Image Acquisition from Highly Incomplete Information

Justin Romberg, California Institute of Technology

IMA: New Mathematics and Algorithms for 3-D Image Analysis
Minneapolis, Minnesota
January 9, 2006

Collaborators: Emmanuel Candès (Caltech), Terence Tao (UCLA)
Shepp-Logan Phantom

\[ N = 512^2 = 262,144 \text{ pixels} \]
Model Acquisition Problem

Observe a subset $\Omega$ of the 2D discrete Fourier plane

22 radial lines, 10,486 samples, $\approx 4\%$ coverage
“Filtered Backprojection”

Reconstruct $g^*$ with

$$
\hat{g}^*(\omega) = \begin{cases} 
\hat{f}(\omega) & \omega \in \Omega \\
0 & \omega \not\in \Omega 
\end{cases}
$$

Set unknown Fourier coeffs to zero, and inverse transform
Total Variation Reconstruction

Idea: Find an image that

- Fourier domain: *matches observations*
- Spatial domain: has a *minimal amount of oscillation*

\[
\min_g \|g\|_{TV} \quad \text{s.t.} \quad g(\omega) = f(\omega), \omega \in \Omega
\]

\[
\|g\|_{TV} = \sum_{t_1, t_2} \sqrt{|g(t_1 + 1, t_2) - g(t_1, t_2)|^2 + |g_{i,j} + 1 - g_{i,j}|^2} = \sum_{i,j} |(\nabla g)_{i,j}|_{image}
\]

\[\Omega \text{ in Fourier domain}\]
Total Variation Reconstruction

Idea: Find an image that

- Fourier domain: *matches observations*
- Spatial domain: has a *minimal amount of oscillation*

Reconstruct \( g^* \) by solving:

\[
\min_g \text{TV}(g) \quad \text{subject to} \quad \hat{g}(\omega) = \hat{f}(\omega) \quad \text{for} \quad \omega \in \Omega
\]

\[
\text{TV}(g) = \sum_{t_1,t_2} \left( (g_{t_1+1,t_2} - g_{t_1,t_2})^2 + (g_{t_1,t_2+1} - g_{t_1,t_2})^2 \right) = \sum_{t_1,t_2} |(\nabla g)_{t_1,t_2}|
\]

![Image](image.png)  ![Image](Omega.png)

image  \( \Omega \) in Fourier domain
Reconstruct $g^*$ with

$$\min_g \ TV(g) \quad \text{s.t.} \quad \hat{g}(\omega) = \hat{f}(\omega), \ \omega \in \Omega$$

$$TV(g) = \sum_{t_1,t_2} |(\nabla g)_{t_1,t_2}|$$

original  \hspace{1cm} g^* = \text{original} — \text{perfect reconstruction!}
Reconstruction from incomplete measurements

• Given: $K$ linear measurements of an unknown digital object $f(t) \in \mathbb{R}^N$

$$ y_k = \langle f, \phi_k \rangle, \quad k = 1, \ldots, K $$

$y_k$ = “measurements”, $\phi_k(t)$ = “test functions”

• Examples:
  - Delta functions, $\phi_k = \delta(t - t_k) = \{1$ at $t = t_k, \ 0$ elsewhere$\}$, $y_k$ are samples of $f(t)$
  - Complex sinusoids, $\phi_k = \exp(2\pi i \omega_k t / N)$, $y_k$ are Fourier coefficients of $f(t)$
Reconstruction from incomplete measurements

• Given: $K$ linear measurements of an unknown digital object $f(t) \in \mathbb{R}^N$

$$y_k = \langle f, \phi_k \rangle, \quad k = 1, \ldots, K$$

$y_k$ = “measurements”, $\phi_k(t)$ = “test functions”

• Examples:
  - Delta functions, $\phi_k = \delta(t - t_k) = \{1 \text{ at } t = t_k, \ 0 \text{ elsewhere}\}$, $y_k$ are samples of $f(t)$
  - Complex sinusoids, $\phi_k = \exp(2\pi i \omega_k t / N)$, $y_k$ are Fourier coefficients of $f(t)$
  - Line “integrals”, cone “integrals”, ...

• Problem: recover $N$-point signal from $K$ measurements when $K \ll N$

• Impossible?
  - In general, of course it is impossible
  - But, if $f$ is “sparse”, we can recover perfectly surprisingly often
Example: Sampling a superposition of sinusoids

• Suppose \( f \) is sparse in the Fourier domain:

\[
\hat{f}(\omega) = \sum_{b=1}^{B} \alpha_b \delta(\omega - \omega_b) \quad \Leftrightarrow \quad f(t) = \sum_{b=1}^{B} \alpha_b e^{i\omega_b t}
\]

\( f \) is a superposition of \( B \) complex sinusoids.

• Note: frequencies \( \{\omega_b\} \) and amplitudes \( \{\alpha_b\} \) are unknown.

