ADAPTIVE STRUCTURED RECOVERY OF COMPRESSIVE SENSING VIA PIECEWISE AUTOREGRESSIVE MODELING

Xiaolin Wu and Xiangjun Zhang

Dept. of Electrical & Computer Engineering, McMaster University, Hamilton, Canada L8S 4K1

Abstract- In compressive sensing (CS) a challenge is to find a space in which the signal is sparse and hence recoverable faithfully and efficiently. Given the nonstationarity of many natural signals such as images, the sparse space varies in time/spatial domain. As such, CS recovery should be conducted in locally adaptive, signal-dependent spaces to counter the fact that the CS measurements are global and irrespective of signal structures. On the contrary most CS methods seek for a fixed set of bases (e.g., wavelets, DCT, and gradient spaces) for the entirety of a signal. To rectify this problem we propose a new framework for model-guided adaptive recovery of compressive sensing (MARX), and show how a piecewise autoregressive model can be integrated into the MARX framework to adapt to changing second order statistics of a signal in CS recovery. In addition, MARX offers a powerful mechanism of characterizing and exploiting structured sparsities of a signal, greatly restricting the CS solution space. A case study on CS-acquired images shows that the proposed MARX technique can increase the reconstruction quality by up to 8 dB over existing methods.

Index terms: compressive sensing, adaptive modeling, autoregressive process, inverse problem.

1. INTRODUCTION

The recent development of compressive sensing (CS) theory [1, 2] has stirred quite an amount of excitement in signal processing community. The CS theory reveals, in a pleasant surprise, the possibility of reconstructing a signal from a small number of random measurements, if the signal has a sparse representation in some space Ψ. A signal $\mathbf{f} = \{f_n\}_{n=1}^N$ of length N is said to be sparse in space Ψ of bases $\{\psi_n\}_{1\leq n\leq N}$, if the transform coefficients $\langle \mathbf{f}, \psi_n\rangle$, $1 \leq n \leq N$, are mostly zero. The sparsity of $\boldsymbol{f}$ in Ψ is quantified by the number of non-zero coefficients K. The signal can be perfectly recovered from $M = O(K \log(N/K))$ observations with high probability. Given M CS measurements $\mathbf{y} = \Phi\mathbf{f}$, with Φ producing the random projections, the CS recovery of $\mathbf{f}$ from $\mathbf{y}$ is posed as the following constrained optimization problem:

$$\min_{\mathbf{f}} \parallel \Psi^T \mathbf{f} \parallel_{l_1} \quad \text{subject to } \mathbf{y} = \Phi\mathbf{f}, \tag{1}$$

The ℓ_1 minimization problem of (1) can be solved by linear programming [3]. Few other CS recovery algorithms were recently proposed: gradient projection sparse reconstruction [4], matching pursuit [5], and iterative thresholding [6].

A much celebrated property of CS is its ability to compactly encode a signal $\mathbf{f}$ in total blindness of any structures of $\mathbf{f}$, i.e., the same random projections Φ can be performed on all signals, regardless of the differences in their characteristics. However, this does not mean that one can escape from the issue of adapting or optimizing the CS recovery process to the specific signal $\mathbf{f}$ on hand. Indeed, a thorny issue in practice is what space Ψ should be chosen to recover a particular $\mathbf{f}$. For example, in signal compression, while conventional methods strive to encode $\mathbf{f}$ in a transform domain that achieves maximum energy packing of $\mathbf{f}$, CS methods need to recover $\mathbf{f}$ from $\Phi\mathbf{f}$ in a space Ψ in which $\mathbf{f}$ exhibits a high degree of sparsity. Thus, from a system point of view, CS merely transfers the task of signal-dependent code optimization from the encoder to the decoder. Finding a sparse space Ψ for optimal CS recovery of the signal $\mathbf{f}$ poses as much, if not more, a challenge as finding an adaptive transform to completely decorrelate $\mathbf{f}$. This is attested by so far disappointing performance of CS-based compression methods, despite the enthusiasm to change the prevailing practice of ”oversampling followed by massive dumping” in signal acquisition-compression by CS.

