



Compressive spectral imaging via deformable mirror and colored-mosaic detector

MIGUEL MARQUEZ,¹ PABLO MEZA,² HENRY ARGUELLO,³ AND ESTEBAN VERA^{4,*}

¹*Department of Physics, Universidad Industrial de Santander, Bucaramanga, Colombia*

²*Department of Electrical Engineering, Universidad de La Frontera, Temuco, Chile*

³*Department of Systems Engineering, Universidad Industrial de Santander, Bucaramanga, Colombia*

⁴*School of Electrical Engineering, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile*

* esteban.vera@pucv.cl

Abstract: Compressive Spectral Imaging (CSI) is an emerging technology that aims at reconstructing a spectral image from a limited set of two-dimensional projections. To capture these projections, CSI architectures often combine light dispersive elements with coded apertures or programmable spatial light modulators. This work introduces a novel and compact CSI architecture based on a deformable mirror and a colored-filter detector to produce compressive spatio-spectral projections without the need of a grating or prism. Alongside, we propose a tensor-based reconstruction algorithm to recover the spatial-spectral information from the compressed measurements. Experimental results on both simulated and real datasets demonstrate efficacy of the proposed acquisition architecture and the especially crafted inversion algorithms.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Spectral imaging (SI) has become a powerful analytical tool in various fields, such as remote sensing, agriculture, environmental monitoring and medicine [1–3]. The datasets acquired by a SI system are three-dimensional images that comprise a large amount of spatial information across a multitude of wavelengths. Conventional imaging spectroscopy scan adjacent zones of the underlying scene and combine the readings to construct spectral data cubes [4–6]. Specifically, these sampling techniques are based on pushbroom, whiskbroom or spotlight systems, which measure the signal at a certain constant sampling rate on each of the three dimensions [7]. As a result, these sensing techniques require scanning a number of slices of the datacube that grow linearly in proportion to the desired spatial-spectral resolution. Thus, these systems are impractical in dynamic or fast-moving scenarios.

To alleviate these difficulties, compressive spectral imaging (CSI) has emerged as an alternative acquisition framework based on the compressive sensing (CS) theory that captures and compresses simultaneously the relevant information embedded in the spectral image of interest [8]. CSI aims at recovering a spatio-spectral data cube by acquiring two-dimensional projections of encoded and shifted versions of the spatio-spectral source through exploiting the use of coded apertures and optical dispersive elements [7, 9]. In particular, the original coded aperture snapshot spectral imager (CASSI) paradigm was introduced by the seminal dual disperser (DD-CASSI) [10] architecture, which employed two dispersive elements and one coded aperture, leading to measurements that only encode the spectral content for each spatial position. As an evolutionary step, the single disperser (SD-CASSI) [11] has become one of the most popular spectral imaging sampling schemes since it uses only one coded aperture located at the image plane which is later dispersed by a single prism or grating before being projected onto a detector, producing a particular spatio-spectral coded projection onto the detector, but effectively reducing the calibration complexity in contrast with the DD-CASSI. All CASSI schemes rely on CS sparse reconstruction algorithms to recover the data cube by exploiting spatio-spectral redundancies in a transform domain such as Wavelet and DCT. Several variations of the SD-CASSI have

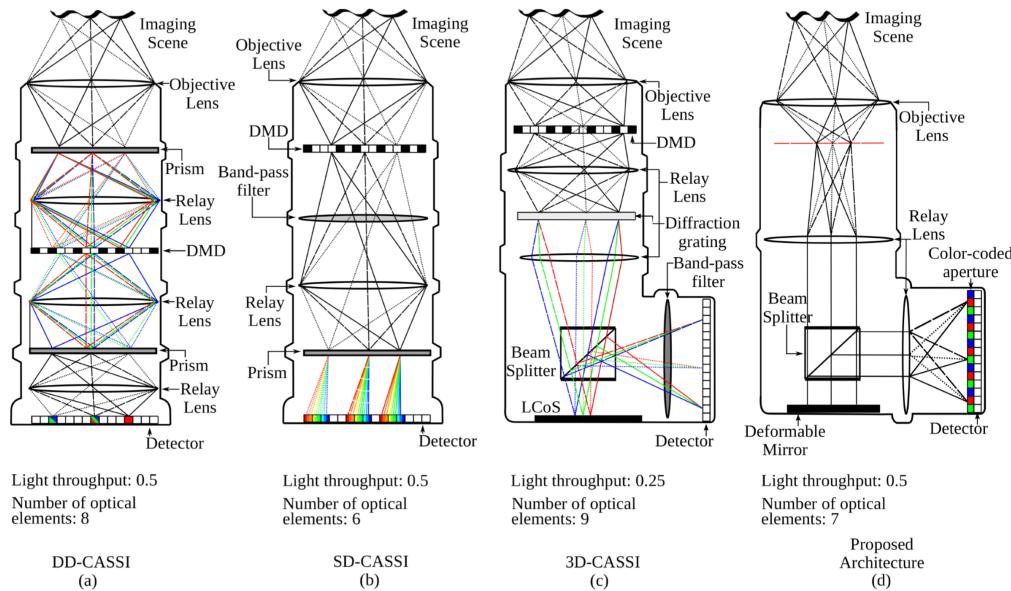


Fig. 1. Diagram of four multishot compressive multispectral cameras: (a) DD-CASSI, (b) SD-CASSI, (c) 3D-CASSI and (d) proposed architecture.

been proposed since its inception, such as the approach in [12] where the prism and coded aperture have been interchanged, generating a different class of spatio-spectral coding and using dictionary learning with overlapped patches for improving the reconstruction quality. Another modification to the CASSI systems that can be used to enhance the reconstruction quality is the use of dynamic optical elements such as a digital micromirror devices (DMDs), which can be used as programmable coded apertures enabling the possibility of taking different projections of the data cube. The schemes for the multishot DD-CASSI [10] and SD-CASSI [11] can be found in Fig. 1(a) and 1(b), respectively. Similarly, a spatial light modulator (SLM) based on LCOS technology enables the ability to produce predetermined spectral filters, and together with the use of the DMD has enabled the development of the colored coded aperture spectral imager (3D-CASSI) [13], providing a rich and flexible multishot spatio-spectral coding scheme, as seen in Fig. 1(c). It is worth noting that the reconstruction quality attained by the 3D-CASSI surpasses the results obtained by the DD-CASSI or SD-CASSI, at the expense of additional optical elements, more sensitivity to calibration errors, and a diminished light throughput.

In this regard, compressive spectral imaging (CSI) systems with more flexibility, versatility, better light throughput, and lower calibration complexity are highly desired [14]. Taking advantage of the simplicity of the RGB Bayer filter design, a static color-coded aperture placed in front of the detector array has been included as a robust solution to snapshot multispectral imaging [15], at the expense of light throughput from the filters. However, spatial resolution is severely affected as more spectral bands are demanded [11], which can be enhanced by properly designing the filters [16], which can also be forced to have a defined throughput. Nonetheless, the usefulness of colored filter arrays can be improved by adding a fixed grating [17] or prism [18]. Similar to the case of the multishot CASSI systems, the performance can be greatly improved by adding variability, which can be achieved by moving the dispersive element [19]. In order to not depend on a dispersive element, in [20] a DMD is used for dynamically coding the spatial domain before reaching an array of Fabry-Perot filters placed before the detector, although these filters are narrow and may significantly reduce the light throughput. Moreover, 3D-CASSI finally provides with a mixed solution where both spatial and spectral coding are dynamic; one by the

DMD while the other provided by the LCOS SLM, although not as arbitrary nor selective as the Fabry-Perot array. Nevertheless, it is well known that using a binary coded aperture scheme, such as what can be provided by a DMD, may jeopardize even further the diminished light throughput of a colored filter array. In this regard, deformable mirrors (DM) have become an attractive solution as programmable phase coding devices to perform wavefront modulation at the pupil plane to compensate for the effects of atmospheric turbulence in astronomy images [21], correct the aberrations of the eye [22], encoding depth-of-field [23], and performing compressive imaging [24], among others, with a low loss of radiant energy flow. In particular for compressive imaging [24], the DM provides with dynamic phase modulation capabilities that can also enable a multiframe coding approach.

