Image Representation by Compressive Sensing for Visual Sensor Networks

Bing Han[*1], *Feng Wu*[2], *Dapeng Wu*[1]

[1] Department of Electrical and Computer Engineering, University of Florida, Gainesville, 32611

[2] Microsoft Research Asia, Beijing, 100080

Correspondence author: Prof. Dapeng Wu, wu@ece.ufl.edu, http://www.wu.ece.ufl.edu

***Abstract* — This paper addresses the image representation problem in visual sensor networks. We propose a new image representation method for visual sensor networks based on compressive sensing (CS). CS is a new sampling method for sparse signals, which is able to compress the input data in the sampling process. Combining both signal sampling and data compression, CS is more capable of image representation for reducing the computation complexity in image/video encoder in visual sensor networks where computation resource is extremely limited. Since CS is more efficient for sparse signals, in our scheme, the input image is firstly decomposed into two components, i.e., dense and sparse components; then the dense component is encoded by the traditional approach (JPEG or JPEG 2000) while the sparse component is encoded by a CS technique. In order to improve the rate distortion performance, we leverage the strong correlation between dense and sparse components by using a piecewise autoregressive model to construct a prediction of the sparse component from the corresponding dense component. Given the measurements and the prediction of the sparse component as initial guess, we use projection onto convex set (POCS) to reconstruct the sparse component. Our method considerably reduces the number of random measurements needed for CS reconstruction and the decoding computational complexity, compared to the existing CS methods. In addition, our experimental results show that our method may achieves up to 2dB gain in PSNR over the existing CS based schemes, for the same number of measurements.**

***Index Terms*— Image representation, compressive sensing, random sampling, projection onto convex sets, convex optimization**

I. INTRODUCTION

According to the sampling theory and the Nyquist-Shannon sample theorem, exact reconstruction of a continuous-

[*] This work has been done while the author is with Microsoft Research Asia.

time signal from its samples is possible if the signal is band-limited and the sampling rate is more than twice the signal bandwidth [1]. In recent years, a new theory Compressive Sensing (CS) also referred as Compressed Sensing or Compressive Sampling, has been proposed as a more efficient sampling scheme.

The theoretical framework of CS was developed by Candes et al. [2] and Donoho [3]. The CS principle claims that a sparse signal can be recovered from a small number of random linear measurements. It means that it is possible to reconstruct the signal x, which is sparse in Ψ domain, by a small number of measurements, $y = \Phi x$, where the measurement ensemble Φ obeys the restricted isometry hypothesis [2]. The recovery procedure is to minimize the l_1 norm of the signal f in Ψ domain, which is shown to be a linear programming problem and could also be cast as a convex optimization problem.

Comparing with the traditional Nyquist-Shannon sampling theory, the CS theory provides a great reduction of sampling rate, power consumption and computational complexity to acquire and represent a sparse signal. R. Baraniuk et al. [36] has proposed hardware to support the new theory of Compressive Imaging (CI). It shows that CI is able to obtain an image with a single detection element while measuring the image/video fewer times than the number of pixels, which can significantly reduce the computation required for video acquisition/encoding.

CS has been connected with many other fields such as information theory [23]-[25], high dimension geometry [26]-[29], statistical signal processing [30]-[32], and data streaming algorithms [33]-[35]. Besides the connections to the existing theories, CS has also been used for some practical areas, such as compressive imaging [36]-[39], medical imaging [40]-[43], distributed compressed sensing [44]-[46] , analog to information conversion [47]-[50] and bio-sensing [51].

Most of the recent papers study two problems of CS. One is to find the optimal sampling ensembles and study the methods for fast implementation of the CS ensembles [4]-[7]. The other one is to develop fast and practical reconstruction algorithms to recover the signal and suppress the noise introduced by CS [5]-[12].

In [5], Donoho et al. report several families of random measurement ensembles which behave equivalently, including random spherical, random signs, partial Fourier and partial Hadamard. The further work on measurement ensemble studies the optimal measurement matrices, which allow us to recover as many entries of the signal as possible with as few as measurements. In [6], Baraniuk et al. prove the existence of optimal CS measurement ensembles and they have certain universality with respect to the sparsity inducing basis. Michael further proposes an average measurement of the mutual-coherence of the effective dictionary and demonstrated that it leads to better CS reconstruction performance [7].

Among the reconstruction algorithms, Basis Pursuit (BP) is the first one proposed to solve this problem [2][3]. Orthogonal Matching Pursuit (OMP) [13] and Tree Matching Pursuit (TMP) [14][15] are proposed for fast reconstruction. In [18], Donoho et al. show that the Homotopy method runs much more rapidly than general purposed linear programming solvers when sufficient sparsity is presented. In [19], Kim et al. describe a specialized interior-point method for solving large-scale l_1-regularized least square programming. In order to suppress the noise and increase the computation speed, Figueiredo et al. [22] propose a gradient projection algorithm for the bound-constrained quadratic programming formulation of CS problem.

In this paper, we use CS to represent visual data and propose a new image representation scheme for visual sensor networks. Comparing to JPEG2000, CS is more suitable for applications in sensor networks because the sensors are resource constrained. Different from the previous work on compressive imaging [30]-[32], which treats the input image as a whole, we decompose the visual data into two components: a dense component and a sparse one. We process the dense component by traditional method and the sparse component by CS. The decomposition helps to generate a sparser signal which is more suitable for CS. Furthermore, there exists a strong correlation between these two components because the decomposition is usually not completely orthogonal. Even if the decomposition is orthogonal, since nature images present a strong visual correlation, some vision technologies could be used to derive useful information from one component to predict the other. Thus, another advantage of our work is that we use the correlation of the two components in the reconstruction procedure, which could help to reduce the measurements needed for reconstruction and to use less computation time for the same accurate result. We implement a POCS based algorithm to use the correlation between dense component and sparse component of the signal in the recovery procedure.

This paper is organized as follows. Section II gives a brief overview of CS. Section III discusses how to apply CS to practical signals. Section IV proposes a scheme for image representation using CS. The proposed reconstruction algorithm is discussed in details. The experimental results are presented in Section V and Section VI concludes this paper.

II. OVERVIEW OF COMPRESSIVE SENSING

This section briefly reviews the CS theory. Suppose we have a one dimensional signal $x \in \mathbf{R}^N$, which could be represented sparsely in a certain domain by the transform matrix $\Psi := [\psi_1, \quad \psi_2, \quad ..., \quad \psi_N]$ with each column vector $\psi_n \in \mathbf{R}^N$, $n = 1, 2, ..., N$. If there are only K non-zero coefficients in the Ψ domain, we can say that x is exactly K-sparse. Then we have the following question: if we already know the signal to be sampled is K-sparse, why should we take N measurements, where $N >> K$.

