DOI: http://dx.doi.org/10.15446/ing.investig.v36n2.56426

Probability of correct reconstruction in compressive spectral imaging

Probabilidad de reconstrucción exitosa en el sensado compreso de imágenes espectrales

Samuel E. Pinilla[1], Héctor M. Vargas[2], and H. Arguello[3]

ABSTRACT

Coded Aperture Snapshot Spectral Imaging (CASSI) systems capture the 3-dimensional (3D) spatio-spectral information of a scene using a set of 2-dimensional (2D) random coded Focal Plane Array (FPA) measurements. A compressed sensing reconstruction algorithm is then used to recover the underlying spatio-spectral 3D data cube. The quality of the reconstructed spectral images depends exclusively on the CASSI sensing matrix, which is determined by the statistical structure of the coded apertures. The Restricted Isometry Property (RIP) of the CASSI sensing matrix is used to determine the probability of correct image reconstruction and provides guidelines for the minimum number of FPA measurement shots needed for image reconstruction. Further, the RIP can be used to determine the optimal structure of the coded projections in CASSI. This article describes the CASSI optical architecture and develops the RIP for the sensing matrix in this system. Simulations show the higher quality of spectral image reconstructions when the RIP property is satisfied. Simulations also illustrate the higher performance of the optimal structured projections in CASSI.

Keywords: Restricted Isometry Property, RIP, CASSI, compressive sensing, spectral imaging, coded aperture.

RESUMEN

El sistema de adquisición de imágenes espectrales de única captura basado en apertura codificada (CASSI), capta información tridimensional (3D) espacio-espectral de una escena, usando un conjunto de medidas bidimensionales (2D) proyectadas en un FPA (Focal Plane Array). Para recuperar el cubo de datos a partir de las proyecciones en el FPA, se usa un algoritmo de reconstrucción basado en la teoría de muestreo compresivo. En CASSI la calidad de la reconstrucción de imágenes espectrales depende exclusivamente de la matriz de sensado, que es determinada por la estructura estadística del código de apertura. La propiedad restringida isométrica (RIP) de la matriz de sensado CASSI es usada para determinar la probabilidad de una correcta reconstrucción de la imagen. Este artículo describe la arquitectura óptima CASSI y desarrolla la RIP para las matrices de muestreo, para la captura de la información del cubo de datos. En efecto, la RIP provee la guía para determinar el mínimo número de capturas FPA necesarias para la reconstrucción de una imagen. Más adelante, la RIP es usada para encontrar la estructura óptima de las proyecciones de los códigos de apertura de CASSI. Las simulaciones muestran alta calidad de la reconstrucción obtenida de las imágenes espectrales cuando se satisface la condición impuesta por la RIP. También muestran el más alto rendimiento obtenido de las estructuras óptimas de las proyecciones CASSI.

Palabras clave: Propiedad Isométrica Restrictiva, RIP, CASSI, muestreo compresivo, imágenes hiperespectrales, códigos de apertura.

Received: April 23rd 2016
Accepted: July 18th 2016

Introduction

Imaging spectroscopy requires sensing a large amount of spatial information across a multitude of wavelengths. The sensed signals are titled multispectral or hyperspectral images. Traditional imaging spectroscopy sensing techniques scan adjacent zones of the underlying spectral scene and merge the results to construct a spectral 3-Dimensional (3D) data cube. Push broom spectral imaging sensors, for instance, capture a spectral data cube with one FPA measurement per

[1] System Engineer, Masters in Math, Universidad Industrial de Santander, Colombia. Affiliation: Department of Computer Science, Universidad Industrial de Santander, Colombia. E-mail: samuel.pinilla@correo.uis.edu.co.

[2] Electronic Engineer, Masters in System Engineering, Universidad Industrial de Santander, Colombia. Affiliation: Department of Computer Science, Universidad Industrial de Santander, Colombia.
E-mail: hector.vargas@correo.uis.edu.co.

[3] Electrical Engineer, Masters in Electrical Power, Universidad Industrial de Santander. Ph.D in Electrical and Computer Engineering, University of Delaware, USA. Affiliation: Department of Computer Science, Universidad Industrial de Santander, Colombia. E-mail: henarfu@uis.edu.co.

How to cite: Pinilla, E., Vargas, M., & Arguello, H. (2016). Probability of correct reconstruction in compressive spectral imaging. *Ingeniería e Investigación*, *36(2)*, 68 – 77. DOI: 10.15446/ing.investig.v36n2.56426.

spatial line of the scene (Gehm *et al.,* 2008). Spectrometers based on optical band-pass filters spectrally scan the source data cube (Gupta, 2008). These sensing techniques obey the well-known Nyquist criterion, which imposes a severe limit on the number of samples needed. More specifically, these methods require scanning a number of zones linearly in proportion to the desired spatial or spectral resolution. As the desired resolution increases, the required number of samples grows considerably, such that the cost of sensing a hyperspectral image is extremely high. Recently, with a mathematical technique for reconstructing sparse signals, called Compressive Sensing (CS), one can use fewer samples much lesser then predicted by the Nyquist criterion. This new technique involves diverse mathematical areas such as numerical optimization, signal processing, random matrix analysis, and statistics. The enormous potential of CS has been recently applied in areas such as microscopy, holography, tomography and spectroscopy (Arce *et al.*, 2014; Brady *et al.*, 2009; Studer et al., 2012; Wagadarikar *et al.*, 2008; Yu and Wang, 2009). CS allows sensing a signal with a fewer number of samples than that required by the Nyquist criterion. Thus, CS has allowed overcoming a diverse number of physical sensing limits. This paper focuses on the application of CS in spectral imaging which is coined Compressive Spectral Imaging (CSI). CSI senses 2D coded random projections of the underlying scene, such that the number of required projections is far less than the linear scanning case. The Coded Aperture Snapshot Spectral Imaging (CASSI) system is an imaging architecture that effectively implements CSI. Thus, CASSI senses the 3D spectral information of a scene using 2D random projections. Figure 1 illustrates the CASSI optical architecture. The main components in CASSI are the coded aperture, the prism and the Focal Plane Array (FPA). The coded aperture patterns are the only varying elements in CASSI, the other optical elements remain fixed during the operation of the instrument.

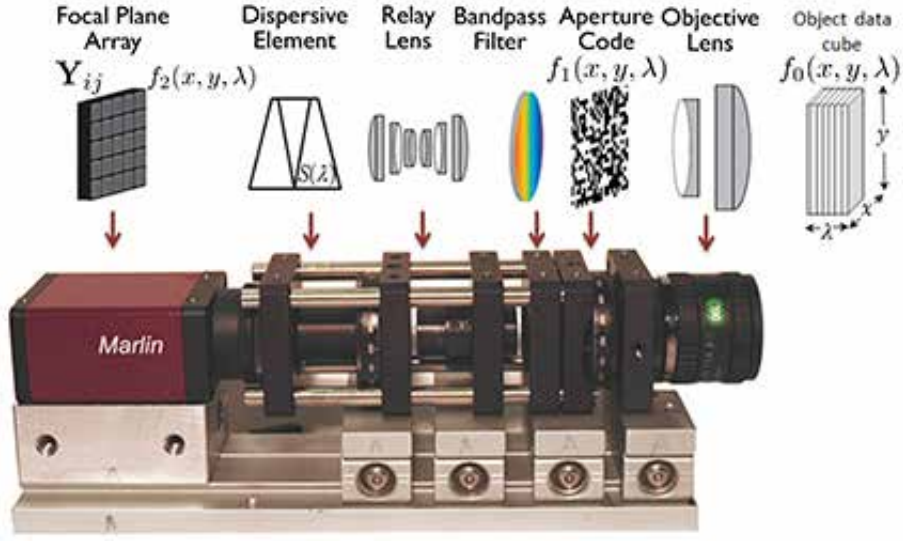

Figure 1. CASSI architecture components (Kittle et al., 2010).

