A NOVEL FRAMEWORK FOR IMAGING USING COMPRESSED SENSING

Pradeep Sen and Soheil Darabi

Advanced Graphics Lab, University of New Mexico

ABSTRACT

Recently, there has been growing interest in using compressed sensing to perform imaging. Most of these algorithms capture the image of a scene by taking projections of the imaged scene with a large set of different random patterns. Unfortunately, these methods require thousands of serial measurements in order to reconstruct a high quality image, which makes them impractical for most real-world imaging applications. In this work, we explore the idea of performing sparse image capture from a single image taken in one moment of time. Our framework measures a subset of the pixels in the photograph and uses compressed sensing algorithms to reconstruct the entire image from this data. The benefit of our approach is that we can get a high-quality image while reducing the bandwidth of the imaging device because we only read a fraction of the pixels, not the entire array. Our approach can also be used to accurately fill in the missing pixel information for sensor arrays with defective pixels. We demonstrate better reconstructions of test images using our approach than with traditional reconstruction methods.

***Index Terms*—** Compressive imaging, sampling/reconstruction

1. INTRODUCTION

It is well known that real-world images are compressible in transform domains, which is the reason for the success of transform-coding compression algorithms such as JPEG and JPEG2000. However, most imaging systems do not take advantage of this compressibility when *capturing* the image. Instead, they measure the information at every pixel and then throw out most of this information during the compression process. Naturally, this raises the question if we can measure only the "important" information in an image directly without wasting effort (time, power, bandwidth, etc.) measuring data that will be thrown away eventually during compression.

To address this issue, there has been a growing amount of interest in recent years in applying results from the field of *compressed sensing* (CS) to imaging applications, an area known as *compressive imaging*. The theory of compressed sensing states that if a signal is sparse in a transform domain, then under certain conditions it can be reconstructed *exactly* from a small set of linear measurements using tractable optimization algorithms [1, 2]. Although an in-depth review of CS is beyond the scope of this paper, we present some of its key ideas in this section to put our contribution into context. Readers seeking more detail are referred to the many papers on the subject found in the Rice University repository [3].

This work was funded by NSF CAREER award #0845396.

To understand how compressed sensing is traditionally used for imaging applications, suppose we have an n-pixel image represented by vector $\mathbf{x} \in \mathbb{R}^n$. We can write the transform of the image as $\hat{\mathbf{x}} = \mathbf{\Psi}\mathbf{x}$, where $\mathbf{\Psi}$ is a matrix whose rows represent the compression basis. We say that $\hat{\mathbf{x}}$ is m-sparse if $\|\hat{\mathbf{x}}\|_0 \leq m$, meaning that it has at most m non-zero coefficients (where $m \ll n$ and $\|\cdot\|_0$ represents the ℓ_0 norm). Since we would like to take advantage of this sparsity in order to improve the imaging process, it seems that we would need to measure the image directly in the transform domain.

Initially, it appears that measuring the m non-zero coefficients of $\hat{\mathbf{x}}$ in the transform domain would still take n measurements, since we do not know which coefficients are the largest ones. Fortunately, the theory of compressed sensing allows us to recover the sparse vector $\hat{\mathbf{x}}$ from a smaller set of measurements under certain conditions. Specifically, we write the process of taking $k < n$ linear measurements as $\mathbf{y} = \mathbf{A}\hat{\mathbf{x}}$, where $\mathbf{y}$ is the $k \times 1$ observation vector composed of the k measurements and $\mathbf{A}$ is the $k \times n$ *measurement* matrix. Here, $\mathbf{A} = \mathbf{S}\mathbf{\Psi}^T$, where $\mathbf{S}$ a $k \times n$ *sampling* matrix which specifies the linear combination of basis functions measured at every step and $\mathbf{\Psi}^T$ is obviously the inverse transform. If the sampling matrix $\mathbf{S}$ and the compression matrix $\mathbf{\Psi}$ are incoherent and the number of measurements $k > 2m$, it has been shown that we can recover $\hat{\mathbf{x}}$ exactly by solving the following ℓ_0-minimization problem [1]:

$$\min \|\hat{\mathbf{x}}\|_0 \text{ s.t. } \mathbf{y} = \mathbf{A}\hat{\mathbf{x}} \tag{1}$$

Unfortunately, this problem is difficult to solve because ℓ_0 algorithms are combinatorial in nature. However, recent work in greedy *matching pursuit* algorithms have shown that they approximate sparsity and are therefore a tractable way to approximately solve the system in Eq. 1 for $\hat{\mathbf{x}}$. In this work, we use one of these techniques called Regularized Orthogonal Matching Pursuit (ROMP) [4], which tries to select the largest m coefficients through a greedy iterative process.