The poor rate-distortion performance of current CS recovery techniques relative to conventional coding techniques is rooted in a fatal flaw of the problem formulation (1) for CS recovery. A natural signal $\mathbf{f}$ is typically nonstationary, and there exists no space Ψ in which all segments of $\mathbf{f}$ exhibit sparsity. The problem is particularly acute for images. For a nonstationary 2D $N_r \times N_c$ image signal $f(x, y) \in \mathbb{N}^{N_r\times N_c}$, in two different areas $\mathcal{A}_i$ and $\mathcal{A}_j$ of the spatial domain, subimages $f_i(x, y)$ and $f_j(x, y)$ can have very different waveforms (e.g., smooth shade vs. strong edge), and hence they are sparse in different spaces Ψ_i and Ψ_j. Thus, performing CS recovery in a fixed space Ψ, such as that of DCT, a wavelet, or total variation, can and do fail in parts of the image. This critique leaves us no choice but seeking for a locally adaptive strategy to recover CS-acquired images.

In this work we propose a new framework of Model-based Adaptive Recovery of Compressive Sensing (MARX) to rectify the flaw in the current CS recovery problem formulation (1). The defining feature of MARX, which distinguishes it from other CS recovery techniques, is a locally adaptive sparse signal representation facilitated by a piecewise autoregressive (PAR) model $\mathbf{f} = \mathbf{A}\mathbf{f} + \mathbf{v}$. For images $\mathbf{f}$ is the vector by stacking all $N = N_r N_c$ pixels of an image $f(x, y) \in \mathbb{N}^{N_r \times M_c}$, and $\mathbf{A} \in \mathbb{R}^{N \times N}$ is a square matrix with all elements on the main diagonal being zero. The term $\mathbf{v} \in \mathbb{R}^N$ is a random vector that is the excitation of the 2D autoregressive process. The n^{th} row vector of $\mathbf{A}$ is denoted by $\mathbf{a}_n$, $a_{nn} \equiv 0$, $1 \leq n \leq N$. Since an image is a random Markov field (RMF) of a modest order, each row vector $\mathbf{a}_n$ is sparse, i.e., only a very small portion of the elements of $\mathbf{a}_n$ are nonzero. The nonzero elements of $\mathbf{a}_n$ constitute the 2D support of the regression relation $f_n = \mathbf{a}_n \mathbf{f} + v$ for pixel f_n. The spatial configuration and the order of the regression support for f_n are given by the image waveform at the pixel location n.

Respecting the fact that a natural image is a nonstationary RMF, MARX allows the PAR model parameters $\mathbf{a}_n$ to vary in n. As such, via its parameter matrix $\mathbf{A}$, the PAR model offers a sparse and yet adaptive representation of image signal $\mathbf{f}$. Therefore, the CS recovery of $\mathbf{f}$ can be formulated as the following problem of ℓ_1 minimization:

$$\min_{\mathbf{A},\mathbf{f}} \sum_{1 \leq n \leq N} \|\mathbf{a}_n\|_{\ell_1} \ \text{ subject to } \ \mathbf{y} = \Phi \mathbf{A}\mathbf{f}. \tag{2}$$

We stress the contrast between the variability of $\mathbf{A}$ in (2) and the predetermined bases of Ψ in (1). The MARX sparsity mechanism can fit image local structures (e.g., edges, textures, smooth shades, etc.) much better than wavelet, curvelet, DCT or whatever predetermined bases of Ψ.

Granted, the proposed MARX objective function is computationally more complex than that in the current CS problem formulation. The former involves joint estimation of the image $\mathbf{f}$ and its underlying PAR model $\mathbf{A}$, rather than estimating $\mathbf{f}$ in a fixed space Ψ. The added search space of $\mathbf{A}$ makes the inverse problem of CS recovery severely underdetermined. In following sections we will develop algorithm techniques to overcome this difficulty, making the MARX solution feasible and robust. In pursuing maximum confinement of the solution space for (2), we exploit structured sparsities due to fractal behavior (self similarities) of natural images. We show how the PAR model can be made a convenient machinery to incorporate the structured sparsities into the framework of MARX. The resulting technique not only makes the MARX process computationally tractable but also greatly improves the performance of existing CS recovery algorithms. Very recently, Baraniuk *et al.* demonstrated theoretically the possibility of recovering a signal using a substantially smaller number of measurements than $M = O(K \log(N/K))$ by leveraging more realistic signal models [7]. Our work offers a fresh and successful example to corroborate the above theory. Extensive experiments on a broad class of natural images, ranging from conventional photographs to biomedical images, establish the superior recovery quality of MARX over other CS methods. The gap in performance can be as much as $1 \sim 8$ dB, with the advantage of MARX being the most prominent in the recovery of local structural information.

