OPTICAL CODING IN COMPRESSIVE SPECTRAL IMAGING SYSTEMS AND THE MULTI-RESOLUTION RECONSTRUCTION PROBLEM

by

Claudia V. Correa-Pugliese

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering

Fall 2017

© 2017 Claudia V. Correa-Pugliese
All Rights Reserved
OPTICAL CODING IN COMPRESSIVE SPECTRAL IMAGING SYSTEMS AND THE MULTI-RESOLUTION RECONSTRUCTION PROBLEM

by

Claudia V. Correa-Pugliese

Approved: ____________________________________________________

Kenneth E. Barner, Ph.D.
Chair of the Department of Electrical and Computer Engineering

Approved: ____________________________________________________

Babatunde A. Ogunnaike, Ph.D.
Dean of the College of Engineering

Approved: ____________________________________________________

Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education
I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:  
Gonzalo R. Arce, Ph.D.  
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:  
Kenneth E. Barner, Ph.D.  
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:  
Javier Garcia-Frias, Ph.D.  
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:  
Henry Arguello, Ph.D.  
Member of dissertation committee
ACKNOWLEDGEMENTS

I want to express my sincere gratitude to my advisor, Dr. Gonzalo R. Arce for his continuous guidance and academic support during these years. My deep gratitude to my co-advisor, Dr. Henry Arguello for the encouragement and advice he has provided throughout my MS and PhD studies. I also want to thank the members of the dissertation committee: Dr. Kenneth Barner and Dr. Javier Garcia-Frias, for accepting the invitation to be part of the committee. I would like to acknowledge Colciencias (the Colombian national science agency) for providing the funds for my doctoral studies. My appreciation extends to my lab-mates at the Computational Imaging Group for the weekly seminars and informal discussions that contributed to my research, and for many hours spent together in the lab. Also, to the friends I’ve made in Newark, the ones who are staying, and those who already left. Last but not least, I want to specially thank my family for their love and invaluable support throughout this journey. My parents for showing me the importance of hard work and commitment, and for being my first teachers. My husband for sharing this journey with me, and being my unconditional partner in every sense. And my brother, for being always there in spite of the distance.
TABLE OF CONTENTS

LIST OF TABLES ........................................... vii
LIST OF FIGURES .......................................... viii
ABSTRACT ...................................................... xv

Chapter

1 INTRODUCTION .................................................. 1

1.1 Spectral Imaging ............................................. 1

1.1.1 Traditional Spectral Imaging Methods .......... 1
1.1.2 Compressive Spectral Imaging Systems ........ 3
1.1.3 Multi-Resolution Compressed Sensing Reconstructions . 5

1.2 Motivation ................................................... 6
1.3 Dissertation Organization ................................. 7
1.4 Contributions ............................................... 8

2 SPATIO-TEMPORAL BLUE NOISE CODED APERTURES FOR
COMPRESSION SPECTRAL IMAGING .......................... 10

2.1 Multi-Shot CASSI Matrix Representation ............ 14
2.2 Restricted Isometry Property in Multi-Shot CASSI .... 17
2.3 Coded Aperture Design Based on RIP .................. 20
2.4 Simulation and Experimental Results .................. 26

2.4.1 Simulations ............................................. 27
2.4.2 Testbed Implementation ............................... 38

2.5 Conclusions ................................................. 40
# LIST OF TABLES

4.1 Average reconstruction quality results for the Toy data cube in terms of PSNR and SSIM, for different number of shots. 81

4.2 Average reconstruction quality results for the Feathers data cube in terms of PSNR and SSIM, for different number of shots. 81

4.3 Quality of interpolated LR reconstructions using a bicubic interpolator for the Toy data cube. 83

4.4 Quality of interpolated LR reconstructions using a bicubic interpolator for the Feathers data cube. 84
LIST OF FIGURES

2.1 Optical elements present in CASSI. The input scene $f_0 (x, y, \lambda)$ is first spatially encoded by the coded aperture $T (x, y)$ and then, the encoded source is dispersed by a prism. Finally, the coded and dispersed scene is captured by the focal plane array. .......................... 11

2.2 The structure of the matrix $H$ is illustrated for $K = 2, N = 4$, and $L = 4$. For illustration, the entries of the coded apertures are $t_{ij} \sim \text{Bernoulli}(0.5)$. Dark elements represent zero values and white elements correspond to the one-valued entries of the coded aperture. 16

2.3 Comparison for different coded aperture ensembles with $K = 4$ in terms of the parameter $\alpha$. (Left) Random coded aperture entries are realizations of a Bernoulli random variable with $p = 50\%$. (Right) Coded aperture intentionally designed to exhibit horizontal separation of one-valued features. ................................. 21

2.4 Incoherence comparison for different coded aperture ensembles with $K = 4$. (Left) Random coded aperture entries are realizations of a Bernoulli random variable with $p = 50\%$. (Right) Coded aperture intentionally designed to exhibit vertical separation of one-valued features. ................................. 21

2.5 Graphic representation of the metric used to determine the concentration of one-valued entries in a window $U_i$ of the code. This metric helps on deciding the best $i$ for setting the one-valued element. 23

2.6 Block diagram of the algorithm that generates the designed coded apertures. This procedure iterates over all $(\Omega_r, \Omega_c)$ points and determines the best coded aperture of the ensemble in which the $(\Omega_r, \Omega_c)$-th pixel is set to one. ................................. 24

2.7 Example of the operations performed by the proposed algorithm. (a) Initial guess of a 8 × 8 coded aperture ensemble for $K = 4$ snapshots. (b) Resulting ensemble after applying the proposed algorithm. . . 25
2.8 (Top) Comparison of one realization of the spatio temporal blue noise (BN) coded apertures with the traditional random, boolean and DBS coded apertures. (Bottom) Zoomed portion of an BN coded aperture ensemble for $K = 4$ measurement snapshots to illustrate low correlation across shots. ................................................................. 26

2.9 Average reconstruction quality measured by the PSNR and SSIM, as a function of the number of captured snapshots. Results for three different data bases are illustrated, using random, boolean, DBS and BN coded apertures. ................................................................. 28

2.10 RGB profiles of the reconstructed images obtained using random, boolean, DBS and BN coded aperture ensembles, and $K = 4$ measurement snapshots, compared to the original data cubes. .... 30

2.11 Comparison of a zoomed portion of the 8 spectral bands from reconstructions in Fig. 2.10 for Database 1. ......................... 31

2.12 Comparison of a zoomed portion of the 8 spectral bands from reconstructions in Fig. 2.10 for Database 2. ............................... 32

2.13 Comparison of a zoomed portion of the 8 spectral bands from reconstructions in Fig. 2.10 for Database 3. ......................... 33

2.14 Spectral signature comparison for two spatial points on each data base. Reconstructions obtained from $K = 4$ snapshots. .......... 34

2.15 Average reconstruction PSNR and SSIM for the Moffett Field data set, as a function of the number of snapshots. .................. 35

2.16 Comparison of a zoomed portion of 6 out of the 32 spectral bands from the reconstructions of the Moffet Field database for $K = 12$. Average PSNR: Random 30.88 dB; Boolean 31.63 dB; DBS 32.33; BN 32.66 dB. ................................................................. 36

2.17 Comparison of the radially averaged power spectral density (RAPSD) for random, boolean, DBS and BN coded apertures. ........ 37

2.18 Testbed implementation of the CASSI system. ..................... 38

2.19 RGB representations of the experimental reconstructions from $K = 2, 4, 6$ for different types of coded apertures. .......... 39
2.20 Selected reconstructed spectral bands for random, boolean, DBS and BN coded apertures with \( K = 4 \).  

2.21 Spectral signature reconstructions of three points of the scene, using \( K = 4 \) snapshot measurements. (Top Right) Orange arm; (Bottom Left) Yellow face; (Bottom Right) Green chest.

3.1 Schematic of (a) the CASSI imager and (b) the snapshot colored compressive spectral imager (SCCSI).

3.2 Example of a discrete color filter array. Each spatial location \((m,n)\) contains the spectral response of the correspondent optical filter. Black features in \( C_{m,n,k} \) block the correspondent wavelengths and, white features correspond to the pass-bands of the filter.

3.3 A source voxel after being dispersed is illustrated. The energy of the voxel splits into three regions that impinge onto three consecutive FPA pixels. The proportion of energy from each voxel in the region \( u \) is represented by the weight \( w_{m,n,k,u} \).

3.4 Schematic representation of the discrete process behind SCCSI. Each spectral band of the input is dispersed according to the function \( S(\lambda) \). Then, the color filter array encodes the dispersed light before it is captured by an intensity detector. The set of filters in the array is also illustrated.

3.5 Dispersion matrix \( \mathbf{D} \), for \( N = 6 \) and \( L = 3 \). The three diagonals represent the regions in which each voxel splits due to the dispersive function \( S(\lambda) \). Different color intensities indicate the weights \( w_{m,n,k,u} \). The weights are arranged by spectral band and region as \( w_k^u \) for \( k = 0, \cdots, L - 1 \) and \( u = 0, 1, 2 \).

3.6 Sensing matrix \( \mathbf{H} \), for \( N = 6 \) and \( L = 3 \). The structure of the matrix accounts for the effects of the dispersion, coding and integration phenomenon occurred behind SCCSI. The three diagonals with different color intensities represent the regions in which each voxel splits due to the dispersive function \( S(\lambda) \) and different weight values.
Test data cubes with 8 spectral bands ranging from 450nm ($\lambda_1$) to 667nm ($\lambda_8$) used for simulations. Each spectral slice has a spatial resolution of 256 $\times$ 256 pixels. (a) Spectral bands with central wavelengths of the first data cube. (b) RGB representation of the Pavia University aerial view. (c) RGB representation of the color palette data cube.

Set of predefined band pass filters used for simulations. A 50% transmittance color filter array is guaranteed when these filters are uniformly distributed in the color filter array. $\{\lambda_1, \cdots, \lambda_8\}$ indicate the cut-off wavelengths which match the spectral bands of F.

Average reconstruction PSNR as a function of the number of colors in the detector (Filters in Fig. 3.8) of (a) data cube in Fig. 3.7 (a), (b) Pavia university in Fig. 3.7(b) and, (c) color palette in Fig. 3.7(c). The average PSNR obtained using a single shot with CASSI and SSCSI are also illustrated.

Zoomed reconstructions of a subset of spectral bands using (Second row) CASSI with average PSNR across bands 21.2 dB, (Third row) SSCSI with average PSNR 25.83 dB, (Fourth row) SCCSI with 4 colors with average PSNR 25.36 dB, and (Fifth row) SCCSI with 5 colors with average PSNR 26.45 dB. The PSNR for each spectral band is also indicated.

RGB reconstruction comparison using CASSI, SSCSI and SCCSI with 4 colors. (a) Reconstructions for the data cube in Fig. 3.7 (b) and, (b) Fig. 3.7 (c).

Reconstruction along the spectral axis for highlighted spatial locations using CASSI, SSCSI and SCCSI with 4 colors.

Eigenvalue distribution of the sensing matrix for the SCCSI with 4 colors, CASSI and SSCSI. The condition number $\kappa = \lambda_{\text{max}}/\lambda_{\text{min}}$ for all cases are also indicated.

(Left) Laboratory implementation of the SCCSI architecture. (Right) RGB representation of the target scene used for the experiments.

Characterization of the double amici prism in the SCCSI architecture. The effect of the dispersion in the spectral range \{448 – 663\} nm is illustrated. The energy of each voxel splits into two regions $R_0, R_1$ and, the corresponding weights are indicated.
3.16 Set of filters used for experiments. A 50% transmittance color filter array is obtained when a uniform distribution of the colors is used. \( \lambda_0 = 448 \text{ nm} \), \( \lambda_1 = 466 \text{ nm} \), \( \lambda_2 = 487 \text{ nm} \), \( \lambda_3 = 516 \text{ nm} \), \( \lambda_4 = 550 \text{ nm} \), \( \lambda_5 = 600 \text{ nm} \), and \( \lambda_6 = 663 \text{ nm} \).

3.17 Recovered spectral bands from SCCSI measurements using filters 1 to 4 in Fig. 3.16.

3.18 Comparison of different spectral bands using SCCSI with 2 and 4 colors in the detector. Zoomed regions of the reconstructions are illustrated.

3.19 Comparison of the spectral signature for the highlighted spatial points, measured with a commercial spectrometer and the SCCSI reconstructions using 2, 4 and 6 colors.

4.1 Schematic comparison of the traditional high-resolution (HR) CSI reconstruction with the proposed multi-resolution (MR) compressive spectral imaging reconstruction approach.

4.2 Diagram of the proposed MR reconstruction approach for compressive spectral imaging. First, a LR reconstruction is obtained using a decimation factor \( \Delta \). The upscaling process consists on finding the edges of the available LR reconstruction to determine a new MR decimation matrix to solve the equivalent inverse problem to recover a data cube with a decimation factor of \( \Delta - 1 \). The upscaling process is repeated until the reconstruction reaches the original resolution.

4.3 Structure of the decimation matrix \( D_\Delta \) for a factor \( \Delta = 1 \), and a data cube with \( N = 4 \) and \( L = 2 \). White elements represent a value of 1, and black entries are zero.

4.4 Low-resolution approximation errors for different decimation factors \( \Delta \). The errors are concentrated around the high frequency regions of the scene.

4.5 Example of the MR decimation matrix generation for \( \Delta - t = 1 \) (\( D_2 \)) from the decimation matrices \( \tilde{D}_1 \) and \( \tilde{D}_2 \). Illustrated matrices correspond to a data cube of spatial resolution \( N = 8 \), and \( L = 1 \) spectral band.
4.6 Example of the single step MR decimation matrix generation for $\Delta = 1$ from the decimation matrix $\hat{D}_1$ and identity matrix $I_N$.
Illustrated matrices correspond to a data cube of spatial resolution $N = 8$, and $L = 1$ spectral band.

4.7 RGB mapping of the test data cubes employed for simulations. (Left) $256 \times 256 \times 8$ toy spectral image. (Right) $512 \times 512 \times 16$ feathers spectral image.

4.8 Multi-resolution reconstructions of the Toy data cube for $K = 4$ snapshots and different values of $\Delta$, along with a zoomed portion of the reconstructions.

4.9 Multi-resolution reconstructions of the Feathers data cube for $K = 4$ snapshots and different values of $\Delta$, along with a zoomed portion of the reconstructions.

4.10 Reconstruction complexity as a function of the spatial dimensions of the data cube to recover $N, L = 5$ spectral bands, and $K = 1$ snapshot, for the full-resolution reconstruction, and MR with $\Delta = 1, 2$. Zoomed portions of the figure are also included.

4.11 Average reconstruction PSNR and SSIM comparison for $64 \times 64, 128 \times 128$ and $256 \times 256$ target resolutions of the Toy data cube using MR reconstruction from $\Delta = 3$ and bicubic interpolations.

4.12 Average reconstruction PSNR and SSIM comparison for $128 \times 128, 256 \times 256$ and $256 \times 256$ target resolutions of the Feathers data cube using MR reconstruction from $\Delta = 3$ and bicubic interpolations.

4.13 Data cube reconstruction comparison between MR and interpolated LR images from $\Delta = 3, 2, 1$ using $K = 4$ snapshots for (Top)$256 \times 256$ Toy database and (Bottom) $512 \times 512$ Feathers database.

4.14 Data cube reconstruction comparison between MR and interpolated LR images from $\Delta = 3, 2$ using $K = 4$ snapshots for (Top)$128 \times 128$ Toy database and (Bottom) $256 \times 256$ Feathers database.

4.15 Comparison of the average reconstruction PSNR and SSIM for the sequential MR reconstruction (MR) from $\Delta = 3$ and the single-step MR reconstruction (SMR) for $\Delta = 1, 2, 3$. 

4.16 Single step MR reconstructions of the high-resolution data cubes along with their zoomed portions, for (Left) $\Delta = 3$, (Center) $\Delta = 2$ and (Right) $\Delta = 1$ using $K = 6$ snapshots. Each image includes the correspondent attained SSIM and PSNR values.
ABSTRACT

Traditional spectral imaging approaches require scanning the scene to construct a 3-dimensional data cube. These methods experience low sensing speed and the resulting amount of data makes its management very challenging. In contrast, compressive spectral imaging (CSI) systems capture the spatial and spectral information of the scene at once, in a 2-dimensional set of coded projections. Compressed sensing (CS) reconstruction algorithms are then used to reconstruct the underlying spectral 3D data cube from an underdetermined system of linear equations. In general, CSI projections consist on integrating encoded and dispersed versions of the input source. Thus, different optical configurations yield different sampling strategies and reconstruction performance. More specifically, image reconstruction quality depends on the employed coding of the input scene.

This dissertation aims at exploring different optical coding strategies in CSI systems. A first strategy considers CSI architectures with coded apertures as the coding element. The structure of the coded apertures is crucial inasmuch as they determine the entries of the sensing matrix. Given that conventional coded aperture patterns are selected at random, leading to suboptimal reconstruction solutions, the proposed strategy exploits the restricted isometry property of the sensing matrix and its incoherence with respect to the sparse representation basis to optimize the coded aperture ensemble.

On the other hand, a different coding strategy is introduced motivated by the complicated optical paths of state-of-the-art CSI systems that compromise their portability. In this case, we propose a new compact CSI architecture that exploits the benefits of colored mosaic FPA detectors and the compression capabilities of CSI sensing
techniques. The optical and the mathematical models of the explored encoding strategies are presented along with testbed implementations of the systems. Simulations and experimental data evaluate the accuracy of the proposed strategies.

In addition to the optical coding strategies, this dissertation studies the multi-resolution (MR) reconstruction problem in compressive spectral imaging. To date, the common CSI reconstruction framework has focused on obtaining an approximation of the underlying spatial and spectral information of a scene from a set of coded projections, where the resolution of the reconstruction is as high as the measurements allow. Even though CSI enables fast multiplexed sensing, the complexity of the inverse problem depends on the spatial and spectral resolution of the data to be recovered. Motivated by such cases where a fast preview reconstruction is often desirable, we propose a multi-resolution reconstruction scheme for CSI that enables the sequential recovery of multiple versions of the same data cube at different spatial resolutions. Simulations are developed to analyze the performance of the proposed MR reconstruction model.
Chapter 1
INTRODUCTION

Spectral imaging or imaging spectroscopy involves the sensing of 2D spatial information along multiple wavelengths. The resulting data set provides a whole spectrum at each spatial position, therefore these images are modeled as data cubes containing the spectral signatures of an area of interest. This characteristic makes spectral imagery suitable for a wide range of applications related to material identification and anomaly detection. For instance, in agriculture and geophysics applications, remote sensed spectral information of soil or water is employed for monitoring wetlands and crops [1, 2, 3, 4], or detection/tracking of oil spills [5, 6]. In medical diagnosis, the reflectance spectra of biological tissue allows the detection of malign cells or tumors [7, 8, 9, 10]. Moreover, in surveillance/defense and industrial applications, spectral imagery provides information for target detection and classification [11, 12, 13, 14]. The advantages of spectral images have been also exploited for artwork restoration [15, 16].