In traditional compressive imaging, we take k serial measurements (with $k < n$) where image $\mathbf{x}$ is projected onto each of the k sampling basis functions of $\mathbf{A}$ and use CS to reconstruct the sparse $\hat{\mathbf{x}}$, which can then be transformed into an approximation $\tilde{\mathbf{x}}$ of the original image $\mathbf{x}$ with relatively little loss. This process requires us to take k photographs of the scene. Unfortunately as we shall see later, the size of k is still substantially large which presents a significant obstacle for a practical implementation of compressive imaging.

2. PREVIOUS WORK

2.1. "Single-pixel" cameras

Single-pixel cameras feature a monolithic photosensor that measures one intensity value at a time, unlike the sensor arrays of conventional cameras where every pixel measures a different value simultaneously. Although single-pixel cameras have been around for a long time (e.g. the "flying spot" camera of the 1920's), the growing interest in applying compressed sensing to imaging problems has been spurred by more recent work in single-pixel camera systems.

In 2005, Sen et al. demonstrated the first single-pixel camera from a DLP projector and a single photosensor, a process which they called "dual photography" [5]. Since the theory of CS had not yet been developed, they implemented an efficient adaptive algorithm that was able to capture high-quality images with less than a thousand patterns. This work was later extended by Sen and Darabi to include compressive sensing, which substantially simplified the acquisition process [6].

In 2006, the DSP group at Rice implemented another version of the single-pixel camera by modulating the image of the scene directly onto the photosensor using a DLP device [7, 8]. This work was novel because it used the ideas of compressed sensing to efficiently capture images for the first time, without requiring an adaptive algorithm. The ability to capture projections of imaged scenes with arbitrary basis functions has led to a flurry of activity into applications of imaging in the compressed sensing research community.

2.2. Compressive Imaging Algorithms

Most compressive imaging algorithms developed to-date take a series of measurements of the image over time by projecting it onto random patterns. Extensions include block-based compressed sensing [9] and CMOS hardware to accelerate the measurement of the coefficients [10]. Haupt and Nowak compare conventional pixel sampling to CS imaging, but in their comparisons they treat the two differently: for conventional imaging they interpolate between the k pixel samples, for compressive imaging they assume k serial projections of the image [11]. Finally, there is other work in compressive imaging where the samples are taken in the frequency domain, e.g. [12]. Unfortunately, these Fourier-domain imaging algorithms are impractical for real camera applications. The fact that the majority of compressive imaging algorithms require k serial samples to reconstruct the image forces us to critically examine the practical consequences of these approaches. We discuss this in detail in the next section.

3. THE NEED FOR NEW CS IMAGING PARADIGMS

Although we are excited about the potential of compressive imaging, we must play devil's advocate and take a hard look at the practical ramifications of many of these algorithms. Unfortunately, the fundamental drawback of most these approaches is that they require k serial measurements, a number which is dependent on the sparsity m of the transformed representation. For standard images, the best-known compression bases (such as the CDF 9/7 used in JPEG2000) usually require about 3% of the coefficients for a faithful reproduction of the image. In practice, CS algorithms require $5\times$ more samples than non-zero coefficients in the compression domain, which means that for an image of reasonable quality, say 10 megapixels, we would need to measure 1,500,000 coefficients to capture a visually acceptable image. Unfortunately, these measurements must all be done *serially*, for example in the case of the "single pixel" cameras by either projecting the appropriate light pattern [6] or modulating the DLP [7] once for each measurement.

This creates a fundamental bottleneck for many of the proposed CS-imaging approaches. A conventional consumer camera, e.g. the $100 Nikon Coolpix S550, can capture a 10 Megapixel image in 1/1500 of a second *losslessly*, despite the wasted "effort" in capturing more information than is needed. The serial CS-imaging approaches, however, would need to modulate the acquisition patterns and make measurements at the rate of 2.3 GHz to acquire a *lossy* version of the same image. Furthermore, since each sampling pattern (the rows of $\mathbf{S}$) is 10 megapixels in size, the bandwidth to drive the sampling modulator is considerable. Assuming the best-case scenario of binary sampling patterns, the resulting bandwidth would be 2.8 PB/sec. This is one of the reasons that high-speed single-pixel camera implementations using CS have very low resolutions and a very small set of fixed patterns. For example, the terahertz single-pixel camera system [13] uses 32×32 sampling patterns physically printed on a PCB. These fundamental bottlenecks present serious engineering challenges to practical implementations of these algorithms.

