedical Optics* **12**(1), 0140,361–0140,367 (2007).
5. M. Descour, C. Volin, E. Dereniak, K. Thorne, A. Schumacher, D. Wilson, and P. Maker, "Demonstration of a high-speed non-scanning imaging spectrometer," *Optics Letters* **22**(16), 1271–1273 (1997).
6. N. Gat, G. Scriven, J. Garman, M. D. Li, and J. Zhang, "Development of four-dimensional imaging spectrometers (4D-IS)," vol. 6302 of *Proc. SPIE* (2006).
7. A. Wagadarikar, R. John, R. Willett, and D. J. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Applied Optics* **47**(10), B44–B51 (2008).
8. M. Gehm, R. John, D. J. Brady, R. Willett, and T. Schulz, "Single-shot compressive spectral imaging with a dual-disperser architecture," *Optics Express* **15**(21), 14,013–14,027 (2007).

9. A. Wagadarikar, N. Pitsianis, X. Sun, and D. Brady, "Spectral Image Estimation for Coded Aperture Snapshot Spectral Imagers," vol. 7076 of *Proc. SPIE* (2008).
  10. X. Sun and N. Pitsianis, "Solving non-negative linear inverse problems with the NeAREst method," vol. 7074 of *Proc. SPIE* (2008).
- 

## 1. Introduction

A spectral imager measures the intensity of light as a function of wavelength  $\lambda$  at each spatial location  $(x, y)$  in an image. Thus, it acquires a three-dimensional (3D) data cube of spatio-spectral information,  $(x, y, \lambda)$ , about the scene being imaged. The knowledge of the spectrum at various points in the image can be useful in identifying the composition and structure of objects in the scene being observed by the imager.

Most spectral imagers require some form of temporal scanning to obtain a spatio-spectral data cube. A dispersive push-broom slit spectrometer measures the spectrum at each point in the scene that is imaged onto a slit and relies on the spatial motion of either the slit or the object to measure the complete data cube. A spectrally filtered imager takes multiple images of the same scene consecutively with different discrete wavelength filters. A Fourier-transform spectral imager requires the scanning of the optical path difference between two arms of a Michelson interferometer. Tomographic spectral imagers [1] gather a sequence of two-dimensional (2D) spatio-spectral projections of the data cube through a direct view prism on to a detector array. While these techniques are adequate for spectral imaging of static scenes, the temporal scanning can result in spatial and/or spectral artifacts in imaging of dynamic scenes.

A snapshot spectral imager has a distinct advantage in spectral imaging of fast changing phenomena. Applications of a snapshot spectral imager include observation of fluorescent probes that track fast (0.01-1 s) cellular metabolic processes under a microscope [2], surveillance of fast moving space objects [3], and snapshot spectral imaging of the retina that is unaffected by patient motion [4]. Examples of existing snapshot spectral imagers include the computed tomography imaging spectrometer (CTIS) [5], the four-dimensional imaging spectrometer (4D-IS) [6], and coded aperture snapshot spectral imagers (CASSI) [7, 8]. CTIS uses a computer generated hologram to capture at once multiple spectrally dispersed 2D projections of the data cube. However, it requires a large detector area to capture the multiple projections and fails to sample a conical volume in the Fourier domain representation of the data cube. The 4D-IS uses reformatter fiber optics to map a 2D image of the scene on to the input slit of a dispersive slit spectrometer, which has very high spectral resolution. However, the length of the slit limits the number of fibers that can be imaged and the size of the fibers limits the spatial resolution of the data cube. CASSI systems utilize a coded aperture and one or more dispersive elements to modulate the optical field from a scene. A 2D detector array captures a single coded 2D projection of the full 3D data cube. This array of measurements is then digitally processed using a numerical estimation algorithm to recover an estimate of the 3D data cube.

In this paper, we demonstrate video-rate spectral imaging using a direct view CASSI. The previous single disperser CASSI system [7] consisted of an objective lens, a coded aperture and a pair of lenses to relay the image from the plane of the coded aperture to a 2D detector array through an equilateral prism. In a snapshot, the measurements at the detector pixels were a coded mixture of spatial and spectral information of the objects in the scene. The system had a limited of view, as it was put together with off-the-shelf components. It also suffered from an anamorphic distortion due to the use of the equilateral prism as the dispersive element. This distortion is a stretch of the image along the dispersion direction as the result of different path lengths between opposite ends of the spectral range. Finally, the numerical estimation algorithm required the specification of a regularization parameter to constrain the spatial content of each spectral slice in the estimate of the data cube to be sparse in the redundant wavelet transform

domain.

This paper reports on a direct view CASSI design. The design uses only one relay lens, thus requiring one less optical element than the previous one. Furthermore, it uses a double Amici prism to provide a direct view configuration that effectively removes the anamorphic distortion. We demonstrate the prototype's superior image quality and its snapshot capability through spectral imaging of a dynamic scene being observed and recorded at video rate (30 frames per second). In addition, a rigorous calibration strategy is presented to improve the quality of spectral image estimates. Finally, a fast spectral image estimation algorithm called NeAREst is presented that does not require the specification of any regularization parameter.

The rest of the paper is organized as follows. In the next section, we describe the optical design of the direct view CASSI system. In section 3, we describe the discretization of a simple mathematical model for light propagation through the system. In section 4, we describe the process of calibrating CASSI to provide the image estimation algorithm with a system-specific model that accounts for additional factors that are absent in the simplified light propagation model. In section 5, we describe the spectral image estimation algorithm used to generate a data cube from CASSI detector measurements. In section 6, we demonstrate a spectral video captured by the CASSI and comment on the result. Section 7 concludes the paper.

## 2. Description of the direct view CASSI system

Figure 1 shows a schematic of the direct view CASSI design. The system consists of (1) an objective lens, (2) a coded aperture, (3) a bandpass filter, (4) an F/8 relay lens, (5) a double Amici prism as a dispersive element, and (6) a monochrome charge-coupled device (CCD) detector.

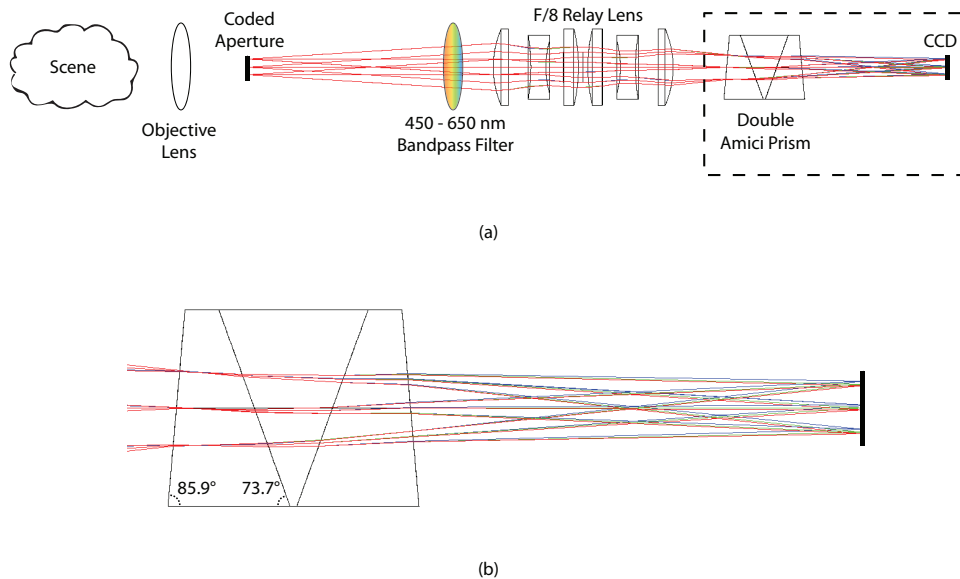


Fig. 1. (a) Schematic of the direct view CASSI. (b) Ray bundles passing through the double Amici prism and being dispersed differently at different wavelengths.