Figure 2 shows the main components of the CASSI architecture. In CASSI, the coding is applied to the image source density $f_0(x, y, \lambda,)$ by the coded aperture $T(x, y)$, where (x, y) are the spatial coordinates. The resulting coded density $f_1(x, y, \lambda,) = T(x, y)\, f_0(x, y, \lambda,)$ is spectrally dispersed by a dispersive element before it impinges in the focal plane array (FPA), resulting in the coded field

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

where $T(x,y)$ is the transmission function representing the coded aperture, $h\left(x'-S\left(\lambda\right)-x, y'-y\right)$ is the optical impulse response of the system, and $S(\lambda)$ is the dispersion induced by the prism. A linear dispersion by the prism is assumed. The compressive measurements across the FPA are obtained by the integration of the field $f_2(x, y, \lambda)$ over the detector spectral range. The discretized FPA measurements can be thus modeled as

$$Y_{j\ell} = \sum_{k=0}^{L-1-a_\ell} F_{j(\ell+k)(k)} T_{j(\ell+k)} + \omega_{j\ell}, \quad (2)$$

where L is the number of spectral bands sensed, $a_\ell = floor\left(\ell / \left(N-L\right)\right)\left(\ell - N + L + 1\right)$, $Y_{j\ell}$ is the intensity measured at the (j,ℓ) position of the FPA with $1 \le j,\ell \le N, F \in \mathbb{R}^{N\times N\times L}$, is the discretized data cube, $T \in \left\{0,1\right\}^{N\times N}$ is the discretized coded aperture with two-dimensional structure, and $\omega_{j\ell}$ is the noise system.

The single-shot CASSI system has been extended to a multi-shot system architecture such that multi-frame coded measurements are acquired as separate FPA measurements, each with a distinct coded-aperture pattern (Kittle *et al.,* 2010).

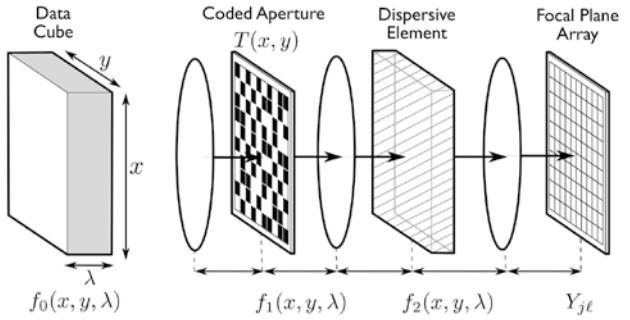

Figure 2. Optical elements present in CASSI.

Equation (2) can be written in matrix form as $\mathbf{y}=\mathbf{Hf}+\boldsymbol{\omega}$, where $\boldsymbol{y}$ is the measurements vector, $\mathbf{H}$ is the matrix that represents the coded apertures and prism effect and $\mathbf{f}$ is the vectorized spectral data cube. Given the compressive measurements $\boldsymbol{y}$, the objective of CS is to recover the signal $\mathbf{f}$. Compressive Sensing (CS) dictates that an n-long signal $\mathbf{f}$ can be recovered from $m \ll n$ random projections $\mathbf{y}=\mathbf{H\Psi\theta}$. Moreover, if it defines the $m \times n$ matrix $\mathbf{A}=\mathbf{H\Psi}$ as the sensing matrix, then the spectral data cube $\mathbf{f}=\mathbf{\Psi\theta}$ is sparse in the base $\mathbf{\Psi}$. The RIP in CASSI is thus fundamental to determine both the minimum number of shots needed for correct reconstruction and for the design and optimization of the coded aperture patterns (Wagadarikar, 2010). CS structured random matrices have been extensively analyzed for Toeplitz, random partial Fourier, and partial circulant matrices (Fornaiser, 2010). Recently a generalization of structurally random matrices has been developed (Do *et al.*, 2012). The RIP in CASSI was first considered in Wagadarikar (2010), where it is assumed that the RIP in CASSI is satisfied. The work in Wagadarikar (2010) then derives conditions on the coded apertures so that the RIP is better satisfied. These results, however, do not develop the explicit parameters for the bounds needed in the RIP, such as the probability of error, or the minimum number of FPA measurements. In this paper the RIP in CASSI is derived, the structure of the CASSI sensing matrix is formulated, and the RIP constants are expressed as a function of the structure of the random coded aperture patterns. A minimum number of random projections are needed so as to satisfy the RIP of the sensing matrix $\mathbf{A}$ (Candes and Tao, 2005). This paper establishes the theoretical probability of correct recovering in CSI. The results of this work can be used in the optimization of coded aperture patterns and to estimate a bound for the minimum number of shots needed for correct reconstruction.

CASSI Sensing Matrix

The goal in this section is to present the discrete model of the optical process in CASSI system in matrix form. Moreover, using this model allows carrying out the RIP analysis on the CASSI sensing matrix. The discretized output at the detector corresponding to the i^{th} coded aperture $T^i\left(j,\ell\right)\in\left\{-1,1\right\}$[1] is given by

$$Y^i_{j\ell}=\sum_{k=0}^{L-1-a_\ell} F_{j(\ell+k)(k)}T^i_{j(\ell+k)}+\boldsymbol{\omega}_{j\ell}, \tag{3}$$

for $i=1,\cdots,K$, where $Y^i_{j\ell}$ is the intensity at the (j,ℓ) position at the detector, whose dimensions are $N\times N+L-1$, F represents an $N\times N\times L$ spectral data cube and $\boldsymbol{\omega}_{j\ell}$ is the white noise of the sensing system. In this work, the entries of the coded aperture are limited to $T^i\left(j,\ell\right)\in\left\{-1,1\right\}$. The varying terms in Equation (3), indexed by i are: first, the discrete coded apertures T^i or its vector representation $\mathbf{t}^i$ whose entries are realizations of a Bernoulli random variable with parameter $p=P\left\{t^i_\ell=1\right\}=1-P\left\{t^i_\ell=1\right\}$ and second, the number of shots K. (See Appendix A).

