applied optics

Compressive spectral testbed imaging system based on thin-film color-patterned filter arrays

Hoover Rueda,[1,*] Henry Arguello,[2] and Gonzalo R. Arce[1]

[1]Department of Electrical and Computer Engineering, University of Delaware, Newark, Delaware 19716, USA
[2]Department of Computer Science, Universidad Industrial de Santander, Bucaramanga 680002, Colombia
*Corresponding author: rueda@udel.edu

Received 19 August 2016; revised 19 October 2016; accepted 24 October 2016; posted 24 October 2016 (Doc. ID 274070); published 18 November 2016

Compressive spectral imaging systems can reliably capture multispectral data using far fewer measurements than traditional scanning techniques. In this paper, a thin-film patterned filter array-based compressive spectral imager is demonstrated, including its optical design and implementation. The use of a patterned filter array entails a single-step three-dimensional spatial-spectral coding on the input data cube, which provides higher flexibility on the selection of voxels being multiplexed on the sensor. The patterned filter array is designed and fabricated with micrometer pitch size thin films, referred to as pixelated filters, with three different wavelengths. The performance of the system is evaluated in terms of references measured by a commercially available spectrometer and the visual quality of the reconstructed images. Different distributions of the pixelated filters, including random and optimized structures, are explored.

OCIS codes: (110.4234) Multispectral and hyperspectral imaging; (110.0110) Imaging systems; (170.1630) Coded aperture imaging; (230.7408) Wavelength filtering devices; (110.4155) Multiframe image processing.

http://dx.doi.org/10.1364/AO.55.009584

1. INTRODUCTION

Compressive spectral imaging (CSI) techniques capture multiplexed and coded projections of a scene. The three-dimensional (3D) underlying data cube is then estimated by exploiting the fact that spectral images are highly correlated and admit sparse representations [1]. Different CSI architectures have been proposed to date [2–5]. All use block-unblock binary-coded apertures and one or two dispersive elements to modulate the optical field from the scene. These coded apertures have been fabricated using materials such as chrome-on-quartz, rendering coded aperture elements that are either opaque or translucent to the whole wavelengths of interest, as illustrated in Fig. 1(a). These coded apertures are referred to as photomasks. Recent coating technologies have allowed the design of patterned arrays of pixelated optical filters [6,7]. Since each pixelated filter attains spatial and spectral coding in a single step, as presented in Fig. 1(b), coding elements based on patterned optical coatings have led to more efficient 3D coding strategies and more compact imaging systems.

There are several fabrication approaches to attain pixelated optical filters: color-dye gels [8,9], Fabry–Perot microstructures [10,11], and thin films [12,13]. Dye gels have been widely used on CCD cameras due to their low cost, while Fabry–Perot and thin-film technology offer much higher filter precision but at an initial higher cost compared to dye gels. The principal difference between these technologies, apart from the cost and fabrication procedure, is the extinction ratio. Dye gels yield broad transition bands, whereas Fabry–Perot and thin films provide a sharper transition that, in turn, enables purer color filtering.

Recently, we have proposed and numerically demonstrated the use of patterned filter arrays, also referred as colored coded apertures, in compressive spectral imaging systems [14,15]. In this paper, we report on the development and fabrication of a thin-film-based patterned filter array and its implementation in a CSI testbed architecture. Moreover, an analysis on the impact of the distribution of the pixelated filters on the quality of the reconstructed 3D spectral images is presented, comparing random against optimized distributions. To evaluate the quality of the reconstructed images, the results are presented in terms of monochromatic spectral images, and also mapped to RGB profiles for spatial resolution and visual inspection. To evaluate the spectral fidelity, spectral profiles of selected representative points in the scenes are compared against references measured by a spectrometer. The attained results show improved quality on the images obtained by the optimized distribution of the pixelated filters over the ones with random distributions.

2. COMPRESSIVE OPTICAL SENSING IMAGE MODEL

We focus on the single-disperser coded aperture snapshot spectral imager (SD-CASSI), which typically uses a photomask to

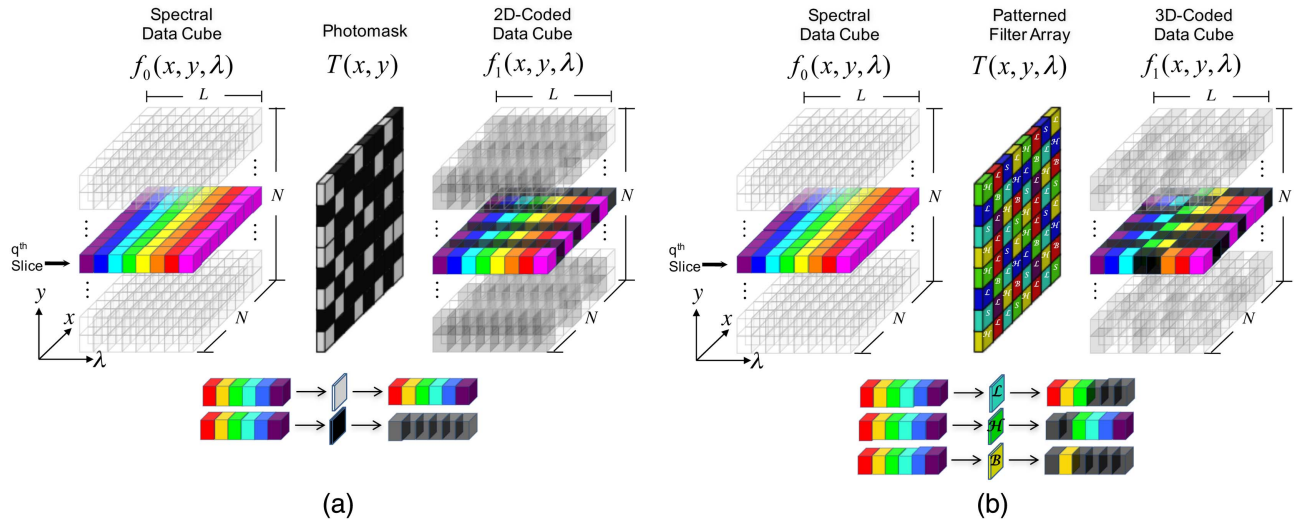


Fig. 1. Comparison of the coding strategies. (a) Photomask coding. (b) Patterned filter array coding. Filter array enables 3D coding due to the spectral response of the pixelated filters, whereas the photomask performs 2D coding over the data cube.

encode the input data. A dispersive element is then used to decompose the coded spectrum and a monochrome sensor captures the multiplexed data [1,2]. In contrast, a 3D patterned filter array is used as the coding element in the proposed optical system, replacing the commonly used photomask, as depicted in Fig. 1. The 3D coding is enabled by the 2D array of micro-optical pixelated filters, each of which may be a low-pass, high-pass, band-pass, stop-band or dichroic filter. The dispersive element can be a reflective element, such as a grating, or a transmissive element, such as a prism. The detection element can be a grayscale 2D detector, a line-detector, or even a single-pixel detector. In particular, the optical system described here uses a custom double Amici prism and a 2D grayscale focal plane array (FPA).