CS gives an answer to the above question that it is possible to recover the K-sparse signal x by taking M random measurements which is much less than N. In order to take CS measurements, we first let Φ denote a M by N matrix with $M << N$. The measurement matrix Φ obeys the restricted isometry hypothesis [8]. Then the measurements are obtained by a linear system

$$y = \Phi x. \tag{1}$$

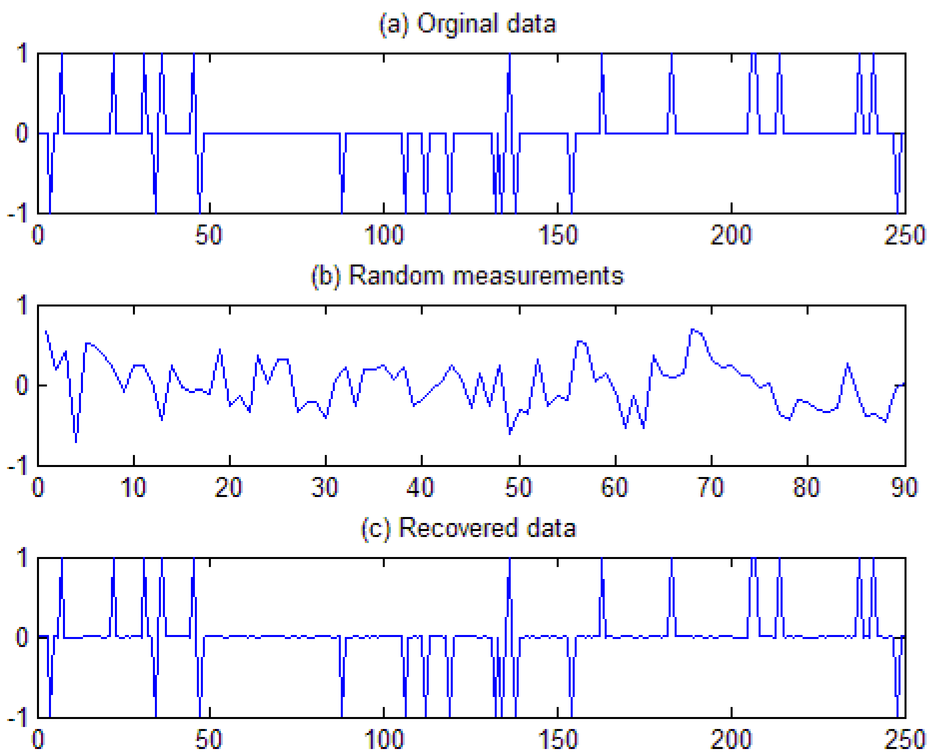
Different from the traditional sampling, the CS measurement measures the information of the whole signal at one time. Therefore, each measurement contains a little information from all elements of the original signal. In this way, CS is able to reduce the sampling rate, power consumption, and computational complexity of the visual sensors. The CS theory states that the signal could be exactly recovered if the number of measurements M satisfies the condition $M \geq Const \cdot K \cdot \log N$ [2], where *Const* is an over-measuring factor more than 1.

Since y is a lower dimension vector compared to x, it is impossible to recover x directly by applying the inverse transform of Φ to y. The signal is reconstructed by solving the following optimization problem

$$(L_1) \quad \min \left\| \Psi^T \tilde{x} \right\|_1 \quad \textit{subject to} \quad y = \Phi \tilde{x} \tag{2}$$

The reconstructed signal $\tilde{x}$ is among all signals generating the same measured data, which has transform coefficients with the minimal l_1 norm. The reconstruction can be cast as a linear programming problem.

Figure 1 gives an example to explain the signal recovery by CS reconstruction. Figure 1 (a) shows the original signal that has 250 samples with only 25 non zeros, which is very sparse. Figure 1 (b) is the CS measurements by a Gaussian ensemble Φ, where there are only 90 measurements. Figure 1 (c) is the recovery result from the CS measurements by a POCS based algorithm. It is clear that the signal is well recovered.



[redacted]

III. THE PROPOSED IMAGE REPRESENTATION

The theoretical results in the previous section indicate that it is possible to recover an unknown sparse signal from a few random measurements taken by a non-adaptive measurement ensemble. However, there are still some issues to be considered when we apply CS into practical applications, such as image representation. First, the visual data is not completely sparse but compressible in a certain transform domain, such as DCT and DWT, which means we have to consider noise in the sample and reconstruction process. Secondly, we need a fast algorithm to apply the CS matrix Φ rapidly because usually the visual data is of high dimension.

In this section, we will discuss the first problem from three aspects: CS reconstruction error bound in noise environment, image decomposition and the correlation between the dense and sparse components. The second problem will be discussed in the next section.

A. Reconstruction error bounds

In general, the visual data, such as natural image, is the reflection of the complicated real world and could not be supported in a transform domain on a set of small size, which means, it could hardly be represented exactly sparsely in any transform domain. However, the visual data is still compressible. We will evaluate the CS reconstruction of compressible signals through the error bound in this sub-section.

Let us still use the example of signal x, we fix an orthonormal basis Ψ and the signal could be represented by $f(x)=\Psi x$. As well known, the compressibility of the signal x relates to the decay rate of the coefficients of f. If the coefficients of f obey a power law, which means $1 \le n \le N$, $|f_n(x)| \le R \cdot n^{-1/p}$, where R and p are constant and depend on the signal, $f_n(x)$ is the n-th largest coefficients in $f(x)$, the difference between x and approximate signal x_K ,which is obtained by keeping the largest K coefficients, obeys

$$\|x-x_K\|_2 \le C_1 \cdot R \cdot K^{-r}, \qquad r=1/p-1/2, \tag{3}$$

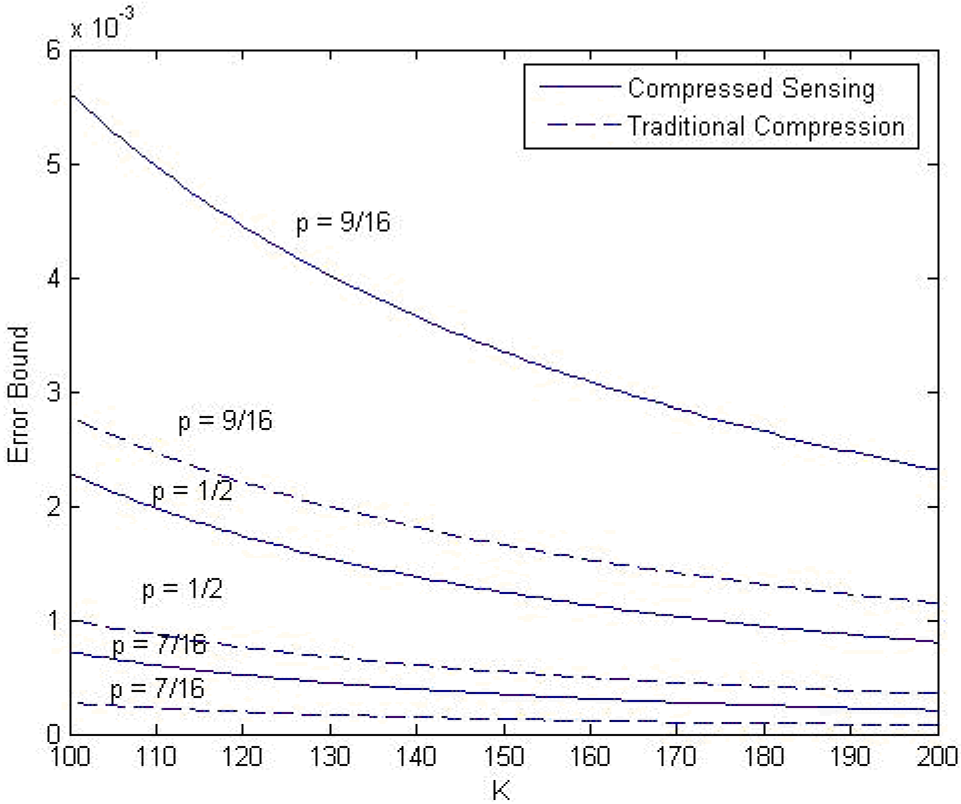
where C_1 is a constant.

