Magic, Syphilis and C.S.I

A friendly introduction to compressive sensing.

Rhian Davies

November 29, 2013

EPSRC
Engineering and Physical Sciences
Research Council

STOR-i
excellence with impact

2	3	6	7	10	11	14	15
18	19	22	23	26	27	30	31
34	35	38	39	42	43	46	47
50	51	54	55	58	59	62	63

16	17	18	19	20	21	22	23
24	25	26	27	28	29	30	31
48	49	50	51	52	53	54	55
56	57	58	59	60	61	62	63

8	9	10	11	12	13	14	15
24	25	26	27	28	29	30	31
40	41	42	43	44	45	46	47
56	57	58	59	60	61	62	63

1	3	5	7	9	11	13	15
17	19	21	23	25	27	29	31
33	35	37	39	41	43	45	47
49	51	53	55	57	59	61	63

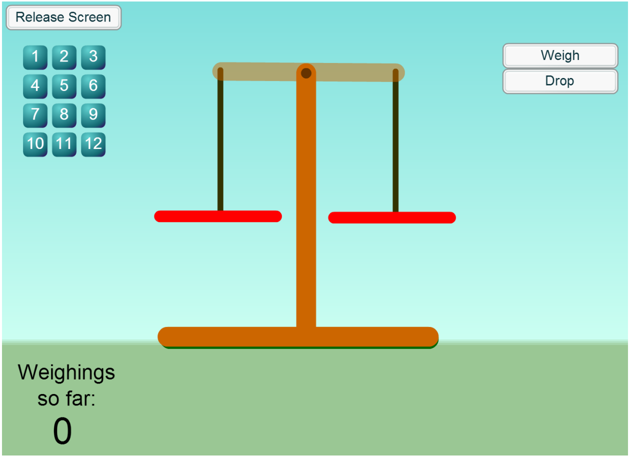
4	5	6	7	12	13	14	15
20	21	22	23	28	29	30	31
36	37	38	39	44	45	46	47
52	53	54	55	60	61	62	63

32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47
48	49	50	51	52	53	54	55
56	57	58	59	60	61	62	63

Syphilis in World War II

- US PHS study WWII
- Rob Dorfman "The detection of defective members of large populations." 1943
- We can combine M blood samples together and test a combined sample to see if at least one recruit in the sample has syphilis.
- If negative we have "saved" $M - 1$ tests.
- If positive, we have "wasted" a test.

The coin puzzle



Non-adaptive solution

1	2	3	4	vs	5	6	7	8
1	4	8	9	vs	2	3	11	12
3	7	9	12	vs	1	2	5	10

What do these group testing examples have in common?

- Sparsity
- Testing in subsets / Reduced number of measurements
- Non adaptive measurements
- Decoding procedure

Compressive Sensing

- $M \times N$ measurement matrix Φ ($M \ll N$)
- Signal $x \in R^N$ which is sparse (contains k nonzero entries).
- Identify the location of k elements using $y = \Phi x$ measurements
- How small can we make M and still recover x using only y?

The encoding process.

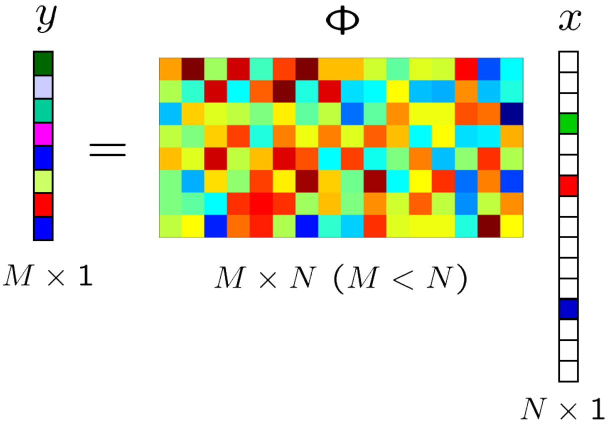


Figure: CS measurement process, courtesy of Volkan Cevher.