In this work, we propose a novel multishot CSI system along with the mathematical framework required to efficiently reconstruct spectral images from compressed measurements. The optical architecture relies on a DM for the spatial coding, and an imaging sensor with a color-coded aperture for the spectral coding, as can be seen in Fig. 1(d). The main technical benefits of the Deformable Mirror (DM) is that it offers high speed modulation with straightforward calibration and high light throughput due to its controlled continuous phase modulation. Apart from allowing multiple shots, the proposed CSI system enables the design of spatial/spectral coding in arbitrary directions in contrast with CASSI systems. We also propose a methodology that includes the spatial-spectral multiplexing sensing protocol and an alternating direction method of multipliers (ADMM)-based tensor reconstruction algorithm.

2. Optical sensing model of the proposed architecture

Let $f_o(x, y, \lambda)$ denotes the 3D function that represents the spatial-spectral source density with (x, y) being the spatial coordinates and λ being the spectral dimension. Light is firstly transmitted through a beamsplitter to a deformable mirror (phase mask) located on the pupil plane, which allows adding controlled aberrations to $f_o(x, y, \lambda)$. This phase modulation can be expressed as

$$f_1(x, y, \lambda) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |h(x - x', y - y', \lambda)|^2 f_o(x', y', \lambda) dx' dy', \quad (1)$$

with

$$h(x, y, \lambda) = \frac{1}{\lambda z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{P}(x', y') e^{i2\pi W(x', y')} e^{-i\frac{2\pi}{\lambda z}(xx' + yy')} dx' dy', \quad (2)$$

where $h(x, y, \lambda)$ represents the PSF, $z \in \mathbb{R}$ the image distance, $\mathcal{P}(x, y)$ the pupil function, and $W(x, y)$ is the wavefront aberration function. Specifically, the wavefront can be expressed as $W(x, y) = \sum_{i=1}^{\infty} a_i Z_i(x, y)$, where $Z_i(x, y)$ represents the i -th Zernike polynomial in the Noll's notation and a_i its the aberration coefficient [25]. Although the DM is achromatic in nature note that the fact of using lenses in the system with potential chromatic aberrations introduce the need to model a spectral depended PSF. The code applied by the color-coded aperture to the spatio-spectral density is $f_1(x, y, \lambda)t(x, y, \lambda)$, with

$$t(x, y, \lambda) = \sum_{i_1, i_2, i_3} T_{i_1, i_2, i_3} \text{rect} \left(\frac{x}{\Delta_c} - i_1, \frac{y}{\Delta_c} - i_2, \frac{\lambda}{\Delta_d} - i_3 \right), \quad (3)$$

where $T_{i_1, i_2, i_3} \in [0, 1]$ is the coding performed on the $(i_1, i_2, i_3)^{th}$ spectral voxel, and Δ_c and Δ_d account for the pixel sizes of the color-coded aperture and focal plane array (FPA) detector, respectively. Furthermore, there exists one-to-one matching between the elements of the color filter array and those of the FPA; thus, the energy in one detector pixel is affected by just one optical filter element. The continuous image on the detector can be then expressed as

$$g(x, y) = \int_{\lambda_1}^{\lambda_2} \int_{-\Delta_d}^{\Delta_d} \int_{-\Delta_d}^{\Delta_d} t(x', y', \lambda) f_1(x', y', \lambda) dx' dy' d\lambda. \quad (4)$$

Equation (4) can be posed as a compressed sensing problem, which is traditionally modeled as the linear system $\mathbf{g} = \Phi\mathbf{f} + \epsilon$. In this sense, \mathbf{g} represents the discrete vectorized version of the measurements, \mathbf{f} is the discrete vectorized version of the spectral image, and $\Phi \in \mathbb{R}^{m \times n}$ is a matrix including the sensing transfer function with $n = NML$ and $m = N_1 N_2 K$, with K is the number of snapshots. Here, for each snapshot, the membrane of the deformable mirror is deflected according to an arbitrary Zernike polynomial $W(x, y)^{(k)} = \sum_{i=1}^{\infty} a_i^{(k)} Z_i$ with $k \in \{1, \dots, K\}$, where $W(x, y)^{(k)}$ represents the membrane deformation for the k -th snapshot. Such linear system requires the product between a big sparse matrix Φ with a vector, demanding a mathematical operation of order $O(mn)$. To alleviate this problem, we model the acquisition process Eq. (4) in tensor form, following the procedure described in Appendix 5, as

$$\underline{\mathbf{G}} = \left[\underline{\mathbf{T}} \circ \left(\prod_{i=1}^3 \left[\underline{\mathbf{W}} \circ \left(\left[\prod_{j=1}^3 \underline{\mathbf{F}} \times_j \mathbf{A}_j \right] \times_4 \mathbf{1} \right) \right] \times_i \mathbf{A}_i^T \right) \right] \times_3 \mathbf{1}^T \quad (5)$$

where $\underline{\mathbf{G}} \in \mathbb{R}^{M \times N \times 1 \times K}$ represents the measurements, $\underline{\mathbf{T}} \in \mathbb{R}^{M \times N \times L \times K}$ is the color-coded aperture, $\underline{\mathbf{W}} \in \mathbb{C}^{M \times N \times L \times K}$ corresponds to the wavefront aberration matrix, $\mathbf{A}_3 = \mathbf{I} \in \mathbb{R}^{L \times L}$ is an identity matrix, and matrices $\mathbf{A}_1 \in \mathbb{R}^{M \times M}$, $\mathbf{A}_2 \in \mathbb{R}^{N \times N}$ are the direct Discrete Fourier transform basis. Note here that the tensorial representation in Eq. (5) preserves the original structure of the data, critical for the analysis of their intrinsic properties. The conventional problem to obtain an estimation of $\underline{\mathbf{F}}$ from the set of compressed measurements $\underline{\mathbf{G}}$, considering Eq. (5), is given by

$$\arg \min_{\underline{\mathbf{F}}} \left\| \underline{\mathbf{G}} - \left[\underline{\mathbf{T}} \circ \left(\prod_{i=1}^3 \left[\underline{\mathbf{W}} \circ \left(\left[\prod_{j=1}^3 \underline{\mathbf{F}} \times_j \mathbf{A}_j \right] \times_4 \mathbf{1} \right) \right] \times_i \mathbf{A}_i^T \right) \right] \times_3 \mathbf{1}^T \right\|_F^2, \quad (6)$$

where $\|\cdot\|_F$ represents the Frobenius-norm. To promote correct solutions of $\underline{\mathbf{F}}$ from its compressive measurements $\underline{\mathbf{G}}$, CSI theory demands the use of sparsity promoting priors. Sparsity indicates that $\underline{\mathbf{F}}$ must be S -sparse in an arbitrary set of orthogonal basis $\Psi_j \in \mathbb{R}^{N_j \times N_j}$ with $j \in \{1, 2, 3\}$ and $\{N_1, N_2, N_3\} = \{M, N, L\}$, such that $\underline{\mathbf{F}} = \prod_{j=1}^3 \underline{\Theta} \times_j \Psi_j$, where $\underline{\Theta} \in \mathbb{R}^{M \times N \times L}$ is the sparse tensor that contains $S \ll MNL$ non-zero coefficients. Therefore, the reconstruction problem described in Eq. (6) can be rewritten as

$$\arg \min_{\underline{\Theta}} \left\| \underline{\mathbf{G}} - \left[\underline{\mathbf{T}} \circ \left(\prod_{i=1}^3 \left[\underline{\mathbf{W}} \circ \left(\left[\prod_{j=1}^3 \underline{\Theta} \times_j \mathbf{A}_j \Psi_j \right] \times_4 \mathbf{1} \right) \right] \times_i \mathbf{A}_i^T \right) \right] \times_3 \mathbf{1}^T \right\|_F^2 + \lambda \|\underline{\Theta}\|_1, \quad (7)$$

where $\|\cdot\|_1$ represents the ℓ_1 -norm.

