

A Novel Progressive Image Scanning and Reconstruction Scheme based on Compressed Sensing and Linear Prediction

Giulio Coluccia, Enrico Magli
Politecnico di Torino

*Dipartimento di Elettronica e Telecomunicazioni
Torino, Italy*

{giulio.coluccia, enrico.magli}@polito.it

Abstract—Compressed sensing (CS) is an innovative technique allowing to represent signals through a small number of their linear projections. In this paper we address the application of CS to the scenario of progressive acquisition of 2D visual signals in a line-by-line fashion. This is an important setting which encompasses diverse systems such as flatbed scanners and remote sensing imagers. The use of CS in such setting raises the problem of reconstructing a very high number of samples, as are contained in an image, from their linear projections. Conventional reconstruction algorithms, whose complexity is cubic in the number of samples, are computationally intractable. In this paper we develop an iterative reconstruction algorithm that reconstructs an image by iteratively estimating a row, and correlating adjacent rows by means of linear prediction. We develop suitable predictors and test the proposed algorithm in the context of flatbed scanners and remote sensing imaging systems. We show that this approach can significantly improve the results of separate reconstruction of each row, providing very good reconstruction quality with reasonable complexity.

Keywords-Image Scanning, Compressed Sensing, Linear Predictor

I. INTRODUCTION

Compressed Sensing (CS) [1], [2] has recently emerged as an efficient technique for sampling a signal with fewer coefficients than the number dictated by classical Shannon/Nyquist theory. The assumption underlying this approach is that the signal to be sampled is sparse or at least “compressible”, i.e., it must have a concise representation in a convenient basis. In CS, sampling is performed by taking a number of linear projections of the signal onto pseudorandom sequences. Therefore, the acquisition presents appealing properties such as low encoding complexity, since the basis in which the signal is sparse does not need to be computed, and universality, since the sensing is blind to the source distribution. Reconstruction of a signal from its projections can be done e.g. using linear programming [2], with a complexity that is $O(N^3)$, with N the number of samples to be recovered.

Recently, CS has been applied to multi-dimensional signals. In this case, a serious problem arises regarding the computational complexity of the reconstruction process. The conventional approach of measuring the signal along all

dimensions at once leads to very large N , making the reconstruction computationally intractable. The authors of [3] showed a way to recast a multidimensional CS problem to a one-dimensional one, by the means of Kronecker products of sensing and sparsity matrices. Reconstruction algorithms for multidimensional signals have also been proposed in [4], [5], [6] for hyperspectral images and multiview video.

In this paper, we consider the application of CS to devices that acquire 2D visual information through progressive scanning. These devices are equipped with a one-dimensional array of detectors, and a 2D image is obtained via the repeated use of the array over different slices of the 2D object to be imaged. This is a very important scenario, which encompasses many applications. Amongst others, it is worth mentioning at least two examples, which we will focus on in the remainder of this paper. The first is given by *flatbed scanners*, where each line of the image is acquired by a 1D optical sensor moving in the orthogonal direction. The second one is represented by *airborne and spaceborne imagers* of the pushbroom type for remote sensing applications. In this case, the 1D sensor is carried on a flying platform such as an airplane or satellite; the sensor looks down at the Earth, and acquires a line-by-line scan of the underlying scene, while each line is oriented in the across-track direction, and the platform flight moves the sensor from one line to the next one. These applications, as well as several other ones, can clearly benefit from CS. CS-based imaging systems have been demonstrated in [7], where the optical sensor is replaced by an array of micromirrors that acquire linear projections of the signal samples via a single detector (*Single-Pixel Camera*). A similar principle can be applied to progressive scanning, where a 1D micromirror array can be used to directly sense lines in the CS format. In the case of the remote sensing imaging system, CS can lead to a simpler and cheaper system, which uses a single detector and produces a reduced number of sampling. Detectors can be costly in the wavelengths outside the visible spectrum, and the reduced number of samples allows to implement simpler onboard processing systems. For the flatbed scanner, CS would be extremely useful in order to develop a scanner of small size, as the CS sensor needs not be of the same physical size as the document being scanned. Moreover, in both cases, processing and data handling would

be greatly reduced, which is important in order to reduce power consumption in the remote sensing case, and in order to enable application to small-sized low-power devices in the flatbed scanner case.

