Universal Distributed Sensing via Random Projections

Marco F. Duarte, Michael B. Wakin, Dror Baron, and Richard G. Baraniuk
Electrical and Computer Engineering
Rice University
Houston, TX 77005

ABSTRACT

This paper develops a new framework for distributed coding and compression in sensor networks based on *distributed compressed sensing* (DCS). DCS exploits both intra-signal and inter-signal correlations through the concept of joint sparsity; just a few measurements of a jointly sparse signal ensemble contain enough information for reconstruction. DCS is well-suited for sensor network applications, thanks to its simplicity, universality, computational asymmetry, tolerance to quantization and noise, robustness to measurement loss, and scalability. It also requires absolutely no inter-sensor collaboration. We apply our framework to several real world datasets to validate the framework.

Categories and Subject Descriptors: E.4 [**Coding and Information Theory**]: Data compaction and compression; J.2 [**Physical Sciences and Engineering**]: Engineering

General Terms: Algorithms, Design, Experimentation, Measurement, Performance, Security, Theory.

Keywords: Sparsity, compressed sensing, linear programming, greedy algorithms, correlation, sensor networks.

1. INTRODUCTION

1.1 Sensor networks and data representation

Recently, a vision has emerged of networks of battery-powered wireless sensor nodes that can sense, process, and communicate information about physical processes [1]. These small, low-cost devices contain sensors for the physical phenomena of interest, an embedded CPU, a radio transceiver for communication with other nearby sensor nodes, a small battery, and so on. When deployed across an area of interest, the sensors self-organize into a network that acquires, processes, and disseminates sensor data.

To realize this sensor network vision, we must surmount a variety of technical challenges, including reducing the consumption of scarce energy and communications bandwidth resources; making the best of inexpensive hardware of limited computational power; communicating over hostile channels with high loss rates; dealing with faulty sensors and noisy measurements; keeping the data secure; etc.

Fortunately, the signals, images, and other data acquired by sensor networks often contain some type of *structure* that enables intelligent representation and processing. For instance, consider the problem of representing and compressing raw sensor network data in order to minimize communication and bandwidth consumption. Current state-of-the-art compression algorithms employ a decorrelating transform such as an exact or approximate Karhunen-Loève transform (KLT) to compact a correlated signal's energy into just a few essential coefficients. Such *transform coders* exploit the fact that many signals have a *sparse* representation in terms of some basis, meaning that a small number K of adaptively chosen transform coefficients can be transmitted or stored rather than the $N \gg K$ signal samples.

1.2 Distributed source coding

While the theory and practice of compression have been well developed for individual signals, sensor network applications involve multiple signals, for which there has been less progress. Fortunately, since the sensors presumably observe related phenomena, the ensemble of signals they acquire can be expected to possess some joint structure, or *inter-signal correlation*, in addition to the *intra-signal correlation* in each individual sensor's measurements. *Distributed source coding* algorithms that exploit both types of correlation promise a substantial savings on communication costs [2–4].

A number of distributed coding algorithms have been developed that involve collaboration amongst sensors [5, 6]. Any collaboration, however, involves inter-sensor communication overhead that can significantly affect the power consumption of the participating nodes. The *Slepian-Wolf* framework for lossless distributed coding [2] offers a collaboration-free approach in which each sensor could communicate losslessly at its conditional entropy rate rather than at its individual entropy rate. Most existing coding algorithms [3, 4] exploit only inter-signal correlations (between samples in each signal) and not intra-signal correlations (between samples across signals); furthermore, there has been only limited progress on distributed coding of so-called "sources with memory." The direct implementation for such sources would require huge lookup tables [7], and approaches combining pre- or post-processing of the data to remove intra-signal correlations combined with Slepian-Wolf coding appear to have limited applicability or require excessive computation [8].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
IPSN'06, April 19–21, 2006, Nashville, Tennessee, USA.
Copyright 2006 ACM 1-59593-334-4/06/0004 ...$5.00.

© ACM, 2006. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in Proceedings of IPSN'06, April 2006.

1.3 Compressed sensing (CS)

A new framework for single-signal sensing and compression has developed recently under the rubric of *Compressed Sensing* (CS) [9, 10]. CS builds on the surprising revelation that a signal having a sparse representation in one basis can be recovered from a small number of projections onto a second basis that is *incoherent* with the first.[1] In fact, for an N-sample signal that is K-sparse,[2] roughly only cK projections of the signal onto the incoherent basis are required to reconstruct the signal with high probability (typically $c \approx 3$ or 4). This has promising implications for applications involving sparse signal acquisition. Moreover, the cK measurements need not be manipulated in any way before being transmitted, except possibly for some quantization. Interestingly, independent and identically distributed (i.i.d.) Gaussian or Bernoulli/Rademacher (random ± 1) vectors provide a useful *universal* measurement basis that is incoherent with *any* given basis with high probability; the corresponding measurements will capture the relevant information of a compressible signal, regardless of its structure.

1.4 Distributed compressed sensing (DCS)

We have recently introduced new theory and algorithms for Distributed Compressed Sensing (DCS) that exploit both intra- *and* inter-signal correlation structures [12]. In a typical DCS scenario, a number of sensors measure signals (of any dimension) that are each individually sparse in some basis and also correlated among sensors. Each sensor *independently* encodes its signal by projecting it onto just a few vectors of a second, incoherent basis (such as a random one) and then transmits the resulting coefficients to a collection point. Under the right conditions, a decoder at the collection point can *jointly* reconstruct all of the signals precisely.

The DCS theory rests on a concept that we term the *joint sparsity* of a signal ensemble. Below we discuss several models for jointly sparse signals (corresponding to a variety of scenarios of signal correlation) and joint reconstruction algorithms for each model. We have derived results on the required measurement rates for signals that have sparse representations under each of the models; while the sensors operate entirely *without collaboration*, we see dramatic savings relative to the number measurements required for separate CS decoding [12].