3. Proposed reconstruction algorithm

The goal of the proposed algorithm, summarized in Algorithm 1, is to reconstruct Θ from its compressive measurements $\underline{\mathbf{G}}$ in Eq. (7). Algorithm 1 solves Eq. (7) via a proximal alternating optimization methodology which finds $\underline{\Theta}$ through

$$\begin{aligned} \underline{\Theta}^{\ell+1} &\leftarrow \arg \min_{\underline{\Theta}} \left\| \underline{\mathbf{G}} - \left[\underline{\mathbf{T}} \circ \left(\prod_{i=1}^3 \left[\underline{\mathbf{W}} \circ \left(\left[\prod_{j=1}^3 \underline{\Theta} \times_j \mathbf{A}_j \Psi_j \right] \times_4 \mathbf{1} \right) \right] \times_i \mathbf{A}_i^T \right) \right] \times_3 \mathbf{1}^T \right\|_F^2 \\ &\quad + \beta \|\underline{\Theta}^{\ell} - \underline{\Theta}\|_F + \lambda_1 \|\underline{\Theta}\|_1 + \lambda_2 \left\| \prod_{j=1}^3 \underline{\Theta} \times_j \Psi_j \right\|_{TV}, \end{aligned} \quad (8)$$

where $\|\cdot\|_{TV}$ represent the total variation regularization, $\lambda_1 > 0$, $\lambda_2 > 0$ and $\beta > 0$ are regularization parameters, where Eq. (8) is solved using an ADMM approach [26]. For this purpose,

Algorithm 1: Spectral image reconstruction

Data: $\underline{\mathbf{W}}, \underline{\mathbf{T}}, \underline{\mathbf{G}}, \mu$
Result: $\underline{\mathbf{F}} \in \mathbb{R}^{M \times N \times L}$

Initialization:

1. $t \leftarrow 0$,
2. $[\tilde{\underline{\Theta}}^t, \underline{\Omega}^t, \underline{\Xi}^t, \underline{\eta}_1^t, \underline{\eta}_2^t] \leftarrow \mathbf{0} \in \mathbb{R}^{M \times N \times L}$

repeat

3. Update $\underline{\Theta}^{t+1}$ via (10)
4. Update $\underline{\Omega}^{t+1}$ via (11)
5. Update $\underline{\Xi}^{t+1}$ via (12)
6. $\underline{\nu}_1^{t+1} = \underline{\nu}_1^t - (\underline{\Omega}^{t+1} - \underline{\Omega}^t)$
7. $\underline{\nu}_2^{t+1} = \underline{\nu}_2^t - (\underline{\Xi}^{t+1} - \prod_{j=1}^3 \underline{\Theta}^{t+1} \times_j \Psi_j)$
8. $t \leftarrow t + 1$
9. $\underline{\mathbf{F}} = \underline{\Xi}^t$

until some stopping criterion is satisfied;

return $\underline{\mathbf{F}}$

we introduce the following functions $\Gamma_1(\underline{\Theta}) = \left\| \underline{\mathbf{G}} - \left[\underline{\mathbf{T}} \circ \left(\prod_{i=1}^3 [\underline{\mathbf{W}} \circ \Gamma_2(\underline{\Theta})] \times_i \mathbf{A}_i^T \right) \right] \times_3 \mathbf{1}^T \right\|_F^2$, $\Gamma_2(\underline{\Theta}) = \left(\left[\prod_{j=1}^3 \underline{\Theta} \times_j \mathbf{A}_j \Psi_j \right] \times_4 \mathbf{1} \right)$, $\Gamma_3(\underline{\Omega}) = \beta \|\underline{\Omega}^t - \underline{\Omega}^t\|_F + \lambda \|\underline{\Omega}^t\|_1$ and $\Gamma_4(\underline{\Xi}) = \|\underline{\Xi}\|_{TV}$ with the splitting variables $\underline{\Omega} = \underline{\Theta}$ and $\underline{\Xi} = \prod_{j=1}^3 \underline{\Theta} \times_j \Psi_j$, such that Eq. (8) can be rewritten as

$$\arg \min_{\underline{\Theta} \underline{\Omega} \underline{\Xi}} \quad \Gamma_1(\underline{\Theta}) + \Gamma_3(\underline{\Omega}) + \Gamma_4(\underline{\Xi}) \quad \text{s.t.} \quad \underline{\Theta} = \underline{\Omega}, \quad \underline{\Xi} = \prod_{j=1}^3 \underline{\Theta} \times_j \Psi_j, \quad (9)$$

where its augmented Lagrangian is given by $\mathcal{L}(\underline{\Theta}, \underline{\Omega}, \underline{\Xi}, \underline{\nu}_1, \underline{\nu}_2) = \Gamma_1(\underline{\Theta}) + \Gamma_3(\underline{\Omega}) + \mu_1 \|\underline{\Theta} - \underline{\Omega} - \underline{\nu}_1\|_F^2 + \mu_2 \|\underline{\Xi} - \prod_{j=1}^3 \underline{\Theta} \times_j \Psi_j - \underline{\nu}_2\|_F^2$, where $\underline{\nu}_1$ and $\underline{\nu}_2$ are the Lagrange multipliers and $\{\mu_1, \mu_2\} \geq 0$ are called the AL penalty parameter [26]. Based on $\mathcal{L}(\underline{\Theta}, \underline{\Omega}, \underline{\Xi}, \underline{\nu}_1, \underline{\nu}_2)$, the alternative form of Eq. (9) can be decoupled in three independent optimization problems

$$\begin{aligned} \tilde{\underline{\Theta}}^{t+1} &= \arg \min_{\underline{\Theta}} \left\| \underline{\mathbf{Y}} - \left[\underline{\mathbf{T}} \circ \left(\prod_{i=1}^3 [\underline{\mathbf{W}} \circ \Gamma_2(\underline{\Theta})] \times_i \mathbf{A}_i^T \right) \right] \times_3 \mathbf{1}^T \right\|_F^2 + \mu_1 \|\underline{\Theta}^t - \underline{\Omega}^t - \underline{\nu}_1^t\|_F^2 \\ &\quad + \mu_2 \|\underline{\Xi}^t - \prod_{j=1}^3 \underline{\Theta}^t \times_j \Psi_j - \underline{\nu}_2^t\|_F^2, \end{aligned} \quad (10)$$

$$\underline{\Omega}^{t+1} = \arg \min_{\underline{\Omega}} \beta \|\underline{\Omega}^t - \underline{\Omega}^t\|_F^2 + \mu_1 \|\underline{\Theta}^t - \underline{\Omega}^t - \underline{\nu}_1^t\|_F^2 + \lambda \|\underline{\Omega}^t\|_1, \quad (11)$$

$$\underline{\Xi}^{t+1} = \arg \min_{\underline{\Xi}} \|\underline{\Xi}\|_{TV} + \mu_2 \|\underline{\Xi}^t - \prod_{j=1}^3 \underline{\Theta}^t \times_j \Psi_j - \underline{\nu}_2^t\|_F^2. \quad (12)$$

To solve the optimization problems in Eqs. (10)-(12), we use the Sherman-Morrison-Woodbury matrix inversion algorithm [27], soft-thresholding, and a fast TV algorithm approach, respectively. The goal of the proposed algorithm, summarized in Algorithm 1, is to reconstruct the datacube $\underline{\mathbf{F}}$ from its compressive measurements $\underline{\mathbf{G}}$.