1.1 Spectral Imaging
1.1.1 Traditional Spectral Imaging Methods

Common spectral imaging techniques can be classified as scanning, snapshot or filter-based methods. Scanning methods capture the information along a subset of dimensions and then merge the data to construct a 3-dimensional (3D) data cube, having two spatial dimensions and one spectral coordinate \((x, y, \lambda)\) [17]. For instance, whiskbroom sensors employ a pinhole aperture for point spectrum acquisition; pushbroom spectrometers, on the other hand, use a slit aperture to acquire the information along the \(x - \lambda\) or \(y - \lambda\) planes; filtered-based cameras capture one monochromatic
image for each desired spectral band using different filters. In general, for the scanning methods, the desired spatio-spectral resolution determines the required amount of scanned areas, which in turn increases the acquisition time. For this reason, these systems experience low sensing speed, and high calibration complexity.

Another type of spectral imaging architectures consists on capturing the spatio-spectral information of the scene in a single snapshot. This type of architectures typically maps all the voxels of the spectral data cube into different regions of a focal plane array (FPA). Single snapshot architectures include the image mapping spectrometers (IMS) [18], field sampling snapshot spectrometers [19], and the image-replicating imaging spectrometers (IRIS) [20]. However, since a single detector captures all the voxels from the spectral scene, these spectrometers require large detectors, where a portion of the area is not used to keep track of the mapping scheme and recover the 3D cube. In addition, most of these architectures require elaborated optical paths, and have high implementation costs [21, 22].

A different spectral imaging approach relies on Fabry-Perot filters or colored mosaic FPA detectors to obtain small sets of spectral bands [21, 23, 24, 25]. In this approach, each pixel in the FPA is assigned an individual spectral response, such that a specific range of wavelengths is captured. Since the colored FPA is the only sensing device, there exists a trade-off between the spatial and the spectral resolutions, which are determined by the number of available filters. Assuming an $N \times N$ FPA with $d$ uniformly distributed filters, there are $\lfloor N^2/d \rfloor$ FPA pixels available to recover each spectral band. In other words, a data cube with $\lfloor N/\sqrt{d} \rfloor$ pixels of spatial resolution can be recovered when no interpolation nor demosaicking method is applied [26]. Thus, increasing the number of colors leads to a larger spectral resolution but smaller spatial resolution. Similarly, decreasing the number of colors increases the spatial resolution but reduces the number of attainable spectral bands. As a consequence of the drawbacks of the aforementioned spectral imaging techniques, different
strategies and optical architectures based on the compressive sensing (CS) theory have been proposed for the acquisition of spatial and spectral data. This field is known as compressive spectral imaging.

1.1.2 Compressive Spectral Imaging Systems

Compressive Spectral Imaging (CSI) is a recently emerged field that allows to sense spatial and spectral information of a scene using 2-dimensional sets of coded projections. Even though CSI devices also capture a single snapshot, their main advantage is that they acquire far fewer samples than the traditional techniques, which leads to improved sensing speed [22]. Furthermore, CSI relies on compressed sensing (CS) principles: sparsity, which characterizes the spectral scenes of interest, and incoherence, which shapes the sensing structure [27, 28]. Specifically, sparsity of spectral images indicates that they can be represented in some basis $\Psi$, using just a small number of coefficients, due to the high correlation among adjacent pixels and also across spectral bands [29]. On the other hand, incoherence refers to the structure of the sampling waveforms used in CS that, unlike the signals of interest, have a dense representation in the basis $\Psi$ [28]. The relevance of the CS framework is that it is possible to design sensing protocols capable of capturing the essential information content in sparse signals with just a small number of compressive measurements. The sensing scheme correlates incoming signals with a small number of fixed waveforms that satisfy the incoherence principle. The signals of interest can then be accurately reconstructed from the compressive measurements using numerical optimization algorithms [30, 31]. In the particular case of CSI devices, the projections occur as the result of the optical phenomena affecting the incoming light as it passes through different optical devices before being integrated by an imaging detector.

One of the most representative CSI architectures is the coded aperture snapshot spectral imager (CASSI), described in detail in Chapter 2, which comprises a dispersive element and a coded aperture [29, 32, 33]. In CASSI, the spatio-spectral 3-dimensional
source $F \in \mathbb{R}^{N \times N \times L}$ is first encoded by a coded aperture \( T \in \mathbb{R}^{N \times N} \); then, the coded source is dispersed by a prism and, finally the coded and dispersed source is captured by an intensity detector. Typically, the coded apertures are block-unblock lithographic masks or spatial light modulators (SLM) that offer good optical sensing properties \cite{22}. 

The binary patterns on the lithographic masks are fabricated using chrome-on-quartz \cite{29}, rendering a 2-dimensional coded aperture whose elements are either transparent or opaque, such that all the spectral bands are blocked or pass through each spatial position of the mask. In this sense, the source is uniformly encoded across a range of wavelengths. A different variation of the CASSI system replaces the binary masks by colored coded apertures. This type of coded apertures consists of a tiling of optical filters that improve the coding scheme of the scene and reduce the required number of projections for correct reconstruction \cite{34, 35}. Other CSI architectures include the dual-coded hyperspectral imager (DCSI) proposed in \cite{36}, that separately encodes the spatial and spectral dimensions by means of a diffraction grating and two spatial light modulators which results in a complicated optical path. The spatial-spectral encoded compressive hyperspectral imaging (SSCSI) presented in \cite{37} uses a diffraction grating in conjunction with a static coding mask that encodes each spectral channel independently, before the spatio-spectral source is captured on the detector. Additional CSI architectures have been proposed in \cite{38, 39}.

The sensing mechanism of these architectures is modeled as the linear system given by $y = Hf$, where $f$ is a vector form of the spatial and spectral data cube, $H$ is the sensing matrix accounting for the optical phenomena induced by the employed devices, and $y$ is the set of coded projections. Because the amount of measurements is less than the number of unknowns, image recovery entails solving an underdetermined system of linear equations. Typically, the inverse problem consists on recovering a sparse representation of the object $\theta = \Psi^T f$ by minimizing the $\ell_1 - \ell_2$ cost function given by

$$\arg \min_\theta ||y - \Psi \theta||_2^2 + \tau ||\theta||_1,$$  \hspace{1cm} (1.1)
where $\| \cdot \|_2$ and $\| \cdot \|_1$ represent the $\ell_2$ and $\ell_1$ norm, respectively; and $\tau$ is a regularization constant. Other approaches employ different cost functions including low-rank formulations and approximate message passing (AMP)\[40, 41, 42, 43].

1.1.3 Multi-Resolution Compressed Sensing Reconstructions

In general, the common CSI reconstruction framework has focused on obtaining approximations of the underlying spatial and spectral information of a scene from a set of coded projections, where the resolution of the reconstruction is as high as the measurements allow. Even though CSI enables fast multiplexed sensing, the computational complexity of the inverse problem depends on the spatial and spectral resolution of the data to be recovered. Several strategies have been proposed to alleviate the computational load of CS reconstructions, including GPU implementations of reconstruction algorithms [44], separable sensing operators [45, 46, 47] and block reconstruction approaches [48]. More recently, multi-resolution (MR) reconstructions have emerged as an alternative for alleviating the computational load of recovering the underlying signals. Specifically, instead of recovering a full-resolution version of the signal, the MR approach consists on recovering lower resolution versions or signal previews. For instance, in [49], a compressed sensing framework that allows fast preview image reconstructions using a sum-to-one (STOne) transform has been proposed. In spite of the benefits of this framework, it can not be directly applied to compressive spectral imaging, because it involves the design of the sensing matrix, while in CSI the structure of the sensing matrices is determined by the optical architecture. Moreover, a more general multi-resolution compressed sensing approach was presented in [50]. More specifically, this approach establishes that low-resolution reconstructions can be obtained from a set of projections of a high-resolution object by introducing a pair of decimation/upscaling matrices such that the reconstruction problem in Eq. (1.1) is solved using the equivalent low-resolution sensing matrix. In general, the decimation matrix involves the creation of super-pixels in the target LR signal. However, a method for obtaining up-scaled versions from the LR reconstruction is not described.
On the other hand, the high dimensionality of spectral images has motivated the development of techniques to alleviate the cost of managing such amounts of data. One of these approaches employs image super-pixels that group similar pixels with respect to a specific criterion, such that the resulting dimension of the signal is reduced and facilitate tasks such as classification and detection [51, 52, 53, 54]. Common super-pixel segmentation techniques include the simple linear iterative clustering (SLIC) [55] and entropy rate segmentation [56]. However, these applications are not focused on compressed sensing of spectral images.

1.2 Motivation

Compressive spectral imaging projections, in general, consist on the integration of encoded and dispersed versions of the input source. To this end, CSI systems entail the use of a dispersive element to decouple the spectral components of the scene, and an optical encoding element that can be either a spatial modulator or a spatial-spectral modulator. Thus, different optical configurations yield different sampling strategies and reconstruction performance. Moreover, implementation costs and complexity of the optical path are directly related to the employed optical devices. For instance, Fabry-Perot based spectral imaging systems do not require additional optical elements such as a coded aperture or a dispersive element, which makes them very compact and portable, but their principal drawback is the limited resolution. CSI systems, on the other hand, alleviate dependencies between spatial and spectral resolutions by means of the coded aperture and the dispersive element.

Given that the employed coding of the input scene is crucial for rendering good quality and accurate image reconstructions, this dissertation explores different optical coding strategies for CSI systems. Specifically, better optical coding schemes result in improved compressive measurements which in turn, yield better spectral image reconstructions.

Moreover, in spite of good CSI optical encoding strategies which enable fast multiplexed sensing, the computational complexity of the inverse problem depends on the
spatial and spectral resolution of the data to be recovered. To date, most of the research in compressed sensing has focused on the recovery of full resolution signals, whose resolution is as high as the measurements allow. Typical compressed sensing reconstruction algorithms employed to recover the underlying data cube perform computations such as matrix pseudo-inverses, sparse basis transformations, and vector-matrix multiplications. Given the high-dimensionality of spectral images, the reconstruction process can be overwhelming, and is usually performed offline. Therefore, this dissertation studies a multi-resolution (MR) reconstruction approach for compressive spectral imaging such that multiple versions of the underlying data cube at different spatial resolutions can be sequentially recovered up to the maximum resolution enabled by the acquired projections.

1.3 Dissertation Organization

This dissertation is organized as follows:

Chapter 2 considers CSI systems with coded apertures as the main encoding device such as CASSI, and presents the spatio-temporal blue noise (BN) coded apertures. In particular, this design is mainly motivated by the fact that the only varying components in these architectures are the coded aperture patterns, whose structure is crucial inasmuch as it determines the minimum number of FPA measurements needed for correct image reconstruction and the corresponding attainable quality. Traditionally, the spatial structure of the coded aperture entries is selected at random leading to suboptimal reconstruction solutions. The BN coded apertures are designed by exploiting the Restricted Isometry Property (RIP) of the sensing matrix and its incoherence with respect to the sparse representation basis. This chapter also presents the CASSI matrix system representation in terms of the ensemble of random projections. Also, the RIP of the CASSI projections is determined as a function of the coded aperture entries. Furthermore, a computational algorithm that implements the BN coded aperture ensembles is developed. Extensive simulations and a testbed implementation illustrate the improvements of the BN coded apertures over the traditionally used random coded...
aperture structures, in terms of spectral image reconstruction peak-signal to noise ratio (PSNR) and structural similarity (SSIM).

Chapter 3 presents a compact CSI optical architecture called snapshot colored compressive spectral imager (SCCSI) that exploits the benefits of colored mosaic FPA detectors and the compression capabilities of CSI sensing techniques. Specifically, SCCSI entails the use of a dispersive element in conjunction with a color-patterned detector to capture the spatial and spectral information of a scene in a single snapshot. This chapter includes the optical and the mathematical models of SCCSI along with a testbed implementation of the system. Simulations and testbed experiments show the accuracy of SCCSI and compare the reconstructions with those of similar CSI optical architectures, such as the CASSI and SSCSI systems.

Chapter 4 presents a multi-resolution (MR) reconstruction approach to recover different versions of a data cube from the originally captured set of coded projections. Each of those versions at a different spatial resolution. Specifically, the proposed approach employs an initial low-resolution (LR) reconstruction of the scene at a small scale to estimate the high frequency components of the next scale reconstruction such that the latter can be recovered by employing super-pixels on the low-frequency areas of the scene. The gradient intensity image is used to develop a decimation matrix that modifies the original sensing matrix to generate the equivalent MR sensing matrix that is later used in a CS reconstruction algorithm. Then, the same process is performed up to reaching the maximum resolution allowed by the captured projections. Simulation results are used to analyze the performance of the proposed MR approach.

1.4 Contributions

The contents of this dissertation have been published on the following journals and conferences:


Chapter 2

SPATIO-TEMPORAL BLUE NOISE CODED APERTURES FOR COMPRESSIVE SPECTRAL IMAGING

The coded aperture snapshot spectral imager (CASSI) is a CSI architecture that effectively captures the 3D information using a single 2D projection measurement. Specifically, the image source density $f_0(x, y, \lambda)$ is encoded by means of a coded aperture $T(x, y)$ as depicted in Fig. 2.1, where $(x, y)$ are the spatial coordinates and $\lambda$ represents the wavelengths [29, 57, 58]. The resulting coded field $f_1(x, y, \lambda)$ is subsequently sheared by a dispersive element before it impinges onto the FPA detector. The compressive measurements across the FPA are realized by the integration of the field $f_2(x, y, \lambda)$ over the spectral range sensitivity of the detector. Assuming that the detector and the coded aperture have a pixel pitch $\Delta_d$, and a band-pass filter of the instrument limits the spectral components between $\lambda_1$ and $\lambda_2$, the number of resolvable bands $L$ is limited by $L = \beta \frac{\Delta_d - \Delta_1}{\Delta_d}$ where $\beta\lambda$ is the dispersion induced by the prism, when a linear dispersion function is considered. Similarly, the spectral resolution is limited by $\frac{\Delta_d}{\beta}$. Thus, the spatial resolution in the horizontal and the vertical axes are limited by $\Delta_d$ and, the number of spatially resolvable pixels of the underlying scene is $N \times N$.

It has been shown that for spectrally rich scenes or very detailed spatial scenes, a single shot CASSI measurement may not provide a sufficient number of compressive measurements [59]. Increasing the number of measurement snapshots, each using a different coded aperture that remains fixed during the integration time of the detector, will rapidly increase the quality of image reconstruction. Given that each CASSI snapshot simultaneously adds $N(N + L - 1)$ compressive measurements, the total number of available measurements when $K$ shots are taken is $m = KN(N + L - 1)$. 
Figure 2.1: Optical elements present in CASSI. The input scene $f_0(x, y, \lambda)$ is first spatially encoded by the coded aperture $T(x, y)$ and then, the encoded source is dispersed by a prism. Finally, the coded and dispersed scene is captured by the focal plane array.

The time-varying coded apertures can be realized by moving a large photomask using a piezo-electric system [59]. Moreover, a more versatile system was developed in [33], in which a digital micromirror device (DMD) was used to vary the random patterns on each snapshot.

The theory of CS is used to recover the 3D spatial-spectral data cube from the set of FPA compressive measurements. In general, the reconstruction procedure consists on finding the sparsest approximation of the data cube with the minimum Euclidean distance to the 2D random projections. Mathematically, let $F$ be the $N \times N \times L$ spectral data cube, or its vector representation $f$, then the $i$–th FPA measurement is obtained as $y^i = H^i f$, where $H^i$ is a $N(N + L - 1) \times N^2 L$ projection matrix representing the effects of the coded aperture and the dispersive element. The set of $K$ FPA measurements, each with a different coded aperture, can be stacked one after another as $y = [(y^0)^T, \ldots, (y^{K-1})^T]^T$, therefore the projection matrix for all the snapshots can be written as $H = [(H^0)^T, \ldots, (H^{K-1})^T]^T$. The underlying data cube is estimated as $\tilde{f} = \Psi (\arg\min_{\theta} ||y - H\Psi\theta||_2 + \tau||\theta||_1)$, where $\theta$ is an $S$-sparse representation of $f$ on the basis $\Psi$, and $\tau$ is a regularization constant.

The product between the projection matrix and the sparse representation basis is known as the sensing matrix and is given by $A \triangleq H\Psi$. The structure of $A$ is a critical component in the inverse problem that estimates $f$ from $y$, because it determines the
attainable quality of the reconstructions. Typically, the sparse representation basis is fixed according to the data to be acquired. Therefore, the structure of $A$ is determined by the structure of $H$, which is partially defined by the optical architecture. In the particular case of the CASSI system, $H$ has a well defined sparse structure, and its non-zero entries are determined by the coded apertures used on each measurement snapshot. Consequently, the objective in CASSI is to design the set of coded apertures so as to forge a structure on $A$ that minimizes the number of FPA snapshots while attaining the highest reconstruction quality. In this work, the Restricted Isometry Property (RIP) is used as the tool to explore the relation between the CASSI sensing matrix and the coded apertures.

Specifically, the RIP provides a characterization of CS sensing matrices. It establishes the necessary conditions for $A$ such that the $\ell_2$ norm of the underlying 3D spectral image is approximately preserved under the transformation $A\theta$. In addition, the RIP also implies a stable recovery of the signal $\theta$ from the projections $A\theta$ using an $\ell_1$ optimization algorithm [60]. More precisely, the restricted isometry constant $\delta_s$ of the matrix $A$ is the smallest constant such that the inequality $(1 - \delta_s)||\theta||_2 \leq ||A\theta||_2 \leq ||\theta||_2(1 + \delta_s)$ holds [60]. To satisfy the RIP, it is required that all $m \times |T|$ column submatrices of $A$, $A_{|T|}$ are well conditioned for all $|T| \leq S$.

