RICE UNIVERSITY

Compressed Sensing for Imaging Applications

by

Dharmpal Takhar

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Kevin F. Kelly, Chairman
Assistant Professor of Electrical and Computer Engineering

Richard Baraniuk
Professor of Electrical and Computer Engineering

Wotao Yin
Assistant Professor of Computational and Applied Mathematics

Houston, Texas

February, 2008
ABSTRACT

Compressed Sensing for Imaging Applications

by

Dharmpal Takhar

Compressed sensing is a new sampling theory which allows reconstructing signals using sub-Nyquist measurements. This can significantly reduce the computation required for both image and video whether during acquisition or encoding, especially at the sensor. Compressed sensing works on the assumption of sparsity of the signal in some known domain, which is incoherent with the measurement domain. We exploit this technique to build a single pixel camera using an optical modulator and a single photosensor. Random projections of the signal (image) are applied to the optical modulator, which has a random matrix displayed on it corresponding to the measurement domain (random noise). This random projected signal is focused and summed at the photosensor and will be later used for reconstructing the signal. In this scheme, a tradeoff between the spatial extent of sampling array and a sequential sampling over time with a single detector is performed. In addition to the single sensor method, we will also demonstrate a new design which allows compressive im-
plementation by parallel collection of many random projections simultaneously with a modified sensor array. Applications of this technique in hyperspectral and infrared imaging will be discussed.
Acknowledgments

I would like to express my sincere gratitude towards my advisor Dr. Kevin Kelly and our collaborators Dr. Richard Baraniuk and Dr. Daniel Mittleman for their invaluable guidance and support given to me during the course of this research. I am grateful to my research group members for their help during my stay here. I would like to thank Jason Laska, Marco Duarte, Ting Sun and Vivek Bansal for their help and invaluable suggestions.
# Contents

Abstract ii  
Acknowledgments iv  
List of Figures ix  

1 Introduction 1  
1.1 Compressed Sensing 1  
1.2 Single-pixel Camera Based on Compressed Sensing 3  
1.3 Extensions of the Single-pixel Camera 4  

2 Compressed Sensing 6  
2.1 Traditional Imaging 6  
2.1.1 Conventional Sampling 6  
2.1.2 Consequences of CS Acquisition 7  
2.2 Compressed Sensing 8  
2.3 Compressed Sensing Theory 9  
2.3.1 Sparsity and Compressibility 11  
2.3.2 Incoherent Projections 13  
2.4 Signal Recovery 14  
2.4.1 $\ell_0$ Optimization 14
2.4.2 $\ell_1$ Optimization ............................................. 15

2.5 Advantages ......................................................... 16
  2.5.1 Applications .................................................. 18

3 Compressed Sensing Based Single-pixel Camera .................... 20
  3.1 The Optical Modulator ......................................... 20
    3.1.1 Digital Micromirror Device (DMD) ....................... 21
    3.1.2 DMD Discovery 1100 Board .............................. 22
    3.1.3 Accessory Light-modulator Package (ALP) ............. 24
  3.2 Experimental Setup ........................................... 26
    3.2.1 Photodiode ............................................... 31
    3.2.2 Measurement Basis ....................................... 33
  3.3 Reconstruction ................................................ 34
    3.3.1 Basis Pursuit ............................................ 34
    3.3.2 Matching Pursuit ........................................ 34
    3.3.3 Fixed-Point Continuation (FPC) Method ................. 34
    3.3.4 Total Variation Minimization ........................... 35
  3.4 Results ......................................................... 35
  3.5 Classification and Detection .................................. 36
  3.6 Structured Illumination ....................................... 38
4 Extensions of the Single-pixel Camera

4.1 Color CS Camera ................................................. 41
4.2 Simultaneous Visible and Infrared Imaging ....................... 42
4.3 Compressive Hyperspectral Camera .............................. 45
  4.3.1 Hyperspectral Reconstruction .............................. 49
4.4 Compressive THz Camera ........................................ 52
  4.4.1 State-of-the-art in THz Imaging ............................ 52
  4.4.2 Lithographic Shutters ..................................... 57
  4.4.3 THz Imaging Setup ...................................... 58
  4.4.4 Reconstruction .......................................... 59
  4.4.5 Alternative Modulators .................................. 60
  4.4.6 Diffraction in THz Camera ................................ 62

5 Spatially Multiplexed CS Camera ................................. 63

5.1 Imaging Systems Based on Lenslet Array ....................... 63
  5.1.1 TOMBO .................................................. 64
  5.1.2 Thin Compressive Imaging Sensors ....................... 65
5.2 Single-shot Compressed Sensing Camera Setup .................. 66
  5.2.1 Static Mask ............................................. 66
  5.2.2 CMOS Camera ............................................ 69
  5.2.3 Lenslet Array ........................................... 69
6 Future Directions

References
List of Figures

3.1 0.7 XGA DDR DMD chipset ........................................... 21
3.2 DMD architecture ...................................................... 23
3.3 Optical modulation by tilting of the micromirrors ............... 23
3.4 DMD discovery 1100 board combined with ALP-1 board .......... 24
3.5 Single-pixel camera schematic ..................................... 27
3.6 Single-pixel camera setup ........................................... 28
3.7 Sequence of 64×64 random patterns ................................ 28
3.8 Typical datalogging for single-pixel camera .................... 30
3.9 Photosensitivity plot of S1223 photodiode ...................... 32
3.10 Reconstruction of ‘R’ at different resolutions .................. 37
3.11 Reconstruction of ‘R’ at 128 × 128 and 256 × 256 resolutions .... 37
3.12 Targets for the classification and detection experiment ...... 39
3.13 Plot of classification results ...................................... 39
3.14 Structured light illumination setup .............................. 40

4.1 Color reconstruction of 32 × 32 mandrill image ............... 43
4.2 Two-color (sandwich) photodiode .................................. 44
4.3 Spectra of the silicon (visible) layer and InGaAs (infrared) layer in the two-color photodiode .............................................. 45
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4</td>
<td>Reconstruction of ‘K’ with the two-color photodiode</td>
</tr>
<tr>
<td>4.5</td>
<td>Transmission curves for the DMD glass window</td>
</tr>
<tr>
<td>4.6</td>
<td>USB4000 spectrometer</td>
</tr>
<tr>
<td>4.7</td>
<td>Compressed sensing hyperspectral camera setup schematic</td>
</tr>
<tr>
<td>4.8</td>
<td>64 channel reconstruction of 32 × 32 mandrill</td>
</tr>
<tr>
<td>4.9</td>
<td>16 channel joint reconstruction of 128 × 128 mandrill</td>
</tr>
<tr>
<td>4.10</td>
<td>64 channel independent reconstruction of 128 × 128 mandrill</td>
</tr>
<tr>
<td>4.11</td>
<td>64 channel joint reconstruction of 128 × 128 mandrill</td>
</tr>
<tr>
<td>4.12</td>
<td>256 channel joint reconstruction of 128 × 128 mandrill</td>
</tr>
<tr>
<td>4.13</td>
<td>Six random patterns etched on a copper plate</td>
</tr>
<tr>
<td>4.14</td>
<td>THz Image Reconstruction</td>
</tr>
<tr>
<td>5.1</td>
<td>Mask of (32 × 32) random patterns</td>
</tr>
<tr>
<td>5.2</td>
<td>Microscope snapshot of the mask at 50X resolution</td>
</tr>
<tr>
<td>5.3</td>
<td>Single-shot compressed sensing setup</td>
</tr>
<tr>
<td>5.4</td>
<td>CMOS image of the letter ‘R’ captured through the lenslet array and the mask</td>
</tr>
</tbody>
</table>
Chapter 1
Introduction

1.1 Compressed Sensing

A new universal theory of sampling, called compressed sensing has emerged [1-3]. This thesis deals with direct application of compressed sensing in the image acquisition scenario which is essentially compressed sensing in two dimensions and higher. Thus our discussion in this thesis will be limited to the field of imaging only. In the conventional imaging schemes, based on Shannon sampling theorem, we sample at the rate of at least twice the bandwidth which in this case would be the spatial frequency [4]. All the commonplace signals we are normally interested in capturing have some structure embedded in them and thus can be sparsely represented in at least one known domain. If there is no structure then this sparsity is lacking and the proposed scheme is not capable of capturing that. Clearly the Shannon sampling scheme does not exploit the possibility that there is extra information and thus some inherent structure to the signal/image. Using this a priori information, it should lead to a potentially huge under-sampling.

Scientists and engineers have come up with a great deal of empirical methods which allowed to sample at the rate lesser than required by the Shannon sampling theorem. In a way all of these methods were exploiting the structure of the desired signal. Compressed sensing is a general scheme of sampling which requires and utilizes
the information about the signal structure and allows greatly under-sampled signal acquisition. The structure we are interested in here is the sparsity of the signal in some known basis which in our case (natural images) would be mostly Fourier or wavelet bases. We call this basis the sparsity basis. This knowledge about the sparsity basis allows us to choose an appropriate measurement basis such that it is incoherent with the sparsity basis. Roughly it would translate as the requirement that no element of either basis can be sparsely represented in terms of the elements of the other basis. One such pair of the sparsity basis and measurement basis would be Dirac function and sinusoids. Gaussian random matrices are universally incoherent with all the known bases. Thus they are the ideal choice as the measurement basis. In the case of imaging, this corresponds to taking projections/inner-products of the image with random matrices. The theory and applications of compressed sensing will be explained in detail in chapter 2.

