Compressed Sensing: Challenges and Emerging Topics

Mike Davies
Edinburgh University Defence Research Collaboration (UDRC)
Edinburgh Compressed Sensing research group (E-CoS)

University of Edinburgh
Compressed sensing

Engineering Challenges in CS:

• What is the right signal model?
  
  Sometimes obvious, sometimes not. When can we exploit additional structure?

• How can/should we sample?
  
  Physical constraints; can we sample randomly; effects of noise; exploiting structure; how many measurements?

• What are our application goals?
  
  Reconstruction? Detection? Estimation?
CS today – the hype!

Papers published in Sparse Representations and CS [Elad 2012]

Lots of papers..... lots of excitement.... lots of hype....

Google About 4,500,000 results (0.37 seconds)
CS today: - new directions & challenges

There are many new emerging directions in CS and many challenges that have to be tackled.

- Fundamental limits in CS
- Structured sensing matrices
- Advanced signal models
- Data driven dictionaries
- Effects of quantization
- Continuous (off the grid) CS
- Computationally efficient solutions
- Compressive signal processing
Compressibility and Noise Robustness
Noise Robustness

CS is robust to measurement noise:

\[ \text{RIP} \Rightarrow \| \Delta (\Phi x + \epsilon) - x \|_2 \leq C_1 \sigma_k(x) + C_2 \| \epsilon \|_2 \]

What about signal errors: \( \Phi(x + e) = y \)?

No free lunch!

- Detecting signals through wide band receiver noise: noise folding!
  - 3dB SNR loss per 2 undersampling [Treichler, et al 2011]
- Imaging: “tail folding” of coefficient distribution
  - Fundamental limits in terms of the compressibility of the pdf [D. and Guo. 2011, Gribonval et al 2012]
SNR/SDR vs Undersampling

Wideband spectral sensing

Bounds guide sampling strategies and provide fundamental limits to CS imaging performance

Adaptive sensing can retrieve lost SNR [Haupt et al 2011]

Trade-off – better dynamic range for SNR loss [Treichler, et al 2011]

Compressible distributions
Sensing matrices
Structured CS sensing matrices

i.i.d. sensing matrices are really only of academic interest. Need to consider wider classes, e.g.:

- Random rows of DFT [Rudelson & Vershynin 2008]

\[ m \sim \mathcal{O}(k \delta^{-2} \log^4 N) \]
Structured CS sensing matrices

i.i.d. sensing matrices are really only of academic interest. Need to consider wider classes, e.g.:

- Random samples of a bounded orthogonal system [Rauhut 2010]

Also extends to continuous domain signals.

$\delta$-RIP of order $k$ with high probability if:

$$m \sim \mathcal{O}(k \mu(\Phi, \Psi)^2 \delta^{-2} \log^4 N)$$

where $\mu(\Phi, \Psi) = \max_{1 \leq i < j \leq N} |\langle \Phi_i, \Psi_j \rangle|$ is called the mutual coherence.
Structured CS sensing matrices

i.i.d. sensing matrices are really only of academic interest. Need to consider wider classes, e.g.:

- Universal Spread Spectrum sensing [Puy et al 2012]

Sensing matrix is random modulation followed by partial Fourier matrix. $\delta$-RIP of order $k$ with high probability if:

$$m \sim \Theta(k \delta^{-2} \log^5 N)$$

*Independent of basis $\Psi$!*
Generalized Dimension Reduction

Compressed sensing matrices can be used to preserve information beyond sparsity. Robust embeddings (RIP for difference vectors):

\[(1 - \delta) \|x - x'\|_2 \leq \|\Phi(x - x')\|_2 \leq (1 + \delta) \|x - x'\|_2\]

hold for many low dimensional sets.

- Sets of \(n\) points [Johnston and Lindenstrauss 1984]
  \[m \sim O(\delta^{-2} \log n)\]

- Arbitrary Union of \(L\) \(k\)-dimensional subspaces [Blumensath and D. 2009]
  \[m \sim O(\delta^{-2} (k + \log L))\]

- Set of \(r\)-rank \(n \times l\) matrices [Recht et al 2010]
  \[m \sim O(\delta^{-2} r(n + l) \log nl)\]

- \(d\)-dimensional affine subspaces [Sarlos 2006]
  \[m \sim O(\delta^{-2} d)\]

- \(d\)-dimensional manifolds [Baraniuk and Wakin 2006, Clarkson 2008]
  \[m \sim O(\delta^{-2} d)\]
Advanced signal models & algorithms
CS with Low Dimensional Models

What about sensing with other low dimensional signal models?

- Matrix completion/rank minimization
- Phase retrieval
- Tree based sparse recovery
- Group/Joint Sparse recovery
- Manifold recovery

... towards a general model-based CS?
[Baraniuk et al 2010, Blumensath 2011]
Matrix Completion/Rank minimization

Retrieve the unknown matrix $X \in \mathbb{R}^{N \times L}$ from a set of linear observations

$$y = \Phi(X), \ y \in \mathbb{R}^{m} \text{ with } m < NL.$$ 

Suppose that $X$ is rank $r$.