The compressive coded projections of the spectral components of the scene are captured as follows. Denote the continuous spatio-spectral scene as $f_0(x,y,\lambda)$, the patterned filter transfer function as $T(x,y,\lambda)$, and the Amici prism dispersion function as $S(\lambda)$. Coding is realized at the image plane of the field focused by an imaging lens, thus creating the 3D-coded field $f_1(x,y,\lambda) = T(x,y,\lambda) f_0(x,y,\lambda)$. After coding, a relay lens transmits the coded scene through the Amici prism, dispersing the field along the horizontal axis, forming the image on the FPA detector. The integrated coded and sheared spectral field along the spectral range sensitivity of the detector (Λ) can be expressed as

$$g(x,y) = \int_{\Lambda} T(x,y,\lambda) f_0(x,y,\lambda) h(x - x' - S(\lambda), y - y') \mathrm{d}\lambda, \quad \textbf{(1)}$$

where $h(x - x' - S(\lambda), y - y')$ accounts for the optical impulse response of the system and the shifting entailed by the prism along the x axis. The coding transfer function of the patterned filter array can be modeled as

$$T(x,y,\lambda) = \sum_{i,j,k} T_{i,j,k} \mathrm{rect}\left(\frac{x}{\Delta_c} - i - \frac{1}{2}, \frac{y}{\Delta_c} - j - \frac{1}{2}, \frac{\lambda}{\Delta_d} - k - \frac{1}{2}\right), \quad \textbf{(2)}$$

where $T_{i,j,k} \in [0,1]$ represents the filtering operation to be performed on the (i,j,k)th data cube voxel; $i,j \in \{0,\ldots,N-1\}$ and $k \in \{0,\ldots,L-1\}$ index the coordinates of an $N \times N \times L$ spectral data cube; and Δ_c, Δ_d account for the pixel sizes of the pixelated filters and the FPA detector, respectively. Notice that the spatial resolution of the resolvable scene is determined by the pitch size of the pixelated filters (Δ_c), while the spectral resolution depends on the pitch size of the FPA pixels (Δ_d) and the dispersion efficiency of the prism $S(\lambda)$ [14].

Let each voxel of the spectral scene be expressed as $F_{i,j,k} = \iiint_{\Omega_{i,j,k}} f_0(x,y,\lambda)\mathrm{d}x\mathrm{d}y\mathrm{d}\lambda = w_{ijk} f_0(x_i, y_j, \lambda_k)$, where $\Omega_{i,j,k}$ represents the (i,j,k)th voxel boundaries, and $w_{i,j,k}$ the voxel mass center weight. Assuming that the response of the dispersive element is linear over the spectral range of the system, and that ideal optical elements are used, the measurement attained on the (n,m)th pixel of the proposed system is given in discrete form by

$$g_{nm} = \sum_k F_{n,m-k,k} T_{n,m-k,k}, \quad \textbf{(3)}$$

where $n \in \{0,\ldots,N-1\}$ and $m \in \{0,\ldots,N+L-1\}$ index the pixels on the rows and columns of the detector, respectively.

A. Multiple-snapshot Matrix Model

A single snapshot may not be enough to attain a certain required quality in the reconstructed data cubes. The proposed system is then extended to include multiple snapshots. Multiple-snapshot sensing leads to less ill-posed inverse problems and consequently improved signal recovery [1]. In such a case, the patterned filter array must change its coding pattern with every new snapshot. As we will describe later, we used a wafer with numerous patterned filter arrays, contiguously placed to each other, so that a new coding pattern is attained by mechanically shifting the wafer. Multiple-snapshot sensing can be modeled by $g^{\ell}_{n,m} = \Sigma_k F_{n,m-k,k} T^{\ell}_{n,m-k,k}$, where $\ell = 0,\ldots,K-1$ accounts for K snapshots, attained by K spatial translations of the patterned filter array wafer.

Denote the kth spectral band of the input spectral data cube as $\mathbf{f}_k \in \mathbb{R}^{N^2}$, and let each ℓth patterned filter array in vectorial form as $\mathbf{t}_k^\ell \in \mathbb{R}^{N^2}$, such that $(\mathbf{f}_k)_z = F_{(z-\lfloor \frac{z}{N} \rfloor N), \lfloor \frac{z}{N} \rfloor, k}$, and $(\mathbf{t}_k^\ell)_z = T^\ell_{(z-\lfloor \frac{z}{N} \rfloor N), \lfloor \frac{z}{N} \rfloor, k}$, respectively, for $z = 0, 1, \ldots, N^2$. Also, let $\mathbf{g}^\ell$ be a column vector holding all of the recorded pixel values, $g^\ell_{n,m}$. Given the lateral dispersion of light, the length of each compressive measurement $\mathbf{g}^\ell$ is $N(N + L - 1)$ such that all light is accounted for on the sensor. Thus, while the patterned filter array is an $N \times N$ pixel array, the sensor has an extra set of L pixel columns. Assuming that $\mathbf{f}$ stays constant over the K snapshots, we can relate $\mathbf{g}^\ell$ to $\mathbf{f}$ in matrix form as

$$\mathbf{g}^\ell = \mathbf{H}^\ell \mathbf{f} = \mathbf{P}\mathbf{T}^\ell \mathbf{f}, \tag{4}$$

where $\mathbf{f} = [\mathbf{f}_0^T \mathbf{f}_1^T \ldots \mathbf{f}_{L-1}^T]^T$, and $\mathbf{H}^\ell$ is the system transfer function of the ℓth snapshot, which accounts for the coding pattern $\mathbf{T}^\ell$ and the dispersion function of the prism $\mathbf{P}$, which remain constant for all the snapshots. Let $Q = N^2 L$ be the total number of data cube voxels, and $U = N(N + L - 1)$ be the total number of FPA pixels. Therefore, $\mathbf{H}^\ell \in \mathbb{R}^{U \times Q}$, $\mathbf{P}$ is an $U \times Q$ rectangular matrix accounting for the shearing of the prism and the integration performed by the FPA. The structure of the matrix $\mathbf{P}$ is given by

$$\mathbf{P} = \begin{bmatrix} \mathrm{diag}(\mathbf{1}_{N^2}) & \mathbf{0}_{N\times N^2} & \cdots & \mathbf{0}_{N\times N^2} \\ \mathbf{0}_{N\times N^2} & \mathrm{diag}(\mathbf{1}_{N^2}) & \cdots & \mathbf{0}_{N\times N^2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N\times N^2} & \mathbf{0}_{N\times N^2} & \cdots & \mathrm{diag}(\mathbf{1}_{N^2}) \end{bmatrix}_{U\times Q}, \tag{5}$$

where $\mathbf{0}_{N\times N^2}$ is a 0-valued $N \times N^2$ matrix, and $\mathrm{diag}(\mathbf{1}_{N^2})$ is a 1-valued $N^2 \times N^2$ diagonal matrix. Notice that the number of $\mathbf{0}_{N\times N^2}$ sub-matrices in each row or column of Eq. (5) is $L - 1$, and there is just one $\mathrm{diag}(\mathbf{1}_{N^2})$ sub-matrix per row or column. Similarly, $\mathbf{T}^\ell$ is a $Q \times Q$ block-diagonal matrix, where each $N^2 \times N^2$ diagonal sub-matrix accounts for the effect of the ℓth patterned filter array on the kth data cube spectral band, as

$$\mathbf{T}^\ell = \begin{bmatrix} \mathrm{diag}(\mathbf{t}_0^\ell) & \mathbf{0}_{N^2\times N^2} & \cdots & \mathbf{0}_{N^2\times N^2} \\ \mathbf{0}_{N^2\times N^2} & \mathrm{diag}(\mathbf{t}_1^\ell) & \cdots & \mathbf{0}_{N^2\times N^2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N^2\times N^2} & \mathbf{0}_{N^2\times N^2} & \cdots & \mathrm{diag}(\mathbf{t}_{L-1}^\ell) \end{bmatrix}_{Q\times Q}, \tag{6}$$

where $\mathrm{diag}(\mathbf{t}_k^\ell)$ is an $N^2 \times N^2$ diagonal matrix, with diagonal elements $\mathbf{t}_k^\ell$.