Measurement Matrix Φ: Null space condition

We require that the signal $\mathbf{x}$ can be uniquely reconstructed. It is possible to show that this will hold true if the null space $\mathcal{N}(\mathbf{\Phi})$ does not contain any vectors in Σ_{2K}.

$$\mathcal{N}(\mathbf{\Phi}) = (\mathbf{z} : \mathbf{\Phi z} = 0).$$

In order to preserve $\mathbf{x} \in \Sigma_K$, it is required $\mathbf{\Phi x} \neq \mathbf{\Phi x'}$ $\forall \mathbf{x'} \in \Sigma_K$, because if $\mathbf{\Phi x} = \mathbf{\Phi x'}$ it would be impossible to distinguish $\mathbf{x}$ from $\mathbf{x'}$ based only on $\mathbf{y}$.

$$\begin{aligned} \text{Consider,} \qquad \mathbf{\Phi x} &= \mathbf{\Phi x'} \\ &\Rightarrow \mathbf{\Phi}(\mathbf{x} - \mathbf{x'}) = 0 \\ &\Rightarrow (\mathbf{x} - \mathbf{x'}) \in \Sigma_{2K} \end{aligned}$$

$\mathbf{\Phi}$ uniquely represents all $\mathbf{x} \in \Sigma_K \iff v \notin \mathcal{N}(\mathbf{\Phi})\ \forall \mathbf{v} \in \Sigma_{2K}$.

Measurement Matrix Φ: Restricted Isometry Property

- $\mathbf{y}$ may be corrupted by noise during the measurement process.
- Matrix $\mathbf{\Phi}$ satisfies the (RIP) of order K if there exists a $\delta_K \in (0,1)$ such that
$$(1-\delta_k)||\mathbf{x}||_2^2 \leq ||\mathbf{\Phi x}||_2^2 \leq (1+\delta_k)||\mathbf{x}||_2^2,$$
for all $\mathbf{x} \in \sum_K = \mathbf{x} : ||\mathbf{x}||_0 \leq K$.
- $\mathbf{\Phi}$ preserves the distance between any pair of K-sparse vectors.
- This gives a stronger guarantee of robustness against noise.
- Both of these conditions will hold true with high probability if $\mathbf{\Phi}$ is selected as a random matrix.

Recovery of sparse transforms

- Solve $\mathbf{y} = \mathbf{\Phi x}$, infinitely many solutions! Fat $\mathbf{\Phi}$ implies underdetermined system.
- We know that $\mathbf{x}$ was **sparse**
- What algorithms can we use to decode?
- Convex Optimisation or Greedy Algorithms or something else...?

ℓ_1 minimisation | Orthogonal Matching Pursuit

Traditional Image Acquisition

We cannot use compressive sensing with this camera!

The single pixel camera.

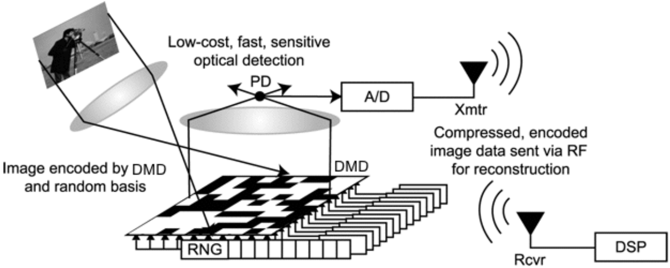
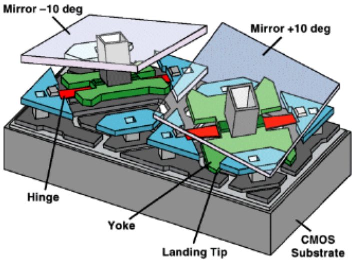


Figure: The single pixel camera

Doesn't acquire a single ray of light per pixel but rather a combination of rays of light (each coming from a different direction or spectral band or both) per pixel. To obtain back a picture that can be understood by the human eye, one needs the reconstruction methods mentioned earlier.

