Compressive Sensing:
A New Framework for Computational Signal Processing

Richard Baraniuk
Rice University
dsp.rice.edu/cs
Pressure is on Signal Processing

• Networked sensing placing increasing pressure on signal/image processing hardware and algs to support

  higher resolution / denser sampling
  » ADCs, cameras, imaging systems, ...

+ large numbers of sensors
  » multi-view target data bases, camera arrays and networks, pattern recognition systems,

+ increasing numbers of modalities
  » acoustic, seismic, RF, visual, IR, SAR, ...

= deluge of data
  » how to acquire, store, fuse, process efficiently?
Antipasto

Sensing by Sampling
Data Acquisition and Representation

- Time: A/D converters, receivers, ...
- Space: cameras, imaging systems, ...

- Foundation: *Shannon sampling theorem*
  - *Nyquist rate*: must sample at 2x highest frequency in signal
Sensing by *Sampling*

- Long-established paradigm for digital data acquisition
  - *sample* data (A-to-D converter, digital camera, ...)
  - *compress* data (signal-dependent, nonlinear)

\[
\begin{align*}
\mathcal{X} \xrightarrow{\text{sample}} N & \xrightarrow{\text{compress}} K \\
& \xrightarrow{\text{transmit/store}} \\
& \xrightarrow{\text{receive}} K \xrightarrow{\text{decompress}} N \xrightarrow{\hat{\mathcal{X}}}
\end{align*}
\]

- *sparse* wavelet transform
Sparsity

- Many signals can be **compressed** in some representation/basis (Fourier, wavelets, ...)

\[ N \text{ pixels} \]

\[ N \text{ wideband signal samples} \]

\[ K \ll N \]

large wavelet coefficients

\[ K \ll N \]

large Gabor coefficients
Sensing by *Sampling*

- **sample** data (A-to-D converter, digital camera, …)
- **compress** data (signal-dependent, nonlinear)
- **brick wall** to performance of modern acquisition systems

\[ f \rightarrow \text{sample} \rightarrow \overset{N \gg K}{\text{compress}} \rightarrow \text{transmit} \]

\[ \overset{K}{\text{receive}} \rightarrow \text{decompress} \rightarrow \overset{N}{\hat{f}} \]

*sparse* wavelet transform
Pasta

Compressive Sensing
From Samples to Measurements

- Shannon was a pessimist
  - worst case bound for any bandlimited data

- **Compressive sensing** (CS) principle
  
  “sparse signal statistics can be recovered from a small number of nonadaptive linear measurements”

  - integrates sensing, compression, processing
  
  - based on new **uncertainty principles** and concept of **incoherency** between two bases
Incoherent Bases

- Spikes and sines (Fourier) (Heisenberg)

$$\psi = I$$

$$\Phi = \text{idct}(I)$$
Incoherent Bases

- Spikes and “random basis”

\[ \Psi = I \]

\[ \Phi = \text{randn}(N, N) \]
Incoherent Bases

- Spikes and “random sequences” (codes)

\[ \Psi = I \]

\[ \Phi \]
Incoherent Bases
Sampling

- Signal \( x \) is \( K \)-sparse in basis/dictionary \( \Psi \)
  - WLOG assume sparse in space domain \( \Psi = I \)

**Samples**

\[ y = \Phi x \]

\( N \times 1 \) measurements

\( N \times 1 \) sparse signal

\( K \) nonzero entries
Compressive Sensing
[Candes, Romberg, Tao; Donoho]

- Signal $x$ is $K$-sparse in basis/dictionary $\Psi$
  - WLOG assume sparse in space domain
  $$\Psi = I$$

- Replace samples with \textit{few linear projections} $y = \Phi x$

- $M \times 1$ measurements

- $K < M \ll N$

- $N \times 1$ sparse signal

- $K$ nonzero entries
Compressive Sensing
[Can
des, Romberg, Tao; Donoho]

• Signal $x$ is $K$-sparse in basis/dictionary $\Psi$
  - WLOG assume sparse in space domain
    $\Psi = I$