Relax!

as with $L_1$ min., we convexify: replace $\text{rank}(X)$ with the nuclear norm

$\|X\|_* = \sum_i \sigma_i$, where $\sigma_i$ are the singular values of $X$.

$$\hat{X} = \arg\min_X \|X\|_* \text{ subject to } \Phi(X) = y$$

Random measurements (RIP) $\rightarrow$ successful recovery if

$$m \sim O(r(N + L) \log NL)$$

e.g. the Netflix prize

– rate movies for individual viewers.
Phase retrieval

Generic problem:

Unknown $x \in \mathbb{C}^n$,
magnitude only observations: $y_i = |A_i x|^2$

Applications

- X-ray crystallography
- Diffraction imaging
- Spectrogram inversion

Phaselift

Lift quadratic $\longrightarrow$ linear problem using rank-1 matrix $X = xx^H$

Solve: $\hat{X} = \arg\min_X \|X\|_* \text{ subject to } A(X) = y$

Provable performance but lifting space is huge!

… surely more efficient solutions?
Tree Structured Sparse Representations

Sparse signal models are type of "union of subspaces" model [Lu & Do 2008, Blumensath & Davies 2009] with an exponential number of subspaces.

\[ \# \text{ subspaces} \approx \left( \frac{N}{k} \right)^k \]  
(Stirling approx.)

Tree structure sparse sets have far fewer subspaces

\[ \# \text{ subspaces} \approx \frac{(2e)^k}{k+1} \]  
(Catalan numbers)

**Example** exploiting wavelet tree structures

Classical compressed sensing: stable inverses exist when

\[ m \sim \mathcal{O}(k \log(N/k)) \]

With tree-structured sparsity we only need [Blumensath & D. 2009]

\[ m \sim \mathcal{O}(k) \]
Algorithms for model-based recovery

Baraniuk et al. [2010] adapted CoSaMP & IHT to construct provably good model-based recovery algorithms.

Blumensath [2011] adapted IHT to reconstruct any low dimensional model from RIP-based CS measurements:

\[
x^{n+1} = P_\mathcal{A}(x^n + \mu \Phi^T(y - \Phi x^n))
\]

where \(\mu \sim N/m\) is the step size, \(P_\mathcal{A}\) is the projection onto the signal model.

Requires a computationally efficient \(P_\mathcal{A}\) operator.
Sparse Multiple Measurement Vector Problem

Find a row sparse matrix \( X \in \mathbb{R}^{N \times l} \) given the multiple measurements \( Y \in \mathbb{R}^{m \times l} \). Applications such as:

- multiband spectral sensing,
- hyperspectral imaging

\( L_0 \) solution:

\[
\hat{X} = \arg\min_X |\text{row supp}(X)| \quad \text{s.t.} \quad \Phi X = Y
\]

Difficulty of inverse problem depends on rank of \( Y \) [Eldar and D. 2013]

When \( \text{rank}(Y) = k \) the problem can be solved with the MUSIC algorithm.
Compressed Signal Processing
Compressed Signal Processing

There is more to life than signal reconstruction:

- Detection
- Classification
- Estimation
- Source separation

May not wish to work in large ambient signal space, e.g. ARGUS-IS Gigapixel camera

CS measurements can be information preserving (RIP)... offers the possibility to do all your DSP in the compressed domain!

Without reconstruction what replaces Nyquist?

\[ H_0 : y = \Phi n \]
\[ H_1 : y = \Phi (s + n) \]
Compressive Detection

The Matched Smashed Filter [Davenport et al 2007]

Detection can be posed as the following hypothesis test:

\[ \mathcal{H}_0 : z = hn \]
\[ \mathcal{H}_1 : z = h(s + n) \]

The optimal (in Gaussian noise) matched filter is \( h = s^H \)

Given CS measurements: \( y = \Phi s \), the matched filter (applied to \( y \)) is:

\[ h = s^H \Phi (\Phi \Phi^H)^{-1} \]

Then

\[ P_D \approx Q \left( Q^{-1}(\alpha) - \sqrt{\frac{m}{N}} \sqrt{\text{SNR}} \right) \]

\( Q \) - the Q-function, \( \alpha \) – Prob. false alarm rate

[SNR=20dB] [Davenport et al 2010]
Joint Recovery and Calibration

Estimation and recovery, e.g. on-line calibration.

Compressed Calibration

Real Systems often have unknown parameters $\theta$ that need to be estimated as part of signal reconstruction.

$$ y = \Phi(\theta)x $$

Can we simultaneously estimate $x$ and $\theta$?

Example – Autofocus in SAR

Imperfect estimation of scene centre leads to phase errors, $\phi$:

$$ Y = \text{diag}(e^{j\phi})h(X) $$

$X$ - scene reflectivity matrix, $Y$ - observed phase histories, $h(\cdot)$ - sensing operator.

Uniqueness conditions from dictionary learning theory [Kelly et al. 2012].
Joint Recovery and Calibration

Compressed Autofocus:

Perform joint estimation and reconstruction:

\[
\min_{X, d} \|X\|_1 \quad \text{subject to} \quad \|Y - \text{diag}(d) h(X)\|_F \leq \epsilon
\]

and \( d_i d_i^* = 1, i = 1, \ldots, N \)

- Fast alternating optimization schemes available
- Provable performance? Open

No phase correction | Post-recon. autofocus | Compressive autofocus
Summary

Compressive Sensing (CS)
- combines sensing, compression, processing
- exploits low dimensional signal models and incoherent sensing strategies

Still lots to do…
- Developing new and better model-based CS algorithms and acquisition systems
- New emerging field of compressive signal processing
- Exploit dimension reduction in computation: randomized linear algebra,… big data!
References

Compressibility and SNR loss


Structured Sensing matrices


References

Information Preserving Dimension Reduction


Structured Sparsity & Model-based CS


References

Compressed Signal Processing