The objective lens images the scene on to the plane of the coded aperture. The coded aperture is lithographically patterned as a chrome coating on a quartz substrate, with an anti-reflective

coating on both sides designed for the 400 – 700 nm range. The code used in all the experiments is a random  $256 \times 248$  element binary pattern as shown in Figure 2, with the smallest code feature being  $2 \times 2$  CCD pixels ( $19.8 \mu\text{m}$  on each side). This limits the spatial size of the reconstructed data cube to a maximum of  $256 \times 248$  spatial elements. The bandpass filter

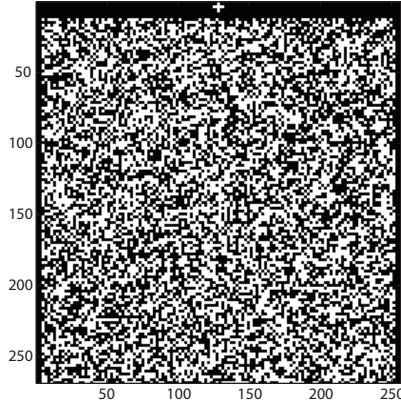


Fig. 2. The aperture code: a random  $256 \times 248$  binary element pattern, with the smallest element being  $2 \times 2$  CCD pixels ( $19.8 \mu\text{m}$  on each side). The cross on the top helps to align the aperture code with the detector pixels during the calibration process.

(Omega Optical) limits the spectral range of the system to 450 – 650 nm. A relay lens (Edmund Optics stock part C45762) relays the image from the plane of the coded aperture to the CCD. Spectral dispersion is introduced using a custom designed (Shanghai Optics Inc.) double Amici prism. This prism consists of three separate prisms cemented together. The first and third prisms are made of SK2, a medium dispersion crown glass (low-medium  $n_d$  and  $\nu_d$  on the glass map), while the second prism is made of SF4, a higher dispersion flint glass (high  $n_d$  and  $\nu_d$  on the glass map). Based on the indices of refraction of each prism, Snell's law can be used to choose angles of these prisms to ensure that rays corresponding to the center wavelength of 550 nm pass through the prism undeviated. Rays corresponding to adjacent wavelengths disperse to either side of the optical axis, resulting in a direct view prism. This configuration is useful because all the components of the system can be placed in a line, which makes system alignment much easier than that of the previous prototype. Furthermore, wavelength dependent anamorphic distortion is minimized because the path lengths of all wavelengths from the prism to the CCD are similar. Figure 1(b) shows an on-axis ray bundle passing through the custom designed double Amici prism. The CCD detector (AVT Marlin, Allied Vision Technologies) is an 8-bit camera with  $9.9 \times 9.9 (\mu\text{m})^2$  pixels and has its strongest response at 500 nm, with a relative response greater than 0.7 between 450 – 650 nm.

All of these system components are mounted on a metal rod/cage system, which serves as the backbone of the system in front of the CCD. The coded aperture and the double Amici prism are placed inside rotation mounts that can be rotated to ensure proper horizontal alignment of the image of the coded aperture and to ensure horizontal dispersion across the CCD.

Ray tracing software was used to optimize the design and layout of the imaging optics and prism. These components included the coded input aperture, the F/8 relay lens, the double Amici prism and the CCD, in that order. The relay lens was not designed for use with any other optical elements. However, ray tracing was used to modify the prescription of the lens with the introduction of the double Amici prism between the relay lens and the CCD, as shown in

Fig. 1(a). After optimization of the design, the relay lens was positioned 39.3 mm from the coded aperture plane and 43.8 mm from the CCD. The overall design had a paraxial magnification of  $-1$  and a working F/# of 8.04.

The total amount of dispersion created by the double Amici prism can be adjusted by translating the prism between the relay lens and the detector array along the optical axis. This translation induces negligible change in the point spread function at the detector array. Figures 3, 4 and 5 show the spot diagrams of rays corresponding to wavelengths of 450, 550 and 650 nm respectively from 9 different field points on the input coded aperture plane mapping to 9 different points on the detector array. The front of the prism is located 8.6 mm from the relay lens and the back of the prism is located 23.2 mm from the CCD. The RMS radii of the spots range between  $7 - 9 \mu\text{m}$  at 450 nm,  $3 - 8 \mu\text{m}$  at 550 nm, and  $3 - 8 \mu\text{m}$  at 650 nm. The radius of the Airy disk ranges between  $4.4 - 6.4 \mu\text{m}$  over 450 - 650 nm. The system is not diffraction limited, as the spot sizes are larger than the size of the Airy disk, a result of the double Amici prism perturbing the imaging properties of the relay lens.

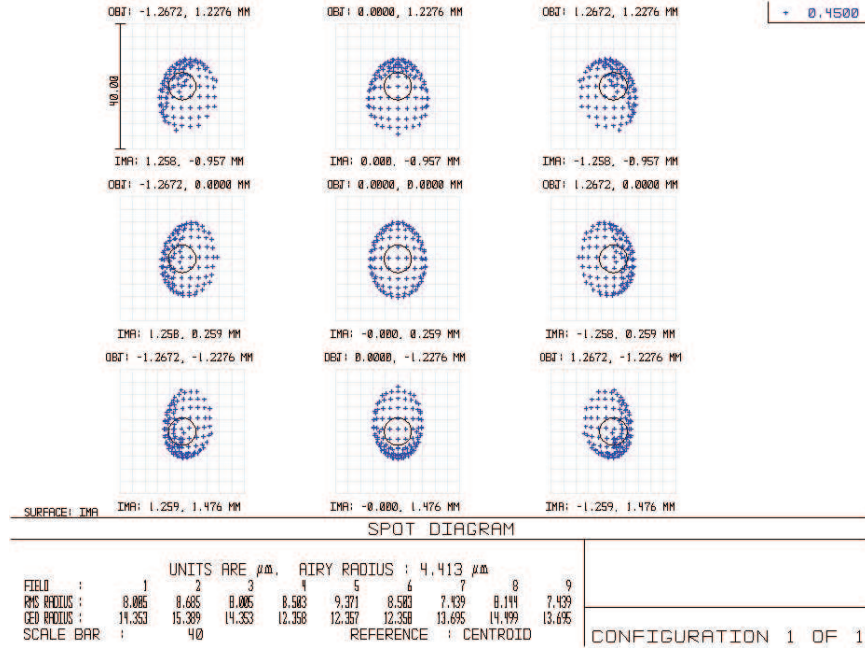


Fig. 3. Spot diagrams of 450 nm rays at 9 different field points at the detector.

### 3. Light propagation model

This section describes the basic CASSI mathematical model. We denote the object's spatial power spectral density at the aperture code as  $f_0(x'', y'', \lambda)$ . Then the power spectral density after the coded aperture is

$$f_1(x', y', \lambda) = T(x', y') f_0(x', y', \lambda) \quad (1)$$

where  $T(x, y)$  is the transmission pattern printed on the coded aperture as shown in Figure 2. The pattern is designed as an array of square features, with each feature being composed of  $2 \times 2$  elements. Each element has the same size as that of a detector pixel,  $\Delta$ . Let  $t_{i,j}$  represent

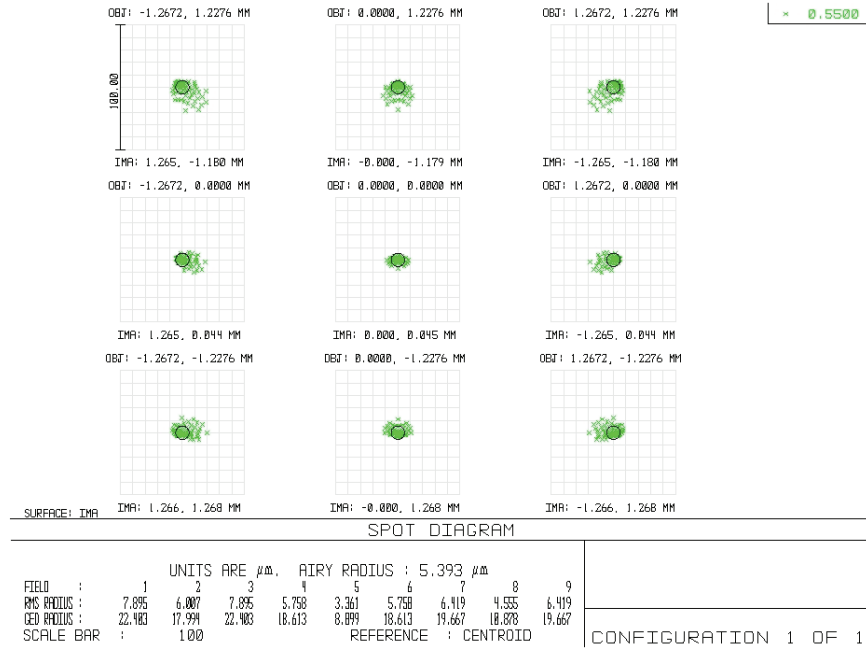


Fig. 4. Spot diagrams of 550 nm rays at 9 different field points at the detector.

