SYNTHETIC CODED APERTURES IN COMPRESSIVE SPECTRAL IMAGING

Laura Galvis[†] *Henry Arguello*[⋆] *Gonzalo R. Arce*[†]

[†] Department of Electrical and Computer Engineering, University of Delaware, Newark, DE, 19716, USA
[⋆]Department of Computer Science, Universidad Industrial de Santander, Bucaramanga, 680002, Colombia

ABSTRACT

Compressive spectral imagers have gained popularity recently due to their ability to sense a three-dimensional (3D) data cube with just a few two dimensional (2D) coded aperture projection snapshots. The coded apertures are realized by digital micromirror devices (DMD) which often do not match the pitch resolution of the focal plane array (FPA). This paper introduces the forward model and associated reconstruction algorithm for such mismatched spectral imagers, without the loss of spectral and spatial resolution. Simulations show the improvements in the reconstructions achieved with the proposed approach yielding up to 12 dB gain in PSNR with respect to traditional.

***Index Terms*—** Spectral imaging, coded aperture, pitch resolution, compressive sensing, hyperspectral imaging.

1. INTRODUCTION

Spectral imaging measures the intensity of light at different wavelengths for each spatial location in a scene. The resulting three-dimensional (3D) dataset is known as spatio-spectral data cube. Different spectral imagers have been developed to capture one dimensional and two dimensional subsets of the data cube [1]. To obtain the complete data cube, a scanning of the remaining dimensions is required. These instruments are adequate to capture static scenes, but their use in capturing non-static scenes is challenging. Furthermore, the amount of data captured, stored, or transmitted is directly related to the amount of sensed data, thus leading to the manipulation of large datasets. In contrast, compressive spectral imaging (CSI) senses 2D coded projections of the underlying scene such that the number of measurements is far less than that used in scanning-type instruments [2].

The coded aperture snapshot spectral imager (CASSI) is an example of a CSI architecture whose main components are a coded aperture and a dispersive element. Figure 1 illustrates the CASSI architecture. It captures multiplexed

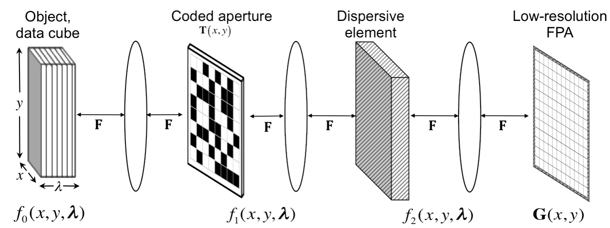

Fig. 1: CASSI Model composed by a coded aperture and a dispersive element. The data cube is coded, spectrally dispersed and integrated on the FPA.

(2D) projections of the spatio-spectral datacube using a snapshot. The data cube is denoted as $\mathcal{F}_{ijk}$ where i and j index the spatial coordinates, and k determines the k^{th} spectral plane. The multiplexed projections in CASSI are given by $g_{mn} = \sum_{k=0}^{L-1} T_{m(n-k)}\mathcal{F}_{m(n-k)k} + \omega_{mn}$, where T represents the coded aperture and ω the noise of the system [3].

The CASSI architecture has been recently modified to allow capturing multiple snapshots, each admitting a different coded aperture pattern [4], [5], [6]. In practice, the coded apertures are usually implemented through photomasks attached to piezoelectric devices [6]. The use of digital micromirror devices (DMD) has been proposed for realizing coded apertures in multi-frame measurements [7]. A limitation for the use of DMDs is that the pixel pitch of the micromirrors Δ_c commonly differs from the pixel size Δ_d of the focal plane array (FPA), so there is no pixel-to-pixel correspondence. Thus, there is a mismatch between the size of the pixels in the coded aperture and the size of pixels in the detector. To circumvent this problem, a general strategy is illustrated in Figure 2 consisting on grouping several pixels in a super-pixel in both, the coded aperture and the detector, such that the relation between them is given by $p_1\Delta_d = p_2\Delta_c$ where $p_1 \neq p_2 \neq 1$ are small integers. This solution, however, reduces significantly the spatial and spectral resolution attained in the reconstructions [2], [8].