Now the samples received by the sensor are not the usual real-space/time samples anymore, instead they are linear projections of the signal. Each linear projection of the signal on the measurement basis elements provides us with one measurement. The number of measurements is far lower than number of samples required by the conventional sampling scheme. The original signal is reconstructed from these linear measurements by nonlinear methods (optimization). Researches are continually coming up with better and faster reconstruction algorithms for CS signal recovery.
The latest one being used (mainly for hyperspectral reconstruction) is fixed-point continuation method for $\ell_1$-minimization [5].

1.2 Single-pixel Camera Based on Compressed Sensing

To implement the compressed sensing idea for imaging we have devised a sensing setup where we can take random projections of the required image [6]. The radiation/light from this projection is captured by a photosensing element which in turn gives us a scalar number (photovoltage). In one method, the random projections can be implemented by an optical modulator. Texas Instruments has built such an optical modulator, the most sophisticated commercial MEMS device, called a digital micromirror array device (DMD) [7]. This device is similar to a CCD (charge-coupled device) array but in this case the pixels are replaced by micromirrors. The DMD employed in our experiments has an XGA resolution applied to the DMD meaning it contains an array of $1024 \times 768$ micromirrors.

The random matrix in this case was a matrix whose elements are ones and zeros (Bernoulli-Rademacher vectors). On average half the elements are ones and rest of them are zeros. Each micromirror of the DMD can tilt by $\pm 12^\circ$ about its diagonal and thus allows us to display the random matrix on the DMD. The light from the scene is focused on the DMD and this random configuration splits the light into two beams (at $\pm 24^\circ$) corresponding to mirrors representing ones and zeros. Either of these beams is then focused on a photodiode which yields a photovoltage. This process is
repeated with different random matrices giving us measurements corresponding to these different random projections. The original image is then reconstructed from these measurements by means of nonlinear optimization techniques. The single-pixel camera and its applications will be detailed in chapter 3.

1.3 Extensions of the Single-pixel Camera

Now that we have demonstrated the proof of concept experimental setup, the next goal would be to exploit it in other schemes where compressed sensing makes the obvious choice to follow. Chapter 4 is devoted to these extensions of the single-pixel camera. In the single-pixel camera setup mentioned in the last chapter the most easily achievable modification would be at the photosensing element. The photodiode can be replaced by a multi-color photodiode or a spectrometer. Currently multi-color photodiodes are commercially available only as two-color (or sandwich) photodiodes. In this sandwich photodiode there are two overlapping layers of sensing elements which respond to different parts of the electromagnetic spectrum. If this photodiode is composed of a silicon top layer and InGaAs bottom layer it will allows us to reconstruct simultaneous visible and infrared images of the scene.

Replacing the photodiode by a spectrometer, we can construct a hyperspectral compressed sensing camera. Depending on the number of sensors in the linear sensing element we can get from a tiny to a really huge hyperspectral datacube, according to the application requirements. This hyperspectral compressed sensing camera achieves
huge undersampling because the sparsity level does not increase at the same rate as the increase in the dimensionality of the dataset. The cases discussed above represent the ultimate trade-off between time and space.

Compressed sensing would be most useful in those sensing problems where currently no sensing/imaging schemes are possible or are economically or physically impractical. Terahertz (THz) imaging is one of the frontiers which belongs to this problem [8]. In this setting we cannot directly employ the DMD because the metallic micromirrors absorb the THz radiation. Therefore as a proof-of-concept, random patterned masks from the copper coated printed circuit boards were used as the modulator. Dr. Daniel Mittleman in our department is a THz imaging authority and this project was done in collaboration with him and his graduate student Wai Lam Chan.

To overcome the problem of a large number of sequential measurements in the single-pixel camera we have designed a single or few-shot compressed sensing camera. Chapter 5 explains the idea behind this new camera and discusses the experimental results thereof. This design uses a CMOS array in combination with a lenslet array and a mask with random patterns lithographically defined on it. The multiple sub-images formed by the lenslet array and randomly coded by the mask can be used to reconstruct higher resolution images. The great advantage with this design is the simultaneous super-resolution achievable by employing few sub-pixel shifts during the measurement.
Chapter 2
Compressed Sensing

2.1 Traditional Imaging

To explain the differences between traditional imaging and compressed imaging it is best to take the example of digital camera. In a digital camera the scene is mapped via optics onto a CCD/CMOS array whose resolution is the same as the final image resolution. The light values are converted into numbers and these numbers are now compressed using a standard DSP chip. For JPEG compression a discrete cosine transform is applied and only a few numbers are saved and rest of the numbers, which are either zero or close to zero, are thrown away. The obvious question that arises out of this scheme of imaging is, why do we capture the 100% measurements (samples) if we are going to throw away most of them? Is it possible to design the acquisition system such that it captures directly only those important coefficients which are saved later? Effectively what we are asking here is acquisition and compression at the same time. That is where compressed sensing (CS) comes to help by allowing to precisely do this, though with a slight oversampling factor [9].

2.1.1 Conventional Sampling

The current technology is based on Shannon sampling theorem which states that we must sample at the rate of at least twice the signal bandwidth [4]. This scheme
is too pessimistic for many signal classes and in reality it is the worst case bound for any bandlimited signal. Though we are rest assured that this sampling scheme will never fail as long as we follow the laid out criteria. The problem arises when the signal at hand puts extreme demands on the acquisition hardware in terms of sampling frequency or power requirements. For example, as the sampling frequency increases, the cost of the A/D converters also rises. Beyond a certain range (few GHz) sampling is not possible. In such a situation, CS becomes ideal because it makes no prior assumption about the bandwidth of the signal. Similarly, when the datasets get larger and larger eventually they become intractable. For example, in hyperspectral imaging the datasets are extremely huge due to the fact that each scene has several images corresponding to different wavelengths. Here again CS provides a solution by exploiting the fact that this hyperspectral datacube would be very sparse in some known domain and very few measurements might be needed to capture the entire datacube.

2.1.2 Consequences of CS Acquisition

The state-of-the-art sampling/acquisition technologies perform well for if they match the imaging system requirements. For example, in the current digital camera due to the cheap silicon processing, it is possible to build CMOS arrays of bigger and bigger resolution and they perform very well. Once we move outside the visible spectrum, the story is totally different because we no longer have access to cheaper
sensing elements. Also the sensor size may become bigger once we move out of traditional camera applications, e.g., photomultiplier tubes (PMT) for low light imaging. They are expensive and very bulky to make a sensor array. These are all few of the examples where a radical imaging paradigm could be very beneficial.

2.2 Compressed Sensing

Compressed sensing in essence states that a small number of linear projections of a compressible signal contain enough information for reconstruction of the signal and further processing. This effectively means undersampling of the signal from the perspective of Shannon sampling theorem. This new method of sampling demands new designs of the imaging systems/camera which are compressing the signal right at the acquisition stage. This is especially important in the realm of increasing pressure on imaging sensors, hardware and algorithms to accommodate ever larger demands. A few examples of this include,

- larger and higher-dimensional datasets (e.g. hyperspectral datacube)
- faster acquisition, sampling and processing rates (e.g. A/D in GHz range and beyond)
- need for radically new sensing modalities to explore unexploited parts of electromagnetic spectrum (e.g. Terahertz)
- integration of sensing and compression resulting in smaller form factor
Compressed sensing promises to help solve all of these problems, but can only be realized with the tremendous improvements in the computing power now available. The idea is to capture few linear projections and reconstruct the signal using nonlinear methods. In essence what CS says is that 'sparse signal statistics can be recovered from a small number of nonadaptive linear measurements [1].' CS integrates sensing and compression and exploits the new sparse data representations. It is based on new uncertainty principles between the sparsity basis and the measurement basis. Signal is local whereas the measurements are global (due to the randomness of the projections). Each measurement extracts out a little information from the signal and that is why each measurement has equal weight.

2.3 Compressed Sensing Theory

Computational signal processing (CSP) is the field where analog signals are sampled in some desired fashion to create intermediate digital representations which are further processed using nonlinear methods. This nonlinear processing yields either the original signal in digital form or extracts some desired information out of it without reconstructing it. CSP exploits the fact that all signals (besides random noise) have some inherent ‘structure’ that allows an intelligent representation of the signal in some domain (to be obtained) and processed accordingly. The structure we are especially concerned with here is the sparsity of the signal in some known domain. A good example is to consider the compression methods which apply an appropriate
transform coder which in turn converts the signal to a sparse representation. This effectively means that the signal can be represented by a few coefficients and the rest of the coefficients are either zero or close to zero. If the original signal has, let’s say, \(N\) samples then after applying the transform we are left with only a small number of \(K\) coefficients \((K \ll N)\) which still accurately represent the signal. The smooth images are sparse in the Fourier domain and piece-wise smooth images are sparse in wavelet domain. This fact is exploited by JPEG which uses these transform coders to compress the images.

