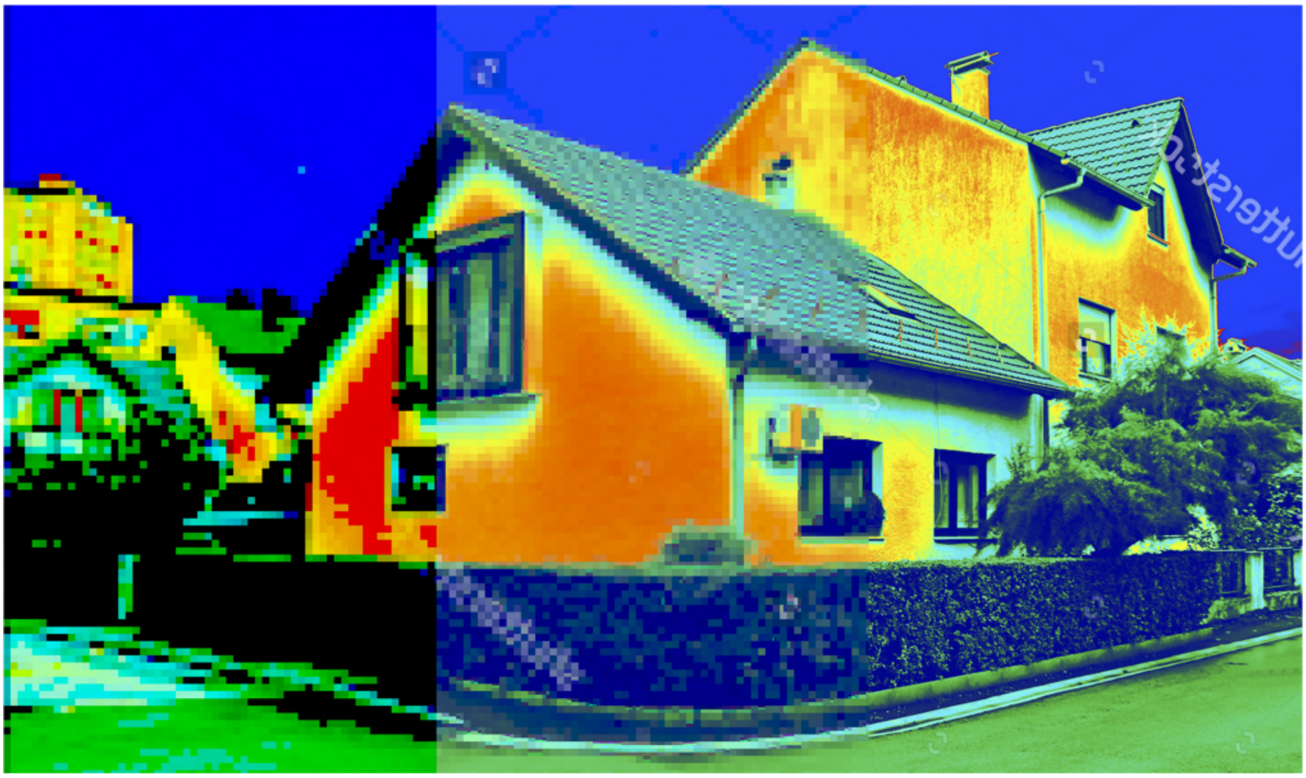
EPSRC
Engineering and Physical Sciences Research Council

InFoMM
Industrially Focused Mathematical Modelling

EPSRC Centre for Doctoral Training in

Industrially Focused Mathematical Modelling



Sparse Matrices for Compressed Sensing in the Presence of Noise

Florian Wechsung

UNIVERSITY OF OXFORD

PA

Contents

1. Introduction

Background

Compressed Sensing is a novel method of recovering signals that has received substantial attention from researchers and professionals in Mathematics, Computer Science, Signal Processing and other fields ever since its introduction in [1] and [2] just over a decade ago. The classical paradigm in Signal Processing is given by the Nyquist-Shannon Sampling-Theorem that has been known since the middle of the 20th century.