This work develops a mathematical model of CASSI with pixel pitch mismatch such that the number of spectral bands and the spatial resolution is dictated by the smaller of the pixel sizes $\min\{\Delta_d, \Delta_c\}$. Using the forward model for the CASSI system with mismatched pixels the spatial and spectral reso-

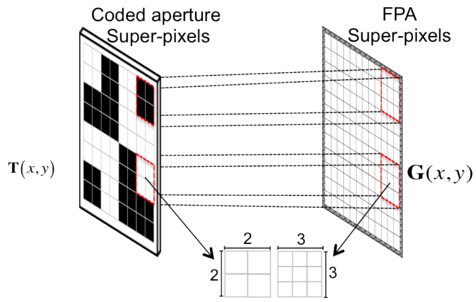

Fig. 2: The pixel pitch mismatch problem: super-pixels are created by grouping several pixels in the coded aperture and in the detector.

lution is increased by a factor of $\max\{p_1, p_2\}$ with respect to the traditional super-pixel approach. Instead of creating a super-pixel, the new technique calculates an equivalent coded aperture referred to as a synthetic coded aperture which accounts for the pixel mismatching.

2. SYNTHETIC CODED APERTURES FOR MISMATCHED PIXELS IN SPECTRAL IMAGING

2.1. CASSI System

The optical elements in the CASSI architecture are depicted in Fig. 1. In this architecture, the spatio-spectral power source density defined as $f_0(x, y, \lambda)$, where (x, y) are the spatial coordinates and λ is the wavelength, is coded by the coded aperture $T(x, y)$ resulting in the coded field $f_1(x, y, \lambda)$. The resulting coded density is then spectrally dispersed by a dispersive element before it impinges on the focal plane array as

$$f_2(x,y,\lambda)=\iint T(x',y')f_0(x',y',\lambda)h(x'-S(\lambda)-x,y'-y)dx'dy', \tag{1}$$

where $h(x' - S(\lambda) - x, y' - y)$ is the optical impulse response of the system, and $S(\lambda)$ is the dispersion induced by the prism along the x-axis. The resulting image at the focal plane array is the integration of the field $f_2(x, y, \lambda)$ over the detector's spectral range sensitivity Λ that can be represented as $g(x, y) = \int_\Lambda f_2(x, y, \lambda)d\lambda$.

A CASSI measurement captured on the $(m, n)^{th}$ pixel of the detector is given by $g_{mn} = \int\int g(x, y)p(m, n; x, y)dxdy$, where $p(m, n; x, y) = rect\left(\frac{x}{\Delta_d} - m, \frac{y}{\Delta_d} - n\right)$ represents the pixel in the detector and Δ_d is the detector pixel pitch.

In discrete notation, the source $f_0(x, y, \lambda)$ can be written as $\mathcal{F}_{ijk}$, where $i \in \{1, \cdots, N\}$ and $j \in \{1, \cdots, N\}$ index the spatial coordinates, and $k \in \{1, \cdots, L\}$ determines the k^{th} spectral plane, the discretized coded aperture is T_{ij}. Notice, that the coded aperture T_{ij} has binary entries such that $T_{ij} \in \{0, 1\}$. The discretized $(m, n)^{th}$ FPA pixel measurement can be then written as

$$g_{mn} = \sum_{k=0}^{L-1} T_{m(n-k)}\mathcal{F}_{m(n-k)k} + \omega_{mn}, \tag{2}$$

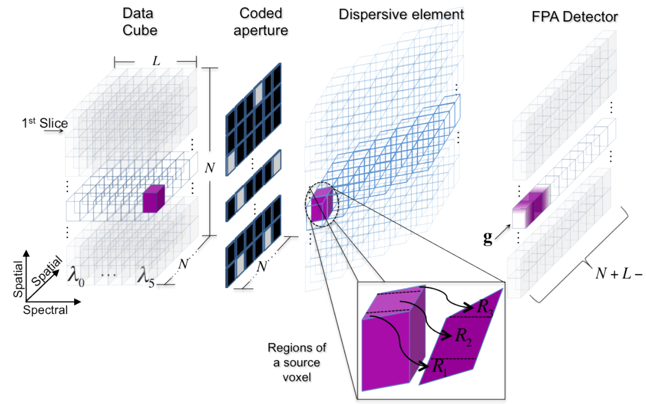

Fig. 3: Illustration of the spatio-spectral data flow in the CASSI architecture. The source is coded by the coded aperture T and dispersed by a prism. The FPA detector integrate the intensities. A source voxel is zoomed to identify the regions R_0, R_1 and R_2 accounted in Eq. 3.

where ω_{mn} is the white noise of the sensing system. Notice that the $N \times M \times L$ data cube is captured by a $N \times (N + L - 1)$ FPA.