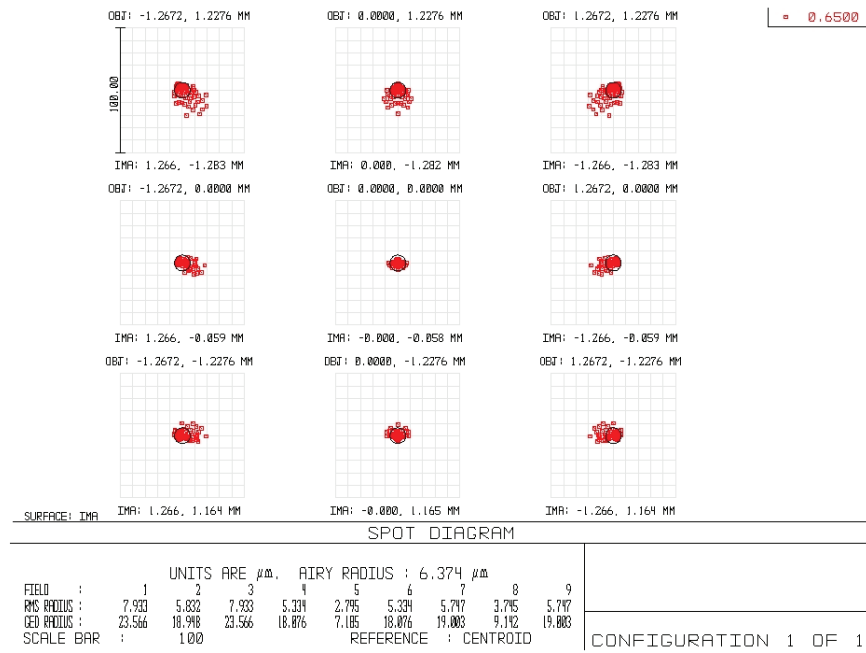


Fig. 5. Spot diagrams of 650 nm rays at 9 different field points at the detector.

the binary value at the  $(i, j)^{th}$  element, with a 1 representing a transmissive code element and a

0 representing an opaque code element. Then,  $T(x,y)$  can be described as

$$T(x,y) = \sum_{i,j} t_{i,j} \tau(i,j;x,y), \quad (2)$$

with

$$\tau(i,j;x,y) = \text{rect}\left(\frac{x}{\Delta} - i, \frac{y}{\Delta} - j\right).$$

After further propagation through the relay lens and the double Amici prism, the power spectral density at the detector plane is

$$\begin{aligned} f_2(x,y,\lambda) &= \iint f_1(x',y',\lambda) h(x' - \phi(\lambda) - x, y' - y, \lambda) dx' dy' \\ &= \iint T(x',y') f_0(x',y',\lambda) h(x' - \phi(\lambda) - x, y' - y, \lambda) dx' dy'. \end{aligned} \quad (3)$$

where  $h(x' - x, y' - y, \lambda)$  represents the shift-invariant optical impulse response of the relay lens and the double Amici prism.  $\phi(\lambda)$  describes the dispersion induced by the double Amici prism. The detector array measures the intensity of incident light rather than the spectral density. As a result, the image on the detector array is the result of an integration process over the spectral range  $\Lambda$ . Furthermore, the detector array is spatially pixelated by the pixel function

$$p(m,n;x,y) = \text{rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n\right),$$

with  $\Delta$  denoting the side length of the pixel in each spatial dimension. As a result, the measurement at the  $(m,n)^{th}$  pixel is

$$\begin{aligned} g_{n,m} &= \int_{\Lambda} \iint f_2(x',y',\lambda) p(m,n;x,y) dx dy d\lambda \\ &= \int_{\Lambda} \iiint T(x',y') f_0(x',y',\lambda) h(x' - \phi(\lambda), x, y', y, \lambda) p(m,n;x,y) dx' dy' dx dy d\lambda \\ &= \int_{\Lambda} \iiint \sum_{i,j} t_{i,j} \text{rect}\left(\frac{x'}{\Delta} - i, \frac{y'}{\Delta} - j\right) f_0(x',y',\lambda) h(x' - \phi(\lambda) - x, y' - y, \lambda) \\ &\quad \times \text{rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n\right) dx' dy' dx dy d\lambda \\ &= \sum_{i,j} t_{i,j} \Omega_{i,j,m,n}. \end{aligned} \quad (4)$$

Since CASSI images the aperture code onto the CCD along the y axis,  $\Omega_{i,j,m,n} = \Omega_{i,n,m,n} \delta_{j,n}$ , so that

$$\begin{aligned} g_{n,m} &= \sum_{i,j} t_{i,j} \Omega_{i,j,m,n} \\ &= \sum_i t_{i,n} \Omega_{i,n,m}. \end{aligned} \quad (5)$$

Here

$$\begin{aligned} \Omega_{i,n,m} &= \int_{\Lambda} \iiint \text{rect}\left(\frac{x'}{\Delta} - i, \frac{y'}{\Delta} - n\right) \text{rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n\right) \\ &\quad \times f_0(x',y',\lambda) h(x' - \phi(\lambda) - x, y' - y, \lambda) dx' dy' dx dy d\lambda. \end{aligned} \quad (6)$$

If we let  $x'' = x' - i\Delta$  and  $y'' = y' - n\Delta$ ,

$$\begin{aligned}\Omega_{i,n,m} &= \int_{\Lambda} \iiint \text{rect}\left(\frac{x''}{\Delta}, \frac{y''}{\Delta}\right) \text{rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n\right) \\ &\quad \times f_0(x'' + i\Delta, y'' + n\Delta, \lambda) h(x'' + i\Delta - \phi(\lambda) - x, y'' + n\Delta - y, \lambda) \\ &\quad \times dx'' dy'' dx dy d\lambda.\end{aligned}\quad (7)$$

If we let  $x''' = x - m\Delta$  and  $y''' = y - n\Delta$ ,

$$\begin{aligned}\Omega_{i,n,m} &= \int_{\Lambda} \iiint \text{rect}\left(\frac{x''}{\Delta}, \frac{y''}{\Delta}\right) \text{rect}\left(\frac{x'''}{\Delta}, \frac{y'''}{\Delta}\right) \\ &\quad \times f_0(x'' + i\Delta, y'' + n\Delta, \lambda) h(x'' + i\Delta - \phi(\lambda) - x''' + m\Delta, y'' - y''', \lambda) \\ &\quad \times dx'' dy'' dx''' dy''' d\lambda.\end{aligned}\quad (8)$$

If we assume that the dispersion by the double Amici prism is approximately linear over the spectral range of the system so that  $\phi(\lambda) = \alpha\lambda$ , and letting  $\lambda' = \lambda + \frac{i\Delta + m\Delta}{\alpha}$ ,

$$\begin{aligned}\Omega_{i,n,m} &= \int_{\Lambda} \iiint \text{rect}\left(\frac{x''}{\Delta}, \frac{y''}{\Delta}\right) \text{rect}\left(\frac{x'''}{\Delta}, \frac{y'''}{\Delta}\right) \\ &\quad \times f_0(x'' + i\Delta, y'' + n\Delta, \lambda' + \frac{m\Delta + i\Delta}{\alpha}) h(x'' - x''' - \alpha\lambda', y'' - y''', \lambda' + \frac{m\Delta + i\Delta}{\alpha}) \\ &\quad \times dx'' dy'' dx''' dy''' d\lambda'.\end{aligned}\quad (9)$$