The remainder of the paper is presented as follows. The detail of the MARX algorithm is described in Section 2. Section 3 develops the MARX algorithm with pattern classification. Experimental results are reported in Section 4.

2. MARX ALGORITHM BASED ON STRUCTURED SPARSITY

The power of adaptive PAR model $\mathbf{A}$ lies in its capability of providing, by varying $\mathbf{a}_n$, different sparse representations for image waveforms in different spatial locations. But one should fully use statistical knowledge of natural images to structure $\mathbf{a}_n$'s in ways to confine the solution space of the underdetermined inverse problem (2). In the CS constraint $\mathbf{y} = \Phi(\mathbf{a}_1\mathbf{f}, \mathbf{a}_2\mathbf{f}, \cdots, \mathbf{a}_N\mathbf{f})^T$, let us examine the PAR prediction of pixel f_n: $\hat{f}_n = \mathbf{a}_n \mathbf{f}$, $a_{n,n} \equiv 0$. For natural images are random Markov field, f_n only depends on pixels in a local window $\mathbf{w}_n = (f_n, f_{n\circlearrowright 1}, f_{n\circlearrowright 2}, \cdots, f_{n\circlearrowright K})$. The subscript $n \circlearrowright k$ denotes the k^{th} neighbor of pixel f_n in the image domain. The ordering of the potentially K effecting pixels $f_{n\circlearrowright 1}, f_{n\circlearrowright 2}, \cdots, f_{n\circlearrowright K}$ is given by a fixed 2D traversal (e.g., raster scan) with respect to pixel location n. Therefore, we set to zero those coefficients in $\mathbf{a}_n$ that correspond to spatial locations outside of the local window $\mathbf{w}_n$, inducing a structure of sparsity in $\mathbf{A}$ that is intrinsic to the physical problem.

A more important and beneficial structured sparsity of $\mathbf{A}$ presents itself if one considers fractal property of natural images. Namely, an image $\mathbf{f}$ often exhibits a similar waveform structure, i.e., having similar second order statistics, in different localities and scales. To characterize the 2D image waveform and noting that the PAR parameters $\mathbf{a}_i$ and $\mathbf{a}_j$ are the same if $\mathbf{w}_i = \mathbf{w}_j + c\mathbf{1}$, we associate each pixel f_n with a mean-removed feature vector

$$\tilde{\mathbf{w}}_n = \mathbf{w}_n - \mu_n \mathbf{1}, \quad \mu_n = \frac{1}{K+1} \sum_{k=0}^{K} f_{n-k}. \tag{3}$$

Observing that f_i and f_j are generated by an underlying PAR model of the same or similar parameters if the two corresponding feature vectors $\tilde{\mathbf{w}}_i$ and $\tilde{\mathbf{w}}_j$ are close to each other, we can strengthen the MARX formulation of (2) to

$$\begin{aligned} &\min_{\mathbf{A},\mathbf{f}} \sum_{1 \leq n \leq N} \|\mathbf{a}_n\|_{\ell_1} \\ &\quad \text{subject to } |\mathbf{a}_n \mathbf{f} - f_{n_i}| \leq \kappa \|\tilde{\mathbf{w}}_n - \tilde{\mathbf{w}}_{n_i}\| \\ &\qquad\qquad 1 \leq n \leq N, \ n_i \in \{\imath \mid \|\tilde{\mathbf{w}}_\imath - \tilde{\mathbf{w}}_n\| < \tau\} \\ &\qquad\qquad \mathbf{y} = \Phi \mathbf{A}\mathbf{f} \end{aligned} \tag{4}$$

where κ is a normalizing factor, and τ is a threshold to select samples having similar 2D waveform as $\tilde{\mathbf{w}}_n$ to learn $\mathbf{a}_n$.

Even with added constraints of structured sparsity, solving (4) directly is still difficult and numerically less stable. We propose an iterative EM approach of constructing the PAR model $\mathbf{A}$ and recovering $\mathbf{f}$ from $\mathbf{y}$ alternatingly in the MARX framework. Specifically, we estimate the PAR model parameters in $\mathbf{A}$ with respect to an initial estimated image $\mathbf{f}^{(0)}$, i.e., solving the constrained ℓ_1 minimization problem (4) given $\mathbf{f} = \mathbf{f}^{(0)}$. The resulting estimated PAR model $\mathbf{A}^{(0)}$ is then used to improve $\mathbf{f}^{(0)}$ to $\mathbf{f}^{(1)}$, and in turn $\mathbf{f}^{(1)}$ used to improve $\mathbf{A}^{(0)}$ to $\mathbf{A}^{(1)}$, and so forth. In iteration t, given $\mathbf{A}^{(t)}$ computed by solving (4), the next estimated image $\mathbf{f}^{(t+1)}$ is obtained by solving the following optimization problem