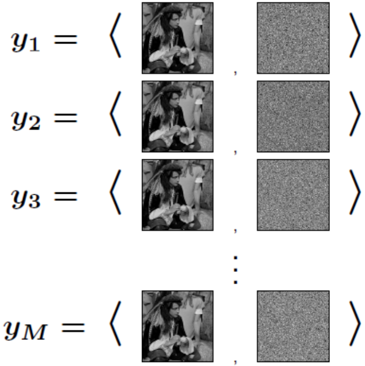
Digital micro-mirror device



- Many very tiny tilt-able mirrors.
- Each mirror can be positioned in two states.
- A random number generator modulates the positions.
- Therefore the light, can be reflected in two distinct directions.

Image Acquisition

- Mathematically - calculating inner products
- Each set of mirror orientations = one measurement.
- Repeat M times.
- Therefore SPC compresses and samples in the measurement process.



Results from SPC

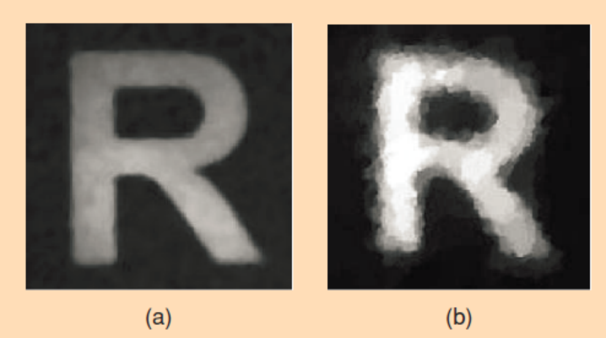


Figure: Reconstructed image taken with the SPC. (a) Conventional image of the target scene. (b) Reconstructed image with M = 1300 measurements.

Why bother?

- Data storage is cheap.
- We can store the information, so why bother?
- What about a fixed sensors on the moon?
- Scenarios where we need simplicity at the sensor, complexity at the analysis hub.

Compressive Sensing is not...

Any Questions?

Sparsity and wavelet transformation

Achievable resolution is dependant on the information content of the image. If an image has low information content it is said to be sparse and can be perfectly reconstructed from a small number of measurements. **Nearly all real world images exhibit this sparsity property when transformed using a wavelet basis.**

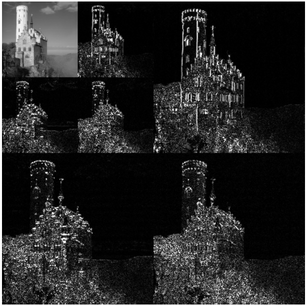


Figure: Wavelet transform

Using ℓ_1 minimization to promote sparsity

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{N} |x_i|$$

Originally used in geophysics to aid detection of sparse spike trends in earthquake data, optimisation based on the l_1 norm can closely approximate compressible signals with high probability.

$$\min_{\mathbf{x}} ||\mathbf{x}||_1 \text{ subject to } \mathbf{y} = \mathbf{\Phi}\mathbf{x}.$$

▸ Recovery Algorithms

Orthogonal Matching Pursuit

Define the columns of $\mathbf{\Phi}$ to be $\varphi_1, \varphi_2, \ldots, \varphi_N$.

Require: $\mathbf{r_0} = \mathbf{y}, \Lambda_0 = \emptyset$ and iteration counter $i = 1$

for $i < T$ **do**

$\lambda_t = \text{argmax}_{\ j=1,\ldots,N} | < r_{t-1}, \varphi_j > |$

{Find the index for the column of Φ with the greatest contribution.}

$\Lambda_t = \Lambda_{t-1} \cup \lambda_t,\ \Phi_t = [\mathbf{\Phi_{t-1}}, \varphi_{\lambda_t}]$

{Keeps track of the columns used.}

$\mathbf{x_t} = \text{argmin}_{\mathbf{x}} ||\mathbf{y} - \mathbf{\Phi_t x}||_2$

{Updates the signal estimate.}

$\mathbf{r_t} = \mathbf{y} - \mathbf{\Phi_t x_t}$

{Updates the measurement residual.}

end for

return $\hat{\mathbf{x}}$