The complete set of K snapshots in Eq. (4) can be assembled into a single vector by concatenating each $\mathbf{g}^\ell$ vector, end to end, to create the $KU \times 1$ vector $\mathbf{g} = [(\mathbf{g}^0)^T, \ldots, (\mathbf{g}^{K-1})^T]^T$, such that $\mathbf{g} = \mathbf{H}\mathbf{f}$, where $\mathbf{H} = [(\mathbf{H}^0)^T, \ldots, (\mathbf{H}^{K-1})^T]^T$ is a $KU \times Q$ matrix. Figure 2 shows a sketch of the $\mathbf{H}$ matrix for an $N \times N = 4 \times 4$ spectral data cube with L bands, using $K = 2$ snapshots, and a patterned filter array with L optical filters. The colored diagonals correspond to the filter array applied to each kth waveband, while the black entries correspond to blocking elements. The yellow circles in the figure highlight the variation of the coding elements along the spectral bands, satisfying the wavelength dependency of the filters within the coded pattern. Note that the two snapshots are vertically stacked, each one with size U, and the dispersion function is modeled by off-setting the diagonal structure of the patterned filters as the wavelength increases from left to right.

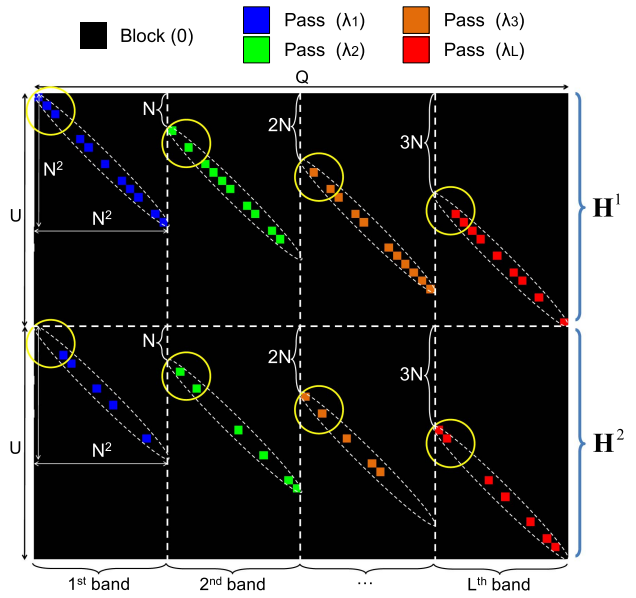


Fig. 2. Sketch of transfer function matrix $\mathbf{H}$ for a $4 \times 4 \times L$ spectral data cube, using $K = 2$ snapshots and a patterned filter array with L filters. $Q = N^2 L$ is the total number of data cube voxels, and $U = N(N + L - 1)$ is the total number of FPA pixels. Note that the coding pattern is wavelength-dependent, and a different distribution of its entries is used in each snapshot.

B. Input Data Cube Reconstruction

Compressive sensing (CS) states that the sensing process can be alternatively expressed as $\mathbf{g} = \mathbf{H}\mathbf{f} = \mathbf{H}\Psi\boldsymbol{\theta} = \mathbf{A}\boldsymbol{\theta}$, such that $\mathbf{A}$ is called the compressive sensing matrix, and $\boldsymbol{\theta} = \Psi^T \mathbf{f}$ is a vector coordinate of $\mathbf{f}$ in a Q dimensional space Ψ. This transformation assumes $\mathbf{f}$ to be S-sparse on Ψ, so that only a small subset, $S \ll Q$, of the basis vectors Ψ can accurately reconstruct $\mathbf{f}$ with little or no distortion. Formally, $\mathbf{f}$ is S-sparse or has sparsity S in a basis Ψ if $\|\boldsymbol{\theta}\|_0 = S$, where $\|\cdot\|_0$ denotes the ℓ_0 pseudo-norm, which simply counts the number of nonzero entries in the vector. In this way, an estimation of the spatio-spectral input data cube can be attained by solving the regularization problem

$$\hat{\mathbf{f}} = \Psi(\underset{\boldsymbol{\theta}'}{\mathrm{argmin}} \|\mathbf{g} - \mathbf{A}\boldsymbol{\theta}'\|_2^2 + \tau\|\boldsymbol{\theta}'\|_1), \tag{7}$$

where $\boldsymbol{\theta}'$ is the estimation of the vector coordinate of $\hat{\mathbf{f}}$ in Ψ, $\|\cdot\|_2^2$ represents the ℓ_2 norm, $\|\cdot\|_1$ the ℓ_1 norm, and τ is a regularization parameter that penalizes the fact of $\boldsymbol{\theta}'$ being sparse, while reducing the error between the estimated $\mathbf{A}\boldsymbol{\theta}'$, and the captured set of compressive measurements $\mathbf{g}$. To solve the optimization problem in Eq. (7), different algorithms have been proposed, including the gradient projection for sparse reconstruction (GPSR) [16], the GPSR-based but faster block-processing approach proposed in [17], the two-step iterative shrinkage/thresholding (TwIST) using the total-variation (TV) as the regularization function [18], Gaussian mixture models (GMM) [19], or the hyperspectral denosing algorithms based on approximate message passing (AMP) [20]. This paper

uses the method in [16], but any of the other methods could also be used.

3. DESIGN AND FABRICATION OF A PATTERNED FILTER ARRAY WAFER

A. Patterned Filter Array Design

The number of different filters to be used, the filters' spectral profile, and the way these filters are spatially distributed have to be carefully designed. First, if the filter array has $N \times N$ pixels of spatial resolution, then it is possible to have at most a different optical filter per pixel; that is, up to N^2 pixelated filters with different cut-off wavelengths. The higher the number of different filters, however, the higher the fabrication cost. Furthermore, we have observed in practice that the gains gradually decrease as the number of filters increases [15]. In fact, only a few different type of optical filters are typically needed to obtain significant improvement over imaging that uses conventional photomask technology. For example [15], only three filters are necessary to improve image reconstruction quality of an eight spectral band data cube by four decibels, in terms of peak signal-to-noise-ratio (PSNR). It is expected, however, that the number of filters used for improved performance should be larger as the spectral bands increase. Based on these results, and taking into account the fabrication costs, we used three filters: a low-pass (denoted by $\mathcal{L}(\lambda)$), a high-pass (denoted by $\mathcal{H}(\lambda)$) complementary filters with cut-off wavelength $\lambda_c = 550$ nm, and a blocking (all-stop, denoted by $\mathcal{B}(\lambda)$) filter, represented in Fig. 3 with colors blue, green, and black, respectively.

Second, the pixelated filters, or coded aperture entries in previous photomasks implementations, have been usually placed in an $N \times N$ grid using simply random spatial distributions. However, we have shown in previous publications that the spatial distribution of the entries in a coded aperture plays an important role in the quality of the reconstructions [1,21,22]. Therefore, in addition to the random distribution of the filters, in this paper we also test and analyze the performance of another two distributions called Boolean and optimized [14] distributions, as explained below.

A.1. Random Distributed Filters

Random distributions of block-unblock filters have been widely used in CASSI. These are justified if high incoherence is attained with the signal representation basis Ψ such as Fourier, wavelet, or cosine [23]. The entries of a coded aperture with random distributions can be modeled as

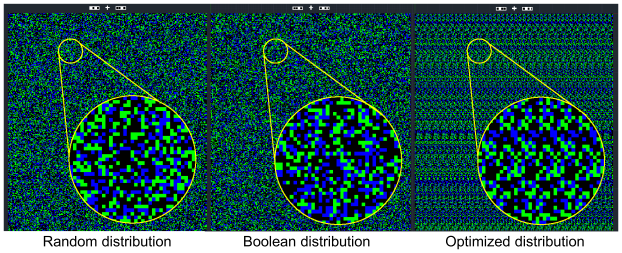


Fig. 3. Three-color patterned filter array design using three different spatial distributions. Blue pixels are low-pass filters, the green are high-pass filters, and the black are all-stop filters.

$$T_{i,j}^{\ell} = \begin{cases} \mathcal{B}(\lambda), & p \le \frac{1}{3} \\ \mathcal{L}(\lambda), & \frac{1}{3} < p \le \frac{2}{3}, \\ \mathcal{H}(\lambda), & p > \frac{2}{3} \end{cases} \tag{8}$$

where p is a random variable with the distribution function $U[0, 1]$. Notice that $T_{i,j}^{\ell_1}$ and $T_{i,j}^{\ell_2}$ $(\ell_1 \neq \ell_2)$ are independently generated from each other.