The Sampling Theorem states that a bandlimited signal can be perfectly reconstructed if it is sampled at a frequency that is at least twice the bandwidth. For an audio signal this means that, if we expect frequencies within the bandlimit of 1Hz and 20000Hz, we need to take at least 40000 measurements per second to recover the full signal. The idea of encoding algorithms like MP3 is that a typical audio signal only has few significant frequencies and that, by only storing these frequencies and their magnitude, a very large compression ratio with negligible losses in quality can be achieved.

A key question that PA Consulting is interested in is whether it is necessary to take a complete set of measurements in the first place if the data is so severely reduced during the compression stage. This problem of redundant data capture also arises in photography and videography; modern image sensors have a resolution of many millions of pixels and therefore there are vast amounts of data associated with taking a picture or video. As a simple example, consider that a 8 Megapixel, 16bit greyscale image results in file sizes of approximately 16 Megabytes and up to a Gigabyte per second for video in colour. Again, encoding algorithms like JPEG or MPEG use sparsity in some basis to reduce the file size by multiple orders of magnitude. With that prior knowledge it may be desirable to avoid acquiring the redundant data, and instead measure the image in a way that allowed the number of measurements needed to be proportional to the information content in a suitable compression format for the given application.

An example where reducing the number of measurements is extremely beneficial is in medical imaging; here reducing the number of measurements taken means reducing the amount of time necessary for a CT or MRI scan and thus helps to keep the radiation that the patient is exposed to at a low level and allows more patients to be scanned in a given time.

Compressed Sensing

Compressed Sensing is a new technology that can be used to reduce the number of measurements to recover a signal.

The power of Compressed Sensing is that it provides a way of measuring information where the number of measurements is only proportional to the actual information content of the signal and not proportional to the bandwidth. Instead of sampling the actual signal at different times (for audio) or different locations (for image) one measures instead different linear combinations of the signal. For a signal x consisting of n entries (e.g. pixels) this means measuring $y_i = a_1x_1 + a_2x_2 + \cdots + a_nx_n$ multiple times and writing the coefficients as rows of a matrix A this is equivalent to measuring $y = Ax$. The goal is now to reconstruct the original values of x from the measured linear combinations y.

However, since the goal is to use fewer measurements than n, we cannot expect to be able to recover our signal, as there are more unknowns in x than measurements of y unless we have extra information. In the case of Compressed Sensing, we assume that the signal x has a lot of redundant information, i.e. that it is sparse in some basis. This means that almost all of the information of the signal is contained in only $k \ll n$ coefficients. If we know what this basis is, then instead of solving for n unknowns x_j we only need to solve for k unknowns. Since k is much smaller than n, we can expect that that is possible from far fewer measurements.

To give an example: A photograph with a resolution of 1 Megapixel is usually represented very well by the 25,000 biggest Wavelet coefficients. If we could measure linear combinations of the individual pixels we can hope to recover the picture from only about 100,000 measurements, which is one order of magnitude less than using all the pixels and would mean a drastic reduction of data rates or of radiation in the case of a CT scanner.

Single Pixel Camera

After having illustrated the potential benefits of using Compressed Sensing, we now describe the essential components of a camera that would utilize Compressed Sending. A *Single Pixel Camera* as seen in Figure 1 consists of two lenses, a photodiode, and a Digital Mirror Device (DMD). The DMD is an array of many (n) small mirrors, that can individually be pointed in one of two directions. Light falls from the object being photographed onto a lens and is focussed on the DMD. Some of the mirrors are pointed towards the photodiode and some are pointed away from it, so the photodiode (sensor) will see the sum of those parts of the picture with mirrors pointing at it. For a given set of mirror configurations on the DMD this gives one measurement from the photodiode. The mirror configurations are then changed to get a second measurement. These DMDs are able to switch patterns extremely quickly and hence many measurements can be made in a small time. The DMD has n mirrors and their configurations, represented by zeros and ones, give the as in y. If m measurements are then made the matrix A is of size m by n.

