IMAGE RECONSTRUCTION FROM PARTIAL FOURIER MEASUREMENTS VIA CURL CONSTRAINED SPARSE GRADIENT ESTIMATION

C. Ravazzi, G. Coluccia and E. Magli

Department of Electronics and Telecommunications (DET), Politecnico di Torino, Italy

ABSTRACT

In this paper, we propose new gradient-based methods for image reconstruction from partial Fourier measurements, which are commonly used in magnetic resonance imaging (MRI) or synthetic aperture radar. Compared to classical gradient recovery methods, a key improvement is obtained by formulating the gradient recovery problem as a compressed sensing problem with the additional constraint that the curl of the gradient field must be zero. Moreover, we formulate the image recovery problem as an inverse problem on graphs. Iteratively reweighted ℓ_1 recovery methods are proposed to recover these relative differences and the structure of the similarity graph. Finally, the image is recovered from the compressed Fourier measurements using least squares estimation. Numerical experiments demonstrate that the proposed approach outperforms the state-of-the-art image recovery methods.

***Index Terms*—** Compressed sensing, Fourier transform, Sparse recovery, Spectral graph theory, Total variation.

1. INTRODUCTION

The recovery of an image from Fourier measurements plays a very important role in several scanning technologies, such as magnetic resonance imaging (MRI, [1]) and synthetic aperture radar [2]. In this context, one would like to reduce the scan time and acquire the smallest number of measurements allowing recovery with the highest quality.

Compressed sensing [3] has emerged in the last few years as a valuable approach to reduce the amount of spectral data needed for reconstruction. Indeed, CS has been shown to effectively recover images from a limited number of samples by taking advantage of the sparse nature of the gradient minimizing total variation (TV, [4]). Among the solvers for TV minimization, RecPF [5] achieves the best CS recovery results. Although TV minimization allows a significant reduction in the number of measurements to be acquired, reconstructed images often suffer from undesirable artifacts and image details tend to be over-smoothed [6].

This work is supported by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement n.279848.

In [6] a new algorithm for image reconstruction, labelled as GradientRec–Diff, has been proposed. In a nutshell, given the set of spectral data, the horizontal and vertical differences are estimated from compressed measurements; the image is then recovered using an integration method. When the number of measurements is very small, RecPF achieves better reconstruction error. However, GradientRec–Diff shows better performance than RecPF, at the price of higher complexity, for low undersampling regimes.

To overcome these drawbacks, in this paper we propose new gradient based methods for image recovery. Inspired by the emerging field of signal processing on graphs, the image recovery problem is formulated as an inverse problem on graphs. More precisely, a graph is defined on the data units of the image: each unit is associated to a graph node and an edge is drawn with a weight depending on the similarity between the image values. We cast the gradient recovery problem as a compressed sensing problem enforcing both the sparsity and the directional continuity in the image gradient domain. Indeed, given incomplete information or presence of noise, the reconstructed gradient field by GradientRec–Diff might be not conservative and, consequently, non-integrable. In the proposed method we enforce that the integral along any closed curve should be equal to zero, as this allows to obtain a more accurate estimation of the image gradient and, consequently, a better image reconstruction quality. Moreover, iteratively reweighted ℓ_1 recovery methods are proposed to recover the relative differences and to infer the structure of the similarity graph. Once the gradient field is estimated, the image is recovered using least squares estimation.

Numerical experiments show that the proposed approach outperforms the state-of-the-art image recovery methods in terms of relative error for several sampling patterns. More precisely, the proposed CCGE algorithms achieve perfect reconstruction with the lowest number of measurements.

2. IMAGE RECOVERY FROM SPECTRAL DATA

2.1. Problem formulation

Let $F \in \mathbb{R}^{m\times n}$ be an image and denote $N = mn$ and $[n] = \{1, \ldots, n\}$. Each pixel of the image can be identified by a pair of indexes $(u_x, u_y) \in [m] \times [n]$ corresponding to the row and to the column. Let $\mathcal{F}:\mathbb{R}^{m\times n}\to\mathbb{C}^{m\times n}$ be the bidimensional DFT of F:

$$[\mathcal{F}(F)]_{\underline{\omega}}=\frac{1}{\sqrt{N}}\sum_{u_x\in[m]}\sum_{u_y\in[n]}F(u_x,u_y)\mathrm{e}^{-2\pi \mathrm{j}\left(\omega_x\frac{u_x}{n}+\omega_y\frac{u_y}{m}\right)}.$$