1.5 From DCS to universal distributed sensing

In this paper we demonstrate the potential of DCS for universal distributed sensing in networks. We develop algorithms for several network signal processing and compression tasks using random measurements and validate them on real sensor network data. The properties of DCS directly address the sensor network challenges outlined above. In particular, DCS algorithms: offer a universal encoding appropriate for any jointly sparse signal ensemble; are completely non-collaborative and involve no communication overhead; can be implemented on the simplest computing hardware on the sensor nodes since they shift nearly all computational complexity to the decoder at the collection point; are inherently fault tolerant, robust to measurement and quantization noise, and secure; are robust to lossy communication links; offer progressively better, tunable reconstruction as measurements stream in; and are applicable to a range of sensor network signal processing tasks, from signal compression to estimation and detection/classification. To coin a term, DCS sensors are "omnescient": they omnisciently capture the relevant signal information despite being nescient (ignorant) of the actual structure.

This paper is organized as follows: Section 2 presents related work. Section 3 provides background on CS. Section 4 outlines three of our models for joint sparsity. Section 5 describes the basic implementation of DCS in wireless sensor networks, and Section 6 highlights the advantages of DCS for sensor networks. Section 7 presents experiments showcasing performance of DCS on real world sensor networks. Finally, Section 8 concludes with directions for future work.

2. RELATED WORK

Several approaches have been proposed for data collection in sensor networks, most of which exploit the correlation among the signals being recorded. DIMENSIONS [13] enables distributed information storage and multiresolution data retrieval; it achieves compression by assuming that the signal at each sensor node features temporal correlation and clustering sensors that observe correlated signals in a hierarchical fashion. The compression of signal ensembles thus requires high computation during clustering, and so the cluster heads must be capable of performing such tasks within their power and computational budgets. Fractional cascading [14] allows queries to be injected at any point in the network. Information is redundantly stored at several sensors, requiring again collaboration and computation to integrate measurements from local groups of sensors.

Other algorithms that exploit correlations in sensor networks include signal compression [15], routing [16], and signal processing tasks [17–19]. The general approach consists of clustering nodes that observe correlated signals and then performing local processing, routing, or compression at a node chosen as a cluster head; the process continues iteratively until a single cluster is obtained. Unfortunately, clustering techniques require collaboration amongst sensors, which increases power consumption for the nodes due to message passing inside clusters. Furthermore, not all sensor network architectures can support the computational complexity of the signal processing algorithms.

Algorithms for non-collaborative signal compression have been proposed [3, 20], but they either have high complexity or do not exploit inter-sensor correlations. One technique [21] exploits CS for joint measurement of a spatial sensor field at a single time instant. This approach uses matched source-channel communication [22] to significantly reduce the required power. However, it neglects intra-sensor correlations – those between the samples of each signal – and it requires both the deployment of sensors on a regular grid and a potentially complicated time and power synchronization of wireless transmitters among the nodes.

In contrast to these approaches, our proposed framework involves no collaboration among the sensors, has low computational complexity, and facilitates easy measurement aggregation. Section 6 elaborates on these and other benefits.

[1] Roughly speaking, *incoherence* means that no element of one basis has a sparse representation in terms of the other basis [9–11].

[2] By K-sparse, we mean that the signal can be written as a sum of K basis functions.

3. COMPRESSED SENSING

We briefly explain the Compressed Sensing (CS) framework proposed in [9, 10]. Suppose that $x \in \mathbb{R}^N$ is a signal, and let $\Psi = \{\psi_1, \psi_2, \ldots, \Psi_N\}$ be a basis of vectors spanning $\mathbb{R}^N$; the theory can be extended to frames and dictionaries of vectors. When we say that x is *sparse*, we mean that x is well approximated by a linear combination of a small set of vectors from Ψ. That is, there exists a set of indices $\{n_1, \ldots, n_K\} \subset \{1, \ldots, N\}$ for small $K \ll N$ such that $x \approx \sum_{i=1}^{K} \theta_{n_i} \psi_{n_i}$; we say that x is *K-sparse* in Ψ, and we call Ψ the *sparse basis*. The CS theory states that it is possible to construct an $M \times N$ *measurement* matrix Φ where $M \ll N$, yet the measurements $y = \Phi x$ preserve the essential information about the K-sparse signal x. For example, let Φ be a $cK \times N$ random matrix with i.i.d. Gaussian entries, where $c = c(N, K)$ is an *overmeasuring factor*. Using such a matrix it is possible, with high probability, to recover any signal that is K-sparse in the basis Ψ from its image under Φ. Moreover, for signals that are not K-sparse but *compressible*, meaning that their coefficient magnitudes decay rapidly, there are tractable algorithms that achieve not more than a multiple of the error of the best K-term approximation of the signal. Most natural and man-made signals are compressible in some basis, including audio, images, and video signals.

Several algorithms have been proposed for recovering x from the measurements y, each requiring a slightly different constant c. The canonical approach [9, 10] uses linear programming to solve the ℓ_1 minimization problem

$$\widehat{\theta} = \arg\min_{\theta} \|\theta\|_1 \quad \text{subject to} \quad \Phi\Psi\theta = y$$

and then sets $\widehat{x} = \Psi\widehat{\theta}$. This problem requires $c \approx \log_2(1 + N/K)$ [12] but has somewhat high computational complexity. Greedy pursuit methods have also been proposed for CS reconstruction, including Orthogonal Matching Pursuit (OMP), which tend to require fewer computations but at the expense of slightly more measurements [11].

4. JOINT SPARSITY MODELS

In this section, we generalize the notion of a signal being sparse in some basis to the notion of an ensemble of signals being *jointly sparse*. We consider three different Joint Sparsity Models (JSMs) that are inspired by different real world situations. In the first two models, each signal is itself sparse, and so we could use the CS framework from above to encode and decode each one separately, yet there also exists a framework wherein a *joint representation* for the ensemble uses fewer total vectors. In the third model, no signal is itself sparse, yet there still exists a joint sparsity among the signals that allows recovery with significantly fewer measurements per sensor. We note that for different real world settings, different models for sparsity-based signal ensemble structure can be posed, together with appropriate reconstruction algorithms.

We use the following notation for signal ensembles. Denote the *signals* in the ensemble by x_j, $j = 1, 2, \ldots, J$ where each $x_j \in \mathbb{R}^N$. We assume that there exists a known *sparse basis* Ψ for $\mathbb{R}^N$ in which x_j can be sparsely represented.