$$\min_{\mathbf{f}^{(t+1)}} \left\{ \sum_{1\le n\le N} \|\mathbf{f}^{(t+1)} - \mathbf{A}^{(t)}\mathbf{f}^{(t+1)}\| \right\} \quad \text{subject to } \boldsymbol{y} = \Phi\mathbf{f}^{(t+1)} \tag{5}$$

The initial estimated image $\mathbf{f}^{(0)}$ can be obtained by a non-adaptive CS recovery technique, for instance, the total variation (TV) method [4]

$$\min \sum_n |\nabla f_n^{(0)}| \quad \text{subject to } \boldsymbol{y} = \Phi\mathbf{f}, \tag{6}$$

where ∇ is the 2D Laplacian operator. Other CS recovery methods can also be used to produce $\mathbf{f}^{(0)}$, such as wavelet-based method, DCT-based method, matching pursuit, etc.

The impact of the estimation errors $\mathbf{e}_t = \mathbf{f} - \mathbf{f}^{(t)}$ on the solution of (4), $t = 0, 1, \cdots$, is greatly reduced by imposing the CS constraints on $\mathbf{f}^{(t)}$ in (4), because error terms in $\mathbf{e}_t$ are approximately i.i.d. and have zero mean in practice. Indeed, it follows from $\mathbf{A}\mathbf{e}_t \approx \mathbf{0}$ that

$$\begin{aligned} \mathbf{y} &= \Phi\mathbf{A}\mathbf{f}^{(t)} \\ &= \Phi\mathbf{A}\mathbf{f} - \Phi\mathbf{A}\mathbf{e}_t \\ &\approx \Phi\mathbf{A}\mathbf{f}, \end{aligned} \tag{7}$$

meaning that the use of $\mathbf{y} = \Phi\mathbf{A}\mathbf{f}^{(t)}$ has almost the same effect as though the true $\mathbf{f}$ was known.

As to the termination criterion for the EM-based MARX algorithm, we monitor the successive model fit errors $\delta_t = \|\mathbf{f}^{(t+1)} - \mathbf{A}^{(t)}\mathbf{f}^{(t+1)}\|$, which are byproduct of solving (5). The MARX algorithm terminates when $\delta_{t-1} - \delta_t < \epsilon$.

In (4) the N sets of PAR parameters (all N row vectors of $\mathbf{A}$) are estimated jointly. The complexity of the MARX algorithm can be greatly reduced without materially affecting the performance by estimating $\mathbf{a}_n^{(t)}$ individually, $n = 1, 2, \cdots, n$, in iteration t of the EM estimation process. Estimating $\mathbf{A}$ one row vector at a time reduces the number of unknowns by N folds. Denote by $\tilde{\mathbf{w}}_n^{(t)} = (f_n^{(t)}, f_{n\circlearrowleft 1}^{(t)}, \cdots, f_{n\circlearrowleft K}^{(t)}) - \mu_n\mathbf{1}$ the mean-removed feature vector of the t^{th} estimated pixel $f_n^{(t)}$, and let

$$\mathbf{A}_n^{(t)} = \begin{pmatrix} \mathbf{a}_1^{(t-1)} \\ \mathbf{a}_2^{(t-1)} \\ \vdots \\ \mathbf{a}_n^{(t)} \\ \vdots \\ \mathbf{a}_N^{(t-1)} \end{pmatrix} \tag{8}$$

Then, the original sparse estimation problem of (4) can be broken into N ℓ_1 minimization problems of much smaller size, and solved approximately:

$$\begin{aligned} &\min_{\mathbf{a}_n^{(t)}} \|\mathbf{a}_n^{(t)}\|_{\ell_1} \\ &\text{subject to } |\mathbf{a}_n^{(t)}\mathbf{f} - f_i^{(t)}| \le \kappa\|\tilde{\mathbf{w}}_n^{(t)} - \tilde{\mathbf{w}}_i^{(t)}\|, \\ &\qquad i \in \{\imath \mid \|\tilde{\mathbf{w}}_i^{(t)} - \tilde{\mathbf{w}}_n^{(t)}\| < \tau\} \\ &\qquad \mathbf{y} = \Phi\mathbf{A}_n^{(t)}\mathbf{f}^{(t)} \\ &\qquad n = 1, 2, \cdots, N \end{aligned} \tag{9}$$