In addition, the RIP imposes that all the eigenvalues of the matrices $A_{|T|}^TA_{|T|}$ are in the interval $[1 - \delta_s , 1 + \delta_s]$. The probability of satisfying this condition is calculated by estimating the statistical distribution of the maximum eigenvalue $\lambda_{max}$ of the gram matrices $A_{|T|}^TA_{|T|} - I$, where $I$ is an identity matrix. The concentration of measure for random matrices developed in [61] can be used to estimate the distribution of the maximum eigenvalue $\lambda_{max}$ of $A$.

Given that random matrices have well established RIP properties, CS sensing matrices are frequently constructed from realizations of common random variables, including unitary dense Gaussian random matrices whose entries are drawn from zero mean Gaussian random variables with variance $1/m$; Bernoulli random matrices with entries taking equiprobable values $\{-1/\sqrt{m}, 1/\sqrt{m}\}$. Specific applications employ
structured random sensing matrices for which the RIP has been also previously analyzed; for instance in sparse channel estimation and multi-user detection the sensing matrix can be a Toeplitz matrix with random Gaussian entries [62, 63], a random circulant matrix [60], or a partial random Toeplitz matrix [64]. Sensor networks and multiple view imaging employ block diagonal random matrices (RBD), whose RIP has been studied in [65]. Furthermore, a more general framework for structurally random matrices (SRM) has been developed in [66], where SRM matrices are defined as the product $\Phi = \sqrt{n/m} D FR$, where $R$ is a random permutation matrix, $F$ is a fast computable matrix, $n = N^2 L$, and $D$ is a subsampling matrix. SRM can be designed with high flexibility using different combinations of fast computable matrices and block-based processing. The RIP for other matrix structures has been studied in [67, 64].

However, given that the CASSI sensing matrix is highly sparse and structured, the RIP characterization for dense sensing matrices is not applicable. To date, different studies have focused on the RIP analysis for the CASSI system. The first approach to characterize the RIP in CASSI was presented in [57], where it was beforehand assumed that the RIP for the matrix $A$ is satisfied for some constant $\delta_s$, then conditions on the coded apertures were determined so that the RIP is better satisfied. These results, however, do not involve the RIP in the optimal design of the coded apertures. In fact, frequently used coded apertures in CASSI are: Hadamard matrices $H_N$ whose entries are $(H_N)_{ij} \in \{-1,1\}^{N \times N}$, Hadamard S matrices $S_N$ where $S_N = 1/2(1 - H_N)$ [68], cyclic S-matrices consisting of cyclic permutations of a single master codeword [69], and Bernoulli random matrices [70, 33]. The use of these coded apertures in CASSI has been principally motivated by the fact that they are well conditioned when used in a least squares estimation [68, 71]. Nonetheless, these coded apertures do not fully exploit the rich theory of CS. In particular, they do not exploit the RIP condition or the concentration of measure of the respective random submatrices of $A$ to define optimal coded aperture sets. A more recent coded aperture design, based on the concentration of measure can be found in [22], where the correlation between measurement snapshots is taken into account to better satisfy the RIP. Despite the improved reconstructions
that can be obtained with this design, its main weakness is that the spatial correlation of the one-valued entries is not considered. The design in [22] was coined boolean coded apertures.

In this chapter, the structure of the multi-shot CASSI sensing matrix is formulated, then, the RIP for CASSI sensing matrices is derived and, the corresponding RIP constants are expressed as a function of the structure of the coded aperture patterns. It will be shown that the reconstruction quality in CASSI can be improved either by increasing the number of FPA projections or designing the coded aperture set, which results in better satisfying the RIP conditions. The resulting optimal set of coded apertures are coined blue noise (BN) coded apertures, since their distribution exhibits spatio temporal characteristics of blue noise patterns that suppress low frequency components of noise [72, 73, 74]. These coded apertures allow uniform sensing across spatial and spectral dimensions of the scene within different measurement snapshots. BN coded apertures are then compared with traditional random codes apertures used in CASSI. Simulations are shown to illustrate the benefits of this design. In addition, experimental reconstructions using BN coded apertures in a testbed implementation of the CASSI system are presented.

2.1 Multi-Shot CASSI Matrix Representation

Let \( T_{ij} \) be the \((j, \ell)\)-th spatial index of a discretized representation of the coded aperture used to sense the \(i\)-th FPA measurement, and let \( F_{j,\ell,k} \) be the spatio-spectral source with \(k\) determining the \(k\)-th spectral plane. The discrete output at the detector is given by [29, 59, 70]

\[
Y_{ij} = \sum_{k=0}^{L-1} F_{j,(\ell-k),k} T_{ij,(\ell-k)} + \omega_{ij} \quad i = 0, \ldots, K - 1, \tag{2.1}
\]

where \(i\) indexes the FPA snapshot, \(Y_{ij}\) is the intensity measured at the pixel \((j, \ell)\) of the FPA detector, and \(\omega_{ij}\) is the noise of the system. The coded apertures \(T_{ij}\) are time-varying and indexed by \(i\). The output of the system \(Y^i\) in (2.1) is a \(N \times V\)
signal, with \( V = N + L - 1 \), yielding the compression ratio \( NV/N^2L \approx 1/L \). The FPA measurement can be written in vector notation as

\[
y^i = H^i f + \omega^i
\]  

(2.2)

where \( y^i \in \mathbb{R}^{NV} \) is a \( NV \)-long vector representation of \( Y^i \), \( H^i \in \{0, 1\}^{NV \times n} \) represents the effects of the coded aperture and the dispersive element, and \( f = \text{vec}([f_0, \ldots, f_{L-1}]) \) is the vector representation of the data cube \( F \), where \( f_k \) is the vector form of the \( k \)-th spectral band. More specifically, the entries of \( f_k \) can be expressed as \( (f_k)_{\ell} = F_{(\ell-rN),r} \) for \( \ell = 0, \ldots, N^2 - 1, \ k = 0, \ldots, L - 1 \), where \( r = \lfloor \frac{\ell}{N} \rfloor \). Using this definition of \( r \), the vectorization of the coded aperture \( T^i_{j,\ell} \) can be defined as \( (t^i)_{\ell} = T^i_{(\ell-rN),r} \) for \( \ell = 0, \ldots, N^2 - 1, \ i = 0, \ldots, K - 1 \). Similarly, the vector representation of the output \( Y^i \) is written as \( (y^i)_{\ell} = Y^i_{(\ell-rN),r} \) for \( \ell = 0, \ldots, NV - 1, \ i = 0, \ldots, K - 1 \). Based on the above matrix representation, the output \( y^i \) in (2.2) can be expressed as

\[
y^i = \begin{bmatrix}
\text{diag}(t^i) & 0_{N(1) \times \ldots \times N(L-1) \times N^2} & \ldots & 0_{N(L-1) \times N^2} \\
0_{N(L-1) \times N^2} & \text{diag}(t^i) & \ldots & \vdots \\
0_{N(L-1) \times N^2} & 0_{N(L-2) \times N^2} & \ldots & \text{diag}(t^i)
\end{bmatrix}
\begin{bmatrix}
f_0 \\
f_1 \\
\vdots \\
f_{L-1}
\end{bmatrix},
\]

(2.3)

where \( \text{diag}(t^i) \) is an \( N^2 \times N^2 \) diagonal matrix whose entries are the elements of the vectorized coded aperture \( t^i \), \( 0_{N(1) \times N^2} \) and \( 0_{N(L-1) \times N^2} \) are \( N(1) \times N^2 \) and \( N(L - 1) \times N^2 \) zero-valued matrices, respectively. The matrices for multiple snapshots \( H^i \) are assembled in the matrix \( H \) given by \( H = [(H^0)^T, \ldots, (H^{K-1})^T]^T \). Figure 2.2 illustrates the structure of the sensing matrix \( H \) for \( N = 8, \ L = 4 \) and \( K = 2 \), in which the entries \( (t^i)_{\ell} \) are realizations of a Bernoulli random variable with parameter \( p = 0.5 \). In practice, the dimensions of the matrices \( H \) and \( A \) are much larger [70, 59].
be observed in Fig. 2.2 that the $j-$th row of $H$ contains at most $L$ non-zero elements. Further, the entries of the $j-$th row of the sensing matrix $H$ can be written as

$$
(h_j)_\ell = \begin{cases} 
t_{j-rN}, & \text{if } \ell = (j \mod N') \text{ and } j - rN \geq 0 \\
0, & \text{otherwise}, 
\end{cases} \quad (2.4)
$$

for $\ell = 0, \ldots, N^2L - 1$, where $r = \left\lfloor \frac{\ell}{N^2} \right\rfloor$, $N' = N^2 - N$ and $i = \lfloor j/NV \rfloor$.

**Figure 2.2:** The structure of the matrix $H$ is illustrated for $K = 2, N = 4$, and $L = 4$. For illustration, the entries of the coded apertures are $t_{ij} \sim \text{Bernoulli}(0.5)$. Dark elements represent zero values and white elements correspond to the one-valued entries of the coded aperture.

The ensemble of CASSI outputs $y = [(y^0)^T, \ldots, (y^{K-1})^T]^T$ can be rewritten as

$$y = A\theta = H\Psi\theta + \omega, \quad (2.5)$$

where $A = H\Psi$ is the CASSI sensing matrix, $\theta$ is a sparse representation of $f$ in the basis $\Psi$, and $\omega$ represents the noise of the system. Notice that $A \in \mathbb{R}^{m \times n}$, where $m = KNV$, $n = N^2L$, and $m \leq n$. For instance, the sensing matrix $A$ can be a
Kronecker representation basis given by $\Psi = \Psi^{2D} \otimes \Psi^C$, where $\Psi^C$ is the 1D cosine matrix transform, $\Psi^{2D}$ is the 2D Symmlet Wavelet transform, which treats each spectral band independently and $\otimes$ is the Kronecker product. In this case, the entries of $\theta = \Psi^T f$ account for the correlation among all the elements in the data cube $f$. In summary, given that $\Psi$ is fixed, the only variable elements in $A$ correspond to the coded aperture entries $T^i_{j,\ell}$ that in turn, determine $H$. Thus, the goal of this work is to design $T^i_{j,\ell}$ such that the sensing matrix $A$ is better conditioned in order to improve the attainable reconstruction quality.

2.2 Restricted Isometry Property in Multi-Shot CASSI

The restricted isometry property (RIP) provides guidelines to determine the minimum number of measurements needed for signal reconstruction in CS. In CASSI, the RIP is critical to design optimal coded aperture ensembles in order to maximize the quality of the reconstructions. In particular, the RIP of the CASSI sensing matrix $A$ of order $S$ is defined as the smallest constant $\delta_S$ such that $$(1 - \delta_S) \|\theta\|_2^2 \leq \|\theta\|_2^2 \leq (1 + \delta_S) \|\theta\|_2^2$$ holds for all $S$-sparse vectors $\theta$. The RIP constant $\delta_S$ is given by

$$\delta_S = \max_{T \subset [n], |T| \leq S} \|A^T_{|T|}A_{|T|} - I\|_2, \quad (2.6)$$

the operator $\|\cdot\|_2^2$ is the squared norm from $\ell_2$ into $\ell_2$, $A_{|T|}$ is a $m \times |T|$ matrix whose columns are $|T|$ columns of the CASSI matrix $A$ indexed by the set $\Omega$, and $I$ is an identity matrix [75, 60]. In other words, the RIP determines whether all column submatrices $A_{|T|}$ of $A$ are well conditioned. The RIP constant in (2.6) can be rewritten in terms of the eigenvalues of $A_{|T||T|} = A^T_{|T|}A_{|T|}$ as

$$\delta_S = \max_{|T| \subset [n], |T| \leq S} \sqrt{\lambda_{\text{max}}(A_{|T||T|} - I)}, \quad (2.7)$$

where $\lambda_{\text{max}}(.)$ denotes the largest eigenvalue [60]. Given that the sensing matrix $A$ is determined by the matrix $H$ and the basis representation matrix $\Psi$, the entries of $A_{|T|}$ can be expressed as the product of the entries of $\Psi$ and $H$. Let the columns of $\Psi$ be...
written as $[\psi_0, \ldots, \psi_{n-1}]$. Using the structure of $H$ in (2.4), the entries of $A_{|T|}$ can be expressed as the product of the rows of $H$ and the columns of $\Psi$ indexed by the set $\Omega$

$$(A_{|T|})_{j,k} = h_j \psi_{\Omega_k} = \sum_{r=0}^{L-1} T^i_{[j-rN_{-1}], j-rN_{-1}} \psi_{j+rN_{-1}, \Omega_k}, \quad (2.8)$$

for $j = 0, \ldots, m-1$, $k = 0, \ldots, |T| - 1$, where $i = \lfloor j/N \rfloor$, $N' = N^2 - N$, and $\Omega_k \in \{0, \ldots, n-1\}$.

Given that the entries of $A_{|T||T|}$ are given by the product of two columns of $A_{|T|}$, (2.8) can be used to express the elements of $A_{|T||T|}$ as

$$(A_{|T||T|})_{j,k} = \sum_{i=0}^{K-1} \sum_{\ell_1=0}^{N-1} \sum_{\ell_2=0}^{L-1} \sum_{r=0}^{L-1} \sum_{u=0}^{L-1} T^i_{i, \ell_1, \ell_2 - r} T^i_{i, \ell_1, \ell_2 - u} \times \psi_{\ell_2 N + \ell_1 + rN', \Omega_j} \psi_{\ell_2 N + \ell_1 + uN', \Omega_k}, \quad (2.9)$$

for $j, k = 0, \ldots, |T| - 1$.

A necessary condition for satisfying the RIP is that the expected value of the diagonal terms of $A_{|T||T|}$ be equal to one, $E \left( (A_{|T||T|})_{s,s} \right) = 1$ for $s = 0, \ldots, |T| - 1$. This condition holds when the coded aperture entries are constrained to satisfy $\sum_{i=0}^{K-1} \left( T^i_{\ell_1, \ell_2 - r} \right)^2 = C$, with $C$ a selectable constant. Hence, $A_{|T||T|}$ can be normalized by defining the matrix $B_{|T||T|} = A_{|T||T|}/C$. Using (2.9), the normalized matrix can be written as

$$(B_{|T||T|})_{j,k} = \sum_{i=0}^{N-1} \sum_{\ell_1=0}^{L-1} \sum_{\ell_2=0}^{L-1} \sum_{r=0}^{L-1} \sum_{u=0}^{L-1} \gamma_{\ell_1, \ell_2, u} \phi_{\ell_1, \ell_2, u} / C, \quad (2.10)$$

where $\gamma_{\ell_1, \ell_2, u} = \sum_{i=0}^{K-1} T^i_{\ell_1, \ell_2 - r} T^i_{\ell_1, \ell_2 - u}$ and $\phi_{\ell_1, \ell_2, u} = \psi_{\ell_2 N + \ell_1 + rN', \Omega_j} \psi_{\ell_2 N + \ell_1 + uN', \Omega_k}$.

Equation (2.10) can be analyzed from two perspectives: the first one assumes that $\phi_{\ell_1, \ell_2, u}$ is a constant in order to study the effect of the term $\gamma_{\ell_1, \ell_2, u}$; the second perspective studies the effect of the coded apertures in the coherence between $H$ and
Given that the elements of $\Psi$ are fixed and bounded, the sparse representation basis term $\phi_{\ell_1,\ell_2,r,u}$ is also bounded. More specifically, $|\phi_{\ell_1,\ell_2,r,u}| < C_1$, for all $\ell_1, \ell_2, r$ and $u$. Thus, (2.10) can be rewritten as 

$$(B_{|T||T|})_{j,k} \leq \frac{C_1}{C} \sum_{\ell_1=0}^{N-1} \sum_{\ell_2=0}^{V-1} \sum_{r=0}^{L-1} \sum_{u=0}^{L-1} \gamma_{\ell_1,\ell_2,r,u}.$$ 

In consequence, the elements $(B_{|T||T|})_{j,k}$ are the sum of bounded random variables and they can be modeled as a sub-Gaussian random variable $(B_{|T||T|})_{j,k} \sim \text{Sub}(\alpha^2)$ with parameter given by 

$$\alpha = \max_{j,k} \frac{C_1}{C} \sum_{\ell_1=0}^{N-1} \sum_{\ell_2=0}^{V-1} \sum_{r=0}^{L-1} \sum_{u=0}^{L-1} \gamma_{\ell_1,\ell_2,r,u}. \quad (2.11)$$ 

Previous work on the RIP for sub-Gaussian random variables have established that 

$$P \left( \|B_{|T||T|} - I\| \leq \delta_S \right) \geq 1 - \varepsilon,$$ 

where $\varepsilon = 2 \left( 1 + 2/\rho \right)^2 e^{-\delta_S (2 - (1 + \rho)^2)^2 K N V c_2 / \alpha}$, with $\rho = 2 / (e^3 - 1)$ and $c_2$ a constant independent of $K$ and $V$ [76, 34]. Therefore, the error probability $\varepsilon$ is minimized either when the number of measurement shots $K$ is increased, or when the coded aperture ensemble is designed such that the sub-Gaussian parameter $\alpha$ is minimized. It can be noticed in (2.11), that the minimization of $\alpha$, based on the design of the coded apertures, is directly related to the minimization of the variable $\gamma_{\ell_1,\ell_2,r,u} = \sum_{i=0}^{K-1} T^i_{\ell_1,\ell_2-r} T^i_{\ell_1,\ell_2-u}$, which depends on the horizontal separation of the one-valued entries of $T^i$, and their correlation across frames.

Equivalently to the RIP, the notion of coherence provides an alternative approach for evaluating the properties of $A$ that guarantee the correct recovery of the underlying signal. In particular, CS requires the acquired measurements to be uncorrelated, and the coherence is used to measure the correlation between $H$ and $\Psi$, such that CS conditions hold. Recently, the coherence has been defined as the largest absolute inner product between any two columns of the sensing matrix [77]. More specifically, the coherence of the CASSI sensing matrix can be calculated as

$$\mu \triangleq \max_{i \neq j} \left| \text{TPSF} (i,j) \right| / \sqrt{\text{TPSF} (i,i) \text{TPSF} (j,j)}, \quad (2.12)$$

where TPSF $(i,j) = (\Psi^T H^T H \Psi)_{i,j}$ is a transform point spread function as defined in [78, 79]. In other words, the coherence is calculated as the maximum off-diagonal entry
of the TPSF matrix. Given that low correlation values are desired, one can define the incoherence as $\hat{\mu} = 1 - \mu$ for simplicity in evaluating low correlation values between $\textbf{H}$ and $\Psi$.