Defined a set of $M \ll N$ frequencies $\Omega=\{\underline{\omega}^{(k)}=(\omega_x^{(k)},\omega_y^{(k)}) : k\in\{1,\dots,M\}\}$, our aim is to recover the image F from partial frequency information $y=[\mathcal{F}(F)]_\Omega$. Let $\nabla F=(\nabla_x F,\nabla_y F)\in\mathbb{R}^{m\times n\times 2}$ be the discrete gradient operator: $(\nabla_x F)_{i,\ell}=F_{i,\ell}-F_{i-1,\ell}$, $(\nabla_y F)_{i,\ell}=F_{i,\ell}-F_{i,\ell-1}$. Given a vector field (F_1,F_2) the curl is defined as $\mathrm{curl}(F_1,F_2)=\nabla_y F_1-\nabla_x F_2$.

2.2. Gradient-based image recovery

Using the properties of the Fourier transform, it can be shown that the Fourier measurements of image gradient can be obtained by a diagonal transformation of the Fourier transform of the original image. More specifically,

$$y_x=\Lambda_x y=\mathcal{F}_\Omega(\nabla_x F)\quad y_y=\Lambda_y y=\mathcal{F}_\Omega(\nabla_y F)\qquad(1)$$

where $\Lambda_*=\mathrm{diag}(1-\mathrm{e}^{-2\pi\mathrm{j}\omega_*/N})$. Classical gradient-based recovery algorithms, as GradientRec-Diff [6], estimate the gradients via Basis Pursuit (BP)

$$\begin{array}{ll}\min\|\nabla_x F\|_1, & \text{s.t. } \mathcal{F}_\Omega(\nabla_x F)=y_x\\ \min\|\nabla_y F\|_1, & \text{s.t. } \mathcal{F}_\Omega(\nabla_y F)=y_y\end{array}\qquad(2)$$

Once the gradients have been estimated, the image is reconstructed using an integration method imposing that integral along any closed curve should be zero.

As will be clear next, when the number of measurements is very small, the gradient estimation obtained via BP is affected by errors. Consequently, the estimated gradient field is necessarily not conservative and these errors are not spread uniformly throughout the gradient field but concentrated around the edges of the image. As a consequence, these errors affect also the image recovery.

3. PROPOSED ALGORITHM

3.1. Graph analogy

We now formulate the image recovery problem using graph theory. Each pixel of the image F is labeled with a vertex $u\in\mathcal{V}$ and can be identified by a pair of indexes $(u_x,u_y)\in[m]\times[n]$ corresponding to the row and to the column. We consider the signal f, defined on the set of vertices $f:\mathcal{V}\to\mathbb{R}$ with the vector $f\in\mathbb{R}^N$, where the u-th entry represents the image value at the vertex $f_u=F(u_x,u_y)$.

The gradients are represented as the edges $\mathcal{E}\subseteq\mathcal{V}\times\mathcal{V}$ of an oriented graph $\mathcal{G}=(\mathcal{V},\mathcal{E})$. Each node u belonging to the interior of the image (*i.e.*, not on its boundary) has four edges, connecting it to nodes in north, south, east, and west directions. Therefore, the resulting graph is a grid. The orientation of the edge e connecting nodes u and v is conventionally assumed (u,v) with $u<v$. The graph topology is encoded in the incidence matrix $A\in\{0,\pm1\}^{\mathcal{E}\times\mathcal{V}}$ defined by

$$A_{ew}=\begin{cases}+1 & \text{if } e \text{ enters vertex } w\\ -1 & \text{if } e \text{ leaves vertex } w\\ 0 & \text{if } w \text{ is not a vertex of } e\end{cases}\qquad(3)$$

for every $e\in\mathcal{E}$ and $w\in\mathcal{V}$. We denote

$$g=\begin{bmatrix}g_x\\ g_y\end{bmatrix}=\begin{bmatrix}A_x\\ A_y\end{bmatrix}f=Af$$

where A_x and A_y are the incidence matrices corresponding to the horizontal and vertical directions, respectively, and $g_x=\mathrm{vec}(\nabla_x F)$ and $g_y=\mathrm{vec}(\nabla_y F)$. It should be noticed that with these notations $\mathrm{curl}(\nabla_x F,\nabla_y F)=A_y g_x-A_x g_y$.