In this paper we address this scenario, and tackle the reconstruction problem for 2D images acquired line-by-line. In particular, we propose a simple progressive acquisition algorithm for 2D signals, where rows are acquired independently of each other, but the reconstruction is performed jointly over all rows. Joint reconstruction is achieved through an iterative algorithm that correlates different rows through linear prediction, instead of taking a multidimensional transform as sparsity domain. Prediction allows to exploit correlation in both horizontal and vertical dimensions, even if the acquisition is performed in one direction only. The main concept is to exploit correlation along the vertical direction by iteratively predicting each line and reconstructing the prediction error only, which is sparser than the line itself. Results show that few iterations of the proposed algorithm suffice to significantly improve the MSE of the reconstruction, allowing to obtain high-quality reconstruction results with feasible complexity.

II. BACKGROUND

A. Notation and definitions

We denote (column-) vectors and matrices by lowercase and uppercase boldface characters, respectively. The (m, n) -th element of a matrix \mathbf{A} is $(\mathbf{A})_{m,n}$. The m -th row of matrix \mathbf{A} is $(\mathbf{A})_m$. The n -th element of a vector \mathbf{v} is $(\mathbf{v})_n$. The transpose of a matrix \mathbf{A} is \mathbf{A}^T . The stack operator $\text{vec}\{\mathbf{A}\}$ denotes the column vector obtained by stacking the columns of \mathbf{A} on top of each other, from left to right.

The notation $\|\mathbf{v}\|_{\ell_0}$ denotes the number of nonzero elements of vector \mathbf{v} . The notation $\|\mathbf{v}\|_{\ell_1}$ denotes the ℓ_1 -norm of the vector \mathbf{v} and is defined as $\|\mathbf{v}\|_{\ell_1} \triangleq \sum_i |(\mathbf{v})_i|$. The notation $a \sim \mathcal{N}(\mu, \sigma^2)$ means denotes a Gaussian random variable a with mean μ and variance σ^2 .

B. Compressed Sensing

In the standard CS framework, introduced in [8], a signal $\mathbf{x} \in \mathbb{R}^{N \times 1}$ which has a sparse representation in some basis $\Psi \in \mathbb{R}^{N \times N}$, i.e:

$$\mathbf{x} = \Psi\boldsymbol{\theta}, \quad \|\boldsymbol{\theta}\|_{\ell_0} = K, \quad K \ll N$$

can be recovered by a smaller vector $\mathbf{y} \in \mathbb{R}^{M \times 1}$, $K < M < N$, of linear measurements $\mathbf{y} = \Phi\mathbf{x}$, where $\Phi \in \mathbb{R}^{M \times N}$ is the *sensing matrix*. The optimum solution, requiring at least $M = K + 1$ measurements, would be

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_{\ell_0} \quad \text{s.t.} \quad \Phi\Psi\boldsymbol{\theta} = \mathbf{y} .$$

Since the ℓ_0 norm minimization is a NP-hard problem, one can resort to a linear programming reconstruction by minimizing the ℓ_1 norm

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_{\ell_1} \quad \text{s.t.} \quad \Phi\Psi\boldsymbol{\theta} = \mathbf{y} , \quad (1)$$

provided that M is large enough ($\sim K \log(N/K)$).

The same algorithm holds for signals which are not exactly sparse, but rather compressible, meaning that they (or their representation $\boldsymbol{\theta}$ in basis Ψ) can be expressed only by K significant coefficients, while the remaining ones are (close to) zero.

It has been shown in [9] that extracting the elements of Φ at random from a Gaussian or Rademacher distribution (i.e., ± 1 with the same probability), and, in general, from any Sub-Gaussian distribution, allows a correct reconstruction with overwhelming probability.

III. PROPOSED ALGORITHM

According to typical progressive scanning approaches, like the ones used by commercial flatbed scanners or by remote sensing systems acquiring environmental pictures, an image is acquired by sensing N_{COL} pixels of each row in a progressive fashion, until N_{ROW} rows are acquired. Hence, the acquired image will result as a matrix of pixels of size $N_{\text{ROW}} \times N_{\text{COL}}$, which will be compressed (and, accordingly, decoded) using a conventional technique. This process requires the acquisition (and processing) of $N_{\text{ROW}}N_{\text{COL}}$ pixels. When N_{ROW} and N_{COL} are large, processing of this huge amount of data may represent an issue, especially when dealing with low cost or low complexity devices.

