Author manuscript, published in "SAMPTA'09, Marseille : France (2009)"

hal-00453304, version 1 - 4 Feb 2010

Adaptive compressed image sensing based on wavelet modeling and direct sampling

Shay Deutsch (1), Amir Averbuch (1) and Shai Dekel (2)

(1) Tel Aviv University. Israel
(2) GE Healthcare, Israel
shayseut@post.tau.ac.il, Shai.dekel@ge.com, amir@math.tau.ac.il

Abstract:

We present Adaptive Direct Sampling (ADS), an algorithm for image acquisition and compression which does not require the data to be sampled at its highest resolution. In some cases, our approach simplifies and improves upon the existing methodology of Compressed Sensing (CS), by replacing the 'universal' acquisition of pseudo-random measurements with a direct and fast method of adaptive wavelet coefficient acquisition. The main advantages of this direct approach are that the decoding algorithm is significantly faster and that it allows more control over the compressed image quality, in particular, the sharpness of edges.

1. Introduction

Compressed Sensing (CS) [1, 3, 4, 6] is an approach to simultaneous sensing and compression which provides mathematical tools that, when coupled with specific acquisition hardware architectures, can perhaps reduce the acquired dataset sizes, without reducing the resolution or quality of the compressed signal. CS builds on the work of Candès, Romberg, and Tao [4] and Donoho [6] who showed that a signal having a sparse representation in one basis can be reconstructed from a small number of non-adaptive linear projections onto a second basis that is incoherent with the first. The mathematical framework of CS is as follows:

Consider a signal $x \in \mathbb{R}^N$ that is k-sparse in the basis Ψ for $\mathbb{R}^N$. In terms of matrix representation we have $\Psi x = f$, in which f can be well approximated using only $k << N$ non zero entries and Ψ is called the sparse basis matrix. Consider also an $n \times N$ measurement matrix Φ, where the rows of Φ are incoherent with the columns of Ψ. The CS theory states that such a good approximation of signal x can be reconstructed by taking only $n = O(k \log N)$ linear, non adaptive measurements as follows: [1, 3]:

$$y = \Phi x , \qquad (1.1)$$

where y represents an $n \times 1$ sampled vector. Working under this 'sparsity' assumption an approximation to x can be reconstructed from y by 'sparsity' minimization, such as l_1 minimization

$$\min_{\Phi\Psi^{-1} f = y} \|f\|_{l_1} \qquad (1.2)$$

1.2 The "single pixel" camera

For imaging applications, the CS framework has been applied within a new experimental architecture for a 'single pixel' digital camera [10]. The CS camera replaces the CCD and CMOS acquisition technologies by a ***Digital Micro-mirror Device (DMD)***. The DMD consists of an array of electrostatically actuated micro-mirrors where each mirror of the array is suspended above an individual SRAM cell. In [10] the rows of the CS sampling matrix Φ are a sequence of n pseudo-random binary masks, where each mask is actually a 'scrambled' configuration of the DMD array (see also [2]). Thus, the measurement vector y, is composed of dot-products of the digital image x with pseudo-random masks. At the core of the decoding process, that takes place at the viewing device, there is a minimization algorithm solving (1.2). Once a solution is computed, one obtains from it an approximate 'reconstructed' image by applying the transform Ψ to the coefficients. The CS architecture of [10] has few significant drawbacks:

1. Poor control over the quality of the output compressed image: the CS architecture of [10] is not adaptive and the number of measurements is determined before the acquisition process begins, with no feedback during the acquisition process on the progressive quality.
2. Computationally intensive sampling process: Dense measurement matrices such as the sampling operator of the random binary pattern are not feasible because of the huge space and multiplication time requirements. Note that in the one single pixel camera, the sampling operator is based on the random binary pattern, which requires a huge memory and a high computation cost. For example, to get 512×512 image with 64k measurements (25% sampling rate) a random binary operator requires nearly a gigabyte of storage and Giga-flop operations, which makes the recovery almost impossible [14]. The designing of an efficiently measurement basis was proposed [14, 16] by using highly sparse measurements operators, which solve the infeasibility of Gaussian measurement matrix or a random binary masks such as in the one pixel camera. Note, however, in [16], the trade-off between acquisition time and visual quality. To obtain good visual quality, when using TV minimization (which significantly increase the decoding time, compared to LP decoding time)

hal-00453304, version 1 - 4 Feb 2010

recovery times of a 256×256 'boat' image are around 60 min.