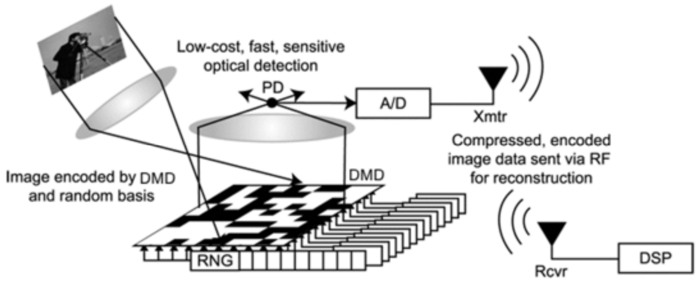


Figure 1: A Single Pixel Camera; image source: http://dsp.rice.edu/cscamera.

One of the main advantages of this approach is that the number of measurements is significantly reduced. This means that data rates from the sensor to the processing chips are lower and hence implies lower power consumption. The data also does not need to be compressed further, as no JPEG transform or anything similar is necessary, instead the image is already sensed in the compressed domain.

Single Pixel Cameras measure the image in a compressed way – thus removing the need for further image compression afterwards

Another main advantage is that since only one photo sensor is necessary; this sensor can easily be replaced by a low cost infrared sensor or ultraviolet sensor, whereas high resolution image sensors for these wavelengths are extremely expensive or not even possible to build. This enables low cost thermal and night imaging and is one of the applications that we focus on in this report.

There are, however, two disadvantages to this approach: firstly, the resolution and the framerate is limited by the number of mirrors on the DMD and by the speed at which the DMD can switch its patterns. Secondly, though the process of taking the picture is now of reduced complexity, the reconstruction of x from the values of y using Compressed Sensing is computationally expensive. The second issue can be addressed by developing faster and more efficient reconstruction algorithms and this is the main purpose of this project.

2. L0-Decoding for Sparse Compressed Sensing

Central to the Compressed Sensing method is choosing suitable values for the entries in A, i.e. in our case how the DMD mirrors should be configured. Classically, the entries in A are either drawn randomly, which implies that the matrix A is "dense", i.e. a large number of entries are nonzero and hence storing these matrices and performing matrix multiplications is very expensive. Since the latter has to be done repeatedly during most recovery algorithms, this usually makes them quite slow.

Sparse matrices contain only few non-zero entries; this makes working with them much faster than working with dense matrices.

The natural step to rectify this is by using sparse binary matrices for A instead. Sparse matrices are matrices with very few non-zero entries and hence storing them and multiplying with them is much faster since all the zero entries can be skipped. However, having many zero entries in the matrix A means that some of the information in x might be lost when multiplying with A. To avoid this, we need our matrix A to have well distributed nonzero entries in order to spread the information from x well over y. Expander Matrices satisfy this property and so we focus only on matrices of this type.

For the case when A is an expander matrix and x is a strictly sparse signal, one can use the l0-decoding algorithm [3] to reconstruct the signal x from y. This algorithm is extremely fast, can be parallelized, and is able to reconstruct signals from few measurements. However, when x is only approximately sparse, or the measurements are subject to noise, this algorithm fails. We have taken the l0-decoding and developed a model for noisy measurements. The result is that we have created the "robust-l0-decoding" algorithm that can cope with additive Gaussian noise.

The difficulty of extending the l0-decoding algorithm to the noisy case can be illustrated as follows: if there is no noise and we want to solve $Ax = y$, then this is equivalent to finding x such that the residual $r = y - Ax$ is exactly zero. The l0-decoding does this by iteratively identifying entries in x that result in more zero entries in the residual. However, if there is noise added to our measurements y, then we can only hope to find x so that $r = Ax - y \approx 0$, hence as we iterate down the values in r, we need to decide whether the entry is small enough that it is just noise, or whether we need to continue to search for entries in x to further decrease this particular residual entry.

The model behind the algorithm assumes additive Gaussian noise and a sparse Gaussian signal; this allows us to use Bayes rule and a Normal approximation to obtain the probability whether the residual is just noise or whether it is significant. We call our algorithm the Bayesian Robust l0-decoding.

