

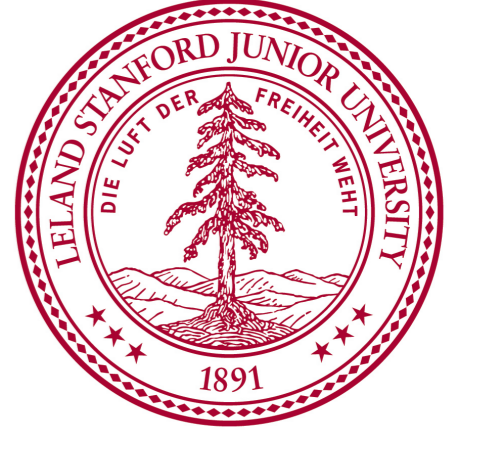


# Toward 0-norm Reconstruction by Convex Iteration

CHRISTINE LAW\* AND GARY GLOVER

Lucas Center, Stanford University, Stanford, CA 94305 USA

\* cslaw@stanford.edu



## Abstract

We repeat Candès' image reconstruction experiment [1] which led to discovery of sparse sampling theorems. But we achieve perfect reconstruction with an algorithm based on vanishing gradient of the compressed sensing problem's Lagrangian; computationally 10X faster than contemporary methods because

1. we formulate the problem as contraction which is solved efficiently by conjugate gradient method,
2. matrix multiplications are replaced by FFT, and
3. the number of constraints is cut in half by sampling symmetrically.

*Convex iteration* for cardinality minimization is introduced which allows perfect reconstruction of the phantom at 4% subsampling rate; 50% Candès' rate. By making neighboring-pixel selection variable, we show image-gradient sparsity of the Shepp-Logan phantom to be 1.9%; 33% lower than previously reported.

## Theory

We demonstrate application of image-gradient sparsification to the  $256^2$  Shepp-Logan phantom, simulating ideal acquisition of MRI data by radial sampling in the Fourier domain. Denoting an unknown image  $\mathcal{U} \in \mathbb{R}^{n \times n}$ , its 2D DFT  $\mathfrak{F}$  is  $\mathfrak{F}(\mathcal{U}) \triangleq FUF$  where  $F$  is a DFT matrix. Vectorized DFT is

$$\text{vec } \mathfrak{F}(\mathcal{U}) \triangleq (F \otimes F) \text{vec } \mathcal{U}$$

where  $\otimes$  is Kronecker product and  $\text{vec}$  is the vectorization operator. IDFT of the incomplete k-space samples  $K$  is

$$(F^H \otimes F^H) \delta(\text{vec } \Phi) (F \otimes F) \text{vec } \mathcal{U} = (F \otimes F) \text{vec } K \quad (1)$$

where  $\delta$  is the main-diagonal operator. Binary sampling mask  $\Phi$  is vertically and horizontally symmetric so that the IDFT is real; number of constraints is thus reduced by half.

Equation 1 is abbreviated as  $P \text{vec } \mathcal{U} = f$ . Total variation operator  $\Psi \in \mathbb{R}^{4n^2 \times n^2}$  computes first-order difference of neighboring pixels in right, left, up, and down directions.

The optimization problem to recover an unknown image by cardinality minimization is

$$\underset{\mathcal{U}}{\text{minimize}} \|\Psi \text{vec } \mathcal{U}\|_0 \quad \text{subject to } P \text{vec } \mathcal{U} = f \quad (2)$$

Problem (2) is hard to solve. The common approach is to replace 0-norm by 1-norm. When enough samples are taken, 1-norm yields results identical to a 0-norm formulation. [2]

Here, instead of 1-norm, we introduce a *direction vector*  $y$  as part of a *convex iteration* method [3] [4] that requires fewer Fourier samples than 1-norm while optimal solution more closely approaches a 0-norm result: *i.e.*, there exists a vector  $y^*$ , having entries  $y_i^* \in \{0, 1\}$ , such that (2) is equivalent to a Lagrangian form

$$\underset{\mathcal{U}}{\text{minimize}} \langle |\Psi \text{vec } \mathcal{U}|, y^* \rangle + \frac{1}{2} \lambda \|P \text{vec } \mathcal{U} - f\|_2^2 \quad (3)$$

Existence of such a  $y^*$ , complementary to an optimal vector  $\Psi \text{vec } \mathcal{U}^*$ , is obvious by definition of global optimality

$$\langle |\Psi \text{vec } \mathcal{U}^*|, y^* \rangle \triangleq 0 \quad (4)$$

under which a cardinality- $c$  optimal objective is assumed to exist.