For this reason, we propose a very simple acquisition scheme, based on CS linear measurements taken on each row, without any further processing. This reduces the amount of data to be acquired and processed. The reconstruction algorithm relies on linear predictors in order to improve the quality of CS reconstruction, by correlating the measurements of adjacent rows in order to exploit their statistical dependencies during the reconstruction stage, largely improving over individual separate reconstruction. The stronger the correlation of pixels within a row and among rows, the better will be the performance of CS reconstruction and of the linear predictor (and hence of the whole reconstruction algorithm).

A. Image Acquisition

The image acquisition algorithm we propose, labelled as Algorithm 1, is very simple and consists in taking linear measurements of each row of the image in a progressive fashion. To minimize the risks of failures in the reconstruction side, a different sensing matrix Φ is drawn for each row.

The image to be measured can be divided into N_{ROW} rows. For each row, M linear measurements are taken, where $M < N_{\text{COL}}$ and N_{COL} is the desired vertical resolution.

In summary, the scene we wish to acquire is represented by the matrix $\mathbf{X} \in \mathbb{R}^{N_{\text{ROW}} \times N_{\text{COL}}}$. For each row of \mathbf{X} , we draw a matrix $\Phi^i \in \mathbb{R}^{M \times N_{\text{COL}}}$ whose elements are Gaussian i.i.d. such that $(\Phi^i)_{kj} \sim \mathcal{N}(0, 1/M)$, with $k = 1, \dots, M$ and $j = 1, \dots, N_{\text{COL}}$. Then, we take M linear measurements of $(\mathbf{X})_i$

Algorithm 1 Proposed acquisition algorithm

Require: the image \mathbf{X} , M

Ensure: the measurement matrix \mathbf{Y}

- 1: **for** $i = 1$ to N_{ROW} **do**
 - 2: Draw Φ^i s.t. $(\Phi^i)_{kj} \sim \mathcal{N}(0, 1/M)$
 - 3: $(\mathbf{Y})_i^\top \leftarrow \Phi^i(\mathbf{X})_i^\top$
 - 4: **end for**
 - 5: **return** \mathbf{Y}
-

which will form the rows of the matrix of measurements $\mathbf{Y} \in \mathbb{R}^{N_{\text{ROW}} \times M}$, namely

$$(\mathbf{Y})_i^\top = \Phi^i(\mathbf{X})_i^\top$$

A more complex algorithm, based on Compressed Sensing and able to capture spatial correlation in both directions (horizontal and vertical), could acquire in a single shot the whole image in a single measurement vector of length M' .

$$\mathbf{y}' = \Phi' \text{vec}\{\mathbf{X}\},$$

where $\text{vec}\{\mathbf{X}\} \in \mathbb{R}^{N_{\text{ROW}}N_{\text{COL}} \times 1}$, $\Phi' \in \mathbb{R}^{M' \times N_{\text{ROW}}N_{\text{COL}}}$, $\mathbf{y}' \in \mathbb{R}^{M' \times 1}$.

Even if this algorithm performed better than the one proposed here since the reconstruction would optimally exploit the correlation in 2 dimensions through a 2D transform matrix, it would require the solution of (1) for a vector of length $N = N_{\text{ROW}}N_{\text{COL}}$. For realistic values of N_{ROW} and N_{COL} , the solution of (1) would be impossible to perform in reasonable time. On the other hand, the proposed approach splits the problem into smaller (and hence tractable) subproblems. However, in doing so, it does not neglect the spatial correlation in vertical direction, which is modeled and employed in the reconstruction process through the use of linear predictors.

B. Image Reconstruction

A trivial reconstruction algorithm based on the acquisition scheme described in section III-A would simply apply the ℓ_1 reconstruction (1) to recover separately each line of \mathbf{X} given the corresponding Φ^i and $(\mathbf{Y})_i$.

Instead, we propose an algorithm using this trivial reconstruction as the initialization step and iteratively improves the current estimate of \mathbf{X} by modelling statistical dependencies between adjacent lines. We label this Algorithm 2. We count the iterations using the index n . The estimation of \mathbf{X} at iteration n is denoted with $\mathbf{X}^{(n)}$.