The current technology is based on ‘sample first, ask questions later.’ It translates to first acquiring all the samples and then only the structure of the signal is addressed. A typical compression transform coding procedure involves these steps:

- acquire all the \(N\) samples needed according to Shannon sampling theory
- apply the transform and find out all the transform coefficients
- choose the \(K\) biggest coefficients and discard rest of them
- encode and save the values and locations of the \(K\) coefficients

It is obvious that much of the data acquired using costly hardware would end up discarded in this process. In the applications where acquiring each extra data point adds significant cost to the acquisition process, this method of transform coding would be very undesirable. Ideally we would like to have such a process which directly
estimates those much sought after $K$ coefficients. This will relax the demands on acquisition side by an order of magnitude. Compressed sensing targets exactly this point. The quintessential theorem of CS states that if a signal is $K$ sparse in one basis (sparsity basis/domain) then it can be acquired and recovered from $cK$ nonadaptive linear projections onto another basis (measurement basis) that is incoherent with the sparsity basis. Here $c$ is a small oversampling factor. In particular, note the linear nature of the acquisition process versus the nonlinearity of the reconstruction process. However this should not be a significant problem if adequate computing power is available.

Compressed imaging works by multiplexing the signal; each measurement is a linear function of several pixels of the image. Single-pixel imaging in the past was done by raster scanning or by using transform coding masks. Compressed sensing provides an advantage over transform coding masks in terms of number of measurements and over raster scanning in terms of both the signal-to-noise ratio (SNR) and the number of measurements.

2.3.1 Sparsity and Compressibility

Let us consider a real-valued signal $x$ of length-$N$ defined as $x(n), n \in 1, 2, ..., N$. Since we are interested in only images, we restrict our attention to 2D signals. We order the $N$ pixels in 1D for simplicity of calculation and representation so $x(n)$ can also represent a 2D image. Assuming that signal $x$ is sparse in some basis $\Psi =$
[ψ₁, ψ₂..., ψ₆], we can write x as a linear combination of elements of Ψ:

\[ x = \sum_{n=1}^{N} \theta(n)ψ(n) = \sum_{l=1}^{K} \theta(n_l)ψ_{n_l}. \]  (2.1)

The linear combination is truncated at K because only that many significant coefficients are needed to accurately represent the original signal. Those K vectors are chosen from Ψ; n_l are the indices of those vectors and θ(n_l) are the corresponding coefficients. In matrix notation the above equation can be written as \( x = Ψθ \), where x is an \( N \times 1 \) column vector, Ψ (the sparse basis matrix) is an \( N \times N \) matrix with the basis vectors \( ψ(n) \) as the columns, and \( θ(n) \) is an \( N \times 1 \) column vector with K nonzero elements. Let us denote the \( ℓ_p \) norm by \( \| \cdot \|_p \),

\[ \| x \|_p \triangleq \left( \sum_{n=0}^{N-1} |x_n|^p \right)^{1/p} \]

then \( ℓ_0 \) norm would be number of nonzero elements which gives us \( \| θ \|_0 = K \). For natural images the common type of sparsifying bases used for compression and representation include Fourier basis, wavelets, Gabor bases and curvelets. It makes sense to restrict our attention to exactly K-sparse signals for the time being though the discussion is also valid for the case of K-compressible signals where the remaining \( N - K \) coefficients decay rapidly to zero but are not exactly zero. In the real application we deal with this second case and that is why robustness of compressed sensing is very important.
2.3.2 Incoherent Projections

During the acquisition we do not directly capture the $K$ significant coefficients. Instead we take $M < N$ measurements which are actually projections of the signal onto another set of basis functions $\phi_m$, $m \in 1, 2, ..., M$. The projections (measurements), $y(m)$, are represented as inner products $y(m) = \langle x, \phi_m^T \rangle$, where $\phi_m^T$ denotes the transpose of $\phi_m$. So, in matrix notation the measurements can be condensed as

$$y = \phi x.$$ 

Since the original signal $x$ is $N \times 1$ and measurement basis matrix $\Phi$ is $M \times N$, $y(m)$ has to be an $M \times 1$ column vector. Each row of the measurement basis matrix $\Psi$ is a basis vector $\phi_m$ which gets multiplied by the signal column vector and the resulting scalar is the corresponding element of measurement matrix $y$.

The choice of measurement basis is dictated by a property called incoherency which in simple terms means that any element of measurement basis cannot be sparsely represented in terms of the elements of the sparsity basis and vice versa. Romberg et al [10] have defined a measure of the mutual coherency of these two orthobases $(\Phi, \Psi)$ as shown below,

$$\mu(\Phi, \Psi) = \sqrt{n} \max_{j,k} |\langle \phi_j, \psi_k \rangle| = \sqrt{n} \max_{j,k} |(M\Psi)_{j,k}|. \quad (2.2)$$

Simply speaking $\mu$ determines the largest correlation between any two elements of $\Phi$ and $\Psi$. Obviously, we want the measurement basis and sparsity basis to be as
incoherent as possible which means that the mutual coherency parameter needs to be as low as possible. Ideally we would like to choose the bases such that $\mu$ is close to its minimum value 1 but the possible range of $\mu$ is $[1, \sqrt{n}]$. Spikes (Dirac-delta) and sinusoids are perfect example of an incoherent pair with $\mu = 1$. Fortunately, it is possible to find many such pairs of bases which are practical from the point of view of their implementation in hardware design.

Compressed sensing states that if the two bases as discussed above are incoherent with each other and number of measurements is large enough (determined by the sparsity of the signal) then it possible to recover the signal.

\section*{2.4 Signal Recovery}

We have the case where $M < N$, i.e., the number of measurements $M$ is far fewer than the number of variables $N$. So recovery of the signal $x$ from measurements $y$ is an ill-posed problem but we have extra information about this system of equation. This extra information is the sparsity of the signal in some known domain and this in turn allows us to reconstruct the signal uniquely.

\subsection*{2.4.1 $\ell_0$ Optimization}

Assuming the signal acquisition satisfies the above mentioned conditions (sparsity and incoherence), it is possible to recover the signal by $\ell_0$ optimization [1]. Since the signal we are considering is K-sparse, we look for the set of coefficients $\theta(n)$ which has
minimum $\ell_0$-norm and agrees with the $M$ measurements. The solution (with high probability), coefficient vector $\theta$ is solved using this $\ell_0$ minimization,

$$\hat{\theta} = \arg\min_\theta \|\theta\|_0 \quad s.t. \quad y = \Phi \Psi \theta.$$  

With CS, we should be able to recover the signal using $\ell_0$ optimization from only $M = 2K$ random measurements. This is not practical though due to the combinatorial complexity of the $\ell_0$ optimization problem. It involves finding the sparsest solution out of the $\binom{n}{k}$ possible sparse subspaces. With a slightly larger number of measurements CS allows us to convert this $\ell_0$ minimization problem into an $\ell_1$ minimization problem.

### 2.4.2 $\ell_1$ Optimization

Due to the incoherence property of the measurement and sparsity bases, it is possible to find a solution which is equivalent to the $\ell_0$ optimization by performing $\ell_1$ minimization. Though in this case the number of measurements required is slightly more than $2K$ by some factor in practical settings. So now the solution for $M > cK$ will be given by,

$$\hat{\theta} = \arg\min_\theta \|\theta\|_1 \quad s.t. \quad y = \Phi \Psi \theta$$  

This is a well known optimization problem and is called the Basis Pursuit (BP). It is much more tractable and can be solved using conventional linear programming techniques.
2.5 Advantages

The fact that we need only one detector to achieve imaging without sacrificing much is extremely important, especially if the cost of detector is prohibitive to make a big array for conventional imaging. The single-pixel camera design based on random projections has some unique features which put it far ahead of current imaging systems [28]:

- Universality: A random (and psuedorandom) measurement basis is universal in nature because it can be paired with any sparsifying basis and we still end up with random (or psuedorandom) measurements.

- Future proof: If in the future somebody comes up with a better sparsifying basis for the signal captured then using the same acquired random projections it is possible to reconstruct an even higher quality image.

- Weak-encryption: Due to the existence of random projections, typically only the person who knows the random seed can reconstruct the signal.

- Robustness and Progressivity: All the random projections carry equal weight or in other words they carry equal amount of information and are democratic in nature. This leads to progressively better reconstruction which means that the addition of every new measurement improves the reconstruction quality. At the same time it also guarantees that there will not be any sudden drop in final
image quality, rather it leads to gradual loss of quality with decreasing number of measurements. This feature is particularly beneficial if there is a possibility of data loss.

- Scalability: In regular digital cameras there is a trade off between resolution and the number of pixels in the imaging sensors. Whereas in single-pixel camera the trade off is between amount of compression desired and acquisition time. It is possible to adaptively select the number of measurements to be computed from this tradeoff.

