HIGH PRECISION DISCRETIZATION MODEL FOR CODED APERTURE-BASED COMPRESSIVE SPECTRAL IMAGING

Henry Arguello[†⋆] *Hoover F. Rueda*[†⋆] *Gonzalo R. Arce*[†]

[†] Department of Electrical and Computer Engineering, University of Delaware,
Newark, DE, 19716, USA
[⋆] Universidad Industrial de Santander, Bucaramanga, Colombia
E-mail: {henarfu, rueda, arce}@udel.edu

ABSTRACT

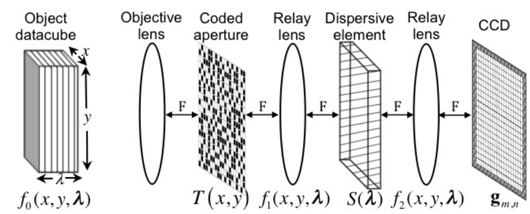
Coded aperture snapshot spectral imaging systems (CASSI) measure the 3D spatio-spectral information of a scene using several compressive 2D focal plane array (FPA) snapshots. The image reconstruction algorithms utilized in CASSI use a first-order approximation of the underlying analog sensing phenomena. A calibration method is then used to compensate for the coarse approximation – an approach not well suited for multishot CASSI systems. This paper develops a more accurate computational model for CASSI which provides a higher quality of image reconstruction. Several simulations are shown to illustrate the performance improvement attained by the new model.

***Index Terms*—** Hyperspectral imaging, compressed sensing, optical imaging, coded aperture

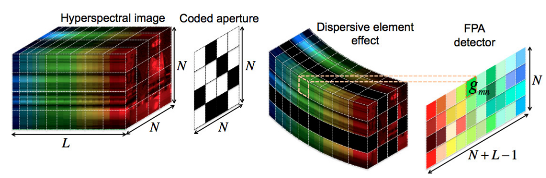
1. INTRODUCTION

Coded aperture snapshot spectral imaging (CASSI) [1, 2] systems as depicted in Fig. 1 capture the spectral information of a scene using a small set of coded focal plane array (FPA) compressive measurements [3, 4]. Compressed sensing (CS) reconstruction algorithms are then used to reconstruct the underlying spectral 3D datacube [5]. Each CASSI measurement is a highly structured random projection of the underlying scene. The structure of these projections is dictated by the CASSI optical system whose only varying components are the aperture code entries. The coded apertures are crucial as they determine the minimum number of FPA measurements needed for correct image reconstruction and the corresponding attainable quality. A hyperspectral datacube $\mathbf{F}$ with $N \times N$ as spatial dimensions, and L spectral planes, can be represented in lexicographical notation as $\mathbf{f} \in \mathbb{R}^{N^2 L}$. Suppose, $\mathbf{f}$ is K-sparse in some basis $\mathbf{\Psi}$ such that $K \ll N^2 L$. CS theory determines that $\mathbf{f}$ can be recovered from m random projections when $m \gtrsim K \log(N^2 L)$. The random projections in CASSI are modeled as $\mathbf{g} = \mathbf{Hf}$, where $\mathbf{H}$ is the optical transmission function which accounts for the coded aperture and the dispersive element operations. The $N \times N \times L$ spectral scene is measured by an $N \times N + L - 1$ FPA detector, thus $\mathbf{H}$ is a $N(N+L-1) \times (N^2 L)$ projection matrix.

This research was supported in part by ONR under the contract NOOO14-1O-C-0 199 and by NSF under the grants EECS-0725422 and CCF-0915800.