The i^{th} FPA measurement $Y^i_{j\ell}$ can be written in matrix notation as

$$\mathbf{y}_i=\mathbf{H}_i\mathbf{f}+\boldsymbol{\omega}, \tag{4}$$

where $\mathbf{y}_i\in\mathbb{R}^v$ is a V-long vector representation of the measurement $\mathbf{Y}^i\in\mathbb{R}^{N\times(N+L-1)}$ in Equation (3), where $V=N(N+L-1)$, $\mathbf{H}_i\in\left\{-1,0,1\right\}$ represents the coded aperture and the dispersive element where $n=N^2L$, $\mathbf{f}=\left[\mathbf{f}_0^T\mathbf{f}_1^T\cdots\mathbf{f}_{L-1}^T\right]$ is the vector representation of the data cube $\boldsymbol{F}$, and $\boldsymbol{\omega}$ is the noise of the system. More specifically, the entries of $\mathbf{f}_k$ can be expressed as

$$(f_k)_\ell=F_{(\ell-rN)rk},\ \text{for}\ \ell=0,...,N^2-1,\ k=0,...,L-1, \tag{5}$$

where $r=floor\left(\ell/N\right)$. The vectorization of the code aperture $\mathbf{T}^i$ is defined as

$$\left(t^i\right)_\ell=T^i_{(\ell-rN)r},\ \text{for}\ \ell=0,...,N^2-1,\ k=0,...,L-1, \tag{6}$$

where $r=floor\left(\ell/N\right)$. Similarly, the vectorization of the output $\mathbf{Y}^i$ is defined as

$$(y_i)_\ell=Y^i_{(\ell-rN)r},\ \text{for}\ \ell=0,...,\mathrm{V}-1,\ k=0,...,L-1, \tag{7}$$

where $r=floor\left(\ell/N\right)$. Using the above matrix representation, the output $\boldsymbol{y}_i$ can be expressed as

$$\mathbf{y}_i=\overbrace{\left[\begin{array}{c|c|c|c} diag(\boldsymbol{t}^i) & \mathbf{0}_{N(1)N^2} & \cdots & \mathbf{0}_{N(L-1)N^2} \\ & diag(\boldsymbol{t}^i) & \cdots & \\ & & \ddots & \\ \mathbf{0}_{N(L-1)N^2} & \mathbf{0}_{N(L-2)N^2} & \cdots & diag(\boldsymbol{t}^i) \end{array}\right]}^{\boldsymbol{H}_i}\left[\begin{array}{c} \boldsymbol{f}_0 \\ \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \\ \vdots \\ \boldsymbol{f}_{L-1} \end{array}\right] \tag{8}$$

where $diag(\mathbf{t}^i)$ is an $N^2\times N^2$ diagonal matrix whose entries are the elements of the vectorized coded aperture $\mathbf{t}^i$,

$\mathbf{0}_{N(1)\times N^2}$ and $\mathbf{0}_{N(L-1)\times N^2}$ are $N(1)\times N^2$ and $N(L-1)\times N^2$ zero-valued matrices, respectively.

The matrix $\mathbf{H}_i$ can be expressed as (Arguello and Arce, 2012)

$$\mathbf{H}_i = \sum_{j=0}^{L-1} (\mathbf{\Theta}_V^N)^j \left[\begin{array}{cc} diag(\mathbf{t}^i) & \mathbf{0}_R \end{array} \right]^T \overline{\mathbf{I}} \left(\mathbf{\Theta}_n^{-N}\right)^{2j}, \quad (9)$$

where $\mathbf{0}_R$ is a $N^2 \times N(L-1)$ zero-valued matrix, the structure of the $V \times V$ permutation matrix $\mathbf{\Theta}_v$ is

$$\mathbf{\Theta}_V = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \quad (10)$$

The $n \times n$ matrix $\mathbf{\Theta}_n$ has the same structure of $\mathbf{\Theta}_v$ in Equation (10) and $\overline{\mathbf{I}}$ is $\overline{\mathbf{I}} = \left[\mathbf{I}\ \mathbf{0} \right]$ where $\mathbf{I}$ is an $N^2 \times N^2$ identity matrix, $\mathbf{\Theta}_V^N = \mathbf{\Theta}_v \mathbf{\Theta}_v \ldots \mathbf{\Theta}_v$, $\mathbf{\Theta}_V^{-N} = \mathbf{\Theta}_V^{-1} \mathbf{\Theta}_V^{-1} \ldots \mathbf{\Theta}_V^{-N}$ are $N-times$ the usual matricial product, and $\mathbf{0}$ is a $N^2 \times N^2(L-1)$ zero valued matrix.

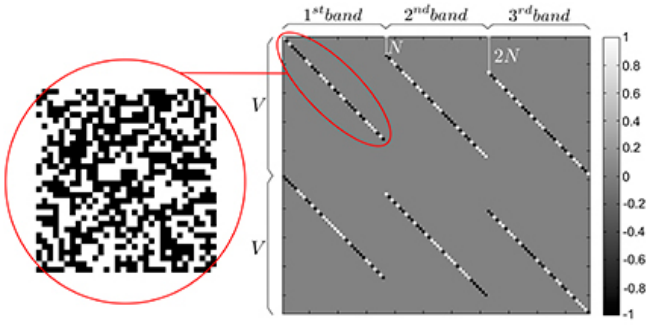

Figure 3. Structure of the matrix $\mathbf{H}$ for K = 2, N = 6 and L = 3. The diagonals in the $\mathbf{H}$ matrix are the vectorized coded apertures.

The ensemble of CASSI outputs $\mathbf{y} = \left[\mathbf{y}_0^T, \ldots, \mathbf{y}_{k-1}^T \right]^T$ can be rewritten as

$$\mathbf{y} = \mathbf{A}\boldsymbol{\theta} = (\mathbf{H}\mathbf{\Psi})\boldsymbol{\theta} + \boldsymbol{\omega}, \quad (11)$$

where the CASSI sensing matrix is $\mathbf{A} = \mathbf{H}\mathbf{\Psi}$, the matrix $\mathbf{H} = \left[\mathbf{H}_0^T, \ldots, \mathbf{H}_{k-1}^T \right]^T$, and $\boldsymbol{\theta}$ is a sparse representation of $\mathbf{f}$ in the base $\mathbf{\Psi}$. Notice that $\mathbf{A} \in \mathbb{R}^{m \times n}$ where $m = KV$, and $m \leq n$. Figure 3 and Figure 4 depict an example of the matrices $\mathbf{H}$ and $\mathbf{A}$ respectively for small values of N, L, and K. In practice, the dimensions of these matrices are much larger (Arguello and Arce, 2011). Figure 4 uses $\mathbf{\Psi} = diag\left(\mathbf{\Psi}_{2D,\ldots,} \mathbf{\Psi}_{2D} \right)$ where $\mathbf{\Psi}_{2D}$ is the 2D Symmlet Wavelet Transform.

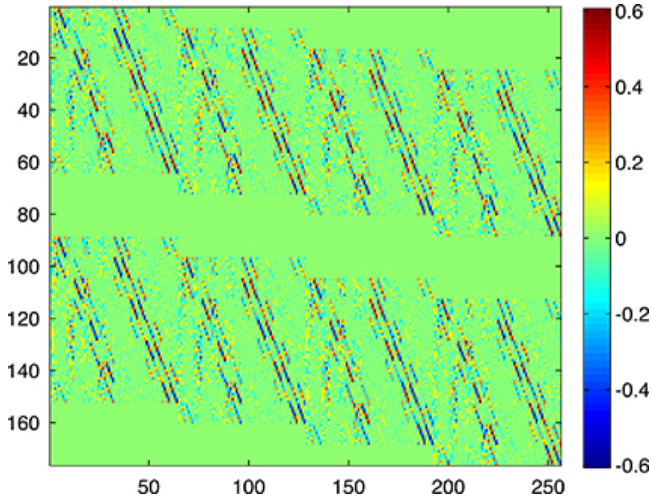

Figure 4. Structure of the matrix $\mathbf{A}$ for K = 2, N = 8 and L = 4 and $\mathbf{\Psi}$ is based on the 2D Symmlet Wavelet Transform.