4.1 JSM-1: Sparse common component + innovations

In this model, all signals share a *common* sparse component while each individual signal contains a sparse *innovations* component:

$$x_j = z + z_j, \quad j \in \{1, 2, \ldots, J\}$$

with

$$\begin{aligned} z &= \Psi\theta_z, \quad \|\theta_z\|_0 = K, \\ z_j &= \Psi\theta_j, \quad \|\theta_j\|_0 = K_j. \end{aligned}$$

Thus, the signal z is common to all of the x_j and has sparsity K in basis Ψ.[3] The signals z_j are the unique portions of the x_j and have sparsity K_j in the same basis.

A practical situation well-modeled by JSM-1 is a group of sensors measuring temperatures at a number of locations throughout the day. The temperature readings x_j have both temporal (intra-signal) and spatial (inter-signal) correlations. Global factors, such as the sun and prevailing winds, could have an effect z that is both common to all sensors and structured enough to permit sparse representation. More local factors, such as shade, water, or animals, could contribute localized innovations z_j that are also structured (and hence sparse). A similar scenario could be imagined for a sensor network recording light intensities, air pressure, or other phenomena. All of these scenarios correspond to measuring properties of physical processes that change smoothly in time and in space and thus are highly correlated.

4.2 JSM-2: Common sparse supports

In this model, all signals are constructed from the same sparse index set of basis vectors, but with different coefficients:

$$x_j = \Psi\theta_j, \quad j \in \{1, 2, \ldots, J\},$$

where each θ_j is supported only on the same $\Omega \subset \{1, 2, \ldots, N\}$ with $|\Omega| = K$. Hence, all signals are K-sparse, and all are constructed from the same K elements of Ψ, but with arbitrarily different coefficients. This model can be viewed as a special case of JSM-1 (with $K = 0$ and $K_j = K$ for all j) but features additional correlation structure that suggests distinct reconstruction algorithms.

A practical situation well-modeled by JSM-2 is where multiple sensors acquire the same Fourier-sparse signal but with phase shifts and attenuations caused by signal propagation. In many cases it is critical to recover each one of the sensed signals, such as in many acoustic localization and array processing algorithms. Another application for JSM-2 is MIMO communication [23]. Section 7 presents a series of experiments applying JSM-2 to environmental and acoustic data.

4.3 JSM-3: Nonsparse common + sparse innovations

This model extends JSM-1 so that the common component need no longer be sparse in any basis; that is,

$$x_j = z + z_j, \quad j \in \{1, 2, \ldots, J\}$$

with

$$z = \Psi\theta_z \qquad \text{and} \qquad z_j = \Psi\theta_j, \quad \|\theta_j\|_0 = K_j,$$

but z is not necessarily sparse in the basis Ψ. We also consider the case where the supports of the innovations are shared for all signals, which extends JSM-2.

A practical situation well-modeled by JSM-3 is where several sources are recorded by different sensors together with a background signal that is not sparse in any basis. Consider, for example, a computer vision-based verification system in

[3] The ℓ_0 norm $\|\theta\|_0$ merely counts the number of nonzero entries in the vector θ.

a device production plant. Cameras acquire snapshots of components in the production line; a computer system then checks for failures in the devices for quality control purposes. While each image could be extremely complicated, the ensemble of images will be highly correlated, since each camera observes the same device with minor (sparse) variations.

5. DISTRIBUTED SENSING USING RANDOM PROJECTIONS

In this section we describe the mechanics of implementing DCS in a sensor network environment. In the next section, we highlight the unique benefits afforded by such an approach.

5.1 Incoherent measurements

We consider a collection of J synchronized sensor nodes that observe signals obeying one of the JSMs or their extensions (as described in Sections 4 and 6). Each sensor *independently* collects a set of incoherent measurements and transmits them to a data sink. The signals are then recovered *jointly* using algorithms discussed in Section 5.3. We emphasize that, thanks to the *universal* nature of random measurements, the sensors need not be informed of the sparsity-inducing basis for the signals; this information is only required to perform reconstruction at the decoder.

We assume that sensor j acquires the N-sample signal x_j observed during a time interval $[t_0, t_0 + T]$ and computes a given number of measurements M_j. (The period $[t_0, t_0 + T]$ could be the complete duration of the signal of interest or could correspond to a length-N block of a longer signal; the above process can be repeated periodically.) We denote the measurement vector by $y_j = \Phi_j x_j$, where Φ_j is the *measurement matrix* for sensor j; Φ_j is $M_j \times N$ and, in general, the entries of Φ_j are different for each j. We denote the vector of measurements as $y_j = [y_{j,1}, \ldots, y_{j,M_j}]^T$ and note that $y_{j,m}$ corresponds to the inner product of x_j with row m of the matrix Φ_j. Since all measurements have the same relevance for signal reconstruction, their values are quantized using the same scheme for each index m; the distortion in the reconstruction due to quantization is bounded [24].

The CS and DCS frameworks require knowledge during reconstruction of the measurement matrix Φ_j for the different sensors $j = 1, \ldots, J$. This can be accomplished by constructing each measurement matrix using a pseudorandom number generator, whose seed could be provided by the data sink or computed as a function of the node ID. While most of the existing theory for CS encoding applies specifically to random Gaussian or Bernoulli measurements, there is active research into developing lower-complexity alternatives [25, 26]. We have strong experimental evidence that structured measurement matrices Φ_j (involving, for example, an FIR filter with pseudorandom taps [25]) can provide suitable incoherence with the sparse basis Ψ.

5.2 Communication to the data sink

Each quantized measurement $\widehat{y}_{j,m}$ is transmitted to the sink together with its timestamp t_0, index m, and node ID j. This is the only information necessary from the sensors to reconstruct the signals. Since the measurements can arrive out of order, they can be sent individually over the network or grouped into packets if desired. Many different options exist for routing the measurements, including TreeCast [27] and DIMENSIONS [13].

5.3 Joint reconstruction

As the measurements are received by the data sink, the measurement matrices Φ_j for the different sensors are built accordingly through the same procedure as in the sensors. Once the data sink receives all M_j measurements from each sensor — or alternatively, once it starts receiving measurements for the next measurement period (beginning at t_0+T) — the data sink can begin reconstructing the signal ensemble.