• Replace samples with few linear projections $y = \Phi x$

\[ M \times 1 \\text{measurements} \]
\[ y \]
\[ = \]
\[ M \times N \]
\[ \Phi \]
\[ x \]
\[ N \times 1 \text{ sparse signal} \]

• Random measurements $\Phi$ will work!
Compressive Sensing

- Measure linear projections onto *incoherent* basis where data is *not sparse/compressible*

\[ x \rightarrow \text{project} \rightarrow \text{transmit/store} \]

- Reconstruct via *nonlinear processing* (optimization) (using sparsity-inducing basis)
CS Signal Recovery

- Reconstruction/decoding: given
  (ill-posed inverse problem) find

\[ y = \Phi x \]

\[ M \times 1 \]
\[ \text{measurements} \]

\[ K < M \ll N \]

\[ M \times N \]

\[ N \times 1 \]
\[ \text{sparse signal} \]

\[ K \]
\[ \text{nonzero entries} \]
CS Signal Recovery

• Reconstruction/decoding: given \( y = \Phi x \) find \( x \)

• \( L_2 \) fast, wrong

\[
\hat{x} = \arg \min_{y=\Phi x} \|x\|_2
\]

\[
x = (\Phi^T \Phi)^{-1} \Phi^T y
\]
CS Signal Recovery

• Reconstruction/decoding: given $y = \Phi x$, find $x$

  (ill-posed inverse problem)

- $L_2$ fast, wrong

- $L_0$ correct, slow
  only $M=K+1$ measurements required to perfectly reconstruct $K$-sparse signal

[Bresler; Rice]

\[
\hat{x} = \arg\min_{y=\Phi x} \|x\|_2
\]

\[
\hat{x} = \arg\min_{y=\Phi x} \|x\|_0
\]

(number of nonzero entries)
CS Signal Recovery

• Reconstruction/decoding: given $y = \Phi x$ find $x$
  (ill-posed inverse problem)

  • $L_2$ fast, wrong
  \[ \hat{x} = \arg \min_{y=\Phi x} \|x\|_2 \]

  • $L_0$ correct, slow
  \[ \hat{x} = \arg \min_{y=\Phi x} \|x\|_0 \]

  • $L_1$ correct, mild oversampling
    \[ \hat{x} = \arg \min_{y=\Phi x} \|x\|_1 \text{ linear program} \]

\[ M = O(K \log(N/K)) \ll N \]
CS Signal Recovery

Why It Works: **Sparsity**

- Many signals can be *compressed* in some representation/basis (Fourier, wavelets, ...)

\[ N \text{ pixels} \]

\[ N \text{ wideband signal samples} \]

\[ K \ll N \text{ large wavelet coefficients} \]

\[ K \ll N \text{ large Gabor coefficients} \]
Sparse Models are Nonlinear

\[ \text{Nonlinear} \]
Sparse Models are *Nonlinear*

\[ x = \sum_{k=1}^{N} \alpha_i \psi_i \]

\[ x \approx \sum_{K \ll N \text{ largest terms}} \alpha_i \psi_i \]

\( \text{\(K \ll N\) large wavelet coefficients} \)
Sparse Models are Nonlinear

$N$ pixels

$K \ll N$ large wavelet coefficients

$x \approx \sum_{K \ll N} \text{largest terms}$

model for all $K$-sparse signals:

**union of subspaces** (aligned with coordinate axes)
Why $L_2$ Doesn’t Work

$$\hat{x} = \text{arg min}_{y=\Phi x'} \|x'\|_2$$

least squares, minimum $L_2$ solution is almost **never sparse**

null space of $\Phi$
translated to $x$
(random angle)
Why $L_1$ Works

$$\hat{x} = \arg \min_{y = \Phi x'} \| x' \|_1$$

minimum $L_1$ solution
= sparsest solution if

$$M = O(K \log(N/K)) \ll N$$
Universality

- Gaussian white noise basis is incoherent with any fixed orthonormal basis (with high probability)
- Signal sparse in time domain: $\Phi = I$

$$y = \Phi x$$
Universality

• Gaussian white noise basis is incoherent with any fixed orthonormal basis (with high probability)
• Signal sparse in frequency domain: $\Psi = \text{idct}$