In particular, the algorithm evaluates a first image reconstruction performing line-by-line separate reconstruction (iteration $n = 0$). Then, the iterations start. The intuition is as follows. For each row, if we are able to reliably predict it using the reconstruction of the upper and lower lines at previous iteration with some linear predictor $P(\cdot, \cdot)$, obtaining \mathbf{x}_P , we can compute the “measurement” \mathbf{y}_P of this prediction by applying matrix Φ^i to \mathbf{x}_P . Then we

Algorithm 2 Proposed reconstruction algorithm

Require: the measurement matrix \mathbf{Y} , the set of Φ^i

Ensure: the estimation $\hat{\mathbf{X}}$

- 1: $n \leftarrow 0$
 - 2: **for** $i = 1$ to N_{ROW} **do**
 - 3: $\hat{\theta} \leftarrow \arg \min_{\theta} \|\theta\|_{\ell_1}$ s.t. $\Phi^i \Psi \theta = (\mathbf{Y})_i^\top$
 - 4: $(\mathbf{X}^{(n)})_i^\top \leftarrow \Psi \hat{\theta}$
 - 5: **end for**
 - 6: **repeat**
 - 7: $n \leftarrow n + 1$
 - 8: **for** $i = 1$ to N_{ROW} **do**
 - 9: **if** $i = 1$ or $i = N_{\text{ROW}}$ **then**
 - 10: $\mathbf{x}_P \leftarrow (\mathbf{X}^{(n-1)})_i^\top$
 - 11: **else**
 - 12: $\mathbf{x}_P \leftarrow P\left((\mathbf{X}^{(n-1)})_{i-1}, (\mathbf{X}^{(n-1)})_{i+1}\right)^\top$
 - 13: **end if**
 - 14: $\mathbf{y}_P \leftarrow \Phi^i \mathbf{x}_P$
 - 15: $\mathbf{e}_Y \leftarrow (\mathbf{Y})_i^\top - \mathbf{y}_P$
 - 16: $\mathbf{e}_\theta \leftarrow \arg \min_{\mathbf{e}} \|\mathbf{e}\|_{\ell_1}$ s.t. $\Phi^i \Psi \mathbf{e} = \mathbf{e}_Y$
 - 17: $\mathbf{e}_X \leftarrow \Psi \mathbf{e}_\theta$
 - 18: $(\mathbf{X}^{(n)})_i^\top \leftarrow (\mathbf{x}_P + \mathbf{e}_X)^\top$
 - 19: **end for**
 - 20: **until** Convergence is reached
 - 21: **return** $\mathbf{X}^{(n)}$
-

calculate the prediction error in the linear measurement domain \mathbf{e}_Y by subtracting this “predicted measurement” from the original measurement row $(\mathbf{Y})_i$. The error \mathbf{e}_Y will be then reconstructed using (1), leading to a prediction error on the signal samples equal to \mathbf{e}_X . Adding \mathbf{e}_X to \mathbf{x}_P provides a new estimate of \mathbf{x} . Since the new estimate is more accurate than the old one, the process can be repeated by estimating a new, more accurate predictor. If the prediction of the row is accurate enough, the prediction error is going to be sparser than the original vector. As a consequence, for an equal number of measurements, the ℓ_1 reconstruction will yield lower MSE.

In section IV-A, we test the performance of several linear predictors $P(\cdot, \cdot)$ and of the overall algorithm. Since (1) is a convex problem and the predictors we test are linear, the overall algorithm can be considered as a projection onto convex sets. This ensures the convergence of the algorithm to the intersection of the constraint sets (if any) [10].

Remark 3.1: We briefly explain here the complexity reduction obtained using Algorithm 2 instead of the standard CS reconstruction algorithm, processing the 2D signal as a whole. For an $N_{\text{ROW}} \times N_{\text{COL}}$ image, the standard CS reconstruction algorithm has an $O(N_{\text{ROW}}^3 N_{\text{COL}}^3)$ complexity. Our algorithm performing N_{ITER} iterations has an $O(N_{\text{ITER}} N_{\text{ROW}} N_{\text{COL}}^3)$ complexity, with, usually, $N_{\text{ITER}} \ll N_{\text{ROW}}, N_{\text{COL}}$.

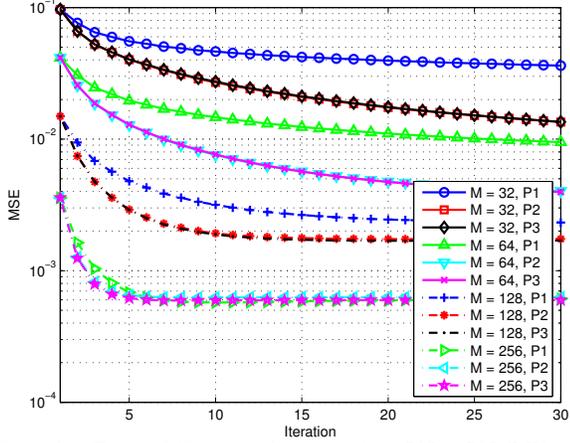


Figure 1. Test of different predictors on *lena* 512×512 B/W image

IV. NUMERICAL RESULTS

A. Choice of the Predictor

First, we start by seeking the linear predictor $P((\mathbf{X})_{i-1}, (\mathbf{X})_{i+1})$ providing fastest convergence and best MSE performance. For this test, we use the standard *lena* black and white image of size 512×512 . M takes the values of 32, 64, 128, 256 and the transform matrix Ψ is the DCT matrix. We denote as \mathbf{x}_P the result of the prediction.