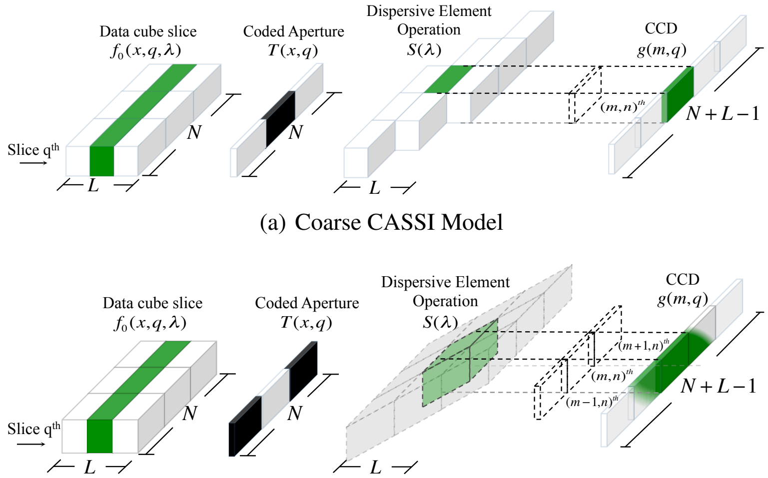
(a) Optical setup



(b) Physical phenomenon

Fig. 1: Coded aperture-based snapshot spectral imaging system (CASSI). (a) Optical elements present in CASSI. (b) Analog physical phenomenon. The hyperspectral scene (datacube) is first spatially modulated by the coded aperture, then dispersed by the dispersive element, and finally integrated in the FPA detector.

The CASSI model in [1] describes the discretization depicted in Fig. 2(a), where a slice of the datacube at a given wavelength is coded by the aperture code, it is sheared linearly by the dispersive element with dispersion $S(\lambda)$ and projected onto the detector. In this model, the dispersion is such that a datacube voxel impinges in a single FPA pixel. We refer to this model as the "coarse model". The projections in CASSI are however the result of an analog phenomenon where discrete spectral bands do not exist. Instead, the datacube contains a continuous set of wavelengths which are all



(a) Coarse CASSI Model

(b) Higher-Order-Precision CASSI Model

Fig. 2: CASSI versus Higher-Order CASSI integration models. (a) In the CASSI model, the dispersion is applied at the spectral band level, where a voxel impinges in a single FPA pixel. (b) In the Higher-Order CASSI model, the dispersion is applied at the sub-band level, such that a voxel impinges into several adjacent FPA pixels.

dispersed according to the prism characteristics. The projection of a single datacube voxel onto the FPA detector is therefore not limited to a single pixel in the detector, but it is projected onto several adjacent pixels as shown in Fig. 2(b). In the coarse model, this effect is treated as blurring across adjacent detector pixels that can be ameliorated by calibration. Our contribution in this paper is to characterize the discretization of the analog physical phenomena more precisely and thus avoid calibration and loss of information. This, in turn, modifies the $\mathbf{H}$ matrix so as to account for the higher order discretization model. Thus, the coarse CASSI model in [1] can be regarded as a first order approximation of the compressive imaging sensing process. In the following sections, we detail the higher-order discretization model in CASSI, and its implications on the CS inverse problem and signal reconstruction.

2. SYSTEM MODEL

2.1. CASSI Model

Denote the analog 3D spatio-spectral power source density as $f_0(x, y, \lambda)$, where x and y index the spatial coordinates and λ indexes the wavelength. As shown in Fig. 1(b), the source is first spatially modulated by the coded aperture $T(x, y)$ resulting in the coded field $f_1(x, y, \lambda)$. After propagation through the coded aperture, the spatially modulated spectral information is spectrally dispersed by a dispersive element. The spectral density at the output of the dispersive element, can be expressed as,

$$f_2(x, y, \lambda) = \iint T(x', y') f_0(x', y', \lambda) \delta(x' - (x - S(\lambda)))\delta(y' - y) dx' dy', \quad (1)$$

where $\delta(x' - (x - S(\lambda)))\delta(y' - y)$ represents the optical impulse response of the system and $S(\lambda) = \alpha(\lambda)(\lambda - \lambda_c)$ represents the dispersion induced by a dispersive element centered at the wavelength λ_c with a dispersion coefficient $\alpha(\lambda)$. The resultant intensity image at the FPA is the integration of the field $f_2(x, y, \lambda)$ over the detector's spectral range sensitivity (Λ), that can be represented as $g(x, y) = \int_\Lambda f_2(x, y, \lambda) d\lambda$. Assuming ideal optical elements, a 2D FPA snapshot can be expressed as,

$$g(x, y) = \int_\Lambda T\left(x - S(\lambda), y\right) f_0\left(x - S(\lambda), y, \lambda\right) d\lambda. \quad (2)$$

Since the detector array is spatially pixelated, in the coarse CASSI model (Fig. 2(a)), the $(m, n)^{th}$ pixel measurement is given by,

$$g_{mn} = \iint p(m, n; x, y) g(x, y) dx dy + \eta_{mn}, \quad (3)$$

where η_{mn} represents additive noise and $p(m, n; x, y)$ represents the detector pixelation, which is given by $p(m, n; x, y) = \text{rect}\left(\frac{x}{\Delta} - m, \frac{y}{\Delta} - n\right)$, with Δ being the pitch of the detector.