On the other hand, one of the main arguments in favor of single-pixel cameras (which has, in turn, spurred research in compressive imaging) has been that they can have more computational/optical processing per-pixel and can do things like multispectral and infrared imaging – a point raised by the first author of this paper in his dual photography work [5]. However, recent improvements in sensor technology (e.g. Krishna et al.'s work on large, focal-plane arrays of quantum dots for multispectral and infrared imaging [14]) allow for fast, high-resolution image capture using conventional "inefficient" methods that measure all n pixels. These recent developments have forced us to reconsider the true usefulness of the traditional approaches in compressive imaging.

Therefore, in this paper we propose a different paradigm for compressive imaging. Rather than applying the projections to the image over time in a serial manner, we propose to perform our sampling spatially and therefore *parallelize* the process so that all the data is captured at a single moment of time. We describe our approach in the next section.

4. SINGLE-IMAGE COMPRESSIVE IMAGING

Our approach is simple. Instead of taking k serial measurements of the imaged scene as with most CS imaging approaches, we measure a random subset of the pixels in the final image in a single moment of time and use the measured pixels to determine the missing pixel values using compressed

Fig. 1. CS imaging with Fourier basis. The **top row** shows the results of using a *sparsified* version of the LENA image. On the left is the input, where all but the 2% largest Fourier coefficients have been forced to zero. On the right is the reconstruction using CS from 25% of pixels samples. The two are virtually identical (PSNR = 87.8dB). On the **bottom row**, we show the problem with using this approach on *real* images. On the left we reconstruct the original Lena image using simple interpolation from 25% of the pixels (PSNR = 32.2dB). On the right, we reconstruct from the same samples using CS and a Fourier basis (PSNR = 27.4dB). These results were initially discouraging, since CS is easily beat by the far simpler interpolation.

sensing. In other words, given k random pixel samples, we use sparsity in a transform domain to determine the values of the missing pixels. To do this, we write our measurements as $\mathbf{y} = \mathbf{S}\mathbf{x}$ where $\mathbf{S}$ is a $k \times n$ point-sampling matrix, a matrix with a single "1" in each row and up to a single "1" in each column. This measurement equation just takes random pixel samples of $\mathbf{x}$ and observes them at $\mathbf{y}$. To get it in the form described in Sec. 1, we substitute $\mathbf{x} = \mathbf{\Psi}^T\hat{\mathbf{x}}$ and get $\mathbf{y} = \mathbf{S}\mathbf{\Psi}^T\hat{\mathbf{x}}$, which we will solve using CS algorithms.

Initially, it seems that we need to use the Fourier basis for $\mathbf{\Psi}$ because the Fourier basis is incoherent with the point-sampling basis $\mathbf{S}$ (wavelets, on the other hand, are not). This works well for artificially-sparsified images and is able to produce near-perfect reconstructions (see top row of Fig. 1). Since this is difficult to do with other interpolation techniques, it motivates our overall approach. However, when dealing with real images that are not sparsified, the algorithm fails, performing worse than bilinear interpolation (see bottom row Fig. 1). The problem is that the sparsity of real images in the Fourier domain is not large enough, and since we typically need about 4 to $5\times$ more samples than the number of sparse coefficients in practice, this approach does not work.

In this paper, we propose a way to overcome this problem and develop an algorithm that performs significantly better than bilinear interpolation. Specifically, we do this by using a wavelet basis for compression, which offers increased

Fig. 2. **(left)** CS reconstruction with wavelets from 25% pixel samples has problems since $\mathbf{\Psi}$ is not incoherent with point-sampled $\mathbf{S}$ (PSNR = 13.4dB). **(right)** Result of using the filtered wavelet formulation presented from the same samples (PSNR = 32.7dB). This new approach beats traditional interpolation techniques.

sparsity. However, we must be careful since wavelets are not incoherent with the spatial point-samples. After all, the better a transform is at defining localized features, the more coherent it will be with the spikes of a point-sample basis and the less likely it will work with the CS framework. The result of trying to use a wavelet basis by itself as $\mathbf{\Psi}$ is shown in Fig. 2.