### 2.3 Coded Aperture Design Based on RIP

Results from the previous section provide useful guidelines for the design of coded aperture ensembles that help in satisfying the RIP and incoherence conditions, such that better sensing matrices can be designed in order to improve image reconstructions. In particular, the following design criteria are taken into account to generate optimized coded aperture ensembles:

a). **Horizontal separation:** The minimization of the sub-Gaussian parameter $\alpha$ in (2.11) is determined by the term $\gamma_{\ell_1,\ell_2,r,u}$. More specifically, given that $r, u \in \{0, \cdots, L - 1\}$, the products $T^i_{\ell_1,\ell_2-r}T^i_{\ell_1,\ell_2-u}$ should be minimized within a horizontal neighborhood of size $L$ of the coded aperture $\textbf{T}^i$. It can be noticed that this minimization is achieved when the one-valued entries of $\textbf{T}^i$ within the same row are maximally separated. This condition, in essence, requires the codes to satisfy spatial blue noise pixel distribution.

To illustrate the effect of this condition, different coded aperture ensembles are employed to determine $\textbf{H}$ and $\alpha$ from (2.11) is evaluated for the corresponding matrices with a fixed $\Psi$. Given that Kronecker product bases have shown good sparse representations of spectral images, a Kronecker product basis $\Psi = \Psi^{\text{2D}} \otimes \Psi^{\text{C}}$ is used, where $\Psi^{\text{2D}}$ is the $N^2 \times N^2$ matrix of the 2D Wavelet transform that is applied to each spectral band and, $\Psi^{\text{C}}$ is the $L \times L$ matrix of the Cosine transform that looks for the sparse representation across the spectral coordinate. The value of $\alpha$ is evaluated for the traditionally employed random coded apertures with 50% transmittance, and it is compared with coded aperture ensembles intentionally designed to exhibit horizontal separation of one-valued entries for $K = 4$. Figure 2.3 shows one realization of these coded apertures and their correspondent average values of $\alpha$. It can be noticed that random coded apertures result in larger values of $\alpha$.

b). **Vertical separation:** The presence of clusters of one-valued elements in the vertical direction of the coded aperture ensemble is related to an increase of the incoherence related to (2.12). This fact is illustrated in Fig. 2.4, where a random coded aperture is compared with a coded aperture intentionally designed to exhibit vertical separation of the one-valued entries. It can be noticed that the random coded aperture results in a lower incoherence value. Furthermore, given
that clusters of one-valued pixels prevent uniform sensing of the scene, one of the design criteria of an optimal coded aperture ensemble is to reduce these vertical cluster occurrences.

**Figure 2.4:** Incoherence comparison for different coded aperture ensembles with $K = 4$. (Left) Random coded aperture entries are realizations of a Bernoulli random variable with $p = 50\%$. (Right) Coded aperture intentionally designed to exhibit vertical separation of one-valued features.

c). **Temporal correlation:** Given that each measurement snapshot employs a different coded aperture, low correlation across shots is desired. An approach to reduce this correlation consists on constraining the set of codes to be complementary. In practical terms, this is equivalent to have just one coded aperture pixel set to one at each particular spatial position of the ensemble. This guarantees that a sensed voxel of the scene is captured just once. This property corresponds to temporal blue noise characteristics.

Based on the previous design criteria, the coded aperture optimization problem is mathematically described below. First, it is easy to see that small regions of the
ensemble can be designed such that the optimization criteria are satisfied, hence, the whole ensemble will fulfill them as well. Thus, let $U_p^i$ be a $\Delta \times \Delta$ window of $T^i$ centered at a specific point $P = (j, \ell)$, defined as $U_p^i = \{T^i_{\ell_1, \ell_2} | \ell_1 \in [x - [\Delta/2], x + [\Delta/2]], \ell_2 \in [y - [\Delta/2], y + [\Delta/2]]\}$, for $i = 0, \cdots, K - 1$. Given that $j = 0, \cdots, N - 1$ and $\ell = 0, \cdots, N - 1$, $P = 0, \cdots, N^2 - 1$. In order to reduce the concentration of one-valued terms within $U_p^i$, let

$$d_{P1}^i = \{T^i_{\ell_1, \ell_2} | \ell_1 \in [x - [\Delta/2], x + [\Delta/2]], \ell_2 = y\},$$  

where $d_{P1}^i$ be the subset of horizontal neighboring pixels within $U_p^i$ around $P$; similarly define

$$d_{P2}^i = \{T^i_{\ell_1, \ell_2} | \ell_1 = x, \ell_2 \in [y - [\Delta/2], y + [\Delta/2]]\},$$  

the subset of vertical neighboring pixels in $U_p^i$ around $P$; finally let the subsets of diagonal neighbors of $P$ in $U_p^i$ be written as

$$d_{P3}^i = \{T^i_{\ell_1, \ell_2} | \ell_1 \in [x - [\Delta/2], x + [\Delta/2]],$$

$$\ell_2 \in [y - [\Delta/2], y + [\Delta/2]], \ell_1 + \ell_2 = \Delta + 1\},$$

$$d_{P4}^i = \{T^i_{\ell_1, \ell_2} | \ell_1 \in [x - [\Delta/2], x + [\Delta/2]], \ell_1 = \ell_2\}.$$  

Figure 2.5 shows a graphic representation of the definitions in (2.13) to (2.16). These subset definitions are then used to define a weighted local metric $S_p^i$ given by

$$S_p^i = \sum_{k=1}^{4} w_k \|d_{P_k}^i\|_1,$$  

that measures the concentration of one-valued elements in $U_p^i$ for $i = 0; \cdots, K - 1$. More specifically, this metric penalizes the appearance of clusters of ones according to the weights $w_k$. Thus, if horizontal and vertical clusters are not desirable, $w_1$ and $w_2$ should be assigned greater values. Equivalently, $w_3$ and $w_4$ should be greater than $w_1, w_2$ if the distance between one-valued elements in the diagonals.

Hence, an optimal coded aperture ensemble can be obtained by minimizing the variance of $S_p^i$, given that all $K$ coded apertures in the ensemble are intended to exhibit
Figure 2.5: Graphic representation of the metric used to determine the concentration of one-valued entries in a window $U_P$ of the code. This metric helps on deciding the best $i$ for setting the one-valued element.

The same transmittance value. The optimal set of coded apertures for compressive spectral imaging can thus be expressed as the solution of the problem

$$\arg\min_{\{T^0, T^1, \ldots, T^{K-1}\}} \text{var}(S_P^i)$$ (2.18)

subject to $\sum_{i=0}^{K-1} T_{\ell_1, \ell_2}^i = 1,$

where $P = 0, \cdots, N^2 - 1$. To generate the optimal set of optimal coded apertures, an algorithm that finds a solution to (2.18) has been developed and is illustrated in Fig. 2.6. The required inputs are: the size of the coded aperture ensemble $N$, the number of measurement shots $K$, the window size $\Delta$, the set of weights $w = \{w_k\}_{k=1}^4$, and a set of randomly ordered spatial coordinates $\Omega = (\Omega_r, \Omega_c)$. Notice that $\Omega$ represents a permuted version of the spatial positions indexed by $P$, such that the procedure is performed in a random raster fashion. Also, an initial guess of the coded aperture is required.

The proposed algorithm moves over the pairs $(\Omega_r, \Omega_c)$ for $r, c = 0, \cdots, N-1$, and for each coded aperture of the ensemble, i.e. $i = 0, \cdots, K-1$, creates a $\Delta \times \Delta$ window, $U_P^i$, centered at $T_{\Omega_r, \Omega_c}^i$ as illustrated in Fig. 2.5. On each window, the algorithm
Select window of $\Delta \times \Delta$ centered at random point $\Omega_r, \Omega_c$

Evaluate amount of ones in $U_p'$:
- horizontally
- vertically
- diagonally

Calculate metric $S_p^i = \sum_{p=1}^{4} \|d_{p}^i\|$

Set $T_{\Omega_r, \Omega_c}^i = 0, \forall i \neq \hat{i}$
$T_{\Omega_r, \Omega_c}^\hat{i} = 1$

Find $\hat{i} = \arg\min_{i} (S_p^i)$

Figure 2.6: Block diagram of the algorithm that generates the designed coded apertures. This procedure iterates over all $(\Omega_r, \Omega_c)$ points and determines the best coded aperture of the ensemble in which the $(\Omega_r, \Omega_c)$-th pixel is set to one.

calculates the concentration of one-valued entries as $\|d_{p1}^i\|_1, \|d_{p2}^i\|_1, \|d_{p3}^i\|_1$ and $\|d_{p4}^i\|_1$ which are then used to calculate $S_p^i$ using (2.17). Hence, the criterion to decide the best codes aperture of the ensemble, i.e. the value of $i$, to insert the one-valued element for the specific spatial position $(\Omega_r, \Omega_c)$ is given by $\hat{i} = \arg\min_{i} (S_p^i)$, which in other words refers to the coded aperture with the minimum concentration of ones in the window around $(\Omega_r, \Omega_c)$. Finally, the algorithm sets the corresponding values of the coded aperture ensemble in the particular point $P = (\Omega_r, \Omega_c)$ as

$$T_{\Omega_r, \Omega_c}^i = 0 \forall i \neq \hat{i}$$
$$T_{\Omega_r, \Omega_c}^\hat{i} = 1.$$  \hspace{1cm} (2.19)

Figure 2.7 illustrates an example of the operations performed by the proposed algorithm. More specifically, a $8 \times 8$ coded aperture ensemble for $K = 4$ snapshots is used for this example. It can be noticed that the initial guess has clusters of one-valued elements and, at the end of the execution of the algorithm, the clusters of ones in the resulting coded aperture ensemble are considerably minimized.
The resulting coded aperture ensemble is coined spatio temporal blue noise (BN). A single realization of the BN coded apertures obtained with the proposed algorithm is illustrated in Fig. 2.8 (Top) along with the average incoherence and RIP parameters, $\tilde{\mu}$ and $\alpha$. Another approach to obtain BN coded apertures is the DBS algorithm, which has been originally used in lithography to optimally represent continuous tone images [80, 72]. In particular, blue noise patterns suppress low frequency components of noise [73, 74], and the characteristics of the patterns generated with the DBS algorithm correspond to the desired design criteria of the proposed BN coded apertures, as shown in Fig. 2.8 (Top). For comparison purposes, Fig. 2.8 (Top) presents one realization of the traditional random coded apertures and, boolean coded aperture which exhibit a spatially random distribution but exploits the temporal correlation restriction of BN patterns. Notice that the clusters of one-valued entries of the random coded aperture are highlighted in the orange rounded square. In addition, given that the spatial distribution of the boolean coded apertures is random, clusters of zero-valued elements are likely to appear, as shown in the yellow circle, as well as clusters of ones. The DBS and the BN coded apertures, on the other hand, have a very similar structure. The correspondent average values of the incoherence and RIP

![Figure 2.7: Example of the operations performed by the proposed algorithm. (a) Initial guess of a 8 × 8 coded aperture ensemble for $K = 4$ snapshots. (b) Resulting ensemble after applying the proposed algorithm.](image)

25
parameters, \( \tilde{\mu} \) and \( \alpha \), respectively, are also indicated for each case. As expected, BN coded apertures exhibit larger incoherence and lower RIP constant values. Furthermore, Fig. 2.8 (Bottom) illustrates the \( 4 \times 4 \) blue highlighted portion of the BN coded aperture ensemble for \( K = 4 \) snapshots, where the temporal correlation restriction is satisfied.

\[
\begin{align*}
\tilde{\mu} &= 0.3698, \alpha = 13 \times 10^4 \\
\tilde{\mu} &= 0.4303, \alpha = 8 \times 10^4 \\
\tilde{\mu} &= 0.4485, \alpha = 6 \times 10^4 \\
\tilde{\mu} &= 0.5014, \alpha = 6 \times 10^4
\end{align*}
\]

**Figure 2.8:** (Top) Comparison of one realization of the spatio temporal blue noise (BN) coded apertures with the traditional random, boolean and DBS coded apertures. (Bottom) Zoomed portion of an BN coded aperture ensemble for \( K = 4 \) measurement snapshots to illustrate low correlation across shots.

### 2.4 Simulation and Experimental Results

The performance of the proposed coded aperture design was tested through simulations and experiments. First, the CASSI acquisition process is simulated. In addition, a testbed implementation of the CASSI system is used for capturing experimental measurements. Reconstruction results and analysis are also presented for each case.
2.4.1 Simulations

The compressive spectral imaging acquisition was simulated using (2.5) for three different spectral data cubes with $256 \times 256$ pixels of spatial resolution and 8 spectral bands with central wavelengths 450nm, 456nm, 484nm, 503nm, 524nm, 549nm, 580nm, and 621nm. The first data cube was acquired by illuminating the scene with a monochromatic light source. The other data cubes are real scenes illustrating a set of colorful balloons and a color palette, available at [81] and [82], respectively. The proposed blue noise coded apertures are compared against the traditional random coded apertures with 50% transmittance, and the boolean coded apertures [22]. In addition, simulations are conducted using blue noise coded apertures obtained with the direct binary search (DBS) algorithm [72]. Recall that the boolean and the DBS coded aperture ensembles are complementary, resulting in low correlation across snapshots. The number of snapshots is varied in the simulations from $K = 2$ to $K = 8$ which matches the number of spectral bands, this is compression ratios from 1:4 to 1:1.

Figure 2.9 shows the average behavior of the reconstruction quality when different types of coded apertures are used, and the number of snapshots is varied. The quality of the reconstructions is measured using two widely known error metrics: the peak signal to noise ratio (PSNR) and the structural similarity index (SSIM). The PSNR measures the log-scale of the inverse of the mean-squared error, and is defined as $20 \log_{10} \left( \frac{\text{max}_I}{\text{MSE}^{1/2}} \right)$, where $\text{max}_I$ is the maximum possible value of the image and $\text{MSE}$ is the mean squared error with respect to the original image.
Figure 2.9: Average reconstruction quality measured by the PSNR and SSIM, as a function of the number of captured snapshots. Results for three different data bases are illustrated, using random, boolean, DBS and BN coded apertures.
On the other hand, the SSIM introduced in [83], compares local patterns of pixel intensities that have been normalized for luminance and contrast. It can be noticed in Fig. 2.9 that the random coded apertures exhibit the poorest reconstructions, while the coded apertures with low correlation across snapshots, i.e. boolean, DBS and BN, yield improved reconstruction quality. In addition, DBS and BN coded apertures provide the best PSNR and SSIM results. More specifically, the proposed BN coded apertures outperform the performance of the random coded apertures in up to 9 dB of PSNR, and the DBS coded apertures in up to 1 dB. Similarly, for the SSIM measure, the BN coded apertures present an improvement of up to 0.14 and 0.04 with respect to the random and DBS coded apertures, respectively. In general, these results indicate that the restrictions on the horizontal, vertical and diagonal separation of the one-valued entries are determinant to obtain improved reconstructions.

Figure 2.10 shows an RGB mapping of the recovered spectral data cubes for the different types of coded apertures, compared to the original scenes. In particular, these reconstructions are obtained from $K = 4$ measurement snapshots, which corresponds to the 50% of the data. In addition, Figs. 2.11, 2.12 and 2.13 illustrate the comparison of a zoomed portion of selected recovered spectral bands from Fig. 2.10, in which the differences between the recovered images are noticeable. Specifically, in spite of the good results obtained with the random, boolean and DBS coded apertures, the proposed BN ensembles provide more detailed and accurate reconstructions. Moreover, to illustrate the spectral accuracy of the BN coded apertures, the spectral signatures for two different points of each scene are illustrated in Fig. 2.14.
Figure 2.10: RGB profiles of the reconstructed images obtained using random, boolean, DBS and BN coded aperture ensembles, and $K = 4$ measurement snapshots, compared to the original data cubes.
Figure 2.11: Comparison of a zoomed portion of the 8 spectral bands from reconstructions in Fig. 2.10 for Database 1.
Figure 2.12: Comparison of a zoomed portion of the 8 spectral bands from reconstructions in Fig. 2.10 for Database 2.
Figure 2.13: Comparison of a zoomed portion of the 8 spectral bands from reconstructions in Fig. 2.10 for Database 3.
Figure 2.14: Spectral signature comparison for two spatial points on each database. Reconstructions obtained from $K = 4$ snapshots.
Additional simulations with a fourth spectral data set captured with AVIRIS sensor [84] were performed to test the BN coded apertures in this type of remotely sensed images. This data set is a $256 \times 256$ portion of the aerial view of Moffett Field with $L = 32$ spectral bands. Figure 2.15 illustrates the average reconstruction PSNR and SSIM for this data base as a function of the number of snapshots. It can be noticed that these results exhibit a similar behavior to those in Fig. 2.9 where DBS and BN perform similarly, both outperforming the results of random an boolean coded apertures. In addition, Fig. 2.16 presents a zoomed portion of the reconstructions from $K = 12$ snapshots for the different types of coded apertures, which corresponds to 37% of the data.

![Figure 2.15: Average reconstruction PSNR and SSIM for the Moffett Field data set, as a function of the number of snapshots.](image-url)
Figure 2.16: Comparison of a zoomed portion of 6 out of the 32 spectral bands from the reconstructions of the Moffet Field database for $K = 12$. Average PSNR: Random 30.88 dB; Boolean 31.63 dB; DBS 32.33; BN 32.66 dB.
Even though, in most applications the coded apertures are usually generated offline, the computation time required to generate a coded aperture ensemble is an additional comparison factor that can be taken into account. Performed simulations show that the DBS algorithm is up to 5 times slower than the proposed algorithm for generating coded apertures with blue noise characteristics. Another metric to compare different types of coded apertures is the radially average power spectrum density (RAPSD) [72], which is used to identify spectral characteristics of lithographic patterns. Given that the designed coded apertures have blue noise characteristics, RAPSD can be used to verify this behavior. In essence, RAPSD relies on estimating the magnitude square of the Fourier transform of the output ensemble, to produce the spectral estimate with the Bartlett’s method of averaging periodograms. Thus, the RAPSD is obtained by partitioning the spectral domain into a series of annular rings, and it is calculated as the average power in the annular ring for each center radius [74, 73]. Figure 2.17 illustrates the RAPSD for random, boolean, DBS and BN coded aperture ensembles, in which, random and boolean designs present a flat behavior. Meanwhile, high frequency characteristics of DBS and BN coded apertures are easily noticeable.