2.2. Higher-Order CASSI Model

The model in Eq. 3 is such that a single voxel impinges on a single FPA pixel. As depicted in Fig. 2(b), however, the analogous sensing phenomena is such that a single voxel impinges onto several adjacent FPA pixels. Hence, a higher-order approximation is proposed as depicted in Fig. 2(b). Using Eq. (2) and defining boundaries for the FPA integration, the measurement at the $(m, n)^{th}$ detector pixel can be rewritten as,

$$g_{mn} = \iiint_{n\Delta, m\Delta, \Lambda}^{(n+1)\Delta, (m+1)\Delta} T(x - S(\lambda), y) f_0(x - S(\lambda), y, \lambda) d\lambda dx dy. \quad (4)$$

The discretization of the analog three-dimensional scene is given by measuring the energy concentrated in each voxel f_{ijk} of the datacube. It can be represented as the integration of the datacube between voxels boundaries as,

$$f_{ijk} = \int_{i\Delta}^{(i+1)\Delta} \int_{j\Delta}^{(j+1)\Delta} \int_{\lambda_k}^{\lambda_{k+1}} f_0(x, y, \lambda) d\lambda dy dx, \tag{5}$$

where i, j, and k are discrete indices representing x, y, and λ respectively. The spectral axis λ is discretized in L spectral bands, and the spatial axis in N rows and N columns. The range of the k^{th} spectral band is $[\lambda_k, \lambda_{k+1}]$ where λ_k is the solution to the equation $S(\lambda_k) - S(\lambda_0) = k\Delta$, for $k = 0, \ldots, L-1$. Using this discretization model, the $(m,n)^{th}$ measurement in Eq. 4 becomes,

$$g_{mn} = \iint_{n\Delta, m\Delta}^{(n+1)\Delta, (m+1)\Delta} \left[\int_{\lambda_0}^{\lambda_1} T(x-S(\lambda), y) f_0(x-S(\lambda), y, \lambda) d\lambda + \cdots + \int_{\lambda_{L-1}}^{\lambda_L} T(x-S(\lambda), y) f_0(x-S(\lambda), y, \lambda) d\lambda \right] dx dy. \tag{6}$$

Figure 3 shows a zoomed version of one voxel of the source after it is sheared by the prism. Notice that its energy will impinge on up to three different FPA pixels. Each voxel at the source can then be partitioned into three different regions denoted as R_0, R_1, and R_2. Depending on the nature of $S(\lambda)$, a voxel may affect more than 3 detector elements. A general model can be achieved using the set of $M \times L$ matrices $\{\mathbf{W}_k^{(u)}\}_{u=0}^{p-1}$ which represent the datacube voxels weights distribution. The weights in a matrix $\mathbf{W}^{(u)}$ matrix can be shown to be,

$$\omega_{mnk}^{(u)} = \frac{1}{\Delta^2} \iiint_{\{x-S(\lambda), y, \lambda\} \in R_u} dx dy d\lambda, \tag{7}$$

for $u = 0, \ldots, p-1$. Here p represents the number of FPA pixels that a single datacube voxel impinges on. It is given by $p = p_s + 1$ for linear dispersion and $p = \max\lceil (m+1)\Delta - S(\lambda_k) \rceil$ for an specific k^{th} spectral band when a non-linear prism is used. Assuming the energy is distributed uniformly

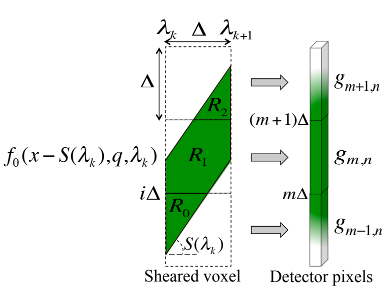


Fig. 3: A single datacube voxel impinges onto 3 adjacent detector pixels

in each spectral band, the integrals in Eq. 6 can then be approximated by,

$$\int_{\lambda_k}^{\lambda_{k+1}} \iint_{\{x-S(\lambda), y\} \in R_u} f(x-S(\lambda), y, \lambda) dx dy d\lambda = \omega_{mnk}^{(u)} t_{(m-u)n} f_{(m-u-k)nk}, \tag{8}$$

for $u = 0, 1, \ldots, p-1$. Using Eq. (8), the portion of the datacube impinging in the $(m,n)^{th}$ detector pixel for the higher-order CASSI, can be succinctly expressed as,

$$g_{mn} = \sum_{k=0}^{L-1} \sum_{u=0}^{p-1} \omega_{mnk}^{(u)} t_{(m-u)n} f_{(m-u-k)nk}. \tag{9}$$