3.2. Gradient estimation

3.2.1. *Curl-constrained ℓ_1 minimization*

In the proposed algorithm, we impose the curl of the reconstructed gradient field to be equal to zero. This yields the solution of the following optimization problem:

$$\underset{g_x,g_y\in\mathbb{R}^{mn}}{\arg\min}\ \|g_x\|_1+\|g_y\|_1 \quad \text{s.t.}\ \left\{\begin{array}{c}y_x=\mathcal{F}_\Omega(g_x)\\ y_y=\mathcal{F}_\Omega(g_y)\\ A_y g_x-A_x g_y=0\end{array}\right. .\qquad(4)$$

Different CS recovery algorithms can be used for edge reconstruction, such as OMP [7], CoSaMP [8], ℓ_p-minimization methods with $p<1$ [9], iteratively reweighted ℓ_1-minimization algorithms [10].

3.2.2. *Iteratively reweighted ℓ_1-CCGE with Gaussian weights*

In order to obtain high quality CS reconstruction, we modify the CCGE problem using both local smoothness and nonlocal self-similarity. More precisely, we consider the following weighted ℓ_1-problem

$$\underset{g_x,g_y\in\mathbb{R}^{\mathcal{V}}}{\arg\min}\sum_{i\in\mathcal{V}}w\left([g_x]_i\right)|[g_x]_i|+w\left([g_y]_i\right)|[g_y]_i| \quad \text{s.t.}\ \left\{\begin{array}{c}y_x=\mathcal{F}_\Omega(g_x)\\ y_y=\mathcal{F}_\Omega(g_y)\\ A_y g_x-A_x g_y=0\end{array}\right. ,\qquad(5)$$

where $w(x)=\mathrm{e}^{-\frac{x^2}{2\theta^2}}$ is the Gaussian kernel weighting function. It should be noted that this non-convex optimization problem is quite difficult to solve directly due to the non-differentiability and non-linearity. In this section, the iteratively reweighted ℓ_1-CCGE algorithm is developed to solve (5). We initialize $g_x^{(0)} = g_y^{(0)} = 0$, then at each iteration $t \in \mathbb{N}$ we compute

$$
\begin{aligned}
w_x^{(t)} &= \mathrm{e}^{-\frac{\left(g_x^{(t)}\right)^2}{2\theta^2}}, \quad w_y^{(t)} = \mathrm{e}^{-\frac{\left(g_y^{(t)}\right)^2}{2\theta^2}} \\
g^{(t+1)} &= \underset{g_x, g_y \in \mathbb{R}^{\mathcal{V}}}{\arg\min} \sum_{i \in \mathcal{V}} w_x^{(t)} |[g_x]_i| + w_y^{(t)} |[g_y]_i| \\
&\text{s.t.} \left\{ \begin{array}{c} y_x = \mathcal{F}_\Omega(g_x) \\ y_y = \mathcal{F}_\Omega(g_y) \\ A_y g_x - A_x g_y = 0 \end{array} \right. .
\end{aligned}
$$

3.3. From gradient to image

Let $\widehat{g} \in \mathbb{R}^{\mathcal{E}}$ be the vector collecting the estimated gradient field. We have $\widehat{g} = Af + \xi$, where ξ is the error obtained on the gradient estimation. Modeling the error ξ as Gaussian noise, we take a least-squares approach for estimating the signal f starting from measurements $\widehat{g}$:

$$\min_{x \in \mathbb{R}^N} \frac{1}{2} \|Ax - \widehat{g}\|_2^2. \tag{6}$$

It should be noticed that, being $A\mathbb{1} = 0$, the solution is not unique. The set of solutions of (6) is described in the following well-known lemma [11].

Lemma 1 (LS estimator)**.** *Let the graph $\mathcal{G}$ be connected and let $\mathcal{L} := A^\top A$ denote the Laplacian of the graph. The following facts hold:*

(a) x is a solution to (6) *if and only if $A^\top A x = A^\top \widehat{g}$;*
(b) there exists a unique minimizer of (6) *$\widehat{x}^{\mathrm{ls}}$ with minimum $\|\widehat{x}^{\mathrm{ls}}\|_2$;*
(c) $\widehat{f}^{\mathrm{ls}} = \mathcal{L}^\dagger A^\top \widehat{g}$;
(d) $\mathbb{E}[\widehat{f}^{\mathrm{ls}}] = \left(I - \frac{1}{N}\mathbb{1}\mathbb{1}^\top\right) f$.