In Donoho's paper [3], he gives the error bound of the CS approximation reconstructed from the non-adaptive measurements. Given M measurements of the signal x, the reconstruction error bound takes the form

$$\|x-x_{cs}\|_2 \le C_2 \cdot R \cdot (M/\log(N))^{-r}, \qquad r=1/p-1/2, \tag{4}$$

where C_2 is constant.

If the signal is exactly sparse, the two error bounds should be at the same level; otherwise there will be a big gap between them. We compare the error bound of CS for signals with different decay rates and draw the error bounds. C_1, C_2, R, M and N are fixed in the experiment and only p is variable, which controls the speed of decay: the smaller p is, the faster it decays. The experimental result is depicted in Figure 2, where K is the number of non-zero coefficients for reconstruction. For each group, the dash line represents the error bound of the traditional best K method and the solid line represents the result of CS. One can observe that, when p = 7/16, the error bound of CS is very close to that of the traditional best K method; however, when p = 9/16, there is a big gap between these two error bounds. Obviously, the performance of CS reconstruction relies on the decay rate of the signal. The faster it decays, the better it recovers.



B. *Image Decomposition*

The above analysis on error bound indicates that sparser signal is more suitable for CS. In this sub-section, we will discuss image decomposition. Generally, we want to decompose the input data into dense and sparse components. The reason is that a sparse signal could be represented more efficiently by CS. For the dense component, it is better to use other method to represent it. A similar work has been done in [52] which decomposes the image into texture and piecewise smooth parts. However, it is not suitable for our problem. Let us first take one dimensional signal as example. Suppose we have a vector t with length N and a pre-chosen basis Ξ, we can represent t as

$$t = \sum_{i=1}^{N} \alpha_i \Xi_i \,. \tag{5}$$

Then we can further decompose it into two components:

$$t = \sum_{i=1}^{N} \alpha_i \Xi_i = \sum_{i=1}^{T} \alpha_{1i} \Xi_{1i} + \sum_{i=1}^{N-T} \alpha_{2i} \Xi_{2i} \tag{6}$$

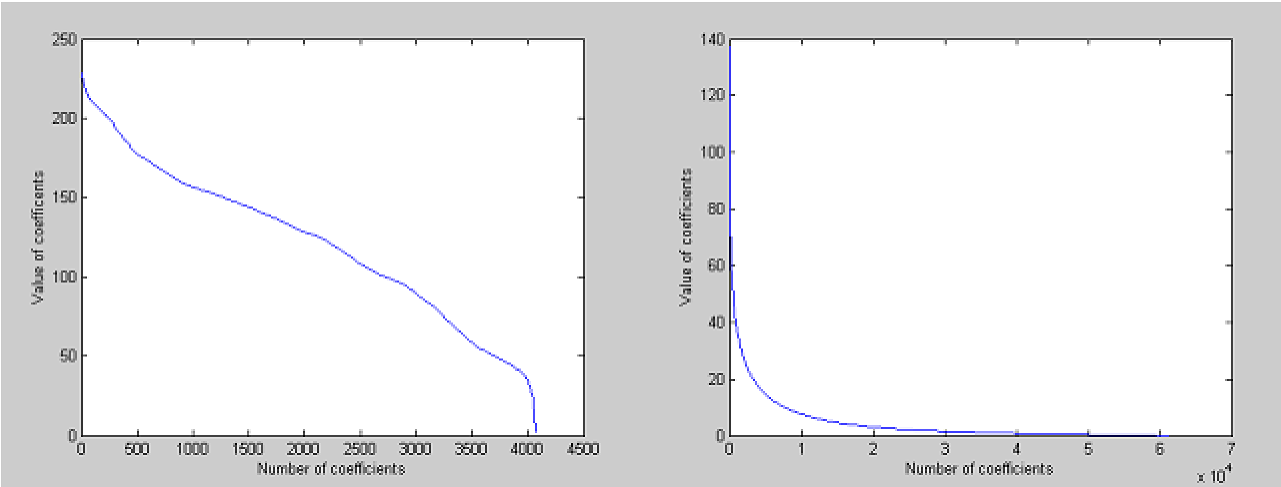
Let $t_D = \sum_{i=1}^{T} \alpha_{1i} \Xi_{1i}$ and $t_S = \sum_{i=1}^{N-T} \alpha_{2i} \Xi_{2i}$ denote dense and sparse components, respectively. However, for natural images, it is hard to find such a sparse component. In this paper, we use the power law model to define the sparsity of the signal. Take t_D for example, we reorder its coefficients α_{1i} and compute the parameter p in the model $|\alpha_{1i}| \le R \cdot i^{-\frac{1}{p}}, 1 \le i \le T$, where R and p are constant and only depend on t_D. With this model, we use parameter p to represent the sparsity of the signal. The smaller p is, the faster the coefficients decay and the sparser the signal is.

The research to find a function Ψ which leads to better decomposition of image is a hot area in recent years [52]~[54]. Previous work has proved that wavelet transform is well designed for sparse representation of natural images. Expand the image I in the wavelet basis

$$I = \sum_k \alpha_{1 j_0,k} W_{j_0,k} + \sum_{j=j_1}^{j_2} \sum_k \alpha_{2 j,k} W_{j,k} , \quad (7)$$

where $W_{j_0,k}$ and $W_{j,k}$ are wavelet at different scales. For simplicity, in this paper, we take the lowest band of wavelet transform $I_D = \sum_k \alpha_{1 j_0,k} W_{j_0,k}$ as dense component and the other bands $I_S = \sum_{j=j_1}^{j_2} \sum_k \alpha_{2 j,k} W_{j,k}$ as sparse component, where j_0 is the coarsest scale, j_1 is the next scale and j_2 is the finest scale.