• Take \( K \) samples of \( f \) at locations \( t_1, \ldots, t_k \).
Sampling example

Time domain $f(t)$

Measure $K$ samples
(red circles = samples)

Frequency domain $\hat{f}(\omega)$

$B$ nonzero components

$\#\{\omega : \hat{f}(\omega) \neq 0\} := \|\hat{f}\|_{\ell_0} = B$
Sparse recovery

• We measure $K$ samples of $f$

\[ y_k = f(t_k), \quad k = 1, \ldots, K \]

• Find signal with \textit{smallest frequency domain support} that matches the measured samples

\[ \min_g \|\hat{g}\|_{\ell_0} \quad \text{subject to} \quad g(t_k) = y_k, \quad k = 1, \ldots, K \]

where $\|\hat{g}\|_{\ell_0} := \# \{\omega : \hat{g}(\omega) \neq 0\}$. 

• Theorem: If $\|\hat{f}\|_{\ell_0} = B$, we can recover $f$ from (almost) any set of $K \geq \text{Const} \cdot B \cdot \log N$ samples.

• The program is absolutely intractable (combinatorial, NP hard).
Sparse recovery

- We measure $K$ samples of $f$

$$y_k = f(t_k), \quad k = 1, \ldots, K$$

- Find signal with *smallest frequency domain support* that matches the measured samples

$$\min_g \|\hat{g}\|_{\ell_0} \quad \text{subject to} \quad g(t_k) = y_k, \quad k = 1, \ldots, K$$

where $\|\hat{g}\|_{\ell_0} := \#\{\omega : \hat{g}(\omega) \neq 0\}$.

- **Theorem:** If $\|\hat{f}\|_{\ell_0} = B$, we can recover $f$ from (almost) any set of

$$K \geq \text{Const} \cdot B \cdot \log N$$

samples.

- The program is absolutely intractable (combinatorial, NP hard).
Convex relaxation

- Convex relaxation: use $\ell_1$ norm as a proxy for sparsity

$$\|\hat{g}\|_{\ell_1} := \sum_\omega |\hat{g}(\omega)|$$

$\ell_1$ norm = “sum of magnitudes”

- Recover from samples $y_k = f(t_k)$ by solving

$$(P1) \min_g \|\hat{g}(\omega)\|_{\ell_1} \text{ subject to } g(t_k) = y_k, \ k = 1, \ldots, K$$

- Very tractable; linear or second-order cone program

- Surprise: $(P1)$ still recovers sparse signals perfectly.
\( \ell_1 \) reconstruction

Reconstruct by solving

\[
\min_{\hat{g}} \| \hat{g} \|_{\ell_1} := \min \sum_{\omega} |\hat{g}(\omega)| \quad \text{subject to} \quad g(t_k) = f(t_k), \; k = 1, \ldots, K
\]

original \( \hat{f} \), \( B = 15 \)  

perfect recovery from 30 samples
A recovery theorem

- **Exact Recovery Theorem**
  - Suppose $\hat{f}$ is supported on set of size $B$
  - Select $K$ sample locations $\{t_k\}$ “at random” with
    
    $$K \geq \text{Const} \cdot B \log N$$
  - Take time-domain samples (measurements) $y_k = f(t_k)$
  - Solve
    
    $$\min_g \|\hat{g}\|_1 \text{ subject to } g(t_k) = y_k, \ k = 1, \ldots, K$$
  - Solution is *exactly* $f$ with extremely high probability.

- In theory, $\text{Const} \approx 20$
- In practice, perfect recovery occurs when $K \approx 2B$ for $N \approx 1000$.
- *In general, minimizing $\ell_1$ finds $f$ from $K \sim B \log N$ samples*
- In total-variation/phantom example, $B=$number of jumps
Nonlinear sampling theorem

- $\hat{f} \in \mathbb{C}^N$ supported on set $\Omega$ in Fourier domain

- Shannon sampling theorem:
  - $\Omega$ is a known connected set of size $B$
  - exact recovery from $B$ equally spaced time-domain samples
  - linear reconstruction by sinc interpolation

- Nonlinear sampling theorem:
  - $\Omega$ is an *arbitrary and unknown* set of size $B$
  - exact recovery from $\sim B \log N$ (almost) arbitrarily placed samples
  - nonlinear reconstruction by convex programming
History and Related Research

- Novel sampling theorems
  - Bresler and Feng (2002); Vetterli et al. (2002–2004)

- Fast algorithms for $B$-term Fourier approximation

- Classical $\ell_1$ reconstruction
  - Santosa and Symes (1986) and others in geophysics
  - Donoho and Stark (1989)

- $\ell_1$ ("Basis Pursuit") for sparse decompositions
  - Chen, Donoho, Saunders (1999); Donoho and Huo (2001)
  - Elad, Gribonval, Nielsen, Fuchs (2001-2004)
Generalized measurements and sparsity

- $f$ is sparse in a known orthogonal system $\Psi$: the $\Psi$-transform is supported on a set of size $B$,

$$\alpha = \Psi^T f, \quad \#\{\omega : \alpha(\omega) \neq 0\} = B$$

- Linear measurements using “test functions” $\phi_k(t)$

$$y_k = \langle f, \phi_k \rangle, \quad \text{or} \quad y = \Phi f, \quad \Phi : K \times N$$

