Compressive hyperspectral imaging by random separable projections in both spatial and spectral domains

Yitzhak August,¹ Chaim Vachman,¹ Yair Rivenson,² and Adrian Stern¹,*

¹ Department of Electro Optical Engineering, Ben-Gurion University of the Negev, Israel

² Department of Electrical & Computer Engineering, Ben-Gurion University of the Negev, Israel

*Corresponding author: stern@bgu.ac.il

An efficient method and system for compressive sensing of hyperspectral data is presented. Compression efficiency is achieved by randomly encoding both the spatial and spectral domains of the hyperspectral datacube. Separable sensing architecture is used to reduce the computational complexity associated with compressive sensing of large data, which is typical to hyperspectral imaging. The system allows to optimize the ratio between the spatial and the spectral compression sensing ratio.

OCIS codes: 110.4155, 110.4190, 110.4234, 110.1758

1 Introduction

Hyperspectral (HS) images are used in numerous fields such as bio-medical imaging applications, remote sensing, food industry, art conservation and restoration and many more. The amount of data typically captured with HS imaging systems is very large and is often highly compressible. This has motivated the application of compressive sensing techniques for HS imaging.
Compressive Sensing, a.k.a. Compressed Sensing, (CS) [1-3] is a fast emerging field in the area of digital signal sensing and processing. CS theory provides a sensing framework for sampling sparse or compressible signals in a more efficient way that is usually done with Shannon-Nyquist sampling scheme. With CS, a compressed version of the signal is obtained already in the acquisition stage, thus canceling the need for digital compressing. Since CS requires fewer measurements it can be applied to reduce the number of sensors or to reduce the acquisition time. One natural implementation of the CS theory is in the field of imaging.

The first implementation of CS for imaging was the single pixel CS camera [4] shown in Fig. 2, block a. Single pixel CS camera architecture has been used for imaging in the visible, in the Terahertz [5, 6] and SWIR [7] spectrum. The use of single pixel CS camera is suitable in cases where large detector arrays are not available or are too expensive. Another use of the single pixel CS camera is in aerospace remote sensing [8, 9]; in this case the motivation is to reduce the cost of data acquisition. Other compressive imaging techniques include the single shot compressive imaging [10, 11], compressive holography [12, 13], progressive compressive imaging [14], compressive motion tracking [15, 16], and CS applications for microscopy [17-19], to name but a few. An overview of CS techniques in optics may be found in [20]. In this work we focus on using CS method for HS imaging. Hyperspectral and multispectral imaging may benefit from CS as HS data is typically highly compressible.

Hyperspectral data is typically organized in the form of a cube, which is a three dimensional digital array as shown in Fig. 1. The xy plane represents the spatial information and the third dimension is for the spectral reflection as a function of wavelength. Each point in the xy plane has its own spectral signature, described by a spectral vector. The number of spectral bands in the HS image is in the range of dozens to thousands where the typical wavelength width of each
spectral band range from 0.5 nm up to 10 nm with some spectral overlap. The common acquisition techniques for HS data are by spectrometer point scanning and spectrometer line scanning [21, 22]. One of the main limitations of these two methods is the relatively slow scanning process. Other limitations arise from the fact that huge amounts of data needs to be processed and transmitted. Compressive sensing inspired methods can help handling these difficulties. The applicability of CS is based on the fundamental notion that data is sparse or at least compressible, properties that HS data typically possess; different studies show that HS cube is sparse and sometimes extremely sparse [27-23]. If we look at only a narrow spectral window, that is if we look on a $xy$ plane, we have a regular image which is typically compressible in the wavelet domain. On the other hand, if we look in the spectral direction $\lambda$, we generally also find the data to be extremely redundant. For example, the spectral signature of green grass is unique, thus all the vectors in the HSI that preset reflection from the grass have the same spectral signature.

![Hyperspectral cube](image)

**Figure 1.** Hyperspectral cube

In the recent years, several types of compressive sensing systems for HS imaging were proposed [28-33]. In [34] compressive sensing HS cube acquisition is accomplished by a method called Coded Aperture Snapshot Spectral Imagers (CASSI). In the CASSI architecture the spatial information is first randomly encoded and then the spectral information is mixed by a shearing
operation. CASSI is suboptimal in terms of CS because CASSI employs random signal multiplexing only in the $xy$ plane, while in the spectral domain it undergoes deterministic uniform transformation.