Comments

- Expander Matrices can be used to speed up recovery algorithms for Compressed Sensing significantly. The l0-decoding algorithm is one of the fastest of these algorithms but is not capable of recovering noisy signals.

- We have created the robust-l0-decoding algorithm that is able to cope with noisy measurements but keeps the speed advantage and the recovery qualities of the noise free l0-decoding algorithm.

3. Numerical Results

Recovery Capability

We will begin by investigating the recovery capabilities of the algorithm. Recall that the signal x we want to recover is a vector of size n with k non-zero entries and we want to reconstruct this signal from m measurements y. We write $\delta = m/n$ for the ratio of measurements to signal size and $\rho = k/m$ for ratio of non-zeros entries in the signal and number of measurements. We compare our Bayesian Robust l0-decoding algorithm against

For small-noise scenarios our new algorithm has superior recovery properties compared to the existing SSMP algorithm.

the Sequential Sparse Matching Pursuit (SSMP) algorithm, an algorithm that is slower than l0-decoding when there is no noise but that can also be used for noisy measurement. A key measure of performance is the "phase transition curve", which shows the value of ρ for which the algorithm is not able to recover the signal anymore and hence a higher phase transition curve means that more non-zero coefficients from the signal can be recovered using the same number of measurements.

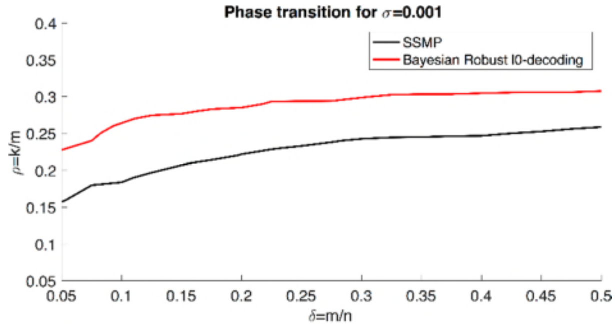


Figure 2: Phase transition curves for the two algorithms for a low noise ($\sigma = 10^{-3}$) scenario.

In Figure 2 we can see that our Algorithm can recover significantly more non-zero values that SSMP when the noise is small compared to the signal strength. For higher noise scenarios we can see in Figure 3 that our algorithm only performs slightly better.

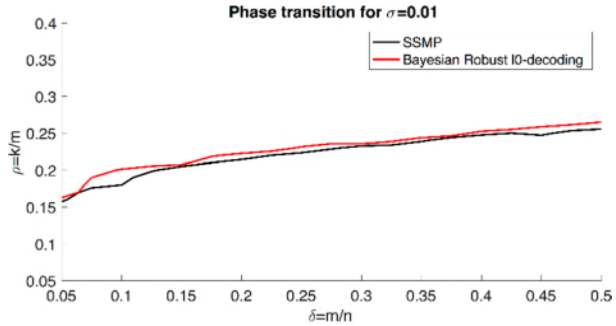


Figure 3: Phase transition curves for the two algorithms for a high noise ($\sigma = 10^{-2}$) scenario.

Run time

In the noise-free case, the l0-decoding algorithm is significantly faster than SSMP and has the advantage that it is parallelizable without loss in quality.

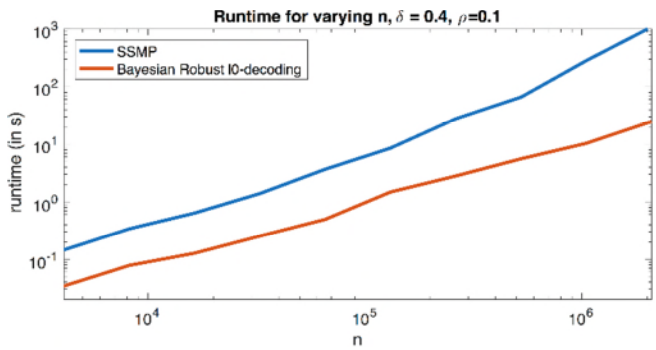


Figure 4: Runtime of SSMP and Bayesian Robust l0-decoding as the problem size increases.

