COMPRESSIVE HYPERSONTIC IMAGING USING PROGRESSIVE TOTAL VARIATION

Simeon Kamdem Kuiteing, Giulio Coluccia*, Alessandro Barducci†, Mauro Barni‡, Enrico Magli*  
* Dipartimento di Ingegneria dell’Informazione, Università di Siena, Italy  
† Dipartimento di Elettronica e Telecomunicazioni – Politecnico di Torino, Italy  
‡ Consiglio Nazionale delle Ricerche – Istituto di Fisica Applicata “Nello Carrara”, Firenze, Italy

ABSTRACT

Compressed Sensing (CS) is suitable for remote acquisition of hyperspectral images for earth observation, since it could exploit the strong spatial and spectral correlations, allowing to simplify the architecture of the onboard sensors. Solutions proposed so far tend to decouple spatial and spectral dimensions to reduce the complexity of the reconstruction, not taking into account that onboard sensors progressively acquire spectral rows rather than acquiring spectral channels. For this reason, we propose a novel progressive CS architecture based on separate sensing of spectral rows and joint reconstruction employing Total Variation. Experimental results run on raw AVIRIS and AIRS images confirm the validity of the proposed system.

Index Terms— Compressed Sensing, Hyperspectral Imaging, Remote Sensing, Total Variation

1. INTRODUCTION

Compressed Sensing (CS) [1, 2] is a new signal acquisition paradigm, which takes advantage from the feature of many natural signals of being highly correlated. A high correlation entails the existence of a domain (usually defined by an integral transform) in which the signal is sparse, and only a small fraction of the transform coefficients are significantly different from zero. CS addresses the problem of collecting a number of measurements smaller than that required by Shannon theorem, but sufficient to allow the reconstruction of sparse signals with an arbitrarily low error.

A promising application of CS theory regards the remote acquisition of hyperspectral imagery for spaceborne and airborne earth observation. On one hand, hyperspectral images exhibit a strong correlation both in the spatial dimension and - more importantly - in the spectral dimension, so they fit perfectly the assumptions underlying CS theory. On the other hand, the use of CS techniques would allow to design sensors requiring a smaller memory buffer, fewer detectors, and a reduced volume of data to transmit.

The application of CS theory to hyperspectral image acquisition is not straightforward mainly due to the complexity of the reconstruction stage. For a fruitful application of CS, in fact, it is necessary that the redundancy of hyperspectral images in both the spatial and spectral dimension is exploited, in a truly 3D fashion. This implies that the measurement process and, more importantly, the reconstruction process are applied to 3D blocks of the image data cube of sufficient size, thus raising the problem of the computational complexity of the reconstruction step, which can quickly become unmanageable. See for example [3], where the properties of Kronecker products are exploited to cast a multidimensional problem to a single dimension vector whose size is the product of the sizes of each source dimension. Other approaches based on compressive projection principal component analysis [4] have been recently proposed in [5] and [6], as well as a method relying on generalized tensor products [7]. See [8] for an overview of compressed sensing methods applied to hyperspectral imaging.

A problem, common to virtually all proposed solutions to the computational complexity issue (noticeably [9]), is that they work in a 2D + 1D fashion assuming that the two spatial dimensions, hereafter indicated as $x$-$y$, are acquired and processed together, and that the spectral dimension $\lambda$ is used in a second phase to progressively refine the reconstruction obtained from $x$-$y$ data, exploiting the correlation along the spectral dimension. Such an approach, however, does not take into account the way hyperspectral images are acquired by onboard sensors. In most cases, in fact, onboard systems are equipped with a linear array of sensors which, at a given time, acquires a spectral row ($x$ dimension at all wavelengths). The next spectral row is then acquired at the subsequent instant exploiting the motion of the satellite. This acquisition architecture is usually referred to as pushbroom configuration. As a result, the $y$ spatial dimension is essentially a time dimension, making it difficult to process first the images in the $x$-$y$ plane and add the spectral dimension in a second time, since buffering the whole data cube is infeasible. A possible solution would be to apply 2D CS reconstruction to the $x$-$\lambda$ plane and use the $y$ dimension to refine the reconstruction. Unfortunately, previous attempts to do so failed to provide satisfactory results [9].