Another implementation of the CS system for HS imaging was presented in [32]. This method follows the single pixel CS camera technique and was expanded to three dimensional imaging by replacing the standard detector (single photodiode) in the single pixel CS camera with a spectrometer probe. Fig. 2 (b) shows the way in which the single pixel CS camera is expanded into a HS imaging camera; the photodiode detector in the left box (a) is replaced by the spectrometer shown in the right inset (b). With this architecture, the spatial information is encoded while the spectral information remains unchanged. This mechanism can be considered a parallel spectral acquisition, leaving the spectral dimension unmixed and uncompressed.

![Figure 2](image-url) **Figure 2.** (a) Schematic diagram of single pixel CS camera and its photodiode detector. (b) The expansion to multispectral imaging using a grating and a CCD vector.
In this work we present a new method for HS image acquisition using CS separable encoding both in the spatial and spectral domains. We propose a scheme for all three dimensional multiplexing using two stages of multiplexing; the first stage is spatial multiplexing which is done by using the classical approach of the single pixel camera, followed by the spectral multiplexing stage introduced in Section 4. The spectral encoding is performed in a single step and thus the proposed method requires the same number of projections as in [27] while benefiting of random multiplexing of the wavelength domain too.

2 Compress sensing

In this subsection we briefly review the CS theory. Compressive Sensing (CS) is a technique to recover sparse signals from significantly less measurements than needed with the traditional sampling theory.

A block diagram of a CS system is depicted in Fig. 3.

![Compressive sensing block diagram](image)

**Figure 3.** Compressive sensing block diagram [10]

In this diagram, $f$ represents a physical signal, e.g. objects intensities. $\alpha$ is a vector of components in the sparsifying domain used to represent $f$. $\alpha$ is a mathematical representation vector that contains mainly zeros or near zero values. In the image acquisition step, the signal
vector \( f \) is sampled using the \( \Phi \) operator yielding the measurement vector \( g \). The final step in Fig. 3 is the image reconstruction, accomplished by estimation of \( f \) using a convex optimization 11 type minimization [1-3].

We assume that a \( N \times 1 \) vector \( f \) that is to be measured can be expressed by \( f = \Psi \alpha \), where the \( N \times 1 \) vector \( \alpha \) contains only \( k << N \) non-zero elements and \( \Psi \) is a sparsifying operator. The measurements vector \( g \in \mathbb{R}^{M} \), is obtained by

\[
g = \Phi f
\]

where \( \Phi \in \mathbb{R}^{M \times N} \) is a sensing matrix. By properly choosing \( M \) and \( \Phi \), and assuming sparsity of \( f \) in the \( \Psi \) domain, the signal \( f \) can be recovered from the measurements \( g \). The crucial step here is to build a sensing matrix \( \Phi \) such that it enables accurate recovery of \( N \) sized \( f \) from fewer \( M \) measurements \( g \). Reconstruction of \( f \) from \( g \) is guaranteed if the number of measurements \( M \) meets the following condition [1, 3]:

\[
M \geq C \mu^2 K \log(N)
\]

(2)

It can be seen that the number of measurements required, \( M \), depends on the size of the signal \( N \), its sparsity, \( k \), and \( \mu \) representing the mutual coherence between \( \Phi \) and \( \Psi \). The mutual coherence it is defined by:

\[
\mu(\Phi, \Psi) = \sqrt{M} \max_{1 \leq i < j \leq M} \left| \langle \Phi_i, \Psi_j \rangle \right|
\]

(3)

where \( \Phi_i, \Psi_j \) are vectors of \( \Phi \) and \( \Psi \) respectively. The value of \( \mu \) is in the range of \( 1 \leq \mu \leq \sqrt{N} \). The lower \( \mu \) is, the better the performance of the system and quality of the product. The original signal \( f \) can be recovered by solving the following problem.
\[
\hat{f} = \Psi \hat{a} \quad \text{subject to} \quad \min_a \left\{ \|g - \Phi \Psi a\|_2^2 + \gamma \|a\|_1 \right\}.
\] (4)

where \(\gamma \|a\|_1\) is \(\ell_1\) norm and \(\gamma\) is a regularization weight.