Our robust version is also parallelizable, which is important since this allows for extremely fast implementations on modern graphics processing units that can recover even very large signals in the blink of an eye. Apart from the benefit of a possible parallel implementation, we can also see that the sequential version of our algorithm is both faster and scales better than SSMP, as shown in Figure 4.

Expanding the model

- Currently the theory behind the new method assumes that both the noise and the signal are centred, independent Gaussian random variables. Our analysis could be undertaken for other families of probability distributions to obtain algorithms suitable for more complicated signals.

4. Real-Time Single Pixel Surveillance Camera

We now apply the l0-decoding to the Single Pixel Camera that was described in the introduction. As explained before, classical reconstruction algorithms are computationally expensive and relatively slow, leading to slow reconstruction times for high resolution video. Even on modern processing units this means that real time video is only possible at very low framerates; however, using our new, fast l0-decoding algorithm, we can achieve significantly higher framerates. Furthermore, we use a second type of redundancy in addition to that within a single frame: for video in general and especially for surveillance cameras, the image changes very little from frame to frame. If we assume that we know the background image that the surveillance camera usually sees, then we can make use of this by initializing our recovery algorithm with the background so that we only need to identify how a new image differs from that background.

We illustrate the performance of l0-decoding in this scenario by simulating measurements from a Single Pixel Infrared Surveillance camera and reconstruct the image using the deterministic Bayesian l0-decoding algorithm.

Figure 5: Recovered image of a thermal camera from 30% measurements.

We are able to recover images from only a fraction of the measurements that a normal camera would need.

In Figure 5 we show the recovered image when using 23040 measurements; since the image has a resolution of $320 \times 240 = 76\,800$ pixel this means a reduction of measurements by a factor of three. The run time of the algorithm to recover this image is about 0.3 seconds using a simple sequential C++ implementation; when using a more optimized, parallel implementation on GPUs we can expect that one can achieve reconstruction in milliseconds and hence achieve significantly higher frame rates with this technique compared to the classical reconstruction methods.

We have considered a single pixel camera, however, the ideas are generic and can easily be extended to multiple pixels each measuring independently a block of the picture. We simulate recording video at $640 \times 360 = 230400$ pixel resolution and splitting it up into 12 blocks of size $160 \times 120 = 19200$ pixels. In Figure 6 we show the reconstructed image from just 2880 measurements, which means that with a modern DMD, using our decoding algorithm, framerates of close to 12fps are possible.

Figure 6: Recovered image of a visible light camera from 15% measurements.

5. Conclusion

We have successfully extended the l0-decoding algorithm to the case of noisy measurements and illustrated its speed advantages over the existing SSMP algorithm. We were also able to show improved recovery capabilities for low noise scenarios with performance similar to SSMP for higher noise scenarios.

The l0-decoding algorithm has been successfully applied to image reconstruction for Single Pixel Cameras and paves the way towards more power efficient cameras and real-time video decoding.

6. Potential Impact

Many Compressed Sensing applications have been on the brink of commercialisation but one problem holding them back has being relatively slow and expensive recovery algorithms. This work brings commercialisation one step closer by demonstrating that fast modern algorithms like the l0-decoding algorithm can be extended to be applicable for real-world signals.

William Carson, Consultant at PA Group, commented: "*Florian embraced the changeable environment of consulting, and demonstrated real innovation and team work in his approach, while ensuring that the project achieved all of its goals: I have no doubt that we will sell multiple pieces of client work because of his project.*"

References

1. E. J. Candes and T. Tao (2005) *Decoding by linear programming.* IEEE Transactions on Information Theory, 51(12): 4203-4215.

2. D. L. Donoho (2006) *Compressed Sensing.* IEEE Transactions on Information Theory, 52(4):1289-1306, April 2006

3. R. Mendoza-Smith and J. Tanner (2015) *Expander parallel-l0 decoding.*