A.2. Boolean Distributed Filters

Boolean distributions keep the properties of random distributions for a single snapshot. However, in multiple-shot sensing each realization is mutually complementary to the previous realizations, thus guaranteeing that each voxel of the data cube is captured only once by the FPA. This characteristic is not guaranteed by random distributions, where each realization is generated independently of the others; therefore, a single voxel could be sensed more than once, or even worse, a single voxel may be always filtered out. Boolean distributions can be modeled as in Eq. (8), but satisfying the mutually complementary constraint $\Sigma_{\ell=0}^{K-1} T_{i,j}^{\ell} = A(\lambda)$, where $A(\lambda)$ is an all-pass filter, which represents the full spectrum of interest. For instance, in a $K = 3$ snapshot sensing, if the (i, j)th filter in the first snapshot is a low-pass $(T_{i,j}^{0} = \mathcal{L}(\lambda))$, then to satisfy the constraint, the second snapshot must use a high-pass filter, and the third snapshot a blocking filter, or vice versa $(T_{i,j}^{1} = \mathcal{H}(\lambda)$ and $T_{i,j}^{2} = \mathcal{B}(\lambda)$; or $T_{i,j}^{1} = \mathcal{B}(\lambda)$ and $T_{i,j}^{2} = \mathcal{H}(\lambda))$.

A.3. Optimized Spatial Distribution of the Filters

We recently proposed and demonstrated an optimization of the distribution of the entries of a patterned filter array [14]. This optimization extracts the benefits of Boolean distributions, but in addition, it exploits the correlations created by the structure of the CS matrix $\mathbf{A}$, such that the restricted isometry property (RIP) is better satisfied. In particular, the optimization of the distribution of the entries can be expressed as

$$\begin{aligned} &\underset{T_{i,0}^{\ell}, T_{i,1}^{\ell} \ldots, T_{i,L-1}^{\ell}}{\operatorname{argmin}} \sum_{k=1}^{L} c_k \gamma_k + \tau_1 U \\ &\text{subject to} \sum_{\ell=0}^{K-1} T_{i,j}^{\ell} = A(\lambda), \end{aligned} \tag{9}$$

for $i = 1, \ldots, N$, where, c_k are the entries of a penalty cost vector, τ_1 is a regularization constant, γ_k forces it to reduce the correlations along the spectral domain, U attempts to reduce the correlations along the spatial domain, and the constraint guarantees the mutual-complementariness along the K multiple-snapshots. Intuitively, the optimization problem seeks to spread as far as possible the passing bands of the pixelated filters along the spectral and the spatial domains, exploiting the correlations entailed by the sensing structure of the optical system. For further detailed information, refer to [14].

Figure 3 shows a single realization of the three spatial distributions as seen in the final design of the wafer, which was developed using AutoCAD. Notice that the random and the Boolean realizations present a pure random distribution of the entries, whereas the optimized realization exhibits a pseudo-random structure since just a single L-long section is designed per row and then horizontally replicated.

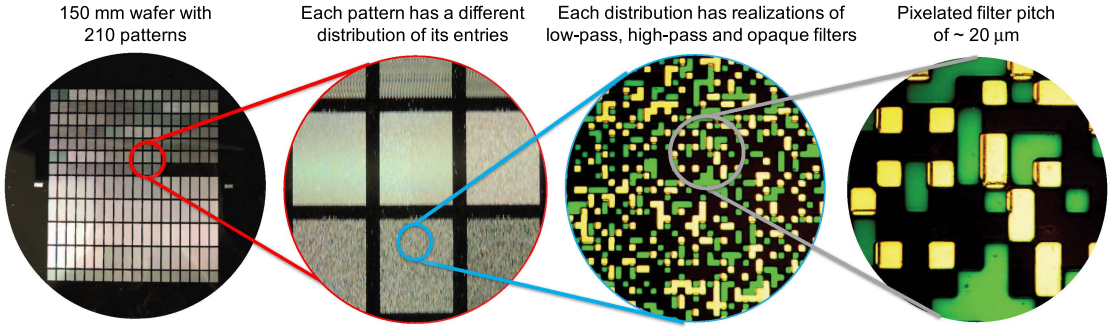
B. Thin-film-based Patterned Filter Array Fabrication

The fabrication process was realized by Pixelteq Inc., and it mainly consisted on deposition and lithographic patterning cycles of the three filters, in a random, Boolean, or optimized fashion [6]. The low-pass filter has an average transmittance of 93% between 350 and 540 nm, and 0.25% between 560 and 750 nm. The high-pass filter has an average transmittance of 0.04% between 350 and 540 nm, and 96% between 560 and 750 nm, as shown in Fig. 4(b). The blocking pixel has a 0.04% transmittance in the entire domain between 350 and 750 nm. The array was fabricated on a 150 mm diameter wafer of Eagle XG substrate, with a 1.1 mm thickness. Each patterned filter array in the wafer was designed to exhibit $N \times N = 256 \times 256$ pixels of spatial resolution, where each pixelated filter has a pixel-pitch size of $\Delta_c = 20$ μm. Due to the micrometer size of the pixelated filters, they are created by multilayer stacks of high- and low-index materials via plasma-assisted deposition or magnetron sputtering [6].

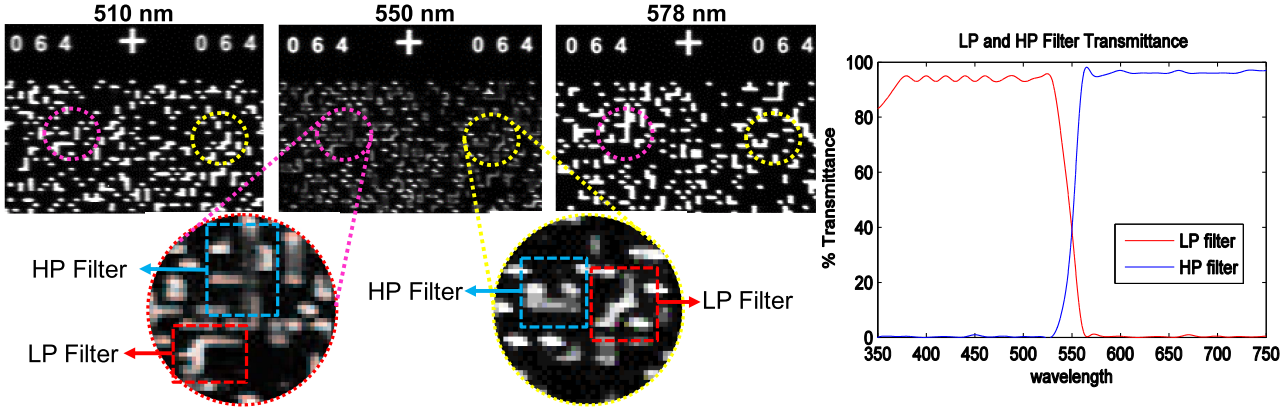
The final product is shown in Fig. 4(a). It contains 210 patterned filter arrays, and zoomed versions show the fine details within the full wafer. The 210 patterned arrays are different realizations of the spatial distributions for a different number of snapshots. Note that the pixelated filters exhibit sloped edges, since the lift-off fabrication process requires the lithography to be removed and the coating in the pixel to remain. These sloped edges cause areas of non-uniformity in the transmission properties of the coatings. Consequently, fine striped chrome areas are used around the pixelated filters, so that only the uniform portion is projected. If these opaque areas are not used, the effective pixel performance will suffer. To analyze the transition band of the pixelated filters, Fig. 4(b) shows three images of a portion from a single patterned filter array illuminated with three different wavelengths (510 nm, 550 nm, and 578 nm). It can be seen that when the pattern is illuminated with monochromatic light at 510 nm, just the low-pass pixelated filters let the light pass; similarly, when it is illuminated with monochromatic light at 578 nm, just the high-pass pixelated filters let the light propagate. Note that when the patterned array is illuminated with monochrome light at the cutoff wavelength of the filters ($\lambda_c = 550$ nm), both high- and low-pass filters let light go through the system, but its intensity is attenuated around 40%, agreeing with the given transition characteristics.