Figure 2.17: Comparison of the radially averaged power spectral density (RAPSD) for random, boolean, DBS and BN coded apertures.

37
2.4.2 Testbed Implementation

The optical setup of the CASSI system illustrated in Fig. 2.18 was constructed to experimentally verify the performance of the proposed BN coded apertures. This setup comprises an objective lens, a custom Amici prism, a relay lens, a monochrome charge-coupled-device, a photomask wafer, and an $x-y$ linear stage. In particular, the wafer is a $17 \times 17 \times 0.015$” chrome-on-quartz photomask that consists of several $128 \times 128$ coded aperture realizations with pixel pitch $19.35 \mu m$. The linear stage moves the photomask accordingly, in order to locate and align the correspondent pattern to be used on each snapshot. The non-linear Amici prism disperses the visible incoming radiation into 10 spectral bands with central wavelengths: 448nm, 460nm, 473nm, 489nm, 507nm, 526nm, 550nm, 579nm, 611nm, and 653nm. Compressive measurements for $K = 2, 4, 6$ measurement snapshots were captured using random, boolean, DBS and BN coded apertures. These sets of measurements and the corresponding calibration data cubes are used to estimate the original scene.

![Testbed Implementation of the CASSI system](image)

**Figure 2.18:** Testbed implementation of the CASSI system.

Figure 2.19 presents a comparison of an RGB mapping of the recovered scenes using the random, boolean, DBS and BN coded apertures, obtained from $K = 2, 4$ and 6 measurement snapshots. It can be noticed that increasing the number of snapshots results in improved and more detailed reconstructions. Furthermore, BN and
DBS coded apertures yield better reconstructions than the random and boolean coded apertures.

Figure 2.19: RGB representations of the experimental reconstructions from $K = 2, 4, 6$ for different types of coded apertures.

To better visualize the improvements in the reconstructions, Fig. 2.20 presents four selected recovered spectral bands, from $K = 4$ measurement snapshots, for random, boolean, DBS and BN coded apertures. Finally, Fig. 2.21 compares the spectral signatures of three different points of the scene with the original spectrum measured with a point spectrometer. This comparison shows that in spite of the good approximation of the spectra obtained with random and boolean coded apertures, the DBS and BN designs provide a better approximation of the bands of interest.
2.5 Conclusions

The spatio temporal blue noise (BN) coded apertures for compressive spectral imaging have been introduced. This design is based on the restricted isometry property of the CASSI sensing matrix, as well as the incoherence with a given sparsifying basis. The mathematical development of the proposed design and a computational algorithm...
Figure 2.21: Spectral signature reconstructions of three points of the scene, using $K = 4$ snapshot measurements. (Top Right) Orange arm; (Bottom Left) Yellow face; (Bottom Right) Green chest.

to generate BN coded apertures have been also presented. Simulations and experiments show that the proposed design results in improved reconstructions with respect to the traditional random coded apertures, while similar quality can be attained with blue noise coded apertures generated with the DBS algorithm. In particular, a gain of up to 9 dB of PSNR and 0.15 in SSIM, are obtained with respect to the random ensembles. However, the proposed algorithm generates BN coded apertures up to 5 times faster than the DBS algorithm. It is worth noting that although BN coded apertures were tested for the CASSI system, the concept can be extended to other CSI architectures such as those in [85] and applications like remote sensing.
Chapter 3

SNAPSHOT COLORED COMPRESSIVE SPECTRAL IMAGER (SCCSI)

This chapter presents a new generation of compressive spectral imaging architectures that couples the resolution capabilities of the CASSI system and the compactness of Fabry-Perot-based architectures. In particular, the design of the snapshot colored compressive spectral imager (SCCSI), illustrated in Fig. 3.1 (b), relies on a colored-patterned FPA detector in conjunction with a dispersive element to capture the spatial and spectral information of a scene in a single compressive snapshot. The colored-patterned FPA employs recent thin film coating advances and lithographic techniques that permit the implementation of multi-patterned arrays of different optical filters [86, 87, 88, 89, 90]. Thus, SCCSI exploits the advantages of recent coating technology to obtain a richer encoding of the spatio-spectral information of the source, allowing a considerably compact design without moving components. In addition, the combination of color filter arrays and a dispersive element increases the attainable spectral and spatial resolution [91].

In the SCCSI architecture, the input source, \( \mathbf{F} \), is first dispersed by a dispersive element and then, it is encoded by an array of optical filters. Finally, the encoded and dispersed source is integrated along the spectral range sensitivity of the detector. In matrix notation, the output of the system is modeled as \( \mathbf{g} = \mathbf{Hf} \), where \( \mathbf{f} \in \mathbb{R}^{N^2L} \) is a vector representation of the source \( \mathbf{F} \) and, \( \mathbf{H} \) is a highly sparse matrix whose non-zero entries are determined by the spectral responses of the color filter array and, its structure represents the effects of the dispersive element. An estimate of the 3D spectral source can be recovered using a compressed sensing reconstruction algorithm that finds
a sparse representation of the source in a given basis $\Psi$, using the measurement set $g$ and the responses of the optical filters on the detector.

The chapter is organized as follows: first, the SCCSI architecture design and the mathematical model of the underlying physical phenomenon are presented. In particular, the mathematical discretization is developed and, a matrix representation of the process is also presented. Then, simulations and result analysis are shown to verify the performance of the proposed system. Simulation results are compared to those of the CASSI system. Finally, a testbed implementation of the SCCSI architecture is described and, reconstructions from experimental data captured in the laboratory are presented.

3.1 Compressive Spectral Imaging with Colored-Patterned Detectors

The SCCSI optical architecture, illustrated in Fig. 3.1(b), comprises an objective lens to capture the incident source, a dispersive element to separate the spectral components of the source and, a FPA detector with a pixelated tiling of optical filters in front of it, to integrate the dispersed and encoded source. For comparison purposes, the CASSI architecture is also illustrated in Fig. 3.1(a) [29].

3.1.1 Continuous SCCSI Model

Denote the spatio-spectral source density in Fig. 3.1(b) as $f_0(x, y, \lambda)$, where $(x, y)$ index the spatial coordinates and $\lambda$ indexes the spectral dimension. As it was previously mentioned, the source $f_0(x, y, \lambda)$ is first dispersed by the dispersive element yielding $f_1(x, y, \lambda)$. The resulting dispersed field can be expressed as

$$f_1(x, y, \lambda) = \int \int f_0(x', y', \lambda) h(x' - x - S(\lambda), y' - y) \, dx' \, dy', \quad (3.1)$$

where $h(\cdot)$ is the optical impulse response of the system and, $S(\lambda)$ is the dispersion function of the dispersive element that operates along the $x$ coordinate axis. The dispersed source is then coded by the colored-patterned detector which can be modeled.
Figure 3.1: Schematic of (a) the CASSI imager and (b) the snapshot colored compressive spectral imager (SCCSI).

as an array of optical filters \( C(x, y, \lambda) \), followed by an irradiance detector. More specifically, each spatial location of \( C(x, y, \lambda) \) is associated to a specific spectral response that modulates the incident light in that particular position. Furthermore, there exists a 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. Hence, the coded and dispersed irradiance source impinging on the detector can be written as

\[
f_2(x, y, \lambda) = f_1(x, y, \lambda) C(x, y, \lambda) .
\] (3.2)

The compressive measurement denoted as \( g(x, y) \), is finally obtained by the integration of the field \( f_2(x, y, \lambda) \) over the spectral range sensitivity of the detector (\( \Lambda \)). More specifically, using Eq. 3.1 and Eq. 3.2 the SCCSI compressive measurements can be written as
\[ g(x, y) = \int_{\Lambda} f_2(x, y, \lambda) d\lambda \]
\[ = \int_{\Lambda} \int f_0(x', y', \lambda) h(x' - x - S(\lambda), y' - y) \times C(x, y, \lambda) \, dx' \, dy' \, d\lambda. \] (3.3)

### 3.1.2 Discrete SCCSI Model

Let \( \Delta \) be the detector pixel pitch, such that the area occupied by the \((m, n)\)-th pixel element can be represented as \( p(n, m; x, y) = \text{rect} \left( \frac{x}{\Delta} - n, \frac{y}{\Delta} - m \right) \). Then, the energy captured in the \((m, n)\)-th pixel of the detector can be written as

\[ g_{m,n} = \int \int g(x, y) \, p(n, m; x, y) \, dx \, dy \] (3.4)
\[ = \int_{\Lambda} \int \int f_0(x', y', \lambda) h(x' - x - S(\lambda), y' - y) \times C(x, y, \lambda) \, \text{rect} \left( \frac{x}{\Delta} - n, \frac{y}{\Delta} - m \right) \, dx' \, dy' \, dx \, dy \, d\lambda. \]

On the other hand, the spatio-spectral source can be represented as a discretized data cube by calculating the energy on each voxel with the expression

\[ F_{i,j,k} = \int_{\lambda_k}^{\lambda_{k+1}} \int_{i\Delta}^{(i+1)\Delta} \int_{j\Delta}^{(j+1)\Delta} f_0(x, y, \lambda) \, dx \, dy \, d\lambda, \] (3.5)

where, \( i, j = 0, \cdots, N - 1 \) are the discrete indices for the spatial dimensions and, \( k = 0, \cdots, L - 1 \) indexes the spectral bands. This expression results in a \( N \times N \times L \) data cube, where \( N \) corresponds to the spatial resolution and \( L \) is the number of spectral bands. The width of the spectral bands is determined by the dispersion function of the prism \( S(\lambda) \) and the detector pixel pitch \( \Delta \). Notice that additional columns of pixels are required in the detector due to the dispersion effects along the \( x \)-axis. In general, an \( N \times V \) FPA detector is needed, where \( V = (N + L + 1) \). Similarly, since there is a
1-1 matching between the optical filters and the pixels in the detector, the color filter array can be discretized in terms of the FPA pixel pitch, $\Delta$, as

$$C(x, y, \lambda) = \sum_{m,n,k} C_{m,n,k} \text{rect}\left(\frac{x}{\Delta} - n, \frac{y}{\Delta} - m, \frac{\lambda}{\Delta} - k\right). \quad (3.6)$$

The discretization in Eq. 3.6 results in a 3D array $C_{m,n,k}$ with two spatial and one spectral dimensions. Given the 1-1 matching between the elements of the color filter array and the FPA detector, the spatial dimensions of the former correspond to those of the latter. Furthermore, the spectral responses of the optical filters in $C_{m,n,k}$ at the detector are discretized to $L$ values, as shown in Fig. 3.2. Assuming normalized spectral responses of the filters, the entries of the color filter array, $C_{m,n,k}$, are in the set $\{0, \cdots, 1\}$. In contrast, when ideal spectral responses are assumed, the entries of $C_{m,n,k}$ are either zero or one. The transmittance of the color filter array which is the amount of energy that passes through, can be written as

$$Tr = \frac{1}{NVL} \sum_{m,n,k} C_{m,n,k}. \quad (3.7)$$

Each spatial pixel in the detector can be associated to a different optical filter. In other words, a $N \times V$ color filter array can be seen as a tiling of $NV$ optical filters. In practice, such amount of different filters is not attainable since the maximum number of possible responses is determined by the number of spectral bands $L$ and, is given by $\sum_{i=1}^{L} \binom{L}{i}$. Thus, $C_{m,n,k}$ is limited to a set of available filters, with typically a maximum of $L$ different filters that are used to recover the same amount of spectral bands.

Using Eq. 3.6 and, assuming an ideal PSF, the output at the detector in Eq. 3.4 can be rewritten as

$$g_{m,n} = \sum_{m,n,k} C_{m,n,k} \int_{\Lambda} \int \int f_0(x + S(\lambda), y, \lambda) \times \text{rect}\left(\frac{x}{\Delta} - n, \frac{y}{\Delta} - m, \frac{\lambda}{\Delta} - k\right) dxdyd\lambda. \quad (3.8)$$
Figure 3.2: Example of a discrete color filter array. Each spatial location \((m,n)\) contains the spectral response of the correspondent optical filter. Black features in \(C_{m,n,k}\) block the correspondent wavelengths and, white features correspond to the pass-bands of the filter.

Now, using Eq. 3.5 and taking into account the dispersive element transfer function, Eq. 3.8 can be in turn expressed as

\[
g_{m,n} = \sum_{m,n,k} F_{m,(n-k),k} C_{m,n,k} \times \iiint \text{rect} \left( \frac{x}{\Delta} - n, \frac{y}{\Delta} - m, \frac{\lambda}{\Delta} - k \right) \, dx \, dy \, d\lambda.
\]

Notice that the shifting in the indices of \(F\) represent the effect of the dispersive element. Moreover, the dispersion function of the prism \(S(\lambda)\) has a critical effect in the energy incidence from a source voxel to each of the detector pixels. More specifically, after dispersion, the incoming squared cube voxel is sheared such that its energy will impinge not just into a single FPA pixel but on the set of up to three neighboring pixels. This phenomenon is depicted in Fig. 3.3 where a single voxel splits in up to three regions \(R_0, R_1,\) and \(R_2\). The corresponding proportion of energy from each region is indexed by the weights \(w_{m,n,k,u}\), where \((m,n)\) are the indices of the spatial coordinates, \(k\) is the index of the spectral dimension and, \(u = 0,1,2\) corresponds to
the regions $R_u$ [92]. The discretized output in the FPA detector can thus be succinctly expressed as

$$g_{m,n} = \sum_{k=0}^{L-1} \sum_{u=0}^{2} w_{m,n,k,u} F_{m,(n-k-u),k} C_{m,n,k}. \quad (3.10)$$

**Figure 3.3:** A source voxel after being dispersed is illustrated. The energy of the voxel splits into three regions that impinge onto three consecutive FPA pixels. The proportion of energy from each voxel in the region $u$ is represented by the weight $w_{m,n,k,u}$.

A representation of the physical phenomenon behind the SCCSI sensing is presented in Fig. 3.4. Notice that the separation between the color filter array and the focal plane array is included only for illustration purposes. In practice, the color filter array consists of a coating attached to the FPA.

3.1.3 Matrix Model

The set of compressive measurements from Eq. 3.10 can be expressed in vector notation as $\mathbf{g}$, such that the compressive spectral imager with colored-patterned
Figure 3.4: Schematic representation of the discrete process behind SCCSI. Each spectral band of the input is dispersed according to the function \( S(\lambda) \). Then, the color filter array encodes the dispersed light before it is captured by an intensity detector. The set of filters in the array is also illustrated.

detectors can be modeled by

\[
g = Hf, \quad (3.11)
\]
in which \( f \) is the vector representation of the spatio-spectral input source \( F \) and, \( H \) is a \( NV \times N^2L \) sparse matrix whose non-zero entries are given by the spectral responses of the optical filters in the color filter array and, its structure accounts for the effects of the dispersive element. Furthermore, the sensing matrix \( H \) can be written as the product of three matrices, each one representing an individual physical phenomenon

\[
H = S\Gamma\Delta, \quad (3.12)
\]
where \( D \) is the dispersion matrix, \( \Gamma \) is the coding matrix and, \( S \) is the integration matrix. More specifically, \( D \), illustrated in Fig. 3.5, is a \( NVL \times N^2L \) matrix that accounts for the dispersion phenomenon. The three diagonals represent the regions

49
of energy $R_u$ in which a single voxel splits due to the dispersion function and, different color intensities indicate the correspondent weights for these regions $w_{m,n,k,u}$. In particular, the weights for all $(m,n)$ coordinates, corresponding to the $k$-th spectral band and the region $u$ can be arranged in the vector $w_k^u$. Thus, $(w_k^u)_{mn} = w_{m,n,k,u}$ as illustrated in Fig. 3.5. On the other hand, $\Gamma$ in Eq. 3.12 is a $NVL \times NVL$ diagonal matrix, whose entries are the spectral responses of the optical filters in the color array $C_{m,n,k}$. Mathematically, $\Gamma$ can be expressed as $\Gamma = \text{diag}(c)$, where $c = [c_0^T, c_1^T, \ldots, c_{L-1}^T]^T$ is a vector representation of the color filter array $C_{m,n,k}$ in Eq. 3.6, with $c_k = [C_{0,0,k}, C_{1,0,k}, \ldots, C_{0,1,k}, \ldots, C_{N-1,V-1,k}]^T$ representing the elements of the color filter array corresponding to the $k$-th spectral band. Finally, $S$, which performs the integration of the encoded and dispersed source, is a $NV \times NVL$ matrix given by $S = S_{NV} \otimes 1_L$, where $S_{NV}$ is an $NV \times NV$ identity matrix, $1_L$ is an $L$-long one-valued row vector and, $\otimes$ represents the Kronecker product.

Alternatively, the entries of the $j$-th row in the sensing matrix $H$ can be succinctly written as

$$
(h_j)_\ell = \begin{cases} 
\sum_{u=0}^{2} (w_k^u)_{j-(u+k)N}(c_k)_{j-uN}, & \text{if } \ell=kN'+uN=j \\
0, & \text{otherwise}
\end{cases} \quad (3.13)
$$

where $j = 0, \ldots, NV - 1$ and $\ell = 0, \ldots, N^2L - 1$ index the rows and columns of $H$ respectively, $k = [\ell/N^2]$ is an index for the spectral bands and, $N' = N^2 - N$ accounts for the column-wise shifting induced by the dispersion. In addition, the sum along $u$ in Eq. 3.13 accounts for the regions $R_0, R_1$ and, $R_2$. An example of the resulting sensing matrix $H$ for $N = 6$ and $L = 3$ is shown in Fig. 3.6. Notice that this matrix reflects the dispersion effect as well as the three diagonals representing the regions in which each voxel splits due to the dispersive element function. Also, the patterns in the non-zero entries exhibit the color filter array effect.
Figure 3.5: Dispersion matrix $D$, for $N = 6$ and $L = 3$. The three diagonals represent the regions in which each voxel splits due to the dispersive function $S(\lambda)$. Different color intensities indicate the weights $w_{m,n,k,u}$. The weights are arranged by spectral band and region as $w_k^u$ for $k = 0, \ldots, L - 1$ and $u = 0, 1, 2$.