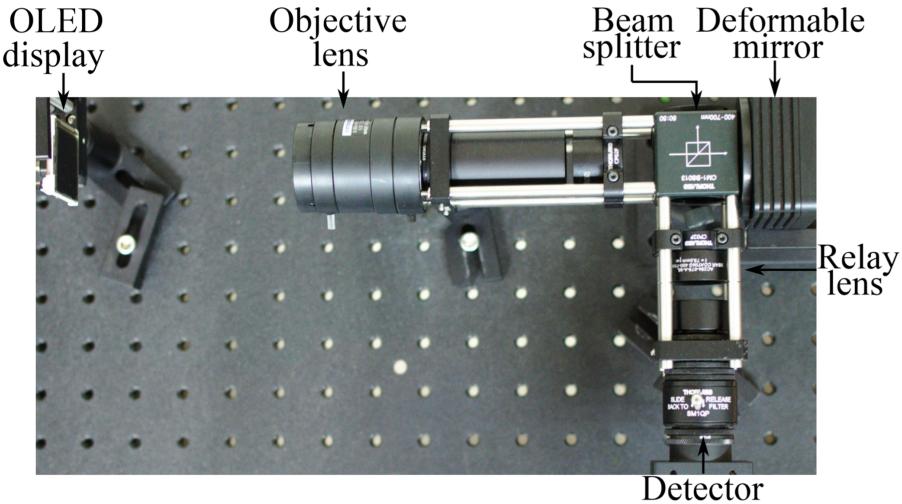


Fig. 2. Experimental setup implemented in the laboratory.

4. Simulations and experimental results

In this section, the performance of the proposed CSI architecture is evaluated using synthetic and experimental compressive measurements. First, the results obtained with the compressive sensing (CS) reconstruction algorithm are presented using the simulated acquisition process of the CSI architecture. Then, a testbed implementation of the proposed system is used for capturing experimental measurements, as can be seen in Fig. 2.

4.1. Quality metrics

To compare the quality of the rendered spectral images obtained with the proposed architecture those obtained by against the SD-CASSI, DD-CASSI, and 3D-CASSI architectures, two metrics were selected, the peak signal-to-noise-ratio (PSNR), defined as,

$$\text{PSNR} = \frac{1}{L} \sum_{i_3=1}^L \left[10 \log_{10} \left(\frac{\max(\tilde{\mathbf{F}}_{:,i_3})^2}{\text{MSE}(\mathbf{F}, \tilde{\mathbf{F}})} \right) \right], \quad (13)$$

and the spectral angular mapper (SAM), defined as,

$$\text{SAM} = \frac{1}{MN} \sum_{i_1=1}^M \sum_{i_2=1}^N \cos^{-1} \left(\frac{\sum_{i_3=1}^L \mathbf{F}_{i_1,i_2,i_3} \tilde{\mathbf{F}}_{i_1,i_2,i_3}}{\|\mathbf{F}_{:,i_3}\|_2 \cdot \|\tilde{\mathbf{F}}_{:,i_3}\|_2} \right), \quad (14)$$

where $\max(\cdot) : \mathbb{R}^{M \times N} \rightarrow \mathbb{R}$ is a function that returns the maximum possible value of $\tilde{\mathbf{F}}_{:,i_3}$, MSE is the mean squared error between $\mathbf{F}_{:,i_3}$ and its approximation $\tilde{\mathbf{F}}_{:,i_3}$, $\|\tilde{\mathbf{F}}_{:,i_3}\|_2 = \|\text{vec}(\tilde{\mathbf{F}}_{:,i_3})\|_2$, and $\sum_{i_3=1}^L \tilde{\mathbf{F}}_{i_1,i_2,i_3} \mathbf{F}_{i_1,i_2,i_3} = \langle \text{vec}(\tilde{\mathbf{F}}_{:,i_3}), \text{vec}(\tilde{\mathbf{F}}_{:,i_3}) \rangle$ with $\langle \cdot, \cdot \rangle$ representing the dot product between two vectors. Thus, the PSNR metric is defined as the ratio between the maximum possible power of a signal and the power of the corrupting noise related to the fidelity of the reconstruction, while the SAM metric provides the spectral angle between the spectral signature of the reference image and its approximation.

4.2. Simulated acquisition process

In order to verify the proposed CSI architecture, the compressive measurements are simulated using the forward model in Eq. (5), adding Gaussian noise with zero mean and variance given

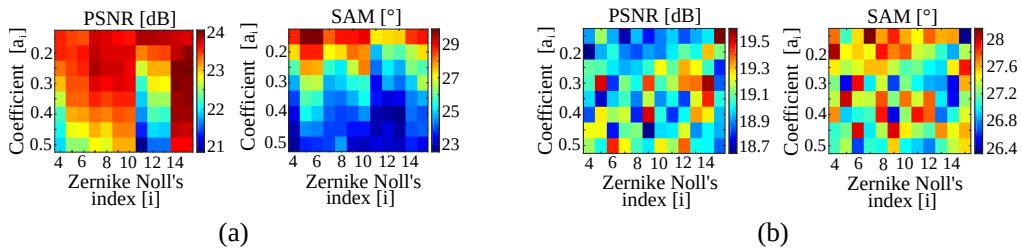


Fig. 3. Spectral image reconstruction with $K = 2$ and $SNR = 50[dB]$. Here, for the first snapshot, in the two cases (a) and (b), the measurement is generated by using $W(x, y) = a_i Z_i$, where $i \in \{4, \dots, 15\}$ and $a_i \in \{0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5\}$. For the second snapshot, the measurement is generated by using (a) $W(x, y) = 0$ and (b) $W(x, y) \sum_{i=4}^{15} a_i Z_i$.

by a specific signal-to-noise-ratio (SNR). The spectral images used to test the performance of the proposed architecture are Beads [28], Feathers [28], Ribeira [29], and Flowers [29]. The spectral images have a spatial resolution of $M \times N = 512 \times 512$ and $L = 26$ spectral bands. The transmission function of the fixed color-coded aperture is generated using a Bernoulli distribution with $p = 0.5$. For each data set, we test three noise scenarios with signal-to-noise ratios (SNR) of $\{20, 30, 50\}$ [dB], and a compression level of $c = 100 \times (1 - \frac{K}{L}) \sim \{93\%, 90\%, 85\%\}$, obtained using $K = \{2, 3, 4\}$ snapshots, respectively. The latter means that the spectral estimation is attained from only $\sim \{7\%, 10\%, 15\%$ of the original data. To select the set of Zernike polynomials that exhibit better performance in the proposed architecture, a cross-validation analysis by varying the used of two Zernike polynomials for the reconstruction of the spectral image was developed, as can be seen in Fig. 3. Specifically, this analysis was developed in two aspects. First, one of the measurements is simulated by deforming the surface of the DM from using a pure Zernike polynomial $W(x, y) = a_i Z_i$, and the second measurement is simulated by setting the DM without aberration in its surface $W(x, y) = 0$, as can be seen in Fig. 3(a). Second, one of the measurements is simulated by deforming the surface of the DM from using random Zernike coefficients $W(x, y) = \sum_{i=4}^{15} a_i Z_i$, and the second measurement is simulated by