Predictor labelled as *P1* estimates the current¹ line to be predicted as the average of the upper and lower lines:

$$\mathbf{x}_P = \frac{1}{2} ((\mathbf{X})_{i-1} + (\mathbf{X})_{i+1})^T$$

Predictor labelled as *P2* predicts each pixel of current line as the average of adjacent pixels of upper and lower lines

$$(\mathbf{x}_P)_j = \frac{1}{6} [(\mathbf{X})_{i-1, j-i} + (\mathbf{X})_{i-1, j} + (\mathbf{X})_{i-1, j+1} + (\mathbf{X})_{i+1, j-i} + (\mathbf{X})_{i+1, j} + (\mathbf{X})_{i+1, j+1}] .$$

Finally, predictor labelled as *P3* predicts each pixel of current line as the *weighted* average of adjacent pixels of upper and lower lines. Weights depend on the distance from the pixel to be predicted, namely

$$(\mathbf{x}_P)_j = [a(\mathbf{X})_{i-1, j-1} + b(\mathbf{X})_{i-1, j} + a(\mathbf{X})_{i-1, j+1} + a(\mathbf{X})_{i+1, j-1} + b(\mathbf{X})_{i+1, j} + a(\mathbf{X})_{i+1, j+1}] ,$$

with $a = \frac{2-\sqrt{2}}{4}$ and $b = \frac{\sqrt{2}-1}{2}$.

Fig. 1 shows the MSE performance of the overall system for different values of M and using the predictors described above. Results show that the convergence is reached for each value of M . The bigger M , the faster the convergence and the smaller is the MSE at convergence. In any case, it can be noticed that the best performance is obtained for each value of M using predictor labelled as *P3*, i.e. the weighted average. Hence, we will use this predictor in our further tests, omitting to mention it from now on.

¹Here and in the following equations, we omit the index (n) denoting current iteration

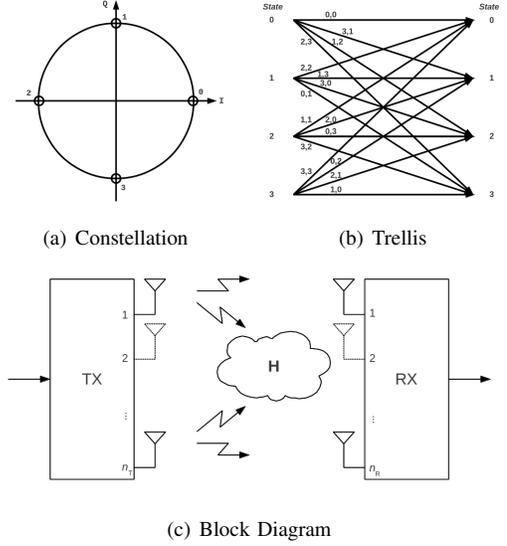


Figure 2. The graphics used as test image for *flatbed scanner* scenario

For $M = 64$, the MSE obtained with separately recovered lines is $4.16 \cdot 10^{-2}$. After 30 iterations, an MSE of $3.96 \cdot 10^{-3}$ is obtained, with a gain of 10.2 dB. The convergence in this case is quite slow, but the MSE is decreased as much as one order of magnitude. Faster convergence is obtained with $M = 128$, as after 15 MSE is decreased from $1.49 \cdot 10^{-2}$ to $1.72 \cdot 10^{-3}$, with a gain of 9.38 dB. Finally, with $M = 256$ the MSE decreases from $3.59 \cdot 10^{-3}$ to $6.18 \cdot 10^{-4}$ in 5 iterations only, with a gain of 7.64 dB.