Problem (3) is not computable because  $P$  is dimensionally too large. Since (3) is an unconstrained convex problem, a zero in the objective function's gradient is necessary and sufficient for optimality; *i.e.*,

$$\Psi^T \delta(y) \text{sgn}(\Psi \text{vec } \mathcal{U}) + \lambda P^H (P \text{vec } \mathcal{U} - f) = \mathbf{0}.$$

Because  $P$  is idempotent and Hermitian and  $\text{sgn}(\kappa) = \kappa/|\kappa|$ , for  $\lambda \gg 0$  and  $0 \preceq y \preceq \mathbf{1}$ ,

$$\text{vec } \mathcal{U}_{t+1} = (\Psi^T \delta(y) \delta(|\Psi \text{vec } \mathcal{U}_t|)^{-1} \Psi + \lambda P)^{-1} \lambda P f \quad (5)$$

is a contraction in  $\mathcal{U}_t$  that can be solved recursively in  $t$  by conjugate gradient method.

Matrix  $P$  is never actually formed; it is replaced by 2D FFT, binary masking, and then IFFT.

## Convex Iteration

Direction vector  $y$  is initialized to  $\mathbf{1}$  until the first fixed point is found; which means, the first sequence of contractions begins calculating the (1-norm) solution  $\mathcal{U}^*$  via problem (3). Once  $\mathcal{U}^*$  is found, vector  $y$  is updated according to an image-gradient cardinality estimate  $c$ : Sum of the  $4n^2 - c$  smallest entries of  $|\Psi \text{vec } \mathcal{U}^*| \in \mathbb{R}^{4n^2}$  is the optimal objective value from a linear program, for  $0 \leq c \leq 4n^2 - 1$

$$\sum_{i=c+1}^{4n^2} \pi(|\Psi \text{vec } \mathcal{U}^*|)_i = \underset{y \in \mathbb{R}^{4n^2}}{\text{minimize}} \langle |\Psi \text{vec } \mathcal{U}^*|, y \rangle$$

$$\text{subject to } 0 \preceq y \preceq \mathbf{1} \quad (6)$$

$$y^T \mathbf{1} = 4n^2 - c$$

where  $\pi$  is a nonlinear permutation-operator sorting its vector argument into nonincreasing order.

An optimal solution  $y$  to (6), that is an extreme point of its feasible set, is known in closed form: it has 1 in each entry corresponding to the  $4n^2 - c$  smallest entries of  $|\Psi \text{vec } \mathcal{U}^*|$  and has 0 elsewhere. The updated image  $\mathcal{U}^*$  is assigned to  $\mathcal{U}_t$ , the contraction sequence is recomputed solving (3), direction vector  $y$  is updated again, and so on until convergence; which is guaranteed by virtue of a monotonically nonincreasing real sequence of objective values.

Optimal vector  $y^*$  is not necessarily uniformly distributed across right, left, up, and down differences represented by  $\Psi$ ; this means an estimate of image-gradient is not locked into four adjacent points per pixel, and can be fewer. In fact, we find image-gradient cardinality is 1.9% ( $c=5092$ ).

## Results

Figure 1 shows aliasing of phantom resulting from 4.1% (10 lines) k-space subsampling (Fig.2). Reconstruction (Fig.3), from this aliased image, is realized in MATLAB 20007b: run-time is 96s on IBM laptop (1.83GHz, 3GB RAM). Calling original phantom  $\mathcal{I}$ ,

$$\text{SNR} \triangleq 20 \log_{10}(\|\mathcal{I}\|_F / \|\mathcal{I} - \mathcal{U}^*\|_F) = 104\text{dB}$$

## Discussion

Perfect reconstruction of the Shepp-Logan phantom is quickly achieved with 4.1% subsampled data; well below an 11% least lower bound predicted by the sparse sampling theorem. [5] Because reconstruction approaches optimal solution to a 0-norm problem, minimum number of Fourier-domain samples is bounded below by cardinality of the image-gradient at 1.9%.