One of the difficulties of using the CS method for HS imaging is the huge size of matrices \(\Phi\) required for representing the sensing operation. Signals in the CS theory are represented by vectors with \(N\) components. The measurements data is \(M\) dimensional, consequently the sensing matrix is of size \(\Phi \in \mathbb{R}^{M \times N}\). Hyperspectral imaging involves 3D signals \(F \in \mathbb{R}^{N_1 \times N_2 \times N_3}\) which can be converted to vectors by lexicographic ordering to \(N\) - length vector \((f = \text{vec}(F))\). Since \(N = N_1 \times N_2 \times N_3\) the sensing matrix size has the order of \((N_1 \times N_2 \times N_3)^2\). For instance let us consider the computational aspects of random encoding a 3D data HS cube of \(F \in \mathbb{R}^{N_1 \times N_2 \times N_3}\), with \(N_1 = N_2 = N_3 = 256\). In this case the sensing matrix \(\Phi\) will be \(\Phi \in \mathbb{R}^{2^{24} \times 2^{24}}\).

Such matrices cannot be handled in standard computational systems because of their challenging storage and memory requirements. The optical implementation and sensor calibration of such systems also present a great challenge because the realization of random \(\Phi\) requires the system to have \(N \times M\) nearly independent modes (degrees of freedom).

### 3 Separable compressive sensing

Separable sensing operators are common in many optical systems (e.g. wave propagation) and often applied in image processing tasks. Separable CS was proposed in [35-37], in order to overcome the practical limitations in CI implementations involving large data, and for the reason that often separable sensing operators arise naturally in multi-dimensional signal processing. As shown in [35-37], a separable system matrix significantly reduces the implementation...
complexity at the expense of some compression efficiency loss, i.e., more samples are required compared to non-separable CS, in order to accurately reconstruct the signal.

A separable sensing operator $\Phi$ can be represented in the form of $\Phi_y \otimes \Phi_x$, where the symbol $\otimes$ denotes the Kronecker product, also referred to as the direct product or the tensor product. If $\Phi_y = [\phi_{y1}, \phi_{y2}, \cdots, \phi_{yn}]$ is an $n \times p$ matrix and $\Phi_x$ an $m \times q$ matrix, then the Kronecker product between $\Phi_y$ and $\Phi_x$ is given:

$$
\Phi_{yx} = \Phi_y \otimes \Phi_x = \begin{bmatrix}
\phi_{y1} \Phi_x & \phi_{y1,2} \Phi_x & \cdots & \phi_{y1,p} \Phi_x \\
\phi_{y2,1} \Phi_x & \phi_{y2,2} \Phi_x & \cdots & \phi_{y2,p} \Phi_x \\
\vdots & \vdots & \ddots & \vdots \\
\phi_{yn,1} \Phi_x & \phi_{yn,2} \Phi_x & \cdots & \phi_{yn,p} \Phi_x 
\end{bmatrix}.
$$

(5)

As we described in the previous subsection, in the case of $n$ dimensional signal, we use the $\text{vec}(\cdot)$ operator in order to create a column vector from a matrix $F$ by stacking the column:

$$
\text{vec}(F) = \begin{bmatrix}
f_1 \\
f_2 \\
\vdots \\
f_n
\end{bmatrix}.
$$

(6)

Let us consider the two dimensional (2D) signal $F = [f_1, f_2, \cdots, f_n]$ and measurement $G = [g_1, g_2, \cdots, g_n]$. $F$ and $G$ are a matrix representation of $f$ and $g$. In such a case (1) can be written in form [35]

$$
\text{vec}(G) = \Phi_{yx} \times \text{vec}(F) = \left( \Phi_y^T \otimes \Phi_x \right) \times \text{vec}(F).
$$

(7)
and using properties of the Kronecker product, we can write:

\[ G = \Phi_y F \Phi_x. \] (8)

Consequently, (4) can be re-written to solve:

\[ \hat{F} = \Psi \hat{A} \Psi^T \quad \text{subject to} \quad \min_\alpha \left\{ \| \text{vec}(G) - \text{vec}(\Phi_y \Psi \Psi^T \Phi_x) \|_2 + \gamma \| \text{vec}(A) \|_1 \right\} \quad \alpha = \text{vec}(A) \] (9)

(9) provides simple way to handle the huge matrix vector multiplication of (4). For example, if the size of \( \Phi_y, \Phi_x, F \) is \( \sim 1000 \times 1000 \) entries, (9) requires operations with matrices of same order, whereas the standard compressive sensing recovery problem (4) involves algebraic manipulations with matrices of the order of \( \sim 10^6 \times 10^6 \).