3.1.4 Reconstruction Process

Since the number of pixels captured in the FPA is much smaller than the total number of voxels in the input data cube, $NV \ll N^2L$, the reconstruction problem cannot be resolved by inverting the system in Eq. 3.11. The theory of compressed sensing is then used to solve the inverse problem that consists in recovering the underlying signal $f$ from the measurement set $g$, assuming that $f \in \mathbb{R}^{N^2L}$ has a $S$-sparse representation in a given basis $\Psi$. In other words, with this assumption $f$ can be approximated as a linear combination of only $S \ll N^2L$ columns of $\Psi$. The measurement set $g$ in Eq. 3.11 can be expressed as $g = H\Psi\theta$ where, $f = \Psi\theta$. Compressed sensing reconstruction algorithms
look for an sparse approximation of the spatio-spectral data cube. Mathematically, the optimization problem can be written as $f = \Psi \{ \arg\min_\theta \| g - H\Psi\theta \|_2 + \tau \|\theta\|_1 \}$, where $\tau$ is a regularization constant.

### 3.2 Simulations and Experiments

Simulations with three test data bases and, experiments with measurements obtained from a testbed, were realized to analyze the performance of the proposed architecture.

#### 3.2.1 Simulations

Three test data cubes $F$ with $256 \times 256$ pixels of spatial resolution and $L = 8$ spectral bands in the range 450nm to 667nm were used for simulations. The first test data base was obtained by illuminating the scene with a monochromatic light source. Figure 3.7(a) shows the spectral bands of this test data cube at their central wavelengths. The second data base is a portion of an aerial view of Pavia University and, the third data cube is a color palette [82]. RGB representations of these data bases are

**Figure 3.6**: Sensing matrix $H$, for $N = 6$ and $L = 3$. The structure of the matrix accounts for the effects of the dispersion, coding and integration phenomenon occurred behind SCCSI. The three diagonals with different color intensities represent the regions in which each voxel splits due to the dispersive function $S(\lambda)$ and different weight values.
shown in Fig. 3.7 (b) and (c), respectively.

Figure 3.7: Test data cubes with 8 spectral bands ranging from 450nm ($\lambda_1$) to 667nm ($\lambda_8$) used for simulations. Each spectral slice has a spatial resolution of $256 \times 256$ pixels. (a) Spectral bands with central wavelengths of the first data cube. (b) RGB representation of the Pavia University aerial view. (c) RGB representation of the color palette data cube.

The color filter array pattern is generated by randomly tiling a set of predefined optical filters using a uniform distribution. CSI devices usually employ coding patterns with transmittance $Tr = 50\%$. In SCCSI, light efficiency depends directly on the spectral responses of the optical filters. To guarantee 50% transmittance in the color filter array when a uniform distribution of the filters is used, the optical filters were
selected such that their band-pass cover half of the spectral bandwidth in $\mathbf{F}$. Figure 3.8 illustrates the set of 8 predefined filters employed in the simulations. The number of filters (colors) is varied in the simulations to determine its effect in the reconstructed image quality. Simulations were performed varying the number of colors from 2 to 8, which matches the number of spectral bands.

\[ \lambda_1, \cdots, \lambda_8 \] indicate the cut-off wavelengths which match the spectral bands of $\mathbf{F}$.

Compressive measurements are constructed by using the test data cubes and Eq. 3.10. The simulated set of compressive measurements and the color filter array are used as an input to a compressed sensing reconstruction algorithm to obtain the reconstruction of the spectral source. In particular, the results in this paper were attained with the GPSR algorithm [93], but the use of other reconstruction methods such as total variation does not reflect a significant variation in the behavior of the results. The sparse basis is set to be a Kronecker product between a 2D Wavelet Symmlet 8 basis for the spectral dimensions and a 1D DCT basis for the spectral domain, $\Psi = \Psi^{2DW} \otimes \Psi^{DCT}$. Figure 3.9 shows the average reconstruction PSNR for the three test data cubes, when different number of optical filters (colors) from Fig. 3.8 are used. These results are compared with the reconstructions obtained.
from a single snapshot of the CASSI system, with a binary coded aperture whose entries follow a Bernoulli distribution with $p = 0.5$ [29, 33] and, the SSCSI architecture with a random binary mask [37]. In general, the SCCSII architecture greatly improves the results obtained with CASSI measurements. More specifically, a gain up to 6 dB in PSNR is attained in simulations. Moreover, simulations show that SCCSII and SSCSII architectures provide similar reconstructions. In some cases, SCCSII outperforms SSCSII, and in some other cases, SSCSII results in better reconstructions. Also, it can be seen that, in general, increasing the number of colors (filters) in SCCSII leads to an improvement in the reconstruction PSNR. However, this behavior reaches a steady state when 5 optical filters are used for an 8 spectral band data cube. In other words, these results show that increasing the amount of optical filters in the detector to more than 5 colors does not significantly improve the image quality in the reconstruction of eight spectral image bands.

The selection of an appropriate set of optical filters is critical to obtain high-quality reconstructions. In particular, two conditions are required for choosing this set: the first is that all the spectral range of the input source needs to be covered and, the second is that the bandwidth of the filters in the set must preserve the light efficiency in the system. Thus, the color filter array distribution and the bandwidth of the optical filters determine the transmittance of the color filter array and, in turn, are critical parameters in order to obtain high-quality reconstructed images.

To visualize the reconstructed input source, Fig. 3.10 illustrates a zoomed area of a subset of reconstructed bands using CASSI, SSCSII and SCCSII with 4 and 5 colors. In addition, Fig. 3.11 shows the RGB representations of the reconstructed data cubes in Fig. 3.7(b) and (c), using CASSI and SSCSII measurements compared to SCCSII results obtained with 4 colored-measurements. It is possible to observe in these figures that SCCSII and SSCSII architectures provide similar reconstructions, in both cases attaining more details than the images recovered from CASSI measurements. However, in Fig. 3.11(c) some artifacts can be noticed on the SSCSII image, while SCCSII provides an artifact-free reconstruction.
### Figure 3.9

Average reconstruction PSNR as a function of the number of colors in the detector (Filters in Fig. 3.8) of (a) data cube in Fig. 3.7(a), (b) Pavia university in Fig. 3.7(b) and, (c) color pallette in Fig. 3.7(c). The average PSNR obtained using a single shot with CASSI and SSCSI are also illustrated.
Figure 3.10: Zoomed reconstructions of a subset of spectral bands using (Second row) CASSI with average PSNR across bands 21.2 dB, (Third row) SSCSI with average PSNR 25.83 dB, (Fourth row) SCCSI with 4 colors with average PSNR 25.36 dB, and (Fifth row) SCCSI with 5 colors with average PSNR 26.45 dB. The PSNR for each spectral band is also indicated.
Figure 3.11: RGB reconstruction comparison using CASSI, SSCSI and SCCSI with 4 colors. (a) Reconstructions for the data cube in Fig. 3.7 (b) and, (b) Fig. 3.7 (c)
To verify the accuracy of the proposed SCCSI, Fig. 3.12 shows the spectral signatures for three spatial points in the reconstructions obtained from CASSI, SSCSI and SCCSI with 4 colors. These results show that the SCCSI spectral reconstruction is more accurate than that of the CASSI architecture. Also, despite of the good spectral approximations obtained with SSCSI, it can be noticed that SCCSI spectral reconstructions are closer to the original spectrum.

![Figure 3.12: Reconstruction along the spectral axis for highlighted spatial locations using CASSI, SSCSI and SCCSI with 4 colors.](image)

The eigenvalue spread of the sensing matrix and, the condition number $\kappa = \lambda_{max}/\lambda_{min}$ is also used to determine the well-condition of the sensing matrices. Figure 3.13 illustrates the eigenvalue distribution for the sensing matrix in CASSI, SSCSI and, SCCSI with 4 colors. It can be noticed that the SCCSI architecture leads to a better conditioned sensing matrix when compared with CASSI. Also, the eigenvalue spread of SCCSI and SSCSI develops a similar decay behavior which, in turn explains the
similarity in the results obtained with both architectures. However, SCCSI exhibits a slightly lower condition number, indicating a better conditioned sensing matrix.

![Eigenvalue distribution of the sensing matrix for the SCCSI with 4 colors, CASSI and SSCSI. The condition number \( \kappa = \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} \) for all cases are also indicated.](image)

**Figure 3.13:** Eigenvalue distribution of the sensing matrix for the SCCSI with 4 colors, CASSI and SSCSI. The condition number \( \kappa = \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} \) for all cases are also indicated.

Simulation results show that equivalent reconstructions can be attained using SCCSI and SSCSI optical architectures. Even though, the cost of using a binary mask in SSCSI is less than that of a colored-patterned detector, SCCSI exhibits a considerably more compact design with a less number of optical elements, thus, alleviating the calibration and alignment procedures. SSCSI, on the other hand, experience challenging alignment due to the offset location of the mask with respect to the detector and, the different pixel pitches. Moreover, colored detectors have recently become more available and are gaining popularity due to the one-step spatio-spectral coding they offer, thus allowing the development of simpler optical designs in several applications in medicine, defense and industry [24, 94, 95]. Mathematically, the use of color filter arrays yields to lower correlation between the entries of the sensing matrix, which allows to better exploit CS theory.
3.2.2 Experiments

An optical setup of the SCCSI architecture was built in the laboratory in order to verify its capabilities. This setup comprises a pair of visible achromatic lenses acting as the objective lens, with a focal length of 100 mm (Thorlabs AC254-100-A-ML), a custom double Amici prism designed to not disperse $\lambda = 550$ nm, a monochrome Stingray F033B CCD camera with pixel size $9.9 \mu$m and, 8 bits of intensity levels and, a set of optical filters. Figure 3.14 (Left) illustrates the described setup. The spectral resolution of the system depends on the pixel pitch of the camera and the dispersion function of the prism. The double Amici prism dispersion function is illustrated in Fig. 3.15. This characterization can resolve up to 24 spectral bands. Due to computational issues we have reduced them to 6 different spectral bands as follows: $\{448 – 465\}$, $\{466 – 487\}$, $\{488 – 516\}$, $\{517 – 550\}$, $\{551 – 600\}$ and, $\{601 – 663\}$ nm. Figure 3.16 illustrates the set of filters used in the experiments. Notice that the filters cover the full spectrum and the light efficiency is fixed at 50%. Different bandwidths are due to the non-linear dispersion function of the prism $S(\lambda)$. It can be noticed in Fig. 3.15 that the energy of each voxel splits into two regions that affect two consecutive detector pixels. The target shown in Fig. 3.14 (Right) was separately captured using each of the optical filters in the set and then, the SCCSI measurements were generated according to the desired color filter array pattern.

Figure 3.14: (Left) Laboratory implementation of the SCCSI architecture. (Right) RGB representation of the target scene used for the experiments.
Figure 3.15: Characterization of the double amici prism in the SCCSI architecture. The effect of the dispersion in the spectral range \{448 – 663\} nm is illustrated. The energy of each voxel splits into two regions $R_0, R_1$ and, the corresponding weights are indicated.

Figure 3.16: Set of filters used for experiments. A 50% transmittance color filter array is obtained when a uniform distribution of the colors is used. $\lambda_0 = 448$ nm, $\lambda_1 = 466$ nm, $\lambda_2 = 487$ nm, $\lambda_3 = 516$ nm, $\lambda_4 = 550$ nm, $\lambda_5 = 600$ nm, and $\lambda_6 = 663$ nm.

To determine the impulse response of the system, the calibration process consists on illuminating each optical filter with a broadband xenon lamp, and the filtered light is measured using an Ocean Optics USB2000+ spectrometer. These non-ideal filter
responses are used to construct the color filter array $C_{m,n,k}$, which is later used to build the sensing matrix $H$. In this way, the spectral responses of the filters are integrated to the inverse problem. The reconstructions of the input source were obtained using the GPSR algorithm. Figure 3.17 shows the obtained reconstruction from SCCSI measurements using 4 colors. In this case, the first 4 filters from Fig. 3.16 were used. In addition, Fig. 3.18 presents a comparison of a zoomed area from the reconstructions obtained with SCCSI measurements using 2 and 4 colors. The differences between these reconstructions are not easily noticeable. For this reason, the spectral accuracy of the architecture was experimentally tested by comparing the spectral signatures of the reconstructions with those measured with a commercially available point spectrometer (Ocean Optics, USB2000+) used as a reference. The results for three spatial locations are shown in Fig. 3.19, in which the reconstruction using the set of filters in Fig. 3.16 was also included. It can be noticed there that using 4 and 6 colors provide similar accurate spectral reconstructions, while using just 2 colors provides a less accurate spectral signature.

![Recovered spectral bands from SCCSI measurements using filters 1 to 4 in Fig. 3.16.](image)

**Figure 3.17:** Recovered spectral bands from SCCSI measurements using filters 1 to 4 in Fig. 3.16.
Figure 3.18: Comparison of different spectral bands using SCCSI with 2 and 4 colors in the detector. Zoomed regions of the reconstructions are illustrated.

Figure 3.19: Comparison of the spectral signature for the highlighted spatial points, measured with a commercial spectrometer and the SCCSI reconstructions using 2, 4 and 6 colors.
3.3 Conclusions

A new generation of compressive spectral imaging architectures based on colored-patterned detectors has been introduced. The mathematical model for the snapshot colored compressive spectral imager (SCCSI) was developed in this paper. Simulations show that SCCSI overcomes both the spatial and spectral CASSI reconstructions even when a low number of colors is used. Moreover, an improvement of up to 6 dB in PSNR has been achieved in simulations. In addition, a gain of 1 dB can be attained with the proposed SCCSI system with respect to SSCSI and, more accurate spectral approximations can be obtained with the proposed architecture. Experimental results demonstrate the spatial and spectral accuracy of the system. It has also been shown that increasing the number of colors in the detector leads to improved image quality in the reconstructions.
Chapter 4
THE MULTI-RESOLUTION RECONSTRUCTION PROBLEM

The inverse problem in compressive spectral imaging entails the reconstruction of the spatial and spectral information of the underlying scene from the captured set of projections. Specifically, for a data cube $F \in \mathbb{R}^{N \times N \times L}$, the amount of projections captured with a CSI system such as CASSI is $KN(N + L - 1)$, where $K$ indicates the number of snapshots. As mentioned in the previous chapters, given that the number of projections is less than the number of image voxels to recover, i.e. the number of equations is less than the number of unknowns, CSI reconstructions involve solving an underdetermined system of linear equations. Therefore, in order to estimate the underlying signal $f \in \mathbb{R}^{N^2L}$ which produced the measurements $y = Hf$, a sparsity condition on $f$ is assumed such that it can be recovered by solving the problem given by

$$\arg \min_\theta ||y - H\Psi\theta||_2^2 + \tau ||\theta||_1.$$  \hspace{1cm} (4.1)

In general, CSI reconstructions have to date focused on the recovery of the full-resolution $N \times N \times L$ data cube whose resolution is limited by that of the acquired measurements. However, the computational complexity of the inverse problem grows exponentially with the spatial dimensions of the image to recover \[48\]. An alternative to alleviate the complexity of these reconstructions is to obtain lower resolution approximations of the signal. Recently, several multi-resolution (MR) compressed sensing reconstruction models have been proposed. For instance, a MR model and a CS recovery algorithm based on approximate message passing (AMP) have been proposed in \[50\]. However, such model only considers the reconstruction of a low resolution version
of the desired signal employing traditional Gaussian compressed sensing matrices and establishes that the target HR image can be obtained by upsampling or interpolation. Moreover, additional MR approaches propose to design the sensing matrices such that an accurate low-resolution version of the object can be recovered, yet these approaches are not applicable to CSI because the sensing matrix structure is imposed by the optical configuration of the system.

This chapter presents a multi-resolution reconstruction scheme for compressive spectral imaging, in which a set of high resolution measurements is used to sequentially recover multiple versions of the same data cube at different spatial resolutions as illustrated in Fig. 4.1.

![Figure 4.1](image)

**Figure 4.1:** Schematic comparison of the traditional high-resolution (HR) CSI reconstruction with the proposed multi-resolution (MR) compressive spectral imaging reconstruction approach.

The proposed model exploits the gradient intensity images of a low resolution reconstruction to generate a decimation matrix $D_\Delta$, in order to solve the inverse problem from Eq. (4.1) for the desired resolution using the equivalent MR sensing matrix
HD$^T_{\Delta}$, as illustrated in Fig. 4.2. Specifically, the first step of the proposed MR approach consists on recovering a low-resolution (LR) approximation of the data cube $f_\Delta$ whose spatial dimensions are $\frac{1}{2^\Delta}$ the dimensions of the full-resolution data cube, while the number of spectral bands remains unchanged; $\Delta$ is an input parameter for the algorithm. Then, the map of intensity gradients of $f_\Delta$ is used to generate a MR decimation matrix such that an upscaled version of the data cube can be recovered. The upscaled version $f_{\Delta-1}$ is a data cube whose spatial resolution will be $\frac{1}{2^{\Delta-1}}$ of the original on each spatial dimension. The new decimation matrix is calculated such that the pixels in smooth regions are decimated by a factor $2^\Delta$ and the pixels indexed by the map of intensity gradient are decimated by a factor $2^{\Delta-1}$. The process can be then repeated as indicated by the iteration variable $t$, until the dimensions of the high resolution object are reached. It it worth noting that in this work, for simplicity, the decimation factor is expressed as a power of 2. The following sections describe each step of the proposed method in detail.

Figure 4.2: Diagram of the proposed MR reconstruction approach for compressive spectral imaging. First, a LR reconstruction is obtained using a decimation factor $\Delta$. The upsampling process consists on finding the edges of the available LR reconstruction to determine a new MR decimation matrix to solve the equivalent inverse problem to recover a data cube with a decimation factor of $\Delta - 1$. The upsampling process is repeated until the reconstruction reaches the original resolution.
4.1 Low-Resolution Reconstruction Problem

The forward model in compressive spectral imaging for acquiring a $N \times N \times L$ data cube is given by the linear system

$$y = Hf,$$  \hspace{1cm} (4.2)

where $f \in \mathbb{R}^v$ is a vector form of the data cube with $v = N^2L$ voxels. Let $D_\Delta \in \mathbb{R}^{\frac{v}{\Delta^2} \times v}$ be a decimation matrix such that a $\frac{N}{\Delta^2} \times \frac{N}{\Delta^2} \times L$ low-resolution version of the object can be obtained as

$$f_\Delta = D_\Delta f,$$  \hspace{1cm} (4.3)

where $\Delta$ represents an integer decimation factor. Figure 4.3 illustrates an example of the structure of $D_\Delta$ for a small data cube with $N = 4$, $L = 2$, and a decimation factor $\Delta = 1$, which means that the resulting data cube exhibits $1/2$ the original spatial resolution.