The forward model of CASSI was recently extended to account for the non-linearity of the dispersive element. Eq. 2 can thus be modified in the discretization of the dispersion curve [4]. This discretization differs from the model in Eq. 2 inasmuch as the energy from a single voxel is mapped onto three detector pixels, such that each source voxel can be split into three regions R_0, R_1 and R_2. Figure 3 illustrates a zoomed version of the regions of the source voxel affecting one pixel on the detector. The corresponding energy of each region that impinges in the $(m, n)^{th}$ detector pixel is represented by the weights w_{mnku}, where m, n index the spatial coordinates, k the spectral dimension and u accounts for the region R_0, R_1 and R_2 of the source voxel. More specifically, $w_{mnku} = (\iiint_{R_u} dxdyd\lambda)(\iiint_{R_1\cup R_2\cup R_3} dxdyd\lambda)^{-1}$. The FPA measurement is then reformulated as in [4]

$$g_{mn} = \sum_{k=0}^{L-1}\sum_{u=0}^{2} w_{mnku}T_{m(n-k-u)}\mathcal{F}_{m(n-k-u)k}, \tag{3}$$

where $m, n = 0, 1, \ldots, N - 1$, $k = 0, 1, \ldots, L - 1$ and $u = 0, 1, 2$.

2.2. CASSI with pixel pitch mismatch

The proposed strategy to account for the pixel pitch mismatch problem is to perform a mapping of the coded aperture into one with higher resolution accounting for the mismatching effect. The fact that some pixels in the detector will capture the effects of the coded aperture features completely whereas others will capture their effect only partially is used in the formulation of the synthetic coded aperture model. Formally, represent the ratio between the coded aperture and the FPA pixel pitch as $r = \frac{\Delta_c}{\Delta_d}$. Define $\hat{T}(x, y)$ as the synthetic coded aperture with discrete entries $\hat{T}_{ij}$. Notice, that the mismatched pixel model is expressed in terms of the Eq. 3. Denote the pixel size of the synthetic coded aperture as $\Delta_{c'} = \min\{\Delta_d, \Delta_c\}$ and the parameters α and β, accounting for the horizontal and vertical fraction of Δ_c that will be reflected in the new synthetic pixel of $\hat{T}(x, y)$. These fractions can be expressed as,

$$\alpha = \begin{cases} B, & B > 0 \\ 1, & B = 0 \end{cases} \tag{4}$$

$$\beta = \begin{cases} C, & C > 0 \\ 1, & C = 0 \end{cases}, \tag{5}$$

where B and C are defined as

$$B = \left\lfloor \frac{(n+1)}{r} \right\rfloor r - n. \tag{6}$$

$$C = \left\lfloor \frac{(m+1-k-u)}{r} \right\rfloor r - (m-k-u), \tag{7}$$

For $0 < \alpha, \beta < 1$ the above accounts for a pixel in the coded aperture $T(x, y)$ that partially affects a pixel on the detector, whereas $\alpha, \beta = 1$ assumes the coded aperture pixel maps entirely onto a pixel on the detector.

Figure 4 (a) presents an example, where the mismatching relation is $p_1\Delta_d = p_2\Delta_c$ where $p_1 = 3$ and $p_2 = 2$ and super-pixels are created. Figure 4 (b) presents the proposed approach where a black pixel of the coded aperture is analized. Notice that the synthetic coded aperture pixel $\hat{T}_{ij}$ reflects the effects of the coded aperture pixel T_{ij} by the fractions α and β. In contrast, the pixel $\hat{T}_{i,j+1}$ depicts the effects of the same T_{ij} pixel by $(1-\alpha)$ and by β. To calculate the pixel $\hat{T}$, an evaluation of the neighbours of the pixel (i, j) of the coded aperture T is required. In Figure 4, the neighbors denoted as $T_{D,E}, T_{D,E+1}, T_{D+1,E}, T_{D+1,E+1}$ are evaluated. The positions D and E are defined as

$$D = \left\lfloor \frac{n}{r} \right\rfloor \tag{8}$$

$$E = \left\lfloor \frac{(m-k-u)}{r} \right\rfloor. \tag{9}$$

Finally, the synthetic coded aperture $\hat{T}_{ij}$ can be succinctly expressed as

$$\hat{T}_{ij} = \alpha\,(\beta\,T_{D,E} + (1\text{–}\beta)\,T_{D,E+1}) + (1\text{–}\alpha)\,(\beta\,T_{D+1,E} + (1\text{–}\beta)\,T_{D+1,E+1}). \tag{10}$$