Applying the power law model to α_1 and α_2, respectively, we can find that there is a big difference between p_1 and p_2, which means the sparsity of the two components differs tremendously. Figure 3 shows the result of wavelet transform of Lena by three-level decomposition. Figure 4 depicts the decay curves of dense and sparse components, respectively. The left curve represents coefficients α_1 of signal I_D and shows that α_1 decays slowly. On the contrary, the curve on the right side representing α_2 indicates that α_2 decays much faster than α_1. It means that component I_S is much 'sparser' than component I_D and is also more suitable for Compressed Sensing.



C. Correlation between Sparse and Dense Components

Although the input image could be decomposed into the dense and sparse components, one can still observe that there exists a strong visual correlation between them. Therefore, it is possible to use the dense component to predict the original image, as well as the sparse component. The recent development on adaptive interpolation provides an effective tool to solve this problem. The adaptive interpolation describe the image as a 2D piecewise autoregressive (PAR) model, namely,

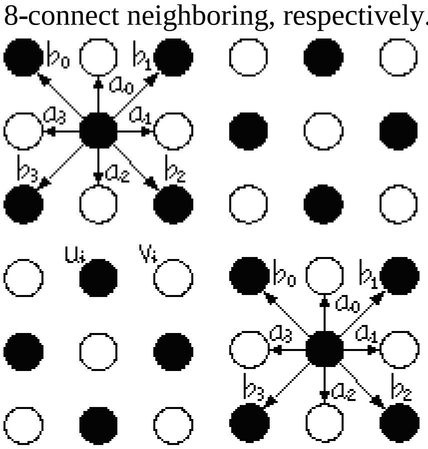
$$X_{i,j} = \sum_{(m,n)\in B_{i,j}} \beta_{m,n} X_{i+m,j+n} + \upsilon_{i,j} \,, \tag{8}$$

where (i, j) is the pixel to be interpolated, $B_{i,j}$ is the window centered at pixel (i, j) and $\upsilon_{i,j}$ is a random perturbation independent of pixel (i, j) and the image signal.

In Zhang et al.'s paper [55], they formulate the interpolation problem as an optimization problem:

$$\min_{a,b,v} \left\{ \sum_{i\in B} \left\| u_i - \sum_{t=0}^{3} \left(a_t v_{i-t}^{(4)} + b_t u_{i-t}^{(8)} \right) \right\|_2 + \sum_{i\in B} \left\| v_i - \sum_{t=0}^{3} \left(a_t u_{i-t}^{(4)} + b_t v_{i-t}^{(8)} \right) \right\|_2 \right\} \tag{9}$$

where I_u is the image to be interpolated and the I_v is the original image. $u_i \in I_u$ and $v_i \in I_v$ are the pixels of the image I_u and I_v respectively. B is the window size. The superscripts (4) and (8) indicate 4-connect neighboring and 8-connect neighboring, respectively. Fig. 5 depicts the sample relationships in (9).



Zhang et al. [55] also give a linear least-square solution to this problem, which estimate neighboring pixels simultaneously in window B.

$$\min_{v}\left\{\sum_{i\in B}\left\|u_i-\sum_{t=0}^{3}\left(\hat{a}_t v_{i-t}^{(4)}+\hat{b}_t u_{i-t}^{(8)}\right)\right\|_2+\sum_{i\in B}\left\|v_i-\sum_{t=0}^{3}\left(\hat{a}_t u_{i-t}^{(4)}+\hat{b}_t v_{i-t}^{(8)}\right)\right\|_2\right\} \tag{10}$$

where $\hat{a}$ and $\hat{b}$ are estimated from I_u.

In Section 3.2, we have discussed the decomposition of a natural image signal I into a dense component I_D and a sparse component I_S. With the above adaptive interpolation algorithm, we can take I_D as a sub-sample of the original image I and interpolate it. From the above discussion, we can get a prediction of I by solving (9). This prediction $\hat{I}$ recovers most of the high frequency information and could be again expanded by wavelet basis as in (7)

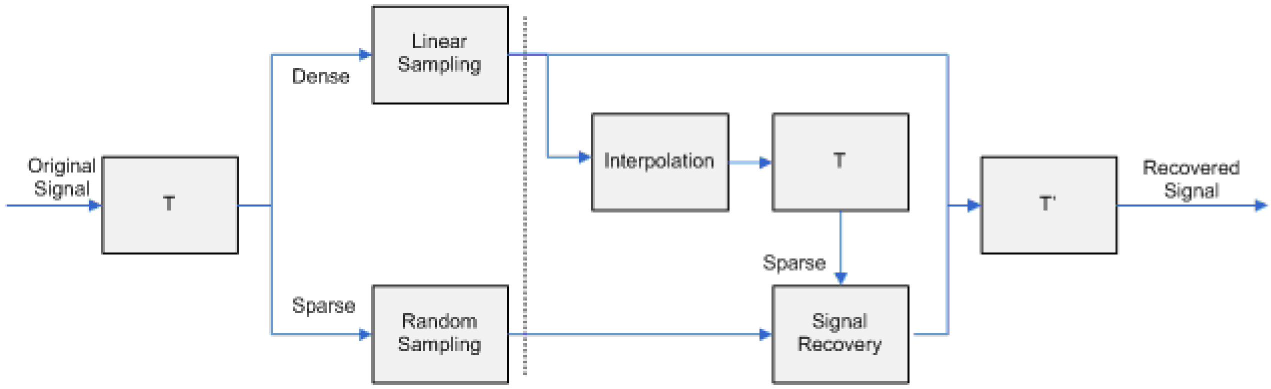
$$\hat{I}=\sum_{k}\hat{\alpha}_{1j_0,k}W_{j_0,k}+\sum_{j=j_1}^{j_2}\sum_{k}\hat{\alpha}_{2j,k}W_{j,k} \tag{11}$$

The new sparse component $\hat{I}_S=\sum_{j=j_1}^{j_2}\sum_{k}\hat{\alpha}_{2j,k}W_{j,k}$ can be used as a good prediction for $I_S=\sum_{j=j_1}^{j_2}\sum_{k}\alpha_{2j,k}W_{j,k}$.

IV. THE PRACTICAL SIGNAL RECONSTRUCTION

A. *Image Representation Scheme*

The theoretical results of CS indicates that it is possible to recover the signal from M random measurements as well as from the best K coefficients in the sparsity domain Ψ. In this section, we present a new image representation scheme based on CS. Different from the previous CS based image representation methods [9][38], which take the input image as a non-separate signal, in our scheme, the input image is decomposed into two components, dense and sparse components. Then the two components are sampled using different methods respectively. For the dense component, we use the traditional approach. In other words, we sample the dense component pixel by pixel. While for the sparse component, we apply CS by random sampling. The proposed scheme is depicted in Figure 6.