If we define a discrete, sampled version of the data cube as

$$\begin{aligned}f_{i,n,m} &= \int_{\Lambda} \iiint \text{rect}\left(\frac{x''}{\Delta}, \frac{y''}{\Delta}\right) \text{rect}\left(\frac{x'''}{\Delta}, \frac{y'''}{\Delta}\right) \\ &\quad \times f_0(x'' + i\Delta, y'' + n\Delta, \lambda' + \frac{m\Delta}{\alpha}) h(x'' - x''' - \alpha\lambda', y'' - y''', \lambda' + \frac{m\Delta - i\Delta}{\alpha}) \\ &\quad \times dx'' dy'' dx''' dy''' d\lambda',\end{aligned}\quad (10)$$

so that  $f_{i,n,m+i} = \Omega_{i,n,m}$ . As a result, the array of detector measurements can be written as

$$\begin{aligned}g_{n,m} &= \sum_i t_{i,n} f_{i,n,m+i} \\ &= \sum_k t_{k-m,n} f_{k-m,n,k}.\end{aligned}\quad (11)$$

This equation provides a simple interpretation of the CASSI measurement of the sampled data cube. First, each spectral slice of the sampled data cube is spatially modulated by an element-wise multiplication with the aperture code. This spatially modulated data cube is then sheared by passing through the double Amici prism, which results in the shift in the first index. Finally, this spatially modulated and sheared data cube is integrated over the wavelength dimension to produce the array of measurements on the detector. This measurement can also be interpreted as an operation on a sheared data cube  $f'_{m,n,k} = f_{m-k,n,k}$  so that

$$g_{n,m} = \sum_k t_{m-k,n} f'_{m,n,k}$$

This can also be written as a matrix-vector expression

$$\mathbf{g} = \mathbf{H}\mathbf{f}.\quad (12)$$



The matrix  $\mathbf{H}$  in (12) represents the CASSI system operator. It is a non-negative, binary matrix that maps voxels of the three-dimensional sampled and sheared data cube to pixels of the detector array. As the number of pixels on the detector used for the measurement is smaller than the number of aperture code elements multiplied by the number of spectral channels, the system of equations is under-determined.

We interpret (10) as the object data cube,  $f_0(x', y', \lambda)$ , being filtered by the CASSI system transfer function, which ultimately determines the spatial and spectral resolution of the data cube estimate. The spatial resolution of the data cube depends on the point spread function,  $h(x', x, y', y, \lambda)$ , of the relay optics and double Amici prism optical transfer function, the pixel width,  $\Delta$ , numerical reconstruction effects, and the size of a feature on the coded aperture. However, if we ignore reconstruction effects, the spatial resolution is approximately given by the width and height of the smallest feature on the coded aperture, or a  $2 \times 2$  block of detector pixels ( $19.8 \mu\text{m} \times 19.8 \mu\text{m}$ ). Using larger features on the coded aperture will produce larger transmissive areas, but will also rely more heavily on numerical reconstruction to estimate the spatial content of objects in the data cube that are being imaged on to opaque areas of the coded aperture.

The spectral resolution of the data cube depends on the optical transfer function, the pixel width, reconstruction numerical reconstruction effects, the amount of dispersion induced by the double Amici prism, and the size of smallest feature on the coded aperture. The position of the double Amici prism between the relay lens and the CCD will determine the amount of dispersion (in pixels) across the CCD. The spectral resolution will be limited by the smallest feature size on the coded aperture. If we assume that the dispersion is in the horizontal direction, 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 CCD. Thus, the spectral resolution of the system, i.e. the separation between the spectral channels, is determined by the amount of dispersion (in nm) across the 2 detector pixels. Furthermore, as the dispersion by the double Amici prism is non-linear, the spectral resolution of the CASSI varies as a function of wavelength.

#### 4. CASSI CALIBRATION

In getting to (12), we assumed that the point spread function  $h(x, x', y, y', \lambda)$  was shift invariant, the dispersion by the double Amici prism was linear, and that there was one-to-one mapping between elements of the aperture code to the detector pixels. However, in reality, the point spread function varies across the field, the dispersion is non-linear over CASSI's spectral range, and there are subpixel misalignments between the aperture code features and detector pixels. As a result, the elements of  $\mathbf{H}$  are not necessarily binary as in (12). The quality of the data cube estimate generated by any numerical estimation algorithm strongly depends on how well this matrix is characterized.

After the CASSI has been properly aligned, system calibration is a critical process to fill the gap between an abstract system operator as in (12) and any system-specific operator used for spectral image estimation. This process accounts for the non-ideal imaging of the aperture code onto the detector, any fabrication errors in the manufactured code, and non-linear dispersion by the double Amici prism.

Element-wise calibration of the CASSI system operator would require the capture of the detector response to a spatio-spectral scan of a point source. As this process would require considerably more equipment, we perform a slightly simplified calibration process that does not fully capture the spatio-spectral, shift-variant, point spread function that is suggested by the spot diagrams in section 2. Specifically, we take CCD measurements at 231 equally spaced wavelengths from 440 to 670 nm after uniformly illuminating the coded aperture with each

monochromatic wavelength of light within and around the bandpass (450 – 650 nm) of the system. These calibration measurements were used to build a modified system operator that accounts for the optical blur and non-linear dispersion. Section 5.2 describes how this modified operator was used by the reconstruction algorithm to estimate a spectral image from CASSI measurements.

This set of measurements is obtained with careful efforts to reduce or remove certain data corruptive processes, including (i) the dark noise on the CCD, (ii) the non-uniform spectral intensity of the calibrating light source, and (iii) the non-uniform spectral sensitivity of the CCD.

#### 4.1. Calibration process

We give a brief summary of key processes performed during calibration of CASSI:

1. **Illumination control:** Every effort was made to illuminate the aperture code with the light from the monochromator as uniformly as possible.
2. **Shot-noise reduction:** At each wavelength, 10 CCD frames were captured and averaged to reduce the impact of shot and readout noise.
3. **Background subtraction:** At each wavelength, 10 dark frames were captured at the same exposure time as the bright frames and averaged. The averaged dark frame at each wavelength was then subtracted from its corresponding bright calibration frame.
4. **Exposure time adjustment:** To improve the signal-to-noise ratio (SNR) of the aperture code image at each wavelength, the exposure time at each wavelength was scaled so that the mean counts over the coded aperture in the CCD measurements at all wavelengths were similar.
5. **Light source spectral intensity distribution:** Light from the source at each wavelength was measured with a photodiode having a known responsivity curve to obtain a calibration curve for the non-uniform spectral intensity of the light source.

The wavelength dependent images of the aperture code obtained after background subtraction were normalized by the non-uniform exposure time curve and the non-uniform light source spectral intensity distribution function.

#### 4.2. Calibration results

The procedure described above was performed to capture 231 monochromatic images of the aperture code in 1 nm increments between 440 – 670 nm. We describe two important results of the calibration process. Figure 6 shows 33 monochromatic images of the coded aperture, each displaced by one column of CCD pixels. As the wavelength increases, the image of the aperture code shifts from right to left due to the dispersion by the double Amici prism. These 33 wavelengths define the centers of the 33 spectral channels that the reconstruction algorithm attempts to estimate the content of.

Although the center wavelengths of the chosen spectral channels are separated by one column of CCD pixels, each channel has a different bandwidth due to the non-linear dispersion of the double Amici prism. The purple dots in Figure 7 track the position of the cross on top of the monochromatic image of the coded aperture as a function of wavelength. In essence, this non-linear curve discretely captures the dispersion coefficient  $\phi(\lambda)$  for the double Amici prism. The dispersion at the blue end of the spectrum is much greater than that at the red end. The black crosshairs in Figure 7 identify the 33 wavelengths that define the centers of the 33 spectral channels.