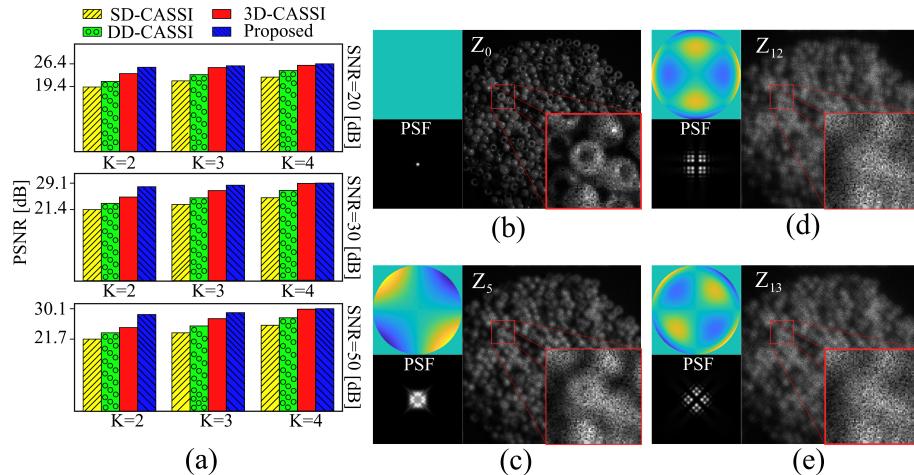


Fig. 4. Reconstruction quality analysis for $K = \{2, 3, 4\}$ and $SNR = \{20, 30, 50\}$. Furthermore, a measurement example generated from the Zernike polynomials (b) $W = 0$, (c) $W = Z_5$, (d) $W = Z_{12}$ and (e) $W = Z_{13}$ are illustrated.

deforming the surface of the DM from using a pure Zernike polynomial $W(x, y) = a_i Z_i$, as can be seen in Fig. 3(b). For the second case, one hundred runs were performed and the mean values were calculated. Based on the simulation results, illustrated in Fig. 3, the proposed architecture exhibit better performance when the deformation in the first snapshot is $a_i = 0$ and the following snapshots are acquired with $i = \{12, 13, 5, 14, 15\}$ and $a_i = 0.4$, respectively. This sequence of Zernike polynomials is select taking into account the trade-off between PSNR and SAM performance.

To evaluate the impact of the total of acquisitions and noise level on the quality of the reconstruction, we analyze the performance of the SD-CASSI, DD-CASSI, 3D-CASSI and the proposed architecture in terms of the PSNR metric. In Fig. 4, it can be seen that the performance of the proposed novel approach is the best in all cases, and the improvement in performance of the proposed architecture in contrast with the traditional architectures diminishes as the amount of snapshots grow. Specifically, the Zernike polynomials used in Fig. 4 were $\{Z_0, Z_{12}\}$, $\{Z_0, Z_{12}, Z_{13}\}$ and $\{Z_0, Z_{12}, Z_{13}, Z_5\}$, for $K = \{2, 3, 4\}$, respectively. Moreover, the compressive measurements generated with the aforementioned Zernike polynomials are illustrated in Z_0 (Fig. 4(b)), Z_5 (Fig. 4(c)), Z_{12} (Fig. 4(d)) and Z_{13} (Fig. 4(e)). Here, it is important to note that even if the proposed architecture improves its performance as more snapshots are taken, an optimum trade-off between the compression ratio and reconstruction quality has been found when $K = 3$. As such, the following experiments are developed for $K = 3$, where the compressive measurements are generated from the Zernike polynomials Z_0 , Z_{12} and Z_{13} with $a_i = 0.4$.

Table 1 illustrates the averaged spatial and reconstruction quality measured in terms of both metrics. For clarity, the best results are bold-faced and the second best are underlined. There, it can be seen that the proposed architecture overcomes the results obtained with the 3D-CASSI, DD-CASSI and SD-CASSI architecture overall in all the noise levels. Specifically, the proposed architecture gains up to 0.91, 1.03 and 5.4 dB in averaged PSNR, against the 3D-CASSI, DD-CASSI and SD-CASSI architecture, and attains 1.6°, 2.7° and 2.8° less spectral degrees in averaged SAM against the 3D-CASSI, DD-CASSI and SD-CASSI architecture, respectively. Figure 5(a) illustrates an RGB composite of the attained reconstructions for the SD-CASSI, DD-CASSI, 3D-CASSI and proposed approaches with a compression ratio of $c \sim 90\%$ ($K = 3$) and a noise level of $SNR = 20$ [dB]. Further results with all the datasets can be appreciated online in the supplemental [Visualization 1](#). To verify the accuracy of the experimental results, Fig. 5(b) shows the cumulative absolute error between the ground-truth Beads image and the reconstruction. It can be observed that the error obtained by the proposed approach is lower than that of 3D-CASSI, DD-CASSI, and SD-CASSI, approaches, respectively. Finally, Fig. 5(c) shows a visual comparison between 8 out of the 26 spectral bands. To evaluate the spectral accuracy of the reconstructions, three different points of each scene are illustrated in Fig. (6). This comparison shows that despite the good approximation of the spectra obtained with the SD-CASSI, DD-CASSI, and 3D-CASSI approaches, the proposed architecture provides better approximations of the spectral information. Overall, in Fig. 5 and Fig. 6 can be observed that the proposed architecture overcomes the reconstructions obtained from the rest of the evaluated architectures.

4.3. Experimental results

To experimentally evaluate the effectiveness of the proposed architecture, we built a proof-of-concept CSI system using a monochrome detector as seen in Fig. 2. The laboratory prototype is composed of a main objective lens coupled with a 4f system that has a phase modulating element at 2f. The detector array (Point Grey Grasshopper GS3-U3-23S6M-C) is placed at the image plane. In detail, the intermediate image plane is formed by a 16mm objective lens (Computar M1614-MP2) which is relayed by a pair of 75mm Fourier transforming lenses (Thorlabs AC254-075-A-ML). Using a beamsplitter (BS, Thorlabs CCM1-BS013), we placed a deformable