- Computational asymmetry: Since the process of random projection is linear and number of random projections is far less than the true resolution, most of the computational requirements are passed onto the decoding side which usually contains the bulk of the computation resources. By putting the computational complexity at the reconstruction side, the demand on the imaging sensors and DSP chips at the acquisition front is relaxed.

Random projections have been long known for dimensionality reduction in the mathematical and computer science fields. Similarly $\ell_1$-minimization was known to be connected with the sparsity of the signal. The contribution of CS is to bring them together to demonstrate that by taking random projections and later performing $\ell_1$-minimization allows one to recover the original signal accurately.
2.5.1 Applications

Since its inception in 2004, CS has already found several applications in different fields of research and industry. The most significant developments are in the field of medical imaging. Michael Lustig et al [11] achieved significant undersampling of \( k-t \) space by random ordering of the phase encodes in magnetic resonance imaging (MRI). They were able to reduce the number of measurements required by almost an order of magnitude. Similar research has proved CS to be extremely beneficial in EEG [12] and photoacoustic imaging.

Prof. Richard Baraniuk and collaborators have been working on the implementation of CS for 1-D signals [13]. They are mainly interested in signals of frequencies for which no suitable A/D currently exist. They have shown that CS is able to sample and reconstruct signals in the given frequency range. Application of CS in radar imaging and synthetic aperture radar imaging has also been demonstrated [14, 15]. The same group is also designing compressed sensing DNA microarrays so that each sensor target now responds to a group of targets instead of a single target [16].

Compressed sensing also has applications in the earth sciences. Geophysicists know that seismic wavefields are sparse in curvelet domain. The eigenfuntions of the Helmholtz operator for the acoustic waves are incoherent with curvelets. This is the perfect setting for application of CS as discussed by Felix Herrmann [17].

Besides the above mentioned applications, all the imaging systems developed in
this thesis are based on compressed sensing. A single-pixel camera is designed and successfully implemented which reconstructs an image with significant undersampling [6, 18]. The main components of this camera are an optical modulator and a single photosensor. This single photosensor can be replaced with a two-color photodiode which yields simultaneous visible and infrared images of a scene. Replacement of the single photosensor by a spectrometer converts this camera into a compressive hyperspectral camera with tremendous reduction in the size of the hyperspectral datacube. A spatially multiplexed CS camera is also developed which utilizes a lenslet array and random patterned mask to achieve super-resolution.

All of these applications target the visible and infrared regime but CS based imaging is even more powerful beyond this regime. A CS based THz camera is developed where random pattern etched printed circuit boards are used as modulators. In some applications it is not desired to reconstruct the original signal, rather a decision is to be made from the measurements. Classification and detection problems are prime examples of this kind of imaging. An experiment performed with the single-pixel camera shows that 100% classification rates can be achieved from as little as 6 measurements for a 128 × 128 image [19].
Chapter 3
Compressed Sensing Based Single-pixel Camera

Compressed sensing (CS) indeed seems to be a very powerful theory but its utility still requires experimental verification. The goal of this thesis was the design and implementation of a proof-of-concept imaging system based on compressed sensing. As will be detailed, the experiment successfully proved the validity of compressed sensing for imaging applications. In this imaging setup random projections are applied to the given scene (or an image) and the resultant randomly sampled light is collected on a photodiode. These photovoltage measurements are used to reconstruct the original scene through the use of nonlinear optimization techniques. Further, this imaging system was extended to simultaneous visible and infrared imaging as well as hyperspectral imaging.

3.1 The Optical Modulator

For implementation of the compressed sensing to 2-D signals (images), the most crucial part of the experiment would be the application of random projections. Fortunately there exists a spatial light modulator referred to as a digital micromirror device (DMD) which allows us to apply those random projections [7]. Also additional control unit called Accessory Light Package (ALP) is required for the management of the individual micromirrors that compose the DMD [20].
3.1.1 Digital Micromirror Device (DMD)

The DMD is a ubiquitous device in display technology and the commercialization of this chip is marketed as digital light processing (DLP) by Texas Instruments (TI). Most of the current laptop projectors and all DLP high definition TVs have this DMD chip inside them. Commercial DMDs are available with wide range of resolutions. TI has special developer’s kits available for research purposes called DMD Discovery board. Fig. 3.1 shows a 0.7 XGA DDR DMD chipset composed of an array of electrostatically actuated 1024 × 768 micromirrors which was used in our experiment. Each micromirror is controlled individually by changing the binary state of the CMOS control circuitry underlying the micromirror. This chipset is fabricated by silicon
semiconductor processing. The flat top surface of each micromirror is coated with aluminum in order to convert it into a reflecting surface.

The architecture of a single micromirror is illustrated in Fig. 3.2. Each mirror is 13.68 μm by 13.68 μm in size with a 1 μm separation between mirrors. This separation between mirrors and the tiny square hole at the center of every micromirror reduce the efficiency of the DMD by bringing the optical fill factor down to 85%. TI plans to fill this square hole in the newer generation DMD chipsets and also the separation between mirrors has been constantly shrinking with each new generation increasing the fill factor. In addition, there exists light loss from the DMD window transmission which can pose additional problems. The light has to pass through the transmission window twice leading to a significant loss at certain wavelengths. The micromirrors are hinged and can tilt by ±12° about their diagonal axes as shown in Fig. 3.3. This tilting of individual micromirrors allows applying random projections to a scene focused on the DMD by randomly tilting the micromirrors.

3.1.2 DMD Discovery 1100 Board

TI had restricted commercial application of DMD to the development of projectors and high-definition TVs. The immense possibilities of the DMD as an optical modulator in the research community are yet to be fully addressed. Once the businesses based on the DMD were successful only then did TI focus on furthering the applicability of DMD to other areas. To help enable research and development, TI
Figure 3.2  DMD architecture.

Figure 3.3  Optical modulation by tilting of the micromirrors.
made available a variety of DMD components and accessories. Later TI came up with DMD Discovery platform for scientists and engineers to develop new engineering applications. Their first release was DMD discovery 1100 development kit. One of the earliest applications of this board was for microscopy research, e.g., in confocal microscopy where an individual mirror can be used for both pinhole excitation and pinhole detection in a raster scanning technique [26, 27].

3.1.3 Accessory Light-modulator Package (ALP)

The ALP controller board is an accessory for DMD Discovery 1100 board to realize the true potential of the DMD. This board provides direct access to the DMD
through the Discovery 1100 board. The ALP board is made by Vialux who designed it primarily as a component of a projector for building a high speed 3-D camera. At the time we received the ALP board used in our setup, it was sold as ALP. Later this board was renamed ALP-1, which was the first generation board made for DMD-1100 Discovery boards. ALP-1 is connected to the DMD Discovery 1100 board using parallel interface as shown in Fig. 3.4.

A new interface is now created between the computer and ALP through a USB 2.0 connection, though, the combined boards are still powered by the DMD discovery board. The ALP board is designed for high speed DMD operation supporting both binary and gray-scale XGA patterns. This allows us to display both the Rademacher 0/1 vectors as well as Gaussian white noise. The installation CD for the ALP board comes with an application programming interface (API). This API was later used as a template for the implementation of random patterns on DMD in collaboration with Jason Laska from Prof. Richard Baraniuk’s group.

ALP gives control over individual micromirrors and allows us to load random patterns on the DMD with significantly high speed. Though it has its share of problems, the biggest being very small on-board memory. The ALP board is capable of displaying a preloaded frame 8,000 times per second but it is not possible to run different frames that faster. In order to run different random patterns we have to load the random patterns on the on-board memory which later displays these patterns on the
DMD. While one set of patterns is displaying, the next set of patterns is loaded into a buffer. This limits our ability in displaying fast random patterns to build a practical compressed imaging setup. The current generation of the ALP boards is ALP-3 which is a high-speed board providing 13,333 frames/s accessed from on-board memory. Also it can be customized for higher on-board memory up to 128 Gbits. Also the new DMD is FPGA programmable so it is possible to generate the random patterns on the DMD itself. This should allow us to store all the random patterns in the memory and then display them at a faster rate than the current setup.

3.2 Experimental Setup

A schematic of the experimental configuration is shown in Fig. 3.5. As illustrated in this figure, a biconvex lens focuses the light from the scene onto the DMD and the light reflected from the micromirrors of the DMD along one of the two possible paths is collected via a biconvex lens on a photodiode. The resultant photovoltage is a measurement of the random projection or inner product between the scene and the random matrix displayed on the DMD. The actual implementation in the lab is displayed in Fig. 3.6. The random basis chosen for the experiment was generated by psuedorandom 0/1 Bernoulli vectors. In this random matrix on average half of the elements are zeros and the other half are ones. The micromirrors which tilt by $+12^\circ$ can be assigned to be 'ones' and those corresponding to $-12^\circ$ can be assigned to be 'zeros'. The light is reflected in either direction and can be captured from either for
Figure 3.5 Single-pixel camera schematic.

our purpose because they both carry equivalent information. This way a total of $M$ random measurements are acquired for an $N$-pixel scene. In theory, these $M$ random measurements carry sufficient information to reconstruct the scene.