- Product $\Phi \Psi$ remains Gaussian white noise
Pesce

Compressive Sensing in Action
Single-Pixel CS Camera

Image encoded by PMM and random basis

random pattern on DMD array

single photon detector

Low-cost, fast, sensitive optical detection

PD

A/D

Compressed, encoded image data sent via RF for reconstruction

Rng

Xmtr

Rcvr

DSP

image reconstruction

w/ Kevin Kelly and students
TI Digital Micromirror Device (DMD)
Single Pixel Camera

Low-cost, fast, sensitive optical detection

Image encoded by $DMD$ and random basis

RNG

DMD

PD

A/D

Compressed, encoded image data sent via RF for reconstruction

Xmtr

Rcvr

DSP

$1, 2, \ldots, M$
Single Pixel Camera

Potential for:

- **new modalities** beyond what can be sensed by CCD or CMOS imagers
- **low cost**
- **low power**
First Image Acquisition

ideal 128x128 pixels

image at DMD array

6x sub-Nyquist
Second Image Acquisition

8x sub-Nyquist
World’s First Photograph

- 1826, Joseph Niepce
- Farm buildings and sky
- 8 hour exposure
- On display at UT-Austin
Analog-to-Digital Conversion

• Many applications – particularly in RF – have hit an A/D performance **brick wall**
  - limited bandwidth (# Hz)
  - limited dynamic range (# bits)
  - deluge of bits to process downstream

• “Moore’s Law” for A/D’s: doubling in performance only every 6 years
  - “analog-to-information” conversion
    - analog CS
A2I via Random Demodulation

- Leverage extant spread spectrum and UWB concepts and hardware
- Successfully simulated at 6x sub-Nyquist
CS Hallmarks

• CS changes the rules of the data acquisition game
  – exploits a priori signal sparsity information
  – slogan: “sample less, compute more”

• **Universal**
  – same random projections / hardware can be used for any compressible signal class (generic)

• **Democratic**
  – each measurement carries the same amount of information
  – simple encoding
  – robust to measurement loss and quantization

• **Asymmetrical** (most processing at decoder)

• Random projections weakly encrypted
Carne

Distributed Compressive Sensing
Sensor Networks

- Measurement, monitoring, tracking of *distributed physical phenomena* ("macroscope") using wireless embedded sensors
  - environmental conditions
  - industrial monitoring
  - chemicals
  - weather
  - sounds
  - vibrations
  - seismic
  - wildfires
  - pollutants
  ...

![Map of sensor network with red dots indicating sensor locations]
Sensor Networks

- Measurement, monitoring, tracking of *distributed physical phenomena* (“macroscope”) using wireless embedded sensors
  - environmental conditions
  - industrial monitoring
  - chemicals
  - weather
  - sounds
  - vibrations
  - seismic
  - wildfires
  - pollutants
  ...

E. Charbon, M. Vetterli, EPFL
Sensor Networks

- Measurement, monitoring, tracking of distributed physical phenomena (“macroscope”) using wireless embedded sensors
  - environmental conditions
  - industrial monitoring
  - chemicals
  - weather
  - sounds
  - vibrations
  - seismic
  - wildfires
  - pollutants
  ...
Sensor Networks

- Measurement, monitoring, tracking of *distributed physical phenomena* ("macroscope") using wireless embedded sensors
  - environmental conditions
  - industrial monitoring
  - chemicals
  - weather
  - sounds

![Diagram of sensor network with data and light nodes connected to a fusion center](image-url)
New Hardware, Software

- **Hardware platforms**
  - sensing, DSP, networking, communications, power
  - comm standards: 802.15.4 (Zigbee), Bluetooth, ...
  - Crossbow motes
  - Berkeley motes
  - Smart Dust
  - MoteIV
  - Rice Gnomes
  - ...