3. MARX WITH PATTERN CLASSIFICATION

In the previous section the MARX algorithm employs N distinct PAR models, one for each pixel f_n, $1 \le n \le N$. The intent is to allow the maximum degree of freedom in modeling. However, for natural images, the second order statistics may change spatially but the change is smooth. Furthermore, the change may be periodic so that a waveform can repeat itself in different locations. Therefore, we can induce an even stronger structure in the spare matrix $\mathbf{A}$ by classifying the set of feature vectors $\{\tilde{\mathbf{w}}_n\}_{1\le n\le N}$ into C representative waveforms, and describe the image $\mathbf{f}$ in the MARX framework by $C \ll N$ PAR models, one per classified waveform. The classification is performed by a vector quantizer $Q : \mathbb{R}^{K+1} \to \{1, 2, \cdots, C\}$. In the vector quantization the difference metric $D : \mathbb{R}^{(K+1)\times(K+1)} \to [0, \infty)$ between two mean-removed feature vectors $\tilde{\mathbf{w}}_i$ and $\tilde{\mathbf{w}}_j$ is defined as

$$D = \eta\|\tilde{\mathbf{w}}_i, \tilde{\mathbf{w}}_j\|_{\ell_2} + (1-\eta)d(i,j) \tag{10}$$

where $\|\tilde{\mathbf{w}}_i, \tilde{\mathbf{w}}_j\|_{\ell_2}$ is the ℓ_2 distance between two waveforms $\tilde{\mathbf{w}}_i$ and $\tilde{\mathbf{w}}_j$, and $d(i,j)$ is a geometric distance between pixels i and j. The affine weight η is used to adjust the relative importance of the two distance components.

Let $q_n = Q(\tilde{\mathbf{w}}_n) \in \{1, 2, \cdots, C\}$ be the index of the feature class that $\tilde{\mathbf{w}}_n$ falls into, $\hat{\mathbf{a}}_c$ be the model parameter vector for class c. Define the permutation π_n on the elements of $\hat{\mathbf{a}}_c$ such that $f_n = \pi_n(\hat{\mathbf{a}}_c)\mathbf{f} + v$. Now the coefficient matrix of the C PAR models becomes

$$\hat{\mathbf{A}} = \begin{pmatrix} \pi_1(\hat{\mathbf{a}}_{q_1}) \\ \pi_2(\hat{\mathbf{a}}_{q_2}) \\ \vdots \\ \pi_n(\hat{\mathbf{a}}_{q_n}) \\ \vdots \\ \pi_N(\hat{\mathbf{a}}_{q_N}) \end{pmatrix}, \tag{11}$$

in which there are only $C \ll N$ distinct row vectors up to the permutation as the same model $\hat{\mathbf{a}}_c$ acts on all pixels $n \in \{\imath \mid Q(\tilde{\mathbf{w}}_\imath) = c\}$. The feature classifier Q imposes strong signal-dependent structures in the sparse model parameter matrix $\mathbf{A}$,

Table 1. The PSNR (dB) results for different CS recovery methods.

Image	Method	Number of measurements			
		10 %	20 %	30 %	40 %
Vessels	TV	21.81	26.23	29.96	33.18
	MARX	26.10	32.59	36.30	39.06
MRA-Brain	TV	20.21	23.55	26.40	29.15
	MARX	22.31	26.54	29.49	31.87
Micrograph image 1	TV	20.08	22.66	24.84	26.76
	MARX	21.69	25.86	28.62	30.67
Micrograph image 2	TV	19.12	21.15	23.12	24.98
	MARX	20.42	23.87	26.53	28.67
Barb part (256×256)	TV	19.45	20.14	21.03	22.19
	MARX	19.46	23.65	27.87	30.20
Boats	TV	25.68	28.89	31.46	33.74
	MARX	27.70	31.82	34.67	36.90
Monarch	TV	24.36	28.82	32.03	34.84
	MARX	27.05	31.40	34.28	36.79