To simplify the initial measurements, we consider only square shaped objects and as a consequence the random patterns on the DMD are also square. To display the square patterns on the DMD, the top and bottom 128 rows of the DMD were not used which gave us the central square of $768 \times 768$ micromirrors. Lower resolution measurement matrices were displayed by moving the mirrors in blocks. For example, mirror blocks of $24 \times 24$ were formed to display $32 \times 32$ random matrices. Fig. 3.7 shows the sequence of $64 \times 64$ random patterns for which each pixel would correspond to a block of $12 \times 12$ micromirrors.

The dimensions of the $768 \times 768$ square region of the DMD are $10.506 \text{ mm} \times 10.506 \text{ mm}$. The original image/scene (of dimension $4 \text{ cm} \times 4 \text{ cm}$) was focused on this
Figure 3.6  Single-pixel camera setup.

Figure 3.7  Sequence of 64×64 random patterns.
square region of the DMD by a 2 inch diameter biconvex lens of focal length 60 mm. The distance from the object to the lens was 30 cm and from the lens to the DMD was 7.5 cm, corresponding to about four times demagnification. As mentioned earlier the random patterns split the light from the scene into two parts according to the tilt of the micromirrors. One of these randomly projected light beams is focused by a similar 2 inch diameter biconvex lens onto the photodiode. This time the distance from the DMD to the lens was 36 cm and from the lens to the photodiode was 7.2 cm. These measurements correspond to five times demagnification in the image size. The photovoltage signal generated at the photodiode is sent to a 16-bit A/D converter (USB6210 DAQ by National Instruments). The output of the A/D converter was logged using VI Logger, a datalogging software provided with the USB6210 data acquisition device.

Fig. 3.8 shows the typical data logged during the acquisition of random projections of an image. The code for the Mersenne random patterns generates the patterns in a loop which are then sent to the ALP. Each loop generates a total of 12 random patterns in two parts of 6 patterns each which are separated by a buffer frame. So essentially it is a repetition of 6 random patterns followed by a buffer switching time. Since the signal is noisy (mainly 60 Hz noise) and random projections are closely distributed about the mean value, sometimes it is difficult to differentiate between two consecutive measurements. This problem is dealt by putting an extra total-white
frame after every other random pattern to better distinguish the individual random patterns. Acquisition of 100\% measurements for a $32 \times 32$ would require running the loop for 86 times.

The 16-bit A/D converter has dynamic range of ±10 volts. Since the random projections are distributed in a very close range, the true quantization is far lower. For a typical $32 \times 32$ image the difference between maximum and minimum photovoltage values was about 0.2 volts. This corresponds to $1/100$th of the true quantization. Also the dark current in the experiments was in the range of 0.1 to 0.2 volts. This value needs to be subtracted from the measurements to calculate true random projections.

Though the display of the random patterns can be very fast and is essentially

**Figure 3.8** Typical datalogging for single-pixel camera.
limited by the photodiode characteristics, the buffer time is the limiting factor in our setup. The buffer time is fixed for a given set of random patterns and it is more or less fixed. This leads to longer acquisition times than necessary and makes the device impractical for certain applications at this time. The newer generation ALP and DMD board possess advanced functionalities which would allow to acquire measurements fast enough to make the device practical.

For alignment purposes and for comparison with traditional imaging, raster images were also acquired. During raster (pixel by pixel) scans only one block of mirrors tilts in tandem toward the photodiode (corresponding to that pixel).

3.2.1 Photodiode

The photodiode used in this experiment was a model S1223 from Hamamatsu Photonics. The S1223 is a silicon PIN (p-type : insulator : n-type) photodiode with the photosensitivity response as shown in Fig. 3.9. In random projections on average half of the light is reflected by the DMD, therefore the photovoltage values are closely distributed around the mean value. The range of the values is very small and for practical purposes we can safely assume that we are operating in the linear regime of the photodiode response curve. Also to achieve faster response times, we have also reverse biased the photodiode. An SRS preamplifier (SR560) was incorporated to filter the problem of continuous 60 Hz noise and to apply additional signal gain.
Figure 3.9 Photosensitivity plot of S1223 photodiode
3.2.2 Measurement Basis

The two main measurement bases employed for compressed imaging experiment were matrices of Bernoulli-Rademacher pseudorandom vectors and permuted Walsh-Hadamard patterns. The first kind of measurements were generated in the computer by the Mersenne Twister method [23]. The random matrices are the best for compressed sensing because they are universally incoherent with all other known bases. At the same time these patterns are undesirable in the signal recovery stage because it is computationally expensive to deal with them due to non-availability of a readily available inverse transform. There exist fast algorithms to compute the Walsh-Hadamard inverse transform and this leads to faster signal reconstruction.

The Walsh-Hadamard matrices are permuted randomly so that they can be used as the measurement basis. First the elements of each Walsh-Hadamard frame (matrix) are permuted by bit reversal ordering. Then the frames themselves are permuted by randomly ordering the matrices. This double permutation makes the Walsh-Hadamard matrices random enough to be incoherent (within a certain degree) with the sparsity basis. Interestingly, the larger the size of the matrix, the higher the degree of randomness of the permuted Walsh-Hadamard patterns. Thus permuted Walsh-Hadamard patterns are extremely beneficial for acquiring higher resolution images.
3.3 Reconstruction

Below are discussed the various reconstruction methods employed for image recovery.

3.3.1 Basis Pursuit

The idea behind basis pursuit (BP) is to have a highly overcomplete waveform dictionary and then decomposing the signal into an 'optimal' superposition of the dictionary elements [21]. The 'optimal' superposition is found by taking the smallest $\ell_1$ norm of coefficients among all such superpositions.

3.3.2 Matching Pursuit

Matching pursuit (MP) uses a redundant dictionary of waveforms to decompose any signal into a linear expansion of dictionary elements [22]. The waveforms are chosen such that they have the best match with the signal. Matching pursuit thus extracts the signal structures that are coherent with respect to a given dictionary. It is a general procedure to compute adaptive signal representation.

3.3.3 Fixed-Point Continuation (FPC) Method

This method is for solving minimization problems with $l_1$ regularization,

\[
\min \|x\|_1 + \mu f(x). \tag{3.1}
\]

This problem can be solved with a fixed-point globally-convergent iteration scheme [5]. The approach is based upon operator-splitting and continuation. The function
f(x) is a weighted least-squares term for compressed sensing signal reconstruction. The problem is solved for an increasing sequence of $\mu$ values. The continuation approach consists of using the solution at the last $\mu$ value as the starting point for the next $\mu$ value. This mechanism gives the method its name, Fixed-Point Continuation (FPC).

### 3.3.4 Total Variation Minimization

Total variation (TV) minimization is based on minimization of the intensity variation among neighboring pixels and it can be represented mathematically by,

$$TV(X)_{i,j} = |\nabla X|_{i,j} = \sqrt{(X_{i+1,j} - X_{i,j})^2 + (X_{i,j+1} - X_{i,j})^2}. \quad (3.2)$$

Here $X_{i+1,j}$ and $X_{i,j+1}$ are the neighboring pixels of $X_{i,j}$.

### 3.4 Results

The initial test image was a letter ‘R’ printed on paper. Random projections (Mersenne) of this test object were acquired at four different resolutions: $16 \times 16$, $32 \times 32$, $64 \times 64$ and $128 \times 128$. At first Basis Pursuit (BP) was used for reconstruction and as expected, it was computationally expensive. Next Matching Pursuit (MP) was employed for reconstruction which was faster than BP by almost an order of magnitude but still slower for any real application. Matching pursuit was used as the reconstruction algorithm for results shown in Fig. 3.10.

As is evident from Fig. 3.10, the increase in resolution is accompanied by a sharp drop in the number of measurements required for the same quality of image.
reconstruction. This could be attributed to the fact that the sparsity of the signal increases with the increasing resolution and requires lesser number of measurements. Similarly Fig. 3.11 shows the progressivity nature of the random projections. The quality of reconstruction increases with the increase in number of measurements.

These results in Fig. 3.11 correspond to permuted Walsh-Hadamard random projections instead of Mersenne random projections. Given the limited computational resources, it was not possible to reconstruct a $256 \times 256$ image with Mersenne random projection and matching pursuit. Also total variation minimization was used to reconstruct these higher resolution images which results in smoother pictures. Significantly only $2\%$ of the measurements are needed for the $256 \times 256$ image as compared to $10\%$ for the $128 \times 128$ image. Research groups have been developing better and faster reconstruction algorithms to deal with the signal recovery from random projections. The latest one is fixed-point continuation method, as described above, which gives satisfactory reconstruction at extremely fast rate (the current record holder).

3.5 Classification and Detection

In many imaging applications the goal is to make a decision about the object instead of reconstructing the original signal. Specifically, classification and detection problems fall under this area. For classification and detection purposes structured measurements are taken which are very small in number when compared to the overall resolution of the image. Mark Davenport et al, combine the generalized maximum
Figure 3.10  Reconstruction of ‘R’ at different resolutions.

Figure 3.11  Reconstruction of ‘R’ at 128 × 128 and 256 × 256 resolutions.
likelihood hypothesis (matching filter) with compressed sensing and arrive at a new technique termed the ‘smashed filter’ [19]. This smashed filter is applied in the experiment to classify the targets shown in Fig. 3.12.