Given the measurements $\mathbf{y}$, the signal $\boldsymbol{f}$ is estimated by solving

$$\mathbf{f} = \mathbf{\Psi} \left(\arg\min_{\theta} \left\| \mathbf{y} - \mathbf{H}\mathbf{\Psi}\boldsymbol{\theta} \right\|_2^2 + \tau \left\| \boldsymbol{\theta} \right\|_1 \right) \quad (12)$$

where τ is the regularization constant. Several algorithms have been developed to solve this $l_2 - l_1$ optimization. In this paper the Gradient Projection for Sparse Reconstruction (GPSR) algorithm is used (Figueiredo *et al.*, 2007).

Table 1. Variables summary.

Variable	Size	Description
L	1	Number of spectral bands
N	1	Spatial dimension of the data cube
K	1	Number of CASSI shots
$\mathbf{f}$	$n = N^2 L$	Vectorized form of the $N \times N \times L$ spectral data cube $\mathbf{F}$
$\mathbf{T}^i$	$N \times N$	i^{th} two dimensional coded aperture
$\mathbf{t}^i$	N^2	Vectorized form of $\mathbf{T}^i$
$\mathbf{y}_i$	$V = N(N+L-1)$	i^{th} vectorized CASSI output
$\mathbf{H}_i$	$V \times n$	CASSI matrix representation for the i^{th} shot
$\mathbf{H}$	$KV \times n$	CASSI matrix for the K shots
$\mathbf{A} = \mathbf{H}\mathbf{\Psi}$	$KV \times n$	CASSI Sensing matrix
$\mathbf{A}_s$	$KV \times S$	Matrix with S columns of $\mathbf{A}$ chosen at random $\mathbf{A}$

The RIP in CASSI

The *sth* restricted isometry property (RIP) $\delta_s = \delta_s(\mathbf{A})$ of the matrix $\mathbf{A} \in \mathbb{R}^{KV \times n}$ is defined as the smallest δ such that

$$(1-\delta)\left\| \boldsymbol{\theta} \right\|_2^2 \leq \left\| \mathbf{A}\boldsymbol{\theta} \right\|_2^2 \leq (1+\delta)\left\| \boldsymbol{\theta} \right\|_2^2, \quad (13)$$

for all s-sparse vectors $\boldsymbol{\Theta}$. Notice that equivalently Equation (13) is given by

$$\left| \left\| \mathbf{A}_S \boldsymbol{\theta} \right\|_2^2 - \left\| \boldsymbol{\theta} \right\|_2^2 \right| \leq \delta \left\| \boldsymbol{\theta} \right\|_2^2, \tag{14}$$

for all $S \subset \{1,\ldots,n\}$, $|S| = card(S) \leq s$ and all $\boldsymbol{\theta} \in \mathbb{R}^{|S|}$ and $\mathbf{A}_s$ is a $m \times S$ matrix whose columns are equal to S columns of the matrix $\mathbf{A}$. We then observe that, for $\boldsymbol{\theta} \in \mathbb{R}^{|S|}$

$$\begin{aligned} \left| \left\| \mathbf{A}_S \boldsymbol{\theta} \right\|_2^2 - \left\| \boldsymbol{\theta} \right\|_2^2 \right| &= \left| \langle \mathbf{A}_S \boldsymbol{\theta}, A_S \boldsymbol{\theta} \rangle - \langle \boldsymbol{\theta}, \boldsymbol{\theta} \rangle \right| \\ &= \left| \langle (\mathbf{A}_S^H \mathbf{A}_S - \mathbf{I}_d)\boldsymbol{\theta}, \boldsymbol{\theta} \rangle \right|, \end{aligned} \tag{15}$$

where $\mathbf{A}_S^H$ is the conjugate transpose of $\mathbf{A}_s$ and $\mathbf{I}_d \in \mathbb{R}^{S \times S}$ is the identity matrix. Since the matrix $\mathbf{A}_S^T \mathbf{A}_S - \mathbf{I}_d$ is Hermitian, we have from the Equation (15)

$$\max_{\boldsymbol{\theta} \in \mathbb{R}^S \setminus \{\mathbf{0}\}} \frac{\left| \langle (\mathbf{A}_S^H \mathbf{A}_S - \mathbf{I}_d)\boldsymbol{\theta}, \boldsymbol{\theta} \rangle \right|}{\left\| \boldsymbol{\theta} \right\|_2^2} = \left\| \mathbf{A}_S^H \mathbf{A}_S - \mathbf{I}_d \right\|_{2 \to 2} \leq \delta. \tag{16}$$

Thus, from the definition of the RIP in Equation (13) and Equation (16) we have that the constant δ_s is given by

$$\delta_s = \max_{S \subset [n], |S| \leq s} \left\| \mathbf{A}_S^T \mathbf{A}_S - \mathbf{I}_d \right\|_2^2. \tag{17}$$

Defining the matrix $\mathbf{A}_{TT} = \mathbf{A}_S^T \mathbf{A}_S$, Equation (17) can be rewritten as

$$\delta_s = \max_{S \subset [n], |S| \leq s} \sqrt{\lambda_{\max}(\mathbf{A}_{TT} - \mathbf{I}_d)}, \tag{18}$$

where $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue (Fornaiser, 2010). A small value of δ_S in Equation (18) for a large value of S indicates that the RIP condition is satisfied. The RIP condition implies a stable recovery of the signal $\boldsymbol{\Theta}$ from the projections $\mathbf{A}\boldsymbol{\Theta}$ using a l_j optimization algorithm (Fornaiser, 2010). Let the entries of $\boldsymbol{\Psi}$ be $\boldsymbol{\Psi}_{j,k}$, then using the structure of the matrices $\mathbf{H}_i$ in Equation (9), the entries of $\mathbf{A}_S$ can be written as

$$(A_S)_{jk} = \sum_{r=0}^{L-1} (t^i)_{j-rN} \boldsymbol{\Psi}_{j+r(N'),\Omega_k} \tag{19}$$

for $j = 0,\ldots,m-1$, $k = 0,\ldots,S-1$, where $i = f\ loor(j/V)$, $N' = N^2 - N$, and $\Omega_k \in \{0,\ldots,n-1\}$. Using Equation (19), the entries of $\mathbf{A}_{TT}$ are denoted as $(\mathbf{A}_{TT})_{jk}$ which are expressed as

$$A_{TT})_{jk} = \sum_{i=0}^{K-1} \sum_{\ell=0}^{V-1} \sum_{r=0}^{L-1} \sum_{u=0}^{L-1} (t^i)_{\ell-rN} (t^i)_{\ell-uN} \boldsymbol{\Psi}_{\ell+rN',\Omega_j} \boldsymbol{\Psi}_{\ell+uN',\Omega_j} \tag{20}$$

for $j=0,\ldots,S-1$ and $j \neq k$. When $j=k$ then $(\boldsymbol{A}_{TT})_{jk}=K$. Without loss of generality, we analyze the case when $\boldsymbol{\Psi} = diag(\boldsymbol{\Psi}_{2D},\ldots,\boldsymbol{\Psi}_{2D})$ and $\boldsymbol{\Psi}_{2D}$ is the 2D Symmlet Wavelet Transform. Given that $\boldsymbol{\Psi}$ is a block diagonal matrix, the sums over u and r in Equation (20) have only one nonzero element, then Equation (20) can be expressed as