When considering separable based compressive sensing scheme, it was shown in [35] that the mutual coherence of the separable sensing system is given by:

\[ \mu(\Phi_y, \Psi_x, \Phi_x, \Psi_y) = \mu(\Phi_y \otimes \Phi_x, \Psi_y \otimes \Psi_x) = \mu(\Phi_y, \Psi_y) \mu(\Phi_x, \Psi_x). \] (10)

which can be shown to be larger than the mutual coherence of a non-separable sensing operator. Thus, according to (2), the number of measurements \( M \) required to accurately reconstruct the signal with the separable sensing scheme is larger. For example, if \( \Phi \) is a random orthogonal matrix, uniformly distributed on the unit sphere, it can be shown that:

\[ \frac{\mu(\Phi_y \otimes \Phi_x, \Psi)}{\mu(\Phi, \Psi)} \approx \frac{2 \log_{10}(\sqrt{N})}{\sqrt{2 \log_{10}(N)}} = \sqrt{\frac{1}{2} \log_{10}(N)}. \] (11)
meaning that \( \frac{1}{\sqrt{2 \log_{10}(N)}} \) times more measurements are required in order to accurately reconstruct the signal using separable sensing operator instead of a non-separable random operator [35]. This is a reasonable cost for gaining the computational facilitation. It was numerically demonstrated in [35] that the loss in compression efficiency is quite moderate and practically smaller than the one predicted in (11).

4 Implementation architecture for spatial and separable spectral encoding for hyperspectral compressive sensing

In this section we present an optical implementation scheme that permits both spatial and spectral random encoding. In Sec. 4.1 we provide a description for the spectral encoding method and In Sec. 4.2, we provide the full description of the system architecture for Compressive HS Imaging by Separable Spatial and Spectral operators (CHISSS).

4.1 Spectral encoding

In this subsection we describe the principle of the separable spectrum sensing operation. Fig. 2 provides a schematic description of the spectral encoding principle. In this description the input signal is the optical spatially multiplexed signal available at the detector S3 in Fig. 2.
Figure 4. Schematic diagram of the spectral separable operator.

Figure 4 shows a mechanism which replaces the detectors a or b in Fig. 2. The input signal S3 is the output of the single pixel CS camera presented in Fig. 2, thus it is a spectral vector which we wish to encode and measure using the photo sensor. In Fig. 4, the input optical signal at S3 passes through a diffractive or dispersive element working as a spectral to spatial convertor. A spatial grating can be used to separate the spectral components in the horizontal y direction, thus converting the light spot into a spectral line. The spectral line in Fig. 4 (along the y direction) is spatially encoded using the coded aperture mask C1. Here C1 is a single line of coded apertures. This operation gives each wavelength its own weight, i.e. each wavelength is multiplied by the local coded aperture transition value. In order to focus and collect the different spectral components we use regular converging lenses. In practice we use a parallel process for the spectral encoding with cylindrical lens, this will be explained in the next subsection. The technique described above provides a single randomly encoded measurement of the spectral component. However, for CS we need M measurements that satisfy (2), where each measurement is a result of different encoding of the data cube. Multiple encoding of the spectral vector can be
achieved by time division multiplexing; i.e., by changing the aperture pattern for each measurement of the image sensor. However, this will result in long acquisition time. Alternatively, the various spectral encoding can be achieved by spatial division multiplexing. The system described in the next subsection shows such spatial division multiplexing which is essentially implemented by duplicating the apparatus described in Fig. 4 in the $x$ direction. The spectral information is multiplied by different random codes and captured by a line array of sensors. This way parallel spectral encoded measurement is achieved within one exposure for a give spectral vector. The ability to measure all the spectral projections with a single exposure provides a way to measure HS image in the same number of spatial measurement that is needed for a monochromatic single pixel CS camera [4].

