CTG Reading Group

Electrical and Computer Engineering

A discussion on "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?" by Candes and Tao (IEEE Trans. IT, Dec 2006).

Anand Oka (Ph.D. Candidate)

29 February 2008

Outline

- Motivation

Questions that this paper answers:

- What is Compressed Sensing?
- How does it work? (answer is easy)
- Why does it work? (answer is subtle)
- Some revolutionary applications

Motivation

- Consider example of a digital camera ...
- Tens of mega-pixels of *raw data* is sampled. Then *transform coded* to say 10 kilo-pixels. Thus most of the raw data is "thrown out" ...
- The optimal transform *depends on type of scene*.
- What if we directly measure relevant linear functionals, perhaps via *analog processing*? Bonus: make the transform universal!
- Result: extremely sensitive SINGLE-PIXEL camera!

anando@ece.ubc.ca

What is CS?

- An efficient way to encode and reconstruct *sparse* signals.
- *Universal Encoder!* Only the decoder needs to know the *sparsity basis*.
- Reconstruction algorithm is *tractable*.
- *Robust* to erasures or errors in encoded data.

anando@ece.ubc.ca

How does it work?

$\mathcal{F} \subset \mathbb{R}^N$ is said to be a sparse family with sparsity $|T|$ in the basis formed by the columns of an $N \times N$ orthonormal matrix Φ if

$$\forall f \in \mathcal{F}, \; ||\Phi^T f||_0 = |T|.$$

Let F_Ω be a $K \times N$ *random mesasurement marix* with i.i.d entries with bounded variance, say $\mathcal{N}(0, 1/N)$.

Let K measurements be obtained for a generic $f \in \mathcal{F}$ as $y = F_\Omega f$.

Let the reconstruction be made by the *basis pursuit* linear program:

$$\hat{f} = \arg \min_{g : F_\Omega g = y} ||\Phi^T g||_1$$

Discussion: Candes-Tao'06
anando@ece.ubc.ca

Guarantees

If

$$|T| \leq \alpha \frac{K}{\log N}$$

where α is a positive constant independent of f, then

$$\hat{f} = f$$

with probability $1 - O(N^{-\rho/\alpha})$, where $\rho > 0$ is a universal constant.

Paraphrase: With a small oversampling factor we can achieve exact reconstruction with high probability.

Discussion: Candes-Tao'06
anando@ece.ubc.ca

What if strict sparsity is not satisfied by signals?

Suppose signals satisfy a power low decay

$$|(\Phi^T f)_n| \leq R\, n^{-1/p}$$

where R, p are some positive parameters.

Then with probability $1 - O(N^{-\rho/\alpha})$ the reconstruction MSE satisfies

$$||f - \hat{f}||_2 \leq C_{p,\alpha}\, R \left(\frac{K}{\log N}\right)^{-r}$$

where $r = 1/p - 1/2$.

Paraphrase: With a small oversampling factor the MSE decays just as in transform coding.

Discussion: Candes-Tao'06
anando@ece.ubc.ca

How does it work? Heuristics ...

- Close analogy to the technique of *holography*.
- Main requirement: sparsity basis should be *incoherent* w.r.t. the measurement ensemble. (Thats weird!)
- The energy of each sparsity basis element should be more or less evenly distributed in all linear measurement functionals.
- CS works because a random measurement ensemble is *universally decoherent w.r.t. any sparsity basis*.

anando@ece.ubc.ca

How does it work?
The math (theorem 1.2) ...

Recall *uncertainty principle* from signal processing: A signal with a small support in time must necessarily have a wide frequency support.
While this holds *automatically* for time-frequency, we can *axiomatize* this property:

F_Ω is said to satisfy the **Uniform Uncertainty Principle** if, with probability $1 - O(N^{-\rho/\alpha})$,

$$|T| \leq \alpha \frac{K}{\lambda} \Rightarrow \frac{1}{2}\frac{K}{N} \leq \lambda_{min}(F_\Omega F_\Omega^T) \leq \lambda_{max}(F_\Omega F_\Omega^T) \leq \frac{3}{2}\frac{K}{N}$$

Discussion: Candes-Tao'06
anando@ece.ubc.ca

How it works? (contd.)

Another related axiom is the Exact Reconstruction Property (ERP). In some cases, but not always, ERP is implied by UUP.

Paraphrase of Theorem 1.2:
CS "works" (in the sense described earlier) for signal families with power law decay provided the measurement ensemble satisfies the UUP and ERP properties.

So remaining job is to prove that a random ensemble does satisfiy UUP and ERP.

Discussion: Candes-Tao'06
anando@ece.ubc.ca

How it works? (contd.)

The good news (Lemma 4.1- 4.3): Random ensembles (Gaussian, binary and others) *do satisfy* UUP for *any sparsity basis*.

Reason: Marchenko-Pastur law

The limiting density (as $K \to \infty, K/N \to \beta$) of eigen values of a $K \times K$ S.P.D. matrix $F_\Omega F_\Omega^T$, where entries of the $K \times N$ random matrix F_Ω are i.i.d. with variance $1/N$, is given by

$$f_\beta(x) = \frac{\sqrt{(x-a)(b-x)}}{2\pi x}$$

on support $[a, b]$ and is identically zero elsewhere, where $a = (1 - \sqrt{\beta})^2$ and $a = (1 + \sqrt{\beta})^2$.

Discussion: Candes-Tao'06
anando@ece.ubc.ca

Some interesting open questions ...

- Will UUP and ERP hold for carefully selected non-random (deterministic) measurement ensembles?
- Will dependencies in ensemble be catastrophic?
- How much can the oversampling factor be reduced if we have knowledge of the *locations* of sparse entries?
- Most important: What is the relation of CS to information theoretic source compression? What about fountain encoders?
- Most important: Can we *prove* uniform robustness to errors/erasures of measurements?

anando@ece.ubc.ca

Some revolutionary applications

- Extremely Sensitive but universal imaging and detection, e.g. in medicine, astromomy etc
(By hugely reducing number of sensors, we can make each sensor ultra sensitive.)
- Data extraction from Wireless Sensor Networks
- Universal and Encrypted compression
- Reliable Micro Array Analysis of gene expression

anando@ece.ubc.ca

Prominent researchers ...

- Candes, Tao, Romberg (Caltech, UCLA)
- Donoho (Stanford)
- Baranuik (Rice)
- Also many from EE and IT field, e.g. Tarokh (Harvard)

Thank You ...!

Discussion: Candes-Tao'06
anando@ece.ubc.ca