B. Flatbed Scanner

In this section, we apply our algorithm to images suitable to a *flatbed scanner* scenario. These are black and white graphics and text, and are depicted in Fig. 2. M takes the values of 8, 16, 32, 64, 128, 256 and (where possible) 512. Since they all have a completely white background (representing paper), they can be considered sparse in the pixel domain. Hence, the matrix Ψ is the identity matrix of size N_{COL} , namely $\mathbf{I}_{N_{COL}}$.

Fig. 2(a) is the simplest graphic, representing a QPSK constellation. Fig. 2(b) represents a slightly more complicated (hence, less sparse) graphic, the trellis of a convolutional code. Fig. 2(c) is a larger figure representing a generic block diagram. Finally, fig. 2(d) depicts a sample of generic text.

Table I
MSE AND CONVERGENCE RESULTS ON SAMPLE GRAPHICS

M	init. MSE	conv. MSE	gain (dB)	steps
<i>Constellation</i> (680×576)				
64	$2.08 \cdot 10^{-2}$	$7.86 \cdot 10^{-3}$	4.23	18
128	$9.99 \cdot 10^{-3}$	$3.05 \cdot 10^{-3}$	5.15	14
256	$4.72 \cdot 10^{-3}$	$7.25 \cdot 10^{-4}$	8.14	10
<i>Trellis</i> (680×576)				
64	$8.56 \cdot 10^{-2}$	$3.98 \cdot 10^{-2}$	3.33	18
128	$7.45 \cdot 10^{-2}$	$2.00 \cdot 10^{-2}$	5.71	13
256	$3.39 \cdot 10^{-2}$	$6.51 \cdot 10^{-3}$	7.17	8
<i>Block Diagram</i> (529×1123)				
64	$8.38 \cdot 10^{-3}$	$7.02 \cdot 10^{-3}$	0.77	7
128	$5.79 \cdot 10^{-3}$	$3.97 \cdot 10^{-3}$	1.64	6
256	$2.79 \cdot 10^{-3}$	$1.66 \cdot 10^{-3}$	2.25	5
512	$1.23 \cdot 10^{-3}$	$4.71 \cdot 10^{-4}$	4.17	5
<i>Sample Text</i> (512×512)				
64	$6.40 \cdot 10^{-2}$	$4.59 \cdot 10^{-2}$	1.44	10
128	$5.46 \cdot 10^{-2}$	$2.99 \cdot 10^{-2}$	2.62	7
256	$3.39 \cdot 10^{-2}$	$1.41 \cdot 10^{-2}$	3.81	4

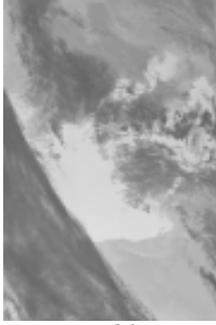


Figure 3. The AIRS sensor gran 9 hyperspectral image, 600-th band

Table I reports the results obtained using the proposed algorithm. The table shows, for each image, the initial MSE (obtained using separate CS reconstruction of each line), the MSE the algorithm converges to, the performance gain, and the number of iterations necessary to reach convergence. Figures confirm the results obtained in the previous section. The more measurements are taken, the faster is the convergence and the lower is the MSE that can be obtained when the algorithm has converged. When the picture is very sparse, it is possible to obtain a reduction of one order of magnitude, while when the picture is less sparse the contribution of Compressed Sensing is weaker, but still a reduction of about 50% in MSE can be obtained.

C. Remote Sensing Image Acquisition

To test the performance of the proposed scheme in a *remote sensing* scenario, we use a spectral band extracted from hyperspectral image “granule 9” of the AIRS sensor. AIRS is an ultraspectral sounder with 2378 spectral channels, used to create 3D maps of air and surface temperature. The spatial size is $N_{\text{COL}} = 90$ pixels and $N_{\text{ROW}} = 135$ lines. The dataset consists in the raw output of the detector, without any processing, calibration or denoising applied. We choose the 600-th band, which is depicted in Fig. 3, but very similar results have been obtained with other bands and are omitted for brevity. M takes the values of 8, 16, 32 e 64. The sparsity

Table II
MSE AND CONVERGENCE RESULTS ON AIRS SENSOR IMAGE

M	init. MSE	conv. MSE	gain (dB)	steps
Basic algorithm				
8	$2.40 \cdot 10^{-1}$	$2.63 \cdot 10^{-2}$ (20-th it.)	9.6	20+
16	$9.57 \cdot 10^{-2}$	$5.21 \cdot 10^{-3}$ (20-th it.)	12.6	20+
32	$2.17 \cdot 10^{-2}$	$1.56 \cdot 10^{-3}$	11.4	10
64	$2.89 \cdot 10^{-3}$	$3.40 \cdot 10^{-4}$	9.29	4
Kronecker improved algorithm				
8	$4.60 \cdot 10^{-3}$	$3.80 \cdot 10^{-3}$	0.83	7
16	$2.62 \cdot 10^{-3}$	$2.02 \cdot 10^{-3}$	1.13	5
32	$1.22 \cdot 10^{-3}$	$9.73 \cdot 10^{-4}$	0.98	3
64	$2.95 \cdot 10^{-4}$	$2.64 \cdot 10^{-4}$	0.48	1

basis Ψ is the DCT.