In this paper we propose a solution to this problem, i.e., we consider the $x$-$\lambda$ plus $y$ configuration and investigate suitable reconstruction algorithms that are able to take advantage of correlations in all image dimensions. In particular, in light of the above issues, we propose a novel CS architecture, based on sensing over $x$-$\lambda$ spectral rows and reconstruction employing Total Variation (TV, [10]) minimization, which is better suited to reconstruct images acquired by a satellite equipped with a sensor working in a pushbroom configuration. Unlike previous approaches, which reconstruct the data cube from a set of separately sensed spectral channels, the proposed algorithm requires separately sensed spectral rows, which is compatible with the structure of pushbroom sensors. We show that the TV prior is effective at capturing the correlation within spectral rows, achieving a reconstruction quality very similar to that obtained by the simpler (but infeasible) conventional approach.

We validate the effectiveness of the proposed approach by testing it on data cubes acquired by the AVIRIS hyperspectral sensor and AIRS ultraspectral sounder. The results we obtain confirm the
validity of our system that, thanks to the possibility of working on larger windows, achieves the same performance in both \(x-y\) plus \(\lambda\) and \(x-\lambda\) plus \(y\) configurations.

2. SATELLITE ONBOARD ARCHITECTURES

With the term imaging spectrometers we refer to instruments able to measure the energy emitted or reflected from an object as a function of two spatial and one spectral coordinate, originating 3D datasets called datacubes. Unfortunately, only few 3D detectors exist, which have coarse spectral resolving power and poor efficiency, and are therefore unsuitable for the realization of spaceborne sensors. Modern imaging spectrometers employ 2D detector arrays, which collect a signal expressed in arbitrary digital units of energy as a function of three indices representing column, row, and exposure [11]. These raw data must be transformed into a standard coordinate/measurement system of at-sensor radiance, cross-track exposure \[11\].

Imaging spectrometers used for earth observation can be categorized into classes based on two main criteria: the technique they adopt to perform spatial sampling and the sensors architecture utilized to induce spectral dispersion/discrimination. Techniques of spatial sampling that have been used in real remote sensing instruments are: whiskbroom, pushbroom, framing, and windowing. Evidently, due to the 3D nature of the signal to be collected, the sampling scheme adopted in the spatial domain is not independent of the sampling scheme utilized for the remaining 1D spectral subset. A whiskbroom scanning instrument employs a 2D spatial field of view (FOV) that scans the object in both the along-track and the cross-track directions. Usually, this FOV covers the entire spectral interval to be sampled; i.e., a 1D detector array is adopted to observe all spectral channels with a single shot. A pushbroom imaging spectrometer scans a 1D FOV in the along-track direction only, covering with a single acquisition the entire spectral range. A framing (also called staring) instrument employs a 2D FOV that remains fixed on the object during acquisition. The term windowing will be used to describe the class of sensors that employ a 2D FOV that scans the target over the along-track direction.

Three main techniques for separating spectral information produced by the observed source can be adopted in an aerospace sensor: wavelength filtering, spectral dispersion, and multiplexing. This last method can be applied into two distinct forms: interferential and dispersive multiplexing. Dispersive instruments use either a prism or a grating to obtain dispersion of the incoming light along a space direction, that is subsequently sampled by a matched detector. The most frequent type of multiplexing spectrometers encountered in aerospace sensors belongs to the class of Fourier-transform spectrometers (FTSs), i.e., two–beam interferometers such as the Michelson, Mach–Zehnder, and Sagnac optical configurations. Multiple-beam interferometers (aka étalons) such as the Fabry–Perot have signal collection abilities that are more similar to those of filtering instruments than of FTSs. The effects of the multiplexing architecture on the Signal–to–Noise (SNR) of the collected signal has been discussed in depth in [12] and [13]. The most common type of imaging spectrometer used for the realization of remote sensing sensors is the pushbroom configuration associated with grating dispersion. Sensors belonging to this class adopt a 2D detector array that samples in a single shot a 2D domain composed by the across-track spatial coordinate on one direction, and the wavelength axis on the other one. A typical example of aerospace sensor having this architecture is the CHRIS–PROBA from the European Space Agency [14]. Therefore, we focus our investigation to assess and compressively sample 2D \(x\lambda\) domains.

![Diagram](image)

Fig. 1. Architecture of an ideal sensor utilizing the CS technology. The sensor modulates (spatial light modulator) the 2D domain output by the imaging spectrometer and focuses (integrates) the modulated domain on the single–point detector.

![Diagram](image)

Fig. 2. Graphical representation of a \([N_x \times N_c \times N_\lambda]\) datacube. A spectral \(x\lambda\) row is highlighted.