Measurement matrix $\Phi$ is formed by stacking rows $\phi_k^T$
Generalized measurements and sparsity

- $f$ is sparse in a known orthogonal system $\Psi$: the $\Psi$-transform is supported on a set of size $B$,

$$\alpha = \Psi^T f, \quad \#\{\omega : \alpha(\omega) \neq 0\} = B$$

- Linear measurements using “test functions” $\phi_k(t)$

$$y_k = \langle f, \phi_k \rangle, \quad \text{or} \quad y = \Phi f, \quad \Phi : K \times N$$

Measurement matrix $\Phi$ is formed by stacking rows $\phi_k^T$

- To recover, solve

$$\min_f \|\Psi^T f\|_{\ell_1} \quad \text{such that} \quad \Phi f = y$$

- Exact recovery if basis $\Psi$ and measurement system $\Phi$ are incoherent

- Random $\Phi$ is incoherent with fixed $\Psi$ with high probability
Random measurements

• Gaussian random matrix ($K \times N$):

  \[ \Phi_{k,n} \sim \text{Normal}(0, 1) \]

• Measure \( y = \Phi f \)

• Theorem (Candès and Tao): If \( f \) is \( B \)-sparse in a known orthobasis \( \Psi \), solving

  \[ \min_f \| \Psi^T f \|_{\ell_1} \quad \text{subject to} \quad \Phi f = y \]

  recover \( f \) exactly when

  \[ K \geq \text{Const} \cdot B \log N. \]

• Once chosen, the same \( \Phi \) can be used to recover all sparse \( f \)

• Finding incoherent measurement matrices is easy!
Compressed sensing

- Can recover $B$-sparse $f$ from $O(B \log N)$ incoherent measurements $\Rightarrow$ number of sensors proportional to inherent complexity of $f$
- The sensing is *not* adaptive, and is simple
- The recovery is flexible
  \[
  \min_f \| \Psi^T f \|_{\ell_1} \quad \text{subject to} \quad \Phi f = y
  \]
  Different $\Psi$ yield different recoveries from same measurements
- Democratic and robust:
  - all measurement are equally (un)important
  - losing a few does not hurt
- Active sensing is secure: incoherent test functions $\phi_k$ have no structure
Stability

• What happens if the measurements are noisy?

\[ y = \Phi f + e, \quad \text{with} \quad \|e\|_2 \leq \epsilon \]

• Recover: \( \ell_1 \) minimization with relaxed constraints

\[ f^\# = \arg \min \| \Psi^T f \|_{\ell_1} \quad \text{subject to} \quad \| \Phi f - y \|_2 \leq \epsilon \]

• Stability Theorem:
  If \( \Phi \) is incoherent w.r.t. \( \Psi \), then solution \( f^\# \) obeys

\[ \| f^\# - f \|_2 \leq \text{Const} \cdot \epsilon \]

• Recovery error is on the same order as the observation error
Examples
Perfect Recovery

1024 \times 1024 \text{ image}

25k \text{ term wavelet approx}

wavelet coeffs
Perfect Recovery

- Take $K = 96000$ incoherent measurements $y = \Phi f_a$
- $f_a = 25$-term wavelet approximation (perfectly sparse)
- Solve
  \[
  \min \| \Psi^T f \|_{\ell_1} \quad \text{subject to} \quad \Phi f = y
  \]
  $\Psi = \text{wavelet transform}$

original \hspace{2cm} perfect recovery
Imaging: Fuel Cells

- “Look inside” fuel cells as they are operating via neutron imaging
- Accelerate process by limiting the number of projections
- Each projection = samples along radial lines in the Fourier domain

Given measurements $y$, solve

$$\min \|g\|_{TV} \quad \text{subject to} \quad \mathcal{P}_\Omega g = y$$

where $\mathcal{P}_\Omega =$ partial pseudo-polar FFT.
Imaging: Fuel Cells

Reconstruction from 20 projections:

original | backprojection | min TV
$\Omega \approx 29\%$ of samples
Quantized observations

measure

quantize

resolution = 1 digit
Quantized observations

Given measurements $y$, recover via

$$\min ||x||_{TV} \quad \text{subject to} \quad ||\Phi x - y|| \leq \epsilon$$

original

recovered from 25k measurements
Conclusions

- Signal with $B$ components can be recovered from $\sim B \log N$ measurements.
- Recovery is stable.
- Recovery is computationally tractable:
  - Optimization problems are second-order cone programs (SOCP).
  - Interior-point methods recover a $256 \times 256 (N = 65,536)$ in a few minutes.
  - Very rough rule-of-thumb: recovery cost $\approx 1000$ FFTs.
- Applications:
  - Tomographic imaging.
  - Random, compressed sensing.
  - Analog-to-digital.
  - Flexible, universal data compression.
  - ...