Table II (Basic Algorithm) summarizes the results obtained applying the proposed algorithm to the 600-th band of the test image. Results show that with $M = 8$ and $M = 16$ the convergence is very slow and is not reached after 20 iterations. On the other hand, when $M = 32$ the convergence is obtained after 10 iterations (reducing from $2.17 \cdot 10^{-2}$ to $1.56 \cdot 10^{-3}$, with a gain of 11.4 dB), while taking $M = 64$ measurements per row implies the convergence after 4 steps only (with MSE reduction from $2.89 \cdot 10^{-3}$ to $3.40 \cdot 10^{-4}$ and a gain of 9.29 dB).

D. Improving performance with Kronecker CS

An improvement to the performance of the algorithm is obtained plugging the so-called *Kronecker Compressed Sensing* (KCS), described in [3], into the algorithm we propose.

KCS allows the acquisition of a multidimensional signal using linear measurement and reconstruction of the entire signal, allowing to exploit sparsity entailed in each signal dimension.

Considering a 2D \mathbf{X} , it is easy to show that, if

$$\mathbf{X} = \Psi_{\text{ROW}}^T \Theta \Psi_{\text{COL}},$$

where $\Theta \in \mathbb{R}^{N_{\text{ROW}} \times N_{\text{COL}}}$ is the matrix collecting the separable 2D transform coefficients and $\Psi_{\text{ROW}} \in \mathbb{R}^{N_{\text{ROW}} \times N_{\text{ROW}}}$ and $\Psi_{\text{COL}} \in \mathbb{R}^{N_{\text{COL}} \times N_{\text{COL}}}$ are the matrices representing the basis applied to rows and columns, respectively, then

$$\text{vec} \left\{ \mathbf{X}^T \right\} = (\Psi_{\text{ROW}}^T \otimes \Psi_{\text{COL}}^T) \text{vec} \left\{ \Theta^T \right\}.$$

Hence, each multidimensional signal can be reshaped as a column vector (in this case, we consider transposed images because we measure image rows), acquired with a “global” sensing matrix Φ' and the reconstruction problem (1) can be recast to multiple dimensions using Kronecker products.

Two KCS sampling techniques are proposed in [3]. The first is similar to the one described in the very last paragraphs of section III-A, where the sensing matrix Φ' is dense. As already stated, the drawback of this approach is the infeasible reconstruction complexity in realistic conditions. The second, instead, allows the acquisition of separate

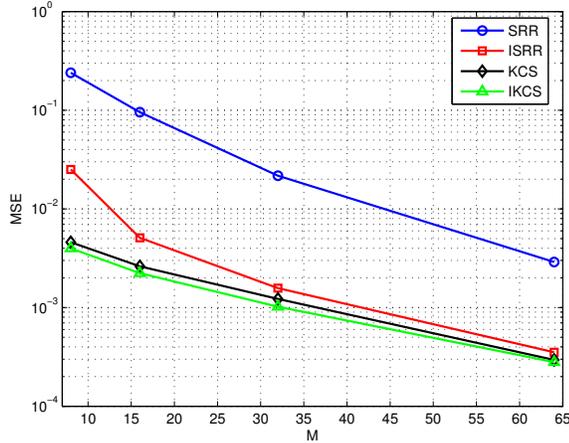


Figure 4. Performance comparison of proposed algorithms vs. M

measurements of each row, but joint reconstruction through the use of the 2D transform $\Psi_{\text{ROW}}^T \otimes \Psi_{\text{COL}}^T$. This means that Φ' is a *block diagonal* $N_{\text{ROW}}M \times N_{\text{ROW}}N_{\text{COL}}$ matrix, where each of the N_{ROW} blocks is the $M \times N_{\text{COL}}$ matrix Φ^i sensing each row. The particular structure of Φ' allows to solve (1) with reasonable complexity.