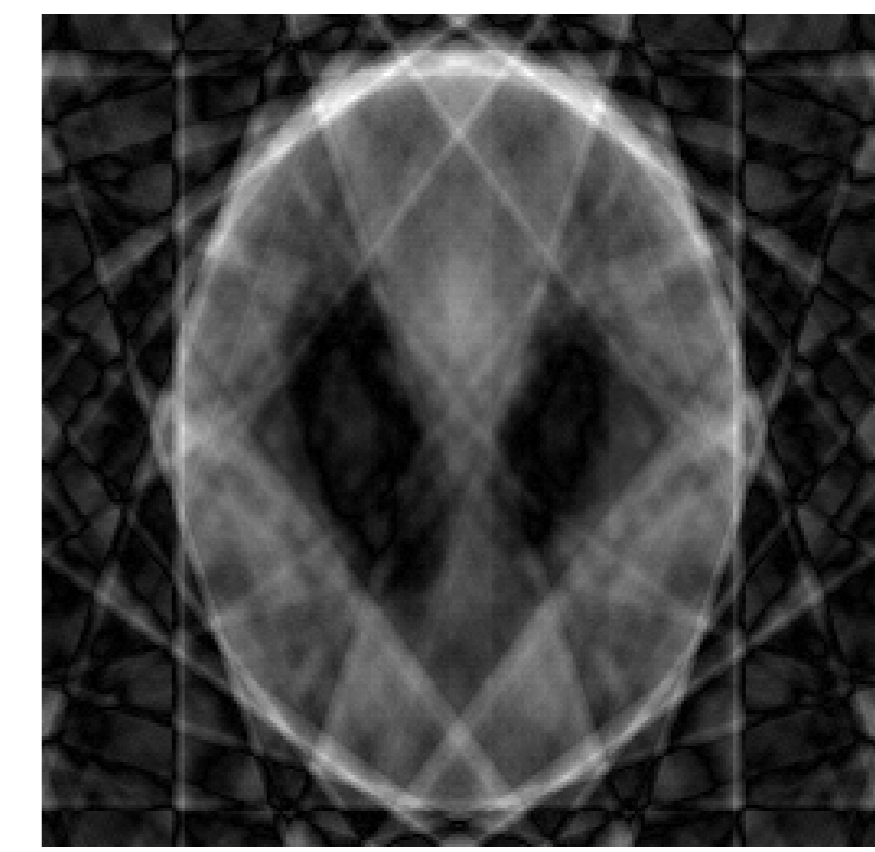


Figure 1: Aliasing of Shepp-Logan phantom

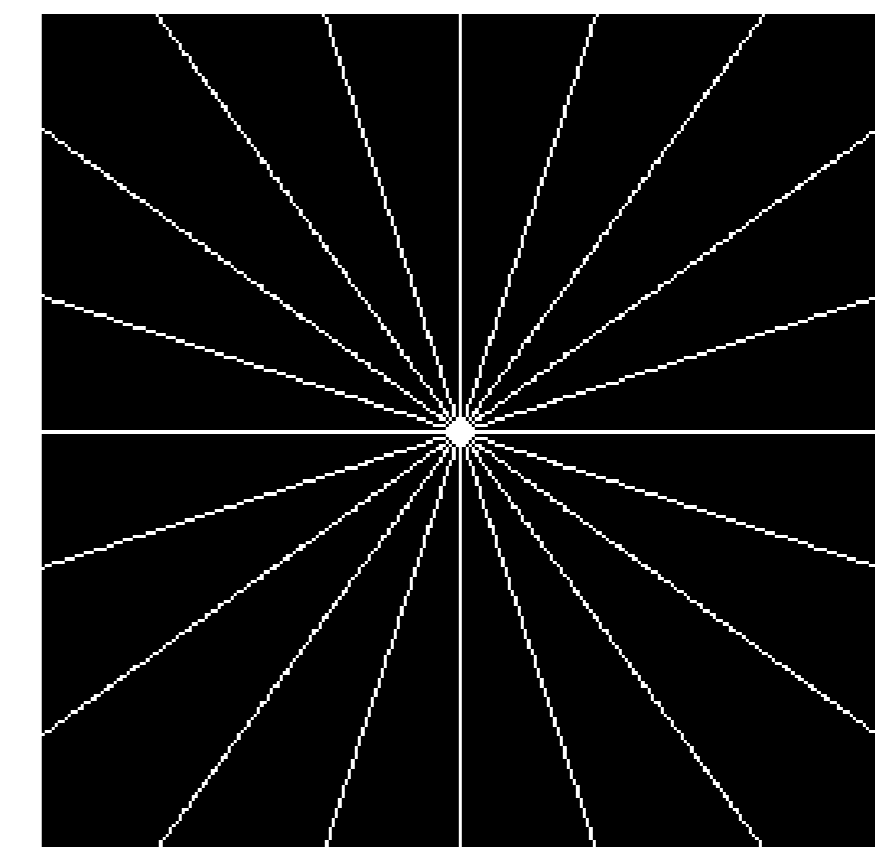


Figure 2: Symmetric binary mask  $\Phi$  samples 4% (10 lines) k-space data



Figure 3: Reconstructed image, SNR=104dB

1. Candès *et al.* IEEE Trans. Info Theory, 52(2):489, 2006.
2. Donoho. IEEE Trans. Info Theory, 52(4):1289, 2006.
3. Dattoro. Convex Optimization & Euclidean Distance Geometry. Meboo, 2005.
4. Chretien. An alternating l1 approach to the compressed sensing problem, arxiv.org/PS\_cache/arxiv/pdf/0809/0809.0660v1.pdf, 2008.
5. Baron *et al.*, Distributed compressed sensing, cite-seerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.85.4789, 2005