The algorithm used for joint signal reconstruction depends on the relevant JSM for the signals observed by the network. In this section we briefly overview our proposed reconstruction techniques for each JSM; more details on the algorithms (and the theoretical requirements on the measurement rates M_j) can be found in [12].

For JSM-1, we have developed an analytical framework inspired by principles of information theory. This allows us to characterize the measurement rates M_j required to *jointly* reconstruct the signals x_j. The measurement rates relate directly to the signals' *conditional sparsities*, in parallel with the Slepian-Wolf theory. The reconstruction technique is based on a single execution of a linear program that seeks the sparsest components $[z;\ z_1;\ \ldots\ z_J]$ that account for the observed measurements. Theoretical analysis and numerical experiments confirm that the rates M_j required for joint CS recovery are well below those required for independent CS recovery of each signal x_j [12].

For JSM-2, we have developed algorithms inspired by conventional greedy pursuit algorithms (such as OMP [11]) that can substantially reduce the number of measurements when compared with independent recovery. In the single-signal case, OMP iteratively constructs the sparse support set Ω; decisions are based on inner products between the columns of $\Phi\Psi$ and a residual. In the multi-signal case, there are more clues available for determining the elements of Ω. We use a simple variant of Simultaneous Orthogonal Matching Pursuit (SOMP) [12, 23]. We have proved that for large J, close to K measurements per signal suffice for joint reconstruction (that is, $c \to 1$ as $J \to \infty$); see Figure 1 for an example of improving performance as J increases. On the contrary, with independent CS reconstruction, perfect reconstruction of all signals requires *increasing* each M_j in order to maintain the same probability of reconstruction of the signal ensemble. This surprise is due to the fact that each signal will experience an independent probability $p \leq 1$ of successful reconstruction; therefore the overall probability of complete success is p^J. Consequently, each sensor must compensate by making additional measurements. We also note that when the supports of the innovations of the signals are small, signals that are well modeled by JSM-1 can also be modeled by JSM-2 by selecting a global support that contains all of the individual supports. Such approximation allows for a simpler reconstruction algorithm, while incurring a slight increase in the number of measurements required for reconstruction.

For JSM-3, no individual signal x_j is sparse, and so recovery of each signal separately would require a full N measurements per signal. To approach the reconstruction problem, we note that the common component z is observed by all sensors. Thus we propose an Alternating Common and Innovation Estimation (ACIE) [12] that alternates between two steps: (1) Estimate the common component z by combining all measurements and treating the innovations

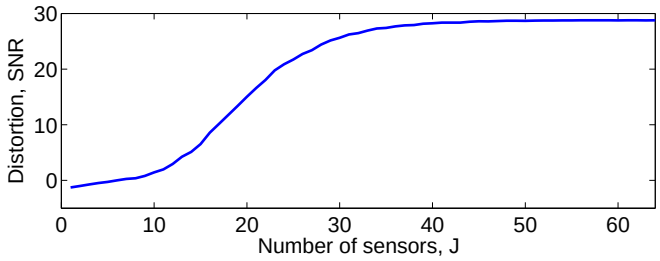

Figure 1: *Joint reconstruction of synthetic JSM-2 signals having length $N = 128$ and sparsity $K = 10$ using $M = 11$ random measurements per sensor. Each measurement is quantized to approximately 5 bits of precision. The reconstruction is robust to quantization and is progressive: as the number of sensors J increases we see improved reconstruction performance.*

z_j as noise that can be averaged out; (2) Estimate the innovations z_j from each sensor by subtracting the estimated common component z and then applying standard CS recovery techniques. We have proved that, asymptotically, each sensor need only measure at the rate dictated by the sparsity K_j [12]. Thus, for large J, the impact of the common nonsparse component z is eliminated.

6. ADVANTAGES OF DCS FOR SENSOR NETWORKS

Our DCS implementation for sensor networks is robust and widely applicable in sensor network scenarios. This section describes in more detail several of the desirable features.

Simple, universal encoding: DCS coding is particularly appealing when we employ random projections at the sensors. Random projections are *universal* in the sense that they are incoherent with any fixed sparsity basis Ψ [9]. In fact, using the same set of random measurements the decoder can attempt to recover the signals using any supposed sparse basis Ψ or JSM. In addition to being universally incoherent, the CS/DCS random measurements are also *future-proof*: if a better sparsity-inducing basis is found (or a better JSM is proposed), then the same random measurements can be used to reconstruct an even more accurate view of the environment without requiring any changes in the deployed sensing hardware. Additionally, DCS can be applied to any number of sensors $J \geq 2$, and furthermore the sensors need not know their physical locations (other than to network their data).

The CS/DCS frameworks, in which measurements can be obtained with low complexity and without collaboration, also shifts the computational load of reconstruction from the sensor network to the data sink or cluster head. Each sensor only needs to compute its incoherent projections of the signal it observes, while the data sink or cluster head reconstructs all of the signals. This *computational asymmetry* is desirable in many sensor networks since data sinks and cluster heads have typically more computational power than sensor nodes.

Robustness, progressivity, and resiliency: DCS enjoys remarkable *robustness* properties thanks to the robustness of the CS framework. CS measurements have been shown to be robust to quantization and noise [21, 24], making the framework applicable to real world settings. Additionally, the incoherent measurements coming from each sensor have equal priority, unlike transform coefficients in current coders. Thus, the CS measurements can be transmitted and received in any order. Signal reconstruction can be attempted using any number of the received measurements — as more measurements are received they allow a *progressively better reconstruction* of the data [9].

In this sense, DCS is automatically *robust to packet loss* in wireless sensor networks; any loss of measurements leads to a graceful degradation in the reconstruction quality. This *loss resiliency* is particularly useful, as errors in wireless sensor network transmissions often cause as many as $10 - 30\%$ of the packets to be dropped [28]. This effect is exacerbated in multi-hop networks.