The targets are plastic models which are mounted on a rotation stage. Measurements are acquired at every 10° resulting in 36 measurement sets for each object (corresponding to 360°). The smashed filter deals with unknown transformations such as translations, scaling and rotations (viewing angle). The theory of smashed filter exploits the fact that the rotated images (different view angles) of an object, and transformed images in general, form a low-dimensional nonlinear manifold in the higher dimensional image space [24]. The 36 acquired measurement sets of each object would represent a point cloud in this manifold. Fig. 3.13 shows the results of this classification experiment. In the low noise case, as little as 6 measurements seem to be enough for 100% classification rate. This number is tiny compared to the overall resolution, 128 × 128 (=16384), of the images. Thus, the smashed filter is very powerful for the classification applications.

3.6 Structured Illumination

The most important component in single-pixel camera setup is the DMD. It is possible to replicate the single-pixel camera setup, with some modifications, in any device which has a DMD built into it. The most obvious device is a DLP projector. We can input the random projections to the projector which in turn randomly illu-
Figure 3.12 Targets for the classification and detection experiment.

Figure 3.13 Plot of classification results.
minates the object of interest. Fig. 3.14 shows such a setup in the lab where the random illumination is focused through a zoom lens onto a photodiode. Here again the output of the photodiode is sent to the USB6210 A/D converter which sends the data to the computer. The photovoltages will be the corresponding measurements for the random projections. Now the original image can be recovered by employing the methods discussed in this chapter earlier for the single-pixel camera setup. Structured illumination is not new in the field of imaging, structured light confocal microscopy being the prime example [29]. The existing imaging systems can be improved by incorporating random projections to achieve compressed imaging.

Figure 3.14  Structured light illumination setup.
This chapter deals with extensions of the single-pixel camera. The extensions involve modifying the current setup by either replacing the photodiode by another photosensing element or by altering the optical path.

The single-pixel camera setup was modified by putting a filter wheel right before the photodiode. Then measurements corresponding to red, green and blue filters were acquired. This change in the setup turns it into a color version of the single-pixel camera. Next the photodiode was replaced with a two-color photodiode which yields simultaneous visible and infrared images of a given scene. Also by replacing the photodiode with a spectrometer the single-pixel camera converts into a compressive hyperspectral camera. The DMD cannot be used for THz imaging because the metallic mirrors absorb the THz radiation. To overcome this, random patterns were etched on copper coated printed circuit boards and used as masks to develop a THz compressed sensing camera.

4.1 Color CS Camera

Typically color imaging is performed by incorporating color filters either in the CCD/CMOS array or separating the color elements with a prism. It is possible to have a tri-layered photodiode where each layer absorbs light only from a certain
wavelength range of the visible spectrum. Foveon Inc. produces such CMOS arrays but there no individual three-color photodiodes available commercially [30]. In the absence of such a photodiode system, we simply placed a color filter wheel in the optical path of our single-pixel camera experiment.

We chose the printout of the standard ‘Mandrill’ image as our object. The reconstructed image from $64 \times 64$ random projections is shown in Fig. 4.1. In the bottom row the gray scale images correspond to separate red, green and blue channels. The top row shows the resulting merged color images by reconstructing 20% and 40% measurements. The availability of three different measurement sets of the same image should have helped in achieving better quality image by joint reconstruction due to enhanced compressibility of the data set. However, we found a negligible difference in this instance.

4.2 Simultaneous Visible and Infrared Imaging

There are no cheap infrared (IR) cameras available commercially even for a $256 \times 256$ resolution, though, there exist higher resolution IR cameras up to 1 megapixel resolution. This is mainly due to the fact that cheap silicon photosensors are sensitive only to the visible part of the electromagnetic spectrum. The infrared photodetectors such as indium gallium arsenide (InGaAs) are very expensive. The next extension of the single-pixel camera setup is the replacement of the photodiode by a sandwich/two-color photodiode which allows simultaneous reconstruction of visible and infrared
Figure 4.1  Color reconstruction of $32 \times 32$ mandrill image.
images. This would be a cheap infrared camera due to the fact that only a single photodiode is required. Also the resolution is only limited by the resolution of the DMD which can be as high as 2 megapixels.

The PIN-DSIn photodiode, sold by OSI Optoelectronics, was chosen for the experiment (Fig. 4.2). This photodiode has two layers: the top layer is silicon with sensitivity from 350 nm to about 1000 nm and the bottom layer is InGaAs with detection range from 950 nm to 1700 nm. The sensitivity curves for both materials are shown in Fig. 4.3. The object in this experiment was a letter 'K' cut out of an aluminum plate. This plate was front-illuminated by a halogen lamp and back-illuminated by a 3×3 array of infrared LEDs. The reconstructed images corresponding to visible and infrared regime are shown in Fig. 4.4. The low quality of the images could be attributed to alignment mismatch in the optics, low light and stray noise. Also the DMD has a window which does not pass infrared very efficiently. As can be seen
Figure 4.3 Spectra of the silicon (visible) layer and InGaAs (infrared) layer in the two-color photodiode.

Figure 4.4 Reconstruction of 'K' with the two-color photodiode.

from Fig. 4.5, the percent light transmission through the DMD glass window in the infrared region is quite attenuated. The light loss from DMD window transmission is amplified by the fact that the light has to pass through this glass window twice.

4.3 Compressive Hyperspectral Camera

One of the most obvious areas for the application of compressed sensing is hyperspectral imaging. The famous Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) camera gives 224 spectral channels with wavelengths ranging from 400 nm
to 2500 nm [31]. The amount of data captured for one scene is tremendous and compression at the acquisition side would be much appreciated. Prof. Dave Brady's group at Duke University has also built a compressive hyperspectral camera [33]. Their camera is based on a shifted Hadamard transform and a unique dual-disperser design. They were able to reconstruct a 15 channel $256 \times 256$ datacube from only $256 \times 256$ measurements.

To implement hyperspectral imaging in our single-pixel camera we replaced the photodiode by a USB4000 Miniature Fiber Optic Spectrometer (Ocean Optics) (Fig. 4.6). It contains a 16-bit A/D resolution spectrometer with signal-to-noise ratio of 300:1 (at full signal). The range for this particular spectrometer is from 350-1000 nm which falls in visible and near-infrared part of the electromagnetic spectrum. This is

Figure 4.5 Transmission curves for the DMD glass window.
in part determined by the operating range of the grating in the spectrometer which is 350-1100 nm. The photosensing element in this spectrometer is a 3648 pixel Toshiba linear CCD array (TCD1304AP) with pixel size of 8 \( \mu \text{m} \times 200 \ \mu\text{m} \) which provides favorable signal-to-noise ratio. The spectrometer is calibrated to provide a linear response to the illumination, the corrected linearity of this spectrometer is more than 99.8\%. The slit through which the light enters the spectrometer is 10 \( \mu \text{m} \) wide.

The setup was modified, with help from Vivek Bansal, to allow for collimation of the light reflected from the DMD as shown in Fig. 4.7. The optics from the object to the DMD was the same as the single-pixel camera setup. The optical path from the DMD to the collector lens of the fiber is shown in the top part of Fig. 4.7. The focal lengths \( f_1 \), \( f_2 \) and \( f_3 \) are 15cm, 6cm and 6cm, respectively. An objective lens was added to further demagnify the image at the diffraction plane. The collimated
Figure 4.7  Compressed sensing hyperspectral camera setup schematic.
light enters the collimating lens and is sent to the spectrometer through a fiber. The spectrometer is connected to the computer through a USB 2.0 interface. The spectrometer output is logged using SPECTRASUITE Operating Software provided with the USB4000 spectrometer. The spectrometer also has trigger functionality but due to some hardware problem the profile of the acquired signal changes. Though, if operating properly, the trigger would decrease the acquisition time. But to avoid the above mentioned problem in our experiments, we did not use the trigger.