2.1. CS Sensor Architecture

The application of the CS technique to remote sensing requires a broadband light modulator that computes random projections of the observed image. It is important that these projections are implemented optically, thus avoiding the acquisition of the entire dataset to digitally perform the random linear combinations. Fig. 1 sketches the conceptual scheme of a CS hyperspectral imager operating in the pushbroom configuration. The direct modulation scheme depicted in Fig. 1 adopts a single element detector, integrating the incoming radiation field as modulated by the Spatial Light Modulator (SLM). This last element is an electro–actuated 2D array of mirrors, crystals, or liquid crystals cells that modulates the available image before the acquisition performed by a single–element detector that integrates the image filtered by the SLM. It must be noticed that the availability of fast detectors and high frame-rate SLMs are critical points for any CS applications. Moreover, it is possible to build up a sensor with a SLM of lower frame rate, provided that a coarse resolution 2D array is utilized in the focal plane for parallelizing the CS of a mosaic of subimages.

3. PROPOSED TECHNIQUE

We propose an acquisition and reconstruction technique based on CS, capable of acquiring separate random projections of each spectral \(x\lambda\) rows (hence complying with the pushbroom sensor configuration described in Section 2) and reconstructing the entire data cube capturing the correlations in both spatial and spectral directions, with manageable complexity.

Referring to Fig. 2, the hyperspectral image \(F \in \mathbb{R}^{N_x \times N_c \times N_\lambda}\) can be represented as a 3D collection of samples, where \(x\) and \(y\) represent spatial dimensions and \(\lambda\) represents the spectral dimension. Hence, \(F\) can be considered as a collection of \(N_\lambda\) spectral rows \(F^i = [F_{1i}, \ldots, F_{Ni}]\), each consisting of a \(N_c \times N_\lambda\) matrix, i.e., \(F^i = [F^i_1, F^i_2, \ldots, F^i_{N_\lambda}]\). We refer to this configuration as \(x\lambda - y\) cube.
3.1. Acquisition
For what concerns the acquisition of the image, it consists in the collection of $y_i \in \mathbb{R}^{M \times 1}$ of $M$ measurements for each spectral row as $y_i = \Phi_i \cdot \text{vec}(F_i)$, where each sensing matrix $\Phi_i \in \mathbb{R}^{M \times N_B}$ is taken as Gaussian i.i.d. and $M < N_B N_C$. For simplicity, $M$ is taken as the same value for all spectral rows. The measurements of all spectral rows are then collected in the matrix $Y$. This setting is amenable to separate spatial reconstruction of each spectral row. However, we expect that separate reconstruction does not yield a sufficiently accurate estimate of the original image, since it lacks modelling of vertical correlation. As a matter of fact, this inaccurate reconstruction will be the initialization point of our proposed reconstruction algorithm of section 3.2. The acquisition procedure is reported in Algorithm 1.

Algorithm 1 Proposed acquisition algorithm

Require: the hyperspectral image $F = [F^1, F^2, \ldots, F^{N_B}]$, the number of measurements per row $M$
Ensure: the measurement matrix $Y$
1: for $i = 1$ to $N_B$ do
2: Draw $\Phi_i \in \mathbb{R}^{M \times N_B}$ s.t. $(\Phi_i)_{k,j} \sim \mathcal{N}(0,1/M)$
3: $(Y)_i \leftarrow \Phi_i \cdot \text{vec}(F^i)$
4: end for
5: return $Y$

3.2. Reconstruction with Iterative Total Variation (ITV)
The idea behind the iterative reconstruction is that if we can obtain a prediction of a spectral row $F^i$, e.g., applying the operator $P(\cdot, \cdot)$ to rows $F^{i-1}$ and $F^{i+1}$ of some initial reconstruction, then we can cancel out the contribution of this predictor from the measurements of $F^i$, and reconstruct only the prediction error instead of the full spectral row. If the prediction filter is accurate, the prediction error is expected to be more compressible than the full signal, and the reconstruction will yield better results [9]. Prediction/reconstruction techniques have also been considered in [15, 16, 17, 18, 19, 20] for different applications. In particular, the iterative procedure starts from the initial reconstruction $F^{(0)}$ of all spectral rows. Even if this initial reconstruction can be obtained using several techniques, in our experiments we reconstruct each spectral row by solving, for each $i = 1, \ldots, N_B$, the following problem

$$F^{(0)}_{i} = \arg \min_{X} TV(X) \text{ s.t. } \Phi_i \cdot \text{vec}(X) = (Y)_i,$$

$$TV(X) = \sum_{k,j} \sqrt{[(X)_{k+1,j} - (X)_{k,j}]^2 + [(X)_{k,j+1} - (X)_{k,j}]^2}.$$ 