It should be noted that determining the signal $\widetilde{f}$ from relative measurements is only possible up to an additive constant. This ambiguity can be avoided by taking $\widetilde{f} = \widehat{f}^{\mathrm{ls}} + \mathbb{1}\mathbb{1}^\top f / N$. It should be remarked that if $0 \in \Omega$ the second term is equal to $y(0)/\sqrt{N}$ leading to $\widetilde{f} = \widehat{f}^{\mathrm{ls}} + y(0)\mathbb{1}/\sqrt{N}$. These preliminary results yield the following theorem, whose proof is omitted for brevity.

Theorem 1. *Let $\widehat{g} \in \mathbb{R}^{\mathcal{E}}$ be the estimated gradient field and assume that $0 \in \Omega$, U eigenvectors of the Laplacian of the square grid, and D the diagonal matrix of corresponding eigenvalues. Then $\widetilde{f} = UD^\dagger U^\top A^\top \widehat{g} + \frac{y(0)}{\sqrt{N}}$.*

The $N = mn$ eigenvalues of the Laplacian of the square grid graph (contained in the diagonal of D) and the corresponding eigenvectors (collected as columns of U) can be evaluated analytically and it can be shown that the eigenvectors of this matrix are exactly the DCT Type II basis vectors (see [12]). Hence, the procedure to recover the image from the estimated gradient field can be efficiently implemented via the bidimensional DCT: given $\widehat{g}$, first $A^\top \widehat{g}$ is resized to match the size of the image; then the direct and inverse DCT are performed as stated in the following theorem.

Theorem 2. *Let $\widehat{g} \in \mathbb{R}^{\mathcal{E}}$ be the estimated gradient field and assume that $0 \in \Omega$, Λ the eigenvalues of the Laplacian of the square grid. Then*

$$\widetilde{f} = \mathrm{IDCT}_2\left[(\Lambda^+) \odot \mathrm{DCT}_2\left[\mathrm{vec}^{-1}(A^\top \widehat{g})\right]\right] + \frac{y(0)}{\sqrt{N}}, \tag{7}$$

where $\Lambda_{s,\ell} = 4\sin^2\left(\frac{\pi s}{2m}\right) + 4\sin^2\left(\frac{\pi l}{2n}\right)$, with $s \in [m]$ and $\ell \in [n]$ and $\odot$ denoted the elementwise product.

4. NUMERICAL RESULTS

In this section we test the following algorithms: (a) ℓ_1-CCGE–LS (*i.e.*, the algorithm recovering the gradient of the image as in Section 3.2.2 (with three iterations) and reconstructing the image using (7)); (b) CoSaMP-CCGE–LS (*i.e.*, the algorithm recovering the gradient of the image solving (4) using CoSaMP and reconstructing the image using (7)); (c) GradientRec–Diff [6], which first separately estimates the gradient on rows and the gradient on columns using (2) , then reconstructs the image from its gradient using an integration method based on diffusion tensors [13]; (d) RecPF [5] for TV minimization. We report, as benchmark, the reconstruction results of an *Oracle* device, which solves (5) in a single step using $w(\mathrm{vec}(\nabla F)_i) = \exp(-\frac{\mathrm{vec}(\nabla F)_i^2}{2\theta^2})$. In this sense, the performance of the Oracle provides the best performance achievable by ℓ_1-CCGE–LS.

We use as test image the 64×64 Shepp–Logan phantom, for different values of the compression ratio $L = M/N \in [0.01, 0.4]$. Three different undersampling patterns are used, namely, a radial sampling pattern, a uniformly-distributed sampling pattern and a variable-density sampling pattern (see [3, 14]). The performance of the algorithms have been evaluated in term of relative error $\mathsf{Err} := \|f - \widetilde{f}\|_2 / \|f\|_2$, where $\widetilde{f}$ is the estimated image and f is the true image.