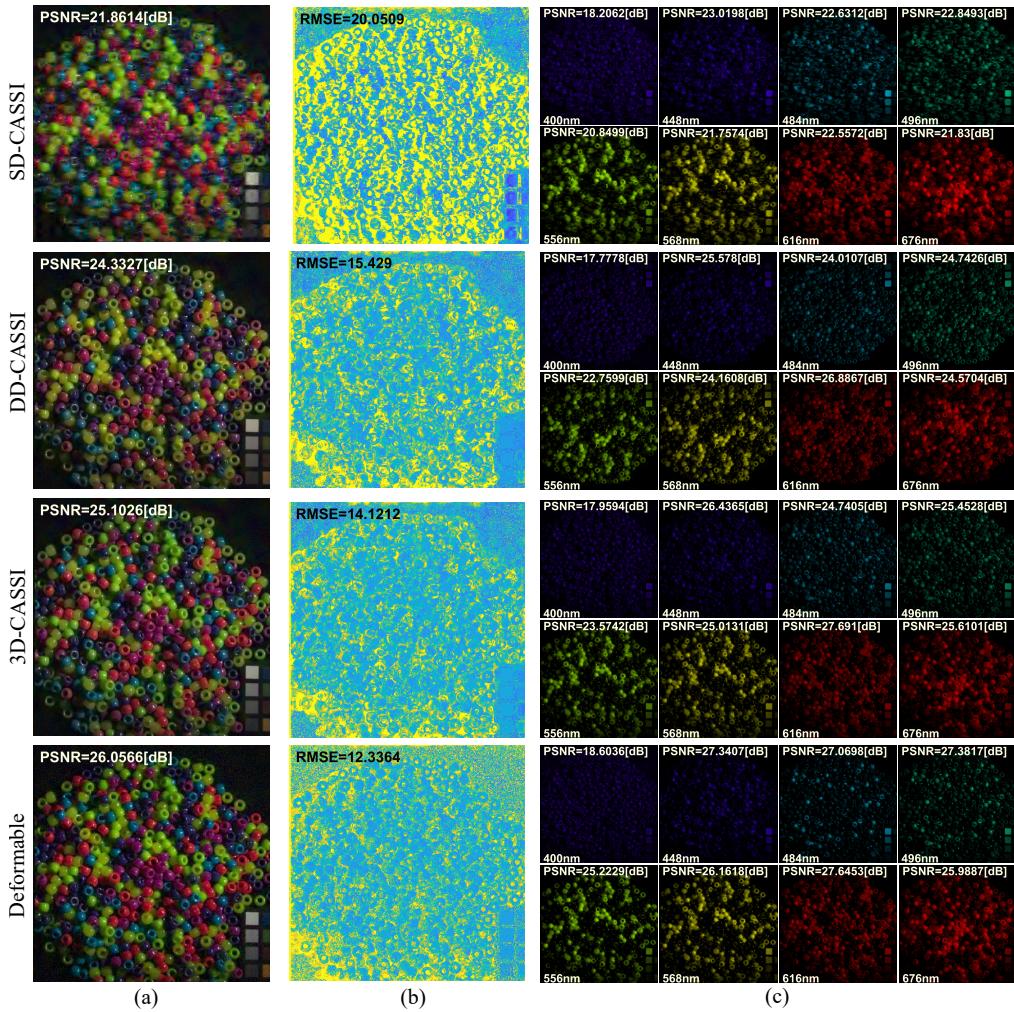


Fig. 5. (Visualization 1) (a) Visual comparison of the spectral reconstructions, mapped to an RGB profile, (b) absolute error of the reconstructions, and (c) comparison of 8 out of the 26 spectral bands. These simulations were developed with the parameters $SNR = 20$ [dB] and $K = 3$.

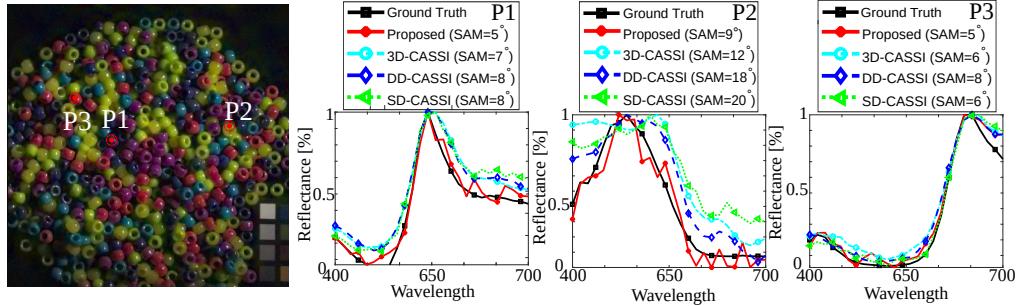


Fig. 6. Visual comparison of spectral signatures for reconstruction from SD-CASSI, DD-CASSI, 3D-CASSI and proposed architecture

Table 1. Reconstruction Results Comparison for Simulations using $K = 3$.

Images	Metrics	Architecture											
		SD-CASSI			DD-CASSI			3D-CASSI			Proposed		
		20	30	50	20	30	50	20	30	50	20	30	50
Beads	PSNR	21.8	22.7	22.8	24.3	27.3	27.6	25.1	27.3	27.9	28.0	28.8	29.8
	SAM	21.3	17.5	17.1	21.4	16.5	16.2	20.7	15.6	12.0	19.9	14.2	9.5
Ribeira	PSNR	26.5	29.2	31.1	29.1	32.6	37.1	29.3	32.5	36.0	29.4	33.1	37.2
	SAM	8.1	5.5	3.6	7.2	5.2	3.1	7.0	5.3	3.1	6.2	4.4	2.8
Flowers	PSNR	27.7	29.4	29.7	30.4	33.81	37.2	30.7	33.8	36.1	31.4	34.7	37.3
	SAM	23.5	15.9	14.3	23.1	16.3	10.6	22.8	16.1	10.5	20.1	13.9	6.9
Feathers	PSNR	28.0	28.2	28.8	30.1	33.6	35.3	31.9	33.6	35.3	30.7	33.8	36.6
	SAM	15.9	11.8	9.0	17.5	11.4	8.2	15.7	11.2	8.0	14.8	11.1	5.5

mirror (DM, Thorlabs DMP40-P01) at the pupil plane at a distance of $2f = 150\text{mm}$ from the intermediate image plane. Finally, the detector array is placed at a distance of $2f = 150\text{mm}$ from the deformable mirror. Also, spectral filters mounted on a filter wheel are used before the camera. Specifically, a set of twelve color filters of $10 \Delta\lambda \text{ nm}$ with steps of $\sim 20 \text{ nm}$ between the 460-680 nm spectral range were used. In this way, we can emulate any potential colored-filter array detector. Similarly to the simulations, the transmission function of the emulated color-coded aperture is generated using a Bernoulli distribution with $p = 0.5$. For the experiments we used two kind of targets to acquire the experimental spectral datasets with the implemented testbed: an OLED screen with a spatial resolution of 128×128 (OLED-128-G2, 4D Systems)

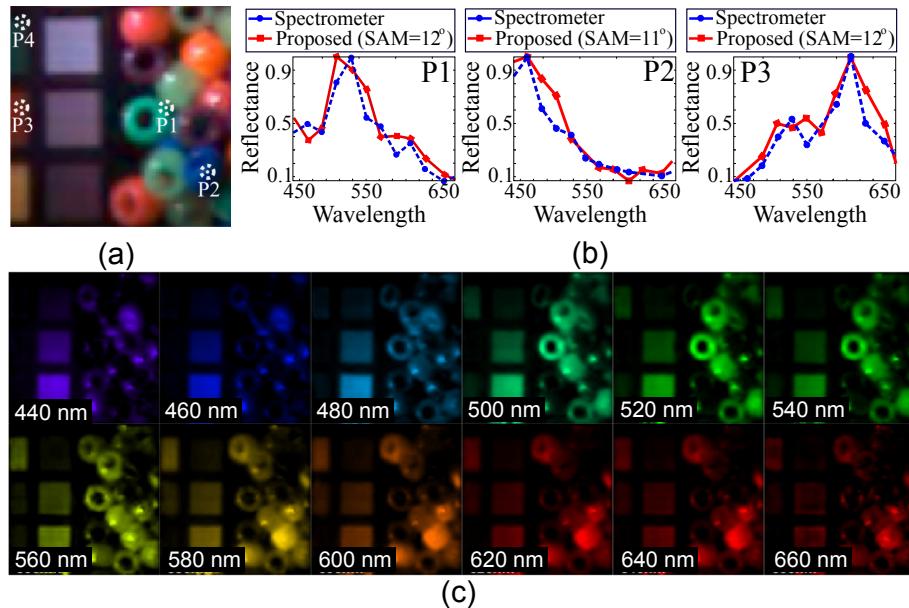


Fig. 7. Experimental results using the “Digital beads” scene, where (a) is an RGB version of the reconstruction and (b) is a visualization of 12 reconstructed spectral bands. Spectral signature reconstructions of four points of the scene, using $K = 3$ snapshot measurements.