The input image I is first decomposed into a dense component I_D and a sparse component I_S through a transform T, where T could be wavelet, curvelet, or any other transforms. In our scheme, we use discrete wavelet transform W to decompose the image. We expand the image I in the following way

$$I = I_D + I_S = \sum_k \alpha_{1j_0,k} W_{j_0,k} + \sum_{j=j_1}^{j_2} \sum_k \alpha_{2j,k} W_{j,k} \,, \tag{12}$$

We take the lowest band of wavelet transform $I_D = \sum_k \alpha_{1j_0,k} W_{j_0,k}$ as dense component and the other bands $I_S = \sum_{j=j_1}^{j_2} \sum_k \alpha_{2j,k} W_{j,k}$ as sparse component, where j_0 is the coarsest scale, j_1 is the next scale and j_2 is the finest scale and $W_{j_0,k}$ and $W_{j,k}$ are wavelet at different scales.

The two components are measured separately. We use $\tilde{I}_D$ and $\tilde{I}_S$ to represent the measurements of the signal components respectively. We take direct measurements to the dense component

$$\tilde{I}_D = \left(\left\langle W_{j_0,k}, I \right\rangle : 0 \le k \le 2^{j_0} \right). \tag{13}$$

where j_0 is the preset coarse level of the wavelet transform. Normally, j_0 is set to be 1.

In order to take measurements of the sparse components I_S, we use a Gaussian random ensemble Φ. As we know, the dimension of the input image is very high, so directly applying the 2D Gaussian random ensemble Φ to the signal I_S is not practical. In order to apply the random ensemble more efficiently, we need to regroup the signal and take a block based sampling strategy. We first divide I_S into several groups by scales and then reorder it into a number of vectors of the same dimension. In this way, we can take random measurements to the vector with moderate size instead of the tremendous size.

$$\tilde{x}_i = \Phi x_i, \quad 1 \le i \le n \tag{14}$$

where n is the number of groups and x_i is the i-th group of I_S.

In the decoder side, we have to recover the signal separately. Since the dense component is measured pixel by pixel, $\tilde{I}_D$ is exactly the wavelet coefficients of I_D. Therefore, we could directly apply inverse transform W' to $\tilde{I}_D$ go get I_D. In order to recover I_S, we need to solve the optimization problem in (2). In Candes's paper [9], they propose a projection onto convex sets (POCS) algorithm to reconstruct the original signal from the random measurements. We follow their approach and improve the algorithm by using prediction of I_S as the starting point of the iterations. The prediction $\hat{I}$ of the input image could be obtained by adaptive interpolation of the dense component using (10). Then we could apply wavelet transform to $\hat{I}$ and get $\hat{I}_S$ as the prediction of I_S. We will discuss the details of the CS recovery algorithm in the next sub-section.

At last, with $\tilde{I}_D$ and the recovered $\tilde{I}_S{}'$, we could recover the original signal through the inverse transform W'.

B. POCS-based Reconstruction

In the first section, we have discussed several shortcomings of iterative interior point methods. In Candes's paper [9], they propose a different recovery procedure, which requires a small amount of priori information of the signal to be recovered but cost less computation in each iteration. In their algorithm, they assume the l_1 norm of the recovered signal is known. We further develop this algorithm and propose a new reconstruction method based on POCS and prediction from adaptive interpolation.

Since x in (2) is the unique solution by the Compressed Sensing principle, we are able to claim that the l_1-ball $B=\{\tilde{x}:\|\tilde{x}\|_{l_1}\leq\|x\|_{l_1}\}$ and the hyperplane $H=\{\tilde{x}:\Phi\tilde{x}=y\}$ meet at exactly one point: $B\bigcap H=\{x\}$. Because both B and H are convex, x can be recovered by an alternate projections onto convex sets algorithm [56].

As we have discussed in last section, we have decomposed the input image into dense and sparse components. Therefore, we are able to utilize the dense component $\tilde{I}_D$ to predict the sparse component I_S by adaptive interpolation. The prediction helps in two aspects: first, it could be used as the initialization of the iteration. As known to all, the initialization is very important to an iterative algorithm and the initial value need be in a certain space for final convergence at local optimal. Secondly, the prediction can be used as a reference which helps for converging more rapidly and accurately.

From the starting point of $\tilde{x}_i$, we iterate by alternating projections onto H, then onto B. The algorithm is guaranteed to converge to a point in $B\bigcap H$ [56].

To find the closest vector $\tilde{x}_i^H$ in H to an arbitrary $\tilde{x}_i$, we apply the equation

$$\tilde{x}_i^H=\tilde{x}_i+\Phi^*(\Phi\Phi^*)^{-1}(y-\Phi\tilde{x}_i)+\lambda(\tilde{x}_i-x_i^p) \tag{15}$$

The step length for $\tilde{x}_i^H$ combines two parts, one is computed from direct projection and the other part is from the difference between $\tilde{x}_i$ and the prediction signal x_i^p. λ is a user specified value and $\lambda\geq 0$. The value of λ stands for the importance of how close the solution is to the prediction signal.

In order to project the vector $\tilde{x}_i^H$ onto the l_1 ball B, we apply a soft thresholding operation

$$\tilde{x}_i^B(n)=\begin{cases}\tilde{x}_i^H(n)-\gamma & \tilde{x}_i^H(n)>\gamma\\ 0 & |\tilde{x}_i^H(n)|\leq\gamma\\ \tilde{x}_i^H(n)+\gamma & \tilde{x}_i^H(n)<-\gamma\end{cases} \tag{16}$$

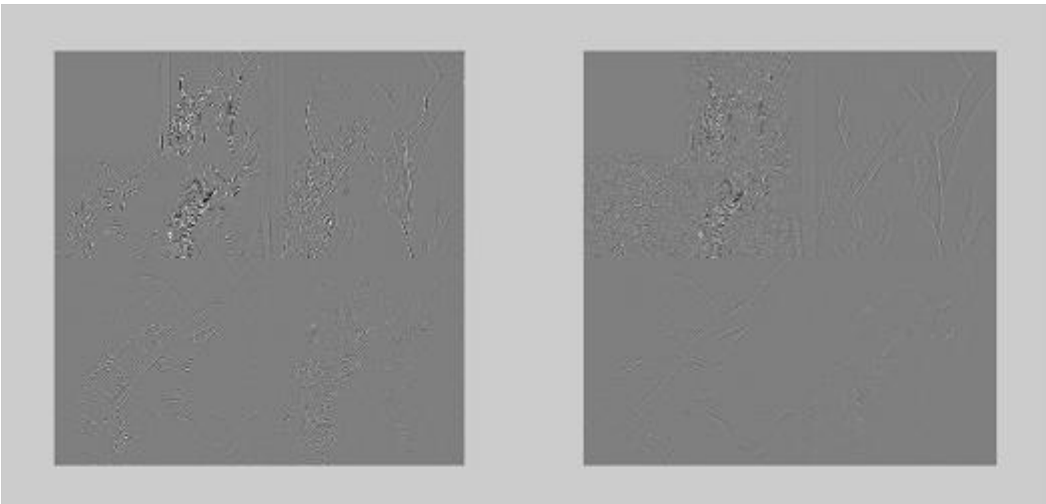
In order to determine the threshold γ such that $\|\tilde{x}_i^B\|_{l_1}\leq\|x_i\|_{l_1}$, we sort the coefficients by absolute value and perform a linear search.