Hence, the mapped coded aperture could be tuned in such that the FPA measurement of the discretized CASSI with the pixel mismatch model is expressed as

$$g_{mn} = \sum_{k=0}^{(L-1)p_1} \sum_{u=0}^{2} w_{mnku} \hat{T}_{mn} \mathcal{F}_{m(n-k-u)k}, \tag{11}$$

where the spatial and spectral resolutions of this measurement are dictated by the factor p_1 such that $m, n = 0, 1, \ldots, (N-1)p_1$ and $k = 0, 1, \ldots, (L-1)p_1$.

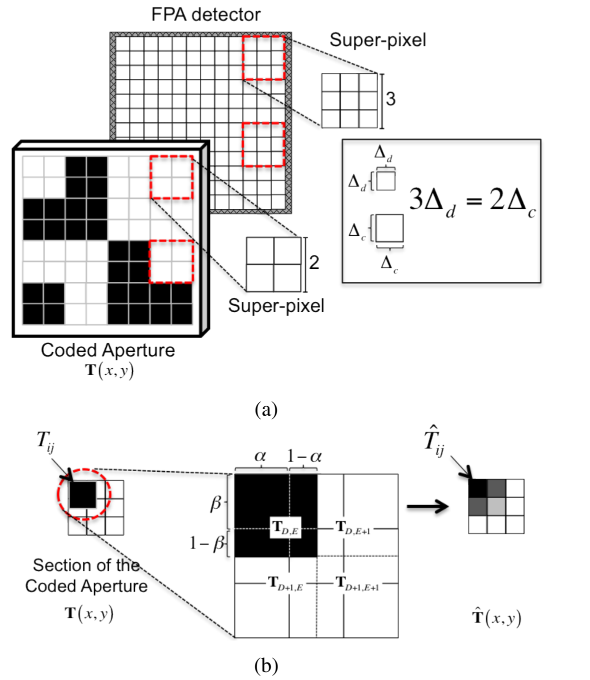

Fig. 4: (a) Superpixels are created to obtain a pixel-to-pixel correspondence; (b) mapping from a section of the coded aperture T to the corresponding section of the synthetic coded aperture $\hat{T}$.

3. SIMULATIONS AND RESULTS

In order to simulate the CASSI with the pixel mismatch model, a set of compressive measurements are calculated using the model in Eq. 11. A test data cube $\mathcal{F}$ with 384×384 pixels of spatial resolution and $L = 24$ spectral bands is used. To construct these measurements, the spectral data cube $\mathcal{F}$ was acquired by a monochromator in the spectral range between $450nm$ and $650nm$. A CCD camera AVT Marlin F0033B, with 656×492 pixels and a pixel pitch size of $9.9\mu m$ is used. One of the most important characteristics of the coded apertures is the transmittance, defined as the amount of light the coded aperture let pass. To analyze this characteristic, two levels of transmittance were selected for the simulations, 25% and 50%. The spatial resolution of the coded aperture T is 256×256 pixels. The corresponding synthetic coded aperture $\hat{T}$ presents a spatial resolution of 384×384 pixels. The GPSR algorithm is used to obtain the reconstructions of the data cube [9]. This algorithm solves the optimization problem $\tilde{\boldsymbol{f}} = \boldsymbol{\Psi}\{\text{argmin}_{\boldsymbol{\theta}} ||\mathbf{g} - \mathbf{H}\boldsymbol{\Psi}\boldsymbol{\theta}||_2 + \tau||\boldsymbol{\theta}||_1\}$, where $\boldsymbol{\theta}$ is an S-sparse representation of $\boldsymbol{f}$ on the basis $\boldsymbol{\Psi}$, and τ is a regularization constant [10]. The basis representation $\boldsymbol{\Psi}$ is set as the Kronecker product between a 2D-Wavelet Symmlet 8 basis and the 1D-Discrete Cosine Transform [11]. The sensing ratio for the simulations is defined as the ratio between the number of measurements and the number of pixels in the reconstructed data cubes. The final reconstruction obtained through the CASSI with mismatched pixels model in Eq. 11 is a $384 \times 384 \times 24$ cube.