One existing approach that is robust to packet drops is multiple description coding [29,30]. These techniques enable data reconstruction at varying levels of quality depending on the number of packets that arrive. Unfortunately, multiple description coding techniques for distributed source coding have not been fully developed [31]. Another approach uses layered coding for unequal bit error protection, where the first layer is highly protected with strong channel coding and is also used as side information when decoding the second layer [32]. This layered approach also increases demands on the system resources because the stronger channel code requires substantial redundancy in terms of channel resources and power consumption.

Security: Using a pseudorandom basis (with a random seed) effectively implements *encryption*: the randomized measurements will themselves resemble noise and be meaningless to an observer who does not know the seed.

Fault tolerance and anomaly detection: DCS reconstruction techniques can be extended to be *fault tolerant*. In the case where a small number of signals may not obey the overall JSM (due to a faulty sensor, for example), the joint reconstruction techniques can be tailored to detect such anomalies. In the case of JSM-2, for example, after running SOMP to determine the common support set Ω, the data sink could examine each sensor's measurements to check for agreement with Ω. Those signals that appear to disagree can then be reconstructed separately from the remaining (JSM-faithful) nodes.

Adaptivity to channel capacity: The DCS measurement and transmission rates can be scaled to adapt to the conditions of the wireless communication channel and the nuances of the observed signals. If, for example, the communication channel capacity is below the required rate to send M_j measurements, then the sensors can perform rate limitation in a similar manner to congestion control algorithms for communication networks. When the data sink detects congestion in the communication channel, it can send a congestion notification (using a trickle of feedback) to the nodes so that the bit rate of the information sent is reduced in one of two ways. First, the sensors could increase the quantization stepsize of the measurements, since the CS/DCS reconstruction is robust to quantization. Second, the sensors could reduce the number of measurements taken for each signal: due to the resiliency of CS measurements, the effect of having few measurements on the reconstruction distortion is gradual. Thus, the CS/DCS measurement process can easily *scale* to match the transmission capacity of the communication channel, which is reminiscent of joint source-channel coding.

Application to compressible and analog signals: DCS can also be applied to signals that do not strictly

obey the JSMs outlined in Section 4. For example, our JSM models can be generalized to *compressible signals*, for which the transform coefficients decay quickly but strictly to zero [9, 10]. The basic requirement is that certain small sets of coefficients approximate the signals with low distortion and that these sets are related among the different sensors. Section 7 demonstrates numerical experiments on actual data sets that only approximately obey our JSMs.

The measurement rates required to capture compressible signals depend on the level of compressibility of the signals, that is, on the decay rate of the signal coefficient magnitudes in the sparse representation. Fortunately, as mentioned above, DCS enables adaptation of the measurement rate for a given signal. Thus, the measurements can be increased to obtain more accurate reconstructions and scaled back to obtain coarser approximations. Several algorithms for CS reconstruction [11, 33] also provide mechanisms to verify the success of the reconstruction, and so the data sink can be aware of when it has received enough measurements.

Finally, DCS can be extended to the acquisition of *analog signals* directly at the physical sensor (bypassing the analog-to-digital converter). We have introduced one technique for CS analog sensing using a camera that directly acquires random projections of the input light field [26]; other techniques are currently under investigation. Transmission for the resulting (digital) measurements would then proceed as in Section 5.2, and reconstruction could be tailored to the measurement structure given by the analog front end, perhaps allowing for faster and/or simpler reconstruction.

Information scalability: Incoherent measurements obtained via DCS can be used to recover *different levels of information* about the sensed signals. It has been shown [34] that the CS framework is *information scalable* beyond signal reconstruction to a much wider range of statistical inference tasks, including estimation, detection, and classification. Depending on the situation, the lower levels of information about the signals can often be extracted using lower computational complexity or fewer incoherent measurements than would be required to reconstruct the signals. For example, statistical detection and classification do not require reconstruction of the signal, but only require an estimate of the relevant *sufficient statistics*. Consequently, it is possible to directly extract such statistics from a small number of random projections without ever reconstructing the signal. As a result, significantly fewer measurements are required for signal detection than for signal reconstruction [34]. Furthermore, as in reconstruction, random measurements are again *universal*, in the sense that with high probability the sufficient statistics can be extracted from them regardless of the signal structure.

As a first example, we consider sensor networks for surveillance applications [17]. Typically, a detection algorithm is executed continuously on the sensed data; when the algorithm returns an event detection, other algorithms such as classification, localization, and tracking are executed. These algorithms require a larger amount of information from the signals than that of detection. In our DCS scheme, we can adapt the measurement rate of the sensor nodes according to the tasks being performed. We apply a low measurement rate for detection; once the detection returns an event, the measurement rate is increased to that required by the other tasks.

As another example, one may be interested in estimating

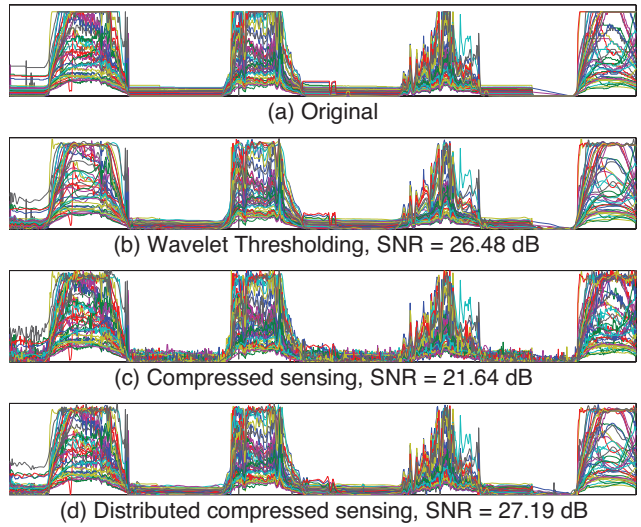

Figure 2: *Reconstruction of light intensity signals from 48 sensors with length $N = 1024$. (a) Original signals; (b) wavelet thresholding using 100 coefficients per sensor, average SNR = 26.48dB; (c) separate reconstruction of each signal using CS from $M = 400$ random projections per sensor, average SNR = 21.64dB; (c) joint reconstruction of the signal ensemble using DCS from $M = 400$ random projections per sensor, average SNR = 27.19dB.*