and reduces the MARX objective function to

$$
\begin{aligned}
&\min_{\hat{\mathbf{a}}_1,\hat{\mathbf{a}}_2,\cdots,\hat{\mathbf{a}}_C,\mathbf{f}} \sum_{1\le c\le C} \|\hat{\mathbf{a}}_c\|_{\ell_1} \\
&\text{subject to } |\pi_n(\hat{\mathbf{a}}_{q_n})\mathbf{f} - f_{n_i}| \le \kappa\|\tilde{\mathbf{w}}_n - \tilde{\mathbf{w}}_{n_i}\| \\
&\qquad\qquad 1 \le n \le N,\ n_i \in \{\imath \mid \|\tilde{\mathbf{w}}_\imath - \tilde{\mathbf{w}}_n\| < \tau\} \\
&\qquad\qquad \mathbf{y} = \Phi\hat{\mathbf{A}}\mathbf{f}
\end{aligned} \tag{12}
$$

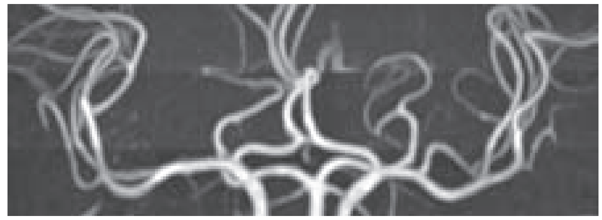
The EM-based MARX algorithm developed in the previous section can be extended to solve the estimation problem of (12), by replacing $\mathbf{A}^{(t)}$ with $\hat{\mathbf{A}}^{(t)}$.

4. EMPIRICAL RESULTS AND REMARKS

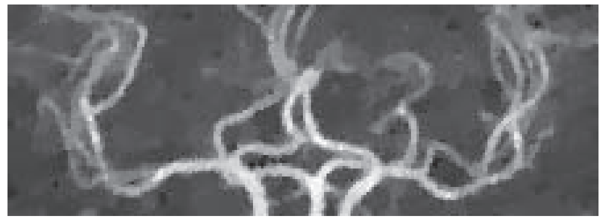
In this section, we report experimental results of the proposed MARX technique and discuss our findings. To highlight the importance of spatial adaptability of a CS recovery method and the efficacy of MARX in this regard, we conduct a comparative study between MARX and the total variation method (TV) method, which is the best in PSNR among all CS methods posted in [8]. The software of the TV methods in our evaluation was downloaded from [9]. A fairly general set of test images was used in our comparative study, including photographic images commonly found in the literature (e.g., Boats, Barb and Monarch) and many biomedical images.

Table 1 lists the PSNR values versus the number of CS measurements (presented as the percentage of the total number of pixels) for the two different CS recovery methods. The MARX algorithm consistently outperforms the TV method over all numbers of CS measurements by significant margins.

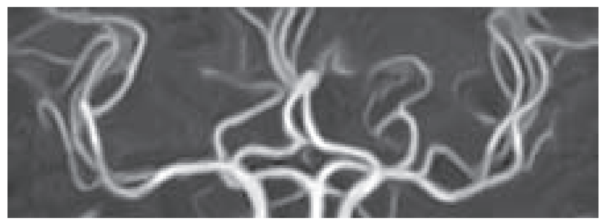
Fig. 1 is for an evaluation of tested methods in terms of visual quality. These results manifest the superiority of the MARX approach in preserving the structures and details of edges and textures. Its structure preserving ability is attributed to the underlying piecewise autoregressive model that adapts to spatially varying second-order statistics of the image signal.



(a) Original



(b) TV recovery with 15% measurements (24.22 dB)



(c) MARX with 15% measurements (30.04 dB)

Fig. 1. CS recovered MRA-Vessels images.

5. REFERENCES

[1] E. Candès, "Compressive sampling," in *Proceedings of the Int. Congress of Mathematics*, Madrid, Spain, 2006, pp. 1433–1452.

[2] D. Donoho, "Compressed sensing," *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[3] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Info. Theory*, vol. 51, pp. 4203–4215, 2005.

[4] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems," *IEEE J. Select Topic in Signal Processing*, vol. 1, no. 4, pp. 568–597, 2007.

[5] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Info. Theory*, vol. 53, no. 12, pp. 4655–1666, Dec. 2008.

[6] I. Daubechies, M. D. Friese, and C. D. Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity contraint," *Commun. Pure Appl. Math.*, vol. 57, pp. 1413–1457, 2004.

[7] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *EEE Transactions on Information Theory*, 2008, submitted.

[8] http://igorcarron.googlepages.com/cscodes.

[9] http://www.acm.caltech.edu/l1magic/.