4.2 The system structure

In this section we describe the CHISSS architecture. The architecture implements optical compressive sensing system using separable operators. In contrast to the previous architecture for CS-HS imaging [29, 38, 39], CHISSS architecture provides a way for encoding both the spatial and spectral domains using separate and random operations with the ability to change the compression ratio between the spectral and spatial domains. The CHSISS system uses two separable random encoding codes, one for the spatial domain and the other of the spectral domain. Figure 5 depicts the proposed CHSISS system.
**Figure 5.** Schematic diagram of CHSISS system for CS HS imaging.

The spatial multiplexing process is performed in a way similar to the "single pixel camera" [32]. As in Fig. 2, the lens $L_1$ is used to image the object on the DMD device $D_1$. A random code of size $N_x \times N_y$ is displayed by $L_1$. The encoded light reflected from $D_1$ is then focused on the central point of the $G_1$ grating using the lens $L_2$. At this point, the spot on the $G_1$ plane contains the same mixed spatial information for the entire spectrum. One can view the process up to this stage as a parallel encoding of the spatial data for each wavelength. Therefore each spectral component is a result of the spatial $xy$ multiplexing (provided by the DMD), where each component undergoes the same multiplexing process.

The spectral multiplexing is achieved by applying a second encoding operator separately. The spectral encoder is based on the method described in Fig. 4. By means of the cylindrical lenses $L_3$ and $L_4$ and the coded aperture $C_1$ the spectral encoding process, described in Sec. 4.1, is performed in parallel. Grating $G_1$ splits and diffracts the beam $S_3$ into $N_\lambda$ spectral spots, which
are spread along parallel rays on the coding device $C_1$ by means of the cylindrical lens $L_3$. The coded aperture $C_1$ has a random reflection pattern therefore each horizontal spectral geometrical line is encoded by a different random pattern. The coded aperture size in Fig. 5 has $M_\lambda$ horizontal elements and $N_\lambda$ vertical elements. Next, $N_\lambda$ spectrally encoded components reflected from the vertical lines of $C_1$ are summed by means of the cylindrical lens $L_4$ and collected by the appropriate pixel in a line array sensor. The different spectral modulations are passing through the $L_4$ cylindrical lens in parallel. Note that the encoding process with the CHISSS system in Fig. 5 is separable in $xy$ and $\lambda$ domain. Since the spectral encoding is performed in parallel in a single step (by space-division-multiplexing) the overall acquisition time is determined solely by the spatial encoding. Therefore the CHISSS acquisition time is similar to that of the single pixel CS camera.

## 5 Simulation results

We have simulated the acquisition process with the CHISSS shown in Fig. 5 and investigated the reconstructions. In order to simulate the system we use a computer procedure that implements the appropriate spatial and spectral separable encoding operators. We used real data from a HS camera. The HS image of the Iris painting (Fig. 6 left) was taken indoors using a halogen light source and the parking lot image (Fig. 6 right) was taken outdoors during the day light. Both images were recorded in 256 spectral bands from 500nm to 657nm where the spectral width of each band is about $0.61 \sim 0.62$ nm. The spatial image size was $256 \times 256$ pixels. We used these two HS cubes as objects and sampled them according to the CHISSS system structure shown in Fig. 5.
As we describe in Fig. 4, each HSI cube was first spatially encoded and then spectrally encoded. In the simulation we used three orthogonal random masks, $\Phi_x, \Phi_y, \Phi_\lambda$, to compose the separable sensing operator. Note that with the CHISSS shown in Fig. 5 (as with the systems in Fig.2) the spatial sampling operator $\Phi_{xy}$ does not have to be separable in $x$ and $y$ directions. However, to remedy the computational burden required for compressive sensing and reconstruction of data of size $N = 256^3$, we chose to use spatial masks obtained from a Kroneker product of $\Phi_x$ with $\Phi_y$.

While a non-separable spatial sensing operator $\Phi_{xy}$ is represented by a matrix of the order $256^2 \times 256^2$, the matrices $\Phi_y$ and $\Phi_x$ are of the order $256 \times 256$ and the system forward model is implemented simply by (3). For the recovery process we used the MATLAB$^{R2012a}$ and the TwIST [40] solver procedure. The programs were run on an Intel i7-2600 3.4 GHz processor with 8GB memory. We used the 3D Haar wavelets as the sparsifying operators $\Psi^T$ together with $\ell_1$ regularization according to (9). Reconstructed images from the simulated CHISSSS are shown.