3. Computationally intensive reconstruction algorithm: It is known that all the algorithms for the minimization (1.2) are very computationally intensive.

2. Direct and adaptive image sensing

Our proposed architecture aims to overcome the drawbacks of the existing CS approach and achieve the following design goals:

1. An acquisition process that captures n measurements, with $n << N$ and $n = O(k)$, where N is the dimension of the full high-resolution image, assumed to be 'k-sparse'. The acquisition process is allowed to adaptively take more measurements if needed to achieve some compressed image target quality.
2. A decoding process which is not more computationally intensive than the existing algorithm in use today such as JPEG or JPEG2000 decoding.

We now present our ADS approach: Instead of acquiring the visual data using a representation that is incoherent with wavelets, we sample directly in the wavelet domain. We use the DMD array architecture in a very different way than in [10]:

1. Any wavelet coefficient is computed from two measurements of the DMD array.

2. We take advantage of the 'feedback' architecture of the DMD where we make decisions on future measurements based on values of existing measurements. This adaptive sampling process relies on a well-known modeling of image edges using a wavelet coefficient tree-structure and so decisions on which wavelet coefficients should be sampled next are based on the values of wavelet coefficients obtained so far [8, 9]. First we explain how the DMD architecture can be used to calculate a wavelet coefficient from two DMD measurements. Modeling an image as a function $f \in L_2(\mathbb{R}^2)$, we have the wavelet representation $f(x) = \sum_{e,j,l} \langle f, \tilde{\psi}^e_{j,l} \rangle \psi^e_{j,l}$, where $e = 1,2,3$ is the subband, $j \in \mathbb{Z}$ the scale and $l \in \mathbb{Z}^2$ the shift. For orthonormal wavelets $\tilde{\psi}^e_{j,l} = \psi^e_{j,l}$. If we consider the Haar basis as an example, than a bivariate Haar wavelet coefficient of type 1 can be computed as follows

$$\langle f, \psi^1_{j,l} \rangle = 2^j \left(\int_{2^{-j} l_1}^{2^{-j}(l_1+1)} \int_{2^{-j} l_2}^{2^{-j}(l_2+1/2)} f(x_1, x_2) dx_1 dx_2 - \int_{2^{-j} l_1}^{2^{-j}(l_1+1)} \int_{2^{-j}(l_2+1/2)}^{2^{-j}(l_2+1)} f(x_1, x_2) dx_1 dx_2 \right), \qquad (2.1)$$

i.e., the difference of pixel sums over two neighboring dyadic rectangles multiplied by 2^j. By Similar computation we can sample the Haar wavelet coefficients of the second and third kinds with two measurements. Moreover, there exist DMD arrays with micro-mirrors that can produce a grayscale value, not just 0 or 1 (contemporary DMD can produce 1024 grayscale value). We can use these devices for computation of arbitrary wavelet transforms, where the computation of each coefficient requires only two measurements, since the result of any real-valued functional g acting on the data can be computed as a difference of two 'positive' g_+, g_- 'functionals', i.e. ,where the coefficients are positive: $g = g_+ - g_-$, $g_+, g_- \geq 0$.

3. Modeling of image edges by wavelet tree-Structures and the ADS algorithm

Most of the significant wavelet coefficients are located in the vicinity of edges. Wavelets can be regarded as multi-scale local edge detectors, where the absolute value of a wavelet coefficient corresponds to the local strength of the edge. We impose the tree-structure of the wavelet coefficients. Due to the analysis properties of wavelets, coefficient values tend to persist through scale. A large wavelet coefficient in magnitude generally indicates the presence of singularity inside its support. A small wavelet coefficient generally indicates a smooth region. We use this nesting property and acquire wavelet coefficients in the higher resolutions if their parent is found to be significant. For further detection of singularities at fine scales, we estimate the Lipschitz exponent.