$$(A_{TT})_{jk} = \sum_{i=0}^{K-1} \sum_{\ell=0}^{V-1} (b^i)_\ell (\boldsymbol{\Psi}_{2D})_{r_\ell,\Omega_j} (\boldsymbol{\Psi}_{2D})_{u_\ell,\Omega_k'} \tag{21}$$

where $r_\ell,\ u_\ell \in \{0,\ldots,N^2-1\}$, and $(b^i)_\ell = (t^i)_{r_\ell} (t^i)_{u_\ell}$ are realizations of a Bernoulli random variable with parameter $p^* = p^2 + (1-p)^2$. Notice that $(\mathbf{x})_\ell = \sum_{i=\rho}^{k-1} (b^i)_\ell$ is a Binomial random variable. More specifically, $(\mathbf{x})_\ell \sim B(K, p^*)$. Notice also that $(\mathbf{y})_\ell = (\boldsymbol{\Psi}_{2D})_{r_\ell,\Omega_j} (\boldsymbol{\Psi}_{2D})_{u_\ell,\Omega_k}$ with Ω_j and Ω_k varying randomly, it can be considered a random variable with $E((\mathbf{y})_\ell) = 0$. Then, Equation (21) can be rewritten as

$$(A_{TT})_{jk} = \sum_{\ell=0}^{V-1} (\mathbf{x})_\ell (\mathbf{y})_\ell \quad j \neq k, \tag{22}$$

and $(A_{TT})_{jk} = K$ for $j = k$. Because $(A_{TT})_{jk}$ for $j \neq k$ is the sum of i.i.d zero mean random variables, and applying the central limit theorem when $V \to \infty$, then $(\mathbf{A}_{TT})_{jk} \to N(0, \sigma^2)$ where $\sigma^2 = V\ Var((\mathbf{x})_\ell (\mathbf{y})_\ell)$. A demonstration of the applicability of the central limit theorem in Equation (22) can be found in Do *et al.* (2012). Figure 5 depicts the matrix $\mathbf{A}_{TT}$ for k=2, N=8, and S=180. It can be seen in this figure that the matrix $\mathbf{A}_{TT}$ is dense and its entries can be seen as a Gaussian random variable. Given that $\mathbf{x}$ and $\mathbf{y}$ in Equation (22) are independent random variables then $V_{xy} = Var((\mathbf{x})_\ell (\mathbf{y})_\ell)$ can be estimated as

$$V_{xy} = (E((\mathbf{x})_\ell))^2 Var((\mathbf{x})_\ell) Var((\mathbf{y})_\ell). \tag{23}$$

Since $(\mathbf{x})_\ell$ is a Binomial random variable, then $Var((\mathbf{x})_\ell) = 4Kp^*(1-p*)$ and $(E((\mathbf{x})_\ell)) = K(2p^* - 1)$. Given the specific structure of $\boldsymbol{\Psi}_{2D}$, the variance $Var((\mathbf{y})_\ell) = 1/N^4$. The variance V_{xy} in Equation (23) can be rewritten as

$$V_{xy} = \frac{K^2(2p^*-1)^2}{N^4} + \frac{4Kp^*(1-p^*)}{N^4}. \tag{24}$$

Given that the entries of the matrices $\mathbf{A}_{TT}$ are Gaussian, then it is possible to apply the concentration inequalities developed by Ledoux to calculate the distribution of λ_{max} in Equation (17) (Candes and Tao, 2005; Ledoux, 2001). More specifically, for each fixed $t > 0$ the concentration inequality is given by

$$P\left(\sigma_{\max}(\mathbf{A}_{TT} - \mathbf{I}_d) > \sqrt{\frac{SV_{xy}}{K^2} + o(1) + t} \right) \leq e^{\frac{-K^2}{2V_{xy}} t^2}, \tag{25}$$

where σ_{max} indicates the maximum singular value and $o(1)$ is a small term tending to zero as $K \to \infty$ which can be estimated as $o(1) = \frac{V_{xy}^{1/3}}{2K^{2/3}} \left(\frac{SV_{xy}}{K^2} \right)^{1/6} \left(1 + \sqrt{\frac{SV_{xy}}{K^2}} \right)^{2/3}$. Using Equation (25) and following a similar approach as that of

Candes in (Candes and Tao, 2005), the RIP in CASSI can be written as

$$P\left(\sigma_{\max}(\mathbf{A}_{TT}-\mathbf{I}_d)>\sqrt{\frac{SV_{xy}}{K^2}+o(1)+t}\right)\leq P_e, \qquad (26)$$

where $Pe=e^{-\frac{K^2}{V_{xy}}t^2/2}e^{-nH(r)}$, $r=S/n$ and $H(r)=-r\log(r)-(1-r)\log(1-r)$. Using Equation (26), the constant δ_s is found as $\delta_s=\left(\sqrt{\frac{SV_{xy}}{K^2}}+o(1)+t\right)^2$.

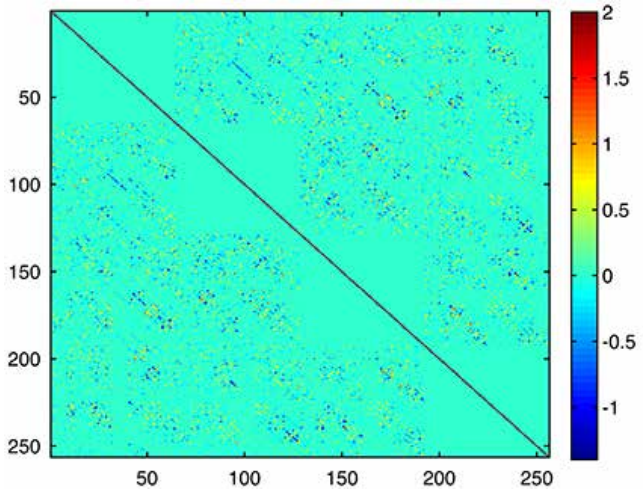

Figure 5. The matrix depicts the matrix $\mathbf{A}_{TT}$ for K=2, N=8, and S=180. The off-diagonal entries of $\mathbf{A}_{TT}$ can be seen as Gaussian entries.

Simulations

In order to verify the RIP in CASSI in Equation (26), an algorithm is proposed based on Monte Carlo Uniform Sampling. Also, two spectral data cubes are used to simulate the CASSI real system reconstructions. All experiments have been carried out in Matlab 2010b on a 3GHz Intel Core i7 and 32GB memory desktop computer.

Algorithm 1. CASSI transmittance analysis.