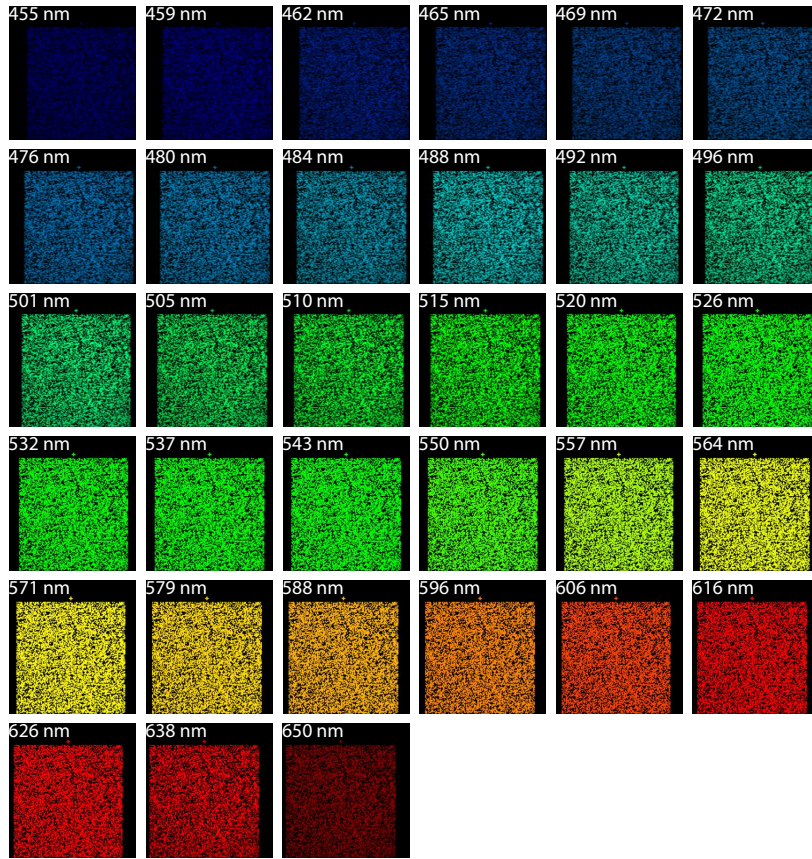


Fig. 6. Demonstrating CASSI's response to uniform illumination at 33 different monochromatic wavelengths. These wavelengths define the centers of CASSI's 33 spectral channels.

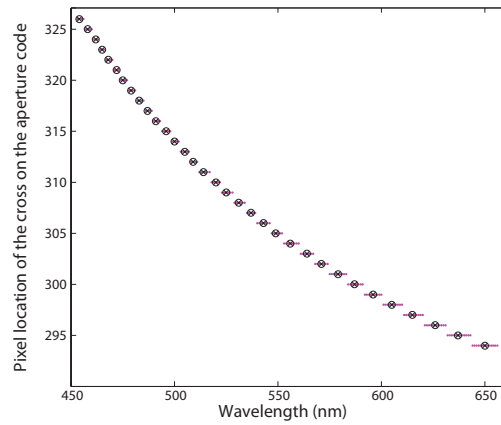


Fig. 7. The purple dots track the location of the cross on the top of the aperture code as a function of wavelength and demonstrate the non-linear dispersion by the double Amici prism. The black crosshairs identify the 33 wavelengths that define the centers of CASSI's 33 spectral channels.

## 5. Numerical estimation of spectral images

In this section, we describe the process of reconstructing a 3D spatio-spectral data cube from a 2D array of CASSI measurements. In particular, we describe the use of the NeAREst algorithm, which stands for Nested Adaptive Refinement Estimation. When the algorithm is applied to CASSI measurements, it does not require human interaction as certain regularization methods do. This property is important for spectral image reconstruction at video rate.

Conventionally, image reconstruction is based on a single system of linear equations, such as that suggested in (12). As the number of voxels in the spatio-spectral data cube to be reconstructed exceeds the number of pixels in the CCD measurements, the system is under-determined, hence having infinitely many solution candidates. In compressive sensing, various encoding schemes in the sensing process and decoding schemes in the reconstruction process have been used to overcome the algebraic limitation on the number of voxels that can be reconstructed and the corresponding detail they reveal. In reconstruction, one utilizes additional information or prescribed conditions to favor a particular solution over the others. A common approach to incorporating the additional condition and selecting the best fitting solution is to reformulate the system of linear equations into an optimization problem with an additive regularization term. For instance, the regularization term may be used to specify the degree of sparsity of an image representation in a mathematically rich and computationally structured function space. Such regularization methods have provided overwhelming evidence in support of the concept and practice of compressive sensing. However, their effective use is limited to the situations that allow human interaction in the reconstruction process to determine a data-dependent regularization parameter, such as one specifying the degree of sparsity of the data cube. The latency of searching, selecting and evaluating the basis functions for the sparse representation must also be considered. We recently reported a comparison of such regularization methods on CASSI measurements [9]. For spectral imaging at video rate, we employ NeAREst, a new method that can be used to perform spectral image reconstruction without human interaction.

### 5.1. A nested set of discretized systems

NeAREst builds and solves a nested set of systems at multiple spatio-spectral scales. Specifically, each of the systems is a discretization of the light propagation model (4) at a particular scale. In principle, the propagation model embeds all the multiscale systems, and the systems are also coupled by a cascading discretization framework. In practice however, one must also incorporate system-specific parameters, which can be calibrated, into the nested systems. The calibration process is necessary to account for two important factors. First, the dispersion in CASSI is non-linear. Second, the effective aperture code that is imaged on to the detector array is not binary valued with perfectly square features as described by the binary function in (2). These factors make practical reconstruction of spectral images from real CASSI measurements more challenging.

We describe the discretization scheme used to construct the system operator for any scale. The measurement at the  $(m, n)$  detector pixel is the integration of the following components, each attributed to a particular wavelength  $\lambda$ ,

$$g(x, y, \lambda) = \sum_{(i,j)} \iint_{A_{i,j}} G_{i,j}(x, y; x', y', \lambda) dx' dy' \quad (13)$$

with

$$G_{i,j}(x, y; x', y', \lambda) = T_{i,j}(x', y') f(x', y', \lambda) h(x' - \phi(\lambda) - x, y' - y, \lambda),$$

where  $\phi(\lambda)$  is the dispersion at wavelength  $\lambda$  and  $A_{i,j}$  is the  $(i, j)^{th}$  virtual sub-aperture at a particular scale. For example, at the finest scale, the size of the sub-aperture is the same size

as the detector pixels, while at a coarser scale, the size of this sub-aperture may be the same size as the smallest aperture code element, i.e.  $2 \times 2$  detector pixels. In (13),  $T_{i,j}(x',y') \geq 0$  represents the effective code function at the  $(i,j)^{th}$  sub-aperture. We note, once again, that even when the sub-aperture is of the same size as the code elements, it is not necessarily equal to the binary-valued function in (2). The intensity measured at detector pixel  $(n,m)$  can be described as

$$g_{n,m} = \sum_k \int_{\Lambda_k} \iint_{P_{n,m}} g(x,y,\lambda) dx dy d\lambda. \quad (14)$$

In this section,  $P_{n,m}$  denotes the spatial support of pixel  $(n,m)$  and  $\Lambda_k$  denotes the  $k^{th}$  spectral channel, which is determined by the calibration process and the scale at which the system of equations is being constructed.

Thus, at a scale characterized by  $\{A_{i,j}\}$  as the sub-apertures and  $\{\Lambda_k\}$  as the spectral channels, there is a system of equations relating a discrete 3D spatio-spectral data cube to the complete 2D array of CASSI measurements,

$$g_{n,m} = \sum_{i,j,k} Q(m,n;i,j,k) f(x_i, y_j, \lambda_k), \quad (15)$$

for some  $(x_i, y_j, \lambda_k) \in A_{i,j} \times \Lambda_k$ , where

$$Q(m,n;i,j,k) = \iint_{P_{n,m}} \int_{\Lambda_k} \iint_{A_{i,j}} T_{i,j}(x',y') h(x' - \phi(\lambda) - x, y' - y, \lambda) dx' dy' dx dy d\lambda,$$

by the generalized mean-value theorem for integration.

We can approximate  $Q(m,n;i,j,k)$  by applying numerical quadrature for integration:

$$Q(m,n;i,j,k) \approx w_{m,n} w_{i,j} w_k T_{i,j}(x'_i, y'_j) h(x'_i - \phi(\lambda_k) - x_m, y'_j - y_n, \lambda_k) \quad (16)$$

where  $w_{m,n}$  and  $w_{i,j}$  are the quadrature weights associated with the spatial discretization and  $w_k$  is the quadrature weight associated with the spectral discretization.