Since the TV is the sum of the magnitudes of the discretized gradient, seeking to minimize the TV norm relies on the assumption that the gradient of the spectral row is approximately sparse, hence the TV norm should be small. Then, for every row we first obtain its prediction from upper and lower rows estimated at previous iteration $F_P = P\left(F^{(i-1)}, F^{(i+1)}\right)$. After that, we compute prediction error measurements as $e_y = y_i - \Phi_i \cdot \text{vec}(F_P)$, using the prediction filter of [9, Section III.C.2]. We use $e_y$ to reconstruct the $i$-th row summing the CS reconstruction of $e_y$ to $F_P$

$$F^{(n)}_{i} = F_P + E_P,$$

$$E_P = \arg \min_{E} TV(E) \text{ s.t. } \Phi_i \cdot \text{vec}(E) = e_y,$$

This process is performed on all spectral rows, and is iterated until convergence. The proposed iterative reconstruction scheme is shown in Algorithm 2.

Algorithm 2 ITV reconstruction algorithm

Require: the measurement matrix $Y$, the set of $\Phi^i$
Ensure: the estimation $\hat{F}$
1: for $i = 1$ to $N_B$ do
2: $F^{(0)}_{i} \leftarrow \arg \min_{X} TV(X) \text{ s.t. } \Phi_i \cdot \text{vec}(X) = (Y)_i$
3: end for
4: $n \leftarrow 0$
5: repeat
6: $n \leftarrow n + 1$
7: for $i = 1$ to $N_B$ do
8: $F_P \leftarrow P\left(F^{(n-1)}_{i-1}, F^{(n+1)}_{i+1}\right)$
9: $y_P \leftarrow \Phi_i \cdot \text{vec}(F_P)$
10: $e_y \leftarrow (Y)_i - y_P$
11: $E_P \leftarrow \arg \min_{E} TV(E) \text{ s.t. } \Phi_i \cdot \text{vec}(E) = e_y$
12: $F^{(n)}_{i} \leftarrow F_P + E_P$
13: end for
14: until Convergence is reached
15: return $F^{(n)}$

4. RESULTS

4.1. Dataset description
Our simulations are mainly based on two hyperspectral images selected among those used as reference for onboard lossy compression in the “multispectral and hyperspectral data compression” working group of the Consultative Committee for Space Data Systems (CCSDS). These images are commonly known as granule 9 (gran9) of AIRS and scene 0 (set0) of AVIRIS (Yellowstone). AVIRIS is a spectrometer with 224 bands, and the size of this image is 512 lines and 680 pixels. AIRS is an ultraspectral sounder with 2378 spectral channels, used to create 3D maps of air and surface temperature. In the CCSDS dataset, only 1501 bands are considered. The unstable channels have been removed as they have little or no scientific interest. The spatial size is 90 pixels and 135 lines. These scenes are widely used in literature so the comparisons with other techniques are easier. Both are raw images i.e the output of the detector, without any processing, calibration or denoising applied. Since our objective is to assess the potential of CS to manufacture hyperspectral sensors, employing the raw image is a more realistic approach.

4.2. Experimental results
In order to evaluate the performance of our proposed algorithm, we present results from a set of experiments, given in terms of Mean Square Error (MSE) as a function of the percentage of measurements $M/N$, where $N = N_B N_C$ for the standard $xy - \lambda$ configuration and $N = N_B$ for the $x\lambda - y$ configuration.

For the sake of comparison with our scheme, we report Fig. 3 which is obtained by considering the standard use of the hyperspectral data cube with $xy$ as spatial dimension and $\lambda$ as spectral one. To keep the computational complexity manageable, a $32 \times 32$ spatial crop of the image across all frequency bands was used. Fig. 3 shows the reconstruction of the AIRS scene performed by Kronecker Compressed Sensing (KCS) [3] and those obtained through Kronecker-iterative compressed sampling (KICS) which relies on the iterative procedure described in [9], with the initial point computed by KCS. Note that both KCS and KICS perform signal recovery using the $\ell_1$-norm minimization process. As we can see, KICS provides quite...
In Figs. 4, 5 and 6, we show the experiments with the $x\lambda - y$ cube. The cube is acquired using the procedure of Alg. 1 and reconstructed using ITV (Alg. 2). In this scenario, it is worth mentioning that KICS performed on the $x\lambda - y$ cube does not converge hence, we are constrained to make a comparison between KCS and our ITV algorithm. In order to carry out a fair one, we focused on a small portion of the hypercube, a $32 \times 32$ spectral $x\lambda$ rows and the whole vertical length ($y$). We repeated the experiment for 7 different windows along with their vertical dimension and averaged the MSE values obtained. Results are illustrated in Fig. 4. As can be seen, the 2D $x\lambda$ TV reconstruction yields very large MSE values, which is inappropriate for practical applications. The proposed ITV reconstruction algorithm converges in about 23 iterations and allows to improve significantly the MSE values. ITV outperforms the KCS for $M < 35\%$ and provides similar behavior to the KCS scheme for higher values of $M$. Moreover, the ITV algorithm allows to reduce drastically the computational complexity up to a factor of 12 with respect to the KCS schemes as shown in Tab. 1 (complexity refer to a Matlab–based implementation running on a Windows operating system environment, equipped with Intel® Core™2 Duo CPU T6500 @ 2.1 GHz processor and 4 GB Ram).