For the algorithms requiring the ℓ_1 norm minimization, *i.e.*, the proposed ℓ_1-CCGE–LS and GradientRec–Diff, SPGL1 [15, 16] is used to solve the Basis Pursuit problem. Finally, it has to be remarked that a failure in the reconstruction, *i.e.*, a solver giving a Not-a-Number (NaN) as output, is treated as the algorithm returned a totally black image (all 0s), corresponding to a relative error of 1.

In Figure 1 the radial sampling pattern is used to collect measurements. It can be noticed that the proposed CCGE–LS algorithms are the ones showing the best performance, with almost perfect reconstruction for $L \geq 0.17$. RecPF reaches its optimum performance for $L \geq 0.2$ while GradientRec–Diff performs better than RecPF for $L > 0.27$, confirming the

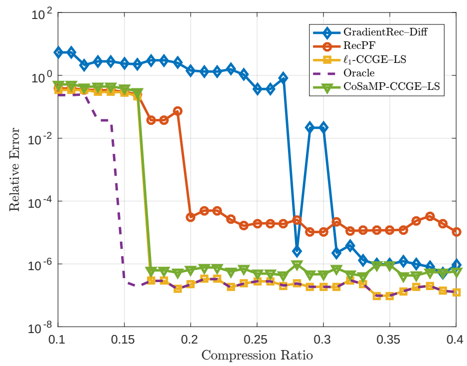

Fig. 1. Phantom 64×64: Relative reconstruction error vs. Compression Ratio for radial sampling patterns.

behavior obtained in [6]. There is a performance gap between the Oracle and the CCGE–LS.

Figure 2 visually shows the reconstruction error for $L = 0.17$. It can be noticed that the reconstruction obtained by the proposed CCGE–LS algorithms are almost perfect, with a relative error of 10^{-7}. RecPF shows an acceptable reconstruction quality (10^{-2}), even if some edge-related artifacts can be noticed. On the other hand, the reconstruction quality of GradientRec–Diff is significantly worse for $L = 0.17$.

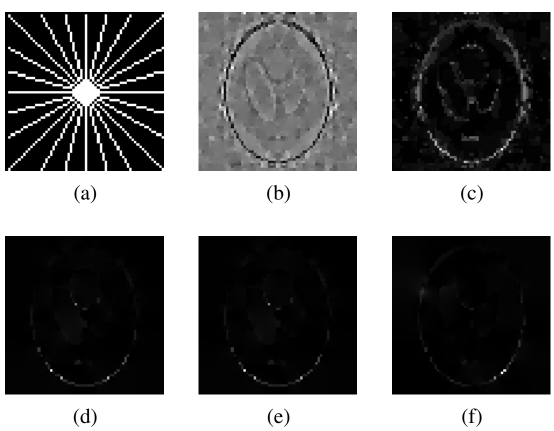

Fig. 2. Reconstruction error of image of Phantom 64×64 image obtained via gradient-based recovery methods. (a) Radial sampling pattern with compression Ratio $L = 0.17$. (b) GradientRec–Diff: Err $= 3.02$. (c) RecPF: Err $= 3.75 \cdot 10^{-2}$. (d) ℓ_1-CCGE–LS: Err $= 2.89 \cdot 10^{-7}$. (e) Oracle: Err $= 2.89 \cdot 10^{-7}$. (f) CoSaMP-CCGE–LS: Err $= 6.24 \cdot 10^{-7}$.

For uniformly distributed sampling pattern the results are shown in Figure 3. The results correspond to the average of 50 tests with different realizations of the sampling pattern. Here, it can be noticed that the ℓ_1-CCGE–LS algorithm is the one performing best, with almost-perfect reconstruction for $L \geq 0.14$ and reduced performance gap with respect to the Oracle. The CoSaMP-CCGE–LS performs slightly worse, with perfect reconstruction for $L \geq 0.19$, while GradientRec–Diff performs as in the radial sampling case, achieving its best result for $L \geq 0.28$. On the contrary, RecPF seems to suffer the uniformly distributed sampling pattern and could not reconstruct the signal.