Equations (15) and (16) describe the discretized system associated with a particular spatio-spectral discretization  $\{A_{i,j}\}$  and  $\{\Lambda_k\}$ . In a CASSI system, pixelization at the detector is uniform on a 2D cartesian grid. We assume from now on that the discretization of the aperture at a given scale is also uniform on a cartesian grid. Under such conditions,  $w_{m,n} w_{i,j}$  is constant over all  $(m,n;i,j)$ .

## 5.2. Accomodation of CASSI system specifics

This section describes how we couple the calibration process and the process of building the system operator  $\mathbf{Q}$  at each discretization scale. Through the calibration process, the spectral channels are determined so that the centers of each channel, at the finest scale, are separated by one column of detector pixels. Thus, the position of a channel corresponds to a fixed dispersion in terms of detector pixels relative to a fixed spectral channel. At a coarser scale, we can relate the dispersion of a spectral channel to the channel index  $k$  in a linear fashion as

$$\phi(\lambda_k) = \alpha_\lambda k. \quad (17)$$

Here,  $\alpha_\lambda$  represents the ratio of the dispersion range per spectral channel to the length of detector pixel. For example, if the centers of two neighboring spectral channels are two pixels apart in dispersion,  $\alpha_\lambda = 2$ . Although (17) is similar to the linear assumption in (9), it holds true for the case of non-linear dispersion under the condition that the channels are not equally partitioned in bandwidth. The non-linearity of dispersion is captured by the non-uniform spectral quadrature weights  $w_k$  across the spectral channels.

Substituting (17) into (16), we have

$$Q(m + \alpha_\lambda k, n; i, j, k) \approx c_k h(x'_i - x_m, y'_j - y_n, \lambda_k) T_{i,j}(x'_i, y'_j), \quad (18)$$

where  $c_k = w_{m,n} w_{i,j} w_k$  varies only with the channel index under the assumption of uniform spatial partitions.

The calibration process is also required to account for the point spread function,  $h$ , of the imaging system from the coded aperture to the detector array, which consists of the relay lens and the double Amici prism. We approximate  $h$  with two factors, one that is spatially shift invariant, while the other which captures the spatial variation at each spectral channel. Specifically, the operator  $Q$  is expressed in terms of the calibrated factors,

$$Q(m + \alpha_\lambda k, n; i, j, k) \approx c_k h^{[k]}(x'_i - x, y'_j - y_n) T_{i,j}^{[k]}(x'_i, y'_j). \quad (19)$$

Here,  $h^{[k]}$  is assumed to be the spatially shift invariant part of the point spread function that depends on the imaging optics and the  $k^{\text{th}}$  spectral channel.  $T_{i,j}^{[k]}$  is the effective grayscale code calibrated at the corresponding channel and compensates for the difference between  $h(x, y, \lambda_k)$  and  $h^{[k]}(x, y)$ .

Equations (15) and (19) describe the system associated with a particular spatio-spectral discretization  $\{A_{ij}\}$  and  $\{\Lambda_k\}$ . They differ from Equations (15) and (16) in that the calibration process is taken into account and the reconstruction process is considered as well. Our experiments show that reconstruction of a data cube using the calibrated operators is much superior to one that uses operators that do not account for the calibration process.

### 5.3. The variational method for reconstruction

Equations (15) and (19) describe the mapping of the 3D discrete data cube  $f(x_i, y_j, \lambda_k)$  at a given scale to a measurement on a detector pixel  $g_{n,m}$  for a calibrated CASSI. This system can be written in a matrix-vector expression,

$$\mathbf{g} = \mathbf{Q}\mathbf{f} = \sum_k \mathbf{Q}_k \mathbf{f}_k, \quad (20)$$

where  $\mathbf{f}_k$  is a 2D slice of the 3D data cube associated with the  $k^{\text{th}}$  spectral channel, and  $\mathbf{Q}_k$  is the sub-matrix relating  $\mathbf{f}_k$  to its contribution to the detector measurements. Across the different discretization scales, the system varies in the degree of freedom, with the degree being lower at a coarser scale. The matrix  $\mathbf{Q}$  at a coarse scale has fewer columns than that at a finer scale, but the former is not necessarily a sub-matrix of the latter. NeAREst transforms the linear systems at each scale into a variational formulation. Specifically, Csiszar's version of the Kullback-Leibler (KL) divergence is used as the objective function,

$$\arg \min_{\mathbf{f} \geq 0} \sum_{m,n} g_{n,m} [\log(g_{n,m}) - \log((\mathbf{Q}\mathbf{f})_{n,m})] - [(g_{n,m} - \text{sgn}(g_{n,m})(\mathbf{Q}\mathbf{f})_{n,m})], \quad (21)$$

with a slight modification to permit zero elements in  $\mathbf{g}$ . In (21), the second term is the algebraic difference between  $\mathbf{g}$  and  $\mathbf{Q}\mathbf{f}$ , while the first term is the weighted difference in the bitwise representation, with  $g_{n,m}$  as the weights. This seems similar to the regularization approaches we have used in the past, including the Gradient Projection for Sparse Reconstruction (GPSR) algorithm and the Two-step Iterative Shrinkage and Thresholding (TwIST) algorithm [9]. The difference is that we use (21) without having to specify a regularization parameter. Nonetheless, there are two regularization means in effect. The first is the non-negativity constraint. The variational method excludes solutions with negative values by design. In contrast, when processing

CASSI measurements with the GPSR and TwIST algorithms, this non-negative condition was enforced by brute force truncation. This non-negativity constraint in (21) substantially narrows the configuration space of solution candidates. It, however, does not necessarily make the solution unique ([10]). The second means of regularization is the use of the nested structure across different scales, governed by the physical model in the continuous form. For example, a solution candidate to the system at a fine scale is disqualified if it does not yield a solution at a coarse scale. In the following sub-section, we show that this can be done efficiently.

#### 5.4. Efficient iterations by exploiting the nested structure of NeAREst

Numerical reconstruction with all the variational formulations mentioned above resort to iterative methods. For any such method, there are three important factors to consider: (a) the computational cost per iteration, (b) the number of iterations used, and (c) any additional human interaction that may be required, such as the need to determine a data-dependent regularization parameter.

At every iteration, matrix-vector multiplication with the matrix  $\mathbf{Q}$  is invoked. Due to the non-uniform spectral weights,  $\mathbf{Q}_k \mathbf{f}_k$  are computed separately.

Specifically, let

$$\mathbf{Q}_k \mathbf{f}_k(m, n) = \sum_{(i, j)} \mathcal{Q}(m, n; i, j, k) f(x'_i, y'_j, \lambda_k). \quad (22)$$

By (19), the computation of  $\mathbf{Q}_k \mathbf{f}_k$  involves a convolution with  $h^{[k]}$  and a masking with  $T_{i, j}^{[k]}$ . Strictly speaking, there is another stage of propagation of light from the scene to the aperture as well, which can fortunately be merged with  $h^{[k]}$ .

We note first that with any variational method being used to reconstruct CASSI measurements, the iterative estimates of  $\mathbf{f}$  are explicitly formed in the image space, regardless of the intermediate representation of  $\mathbf{f}$  in an alternate basis with a multi-scale representation, such as a wavelet basis. Therefore, the computational cost per iteration is at least as much as that for the matrix-vector multiplication with  $\mathbf{Q}$ . In this aspect, NeAREst has the advantage that the matrix size is smaller at a coarser scale, lowering the computational cost per iteration at coarser levels. In comparison, a GPSR-like method accesses the largest matrix entirely at every iteration even when the number of representation scales is small. In addition, the computational cost per iteration increases substantially with an increase in the number of representation scales as well as with an increase in the variety of basis functions used.

Next, NeAREst reduces the total number of iterations by two algorithmic schemes. The first is a relaying scheme that passes the coarse scale solution to the next finer scale as the initial guess. The second scheme is an acceleration scheme that utilizes an efficient approximation to the gradient of the KL divergence within a few iterations being performed at a given scale.