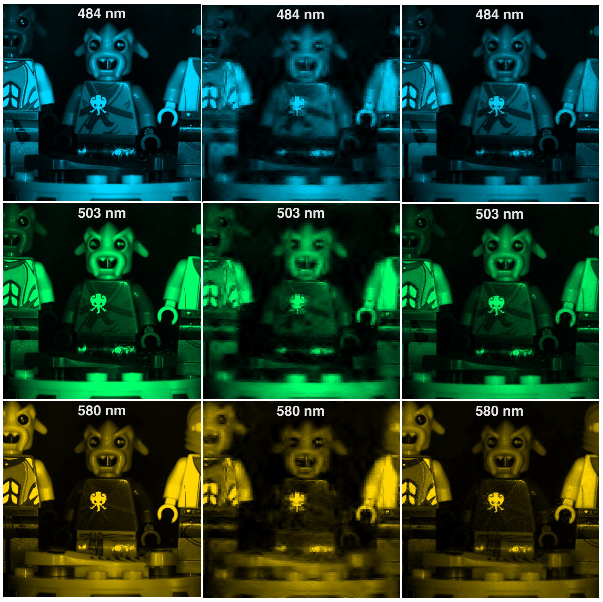

Fig. 5: Reconstructions of bands $k = 3, 4, 7$. Left column depicts the original bands, second column shows the reconstructions using CASSI and the third column shows the reconstructions using the CASSI with synthetic coded apertures model.

The simulations use the relation $p_1\Delta_d = p_2\Delta_c$, where $p_1 = 3,\ p_2 = 2$ and the pitch sizes of the detector and DMD are $\Delta_d = 9.9\mu m,\ \Delta_c = 13.68\mu m$. As a consequence of the mismatch, the reconstruction obtained by the traditional super-pixel approach in Eq. 3 is a $128 \times 128 \times 8$ cube. To compare these reconstructions and those obtained using the CASSI with synthetic coded aperture, an interpolation in the spatial and spectral dimensions should be done. Figure 5 illustrates three original spectral bands of the test data cube and the respective reconstructions obtained with the CASSI, and the CASSI with pixel mismatch using sensing ratios of 50% for each of the cases. The improvement in the spatial quality can be easily noticed. Here, it is important to remark that besides the quality of the reconstructions, the number of spectral bands increases by a factor of p_1 with respect to the super-pixel approach.

A zoomed version of the original data cube $\mathcal{F}$ and the reconstructions using CASSI and CASSI with pixel mismatch respectively can be observed in Figure 6. The scenes are mapped to RGB profiles.

Figure 7 shows a comparison between the mean spatial and spectral PSNR of the reconstructed images for the CASSI and the synthetic coded aperture modeling. Figure 7 (a) and 7 (b) depicts the spatial PSNR with coded apertures having a transmittance of 50% and 25% respectively as a function of different sensing ratios. Figure 7 (c) shows the reconstructed spectral signature for two spatial points, indicated as P1 and P2 in Figure 6 (left), the CASSI with pixel mismatch provides more accurate results. The average spectral PSNR for the spectral bands is presented in Figure 7 (d). The PSNR improvements achieved by the CASSI with pixel mismatch are noticeable.

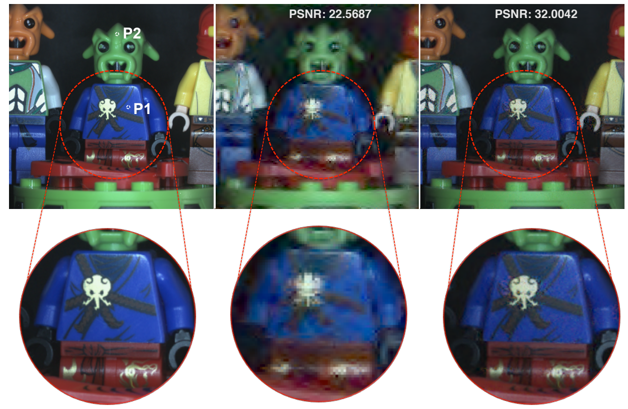

Fig. 6: (Left) Original data cube $\mathcal{F}$ mapped to a RGB profile. (Center) Reconstruction using CASSI and (right), reconstructions using the CASSI with synthetic coded aperture model.