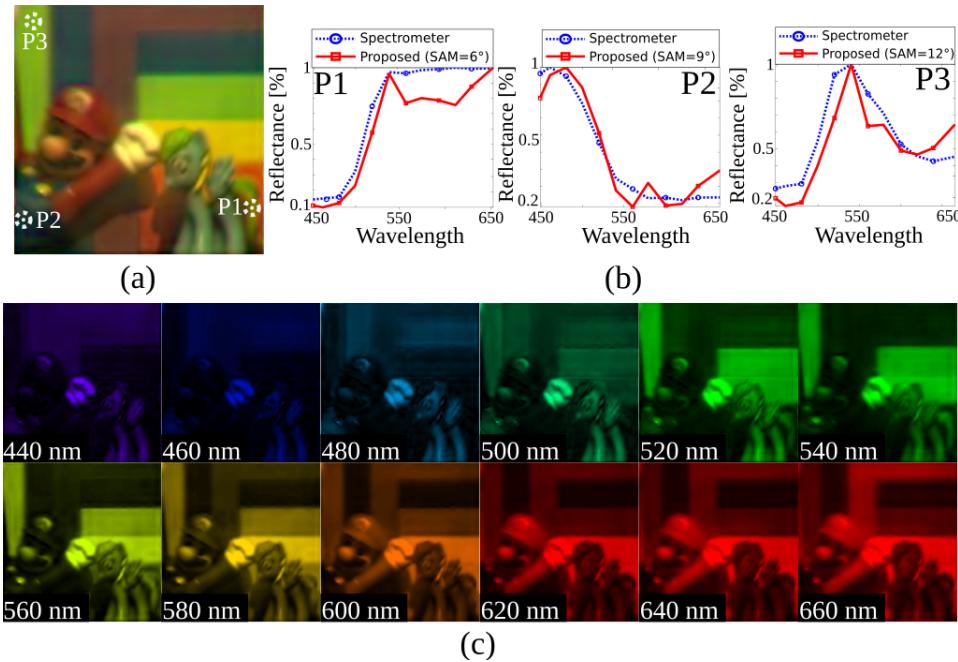


Fig. 8. Experimental result using the “Toys” scene, where (a) is an RGB version of the reconstruction and (b) is a visualization of 12 reconstructed spectral bands. (c) Spectral signature reconstructions of four points of the scene, using $K = 3$ snapshot measurements.

with a displayed RGB image and a constructed scene using a collection of different colored toys. Using the OLED screen, an image composed of several color beads is loaded into the screen and the resulting “Digital beads” dataset exhibit a spatial resolution of $M \times N = 128 \times 128$ pixels, and $L = 12$ spectral bands. On the other hand, the constructed scene was illuminated by a 125 [W] mercury lamp and a calibrated quartz tungsten-halogen (QTH) lamp (Newport QTH 63355), recording the “Toys” dataset with a spatial resolution of $M \times N = 512 \times 512$ pixels, and $L = 12$ spectral bands. Moreover, to validate the spectral reconstruction fidelity, several spatial points of the real scenes were also measured with a spectrometer in the same range of the color filters (SR-2500 Spectral Evolution). In both experiments, we considered a compression ratio of $c \sim 25\%$, obtained when only $K = 3$ snapshots are captured. Figure 7(a) and 8(a) show an RGB spectral-mapped version of the two reconstructed scenes from compressive measurements for the “Digital beads” and “Toys” scenes acquired with the proposed architecture. Although spectral reconstruction can be seen from Fig. 7(c) and 8(c) for several bands, only the specific spectra shown in Fig. 7(b) and 8(b) corroborate the accuracy and spectrum resolution achieved by the experimental reconstructed results when using the proposed architecture with our novel reconstruction algorithm.

5. Conclusion

In this article, we presented a novel and compact multishot CSI architecture based on a deformable mirror and a colored-filter detector to produce arbitrary spatio-spectral projections. The image reconstruction problem was formulated as an inverse problem with two regularization functions, ensuring a smooth reconstructed image and a sparse decomposition of the image in an appropriate representation basis, respectively. We developed an ADMM algorithm in tensor representation to efficiently solve the high-dimensional reconstruction problem. The proposed architecture is

compared with three state-of-the art CSI architectures on four spectral images, where simulations results showed an improvement of up to 5.4 dB in PSNR and attained 2.8 less spectral degrees in averaged SAM when the proposed CSI scheme was using as few as 10% of the data. In contrast with traditional CSI architectures, the proposed spectral imaging system established a better trade-off between reconstruction quality and light throughput. In particular, the obtained simulated and experimental reconstruction results validate the idea that moderate phase coding coupled with a colored filter detector is a useful and practical tool to dynamically modulate information in the spatial and spectral domain, allowing a more compact CSI system.

Appendix

Preliminaries on tensors

In this section, we first summarize some main notations and then introduce some useful product, decomposition, and tensorial representation properties [30, 31]. Here, scalars are denoted as lowercase letters (a, b, \dots), vectors as bold-faced lowercase letters ($\mathbf{a}, \mathbf{b}, \dots$), matrices as bold-faced capital letters ($\mathbf{A}, \mathbf{B}, \dots$), and tensors as underlined bold-faced capital letters ($\underline{\mathbf{A}}, \underline{\mathbf{B}}, \dots$).

- *Tensor notation:* Let $\underline{\mathbf{X}} \in \mathbb{C}^{N_1 \times N_2 \times \dots \times N_T}$ be a T -dimensional tensor with a vector representation $\mathbf{x} = \text{vec}(\underline{\mathbf{X}})$, where $\mathbf{x} \in \mathbb{C}^{\prod_{n=1}^T N_n}$ and $\text{vec}(\cdot) : \{\mathbb{C}^{N_1 \times N_2 \times \dots \times N_T} \rightarrow \mathbb{C}^{\prod_{n=1}^T N_n}\}$.
- *n-Mode product:* Let $\underline{\mathbf{X}} \in \mathbb{C}^{N_1 \times \dots \times N_T}$ and $\mathbf{A} \in \mathbb{C}^{M \times N_n}$ be an T -dimensional tensor and a matrix, respectively. Then, n-mode tensor by the matrix product $(\underline{\mathbf{Y}} = \underline{\mathbf{X}} \times_n \mathbf{A}) \in \mathbb{C}^{N_1 \times N_{n-1} \times M \times N_{n+1} \times \dots \times N_T}$ is defined by

$$y_{i_1, \dots, i_{n-1}, j, i_{n+1}, \dots, i_T} = \sum_{i_n}^{N_n} x_{i_1, \dots, i_n, \dots, i_T} a_{j, i_n}, \quad (15)$$

where $i_k \in \{1, 2, \dots, N_k\}$ with $k \neq n$ and $j \in \{1, 2, \dots, M\}$ with $\{N_k, M\} \in \mathbb{N}$.

- *Tensor decomposition:* Let $\underline{\mathbf{X}} \in \mathbb{C}^{N_1 \times \dots \times N_T}$ and $\mathbf{A}_n \in \mathbb{C}^{N_n \times N_n}$ be a T -dimensional tensor and orthogonal matrices, respectively. Then, a tensor representation of $\underline{\mathbf{X}}$ in the orthogonal basis \mathbf{A}_n can be expressed as

$$\underline{\Theta} = \prod_{n=1}^T \underline{\mathbf{X}} \times_n \mathbf{A}_n, \quad (16)$$

where $\underline{\Theta} \in \mathbb{C}^{N_1 \times \dots \times N_T}$ represents the core tensor which reveals the underlying interactions between the orthogonal matrices \mathbf{A}_i with $\underline{\mathbf{X}} = \prod_{n=1}^T \underline{\Theta} \times_n \mathbf{A}_n^T$, and $\underline{\mathbf{X}} = \prod_{i=1}^T \left(\prod_{j=1}^T \underline{\mathbf{X}} \times_j \mathbf{A}_j \right) \times_i \mathbf{A}_i^T = \prod_{i=1}^T \underline{\mathbf{X}} \times_i \mathbf{A}_i^T \mathbf{A}_i$ with $\mathbf{A}_i^T \mathbf{A}_i = \mathbf{I}_i$.