To reduce the coherence between $\mathbf{S}$ and $\mathbf{\Psi}$, we modify our measurement equation to include an invertible filtering process before sampling: $\mathbf{y} = \mathbf{S}\mathbf{x} = \mathbf{S}\mathbf{\Phi}^{-1}\mathbf{x}_b$, where $\mathbf{\Phi}$ is an invertible blurring filter and $\mathbf{x}_b = \mathbf{\Phi}\mathbf{x}$ is a blurred version of the image. We can now add the wavelet compression basis back in and solve for the sparsest $\hat{\mathbf{x}}_b$ that meets the constraint:

$$\min \|\hat{\mathbf{x}}\|_0 \text{ s.t. } \mathbf{y} = \mathbf{S}\mathbf{\Phi}^{-1}\mathbf{\Psi}^T\hat{\mathbf{x}}_\mathbf{b} = \mathbf{A}\hat{\mathbf{x}}_\mathbf{b} \qquad (2)$$

with the approaches we described earlier. In this work, we set $\mathbf{\Phi}$ to be a Gaussian filter, which we apply by multiplying with a diagonal Gaussian matrix $\mathbf{G}$ in the frequency domain: $\mathbf{\Phi} = \mathcal{F}^T\mathbf{G}\mathcal{F}$, where $\mathcal{F}$ is the Fourier transform and $\mathbf{G}$ has a Gaussian function along its diagonal. To compute our inverse filter, we need to evaluate $\mathbf{G}^{-1}$, which is also a diagonal matrix. Since the inversion of Gaussian curves is prone to noise amplification, we use a linear Wiener filter to invert the Gaussian [15] which means that the diagonal elements of our inverse matrix are $\mathbf{G}^{-1}_{i,i} = \mathbf{G}_{i,i}/(\mathbf{G}^2_{i,i} + \lambda)$. In our experiments, we set $\lambda = 0.4$. The variance σ^2 of the Gaussian function in $\mathbf{G}$ depends on the sampling rate, e.g. for the images of this paper with 25% of samples $\sigma^2 = 3.38 \times 10^3$.

After the addition of the filter, our measurement matrix now is composed of point-samples $\mathbf{S}$ and the filtered wavelet matrix $\mathbf{\Phi}^{-1}\mathbf{\Psi}^T$. Since the coherence between a sampling basis and compression basis can be found by taking the maximum inner product between any two basis elements times $\sqrt{n}$ [16], we can check if our formulation has indeed increased incoherence. For the filtered wavelet, the coherence with the point-sampling basis is 158.3, where the coherence without the filter is 261.6. This reduction in coherence allows us to apply CS to this problem.

5. IMPLEMENTATION AND RESULTS

The framework of Sec. 4 results in a simple algorithm that can be performed in three steps:

Step 1: Measure k random pixels of the image sensor

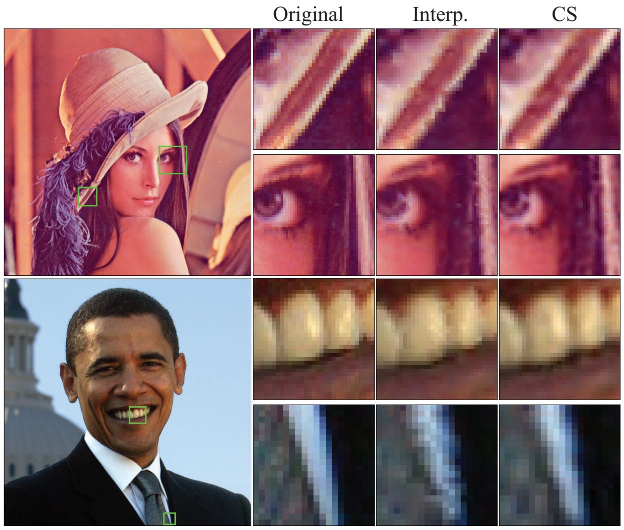

Fig. 3. Images reconstructed from 25% pixels using interpolation and our compressive imaging framework. For each inset we show the original, the interpolated result, and the result of our CS method. The PNSR results are LENA Interp: 31.2dB, CS: 31.4dB, OBAMA Interp: 33.7dB, CS: 33.9dB.

Step 2: Use CS greedy algorithm to estimate $\hat{\mathbf{x}}_b$ in Eq. 2
Step 3: Compute the desired image $\mathbf{x} = \mathbf{\Phi}^{-1}\mathbf{\Psi}^{-1}\hat{\mathbf{x}}_b$

The first step involves the selection of k random pixels. We found that a Poisson-disk distribution [17], where all the pixel samples are separated by at least a fixed distance, works better than completely random samples. Once the pixels were selected, we simulate the imaging process by simply measuring the original image at these pixel locations. We then solve for the missing pixels with Regularized Orthogonal Matching Pursuit (ROMP), using Daubechies-8 (DB-8) wavelets for compression. A MATLAB implementation of ROMP is available from Vershynin's website [4]. Once $\hat{\mathbf{x}}_b$ is found, we can compute the desired $\mathbf{x}$. The entire reconstruction algorithm was written in C and takes 100 seconds on a laptop with 2.2GHz processor to process a 512×512 image.