4. OPTICAL TESTBED IMPLEMENTATION AND EXPERIMENTS

To test the performance of the patterned filter array wafer, the CSI testbed setup presented in Fig. 5 was assembled in our laboratory. A colorful target object was attached to a black surface and illuminated by a broadband Xenon lamp playing the role of ambient light. The wafer of patterned filter arrays was placed at the image plane of the objective lens (Canon EF 50 mm f/1.8 II).



(a) Zoomed versions of the wafer are shown from left to right.



(b) The HP and LP filters appear in the transition band of 550 nm.

Fig. 4. Fabricated wafer of thin-film patterned filter arrays. (a) Microscope was used to capture zoomed versions of the wafer shown from left to right. (b) Analysis of the transition band of the filters. Due to the cutoff wavelength of the filters ($\lambda_c = 550$ nm), both HP and LP filters show up attenuated when the patterned filter is illuminated at that wavelength.

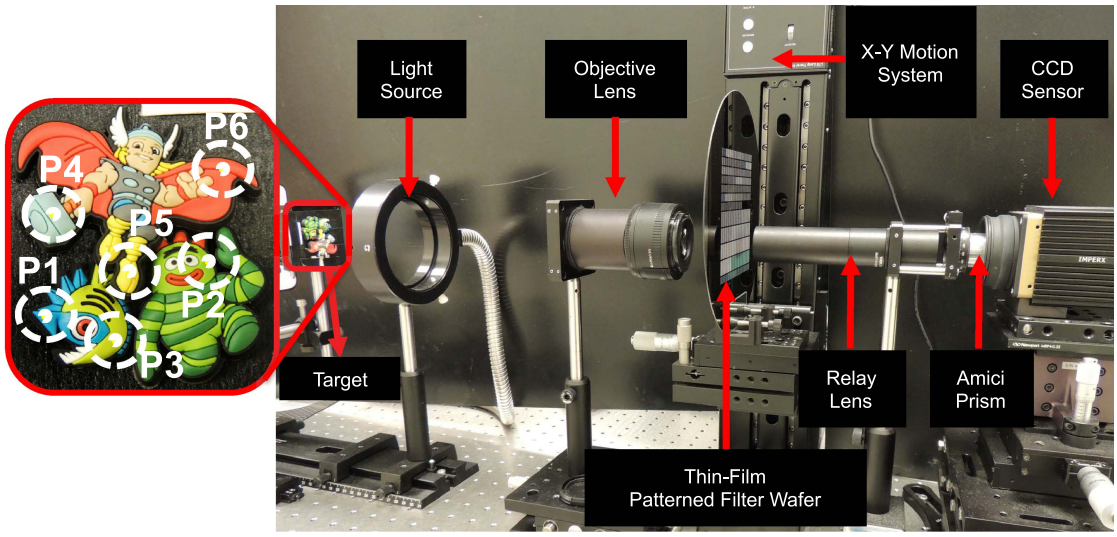


Fig. 5. Thin-film-based compressive spectral imaging testbed. The objective lens images the captured scene onto the thin-film patterned filter wafer, then the relay lens transmits the coded light through the Amici prism before it is integrated by the CCD sensor. The X–Y motion system moves the wafer precisely to access different coding patterns.

Two Thorlabs LTS300 linear translation stages with integrated stepper motor controller were used to precisely move the wafer along the x–y axis, to capture different patterned filter arrays. Each independent linear stage exhibits a velocity of 50 mm/s, on-axis accuracy of 5 μm, and a bidirectional repeatability of 2 μm. A matched achromatic doublet pair lens with 100 mm focal length was used to relay the 3D coded field through the custom designed double Amici prism, which horizontally spreads a single voxel of the data cube into 42 different columns of the FPA. The FPA sensor is a Bobcat B2021 CCD GigE Vision (with PoE), exhibiting a pixel pitch of $\Delta_d = 7.4$ μm and a resolution of 2048 × 2048. A 3 × 3 macro-pixel windowing was realized on the FPA to have an one-to-one correspondence with each pixelated filter; that is, $\Delta_d \approx 3\Delta_c$. Therefore, the attainable resolution in the reconstructions is $N \times N = 256 \times 256$ pixels in the space, and $L = 14$ spectral bands.

To obtain the estimations of the input data cubes, the gradient projection for the sparse reconstruction (GPSR) algorithm [16] was used to solve the optimization problem in Eq. (7). It is important to notice, however, that different strategies exist to do

Fig. 6. Full data cube reconstructions when random, Boolean, and optimized patterned filter arrays are used, and $K = 2$ snapshots are captured. The 14-band data cubes are mapped to false-colored RGB profiles to visually compare the color fidelity with the original target scene.

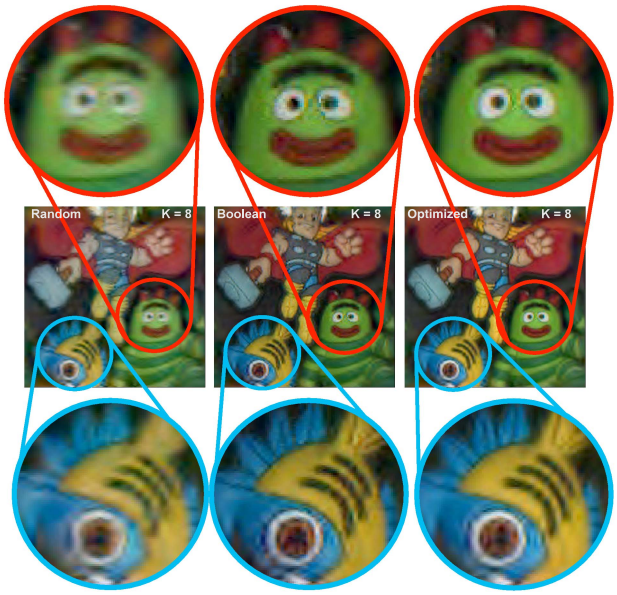


Fig. 7. Full data cube reconstructions when random, Boolean, and optimized patterned filter arrays are used, and $K = 8$ snapshots are captured. The 14-band data cubes are mapped to false-colored RGB profile to visually compare the color fidelity with the original target scene.

this estimation, as mentioned in Section 2.2. GPSR has been broadly used with traditional basis functions Ψ such as the wavelet (Ψ_{W2D}) and discrete cosine transforms (Ψ_{DCT}), as well as pre-trained dictionaries [1,17,19]. However, it is also possible to simultaneously exploit the sparsity properties of the input data cube along each of its dimensions to provide a