V. Experimental Results

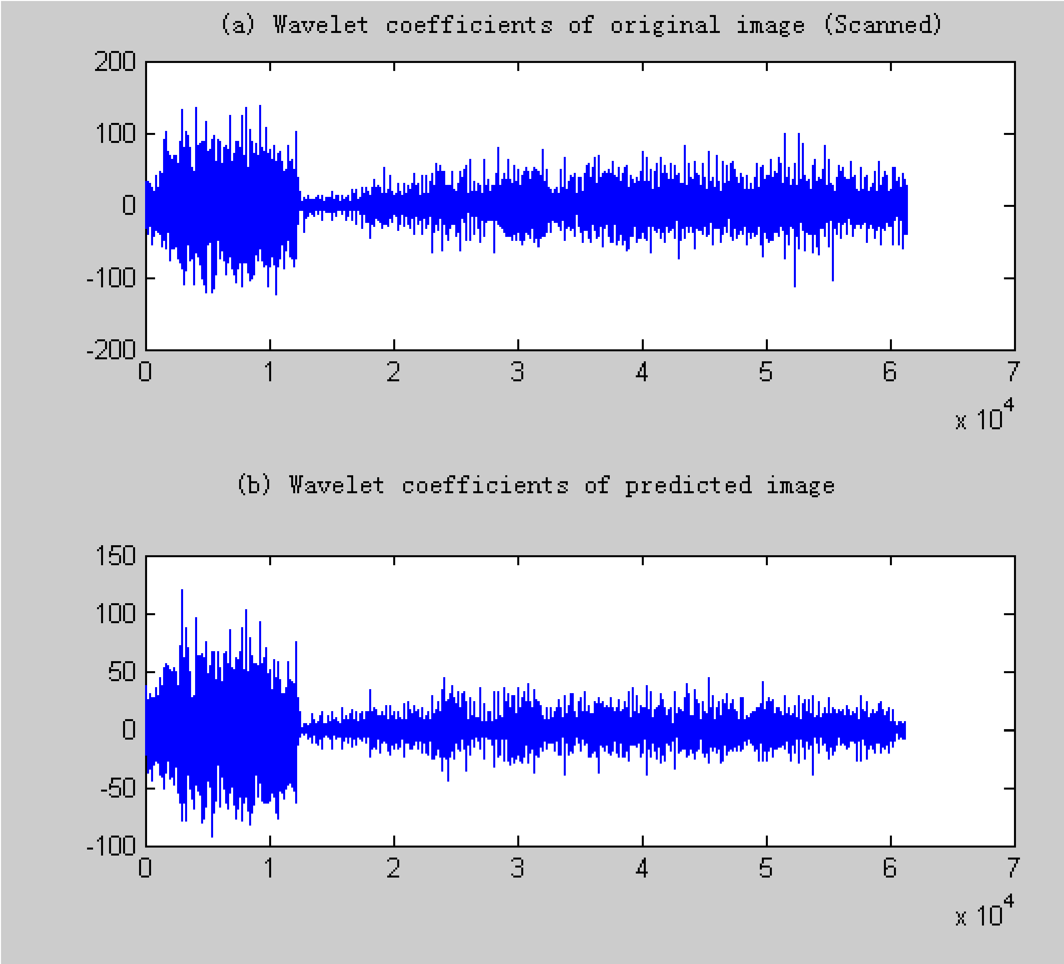
Our experiment includes four parts. The first part verifies the correlation between sparse component and the prediction of it from dense component. In the second part, we test our algorithm on several natural images and compare our results with [9] and [38]. In the third part, we further compare the performance of our algorithm with [9] by error reduction with the same iterations. At last, we show some recovered images.

Figure 7 depicts the comparison between the prediction of the sparse component and the corresponding component.

In order to verify the similarity between the original sparse component and the prediction from the dense component, we first decompose the input signal, where the input data is an image patch from *Lena*, into two components. Figure 7(a) shows the sparse component in wavelet transform domain. Then we interpolate the dense component to get a prediction of the original image patch. The prediction is further decomposed into sparse and dense components. The predicted sparse component is shown in Figure 7(b).



In order to compare them more clearly, we scan the image by rows from low frequency to high frequency. Then the image is scanned into a one-dimensional vector and depicted in Figure 8. The upper subplot shows the original signal and the lower subplot shows the predicted signal. From the comparison in Figure 8, it is clear that the prediction by interpolating the dense component is very close to the original signal. It means that we can use less iteration for convergence in CS reconstruction and it is also possible to use the prediction as a weighted constraint in the reconstruction algorithm. The images we used in our experiment are *Lena*, *Boat*, *Cameraman* and *Peppers* as in [9] and [38]. As we described in Section 4, for each image I, we compute the recovery image $\tilde{I}$ based on (11)~(13). The experiment tests for different sparsity of images and different sizes of measurement ensemble Φ. The recovery error $\|I-\tilde{I}\|_2$ is measured by PSNR in db and tabulated in Table 1.



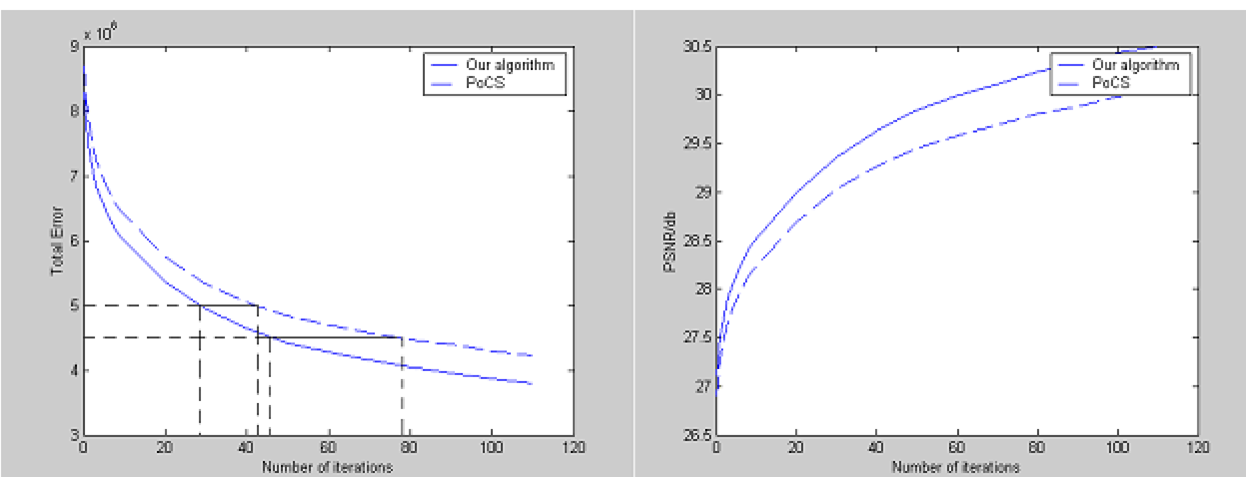
Measurements		10000	15000	20000	25000
Lena	[9]	26.5	28.7	30.4	32.1
	[38]	26.5	28.6	30.6	32.2
	ours	30.0	31.8	33.0	34.2
Boats	[9]	26.7	29.8	31.8	33.7
	[38]	27.0	29.9	32.5	34.8
	ours	29.1	31.0	33.0	34.4
Cameraman	[9]	26.2	28.7	30.9	33.0
	[38]	24.0	26.1	27.9	29.4
	ours	26.3	28.5	29.7	30.7
Peppers	[9]	21.6	25.3	27.5	29.4
	[38]	27.2	30.3	32.7	34.7
	ours	27.4	30.7	32.7	34.6