3.1 The Lipschitz exponent

Our goal is to estimate the significance of wavelet coefficients that were not sampled yet, using values of coefficients that were already sampled. To this end we use the well known characterization of local Lipschitz smoothness by the decay of wavelet coefficients across scales [12]. A function f is said to be Lipschitz α in the neighborhood of (x_1, x_2) if there exists ε_1 and ε_2 as well as $A > 0$ such that for any $h_1 < \varepsilon_1$ and $h_2 < \varepsilon_2$

$$|f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2)| \leq A(h_1^2 + h_2^2)^{\alpha/2} \qquad (3.1)$$

We actually use a subtler, 'directional' notion of local Lipschitz smoothness. So, for example, for the horizontal subband, $e = 1$, we defined local α_1 Horizontal Lipschitz smoothness by the minimal $A > 0$ satisfying for $h_1 < \varepsilon_1$

$$|f(x_1 + h_1, x_2) - f(x_1, x_2)| \leq A h_1^{\alpha_1}.$$

If the function is locally α_e Lipschitz at (x_1, x_2) then for any wavelet $\tilde{\psi}^e_{j,l}$ whose support contains (x_1, x_2), we have that $\left|\langle f, \tilde{\psi}^e_{j,l} \rangle\right| \leq C\left(2^j\right)^{\alpha_e}$. By taking the logarithm we have

$$\log_2 \left|\langle f, \tilde{\psi}^e_{j,l} \rangle\right| \leq \alpha j + \log_2(C). \qquad (3.3)$$

hal-00453304, version 1 - 4 Feb 2010

Thus the Lipschitz exponents can be determined from the slope of the decay of $\log_2 \left| \left\langle f, \tilde{\psi}^{e}_{j,l} \right\rangle \right|$ across scales (see also [15]). These slopes are considered measurements of local singularities, such that when $0 < \alpha_e < 1$ a function f has a directional singularity which increases as $\alpha_e \to 0$. Thus we estimate the existence of local directional singularities and the significance of unsampled coefficients at high scales, using estimates of local directional Lipschitz exponents from wavelet coefficients that were already sampled.

3.2 The ADS Algorithm

Our adaptive CS algorithm works as follows:

1. Acquire the values of all low-resolution coefficients up to a certain low-resolution J. Each computation is done using two DMD array measurements as in (2.1). In one embodiment the initial resolution J can be selected as $\left\lfloor \frac{\log_2 N}{2} \right\rfloor + const$. In any case, J should be bigger if the image is bigger. Note that the total number of coefficients at resolutions $\geq J$ is $2^{2(1-J)}N$, which is a small fraction of N.

2. Initialize a 'sampling queue' containing the indices of each of the four children of significant coefficients at the resolution J. Thus for a significant coefficient with index (e, J, l), we add to the queue the coefficients with indices: $(e, J-1, (2l_1, 2l_2))$, $(e, J-1, (2l_1, 2l_2+1))$, $(e, J-1, (2l_1+1, 2l_2))$ and $(e, J-1, (2l_1+1, 2l_2+1))$.

3. Process the sampling queue until it is exhausted as follows:

a. Sample the wavelet coefficient corresponding to the index (e, j, l) at the beginning of the queue using two DMD array measurements (see Section 2).

b. Add to the end of the queue the indices of the coefficient's four children, only if one of the following holds:

(i) The coefficient is at a resolution $j > J-2$ and the coefficient's absolute value is greater than a given threshold t_{low}.

(ii) The coefficient is at resolution $1 < j \leq J-2$ and the corresponding estimated absolute value of its children using the local Lipschitz exponent method (see Section 3.1) is greater than a given threshold t_{high}.

c. Remove the processed index from the queue and go to step (a).

In a way, our algorithm can be regarded as an adaptive edge acquisition device where the acquisition resolution increases only in the vicinity of edges! Observe that the algorithm is output sensitive. Its time complexity is of the order n where n is the total number of computed wavelet coefficients, which can be substantially smaller than the number of pixels N. The number of samples is influenced by the size of the thresholds used by the algorithm in step 3.b. It is also important to understand that the number of samples is influenced by the amount of visual activity in the image. If there are more significant edges in the image, then their detection at lower resolutions will lead to adding higher resolution sampling to the queue.