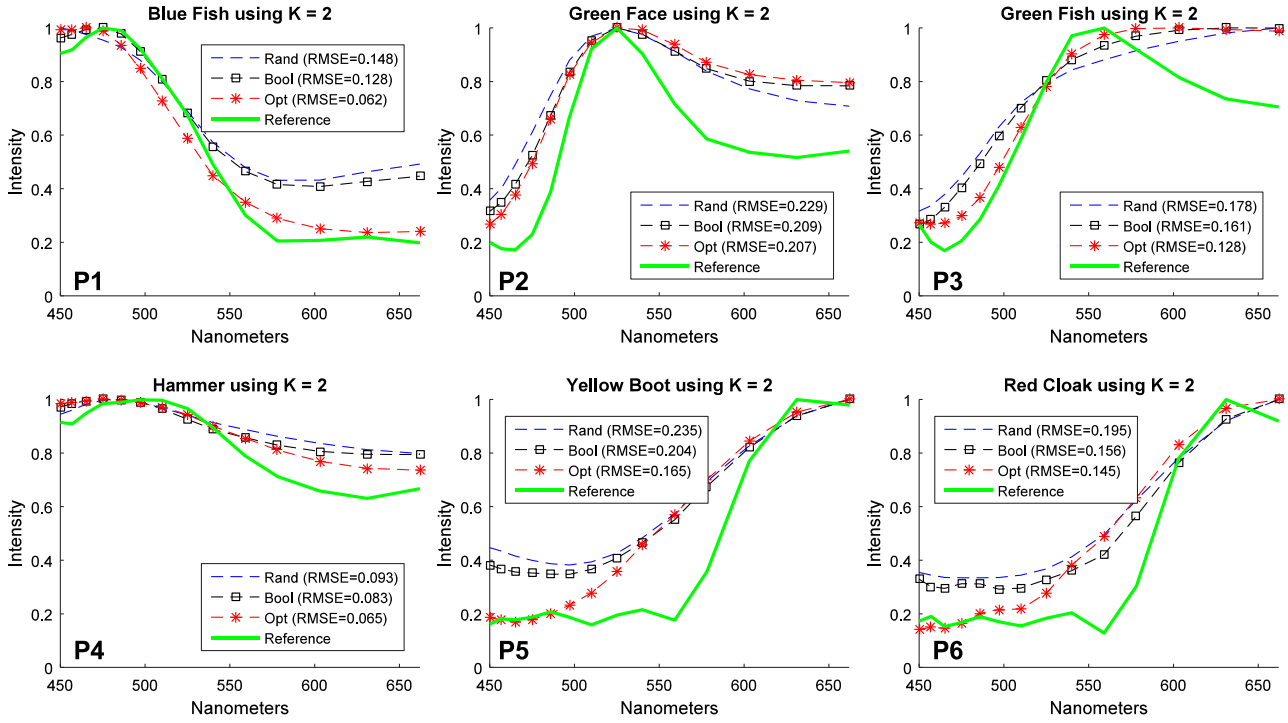


Fig. 8. Comparison of the spectrum reconstructions when $K = 2$ snapshots are captured. Six different points (P1, P2, P3, P4, P5, and P6) from the target scene (Fig. 5) were measured by a spectrometer and compared against the reconstructed data cubes.

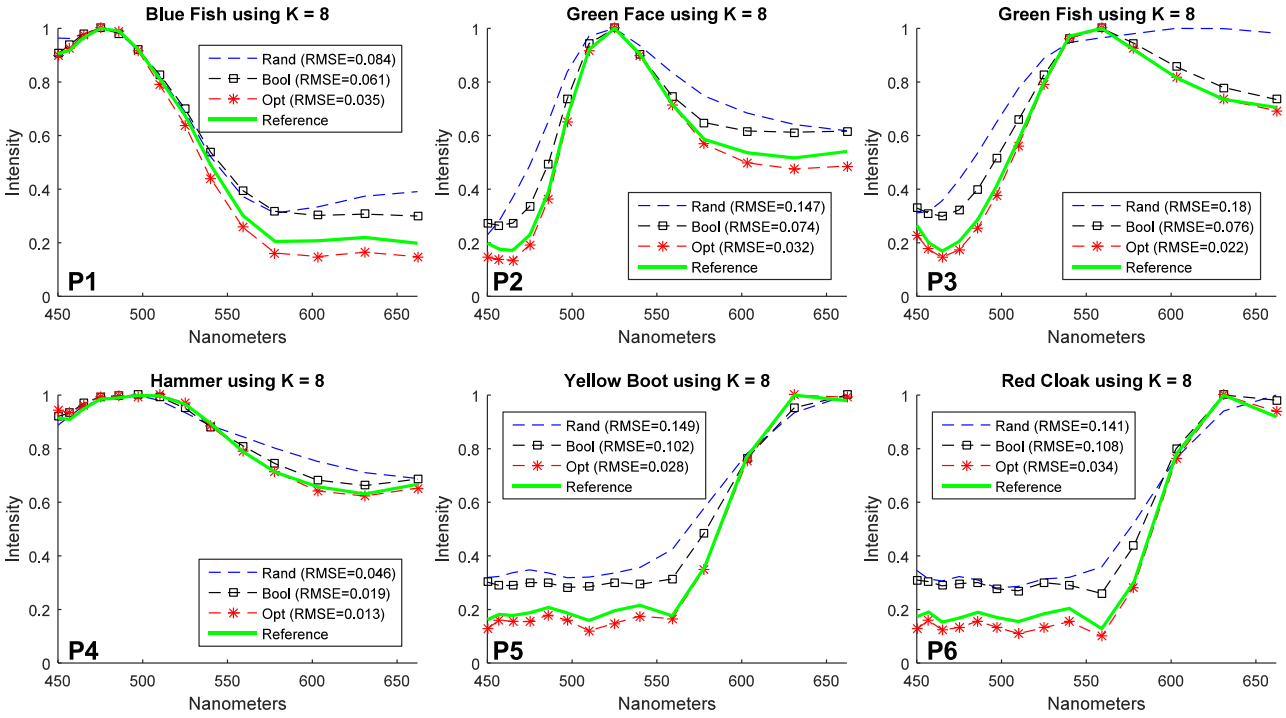


Fig. 9. Comparison of the spectrum reconstructions when $K = 8$ snapshots are captured. Six different points (P1, P2, P3, P4, P5, and P6) from the target scene (Fig. 5) were measured by a spectrometer and compared against the reconstructed data cubes.

new representation for their structure. Doing so, a single sparsity transform for the entire data cube is obtained as the Kronecker product of the bases used for each of its dimensions [4,15,24]. Particularly, we used the basis expressed as the Kronecker product between Ψ_{W2D} and Ψ_{DCT}, where Ψ_{DCT} is used to sparsify the spectral axis while Ψ_{W2D} sparsifies the spatial coordinates.

A single reconstruction with GPSR takes around 90 s to run 300 iterations for $K = 1$, and up to 210 s for $K = 8$ snapshots, when the data is processed in Matlab R2015b on an Intel Xeon CPU E5-1660 v3 at 3.00 GHz and 80GB of RAM. Figures 6–11 summarize the reconstruction results. In these figures, reconstruction results for $K = 2$ (compression ratio of $1/7$) and $K = 8$ (compression ratio of $4/7$) snapshots are shown, for the random, Boolean, and optimized spatial filter distributions.

First, in Figs. 6 and 7 the quality of the spatial reconstructions is evaluated by mapping the set of 14 reconstructed spectral bands to an RGB profile. It can be noticed that the Boolean distribution of the filters attains better quality than the random generated filters, and the optimized reconstructions exhibit the best quality between the three, for both $K = 2$ and $K = 8$ snapshots.

Second, to evaluate the reconstruction results at each wavelength level, Figs. 8 and 9 present the spectral footprint of six different points of the target scene, denoted as P1, P2, P3, P4, P5, and P6 in Fig. 5, which are compared against a reference spectrum measured with an Ocean Optics USB2000+ spectrometer, assumed to be the ground truth. All spectra are averaged in a 5×5 window corresponding to the same color and normalized to the maximum value in their respective curves. The rms errors between the reconstructed and the reference spectra are included in each subplot for ease of interpretation. The curves attained with the optimal (Opt) distribution present a better fit to the reference spectrum than the Boolean (Bool) and random (Rand) reconstructions, and there is an even better fit when the number of captured snapshots increases.