As a consequence, ITV could allow to reconstruct larger spatial rows crops along with all their vertical dimension, a task very difficult to achieve with the KCS. Fig. 5 presents averaged results on three different $128 \times 128$ $x\lambda$ windows along with all their vertical dimension. These results show that the bigger the $x\lambda$ window size the better the performance of ITV. On one hand, by using larger windows, ITV allows to improve significantly the MSE values with respect to the TV reconstruction while keeping the computational time at a very low level. For instance, for $M = 30\%$, the TV takes about 40 minutes to reconstruct the entire 3D signal and a single iteration of the ITV reconstruction algorithm around 20 minutes. On other hand, by using a $128 \times 128$ $x\lambda$ window along with all its $y$ vertical dimension, the KCS problem becomes computationally intractable and as a result the comparison with the ITV is impossible. Experiments on the AVIRIS image are shown in Fig. 6 where the results lead to similar observations to those made on the AIRS image.

5. CONCLUSIONS

In this paper, we proposed an architecture for the acquisition and reconstruction of hyperspectral images. The acquisition uses Compressed Sensing to separately acquire spectral rows, in the same way as actual satellite pushbroom sensors operate. The reconstruction relies on the minimization of Total Variation and a progressive refinement based on linear predictors to jointly process the measurements of each spectral row, in order to exploit both spectral and spatial correlation at the same time, with manageable complexity. Experiments run on AVIRIS and AIRS images show that the proposed approach, allowing to work with larger windows due to lower complexity with respect to existing algorithms, achieves similar performance as simpler but infeasible conventional approaches.

### Table 1. Computational time (min.): AIRS image

<table>
<thead>
<tr>
<th>$M$ %</th>
<th>KCS</th>
<th>TV</th>
<th>ITV</th>
<th>ITV</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>30</td>
<td>4</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>30</td>
<td>98</td>
<td>8</td>
<td>14</td>
<td>43</td>
</tr>
<tr>
<td>50</td>
<td>245</td>
<td>12</td>
<td>21</td>
<td>63</td>
</tr>
</tbody>
</table>

As a consequence, ITV could allow to reconstruct larger spatial rows crops along with all their vertical dimension, a task very difficult to achieve with the KCS. Fig. 5 presents averaged results on three different $128 \times 128$ $x\lambda$ windows along with all their vertical dimension. These results show that the bigger the $x\lambda$ window size the better the performance of ITV. On one hand, by using larger windows, ITV allows to improve significantly the MSE values with respect to the TV reconstruction while keeping the computational time at a very low level. For instance, for $M = 30\%$, the TV takes about 40 minutes to reconstruct the entire 3D signal and a single iteration of the ITV reconstruction algorithm around 20 minutes. On other hand, by using a $128 \times 128$ $x\lambda$ window along with all its $y$ vertical dimension, the KCS problem becomes computationally intractable and as a result the comparison with the ITV is impossible. Experiments on the AVIRIS image are shown in Fig. 6 where the results lead to similar observations to those made on the AIRS image.

5. CONCLUSIONS

In this paper, we proposed an architecture for the acquisition and reconstruction of hyperspectral images. The acquisition uses Compressed Sensing to separately acquire spectral rows, in the same way as actual satellite pushbroom sensors operate. The reconstruction relies on the minimization of Total Variation and a progressive refinement based on linear predictors to jointly process the measurements of each spectral row, in order to exploit both spectral and spatial correlation at the same time, with manageable complexity. Experiments run on AVIRIS and AIRS images show that the proposed approach, allowing to work with larger windows due to lower complexity with respect to existing algorithms, achieves similar performance as simpler but infeasible conventional approaches.
6. REFERENCES