Figure 4.3: Structure of the decimation matrix $D_\Delta$ for a factor $\Delta = 1$, and a data cube with $N = 4$ and $L = 2$. White elements represent a value of 1, and black entries are zero.

Notice that the high-resolution spectral image can be obtained as $f = D_\Delta^{-1}f_\Delta$, where $D_\Delta^{-1}$ represents the upscaling operator, which is not necessarily the inverse of $D_\Delta$. According to [50], using Eq. (4.3) in Eq. (4.2) provides the equivalent forward model for the low-resolution version of the object.
\[ y = H(D_\Delta^{-1}f_\Delta + f - D_\Delta^{-1}f_\Delta) \]
\[ = HD_\Delta^{-1}f_\Delta + H(f - D_\Delta^{-1}D_\Delta f) \]
\[ = HD_\Delta^{-1}f_\Delta + H(I - D_\Delta^{-1}D_\Delta) f \]
\[ = HD_\Delta^{-1}f_\Delta + \epsilon, \quad (4.4) \]

where the term \( \epsilon = H(I - D_\Delta^{-1}D_\Delta) f \) is an approximation error induced by the decimation and upscaling processes. Moreover, the matrix \( HD_\Delta^{-1} \) is the equivalent sensing matrix for the low-resolution scene. Notice that the left-hand side of Eq. (4.4) remains fixed, which means that the encoded projections of the high-resolution scene will be used to solve the low-resolution inverse problem given by

\[ \arg \min_{\theta_\Delta} ||y - HD_\Delta^{-1}\Psi_\Delta \theta_\Delta||^2_2 + \tau ||\theta_\Delta||_1. \quad (4.5) \]

Finally, the LR scene is obtained by applying the inverse sparse transformation \( \hat{f}_\Delta = \Psi_\Delta^{-1}\theta_\Delta \).

Figure 4.4 illustrates an example of the LR approximation errors \( \epsilon \) for \( \Delta = 1, 2, 3 \), which means \( 1/2, 1/4 \) and \( 1/8 \) of the original resolution. It can be noticed that the errors are concentrated around the high frequency pixels (edges) of the scene, which motivates the use of the intensity gradient maps of the LR reconstruction in the upscaling process.

### 4.2 Multi-Resolution Reconstruction Process

The 3D representation of the LR reconstruction \( \hat{f}_\Delta \) can be expressed as \( \hat{\mathbf{F}}_\Delta = [\hat{\mathbf{F}}^1_\Delta, \cdots, \hat{\mathbf{F}}^L_\Delta] \), where the super-index refers to the spectral bands. This version of the data cube is employed as the starting point of the MR upscaling process, which aims at recovering a \( \frac{N}{2^\Delta} \times \frac{N}{2^\Delta} \times L \) data cube at each iteration \( t = 1, \cdots, \Delta \). Intuitively, an interpolation of \( \hat{f}_\Delta \) could be an alternative for the upscaling process. However, for a large initial decimation factor, interpolation methods are not able to provide detailed approximations of high frequency regions of the object. Therefore, in order to obtain
**Figure 4.4:** Low-resolution approximation errors for different decimation factors $\Delta$. The errors are concentrated around the high frequency regions of the scene.

Improved reconstructions of the high frequency details of the scene, the proposed MR reconstruction approach first calculates a map of the intensity gradient of $\hat{F}_\Delta$.

### 4.2.1 Edge Detection

The gradient intensity map can be calculated using any edge detection method. In this work, the Canny edge detection method is employed due to its good performance and low error rate. In general, this method first smooths the input LR data cube using a Gaussian filter. The smoothed spectral bands of $\hat{F}_\Delta$ can be written as

$$S^\ell_{\Delta} = G \ast \hat{F}^\ell_{\Delta},$$  \hspace{1cm} (4.6)

for $\ell = 1, \cdots, L$, where $G$ represents the Gaussian mask and $\ast$ is the convolution operation. Then, the intensity gradient calculation for each spectral band is given by the partial derivatives with respect to the $x$ and $y$ axes as

$$\frac{\partial S^\ell_{\Delta}}{\partial x \partial y} = P \ast S^\ell_{\Delta},$$  \hspace{1cm} (4.7)

where $P$ is any gradient operator. Using Eq. (4.7), and taking into account that spectral images have high spatial correlation across bands, the intensity gradient map
for the whole data cube can be expressed as the sum of the intensity gradients for all the spectral bands as

\[ E_{\Delta} = \sum_{\ell=1}^{L} P \ast S_{\Delta}^\ell. \]  

(4.8)

### 4.2.2 Reconstruction Upscaling

The intensity gradient map \( E_{\Delta} \) from Eq. (4.8) is used to estimate the edges of the data cube with dimensions \( \frac{N_{2\Delta-1}}{2} \times \frac{N_{2\Delta-1}}{2} \times L \), denoted by \( \hat{F}_{\Delta-1} \). Given that the difference between the upscaling factors between two iterations is always two, the edge estimation for \( \hat{F}_{\Delta-1} \) is given by

\[ E_{\Delta-1} = E_{\Delta} \otimes 1_2, \]  

(4.9)

where \( 1_2 \) is an all-ones \( 2 \times 2 \) matrix.

Let \( Q^\ell \) be the set of coordinates of the points corresponding to an edge in \( E_{\Delta-1} \) defined as

\[ Q^\ell = \{(i,j) | (E_{\Delta-1})_{i,j} = 1\}. \]  

(4.10)

Similarly, the points that do not correspond to an edge can be written as the complement set of \( Q^\ell \) as

\[ (Q^\ell)^C = \{(i,j) | (E_{\Delta-1})_{i,j} = 0\}. \]  

(4.11)

It is worth noting at this point, that Eqs. (4.10) and (4.11) only consider the coordinate points for the first spectral band. For convenience, denote the set of linear indices of the points in \( Q^\ell \) and \( (Q^\ell)^C \) as \( C^\ell \) and \( (C^\ell)^C \), respectively, which are given by

\[ C^\ell = \{c | c = (j - 1)N + i + (\ell - 1)N^2, \forall (i,j) \in Q^\ell\}, \]  

\[ (C^\ell)^C = \{c | c = (j - 1)N + i + (\ell - 1)N^2, \forall (i,j) \in (Q^\ell)^C\}. \]  

(4.12)
Using Eq. (4.12), the total sets of indices for the edge and non-edge points are respectively expressed as
\[ C = \bigcup_{\ell=1}^{L} C^\ell \quad \text{and} \quad C^\ell = \bigcup_{\ell=1}^{L} (C^\ell)^\ell. \] (4.13)

The sets \( C \) and \( C^\ell \) in Eq. (4.13) are now used to determine the new decimation matrix for the reconstruction problem with decimation factor \( \Delta - 1 \), denoted \( D_{\Delta-1} \), which allows the recovery of the spectral object \( \hat{F}_{\Delta-1} \). The process to generate \( D_{\Delta-t} \), the decimation matrix for the iteration \( t \) is presented in section 4.2.3.

The reconstruction of \( \hat{F}_{\Delta-1} \) is thus obtained solving the problem from Eq. (4.1) using \( HD_{\Delta-1}^T \). Because the smooth regions have been recovered at a larger decimation factor, the final reconstruction result for each spectral band is obtained by replicating the recovered smooth pixels in a \( 2 \times 2 \) window, and keeping the edge recovered pixels. Denote the non-edge recovered and replicated pixels of the \( \ell \)-th spectral band as \( \left( \hat{F}_{\ell \Delta-1}^\ell \right)^C \). Then, the final reconstruction is described by the expression
\[ \tilde{F}_{\Delta-1}^\ell = \left( \hat{F}_{\ell \Delta-1}^\ell \circ E_{\Delta-1} \right) + \left( \left( \hat{F}_{\ell \Delta-1}^\ell \right)^C \circ (1 - E_{\Delta-1}) \right), \] (4.14)
where \( \circ \) represents point-wise product, and \( 1 \) is an all-ones matrix.

The process is repeated for \( \Delta - t \) with \( t = 0, \cdots, \Delta \). Notice that a decimation factor of 0 is obtained when \( t = \Delta \) and means that no decimation is performed. However, given the formulation of the MR reconstruction algorithm, when the decimation factor reaches the zero value, edge pixels are not decimated but non-edge pixels are decimated by a factor of 2.

4.2.3 Decimation Matrix Generation

The multi-resolution approach in this work can be seen from two perspectives: the first is that the proposed method allows the reconstruction of different versions of the data cube each one at a different spatial resolution. The second is that the reconstruction of each of those data cubes employs two different decimation factors, a
small one for the edge pixels, and a larger factor for the smooth regions. Therefore, the generation of the MR decimation matrix is based on the two matrices that use the same decimation factor for all pixels. Specifically, at iteration $t$, the equivalent sensing matrix is given by $\mathbf{H} \mathbf{D}_{\Delta-t}^T$. Recall that $\mathbf{D}_{\Delta-t}$ involves a decimation by a factor of $\Delta - t + 1$ for the non-edge pixels in the set $C^\ell$ from Eq. (4.13), and a decimation by a factor of $\Delta - t$ for the edge pixels in the set $C$. Thus, let $\tilde{\mathbf{D}}_{\Delta-t}$ be a matrix that decimates all pixels by a factor $\Delta - t$. Similarly, $\tilde{\mathbf{D}}_{\Delta-t+1}$ decimates all pixels by a factor $\Delta - t + 1$. In general, the goal of each $\tilde{\mathbf{D}}_{\Delta-t}$ is to create pixel blocks of size $2^{\Delta-t} \times 2^{\Delta-t}$. In order to generate an expression for $\tilde{\mathbf{D}}_{\Delta-t}$, which has the structure illustrated in Fig. 4.3, define $\mathbf{d} = 1_{1 \times 2^{\Delta-t}} \otimes [1_{1 \times 2^{\Delta-t}}, 0_{1 \times (N-2^{\Delta-t})}]$, where $1_{1 \times 2^{\Delta-t}}$ is a one-valued row vector of length $2^{\Delta-t}$. Let $\mathbf{\Gamma}$ be a circulant permutation matrix given by

$$\Gamma_{\Delta-t} = \begin{bmatrix} 0_{1 \times N 2^{\Delta-t-1}} & 1 \\ I_{N 2^{\Delta-t-1} \times N 2^{\Delta-t-1}} & 0_{N 2^{\Delta-t-1} \times 1} \end{bmatrix}, \quad (4.15)$$

where $I$ is an identity matrix. Now, define the decimation matrix for a single column of image pixels as

$$\tilde{\mathbf{D}}_{\Delta-t} = \begin{bmatrix} \mathbf{d} \\ \mathbf{d} (\mathbf{\Gamma}_{\Delta-t}^T)^{2^{\Delta-t}} \\ \vdots \\ \mathbf{d} (\mathbf{\Gamma}_{\Delta-t}^T)^{(N-1)2^{\Delta-t}} \end{bmatrix}, \quad (4.16)$$

such that the product $\mathbf{d} (\mathbf{\Gamma}_{\Delta-t}^T)^{\delta}$ shifts the elements of $\mathbf{d}$, $\delta$ positions to the right. Using Eq. (4.16), the decimation matrix that decimates all columns of the image by a fixed factor of $\Delta - t$ is given by

$$\tilde{\mathbf{D}}_{\Delta-t} = I_{\frac{NL}{2^{\Delta-t}}} \times \frac{NL}{2^{\Delta-t}} \otimes \tilde{\mathbf{D}}_{\Delta-t}, \quad (4.17)$$

where $I_{\frac{NL}{2^{\Delta-t}}} \times \frac{NL}{2^{\Delta-t}}$ is an identity matrix. The matrices from Eq. (4.17) for $\Delta - t$ and $\Delta - t + 1$ are then used to generate the MR decimation matrix $\mathbf{D}_{\Delta-t}$. More
specifically, the row entries of $D_{\Delta-t}$ indicate the linear indices of the pixels that will be grouped into a particular super-pixel. Consider the case in which the $i-$th pixel belongs to the set of edge pixels in the set $C$, then the $i-$th row of $D_{\Delta-t}$ is obtained by selecting the $i-$th row of $\tilde{D}_{\Delta-t}$. In contrast, when the $i-$th pixel belongs to the non-edge set $C^c$, the correspondent row of $\tilde{D}_{\Delta-t+1}$ is selected as a new row for $D_{\Delta-t}$. The resulting MR decimation matrix has dimensions $\frac{N^2L}{2(\Delta-t)} \times N^2L$. Because of the decimation of the non-edge pixels by a factor of $\Delta - t + 1$, all-zero rows can appear in the MR decimation matrix. Mathematically, the $i-$th row of $D_{\Delta-t}$ can be obtained as

\[
(D_{\Delta-t})_i = \begin{cases} 
(\chi_{\Delta-t})_i \tilde{D}_{\Delta-t} & , i \in C \\
(\chi_{\Delta-t+1})_i \tilde{D}_{\Delta-t+1} & , i \in C^c \\
0 & \text{otherwise}
\end{cases}, \tag{4.18}
\]

where $i' = \left\lceil \frac{u}{2} \right\rceil$ maps the linear index $i$ to the correspondent row index from $\tilde{D}_{\Delta-t+1}$, with $u = i - \text{mod}(i-1, 2) - \left\lfloor \frac{i-1}{N'} \right\rfloor N' + \left\lfloor \frac{i-1}{2N'} \right\rfloor N'$, $N' = \frac{N}{2\Delta-t}$, $\text{mod} (\cdot)$ is the modulo operation, and $(\chi_{\Delta-t})_d$ is a circulant permutation matrix used for selecting the $d-$th row of a matrix defined as

\[
(\chi_{\Delta-t})_d = \Gamma^d \mu (\Gamma^T)^d, \tag{4.19}
\]

where $\mu$ is an all-zero matrix with a 1 in the $(1, 1)$ position, and $\Gamma$ has the same structure as in Eq. (4.15). The size of $\mu$ and $\chi$ matches the size of the correspondent matrix $\tilde{D}$. Figure 4.5 illustrates a small example of the MR decimation matrix in Eq. (4.18) for $N = 8$, $L = 1$, $\Delta = 2$ and $t = 1$, which means that a $4 \times 4$ data cube would be recovered in this case. The dotted rectangle highlights a row corresponding to a non-edge pixel for which the correspondent row from $\tilde{D}_2$ indexed by $i'$ is assigned to the $i-$th row of $D_1$. Similarly, the solid rectangle indicates an edge pixel for which the $i-$th row of $\tilde{D}_1$ is mapped directly to the $i-$th row of $D_1$. All-zero rows appear because dyadic dimensions are required to apply Wavelets for the sparse approximations during the execution of the CS reconstruction algorithm. However, when different
sparse approximation basis are used, or the sparsity prior is not employed for the reconstruction process, these zero rows can be removed.

Furthermore, a single step MR reconstruction can be formulated for the case in which the highest attainable resolution data cube is directly recovered from the initial LR approximation. In this particular case, the edge pixels are not decimated and non-edge pixels are decimated by a factor $\Delta$. Therefore, rows of $D_\Delta$ are either rows of $\tilde{D}_\Delta$ for non-edge pixels or an identity matrix $I_\nu$ for the edge pixels. This MR decimation
matrix can be modeled as

$$(D_{\Delta})_i = \begin{cases} 
(\chi_{\nu})_i I_{\nu}, & i \in C \\
(\chi_{\Delta})_i \tilde{D}_{\Delta}, & i \in C^L \\
0, & \text{otherwise}
\end{cases} \quad (4.20)$$

Figure 4.6 illustrates an example of the single step MR decimation matrix for $N = 8, L = 1,$ and $\Delta = 1$. The orange rectangle highlights the row corresponding to a non-edge pixel and the green circle indicates the corresponding rows for two edge pixels.

Figure 4.6: Example of the single step MR decimation matrix generation for $\Delta = 1$ from the decimation matrix $\tilde{D}_{\Delta}$ and identity matrix $I_{\nu}$. Illustrated matrices correspond to a data cube of spatial resolution $N = 8,$ and $L = 1$ spectral band.

4.3 Simulation Results

The proposed MR reconstruction approach was tested by simulating the measurement acquisition of the CASSI system [29] for two different data cubes. The
first data cube was acquired by illuminating the scene with a monochromatic light source, and exhibits \( L = 8 \) spectral bands of \( 256 \times 256 \) pixels with central wavelengths \( \{450, 466, 484, 503, 524, 549, 580, 621\} \) nm. The second data cube is a \( 512 \times 512 \times 16 \) version of the Feathers data base from [81], with central wavelengths from 400 nm to 700 nm with steps of 20 nm between bands. The number of snapshots is varied in the simulations from \( K = 2 \) to \( K = 8 \), these values of \( K \) result in compression ratios from \( 1 : 4 \) to \( 1 : 1 \) for the first data cube and \( 1 : 8 \) to \( 1 : 2 \) for the second data cube. The sparse basis is set to be a Kronecker product between a 2D Wavelet for the spatial domain and a DCT for the spectral dimension. Simulations were performed employing the blue noise coded apertures from Chapter 2. Reconstruction quality is measured using peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). All the results are the average of 10 reconstructions for each case. Figure 4.7 illustrates an RGB mapping of the two test data cubes.

![Figure 4.7: RGB mapping of the test data cubes employed for simulations. (Left) \( 256 \times 256 \times 8 \) toy spectral image. (Right) \( 512 \times 512 \times 16 \) feathers spectral image.](image)

### 4.3.1 Multi-Resolution Reconstruction Results

This section illustrates the outputs of the proposed MR reconstruction approach, i.e. the set of sequential reconstructions attained for each data cube. For both cases the maximum decimation parameter was fixed to \( \Delta = 3 \) such that the smallest attained
reconstruction exhibits 1/8 the spatial resolution of the ground truths from Fig. 4.7. For comparison purposes, the HR original data cubes were down-sampled to obtain the correspondent ground truth for each case, using Eq. (4.17). Figures 4.8 and 4.9 illustrate the RGB mappings of the attained reconstructions for $K = 4$ snapshots for the Toy and Feathers data cubes, respectively. A zoomed portion of the reconstructions is included in these figures to illustrate the improvements on the detailed areas as the decimation factor is reduced. Notice that due to the small size of the $32 \times 32$ approximation of the Toy data cube, it can only recover the shapes but not the details in the chest.