Finally, Figs. 10 and 11 show seven out of the 14 reconstructed spectral bands, false-colored with the respective wavelength, for $K = 2$ and $K = 8$, respectively. Notice in these figures that with only two snapshots the reconstructions lack details. When eight snapshots are captured, however, the quality of the reconstructions is quite improved and improves even more when either the Boolean or the optimized distribution of the pixelated filters is employed. The corresponding band/color discrimination of the monochrome wavelengths can be contrasted with the RGB target scene on Fig. 5, where the red cloak of the top toy is well characterized by the 662 nm wavelength, the blue fish fins are well resolved at 486 nm, and the green body of the right toy is well reconstructed at 525 nm.

5. DISCUSSION

This paper focuses on the implementation and analysis of patterned filter arrays into a compressive spectral imaging architecture. We did not include comparisons against the use of photomasks in our setup since we already demonstrated in our previous results that the use of patterned filter arrays offers better reconstruction results even just from simply randomly generated pixelated filter patterns [14,15]. In addition, we preferred to compare different distributions of the pixelated filters to evaluate their impact on the reconstruction quality. It was shown that either Boolean or optimized distributions attained better reconstruction results than simply using random distributions. Note that further optimization on the distribution of the entries

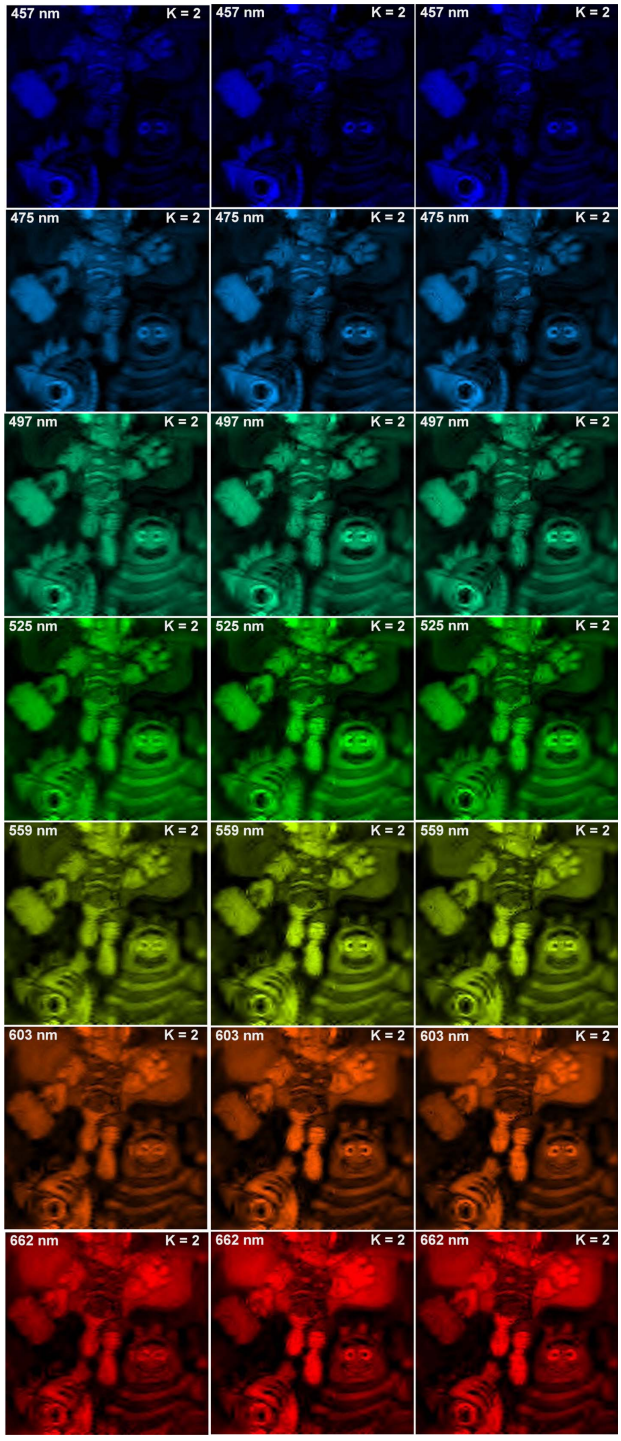


Fig. 10. Comparison of the reconstructions at monochrome level. The 2nd, 4th, 6th, 8th, 12th, and 14th spectral bands for the three different distributions are shown for $K = 2$ snapshots. (First column) random, (Second column) Boolean, and (Third column) optimized distribution.

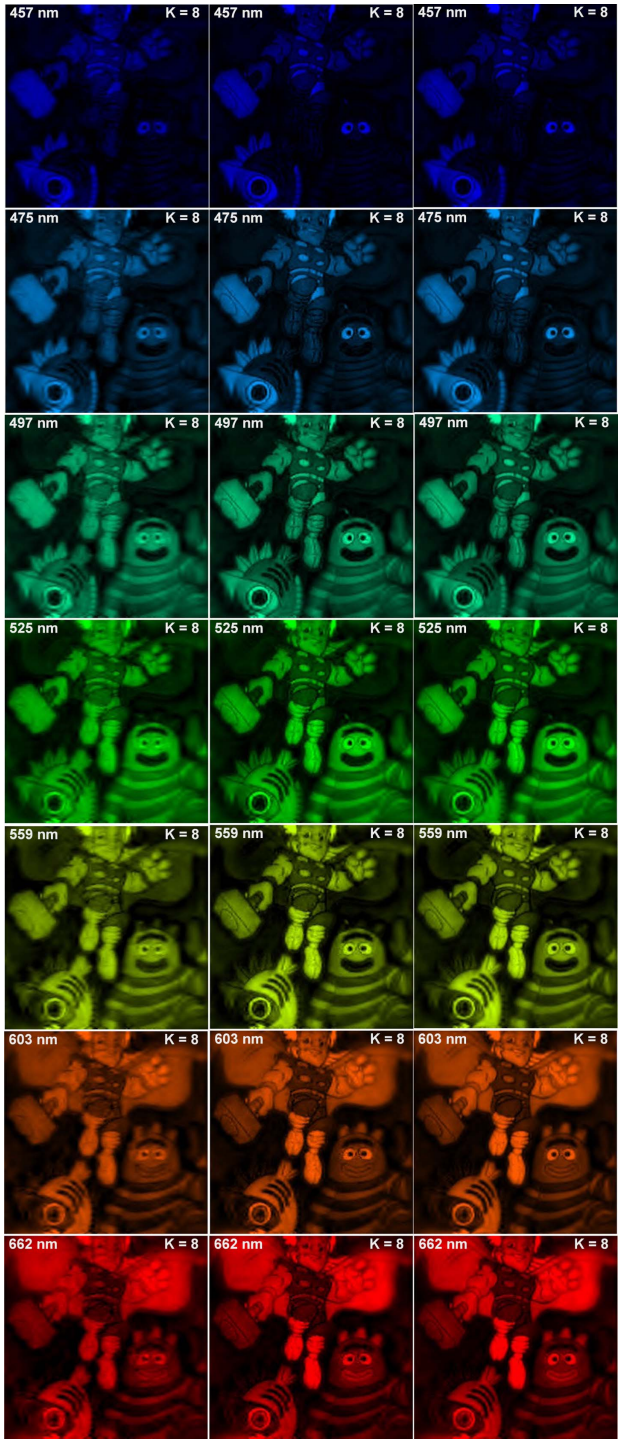


Fig. 11. Comparison of the reconstructions at monochrome level. The 2nd, 4th, 6th, 8th, 12th, and 14th spectral bands for the three different distributions are shown for $K = 8$ snapshots. (First column) random, (Second column) Boolean, and (Third column) optimized distribution.

can be still performed, and a higher-order precision model of the continuous sensing model [25] can be used to achieve even better quality results.