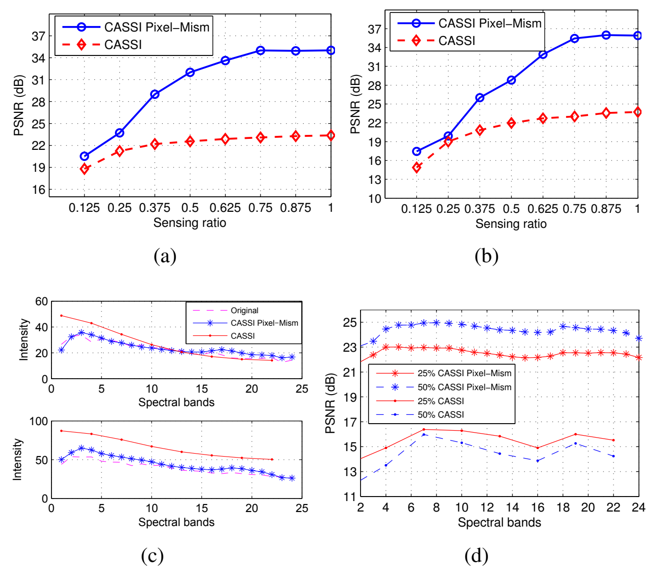

Fig. 7: Spatial PSNR of the reconstructed data cube using the CASSI system and the CASSI with synthetic coded aperture model with a transmittance of (a) 0.5 and (b) 0.25. (c) Reconstructed spectral signatures for two representative spatial points, indicated as P1 and P2 in Figure 6 (left) and (d) averaged spectral PSNR for the different bands.

4. CONCLUSIONS

A forward model of CASSI with pixel mismatching has been developed. The model avoids creating super-pixels to achieve a pixel-to-pixel correspondence between the pixels of the coded aperture and pixels on the detector. Instead, a synthetic coded aperture is defined such that the resolution of the detector is fully utilized. The use of the proposed model increases the spatial and spectral resolution. The achieved improvement for the reconstruction PSNR is up to 12 dB, a three fold improvement in spectral resolution and a more accurate signature profile is obtained compared with the CASSI system where super-pixels are created.

5. REFERENCES

[1] M. E. Gehm, R. John, D. J. Brady, R. M. Willett, and T. J. Schulz, "Single-shot compressive spectral imaging with a dual-disperser architecture," *Opt. Express*, vol. 15, no. 21, pp. 14013–14027, Oct 2007.

[2] Y. Wu, C. Chen, Z. Wang, P. Ye, G. R. Arce, D. Prather, and Garrett J. Schneider, "Fabrication and characterization of a compressive-sampling multispectral imaging system," *Optical Engineering*, vol. 48, no. 12, pp. 123201–123201–12, 2009.

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

[4] H. Arguello, H. Rueda, Y. Wu, D. W. Prather, and G. R. Arce, "Higher-order computational model for coded aperture spectral imaging," *Applied Optics*, vol. 52, no. 10, pp. D12–D21, Apr. 2013.

[5] H. Arguello and G. R. Arce, "Code aperture optimization for spectrally agile compressive imaging," *J. Opt. Soc. Am. A*, vol. 28, no. 11, pp. 2400–2413, Nov 2011.

[6] D. Kittle, K. Choi, A. A. Wagadarikar, and D. J. Brady, "Multiframe image estimation for coded aperture snapshot spectral imagers," *Appl. Opt.*, vol. 49, no. 36, pp. 6824–6833, Dec 2010.

[7] Y. Wu, I. O. 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.

[8] S. K. Nayar, V. Branzoi, and T. E. Boult, "Programmable imaging using a digital micromirror array," in *Computer Vision and Pattern Recognition. CVPR 2004. Proceedings of the IEEE Computer Society Conference on*, 2004, vol. 1, pp. I–436–I–443.

[9] M. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, no. 4, pp. 586–597, 2007.

[10] H. Arguello, C. Correa, and G. R. Arce., "Fast lapped block reconstructions in compressive spectral imaging.," *Applied Optics*, vol. 52, no. 10, pp. D32–D45, 2013.

[11] M. F. Duarte and R. G. Baraniuk, "Kronecker compressive sensing," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 494–504, Feb. 2012.