linear functions of the sensed signals

$$v = \sum_j \omega_j x_j;$$

examples include averages and linear interpolations. Thanks to the linearity of the CS/DCS measurement process, we can extract such information from the incoherent measurements without first reconstructing the signals x_j. More specifically, assuming we use the same measurement process $\Phi_j = \Phi$ at each sensor, we can write

$$\Phi v = \sum_j \omega_j \Phi x_j = \sum_j \omega_j y_j.$$

Assuming that v is sparse, it can be recovered from Φv using standard CS techniques. Thus, by aggregating the measurements y_j using the desired linear function we can *directly* obtain incoherent measurements of v without reconstructing the x_j. We also note that the measurement vectors can be aggregated using matched source-channel communication [22], in which the wireless nodes collaborate to coherently send their measurements so that a receiver directly obtains the weighted sum $\omega_j y_j$. This could enable a significant reduction in power. Such aggregation can also be implemented hierarchically in frameworks such as TreeCast [27] or DIMENSIONS [13].

7. EXPERIMENTS

In this section, we consider four different sensor network datasets. Although the signals we consider are not strictly sparse, we see that the JSM models provide a good approximation for the joint sparsity structure and that DCS offers a promising approach for such sensing environments.

Environmental sensing: The first three datasets [35] contain temperature, humidity, and light readings from a group of 48 nodes deployed at the offices of Intel Research Labs in Berkeley, CA.[4] The signals in Figures 2(a), 3(a) and

[4]For the purposes of our experiments, we select signals of

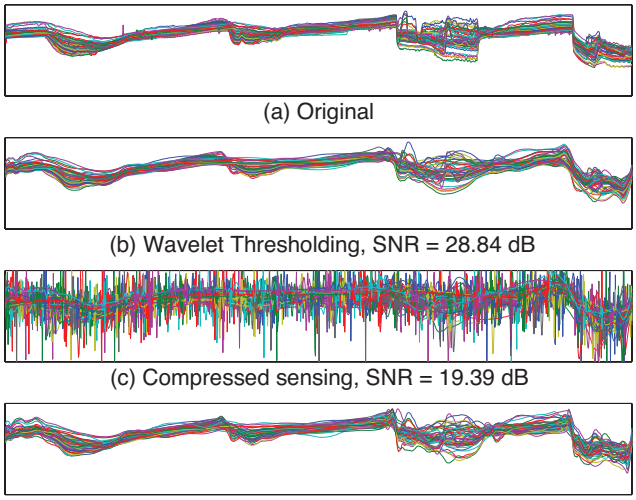

Figure 3: *Reconstruction of humidity signals from 48 sensors with length $N = 1024$. (a) Original signals; (b) wavelet thresholding using 20 coefficients per sensor, average SNR = 28.84dB; (c) separate reconstruction if each signal using CS from $M = 80$ random projections per sensor, average SNR = 19.39dB; (d) joint reconstruction of the signal ensemble using DCS from $M = 80$ random projections per sensor, average SNR = 29.66dB.*

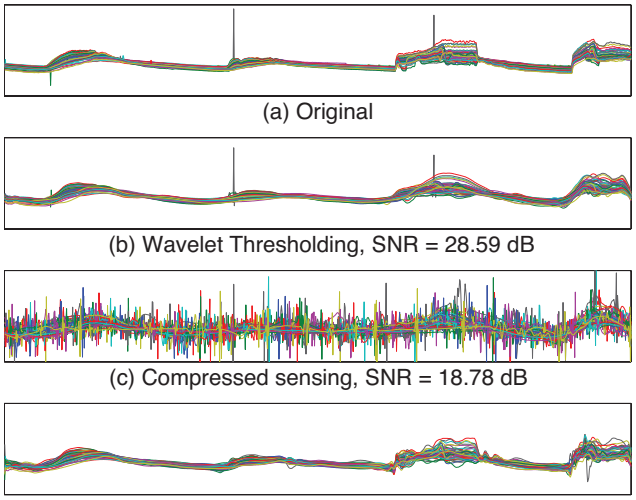

Figure 4: *Reconstruction of temperature signals from 48 sensors with length $N = 1024$. (a) Original signals; (b) wavelet thresholding using 20 coefficients per sensor, average SNR = 28.59dB; (c) separate reconstruction of each signal using CS from $M = 80$ random projections per sensor, average SNR = 18.78dB; (c) joint reconstruction of the signal ensemble using DCS from $M = 80$ random projections per sensor, average SNR = 29.95dB.*

4(a) were recorded in an office environment and therefore exhibit periodic behavior caused by the activity levels during day and night. Furthermore, there are small fluctuations at each one of these states; thus we expect the signals to be compressible both in the Fourier and wavelet domains. Since the signals are observations of physical processes, they are smoothly varying in time and space; this causes the sensor readings to be close in value to each other, a situation well captured by the JSM-1 and JSM-2 models.

We now confirm the joint sparsity of the signals under the JSM-2 model. The top panel in Figure 5 shows the distortion of the best K-term wavelet approximation[5] for each signal in the light dataset as K increases. The figure shows that a modest value of $K = 100$ gives low distortion for all signals. However, the union over all signals of the K best wavelet vectors has size greater than K. The bottom panel in Figure 5 shows the size of this union (the "joint support" for the signals under JSM-2) as K increases. We see that approximately $|\Omega| = 200$ vectors are required to include the $K = 100$ most significant vectors for each signal, which makes the JSM-2 model feasible due to the shared compactness of the representation. Similar results are observed for the other datasets, which are compressible in the wavelet domain as well. Thus, we expect that such datasets can be recovered from incoherent projections using DCS with the appropriate sparsity inducing bases.