3. HIGHER-ORDER MATRIX MODEL

Expanding the model of the $(m,n)^{th}$ pixel to the complete FPA measurement, the general higher-order discretization model is given by

$$\mathbf{g}_i = \mathbf{H}_i \mathbf{f}, \tag{10}$$

for the i^{th} FPA shot, where $\mathbf{f}$ is the N^2L hyperspectral datacube in lexicographical notation, $\mathbf{H}_i$ is a $N(N+L+p-1) \times N^2L$ matrix representing the optical transmission function, and $\mathbf{g}_i$ a $N(N+L+p-1)$ vector accounting for the compressed FPA measurement. The transmission function can be expressed as,

$$\mathbf{H}_i = \mathbf{P}\mathbf{T}_i \tag{11}$$

where $\mathbf{P}$ represents the dispersive element operation and $\mathbf{T}_i$ the coded aperture. Notice that $\mathbf{P}$ is fixed, while $\mathbf{T}_i$ varies each shot. Here $\mathbf{T}_i$ is a block-diagonal matrix given by,

$$\mathbf{T}_i = \begin{bmatrix} \text{diag}(\mathbf{t}_i) & \mathbf{0}_{N^2} & \cdots & \mathbf{0}_{N^2} \\ \mathbf{0}_{N^2} & \text{diag}(\mathbf{t}_i) & \cdots & \mathbf{0}_{N^2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N^2} & \mathbf{0}_{N^2} & \cdots & \text{diag}(\mathbf{t}_i) \end{bmatrix}, \tag{12}$$

where $\mathbf{t}_i$ is the i^{th} coded aperture in lexicographical notation and $\mathbf{0}_{N^2}$ is a $N^2 \times N^2$ zero-matrix. The matrix $\mathbf{P}$ is given by $\mathbf{P} = \sum_{u=0}^{d} \mathbf{P}_u$, such that

$$\mathbf{P}_u = \begin{bmatrix} & \mathbf{0}_{Nu \times N^2L} & & \\ \text{diag}(\mathbf{W}_0^u) & \mathbf{0}_{N \times N^2} & \cdots & \mathbf{0}_{N \times N^2} \\ \mathbf{0}_{N \times N^2} & \text{diag}(\mathbf{W}_1^u) & \cdots & \mathbf{0}_{N \times N^2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N \times N^2} & \mathbf{0}_{N \times N^2} & \cdots & \text{diag}(\mathbf{W}_{L-1}^u) \\ & \mathbf{0}_{N(p-u) \times N^2L} & & \end{bmatrix}. \tag{13}$$

Equation (10) can be succinctly expressed as,

$$\mathbf{g} = \mathbf{H}\mathbf{f}, \tag{14}$$

where $\mathbf{H} = \left[\mathbf{H}_1^T, \mathbf{H}_2^T, \ldots, \mathbf{H}_K^T\right]^T$, with K being the number of FPA measurements.

4. SIMULATION RESULTS

A spectral data cube $\mathbf{F}$ with 256×256 pixels of spatial resolution and $L = 8$ spectral bands is experimentally obtained by clustering 170 monochromatic images captured every 1nm in the spectral range {450nm – 620nm}. In order to simulate the analog sensing process, the compressive FPA measurements are obtained using the 170 monochromatic images of the datacube. The reconstruction process aims to recover the spectral information in the 8 spectral intervals $\{450 - 463\}$, $\{464 - 477\}$, $\{478 - 493\}$, $\{494 - 510\}$, $\{511 - 530\}$, $\{531 - 556\}$, $\{557 - 586\}$ and $\{587 - 620\}$. Notice that the intervals widths are not constant, as a non-linear dispersive element is used. Aperture codes of size 256×256 are employed. Their entries are realizations of a Bernoulli random variable with parameter $p = 0.5$. The calibration weights for the proposed model are approximated using Eq. (7) and the dispersion curve. Given the set of compressive measurements, the voxels weights distributions and the set of coded apertures, the spectral datacube is recovered using the GPSR algorithm [6]. This algorithm solves the optimization problem, $\hat{\mathbf{f}} = \mathbf{\Psi}_{3D}\{\text{argmin}_{\theta'}\|\mathbf{g} - \mathbf{H}\mathbf{\Psi}_{3\text{D}}\theta'\|_2^2 + \tau\|\theta'\|_1\}$, where $\tau > 0$ is a regularization parameter that balances the conflicting tasks of minimizing the least square of the residuals, while at the same time, yields a sparse solution. The basis representation $\mathbf{\Psi}_{3D}$ is set as the Kronecker product of three basis $\mathbf{\Psi}_{3D} = \mathbf{\Psi}_1 \otimes \mathbf{\Psi}_2 \otimes \mathbf{\Psi}_3$, where the combination $\mathbf{\Psi}_1 \otimes \mathbf{\Psi}_2$ is the 2D-Wavelet Symmlet 8 basis and $\mathbf{\Psi}_3$ is the Discrete Cosine basis. The reconstructions are performed using the new model in Eq. (9), and the traditional model in Eq. (3). Figure 4 shows the PSNR of the reconstructions for the two models as function of the number of measurement shots. The gain achieved by the new model is quantitatively noticeable by averaging the PSNR of the recovered datacubes. This improvement approaches 4 dB when more than two FPA shots are used. Figure 5 depicts the $2^{nd}, 4^{th}, 6^{th}$ and 8^{th} bands of the reconstructed datacube when 6 shots are captured for both models. It can be observed that the new model recovers the spectral information with higher accuracy.