Why not using bit-rate but rather number of measurements? The reason is that the traditional quantization scheme is not suitable for CS measurements. How to quantize the CS measurements itself is a crucial problem to be solved [5]. So it is out of the scope of this paper. Therefore, instead of using bit-rate, we follow [9] to use number of measurements.

The results are compared to the work in [9] and [38]. It is clear that for *Lena*, *Boats* and *Peppers*, our algorithm outperforms the reconstruction algorithms in [9] and [38]. It means we need fewer measurements to achieve the same PSNR or achieve better PSNR with the same number of measurements. Our number of measurements in Table 1 includes the measurements of both I_D and I_S. There are two reasons for the improvement: first, we remove the dense component from the input data by decomposition and it increases the sampling efficiency; secondly, by introducing the prediction of the recovered data, we are able to constrain the unknown signal in a small space near the original signal and require fewer measurements for convergence. For the image *Cameraman*, our algorithm loses a little bit to [9] in high PSNR end. However, comparing to [38], our recovery result is more accurate.

In this paper, we will not compare our results to JPEG2000. The reason is that the purpose of this paper is to explore an image representation scheme which is specifically for visual sensor networks. As an emerging research area, it consists of tiny wireless-enabled battery operated cameras. Because the visual data incur high computation and communication energy, the sensors remain relatively resource constrained. Therefore, as the most energy consuming compression scheme, JPEG2000 is not suitable for application in sensor networks. Comparing to JPEG2000, CS can compress the data while sampling, so it consumes less energy [16].

In the third experiment, we use *Peppers* as an example to test the convergence time. We compare error reduction by total error and PSNR by iterations separately in Figure 10. The solid line represents our algorithm and the dash line plot the results from [9]. In Figure 9, it shows that we need 28 iterations comparing to 43 with the algorithm in [9] to reduce the error to 5×10^6 and 46 iterations comparing to 78 to reduce the error to the error 4.5×10^6. The prediction greatly reduces the number of iterations for the same PSNR, which means we use less time for signal reconstruction and therefore again saves energy needed.



At last, we give some examples of recovered images in Figure 10 and Figure 11. The recovered *Lena* and *Boats* are both obtained from 20000 measurements in total.

VI. Conclusion

In this paper, we proposed a new image representation scheme for visual sensor networks. From our work, it shows that decomposition of the input data before CS measuring is important and especially useful for reducing the number of measurements and recovery iterations. First, the decomposition removes the unnecessary component, which is not suitable for CS recovery. Second, the prediction by interpolation of the dense component helps the recovery procedure of the sparse component. Comparing to previous methods, our proposed scheme achieves higher PSNR with the same number of measurements and iterations. Since our work is proposed for visual sensor networks, we did not compare our algorithm with JPEG2000, which requires high computational complexity.

ACKNOWLEDGMENT

The authors would like to thank Prof Mingting Sun, Prof Changwen Chen and Prof Xiaolin Wu for many valuable discussions and suggestions.

REFERENCES

[1] A. J. Jerri, "The Shannon sampling theorem – its various extensions and applications: a tutorial review", Proceeding of the IEEE, vol. 65, pp. 1565-1596, 1977.

[2] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information", IEEE Trans. on Information Theory, vol. 52, pp. 489-509, 2006.

[3] D. Donoho, "Compressed sensing", IEEE Trans. on Information Theory, vol. 52, pp. 1289-1306, 2006.

[4] E. Candès, "Compressive sampling", Proc. International Congress of Mathematics, vol. 3, pp. 1433-1452, Madrid, Spain, 2006.

[5] D. Donoho and Y. Tsaig, "Extensions of compressed sensing", Signal Processing, vol. 86, pp. 533-548, March 2006.

[6] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "The Johnson-Lindenstrauss lemma meets compressed sensing", preprint, 2006.

[7] M. Elad, "Optimized projections for compressed sensing", IEEE Trans. on Signal Processing, vol. 55, pp. 5695-5702, December, 2007.

[8] E. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies", IEEE Trans. on Information Theory, vol. 52, pp. 5406-5425, December 2006.

[9] E. Candès and J. Romberg, "Practical signal recovery from random projections", Wavelet Applications in Signal and Image Processing XI, Proc. SPIE Conf., pp. 5914, 2004.

[10] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best k-term approximation", preprint.

[11] G. Peyré, "Best basis compressed sensing", Scale Space and Variational Methods in Computer Vision, Springer Berlin / Heidelberg, vol. 4485, pp. 80-91, 2007.

[12] H. Rauhut, K. Schass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries", preprint, 2006.

[13] J. Tropp and A. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit", preprint, 2005.

[14] M. Duarte, M. Wakin, and R. Baraniuk, "Fast reconstruction of piecewise smooth signals from random projections", Proc. SPARS Workshop, November 2005.

[15] C. La and M. Do, "Signal reconstruction using sparse tree representations", Proc. SPIE Wavelets XI, San Diego, California, September 2005.

[16] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Compressive wireless sensing", Int. Conf. on Information Processing in Sensor Networks (IPSN), Nashville, Tennessee, April 2006

[17] S. Sarvotham, D. Baron, and R. Baraniuk, "Sudocodes - Fast measurement and reconstruction of sparse signals", Proc. IEEE Int. Symposium on Information Theory, Seattle, Washington, July 2006.

[18] D. Donoho and Y. Tsaig, "Fast solution of ell-1-norm minimization problems when the solution may be sparse", preprint, 2006.

[19] S. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "A method for large-scale ell-1-regularized least squares problems with applications in signal processing and statistics", preprint, 2007.

[20] T. Blumensath and M. E. Davies, "Iterative thresholding for sparse approximations", The Journal of Fourier Analysis and Applications, accepted for publication.

[21] T. Blumensath and M. E. Davies, "Gradient pursuits", preprint, 2007.

[22] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems", IEEE Journal of Selected Topics in Signal Processing, vol.1, no.4, pp.586-597, Dec. 2007.