We now consider a hypothetical implementation of DCS for these signals. For the light intensity signal we take $M = 400$ random Gaussian measurements per sensor and compare DCS reconstruction (via SOMP using wavelets as the sparsity basis) with separable OMP reconstruction. For comparison, we also compare to wavelet thresholding at each signal using 100 terms. Figure 2 shows the reconstruction of the light intensity signal ensemble. We see average SNRs of 26.48dB, 21.64dB, and 27.19dB for wavelet thresholding, separate CS, and DCS reconstruction, respectively. The DCS reconstruction algorithm identifies the common structure emphasized by JSM-2, recovering salient common features for all signals in the ensemble in addition to many of the distinct features in each signal. Similar results are seen for the humidity and temperature datasets in Figures 3, 4, and 6. To illustrate progressivity, Figure 7 also plots the CS (OMP) and DCS (SOMP) reconstruction errors for the temperature signal ensemble at a variety of measurement rates M. SOMP reconstruction is superior at low and moderate rates, yet it is surpassed by OMP at high rates. This illustrates the applicability of the JSM-2 model, which becomes less valid as the very fine features of each signal (which vary between sensors) are incorporated. A joint reconstruction algorithm tailored to this fact would likely outperform both approaches.

Acoustic sensing: Our fourth dataset [18] contains audio recordings of military vehicles from a 16-microphone sensor network array from the University of Wisconsin-Madison. The audio signals are compressible in the Fourier domain and follow the JSM-2 model (see Figure 8). Figure 9 shows an example DCS reconstruction (using SOMP with the Fourier sparse basis); the results are similar to those seen in the previous datasets.

8. CONCLUSIONS AND FURTHER WORK

In this paper we have demonstrated the potential of DCS for universal distributed sensing in networks. Specifically, our DCS framework provides such attractive properties as universality, resiliency, scalability, adaptability, and reduced computational complexity at the sensors. DCS enables *omnescient* sensor networks, in the sense that the nodes acquire accurate information about the signal regardless of its specific structure; the only requirement is that the signals be

length $N = 1024$ and interpolate small amounts of missing data.

[5]We use Daubechies-8 wavelets throughout this section.

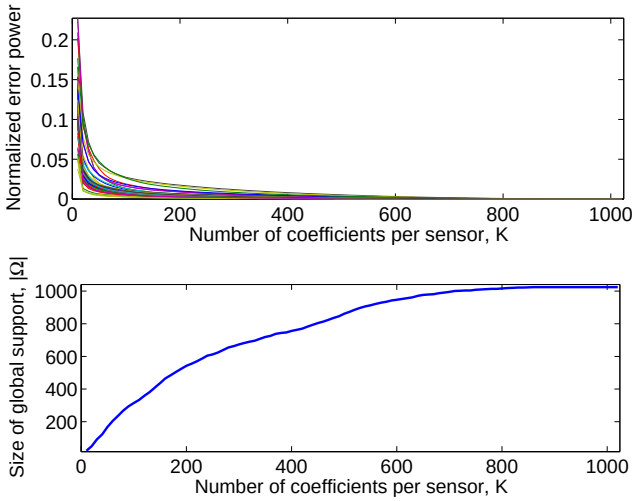

Figure 5: *Top: Quality of approximation of light intensity signals as a function of the number K of wavelet coefficients used per sensor. When $K \geq 100$, the approximations yield low distortion; thus the signals are* compressible. *Bottom: Number of wavelet vectors required to include the K largest wavelet coefficients for each signal. The slope of the curve is small, meaning that the supports of the compressible signals overlap, and that the ensemble is well represented by the JSM-2 model.*

sparse or compressible in some representation, which holds true for many signals of interest.

In ongoing work, we are studying additional joint sparsity models that will allow for the DCS framework to be applied to a broader class of sensor network settings, as well as the effect of quantization in distributed sensing settings. We are also developing fast acquisition and reconstruction algorithms for signals that are jointly sparse. Finally, we are considering new sensing devices and sensor network architectures that apply the CS/DCS principles at the hardware level.

Acknowledgements

Thanks to Shriram Sarvotham for helping develop the DCS theory, to Rob Nowak for helpful comments, and to Farinaz Koushanfar for providing the experimental data. This work was supported by NSF, NSF-NeTS, ONR, and AFOSR.

9. REFERENCES

[1] D. Estrin, D. Culler, K. Pister, and G. Sukhatme, "Connecting the physical world with pervasive networks," *IEEE Pervasive Computing*, vol. 1, no. 1, pp. 59–69, 2002.

[2] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, pp. 471–480, July 1973.

[3] S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," *IEEE Trans. Inform. Theory*, vol. 49, pp. 626–643, Mar. 2003.

[4] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Mag*, vol. 21, pp. 80–94, Sept. 2004.

[5] H. Luo and G. Pottie, "Routing explicit side information for data compression in wireless sensor networks," in *Int. Conf. Distirbuted Computing in Sensor Systems (DCOSS)*, Marina Del Rey, CA, June 2005.

[6] M. Gastpar, P. L. Dragotti, and M. Vetterli, "The distributed Karhunen-Loeve transform," *IEEE Trans. Inform. Theory*, Nov. 2004, Submitted.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley and Sons, New York, NY, 1991.

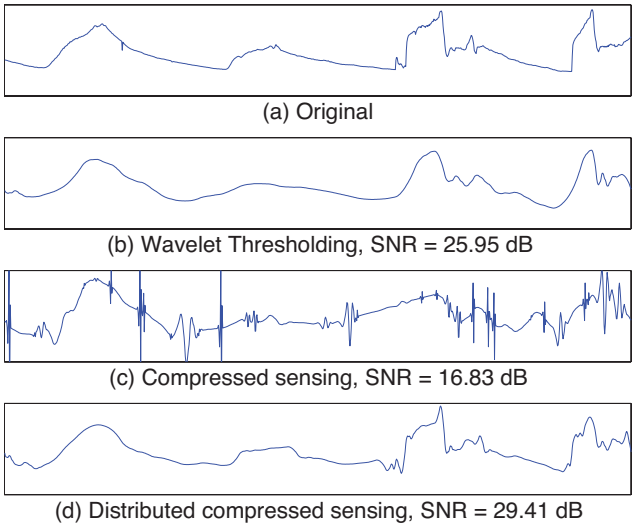

Figure 6: *Reconstruction of temperature signal #41 (extracted from Figure 4). (a) Original signal; (b) separate reconstruction of each signal using wavelet tresholding, SNR = 25.95dB; (c) reconstruction using CS, SNR = 16.83dB; (d) joint reconstruction of the signal ensemble using DCS, SNR = 29.41dB.*

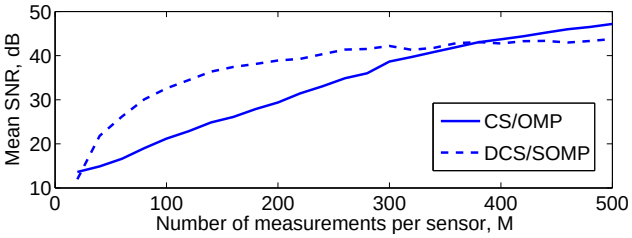

Figure 7: *Average SNR of temperature signals reconstructed from M measurements per sensor using CS (OMP) and DCS (SOMP).*

[8] T. Uyematsu, "Universal coding for correlated sources with memory," in *Canadian Workshop Inform. Theory*, Vancouver, Canada, June 2001.