#### 5.5. Reconstruction of a spectral image video from CASSI measurements

The use of NeAREst enabled frame-by-frame reconstruction of a spectral video of a dynamic scene captured by CASSI, free of human interaction. The scene, which is illustrated in the next section, consisted of colorful and lives candle flames. Each video frame was reconstructed with the same set of system operators and the same algorithm, without the need to specify a frame-dependent regularization parameter.

At the finest scale, CASSI could measure 33 spectral channels, each separated by a column of detector pixels. We built two other discrete systems at a coarser scale, each with 17 spectral channels. A combination of the solutions at the coarse level, based on an adaptive quadrature rule, was used as the initial guess at the fine scale. Experimental results demonstrated that this setting produced efficient results that revealed spatial detail which the other algorithms failed to recover. For both convenience and efficiency during the reconstruction process, the

3D spatio-spectral data cube was represented directly in the image space. To solve the KL objective function, we used the Richardson-Lucy iterative method, which has the advantage of keeping the iterative solutions non-negative, without requiring any artificial truncation that can potentially induce additional artifacts.

The numerical reconstruction reported in the next section was carried out in single-threaded MATLAB on a SunFire X4100 M2 workstation with a Dual-Core AMD Opteron 2220 running at 2.8 GHz. The processing of each frame started afresh, and took approximately 45 seconds per frame, rendering better image quality in a shorter time, in comparison to the other two algorithms we experimented with. With modern multi-core processors and with special-purpose hardware, it is feasible for the reconstruction process to be real time so that it matches the video rate of capture from the CASSI detector array.

For spectral imaging of the live candles, one may also exploit the temporal continuity by using the data cube solution to the previous frame as the initial guess of the current one. However, there is also reason not to do so. As an indispensable component of the CASSI system, the reconstruction algorithm must be able to cope with the situation where the changes of a dynamic scene are not necessarily slower than the video rate.

### 5.6. Additional Remarks

The NeAREst framework is applicable to a compressive sensing system in the following situation: (1) the system model permits discretization at an arbitrarily chosen scale; (2) the effective ratio between the number of measurements and the number of voxels to reconstruction is not known a priori; and (3) human interaction in data-dependent regularization is either infeasible or impractical. Reconstruction of a spectral video from CASSI measurements is an example of such a situation.

The construction of nested systems in the NeAREst framework does not require multiple measurements of a scene. The systems relate the complete set of CASSI detector measurements to multiple discretized representations of the same scene. The system operators at each scale that are built using the calibration measurements are system-specific and independent of the data in each frame. This is in contrast to other regularization methods where certain parameter(s) must be reset from one frame to another.

From the reconstruction point of view, NeAREst provides an efficient approach for exploiting local smoothness and local low dimensionality in a high dimensional reconstruction. For example, with GPSR, the spatial sparsity representation of one spectral slice of the 3D data cube is not necessarily similar to that of the neighboring spectral slices. NeAREst does not suffer from such combinatorial discord along the spatial dimensions or the spectral dimension, thanks to the nested structure of its multi-scale systems and the cascading feature of its reconstruction algorithm.

Finally, NeAREst is still in its early stage of development. It will undergo further development and provide greater understanding of compressive sensing systems.

## 6. EXPERIMENTAL RESULTS & DISCUSSION

Five candles with colored flames were arranged on a table, against a black backdrop, as shown in Figure 8(a) (Media 1). The experiment was performed under normal fluorescent lights in the ceiling of the laboratory room. 300 raw CASSI frames were captured over 10 seconds at 30 frames per second. Note that CASSI could also use a CCD with a faster frame rate as long as there is enough light collected in each frame. The aperture code is an array of  $256 \times 248$  elements with each element the same size as the detector pixel. The feature size is twice as large as the detector pixel on each side. The CASSI detector array measurement includes  $256 \times 280$  pixels, as shown in Figure 8(b) (Media 1). One observes easily the spatio-spectral overlap, due



to the dispersion, of the aperture code modulated images of the candles and their respective flames.

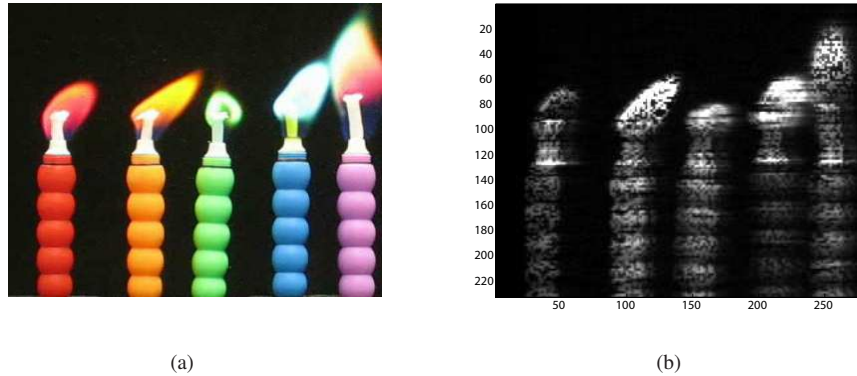


Fig. 8. Single frame excerpts of birthday candles with colored flames as viewed by (a) a Canon SD300 digital camera, and (b) from the CASSI CCD detector array. (Media 1)

Media 1 shows videos of the birthday candles with their live flames as viewed by a Canon SD300 digital camera and CASSI. We note that the flame on the blue candle was often bright enough to saturate the 8-bit CASSI CCD. This would result in errors in the spectral reconstruction of that flame as the various spectral channels are multiplexed on top of each other. We also note that between frames 128 – 141, the blue and purple flames briefly combine, during which time the left edge of the purple flame also saturates the CCD.

The 300 frames were individually processed using the NeAREst algorithm described in section 5. Figure 9 (Media 2) shows an example of a spectral image estimate for frame 107. The figure shows the spatial content of 33 spectral channels between 455 nm and 650 nm. We note that the spatial modulation by the aperture code on each object visible in Fig. 8(b) has been effectively removed from all the spectral channels.

As expected, frames 127 – 141 in Media 2 show the blue flame and the left edge of the purple candle flame appearing in virtually all the spectral channels. This is because during these time instances, these flames were burning brightly enough to saturate the CCD of the CASSI. Figure 10 shows enlarged versions of spectral channels 11, 22, 27, and 30 to demonstrate the spatial detail that the NeAREst algorithm is able to recover.