Input:	***p,K*** ***maxIter*** $\mathbf{\Psi}\in\mathbb{R}^{n\times n}$	**(Transmittance and number of shots)** **(Number of trials)** **(3D Orthonormal basis)**
Output:	**Probability of reconstruction P_r** **RIP Constant δ_s**	
1:	$\mathbf{T}^i\in\{1,-1\}$, with $\sim B(1,p)$ and $i\in\{0,\ldots,K-1\}$	
2:	Create CASSI matrix $\mathbf{H}_i$ from Equation (9) and concatenate as $\mathbf{H}=\left[\mathbf{H}_0^T,\ldots,\mathbf{H}_{k-1}^T\right]^T$	
3:	Create sensing matrix as $\mathbf{A}=(\mathbf{H}\mathbf{\Psi})/K$	
4:	**For** i=1 *to maxIter* **do**	
5:	Generate $S\sim U(1,n)$, $\lvert S\rvert\leq s$	
6:	$C^{(i)}=\sqrt{\lambda_{\max}\left(\mathbf{A}_S^T\mathbf{A}_s-\mathbf{I}_d\right)}$	
7:	**If** $C^{(i)}<1$ **then**	
8:	$count=count+1$	
9:	**end if**	
10:	**end for**	
11:	$\delta_s=\max(c)$	
12:	$Pr=count/maxIter$	

Estimate numerically for δ_s with p and K

To estimate numerically the restricted isometry constant δ_s of order S from matrix $\mathbf{A}$, Monte Carlo simulations are computed by sampling the sub-matrices $\left(\mathbf{A}_S^T\mathbf{A}_s-\mathbf{I}_d\right)$ from Equation (17).

An important parameter of the CASSI system is the transmittance of the coded apertures. Algorithm 1 was developed in order to determine the relationship between the transmittance of the coded apertures and the RIP constant δ_s. The objective of this section is to determine the transmittance that provides the best value of the constant δ_s.

The steps 1-3 in Algorithm 1 are used to construct the CASSI sensing matrix with parameters p and K. The coded apertures $\mathbf{T}^i$ for $i\in\{0,\ldots,K-1\}$ are $N\times N$ symmetric matrices whose entries $\{1,-1\}$ follow a Bernoulli distribution with probability p. The CASSI Sensing matrix $\mathbf{A}$ is formed by multiplying the CASSI matrix $\mathbf{H}$ with the sparsity basis $\mathbf{\Psi}$. Subsequently, an iterative method is implemented to find the maximum eigenvalue of the possible sub-matrices $\left(\mathbf{A}_S^T\mathbf{A}_s-\mathbf{I}_d\right)$.

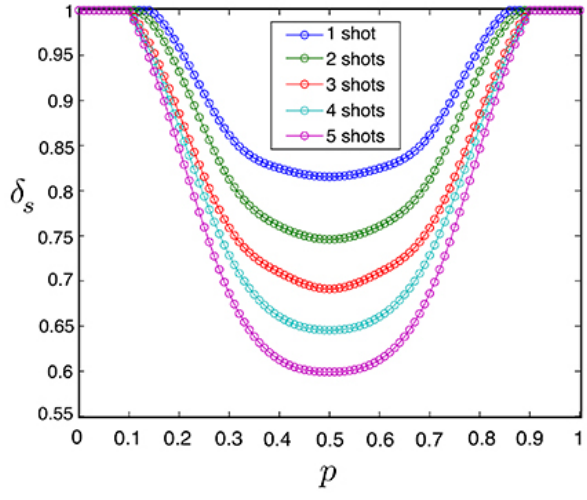

Figure 6. Constant δ_s vs the code aperture transmittance (parameter p), for various number of FPA measurement shots K, and S/n=0,005.

The number of iterations in Algorithm 1 is defined as $maxIter = n!/(n-s)!s!$. Since the number of all possible combinations of the sub-matrices is very large, the constant RIP δ_s can be determined using Monte Carlo simulations. Therefore the index subspace S is sampled by using a uniform distribution (step 5) with size $|S| \leq s$. The eigenvalues in all experiments are stored in the vector C. Finally, the probability of correct reconstruction is calculated from probability distribution of C by determining the number of eigenvalues smaller than 1. In this work, the parameters used in Algorithm 1 were $N=16$, $L=8$, $maxIter=1\,000\,000$ and $\mathbf{\Psi} = \mathbf{\Psi}_{2D} \otimes \mathbf{I}_L$, where $\mathbf{\Psi}_{2D} \in \mathbb{R}^{N^2 \times N^2}$ is the 2D-Wavelet Symmlet 8 basis and $\mathbf{I} \in \mathbb{R}^{L \times L}$ is the identity matrix.

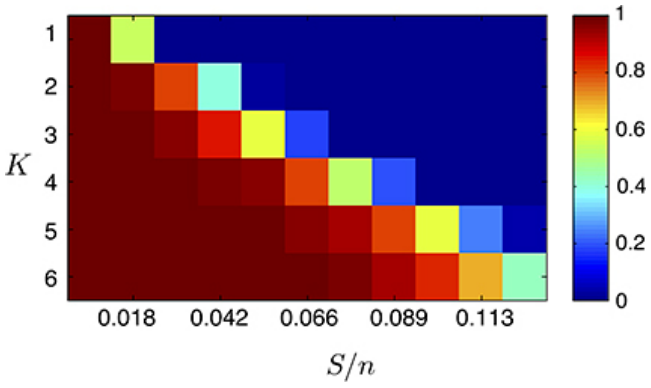

Figure 7. Probability of correct reconstruction $1-P_e$ vs the number of shots K and the sparsity S with transmittance $p=0,5$.

Figure 6 shows the variation of δ_s for different number of shots K, p, transmittances and $S/n=0,005$. The constant δ_s is minimum for $P\{(\mathbf{t})_\ell = 1\} = p = 0,5$, or equivalently when the transmittance of the code apertures is 50 %. Notice that a value of $\delta_s=1$ indicates that reconstruction is not possible.

Similar performance is obtained for other values of S/n. The probability of reconstruction $1-P_e$ (see Equation (26)) is also compared to the results provided by a set of Monte-Carlo simulations. The results are shown in Figure 7, where a high probability depends on small sparsity and a higher number of shots.

Experiments with real data

In this section, the performance of CASSI sensing matrix is evaluated in the reconstruction of three spectral images (Foster *et al.*, 2006). An example of the first database of 533 nm is shown in Figure 8a. This is a natural scene with 256×256 pixels of spatial resolution and $L=33$ spectral bands ranging from 400 nm to 720 nm. An example of the second spectral database at 533 nm is shown in Figure 9a. The second spectral image is a city scene with 256×256 pixels of spatial resolution and $L=33$ spectral bands ranging from 400 nm to 720 nm. The third database is a color scene with 256*256 pixels of spatial resolution and $L=31$ spectral bands ranging from 400 nm to 700 nm at 10 nm steps. In In Figure 10a, an example is shown where the spectral band selected is 500 nm. For all the databases, the reconstruction is compared for different sparsity values using Bernoulli distributed coded apertures in CASSI with transmittance 50 %.

The reconstruction is realized using the GPSR algorithm (Figueiredo et al., 2007). The base representation $\mathbf{\Psi} = diag(\mathbf{\Psi}_{2D}, \ldots, \mathbf{\Psi}_{2D})$ where $\mathbf{\Psi}_{2D}$ is the 2D-Wavelet Symmlet 8 basis. The data cube $\mathbf{F}$ was approximate for different values of the sparsity S/n by removing the $n-S$ less significant coefficients in magnitude from the sparse representation $\boldsymbol{\theta}$. The performance of matrices is analyzed under different sparsity levels. Figure 8d, Figure 9d and Figure10d indicate the performance (PSNR) analysis as a function of the sparsity and number of shots.