- **Operating systems**
  - TinyOS
  - MagnetOS
  - SOS
  - Pumpkin
  - ...
Challenges

- **Computational/power asymmetry**
  - limited compute power on each sensor node
  - limited (battery) power on each sensor node

- Must be *energy efficient*
  - minimize communication

- Hostile *communication* environment
  - multi-hop
  - high loss rate
Distributed Sensing

- Transmitting raw data can be inefficient
Correlations

- Can we exploit *intra-sensor* and *inter-sensor* correlation to *jointly compress? jointly process?*
Collaborative Sensing

- Output *results* rather than *raw data*
- In-network data processing
Collaborative Sensing

- Output *results* rather than *raw data*
- In-network data processing
- Collaboration introduces
  - inter-sensor *communication overhead*
  - complexity at sensors
Independent Compressive Sensing

- Take incoherent measurements at each sensor
- Reconstruct *individually*
- Exploit *intra-sensor* correlations
Joint Compressive Sensing

- Take incoherent measurements at each sensor
- Reconstruct *jointly*
- Exploit *intra- & inter-sensor* correlations
- Zero communication overhead
- Any communication protocol
- Analogy w/ Slepian-Wolf coding
Common Sparse Supports Model

Ex: audio signals

- sparse in Fourier Domain
- same frequencies received by each node
- different attenuations and delays (magnitudes and phases)
Common Sparse Supports Model

• Measure $J$ signals, each $K$-sparse

• \textit{Signals share sparse components but with different coefficients}

\[ x_j = \sum_{\omega \in \Omega} x_{j,\omega} \psi_\omega, \]
\[ |\Omega| = K \]
Common Sparse Supports Model

\[ y_1 = \Phi_1 \times x_1 \]

\[ y_2 = \Phi_2 \times x_2 \]

\[ y_3 = \Phi_3 \times x_3 \]

\[ \ldots \]

\[ y_J = \Phi_J \times x_J \]
Ensemble Reconstruction Comparison

- Separate reconstruction using linear programming
  - measurements per sensor: $O(K \log(N/K))$

- Simultaneous Orthogonal Matching Pursuit (SOMP)
  - extends greedy algorithms to signal ensembles sharing a sparse support
    [Tropp, Gilbert, Strauss; Temlyakov]
  - measurements per sensor: $cK$

$$\lim_{J \to \infty} c = 1$$
Simulation

$K=5$
$N=50$

Separate Joint

Probability of Exact Reconstruction vs. Number of Measurements per Signal, $M$
Real Data Example

- Environmental Sensing in Intel Berkeley Lab
  - $J = 49$ sensors, $N = 1024$ samples each

- Compare:
  - transform coding approx $K$ largest terms per sensor
  - independent CS $4K$ measurements per sensor
  - DCS $4K$ measurements per sensor
Light Intensity – Wavelets, $K = 100$

(a) Original

(b) Transform Coding, SNR = 26.4842 dB

(c) Compressed Sensing, SNR = 21.6426 dB

(d) Distributed Compressed Sensing, SNR = 27.1906 dB
Temperature – Wavelets, $K = 20$

(a) Original

(b) Transform Coding, SNR = 25.9499 dB

(c) Compressed Sensing, SNR = 16.8255 dB

(d) Distributed Compressed Sensing, SNR = 29.4149 dB
DCS Benefits

• Random projections for *sensing and encoding*
  – exploit both intra- and inter-sensor correlations
  – *joint source/channel coding*

• Universality
  – generic hardware
  – “future-proof”

• Simple quantization

• Robust
  – to noise, quantization, loss
  – progressive

• Zero inter-sensor collaboration
Dessert

Conclusions
Conclusions

• **Compressive sensing**
  - exploits signal sparsity/compressibility information
  - based on new uncertainty principles
  - integrates sensing, compression, processing
  - natural for sensor network applications

• Ongoing research
  - new kinds of *cameras* and *imaging* algorithms
  - new “*analog-to-information*” converters (analog CS)
  - new algs for *distributed source coding* (Slepian-Wolf)
    (sensor nets content distribution nets)
  - *fast algorithms* based on LDPC code matrices and BP
  - *R/D* analysis of CS (quantization)
  - CS meets *Johnson-Lindenstrauss*
  - manifold CS for multiple signals/images

dsp.rice.edu/cs