To compare our results, we need to interpolate the entire image from the non-uniform pixel locations. One traditional way to do this is to tessellate the samples into a triangular mesh using Delaunay triangulation and then bilinearly interpolate across each triangle to fill in the missing pixels. We show our results for two color images in Fig. 3. A graphical comparison showing the PSNR for the two algorithms is shown in Fig. 4. We can see from these results that the proposed algorithm works better than the traditional approach.

6. DISCUSSION

The proposed algorithm raises the possibility of practical, high-resolution camera systems which only measure a fraction of the pixel samples in order to reduce the bandwidth of the read-out circuitry. Furthermore, it could also find values of missing pixels caused by defects in the sensor, reducing the cost of manufacture. Instead of using traditional interpolation techniques to determine these unmeasured pixels, we use compressed sensing to leverage the compressibility of the image in the wavelet domain. The resulting images have crisper edges than the interpolated results. We hope that these initial results encourage others to explore paradigms for compressive imaging based on a single image measurement.

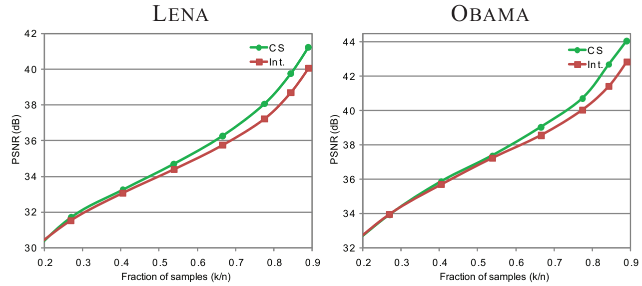

Fig. 4. Error curves as a function of the number of samples for two test images. Our approach results in better reconstruction than bilinear interpolation as the sampling rate increases.

7. REFERENCES

[1] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Info. Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[3] "Rice University Compressive Sensing Resources website," 2009, http://www.dsp.ece.rice.edu/cs/.

[4] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," 2007, Preprint.

[5] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik Lensch, "Dual Photography," *ACM Transactions on Graphics*, vol. 24, no. 3, pp. 745–755, 2005.

[6] Pradeep Sen and Soheil Darabi, "Compressive Dual Photography," *Computer Graphics Forum*, vol. 28, no. 2, pp. 609 – 618, 2009.

[7] M.B. Wakin, J.N. Laska, M.F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K.F. Kelly, and R.G. Baraniuk, "An architecture for compressive imaging," in *IEEE Intl. Conf. on Image Processing (ICIP)*, Oct. 2006.

[8] D. Takhar, J. Laska, M. Wakin, M. Duarte, D. Baron, S. Sarvotham, K. Kelly, and R. Baraniuk, "A new compressive imaging camera architecture using optical-domain compression," in *Proc. of Computational Imaging IV*. SPIE, Jan. 2006.

[9] Lu Gan, "Block compressed sensing of natural images," in *Proc. of Conf. on Digital Signal Processing (DSP)*, 2007.

[10] L. Jacques, P. Vandergheynst, A. Bibet, V. Majidzadeh, A. Schmid, and Y. Leblebici, "CMOS compressed imaging by random comvolution," 2008, Preprint.

[11] J. Haupt and R. Nowak, "Compressive sampling vs conventional imaging," in *IEEE Intl. Conf. on Image Processing (ICIP)*, 2006.

[12] K. Egiazarian, A. Foi, and V. Katkovnik, "Compressed sensing image reconstruction via recursive spatially adaptive filtering," in *IEEE Intl. Conf. on Image Processing (ICIP)*, 2007.

[13] W. L. Chan, K. Charan, D. Takhar, K. Kelly, R. G. Baraniuk, and D. M. Mittleman, "A single-pixel terahertz imaging system based on compressed sensing," *Applied Physics Letters*, vol. 93, no. 12, 2008.

[14] S. Krishna, S. D. Gunapala, S. V. Bandara, C. Hill, and D. Z. Ting, "Quantum dot based infrared focal plane arrays," *Proceedings of the IEEE*, vol. 95, no. 9, pp. 1838–1852, Sep 2007.

[15] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001.

[16] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. on Info. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov 2001.

[17] Robert Cook, "Stochastic sampling in computer graphics," *ACM Trans. Graph.*, vol. 5, no. 1, pp. 51–72, 1986.