Hence, we use KCS to initialize the iterative algorithm we propose in this paper (instead of separate linewise reconstruction) and apply it to the *remote sensing* scenario. The performance of this modified version of the algorithm are reported in Table. II (Kronecker improved algorithm). The figures show two effects. First, the initial MSE is much lower than in the separate reconstruction case. This gain can be noticed in particular when M is small and is due to the better performance of KCS reconstruction with respect to separate reconstruction; second, the iterative algorithm slightly improves the overall performance and converges in very few steps. This is due to the fact that KCS captures also correlation in vertical direction, making the contribution of each iteration less effective.

Figure 4 summarizes the best MSE performance obtained by Separate Row Reconstruction (SRR), our Iterative algorithm initialized with Separate Row Reconstruction (ISRR), the Kronecker Compressed Sensing (KCS) and our Iterative algorithm with KCS initialization (IKCS) vs. the number of measurements M . Best performing algorithms are the ones implementing KCS. Plain KCS shows a gain of 7.37 dB over ISSR when $M = 8$, and 1.10 dB when $M = 32$. When using IKCS, roughly 1 dB of additional gain can be obtained with very few iterations.

Finally, as a term of comparison, we report here the MSE performance of a simple reconstruction algorithm named *Orthogonal Matching Pursuit* [11], whose complexity is linear in the number of samples of the original signal ($N_{\text{ROW}}N_{\text{COL}}$ in this case). We acquire and reconstruct the entire image as a whole using $M = 32 \cdot N_{\text{ROW}}$ and $M = 64 \cdot N_{\text{ROW}}$ measurement, to be compared with the performance of our algorithm with $M = 32$ and $M = 64$, respectively. For $M = 32 \cdot N_{\text{ROW}}$, we obtain an MSE of $1.9 \cdot 10^{-3}$, while for

$M = 64 \cdot N_{\text{ROW}}$ we obtain an MSE of $1.7 \cdot 10^{-3}$. Hence, our algorithm with $M = 32$ performs 3 dB better than OMP with the same total amount of measurements, while with $M = 64$ the gain is 8 dB.

V. CONCLUSIONS

In this paper, we proposed a simple and effective algorithm to acquire a 2D signal entailing correlation in both horizontal and vertical directions, like an image. The acquisition is performed by separately acquire each line of the image, taking a number of random linear projections smaller than the number of pixels composing the row itself, as in the Compressed Sensing paradigm. The reconstruction process consists of an iterative algorithm based on the linear prediction of a line and the CS reconstruction of the prediction error, which is supposed to be sparser than the original vector.

We apply this algorithm to two scenarios: flatbed scanners and remote sensing applications. We show that applying our algorithm to images typical of these scenarios it is possible to improve in few iterations the quality obtained reconstructing each row separately.

REFERENCES

- [1] E. Candes, "Compressive sampling," in *Proceedings of the International Congress of Mathematicians*, 2006, vol. 1.
- [2] D.L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [3] M.F. Duarte and R.G. Baraniuk, "Kronecker product matrices for compressive sensing," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 3650–3653.
- [4] A. Abrardo, M. Barni, CM Carretti, S.K. Kamdem, and E. Magli, "A compressive sampling scheme for iterative hyperspectral image reconstruction," in *19th European Signal Processing Conference (EUSIPCO 2011)*, 2011, pp. 1120–1124.
- [5] M. Trocan, T. Maugey, E.W. Tramel, J.E. Fowler, and B. Pesquet-Popescu, "Multistage compressed-sensing reconstruction of multiview images," in *Multimedia Signal Processing (MMSP), 2010 IEEE International Workshop on*. IEEE, 2010, pp. 111–115.
- [6] M.B. Wakin, "A manifold lifting algorithm for multi-view compressive imaging," in *Picture Coding Symposium, 2009. PCS 2009*. IEEE, 2009, pp. 1–4.
- [7] M.B. Wakin, J.N. Laska, M. Duarte, D. Baron, S. Sarvotham, D. Takhar, K.F. Kelly, and R.G. Baraniuk, "An architecture for compressive imaging," in *International Conference on Image Processing*, 2006, pp. 1273–1276.
- [8] E.J. Candes and T. Tao, "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?," *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.
- [9] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.
- [10] P.L. Combettes, "The foundations of set theoretic estimation," *Proceedings of the IEEE*, vol. 81, no. 2, pp. 182–208, 1993.
- [11] J.A. Tropp and A.C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655–4666, 2007.