- *Tensor notation to linear-matrix model:* Let $\mathbf{A}_i \in \mathbb{C}^{M_i \times N_i}$ and $\underline{\mathbf{X}} \in \mathbb{C}^{N_1 \times \dots \times N_T}$ a collection of matrices and T -dimensional tensor, respectively. Then, the N-mode product between $\underline{\mathbf{X}}$ and the set of matrices \mathbf{A}_i can be expressed in matrix notation as

$$\mathbf{y} = \text{vec}(\underline{\mathbf{Y}}) = \text{vec} \left[\prod_{n=1}^T \underline{\mathbf{X}} \times_n \mathbf{A}_n \right] = \mathbf{A} \text{vec}(\underline{\mathbf{X}}), \quad (17)$$

where $\mathbf{A} \in \mathbb{C}^{(\prod_i^T M_i) \times (\prod_i^T N_i)}$ with $\mathbf{A} = (\mathbf{A}_T \otimes \mathbf{A}_{T-1} \otimes \dots \otimes \mathbf{A}_1)$, $\mathbf{y} \in \mathbb{C}^{\prod_{i=1}^T M_i}$ and $\mathbf{y} \in \mathbb{C}^{\prod_{i=1}^T M_i}$ are the vector versions of $\underline{\mathbf{Y}}$ and $\underline{\mathbf{X}}$, respectively.

Funding

Fondo Nacional de Ciencia y Tecnología (FONDECYT) (1181943); Redes Etapa Inicial (REDI) (REDI170539); Departamento Administrativo de Ciencia, Tecnología e Innovación (COLCIENCIAS) (FP44842-309-2018); Programa regional MATH-AmSud (MATHAMSUD)(10-2017)

References

1. M. Kim, Y. Chen, and P. Mehl, "Hyperspectral reflectance and fluorescence imaging system for food quality and safety," *Transactions of the ASAE* **44**(3), 721–729 (2001).
2. C. Nansen, K. Singh, A. Mian, B. J. Allison, and C. Simmons, "Using hyperspectral imaging to characterize consistency of coffee brands and their respective roasting classes," *Journal of Food Engineering* **190**, 34–39 (2016).
3. G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *Journal of Biomedical Optics* **19**(1), 010901 (2014).
4. Y. Schechner and S. Nayar, "Generalized mosaicing," in *IEEE International Conference on Computer Vision*, (IEEE, 2001), pp. 17–24.
5. N. Gat, "Imaging spectroscopy using tunable filters: a review," *Proc. SPIE* **4056**, 50–64 (2000).
6. R. Green, M. Eastwood, C. Sarture, T. Chrien, M. Aronsson, B. Chippendale, J. Faust, B. Pavri, C. Chovit, M. Solis, and W. Orleska, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris)," *Remote Sensing of Environment* **65**(3), 227–248 (1998).
7. X. Cao, T. Yue, X. Lin, S. Lin, X. Yuan, Q. Dai, L. Carin, and D. J. Brady, "Computational snapshot multispectral cameras: Toward dynamic capture of the spectral world," *IEEE Signal Processing Magazine* **33**(5), 95–108 (2016).
8. S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Transactions on Signal Processing* **56**(6), 2346 (2008).
9. G. Arce, D. Brady, L. Carin, H. Arguello, and D. Kittle, "Compressive coded aperture spectral imaging: An introduction," *IEEE Signal Processing Magazine* **31**(1), 105–115 (2014).
10. M. Gehm, R. John, D. Brady, R. Willett, and T. Schulz, "Single-shot compressive spectral imaging with a dual-disperser architecture," *Opt. Express* **15**(21), 14013–14027 (2007).
11. A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Appl. Opt.* **47**(10), B44–B51 (2008).
12. X. Lin, Y. Liu, J. Wu, and Q. Dai, "Spatial-spectral encoded compressive hyperspectral imaging," *ACM Transactions on Graphics* **33**(6), 233 (2014).
13. X. Lin, G. Wetzstein, Y. Liu, and Q. Dai, "Dual-coded compressive hyperspectral imaging," *Opt. Lett.* **39**(7), 2044–2047 (2014).
14. X. Yuan, T. H. Tsai, R. Zhu, P. Llull, D. Brady, and L. Carin, "Compressive hyperspectral imaging with side information," *IEEE Journal of Selected Topics in Signal Processing* **9**(6), 964–976 (2015).
15. Y. Murakami, M. Yamaguchi, and N. Ohyama, "Hybrid-resolution multispectral imaging using color filter array," *Opt. Express* **20**(7), 7173–7183 (2012).
16. H. Arguello and G. Arce, "Colored coded aperture design by concentration of measure in compressive spectral imaging," *IEEE Transactions on Image Processing* **23**(4), 1896–1908 (2014).
17. C. V. Correa, H. Arguello, and G. R. Arce, "Snapshot colored compressive spectral imager," *J. Opt. Soc. Am. A* **32**(10), 1754–1763 (2015).
18. S. Baek, I. Kim, D. Gutierrez, and M. Kim, "Compact single-shot hyperspectral imaging using a prism," *ACM Transactions on Graphics* **36**(6), 217 (2017).
19. C. Correa, C. Hinojosa, G. Arce, and H. Arguello, "Multiple snapshot colored compressive spectral imager," *Optical Engineering* **56**(4), 041309 (2016).
20. K. Degraux, V. Cambareri, B. Geelen, L. Jacques, and G. Lafruit, "Multispectral compressive imaging strategies using fabry–pérot filtered sensors," *IEEE Transactions on Computational Imaging* **4**(4), 661–673 (2018).
21. P. Madec, "Overview of deformable mirror technologies for adaptive optics and astronomy," *Proc. SPIE* **8447**, 844705 (2012).
22. E. Fernández and P. Artal, "Membrane deformable mirror for adaptive optics: performance limits in visual optics," *Opt. Express* **11**(9), 1056–1069 (2003).
23. M. D. Robinson, G. Feng, and D. G. Stork, "Spherical coded imagers: improving lens speed, depth-of-field, and manufacturing yield through enhanced spherical aberration and compensating image processing," *Proc. SPIE* **7429**, 74290M (2009).
24. E. Vera and P. Meza, "Snapshot compressive imaging using aberrations," *Opt. Express* **26**(2), 1206–1218 (2018).
25. R. J. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Am. A* **66**(3), 207–211 (1976).
26. M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Transactions on Image Processing* **20**(3), 681–695 (2011).
27. M. Akgün, J. Garcelon, and R. Haftka, "Fast exact linear and non-linear structural reanalysis and the sherman–morrison–woodbury formulas," *International Journal for Numerical Methods in Engineering* **50**(7), 1587–1606 (2001).
28. F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," *IEEE Transactions on Image Processing* **19**(9), 2241–2253 (2010).

29. D. H. Foster, K. Amano, S. M. Nascimento, and M. J. Foster, "Frequency of metamerism in natural scenes," *J. Opt. Soc. Am. A* **23**(10), 2359–2372 (2006).
30. S. Rabanser, O. Shchur, and S. Günnemann, "Introduction to tensor decompositions and their applications in machine learning," <https://arxiv.org/abs/1711.10781>.
31. N. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Transactions on Signal Processing* **65**(13), 3551–3582 (2017).