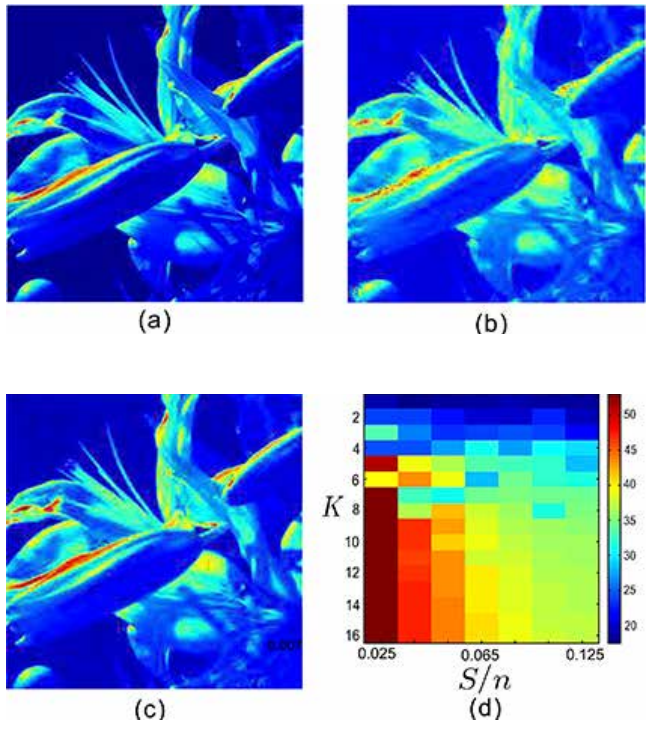

Figure 8. (a) The slice of the spectral band 533 nm. The respective reconstruction is illustrated for $S/n=0,065$ for (b) 4 shots and (c) 9 shots. (d) PSNR as a function of the number of shots and sparsity.

The results of 100 reconstructions experiments given K and S/n are condensed in Figure 8d and Figure 9d (with $S/n=0,065$ in the first image and $S/n=0,125$ in the second and third images), where the PSNR of the reconstruction images verifies that the performance (PSNR) increases as the sparsity and the number of shots K increases.

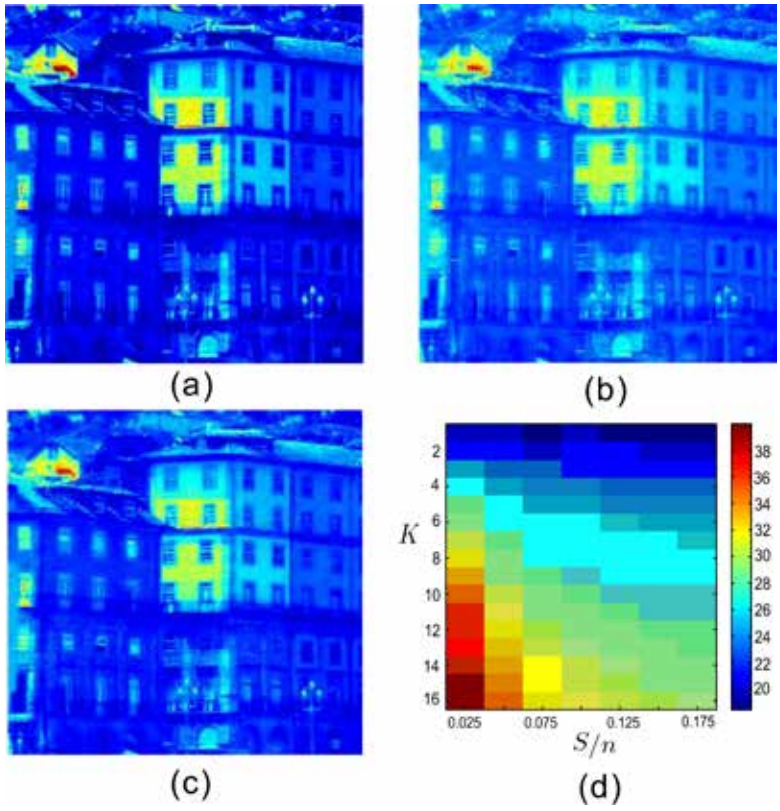

Figure 9. (a) The slice of the spectral band 533 nm. The respective reconstruction is illustrated for S/n = 0,125 for (b) 12 shots and (c) 16 shots. (d) PSNR as a function of the number of shots and sparsity.

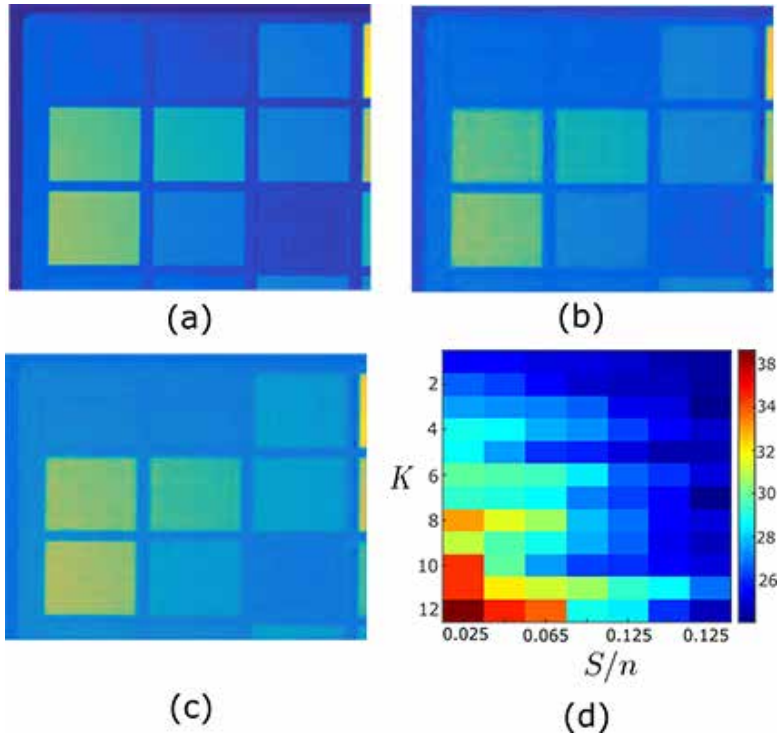

Figure 10. (a) The slice of the spectral band 500nm. The respective reconstruction is illustrated for S/n = 0,125 for (b) 12 shots and (c) 8 shots. (d) PSNR as a function of the number of shots and sparsity.

Minimum number of shots for reconstruction

While the Nyquist Shannon sampling theorem states that a certain minimum number of samples is required for sampling the spectral data cube, compressive sensing can reduce the number of measurements that need to be stored by the prior knowledge that spectral signal is sparse in a known basis and that sensing matrix satisfied the RIP. Figure 8d, Figure 9d and Figure 10d indicate the respective RIP constant δ_s for correct reconstruction. The results indicate clearly that there are a minimum number of shots for correct reconstruction as a function of the sparsity for a noiseless case. The reconstruction of the band depicted in Figure 8a is shown in Figure 8b and Figure 8c for 4 and 9 shots respectively. Figure 8b illustrates the effect of using a number of shots less than that established by the RIP where the PSNR was 25 dBs with sparsity $S/n = 0{,}065$. For Figure 8c the PSNR was 31 dBs with sparsity $S/n = 0{,}065$. Notice that the number of shots used in Figure 8c is higher than the minimum recommended in Figure 8d. The number of shots used in Figure 8d is similar than the minimum recommended in Figure10d. The artifacts in the reconstructed image in Figure 8b, Figure 9b and Figure10c, with low PSNR, are clearly noticeable. On the contrary, Figures 8c, 9c and 10b show high quality in the reconstructed image. For Figure 10b and 10c the reconstruction performance are 28 and 24 dBs respectively, for the given number of shots 12 and 8.