Figure 4.8: Multi-resolution reconstructions of the Toy data cube for $K = 4$ snapshots and different values of $\Delta$, along with a zoomed portion of the reconstructions.

In order to obtain a quantitative measure of the reconstruction quality of the attained results, the PSNR and SSIM were calculated for each case. Specifically, each recovered data cube was compared with respect to the correspondent decimated version of the ground truth using the matrices $\tilde{D}_\Delta$ from Eq. (4.17) for the correspondent
values of $\Delta$. Tables 4.1 and 4.2 present the average PSNR and SSIM as a function of the number of captured snapshots for the Toy and Feathers data bases, respectively. These results correspond to the specific case in which the maximum decimation is determined by $\Delta = 3$ which means that the smallest reconstruction exhibits 1/8 the resolution of the original image. Notice that these results are not comparable among them because the quality measures are calculated with respect to a different version of the ground truth for each resolution. In general, these results present the typical behavior of CSI reconstructions in which increasing the number of snapshots yields improved quality images. Moreover, it can be noticed that the SSIM values for the $512 \times 512$ Feathers reconstructions are considerably smaller than the SSIM values of the largest recovered Toy data cube. The compression ratio provides a reasonable explanation for this situation. Recall that for the Toy data cube, $K = 8$ is equivalent to a compression ratio $1 : 1$, while the same amount of snapshots for the Feathers data

Figure 4.9: Multi-resolution reconstructions of the Feathers data cube for $K = 4$ snapshots and different values of $\Delta$, along with a zoomed portion of the reconstructions.
cube represents a compression ratio $1 : 2$.

### Table 4.1: Average reconstruction quality results for the Toy data cube in terms of PSNR and SSIM, for different number of shots.

<table>
<thead>
<tr>
<th>Resolution</th>
<th>$K = 2$</th>
<th>$K = 4$</th>
<th>$K = 6$</th>
<th>$K = 8$</th>
<th>$K = 2$</th>
<th>$K = 4$</th>
<th>$K = 6$</th>
<th>$K = 8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>32 × 32</td>
<td>29.5874</td>
<td>33.1640</td>
<td>34.7769</td>
<td>35.4658</td>
<td>0.9499</td>
<td>0.9799</td>
<td>0.9872</td>
<td>0.9897</td>
</tr>
<tr>
<td>64 × 64</td>
<td>25.8999</td>
<td>27.3594</td>
<td>27.7039</td>
<td>27.8216</td>
<td>0.8338</td>
<td>0.8786</td>
<td>0.8891</td>
<td>0.8931</td>
</tr>
<tr>
<td>128 × 128</td>
<td>26.2552</td>
<td>29.2768</td>
<td>30.3123</td>
<td>30.7339</td>
<td>0.7332</td>
<td>0.8193</td>
<td>0.8477</td>
<td>0.8614</td>
</tr>
<tr>
<td>256 × 256</td>
<td>27.7183</td>
<td>29.7584</td>
<td>31.4789</td>
<td>32.5624</td>
<td>0.6007</td>
<td>0.7116</td>
<td>0.7570</td>
<td>0.7827</td>
</tr>
</tbody>
</table>

### Table 4.2: Average reconstruction quality results for the Feathers data cube in terms of PSNR and SSIM, for different number of shots.

<table>
<thead>
<tr>
<th>Resolution</th>
<th>$K = 2$</th>
<th>$K = 4$</th>
<th>$K = 6$</th>
<th>$K = 8$</th>
<th>$K = 2$</th>
<th>$K = 4$</th>
<th>$K = 6$</th>
<th>$K = 8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>64 × 64</td>
<td>29.6963</td>
<td>33.7488</td>
<td>35.7748</td>
<td>37.1863</td>
<td>0.6866</td>
<td>0.7811</td>
<td>0.8262</td>
<td>0.8523</td>
</tr>
<tr>
<td>128 × 128</td>
<td>26.4417</td>
<td>28.6165</td>
<td>29.3200</td>
<td>29.6176</td>
<td>0.5739</td>
<td>0.6683</td>
<td>0.7057</td>
<td>0.7257</td>
</tr>
<tr>
<td>256 × 256</td>
<td>27.9920</td>
<td>28.6725</td>
<td>30.7557</td>
<td>31.9634</td>
<td>0.4920</td>
<td>0.6002</td>
<td>0.6569</td>
<td>0.6907</td>
</tr>
<tr>
<td>512 × 512</td>
<td>28.8410</td>
<td>32.2385</td>
<td>33.3903</td>
<td>34.0898</td>
<td>0.3406</td>
<td>0.4842</td>
<td>0.5519</td>
<td>0.5945</td>
</tr>
</tbody>
</table>

### 4.3.2 Reconstruction Complexity

The computational complexity of compressive spectral imaging reconstructions depends on the size of the acquired data which is related to the dimensionality of the sensing matrix. Typical compressed sensing reconstruction algorithms used to solve the inverse problem in Eq. 4.1 involve calculation of inner products, matrix pseudo inverses, and sparse transformations. The cost of each operation depends on the dimensionality of the signal to recover. For instance, the amount of floating point operations per GPSR iteration is $O(KN^4L)$ [48, 22], when the data cube to recover has $N \times N$ pixels and $L$ spectral bands. In contrast, when the MR reconstruction approach is employed, the dimensions of the data cube to recover are $\frac{N}{2^t} \times \frac{N}{2^t} \times L$, for each value of $t = 0, \cdots , \Delta$. Therefore, the resultant complexity of the MR reconstruction approach is $O\left(\frac{K N^4 L}{16^\Delta}\right)$, where the value of $\Delta$ varies according to the iteration $t$. Figure
4.10 illustrates a comparison of the computational complexity for the full-resolution reconstruction in Eq. 4.1 and the MR reconstruction from Eq. 4.5 for \( \Delta = 1, 2 \), which means 1/2 and 1/4 the original resolution. To better visualize the difference between the curves, top-left inner square shows a zoom portion of the graphic for \( N = 32 \) to \( N = 128 \), where the exponential behavior of the traditional reconstruction is clearly seen. Similarly, the bottom-right figure shows a zoom comparison between the MR curves from \( N = 512 \) to \( N = 2048 \).

**Figure 4.10:** Reconstruction complexity as a function of the spatial dimensions of the data cube to recover \( N \), \( L = 5 \) spectral bands, and \( K = 1 \) snapshot, for the full-resolution reconstruction, and MR with \( \Delta = 1, 2 \). Zoomed portions of the figure are also included.

### 4.3.3 Multi-resolution Reconstructions versus Interpolation

Given that interpolation is an alternative solution to obtain up-scaled versions of the data cube, the main goal of this experiment is to compare the reconstructions obtained using the MR approach and the results from LR interpolated reconstructions. For this purpose, the attained MR reconstructions for both data cubes were up-scaled to the larger dimensions using a bicubic interpolator. The average PSNR and SSIM were calculated with respect to the correspondent decimated version of the ground truth. These results are summarized in Tables 4.3 and 4.4, and it can be noticed that,
as expected, interpolation is not a good alternative for up-scaling the LR recovered spectral images when the ratio between the initial and target resolutions is large. For instance, consider the extreme case in which the $256 \times 256$ approximation is obtained as the interpolation of the $32 \times 32$ reconstruction, very low PSNR and SSIM values are obtained, however these quality measures start increasing as the target resolution approaches that of the LR available data cube. The results for the Feathers data cube present similar behavior.

**Table 4.3:** Quality of interpolated LR reconstructions using a bicubic interpolator for the Toy data cube.

<table>
<thead>
<tr>
<th>Initial Resolution</th>
<th>Shots</th>
<th>Target Resolution</th>
<th>PSNR</th>
<th>SSIM</th>
<th>PSNR</th>
<th>SSIM</th>
<th>PSNR</th>
<th>SSIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>32 × 32</td>
<td>2</td>
<td>64 × 64</td>
<td>24.81</td>
<td>0.7725</td>
<td>23.04</td>
<td>0.5040</td>
<td>22.06</td>
<td>0.2949</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>64 × 64</td>
<td>25.29</td>
<td>0.8144</td>
<td>23.33</td>
<td>0.5400</td>
<td>22.29</td>
<td>0.3133</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>64 × 64</td>
<td>25.34</td>
<td>0.8232</td>
<td>23.35</td>
<td>0.5483</td>
<td>22.31</td>
<td>0.3174</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>64 × 64</td>
<td>25.33</td>
<td>0.8260</td>
<td>23.34</td>
<td>0.5516</td>
<td>22.31</td>
<td>0.3190</td>
</tr>
<tr>
<td>64 × 64</td>
<td>2</td>
<td>64 × 64</td>
<td>24.44</td>
<td>0.6053</td>
<td>23.20</td>
<td>0.3698</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>64 × 64</td>
<td>24.99</td>
<td>0.6444</td>
<td>23.63</td>
<td>0.3966</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>64 × 64</td>
<td>25.09</td>
<td>0.5637</td>
<td>23.71</td>
<td>0.4037</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>64 × 64</td>
<td>25.09</td>
<td>0.5669</td>
<td>23.71</td>
<td>0.4060</td>
<td></td>
<td></td>
</tr>
<tr>
<td>128 × 128</td>
<td>2</td>
<td>128 × 128</td>
<td>-</td>
<td>-</td>
<td>25.71</td>
<td>0.5234</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>128 × 128</td>
<td>-</td>
<td>-</td>
<td>27.40</td>
<td>0.5944</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>128 × 128</td>
<td>-</td>
<td>-</td>
<td>27.79</td>
<td>0.6172</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>128 × 128</td>
<td>-</td>
<td>-</td>
<td>27.90</td>
<td>0.6288</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results from Tables 4.3 and 4.4 can be compared to those obtained with the MR reconstruction approach, in order to evaluate the performance of the proposed method with respect to the interpolation. Therefore, the reconstructions from both methods are individually compared for each target resolution. In other words, these comparisons are intended to find out whether there is an improvement in the reconstructions when the MR approach is used. Figures 4.11 and 4.12 present the average PSNR and SSIM curves as a function of the number of shots $K$ for each target size of the Toy and Feathers data cubes, respectively. In general, the results demonstrate
that the multi-resolution approach provides better up-scaled reconstructions than interpolation.

**Table 4.4:** Quality of interpolated LR reconstructions using a bicubic interpolator for the Feathers data cube.

<table>
<thead>
<tr>
<th>Initial Resolution</th>
<th>Shots</th>
<th>128 × 128</th>
<th>256 × 256</th>
<th>512 × 512</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PSNR</td>
<td>SSIM</td>
<td>PSNR</td>
</tr>
<tr>
<td>64 × 64</td>
<td>2</td>
<td>26.73</td>
<td>0.4673</td>
<td>25.39</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>27.91</td>
<td>0.5512</td>
<td>26.27</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>28.27</td>
<td>0.5914</td>
<td>26.54</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>28.44</td>
<td>0.6152</td>
<td>26.67</td>
</tr>
<tr>
<td>128 × 128</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>26.10</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>27.40</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>-</td>
<td>-</td>
<td>27.79</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>-</td>
<td>-</td>
<td>27.94</td>
</tr>
<tr>
<td>256 × 256</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 4.13 presents a visual comparison of the MR reconstructions mapped to their RGB profiles with respect to the interpolated versions for $\Delta = 3, 2, 1$ using $K = 4$ snapshots and both databases. Similarly, Fig. 4.14 illustrates the comparison for the 128 × 128 Toy data cube and the 256 × 256 Feathers reconstruction.
Figure 4.11: Average reconstruction PSNR and SSIM comparison for $64 \times 64$, $128 \times 128$ and $256 \times 256$ target resolutions of the Toy data cube using MR reconstruction from $\Delta = 3$ and bicubic interpolations.

Figure 4.12: Average reconstruction PSNR and SSIM comparison for $128 \times 128$, $256 \times 256$ and $256 \times 256$ target resolutions of the Feathers data cube using MR reconstruction from $\Delta = 3$ and bicubic interpolations.
Figure 4.13: Data cube reconstruction comparison between MR and interpolated LR images from $\Delta = 3, 2, 1$ using $K = 4$ snapshots for (Top) $256 \times 256$ Toy database and (Bottom) $512 \times 512$ Feathers database.
Figure 4.14: Data cube reconstruction comparison between MR and interpolated LR images from $\Delta = 3, 2$ using $K = 4$ snapshots for (Top) $128 \times 128$ Toy database and (Bottom) $256 \times 256$ Feathers database.
4.3.4 Sequential MR Reconstructions versus Single Step MR

Besides recovering sequential versions of the data cube at different spatial resolutions, the multi-resolution approach can be employed to recover a single up-scaled version of a lower-resolution reconstruction as described in section 4.2.3. Suppose that the initial LR approximation is obtained at $\Delta = 2$, according to the sequential MR approach the next reconstruction will be obtained at $\Delta = 1$. Instead, the single step MR will recover the high-resolution data cube, i.e. $\Delta = 0$. Intuitively, larger values of $\Delta$ will result in poor final reconstruction quality because the mapping of edge pixel coordinates to the HR grid will yield to coarse edges. Furthermore, non-edge recovered pixel values will be replicated to large areas in which edge pixel can be easily lost. Simulations were performed in order to compare the sequential MR reconstructions with the single step MR approach. Specifically, both test data cubes were recovered using single step MR decimation matrices from Eq. 4.20 for $\Delta = 1, 2, 3$. Figure 4.15 presents the average reconstruction PSNR and SSIM as a function of the number of snapshots for the sequential MR and single step MR (SMR) reconstructions. These results show that, as expected, the reconstruction quality of SMR increases as the value of $\Delta$ is reduced. Also, it can be noticed that PSNR curves for MR and SMR-$\Delta = 1$, are very similar in both data cubes. However, a larger difference can be noticed in the SSIM curves for the Toy data cube, which implies that the sequential MR provides extra information in order to obtain more accurate reconstructions. The SSIM for the Feathers data cube does not present this behavior, but recall that only 25% of the data is used in these reconstructions, since the Feathers data cube contains $L = 16$ spectral channels.

Figure 4.16 illustrates a comparison of the full-resolution test data cubes using the single step MR approach for $\Delta = 3, 2, 1$ and $K = 6$ snapshots. The attained SSIM and PSNR values are indicated on each image and a zoomed portion for each case is also included. The artifacts in the reconstructions for $\Delta = 3$ are due to the coarse edge estimation in the initial LR reconstruction but they start to vanish as the value of $\Delta$ is reduced. However, there exists a trade-off between the value of $\Delta$ and
the complexity of the inverse problem, therefore a correct selection of the decimation factor should take into account the amount of computational resources available and the application/user tolerance to artifacts for the image preview.

4.4 Conclusions and Directions for Future Work

A MR reconstruction approach for compressive spectral imaging has been proposed. The mathematical model for the MR methodology has been developed. Simulations show that, in contrast to interpolated solutions, the sequential MR approach can better approximate the high-frequency pixels of the scene. In addition, the single step MR reconstruction has been tested. Simulations show that reconstruction quality of the single step MR approach improves as the decimation factor decreases. Further, sequential MR reconstructions and the single step MR for $\Delta = 1$ provide similar results.

Figure 4.15: Comparison of the average reconstruction PSNR and SSIM for the sequential MR reconstruction (MR) from $\Delta = 3$ and the single-step MR reconstruction (SMR) for $\Delta = 1, 2, 3$. 

89
Figure 4.16: Single step MR reconstructions of the high-resolution data cubes along with their zoomed portions, for (Left) $\Delta = 3$, (Center) $\Delta = 2$ and (Right) $\Delta = 1$ using $K = 6$ snapshots. Each image includes the correspondent attained SSIM and PSNR values.
However, since there exists a trade-off between the decimation factor and the complexity of the inverse problem, the selection of $\Delta$ depends on the available computational resources and the desired accuracy of the image preview for a specific application.

The MR reconstruction approach is prone to several improvements, and there is plenty of research directions in this topic. For instance, one research direction involves the use of other CS reconstruction problem formulations such that the dyadic constraint of the Wavelet dictionary can be avoided. Therefore, the MR sensing matrix from Eq. (4.18) does not include the rows of zeros, which will accelerate the reconstruction process. Another approach to avoid the dyadic constraint of the Wavelet transformation but still using the sparsity prior, is to employ dictionary learning techniques. Moreover, a 3-dimensional decimation approach can be considered in future research works, given that this dissertation only analyzed the case in which the decimation is applied to the spatial dimensions. Three-dimensional decimation provides smaller MR sensing matrices which results in faster previews of the spectral image. However, the 3-dimensional decimation also imposes new challenges since the maximum decimation scale should be carefully selected to avoid the loss of spectral characteristics that could affect material identification accuracy.

On the other hand, the proposed MR approach can be employed in different spectral imaging applications such as classification, detection and coded aperture design. Specifically, classification and detection analysis can be performed on lower-resolution approximations of the data cube, instead of relying on the HR version. In addition, coded aperture design tools such as coherence, restricted isometry property and gradient-based methods involve the analysis of the system sensing matrix. One of the main limitations of these methods is the high computational burden to calculate matrix products even for relatively small image dimensions. For instance, the dimensions of the CASSI sensing matrix associated to a $128 \times 128 \times 8$ spectral image are $17280 \times 131072$. Therefore, further research can focus on the development of coded aperture design methods employing MR sensing matrices, such that the computational load of the system matrix analysis is considerably reduced.
BIBLIOGRAPHY


Appendix A

COPYRIGHT NOTICE

This appendix contains the copyright notices for the contents of this dissertation.

Chapter 2 was published in the Journal of the Optical Society of America A (JOSA A). ©2016 Optical Society of America.

Chapter 3 was published in the Journal of the Optical Society of America A (JOSA A). ©2015 Optical Society of America.

One print or electronic copy of these papers may be made for personal use only. Systematic reproduction and distribution, duplication of any material in these papers for a fee or for commercial purposes, or modifications of the content of these papers are prohibited. For further information regarding Optical Society of America copyright and licensing information go to https://www.osapublishing.org/submit/review/copyright_permissions.cfm#posting
Appendix B

DISCLAIMER

Figures presented in this dissertation were not endorsed by the trademark owners and are used here to illustrate the quality of obtained reconstructions from compressive spectral imaging measurements and the proposed methods.
LEGO® is a trademark of the LEGO Group of companies which does not sponsor, authorize or endorse the images in this dissertation. ©2017 The LEGO Group. All Rights Reserved.