5. CONCLUSIONS

A higher-order precision model for coded-aperture-based spectral imaging systems has been developed. The new approach provides a more accurate discretization model of the analog phenomena in CASSI. Simulation results show a reconstruction improvement of up to 4 dB compared to that attained with the coarse traditional CASSI model.

6. RELATION TO PRIOR WORK

This work provides a more accurate discretization model for spectral imaging systems previously proposed in [1]-[4].

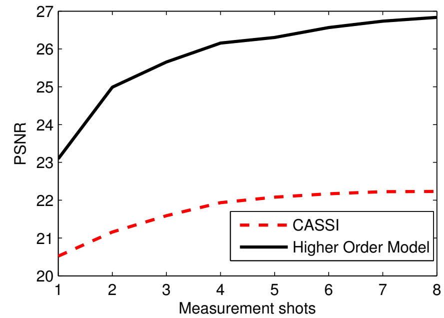


Fig. 4: Averaged PSNR of the reconstructed datacubes as function of the measurement shots. The traditional and the higher order precision models are compared.

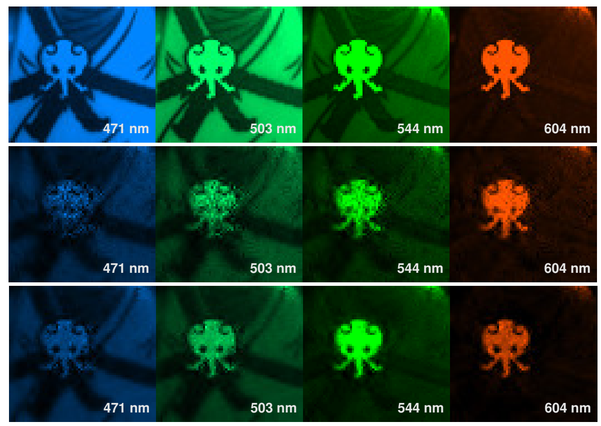


Fig. 5: Reconstruction of the $2^{nd}, 4^{th}, 6^{th}$ and 8^{th} bands when using 6 FPA shots. First row are the original bands. Second and third row are the reconstructions for the traditional and higher-order CASSI models. The PSNR averages across the 8 bands are 22.3 dB and 26.85 dB, respectively.

7. REFERENCES

[1] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Applied Optics*, vol. 47, no. 10, pp. B44–B51, 2008.

[2] Y. Wu, I. Mirza, G. R. Arce, and D. W. Prather, "Development of a digital-micromirror-device-based multishot snapshot spectral imaging system," *Opt. Lett*, vol. 36, pp. 2692–2694, 2011.

[3] H. Arguello and G. R. Arce, "Code aperture optimization for spectrally agile compressive imaging," *JOSA A*, vol. 28, pp. 2400–2413, 2011.

[4] H. Arguello and G. R. Arce, "Rank minimization code aperture design for spectrally selective compressive imaging," *IEEE Trans. Image Process.*, vol. 22, no. 3, pp. 941–954, 2012.

[5] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans Inf Theory*, vol. 52, no. 2, pp. 489, 2006.

[6] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Topics Signal Process*, vol. 1, no. 4, pp. 586–597, 2007.