Conclusions

The structured CASSI sensing matrix has been formulated and its RIP has been demonstrated. Simulations verified the theoretical RIP CASSI bounds. For optimal reconstruction, the coded aperture patterns must be designed with $p = 0{,}5$ or equivalent with a transmittance of fifty percent. Given a sparsity condition for signal representation, the minimum number of shots for reconstruction is determined by the variance of the representation basis and the variance of the random structure of the code apertures. An algorithm based on Monte Carlo simulations was developed to determine the optimal transmittance of the coded apertures. The experiments indicate the influence of the sparsity level and the number of shots in the performance of the system.

References

Arce, G. R., Brady, D. J., Carin, L., Arguello, H., & Kittle, D. S. (2014). Compressive coded aperture spectral imaging: An introduction. *Signal Processing Magazine,* Vol 31. IEEE 105–115. IEEE. DOI: 10.1109/MSP.2013.2278763.

Arguello, H., & Arce, G. (2012). Spectrally Selective Compressive Imaging by Matrix System Analysis, Proc. Imaging and Applied Optics Congress. Monterey, California. DOI: 10.1364/COSI.2012.CM4B.5.

Arguello, H., & Arce, G. R. (2011). Code aperture optimization for spectrally agile compressive imaging. *J. Opt. Soc. Am*, Vol 28. A 2400–2413. DOI: 10.1364/josaa.28.002400.

Brady, D. J., Choi, K., Marks, D. L., Horisaki, R., & Lim, S. (2009). Compressive holography. Optics express, Vol 17. 13040–13049. Optical Society of America. DOI: 10.1364/OE.17.013040.

Candes, E., & Tao, T. (2005). Decoding by linear programming, Information Theory. IEEE Transaction, Vol 51. 4203–4215.

Correa-Pugliese, C. (2013). Lapped-windowed reconstructions in compressive spectral imaging. University of Delaware.

Diaz, N., Rueda, H., & Arguello, H. (2015). High-dynamic range compressive spectral imaging by grayscale coded aperture adaptive filtering. Ingeniería E Investigación, Vol 35(3). 53–60. DOI: 10.15446/ing.investig.v35n3.49868.

Do, T. T., Gan, L., Nguyen, N., & Tran, T. (2012). Fast an Efficient Compressive Sensing Using Structurally Random Matrices, Signal Processing. Vol 60. IEEE Transaction on 139–154.

Figueiredo, M., Nowak, R., & Wright, S. (2007). Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. IEEE J. of Selected Topics in Signal Processing, Vol 1. 586–597. DOI: 10.1109/jstsp.2007.910281.

Fornaiser, M. (2010). Theoretical Foundations and Numerical Methods for Sparse Recovery. (M. Fornaiser, Ed.) De Gruyter. DOI: 10.1515/9783110226157.

Foster, D. H., Amano, K., Nascimento, S. M. C., & Foster, M. (2006). J. Frequency of metamerism in natural scenes, Journal of The Optical Society of America A-optics Image Science and Vision. Vol 23. 2359–2372. DOI: 10.1364/JOSAA.23.002359.

Gehm, M. E., Kim, M. S., Fernandez, C., & Brady, D. J. (2008). High-throughput, multiplexed pushbroom hyperspectral microscopy. Optics Express, Vol 16. DOI: 10.1364/OE.16.011032.

Gupta, N. (2008). Acousto-optic tunable filter based spectropolarimetric imagers. DOI: 10.1117/12.782800.

Kittle, D., Choi, K., Wagadarikar, A. A., & Brady, D. J. (2010). Multiframe Image Estimation for Coded Aperture Snapshot Spectral Imagers. Appl. Opt, Vol 49. 6824–6833. DOI: 10.1364/AO.49.006824.

Kittle, D., Choi, K., Wagadarikar, A., & Brady, D. J. (2010). Multiframe image estimation for coded aperture snapshot spectral imagers. Applied optics, Vol 49. 6824–6833. OSA. DOI: 10.1364/AO.49.006824.

Ledoux, M. (2001). The concentration of measure phenomenon. American Mathematical Society.

Studer, V., Bobin, J., Chahid, M., Mousavi, H. S., Candes, E., & Dahan, M. (2012). Compressive fluorescence microscopy for biological and hyperspectral imaging. Proceedings of the National Academy of Sciences, Vol 109. E1679–E1687. National Acad Sciences. DOI: 10.1073/pnas.1119511109.

Oiknine, Y., August, Y., & Stern, A. (2014). Reconstruction algorithms for compressive hyperspectral imaging systems with separable spatial and spectral operators. In *SPIE Optical Engineering Applications* (p. 921703).

Wagadarikar, A. A. (2010). *Compressive Spectral and Coherence Imaging*.

Wagadarikar, A. A., John, R., Willett, R., & Brady, D. (2008). Single disperser design for coded aperture snapshot spectral imaging. Appl. Opt, Vol 47. B44–B51. DOI: 10.1364/AO.47.000B44.

Yu, H., & Wang, G. (2009). Compressed sensing based interior tomography. Physics in medicine and biology, Vol 54. 2791. IOP Publishing. DOI: 10.1088/0031-9155/54/9/014.

Appendix A

Let $\boldsymbol{\Gamma}$ be a coded aperture such that $\left(\mathbf{F}_{j\ell}\right)=1$ for all j,ℓ. Suppose that $\mathbf{y}_{\Gamma}$ is the FPA measurement when the aperture $\boldsymbol{\Gamma}$ is used, $\mathbf{H}'$ be the CASSI matrix associated to the coded $\boldsymbol{\Gamma}$ and thus $y_{\Gamma}=\mathbf{H}'\mathbf{f}$. Let $\overline{y}_l=\mathbf{H}'\mathbf{f}$ be the CASSI output when the coded aperture $T^i\left(j,\ell\right)\in\left\{0,1\right\}$ is used. Define the FPA measurement $\mathbf{y}_i$ as

$$\begin{aligned}\mathbf{y}_i &= 2\overline{\mathbf{y}}_l-\mathbf{y}_{\Gamma}\\ &= 2\overline{\mathbf{H}}^i\mathbf{f}-\mathbf{H}'\mathbf{f}\\ &= \left(2\overline{\mathbf{H}}^i-\mathbf{H}'\right)\mathbf{f}=\mathbf{H}^i\mathbf{f}\end{aligned} \tag{27}$$

where it can be noticed that $\overline{\mathbf{H}}^i\in\left\{-1,0,1\right\}$. Thus, the binary coded apertures can be implemented by subtracting the reference signal $\mathbf{y}_{\Gamma}$ from boolean output $2\overline{\mathbf{y}}_l$. Notice that this type of coded aperture cannot be implemented directly. That means that the implementation of these coded apertures requires one additional FPA CASSI measurement and the noise in the emulated measurements $\mathbf{y}_i$ is 5 times higher than the noise in the original measurements $2\overline{\mathbf{y}}_l$.