4.3.1 Hyperspectral Reconstruction

Compared to the single-pixel camera described in earlier chapter, the logged data corresponds to spectra (represented as a linear vector) for each random projection instead of a single number reflecting the photovoltage. This data was then reconstructed in collaboration with Marco Duarte by two different schemes: independent spectral slices and joint datacube reconstruction. In the independent spectral slice reconstruction, we separately reconstruct the image corresponding to each channel. Here we are able to exploit the 2D compressibility (or sparsity) in each channel. In this scheme we do not exploit the fact that all these images from different channels belong to the same object and share common structures such as edges. However, it is also possible to perform joint reconstruction of the whole datacube. If $\Psi_1$ is the sparsity basis in spatial domain and $\Psi_2$ is the sparsity basis in spectral (wavelength) domain, then the sparsity basis for the 3D hyperspectral datacube can be computed
by evaluating a tensor product of these two sparsity bases,

$$\Psi_\otimes = (\psi_1 \otimes \psi_2 : \psi_1 \in \Psi_1, \psi_2 \in \Psi_2).$$

The joint datacube reconstruction has $\Psi_\otimes$ as the sparsity basis and permuted Walsh-Hadamard matrices as the measurement basis. The Fig. 4.8 shows 64 spectral channel independent slice reconstruction of a $32 \times 32$ image from 60% random projections (Mersenne). MP was used as the reconstruction algorithm. The spectrum range for this reconstruction was from 470-790 nm which means each channel is an average of spectra over 5 nm. Fig. 4.9 shows a joint reconstruction of 16 channels (450 nm - 850 nm) of a $128 \times 128$ image using 25% of the permuted Walsh-Hadamard projections. Fixed point continuation method was the chosen as the optimization scheme. In this case the images do not appear as smooth as we had expected but increasing the number of channels increases the image quality due to rise in the compressibility/sparsity level. This is obvious in Fig. 4.10 which shows 64 channels of the same $128 \times 128$ image from the same number of measurements. Also Fig. 4.11 shows an independent slice reconstruction of 64 channels of the same $128 \times 128$ image from the same number of measurements. By comparing these two 64 channel images we observe that the image quality has certainly improved through joint reconstruction. To understand whether increasing the number of spectral channels further would improve the image quality, we reconstructed the 256 channel datacube of the same image (Fig. 4.12). Again we observe the image quality is continually improving with
Figure 4.8  64 channel reconstruction of $32 \times 32$ mandrill from 60% measurements.
an increasing number of channels. This improvement is likely due to the increased level of compressibility/sparsity with increase in dimensionality of the data.

A raster scan or full basis scan provides the regular uncompressed hyperspectral datacube. In order to confirm that compressed hyperspectral camera preserves the spectral characteristics of the scene, the compressed datacube reconstruction should match the raster or full basis scan pixel by pixel.

4.4 Compressive THz Camera

Until now we have been concerned with imaging only in the visible domain. The real benefit of compressed sensing lies in developing affordable and efficient imaging devices for other regions of the electromagnetic spectrum. In particular, CS has tremendous potential in the THz domain and its imaging applications.

4.4.1 State-of the-art in THz Imaging

State-of-the-art technology in THz imaging devices is currently limited to a 128 × 128 pixel array where the size of the camera is large and impractical [8]. To reduce the cost and the size of the camera, most of the existing THz cameras employ a single THz sensor and raster scan the object. Of course, rastering results in a compromise on the acquisition speed. Current industry demand is for a terahertz camera which is small enough to be easily transported and at the same time is affordable. Some of the obvious applications for a THz camera would include airport scanners to find
Figure 4.9  16 channel joint reconstruction of $128 \times 128$ mandrill.
Figure 4.10  64 channel independent reconstruction of 128 × 128 mandrill.
Figure 4.11  64 channel joint reconstruction of $128 \times 128$ mandrill.
Figure 4.12  256 channel joint reconstruction of 128 × 128 mandrill.
out hidden weapons, detection of foam insulation defects, illegal drug detection and metal particle detection in package industry.

To overcome the above mentioned deficiencies a new THz camera design is proposed which obviates the need for raster scanning of the object or use of a focal-plane array for detection. This project was planned and executed in collaboration with William Chan from Dr. Daniel Mittelman’s THz imaging laboratory. This group had already demonstrated application of compressed sensing by random sampling in the focal plane [25]. In their experiment a single detector takes far lesser number of measurements (compared to the actual resolution) at pre-selected random locations in the focal plane to perform Fourier space sampling. In essence, it is a Fourier imaging setup which uses compressed sensing (instead of regularized sampling) and phase retrieval for reconstruction of original signal (object). This new design makes a strong promise for real-time pulsed THz imaging with a single detector.

4.4.2 Lithographic Shutters

In case of the optical domain, the DMD implemented the random modulation of the signal. With THz radiation, the DMD is ineffective because the THz radiation gets absorbed by the metal. It would be ideal to have a radiation modulator which selectively reflects/transmits or transmits/absorbs the THz radiation according to the random pattern in question. Printed circuit boards (PCB) are made of fiber-glass coated with copper. Fiber-glass is highly transparent to the THz radiation whereas
copper, being a metal, is an efficient absorber of the THz radiation. Therefore, random patterns were etched on these plates similar to circuits are etched on PCB in regular practice. The PCB plates used for the experiment were 1/16" thick with copper coating of approximately 34.1 \mu m in thickness (far greater than the skin depth of copper for THz radiation).

The random pattern etched plates used for the final experiment were made by an outside company employing standard industry procedures. In this case the pixel size was kept at 1mm \times 1mm assuming that the THz beam size would be wide enough to cover all the 32 \times 32 pixels in a given random pattern. The pixels corresponding to etched copper let the THz radiation transmit through whereas the non-etched pixels block it. Effectively this is a 32x32 random matrix with its elements being only ones and zeros.

4.4.3 THz Imaging Setup

The imaging setup is illustrated in Fig. 4.13. It consists of a THz transmitter/receiver pair at the either end of the setup and in the middle a set of copper plates with etched random patterns and two polyethylene lenses for collimation. The lensmaker formula was used to determine the optimal distances between the elements. The receiver in the setup is a fiber-coupled photoconductive antenna. The object mask is made out of copper tape in shape of letter 'R' on a transparent plastic tape (Fig. 4.14(left)). The size of letter 'R' is 1cm \times 1cm. This copper plate is
mounted on a translation stage so that the movement from one pattern to another is precise to achieve perfect alignment. A THz waveform is recorded corresponding to each random pattern. This waveform is formed by superposition of all the THZ radiation transmitted through the unetched parts of the random pattern. In total 400 measurements were made which is approximately 39\% of the overall resolution.

4.4.4 Reconstruction

Each measurement was calculated by taking the average of the magnitude of the spectrum of the detected THz waveform over a narrow band of frequencies centered around 150 GHz (\(\lambda \approx 2mm\)). The THz beam around this frequency was large enough to cover the whole object (1cm x 1cm) and has the highest spectral amplitude. Total variation (TV) minimization was used as the reconstruction algorithm and Fig. 4.14 shows the reconstruction results corresponding to 200 and 400 measurements respectively. The beam size was smaller than the random pattern size which explains
the fact that the reconstructed letter ‘R’ is much closer to 16 × 16 pixel resolution rather than the anticipated 32 × 32. In spite of this, the features of the object are clearly distinguished. With matching random pattern and beam sizes, it should be possible to achieve a very clear reconstruction. In addition, it should be noted that the laser power fluctuation and some lateral mismatch during alignment of patterns also contributes to the noise in the detected THz waveform which in turn affects the quality of the reconstruction.

The random pattern size was found to be bigger than the THz beam size in the experiment. Thus the natural next step would be to build new random patterns which fit within the beam so that the mask is uniformly illuminated. The proof-of-concept experiment holds the promise toward developing a real-time THz imaging system. If random pattern can be printed on a plastic tape with a metallized ink, it can used like a VCR tape for imaging. Ideally we would like to have a DMD like THz radiation modulator.

4.4.5 Alternative Modulators

Another idea for implementation of random projections in regards to THz radiation involved the standard 35mm B&W camera film as mask substrate. These films have a coating of silver halide which converts to silver colloids at the exposed regions during the ‘developing’ process. The average diameter of silver colloids is about 5 \( \mu \text{m} \) which is far greater than the skin depth of THz radiation in silver metal. We
Figure 4.14 (left) White-light image of object mask with a R-shaped hole. Terahertz image reconstructed via compressed sensing (middle) with 200 measurements (≈ 20%) and (right) with 400 measurements (≈ 40%). Dark areas have small pixel values.

Projected the random patterns and acquired the data using the above mentioned film with random patterns on it. The contrast among the measurements was not good enough for effective measurements. This problem can be overcome by putting more silver through a chemical process or using an enhancing technology such as 'hypersensitization.'

Though the DMD by itself cannot act as the THz modulator, it can still play a part in designing a modulator. By shining a laser of appropriate wavelength a silicon chip can be made transparent to the THz radiation. Here the DMD can be employed to randomly project the laser on the silicon chip which in turn makes those randomly selected parts of chip transparent to THz radiation. Now silicon chip can indirectly act as the THz modulator. Experiments showed that there was not significant contrast in the signals corresponding to the opaque and transparent states.
Vanadium oxide goes through a metal-to-insulator transition at 67°C [32]. The THz radiation transmits through vanadium oxide in insulating state but it is absorbed in the metallic state. Thus, by clever application of localized heaters a vanadium oxide substrate can be made to act as a THz modulator. Also by adding impurities the transition temperature can be manipulated to suitable level.