4. Experimental results

To evaluate our approach, we use the optimal k-term wavelet approximation as a benchmark. It is well known [5] that for a given image with N pixels, the optimal orthonormal wavelet approximation using only k coefficients is obtained using the k largest coefficients

$$\left| \left\langle f, \psi^{e_1}_{j_1,l_1} \right\rangle \right| \geq \left| \left\langle f, \psi^{e_2}_{j_2,l_2} \right\rangle \right| \geq \left| \left\langle f, \psi^{e_3}_{j_3,l_3} \right\rangle \right| \geq \cdots,$$

$$\left\| f - \sum_{i=1}^{k} \left\langle f, \psi^{e_i}_{j_i,l_i} \right\rangle \psi^{e_i}_{j_i,l_i} \right\|_{L_2(\mathbb{R}^2)} = \min_{\#\Lambda = k} \left\| f - \sum_{(e,j,l)\in\Lambda} \left\langle f, \psi^{e}_{j,l} \right\rangle \psi^{e}_{j,l} \right\|_{L_2(\mathbb{R}^2)}.$$

For biorthogonal wavelets this 'greedy' approach gives a near-best result, i.e. within a constant factor of the optimal k-term approximation. One can apply thresholding and construct a k-term approximation using only coefficients whose absolute value is above the threshold, which still requires the order of N computations. In contrast, our ADS algorithm is output sensitive and requires only order of n computations. To simulate our algorithm in software, we first pre-compute the entire wavelet transform of a given image. However, we strictly follow the recipe of our ADS algorithm and extract a wavelet coefficient from the pre-computed coefficient matrix only if its index was added to the adaptive sampling queue. In fig 1(a) we see a 'benchmark' near-best 7000-term biorthogonal [9,7] wavelet approximation of the Lena image, extracted from the 'full' wavelet representation by thresholding. In fig 1(b) we see a 6782-term approximation extracted from an ADS adaptive sampling process with n=12796 sampled wavelet coefficient.

(a) 7000-term

(b) ADS 6782-term

Fig.1. (a) Near-best 7000-term [9,7] approximation computed from the 'full' wavelet representation N=262,144, PSNR=31 dB (b) ADS 6782-term [9,7] approximation, extracted from n=12,796 adaptive wavelet samples, PSNR=28.7 dB.

5. Conclusion

We present an architecture that acquires and compresses high resolution visual data, without fully sampling the entire data at its highest resolution. By sampling in the wavelet domain we are able to acquire low resolution coefficients within a small number of measurements. We then exploit the wavelet tree structure to build an adaptive sampling process of the detail wavelet coefficients. Experimental results show good visual and PSNR results with a small number of measurements. The coefficients acquired by the ADS algorithm can be streamed into a tree-based wavelet compression algorithm whose decoding time is significantly faster then the solution of (1.2).

REFERENCES

1. R. Baraniuk, Compressive Sensing, Lecture Notes in IEEE Signal Processing Magazine, Vol. 24, No. 4, pp. 118-120, July 2007.
2. R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, A simple proof of the restricted isometry property for random matrices, Constructive Approximation 28 (2008), 253-263.
3. E. Candès, Compressive sampling, Proc. International Congress of Mathematics, 3 (2006), 1433-1452.
4. E. Candès, J. Romberg, and T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inf. Theory 52 (2006), 489–509.
5. R. DeVore, Nonlinear approximation, Acta Numerica7 (1998), 50-51.
6. D. Donoho, Compressed sensing, IEEE Trans. Information Theory, 52 (2006), 1289-1306.
7. C. La and M. Do, Signal reconstruction using sparse tree representations, Proc. SPIE Wavelets XI, San Diego, September 2005.
8. A. Said and W. Pearlman, A new fast and efficient image codec based on set partitioning in hierarchical trees, IEEE Trans. Circuits Syst. Video Tech., 6 (1996), 243-250.
9. J. Shapiro, Embedded image coding using zerotrees of wavelet coefficients, IEEE Trans. Signal Process. 41 (1993), 3445-3462.
10 .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, Proc. of Computational Imaging IV , SPIE, 2006.
11. S. Dekel, Adaptive compressed image sensing based on wavelet-trees, report 2008.
12. S. Mallat, "a wavelet tour of signal processing".
13. S. Mallat and W. L. Hwang, "singularity detection and processing with wavelets," IEEE Trans. Inf. Theory 38,617-642(1992).
14. L. Gan, T. Do, T. Tran , Fast compressive imaging using scrambled Hadamard transform ensemble, preprint 2008.
15. Z. Chen and M. A. Karim, Forest representation of wavelet transforms and feature detection, Opt. Eng. 39 (2000), 1194–1202.
16. R. Berinde, P. Indik, sparse recovery using sparse random matrices, Tech. Report of MIT 2008.
17. F Rooms, A. Pizurica and, W. Philips, estimating image blur in the wavelet domain, IEEE Benelux Signal Processing Symposium (SPS-2002).

hal-00453304, version 1 - 4 Feb 2010