We believe we have shown that thin-film technology opens a door to higher-quality and higher-compression-sensing methodologies. Our group has recently demonstrated the use of patterned filter arrays attached to FPAs [26]. Although this approach entails a more compact architecture in terms of size, weight, and power, it only permits single snapshot sensing since the filter array is glued to the FPA surface. However, it can be considered a first step toward the use of conventional dyed-gel-based technology, which has been demonstrated to be both low cost and widely used [27].

In future work, it would be useful to perform a comparison of the use of dyed-gel filters, which are less costly but exhibit a wider extinction ratio, and the impact of their use in coded apertures in compressive spectral imaging. We expect that dyed-gel-based filter arrays will cause a rougher filtering due to the wide response of the gels, and the spectral results, therefore, will not be as good (sharp) as the ones shown here. That is, the wider the response of the filter, the closer the behavior to black-and-white (photomasks) coded apertures. However, the use of dyed gels could help expand the use of this technology because of the cheaper cost. An alternative solution is to use spatial light modulators (SLMs) such as liquid-crystal-based modulators, which can be exploited to perform spatio-spectral coding by tuning each mirror with a different voltage, thus permitting researchers to attain infinite patterns of higher number of filters just by tuning different voltages [3]. This will entail a compromise between no moving parts and the power required to manipulate the SLM.

6. CONCLUSIONS

This paper experimentally demonstrated and analyzed a thin-film-based patterned filter array compressive spectral imaging system. The use of a patterned filter array entailed a greater flexibility in the coding step of the imaging system, allowing us to independently encode the space and the spectrum of the input data cube. This flexibility offered a higher degree of freedom in the design of the sensing operator **H**. This approach, in turn, better satisfied the conditions of the compressive sensing theory, such as the RIP, therefore attaining higher-quality reconstructions.

The spatial distribution of the thin-film filters affects the quality of the attained reconstructions since the optimized distribution outperformed the simply random and Boolean distributions. The experimental results support the promising imaging capabilities introduced in our previous work [14,15], where imaging conditions, such as the variability of the cutoff wavelengths of the color filters and the measurement noise, have also been discussed.

Funding. National Science Foundation (NSF)/Intel Partnership on Visual and Experiential Computing grant (VEC 1538950); Universidad Industrial de Santander (UIS) (VIE 1891); Colciencias-Fulbright scholarship.

Acknowledgment. Hoover Rueda is supported by a Colciencias-Fulbright scholarship.

REFERENCES

1. G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "An introduction to compressive coded aperture spectral imaging," IEEE Signal Process. Mag. **31**(33), 105–115 (2014).

2. D. Kittle, K. Choi, A. A. Wagadarikar, and D. J. Brady, "Multiframe image estimation for coded aperture snapshot spectral imagers," Appl. Opt. **49**, 6824–6833 (2010).
3. T. Tsai, P. Llull, X. Yuan, L. Carin, and D. J. Brady, "Spectral-temporal compressive imaging," Opt. Lett. **40**, 4054–4057 (2015).
4. Y. August, C. Vachman, Y. Rivenson, and A. Stern, "Compressive hyperspectral imaging by random separable projections in both the spatial and the spectral domains," Appl. Opt. **52**, D46–D54 (2013).
5. 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. **36**, 2692–2694 (2011).
6. J. M. Eichenholz and J. Dougherty, "Ultracompact fully integrated megapixel multispectral imager," Proc. SPIE **7218**, 721814 (2009).
7. D. Knipp, R. A. Street, H. Stiebig, M. Krause, J. Lu, S. Ready, and J. Ho, "Vertically integrated thin film color sensor arrays for imaging applications," Opt. Express **14**, 3106–3113 (2006).
8. P. L. P. Dillon, A. T. Brault, J. R. Horak, E. Garcia, T. W. Martin, and W. A. Light, "Fabrication and performance of color filter arrays for solid-state imagers," IEEE Trans. Electron Devices **25**, 97–101 (1978).
9. S. J. Hood and S. G. Chamberlain, "Color-filter arrays for silicon solid-state image arrays sensors," Can. J. Phys. **69**, 543–548 (1991).
10. P. A. Stupar, R. L. Borwick, J. F. DeNatale, P. H. Kobrin, and W. J. Gunning, "MEMS tunable Fabry–Perot filters with thick, two sided optical coatings," in *Proceedings of International Solid-State Sensors, Actuators and Microsystems Conference* (IEEE, 2009), pp. 1357–1360.
11. D. Hays, A. Zribi, S. Chandrasekaran, S. Goravar, S. Maity, L. R. Douglas, K. Hsu, and A. Banerjee, "A hybrid MEMS-fiber optic tunable Fabry–Perot filter," J. Microelectromech. Syst. **19**, 419–429 (2010).
12. P. Buchsbaum, "Method of making dichroic filter arrays," U.S. Patent 5,711,889 (27 January 1998).
13. P. Buchsbaum, "Dichroic filter detector arrays for spectroscopic imaging," U.S. patent 6,638,668 (28 October 2003).
14. H. Arguello and G. R. Arce, "Colored coded aperture design by concentration of measure in compressive spectral imaging," IEEE Trans. Image Process. **23**, 1896–1908 (2014).
15. H. Rueda, H. Arguello, and G. Arce, "DMD-based implementation of patterned optical filter arrays for compressive spectral imaging," J. Opt. Soc. Am. A **32**, 80–89 (2015).
16. M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems," IEEE J. Sel. Top. Signal Process. **1**, 586–597 (2007).
17. H. Arguello, C. Correa, and G. R. Arce, "Fast lapped block reconstructions in compressive spectral imaging," Appl. Opt. **52**, D32–D45 (2013).
18. J. M. Bioucas-Dias and M. A. Figueiredo, "A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration," IEEE Trans. Image Process. **16**, 2992–3004 (2007).
19. A. Rajwade, D. Kittle, T. Tsai, D. Brady, and L. Carin, "Coded hyperspectral imaging and blind compressive sensing," SIAM J. Imaging Sci. **6**, 782–812 (2013).
20. J. Tan, Y. Ma, H. Rueda, D. Baron, and G. R. Arce, "Compressive hyperspectral imaging via approximate message passing," IEEE J. Sel. Top. Signal Process. **10**, 389–401 (2016).
21. H. Arguello and G. R. Arce, "Code aperture optimization for spectrally agile compressive imaging," J. Opt. Soc. Am. A **28**, 2400–2413 (2011).
22. H. Arguello and G. R. Arce, "Rank minimization code aperture design for spectrally selective compressive imaging," IEEE Trans. Image Process. **22**, 941–954 (2013).
23. E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," IEEE Signal Process. Mag. **25**(2), 21–30 (2008).
24. Y. Rivenson and A. Stern, "Compressed imaging with a separable sensing operator," IEEE Signal Process. Lett. **16**, 449–452 (2009).
25. H. Arguello, H. Rueda, Y. Wu, D. W. Prather, and G. R. Arce, "Higher-order computational model for coded aperture spectral imaging," Appl. Opt. **52**, D12–D21 (2013).
26. C. V. Correa, H. Arguello, and G. R. Arce, "Snapshot colored compressive spectral imager," J. Opt. Soc. Am. A **32**, 1754–1763 (2015).
27. H. Rueda, D. Lau, and G. R. Arce, "Multi-spectral compressive snapshot imaging using RGB image sensors," Opt. Express **23**, 12207–12221 (2015).