[23] E. Candès and T. Tao, "Decoding by linear programming", IEEE Trans. on Information Theory, vol. 51, no. 12, pp. 4203 - 4215, December 2005.

[24] S. Sarvotham, D. Baron, and R. Baraniuk, "Measurements vs. bits: Compressed sensing meets information theory", Proc. Allerton Conference on Communication, Control, and Computing, Monticello, IL, September 2006.

[25] M. Wainwright, "Information-theoretic bounds on sparsity recovery in the high-dimensional and noisy setting", Proc. IEEE Int. Symposium on Information Theory, Nice, France, June 2007

[26] D. Donoho, "High-dimensional centrally-symmetric polytopes with neighborliness proportional to dimension", Disc. Comput. Geometry, vol. 35, no. 4, pp. 617-652, 2006.

[27] D. Donoho and J. Tanner, "Neighborliness of randomly-projected simplices in high dimensions", Proc. National Academy of Sciences, vol. 102, no. 27, pp. 9452-9457, 2005.

[28] D. Donoho and J. Tanner, "Counting faces of randomly-projected polytopes when the projection radically lowers dimension", preprint.

[29] R. Baraniuk and M. Wakin, "Random projections of smooth manifolds", Foundations of Computational Mathematics, Springer New York, December, 2007.

[30] J. Haupt, R. Castro, R. Nowak, G. Fudge, and A. Yeh, "Compressive sampling for signal classification", Proc. Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, California, October 2006.

[31] M. Davenport, M. Duarte, M. Wakin, J. Laska, D. Takhar, K. Kelly, and R. Baraniuk, "The smashed filter for compressive classification and target recognition", Proc. of Computational Imaging V at SPIE Electronic Imaging, San Jose, California, January 2007.

[32] M. Duarte, M. Davenport, M. Wakin, and R. Baraniuk, "Sparse signal detection from incoherent projections", Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, Toulouse, France, May 2006.

[33] A. Gilbert, M. Strauss, J. Tropp, and R. Vershynin, "Algorithmic linear dimension reduction in the ell-1 norm for sparse vectors", Proc. of the 44th Annual Allerton Conference on Communication, Control and Computing, Allerton, September 2006.

[34] A. Gilbert, M. Strauss, J. Tropp, and R. Vershynin, "One sketch for all: Fast algorithms for compressed sensing", In Proceedings of the Thirty-Ninth Annual ACM Symposium on theory of Computing (San Diego, California, USA, June 11 - 13, 2007). STOC '07. ACM, New York, NY, 237-246.

[35] S. Muthukrishnan, "Some algorithmic problems and results in compressed sensing", 44th Annual Allerton Conference, September, 2006, Allerton House, Illinois, 2006.

[36] M. Wakin, J. Laska, M. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. Kelly, and R. Baraniuk, "An architecture for compressive imaging", Proc. Int. Conf. on Image Processing, Atlanta, Georgia, October 2006.

[37] 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 at SPIE Electronic Imaging, San Jose, California, January 2006.

[38] L. Gan, "Block compressed sensing of natural images", Proc. Int. Conf. on Digital Signal Processing, Cardiff, UK, July 2007.

[39] 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 at SPIE Electronic Imaging, San Jose, California, January 2006.

[40] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging", Magnetic Resonance in Medicine, vol. 58, pp. 1182-1195, 2007.

[41] M. Lustig, J. M. Santos, D. Donoho, and J. M. Pauly, "k-t SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity", Proc. 14th. Annual Meeting of ISMRM, Seattle, Washington, May 2006.

[42] H. Jung, J. C. Ye, and E. Y. Kim, "Improved k-t BLASK and k-t SENSE using FOCUSS Phys", Med. Biol., vol. 52, pp. 3201 - 3226, 2007.

[43] J. C. Ye, "Compressed sensing shape estimation of star-shaped objects in Fourier imaging", Signal Processing Letters, IEEE, vol.14, no.10, pp.750-753, Oct. 2007.

[44] M. Rabbat, J. Haupt, A. Singh, and R. Nowak, "Decentralized compression and predistribution via randomized gossiping", Proc. Int. Conf. on Information Processing in Sensor Networks Nashville, Tennessee, April 2006.

[45] W. Wang, M. Garofalakis, and K. Ramchandran, "Distributed sparse random projections for refinable approximation", Proc. Int. Conf. on Information Processing in Sensor Networks, Cambridge, Massachusetts, April 2007.

[46] D. Baron, M.l Wakin, M. Duarte, S. Sarvotham, and R. Baraniuk, "Distributed compressed sensing", preprint, 2005.

[47] S. Kirolos, J. Laska, M. Wakin, M. Duarte, D. Baron, T. Ragheb, Y. Massoud, and R. Baraniuk, “Analog-to-information conversion via random demodulation”, Proc. IEEE Dallas Circuits and Systems Workshop, Dallas, Texas, 2006.

[48] J. Laska, S. Kirolos, Y. Massoud, R. Baraniuk, A. Gilbert, M. Iwen, and M. Strauss, “Random sampling for analog-to-information conversion of wideband signals”, Proc. IEEE Dallas Circuits and Systems Workshop, Dallas, Texas, 2006.

[49] J. Laska, S. Kirolos, M. Duarte, T. Ragheb, R. Baraniuk, and Y. Massoud, “Theory and implementation of an analog-to-information converter using random demodulation”, Proc. IEEE Int. Symp. on Circuits and Systems, New Orleans, Louisiana, 2007.

[50] T. Ragheb, S. Kirolos, J. Laska, A. Gilbert, M. Strauss, R Baraniuk, and Y. Massoud, “Implementation models for analog-to-information conversion via random sampling”, 50th IEEE Int’l Midwest Symposium on Circuits and Systems/ 5th IEEE Int’l Northeast Workshop on Circuits and Systems (MWSCAS/NEWCAS), August 2007.

[51] M. Sheikh, O. Milenkovic, and R. Baraniuk, “Compressed sensing DNA microarrays”, Rice ECE Department Technical Report TREE 0706, May 2007.

[52] J. L. Starch, M. Elad, D. Donoho, “Image decomposition via the combination of sparse representations and a variational approach”, IEEE trans. on Image Processing, vol. 14, no 10, pp. 1570-1582, 2005.

[53] M. Elad, A. M. Bruckstein, “A generalized uncertainty principle and sparse representation in pairs of bases”, IEEE trans. on Information Theory, vol. 48, pp. 2558-2567, 2002.

[54] S. Mallat, “A wavelet tour of signal processing”, Academic press, 1999.

[55] X. Zhang, X. Wu, F. Wu, “Image Coding on Quincunx Lattice with Adaptive Lifting and Interpolation”, Data Compression Conference, pp. 193-202, 2007.

[56] L. M. Bregman, “The method of successive projection for finding a common point of convex sets,” Soviet Math. Dokl., vol. 6, pp. 688–92, 1965.