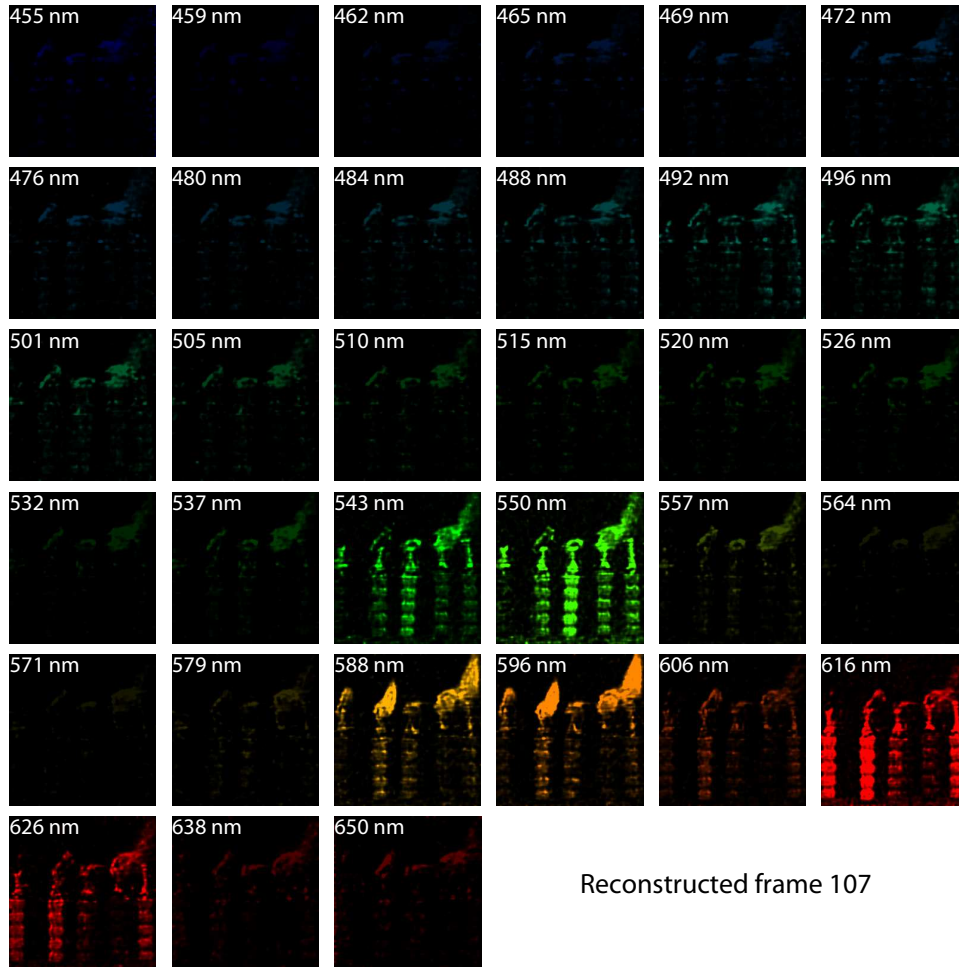


Fig. 9. Spectral image estimate of CASSI frame 107 generated using the NeAREst algorithm. The spatial content in each of 33 spectral channels between 455 and 650 nm is shown. (Media 2)

### 6.1. Quantitative validation with a non-imaging, reference spectrometer

For the purpose of validating the data cube estimate generated with CASSI data, we used the spectral information obtained by a non-imaging spectrometer (Ocean Optics USB2000) as a quantitative reference. The tip of the fiber connected to the spectrometer was brought very close to the body of each candle to obtain a point-wise spectral signature. Figure 11 shows the spectra obtained from the reference spectrometer.

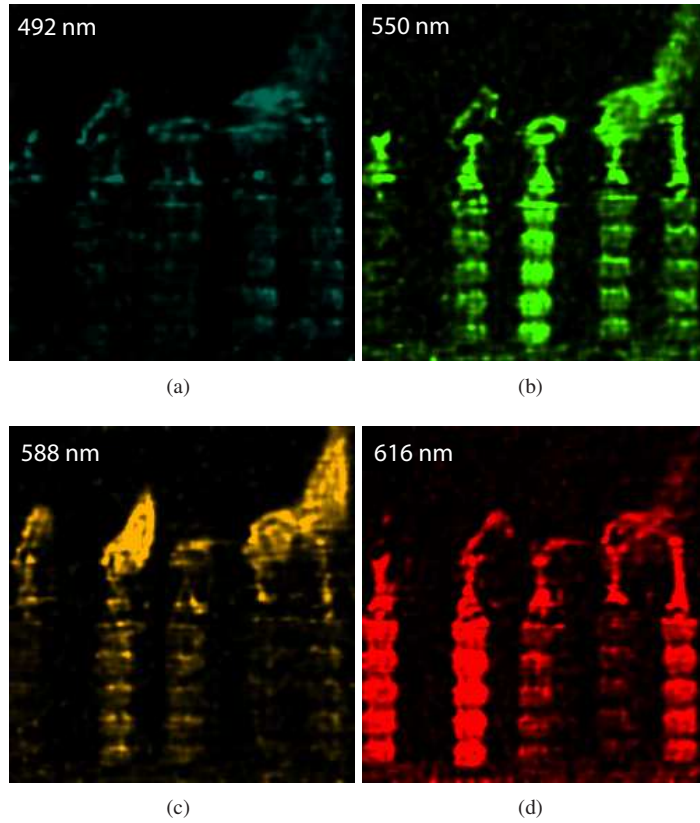


Fig. 10. Enlarged versions of spectral channels 11, 22, 27, and 30.

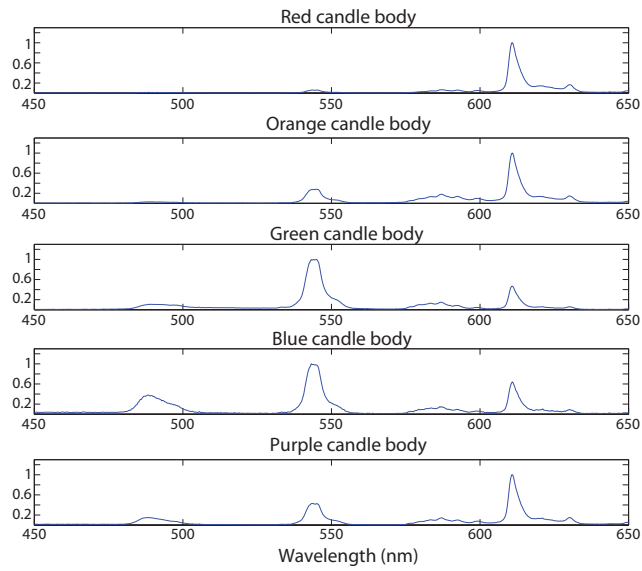


Fig. 11. Spectra of the bodies of the five candles at equally spaced 1 nm wavelengths, as measured by a non-imaging, reference spectrometer.

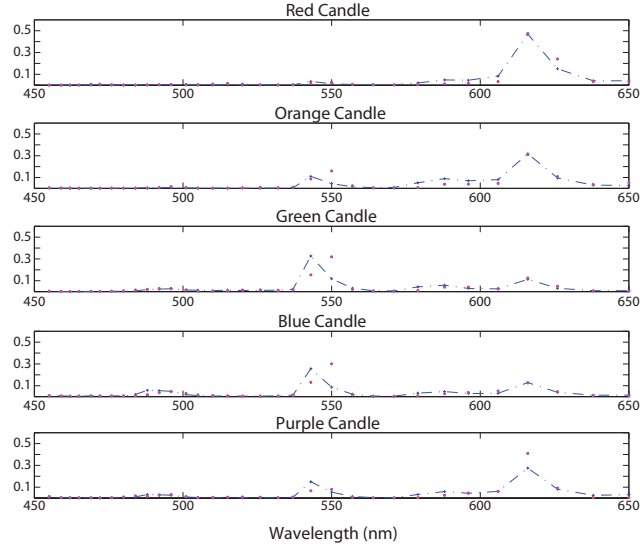


Fig. 12. Channel-wise spectra of the bodies of the five candles. The purple dots are the spectral signatures measured at points on the candle bodies by CASSI. The dashed blue curves are the spectral signatures measured by the reference spectrometer after being integrated into the 33 spectral channels with equivalent bandwidths to the CASSI spectral channels.

These spectral signatures were then integrated into the 33 spectral channels used for our numerical estimation of spectral images from CASSI. The plots in Figure 12 demonstrate the agreement between the reference spectral signatures (in blue dots) and the spectral signature of the candle bodies in the computed estimates using NeAREst (in magenta dots). The spectra are normalized by the total intensity per object. We note that we have not accounted for differences in the system responses of the CASSI and the reference spectrometer in this comparison.

The agreement between the spectra of the candle bodies obtained from the CASSI and the reference spectrometer suggests that the CASSI was also able to correctly measure the spectra of the candle flames. It was not possible to measure reference spectra for the candle flames as they were changing dynamically.

## 7. CONCLUSION

We have reported on significant advances in the optical design and spectral image reconstruction method for a coded aperture snapshot spectral imager using a single disperser. In addition, we have demonstrated the ability of this system to capture spectral images of a dynamic scene at video rate (30 frames per second). The new prototype uses a double Amici prism to introduce spectral dispersion while maintaining a direct view configuration of the overall system. The angles and glass types of the double Amici prism were chosen through optimization of the design of the prototype from the coded aperture plane to the CCD using ray tracing software. A rigorous calibration process was performed to facilitate the generation of system-specific system operators for numerical reconstruction. A new reconstruction method, named NeAREst, was used to process raw frames from CASSI to generate spectral image video of a dynamically changing scene without requiring any specification of a frame-dependent regularization parameter.

## **ACKNOWLEDGEMENTS**

This work was supported by AFOSR grant FA9550-08-1-0151. Ashwin Wagadarikar is partially supported by a postgraduate scholarship from the Natural Sciences and Engineering Research Council (NSERC) of Canada. The authors would like to thank Nathan Hagen for his valuable suggestions on the calibration process.