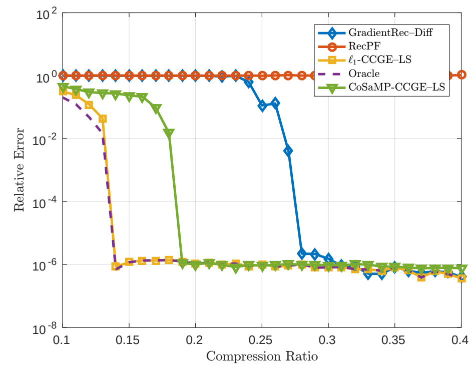

Fig. 3. Phantom 64×64: Relative reconstruction error vs. Compression Ratio. Random uniform sampling.

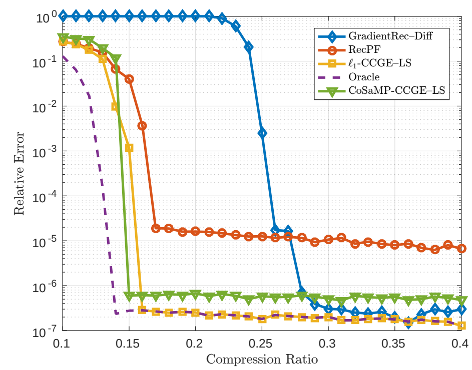

Fig. 4. Phantom 64×64: Relative reconstruction error vs. Compression Ratio. Variable density sampling.

Finally, we show in Figure 4 the results for the variable density sampling pattern. The results correspond to the average of 50 tests with different realizations of the sampling pattern. Again, the proposed CCGE–LS algorithms are the ones performing best, showing perfect reconstruction for $L \geq 0.15$, even if for $L \geq 0.29$ GradientRec–Diff performs slightly better than CoSaMP-CCGE–LS, while RecPF reaches its best performance for $L \geq 0.17$ even if with a higher error floor.

5. CONCLUDING REMARKS

In this paper, we have proposed new gradient based image recovery algorithms which combine constrained CS algorithms using curl information of gradient field with spectral graph filtering. Through extensive simulation, we have shown that the proposed algorithms outperform the state of the art also for small sampling ratio. Moreover, they are the least sensitive to the sampling pattern.

6. REFERENCES

[1] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[2] V. M. Patel, G. R. Easley, D. M. Healy, and R. Chellappa, "Compressed sensing for synthetic aperture radar imaging," in *2009 16th IEEE International Conference on Image Processing (ICIP)*, Nov 2009, pp. 2141–2144.

[3] Michael Lustig, David L Donoho, Juan M Santos, and John M Pauly, "Compressed sensing MRI," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 72–82, 2008.

[4] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, Feb 2006.

[5] Junfeng Yang, Yin Zhang, and Wotao Yin, "A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 288–297, 2010.

[6] Vishal M Patel, Ray Maleh, Anna C Gilbert, and Rama Chellappa, "Gradient-based image recovery methods from incomplete Fourier measurements," *Image Processing, IEEE Transactions on*, vol. 21, no. 1, pp. 94–105, 2012.

[7] Joel A Tropp and Anna C Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655–4666, 2007.

[8] Deanna Needell and Joel A Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[9] Ingrid Daubechies, Ronald DeVore, Massimo Fornasier, and C. Sinan Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Communications on Pure and Applied Mathematics*, vol. 63, no. 1, pp. 1–38, Jan. 2010.

[10] Emmanuel J. Candès, Michael B. Wakin, and Stephen P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 877–905, 2008.

[11] Prabir Barooah and João P Hespanha, "Error scaling laws for linear optimal estimation from relative measurements," *Information Theory, IEEE Transactions on*, vol. 55, no. 12, pp. 5661–5673, 2009.

[12] Gilbert Strang, "The discrete cosine transform," *SIAM Review*, vol. 41, no. 1, pp. 135–147, 1999.

[13] Vishal M Patel, Glenn R Easley, Rama Chellappa, and Dennis M Healy Jr, "Enhancing sparsity using gradients for compressive sensing," in *Image Processing (ICIP), 2009 16th IEEE International Conference on*. IEEE, 2009, pp. 3033–3036.

[14] Felix Krahmer and Rabab Ward, "Stable and robust sampling strategies for compressive imaging," *Image Processing, IEEE Transactions on*, vol. 23, no. 2, pp. 612–622, 2014.

[15] E. van den Berg and M. P. Friedlander, "SPGL1: A solver for large-scale sparse reconstruction," June 2007.

[16] E. van den Berg and M. P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2008.