4.4.6 Diffraction in THz Camera

One point worth discussing is the feature size of the pixels in random patterns during the THz imaging. The wavelength corresponding to 1THz is 0.33mm which is very close to the pixel size, so it may run into the diffraction problems. To avoid this, the object was put right after random pattern etched copper plate so that the random sampling is performed in the near-field.
Chapter 5
Spatially Multiplexed CS Camera

All the experiments in previous chapters with the exception of the hyperspectral camera, utilized the advantage of a single photosensor/antenna design. Such experimental schemes would be ideal in the situations where the sensors are prohibitively expensive or where it is not feasible to have a large array of the sensors. Compressed sensing can also be applied to the traditional camera technology which uses an array of sensors for relatively inexpensive and accessible parts of the electromagnetic spectrum. As a proof-of-concept, the simplest case would involve combining CS with a CCD/CMOS camera to achieve super-resolution. In such an experiment, a lenslet array forms multiple sub-images of the scene which are randomly coded by a mask placed at the image or principal plane. These randomly coded images are then recorded on a CMOS array. Besides achieving super-resolution, this design would also address the major issue of slower measurements with the DMD-based single-pixel camera. In addition, the form factor of such an imaging system would be much smaller compared to the single-pixel camera configuration.

5.1 Imaging Systems Based on Lenslet Array

The idea of lenslet array is inspired by compound-eyes found in certain insects (arthropods). A lenslet array, also called lenticular array or microlens array, is a 2-
dimensional array of microlenses. Each microlens forms a sub-image of the object on the detector plane. The images formed from all the lenlets are shifted relative to each other and the shift is determined by the geometrical parameters of the imaging system. Such multiple aperture imaging systems are essentially performing multichannel sampling.

5.1.1 TOMBO

The most famous imaging system based on lenslet array was demonstrated by a Japanese group [34]. This compact imaging system was named TOMBO (Thin Observation Module by Bound Optics). The most significant advantage of TOMBO imaging system was its small form factor due to very small focal length of the lenslets. A lens with shorter focal length will inherently have a greater depth of field. In this computational imaging system the captured multiple sub-images are processed to reconstruct a super-resolved.

Their reconstruction scheme is explained here. If $H_\odot$ defines the system tensor, matrix $f$ represents the elements of the object and matrix $g$ represents the elements of the image at the photosensor array then,

$$g = H_\odot f.$$

This is essentially an inversion problem where we want to calculate $f$ from the given equation. Now $H_\odot$ can be further split as,
\[ H_\otimes = H_1 H_2, \]

where elements of \( H_1 \) represent the image demagnification for each sub-image and \( H_2 \) represents the point-spread functions of the corresponding lenslet units. \( H_1 \) can be easily calculated by the known optical parameters and \( H_2 \) can be determined experimentally. The super-resolved image is constructed by calculating pseudoinverse from the singular value decomposition of the system tensor \( H_\otimes \).

### 5.1.2 Thin Compressive Imaging Sensors

Prof. David Brady's group at Duke University has developed a similar imaging system with a 3x3 lenslet array, dubbed COMP-I imager [35, 36]. The final image from this system turn out to be upsampled by a factor of 3. They later expanded upon the COMP-I imaging system by adding a static mask to implement focal plane coding. The mask is an array of transform codes and it is aligned with the microlenses resulting in coded sub-images. In their design each transform code is a shifted Hadamard matrix or a quantized cosine transform, with elements ones and zeros where the light corresponding to zeros is blocked by the mask. Later these coded (multiplexed) sub-images are 'inverted' to achieve a higher resolution image. The inversion was based on digital estimation of the relative shifts of the sub-images and maximum likelihood estimators. This imaging system achieves a significant (up to 50%) compression in the acquisition process.
5.2 Single-shot Compressed Sensing Camera Setup

There are three integral parts of the setup: a lenslet array, a randomly coded mask and a CMOS array. The idea of transform coding at the focal plane is borrowed from the above mentioned experiment but in this case the codes are random patterns. The random patterned mask allows implementing compressed sensing in this multi-channel multiplexing scheme.

5.2.1 Static Mask

The random patterns were etched on a standard chrome-on-glass (soda lime) substrate. First the random patterns were written on the substrate by a mask-maker (DWL66 Laser Writer), followed by etching of the chrome from appropriate locations which exposes the glass surface. The light is transmitted through glass and is blocked by the chrome surface. For proof-of-concept purposes, 32 × 32 random patterns were written on the mask which yields coded sub-images of 32 × 32 resolution. The mask contains 1024 (32 × 32) random patterns of 32 × 32 pixels arranged in 32 rows of 32 patterns. The use of 100% measurements allows to study the reconstruction quality of images with increasing number of measurements. The pattern used to write the mask is shown in Fig. 5.1, though the final mask was inverted. The border is left blank intentionally as an aid for the alignment process. At the bottom there are two rows of 8×8 raster scan basis. Fig. 5.2 shows microscope snapshot of the mask at 50X resolution. The pixel feature size on the mask is 4.7μm.
Figure 5.1  Mask of \((32 \times 32)\) random patterns.
Figure 5.2  Microscope snapshot of the mask at 50X resolution.
5.2.2 CMOS Camera

The camera used in the experiment is a USB 2.0 Monochrome 1.3 megapixel CMOS Camera built by Mightex Systems. The resolution of the CMOS camera is $1280 \times 1024$ (5:4) and pixel size is $5.2\, \mu m$. The active area of the CMOS array is $6.66\, \text{mm} \times 5.32\, \text{mm}$.

5.2.3 Lenslet Array

The lenslet array used in this experiment was MLA150-5C (Thorlabs). This lenslet array is $10\, \text{mm} \times 10\, \text{mm}$ in size with lens pitch of $150\, \mu m$ which means that it is an array of $66 \times 66$ microlenses. The lenslet array is made from fused silica and the microlenses have a plano-convex aspherical shape, arranged in a square grid. The microlens array has a chrome coating on the dead space that blocks all the light from being transmitted unless it goes through one of the microlens. This feature is important for better contrast.

Fig. 5.3 shows the initial configuration where the object is placed far away from the lenslet array. The object here is a printed letter ‘R’ which is backlit by a lamp. The lenslet array forms multiple sub-images of the object at the focal plane where these sub-images are randomly sampled by patterns on the mask. The CMOS array is located immediately behind the mask.
Figure 5.3  Single-shot compressed sensing setup.
5.2.4 Alignment

Since the pixel size for CMOS array is 5.2 μm and the feature size of each element of a 32 × 32 random pattern on the mask is 4.7 μm, utmost care is required during the alignment of the optics. The mask is mounted on a 4-axis translation stage and the lenslet array is mounted on a rotation stage which in turn is mounted on a linear translator. First the sub-images were focused on the detector plane in absence of the mask. After this the mask was installed in front of the CMOS array and aligned using the 4-axis translation stage in the plane perpendicular to the optical axis. Fig. 5.4 shows the typical coded image from this setup. The border of the picture corresponds to fully etched areas on the mask and is composed of 32 × 32 non-coded images of the object. Also the bottom two rows of the mask (besides the border) were designed to get an 8 × 8 raster image of the object.

5.3 Reconstruction

Each sub-image in the recorded image is a lower resolution coded image. The images from the border patterns represent a square image of size 150μm × 150μm. By combining a suitable number of coded sub-image we should achieve a higher resolution picture compared to the original image. The sum of the light values from each sub-image would make this system equivalent to a parallel (single-shot) CS camera allowing the reconstruction of an image with a higher number of pixels. In
addition to those summed light values there are the coded sub-images. This extra information should lead to a higher resolution image along with the compression at the acquisition stage.

5.4 Sub-pixel Shifting of the Mask

In their work titled 'Superresolution Digital Image Enhancement by Subpixel Image Translation With a Scanning Micromirror,' Kyoungsik Yu et al have proved that it is possible to achieve super-resolution by sub-pixel shifting [37, 38]. A piezoelectric driver can be employed in the single-shot CS setup in place of the scanning micromirror to implement the subpixel shifting.

5.5 Multilayer Dynamic Mask

Intuitively, it seems possible to break the random patterned mask into two or more overlapped layers. The random patterns are then generated through translations of the layers with respect to each other. This design allows to have more random patterns with the same number of lenslets resulting in an even more super-resolved image.
Figure 5.4  CMOS image of the letter ‘R’ captured through the lenslet array and the mask.
Chapter 6
Future Directions

The biggest obstacle of a practical single-pixel camera is the slower frame transfer rate to the DMD. The newer generation ALP, named ALP-3, in combination with the new DMD Discovery 3000 Board could rectify this problem. The on-board memory of ALP-3 can be customized up to 128 Gbits and this should allow faster acquisition of random projections. Also the new Discovery 3000 board has FPGA functionality which can be exploited to generate the random patterns on the fly. With these modifications it should be possible to acquire random projection for video processing.

The photovoltages corresponding to the random projections are closely distributed around their mean value. This leads to lower dynamic range and this problem gets worse as the resolution increases. By application of an appropriate offset the dynamic range can be expanded and which in turn would allow higher resolution image reconstructions.

The structured illumination setup can be modified to replace the projector lamp by a photodiode. In the lamp cavity we should be able to focus the incoming light onto a photodiode. It is essentially reversing the light path where the light from the object is focused onto the DMD in the projector and then reflected toward the lamp cavity. This design is useful in the situations where structured illumination is not
favorable.
References


13. J.n Laska, S. Kirollos, M. Duarte, T. Ragheb, R. Baraniuk, and Y. Massoud,


Proc. IS&T/SPIE Symposium on Electronic Imaging: Computational Imaging, (San Jose, CA), Jan. 2007.