Figure 6. Left: original image of "Iris painting", and (lower image) its reconstruction from 10% samples. Right: original image of "parking lot", and (lower image) its reconstruction from 10% samples.
in Fig. 6 (lower row). These results are for a total compression ratio \( \frac{M_x \times M_y \times M_z}{N_x \times N_y \times N_z} \) of 10% of the original HS data cube. For the Iris painting the spatial domain (x-y) compressive sensing ratio was set to \( \frac{M_x \times M_y}{N_x \times N_y} = \left( \frac{217}{256} \right)^2 \approx 71.2\% \) while the spectral compressive sensing ratio was \( \frac{M_z}{N_z} = \frac{38}{256} \approx 14.8\% \). For the parking lot the HS data compression ratios are

\[
\frac{M_x \times M_y}{N_x \times N_y} = \left( \frac{181}{256} \right)^2 \approx 49.9\% \quad \text{and} \quad \frac{M_z}{N_z} = \frac{51}{256} \approx 19.9\%
\]

respectively. As we can see in Fig. 6, despite the X10 compression the reconstructions are quite similar to the original images. The reconstruction peak signal to noise ratio (PSNR) for the Iris painting was \( \sim 21\text{dB} \) and for the "parking lot" \( \sim 25\text{dB} \).

The dependence of the reconstruction quality on compressive sensing ratio \( M/N \) is demonstrated in Fig. 7. Figure 7 (a) shows an RGB projection of the HSI source and Fig. 7 (b), (c), (d) and (e) the reconstructions from data compressively sensed with ratios 10%, 38%, 5%, and 13%, respectively. As it can be seen, the results are reasonable even with compression as deep as 5%, while at compression ratios larger than 10% the degradation is hardly noticeable.
Since the sparsity of the HS datacube in the spatial dimension is typically different from that in the spectral dimension it is interesting to investigate the dependence of the CHISSS performance on the spatial and spectral compression ratios. Figure 8 shows the PSNR for the parking lot compressively sampled with various spectral and spatial ratios yielding given overall sampling ratios $M/N$. Points with the same color represent the same total compression ratio.
Figure 8. Reconstruction PSNR calculation for Parking lot HS as function of spatial and spectral compression ratios. Points with same color represent the same overall compression ratio. For visualization purposes a surfaces grid was built by bilinear interpolation.

From Fig. 8, it is evident that, as expected, the PSNR increases as a function of the total sensing ratio. In addition we can also see that the reconstruction PSNR increases as the spectral compression contribution to the total compression ratio is bigger. This reflects the well-known fact the HS cubes are more compressible in the spectral dimension [26, 41, 42].
Figure 9. Reconstruction PSNR contours plots the CSHSS of the "Parking lot". Compression conditions A, and B yield same reconstruction PSNR, while the overall compression of point A is 30 and of B is 19%.

Figure 9 shows the reconstruction PSNR contour lines of the interpolated surface in Fig. 8. The contour lines show, from another perspective, the observation achieved from Fig. 8 that the influence of the spectral compression is larger than that of the spatial. For example, in order to achieve a PSNR of 30 dB, one can choose spatial compression of 42% together with a spectral compression of 72%, yielding a total compression 30%. Alternatively, the same PSNR can be achieved with spatial compression of 75% together with a spectral compression of 25%, yielding a total compression of 19%.

6 Conclusion

We presented a technique for HS compressive imaging using separable random projection in all three dimensions of the HS data. The proposed CHSISS architecture can provide both spatial and
spectral random encoding in a relatively simple structure. The spectral multiplexing is done in parallel and only once per single spatial multiplexing, therefore we can acquire a HS cube for the same number of spatial projections. Simulation results demonstrate the need to balance the compression depths in the spatial and spectral domain in order to optimize the CHISSS performance for a given total compression sensing ratio. Because of higher redundancy in the spectral domain more spatial projections are needed than spectral projections.

**Acknowledgment**

Adrian Stern wishes to thank the Israel Science Foundation (grant No.1039/09). The authors wish to thank Iris Tresman (Arts Department, Ben-Gurion University) for providing her paint for HS imaging. We also acknowledge Prof. Ohad Ben-Shahar research group (The interdisciplinary Computational Vision Lab) for providing the hyperspectral camera.

**References**