[9] E. Candès and T. Tao, "Near optimal signal recovery from random projections and universal encoding strategies," *IEEE Trans. Inform. Theory*, 2004, Submitted.

[10] D. Donoho, "Compressed sensing," 2004, Submitted.

[11] J. Tropp and A. C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," Apr. 2005, Submitted.

[12] D. Baron, M. B. Wakin, M. F. Duarte, S. Sarvotham, and R. G. Baraniuk, "Distributed compressed sensing," 2005, Available at dsp.rice.edu/cs.

[13] D. Ganesan, D. Estrin, and J. Heidemann, "DIMENSIONS: Why do we need a new data handling architecture for sensor networks?," in *Proc. ACM Workshop on Hot Topics in Networks*, Princeton, NJ, Oct. 2002, pp. 143–148, ACM.

[14] J. Gao, L. J. Guibas, J. Hershberger, and L. Zhang, "Fractionally cascaded information in a sensor network," in *Proc. Information Processing in Sensor Networks (IPSN)*, New York, NY, USA, 2004, pp. 311–319, ACM Press.

[15] A. Kashyap, L. A. Lastras-Montano, C. Xia, and Z. Liu, "Distributed source coding in dense sensor networks," in *Proc. Data Compression Conference (DCC)*, Snowbird, UT, Mar. 2005.

[16] S. Pattem, B. Krishnamachari, and R. Govindan, "The impact of spatial correlation on routing with compression in wireless sensor networks," in *Proc. Information Processing in Sensor Networks (IPSN)*, Berkeley, CA, Apr. 2004.

[17] D. Li, K. D. Wong, Y. H. Hu, and A. M. Sayeed, "Detection, classification, and tracking of targets," *IEEE Signal Processing Mag*, vol. 19, no. 2, pp. 17–29, 2002.

[18] M. F. Duarte and Y. H. Hu, "Vehicle classification in

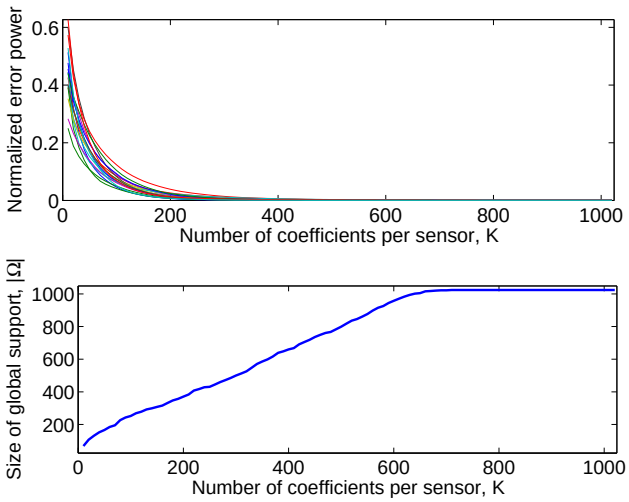

Figure 8: *Top: Quality of approximation of vehicle audio signals as a function of the number K of Fourier coefficients used per sensor. Bottom: Number of Fourier vectors required to include the K largest Fourier coefficients for each signal.*

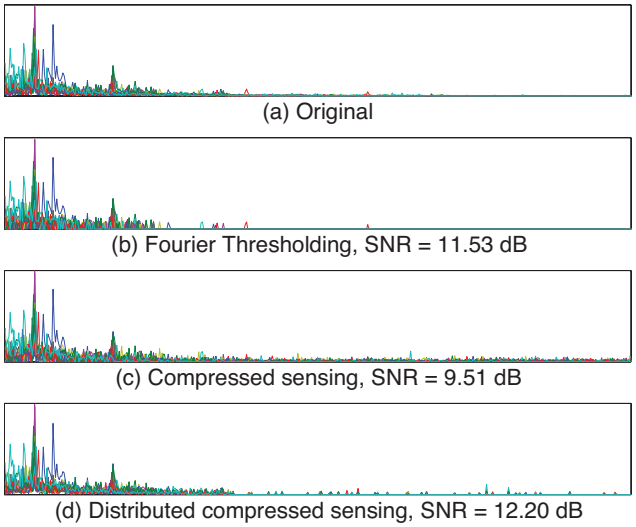

Figure 9: *Fourier coefficients for reconstructions of vehicle audio signals from 18 sensors with length $N = 1024$. (a) Original signals; (b) Fourier thresholding using 100 coefficients per sensor, average SNR $= 11.53$dB; (c) separate reconstruction using CS from $M = 400$ random projections per sensor, average SNR $= 9.51$dB; (c) joint reconstruction using DCS from $M = 400$ random projections per sensor, average SNR $= 12.20$dB.*

distributed sensor networks," *J. Parallel and Distributed Computing*, vol. 64, no. 7, pp. 826–838, July 2004.

[19] A. Sayeed, "A statistical signal modeling framework for wireless sensor networks," Tech. Rep., Univ. of Wisconsin - Madison, Feb 2004.

[20] P. Ishwar, R. Puri, K. Ramchandran, and S. S. Pradhan, "On rate-constrained distributed estimation in unreliable sensor networks," *IEEE J. Select. Areas Commun.*, vol. 23, no. 4, pp. 765–775, 2005.

[21] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Trans. Inform. Theory*, 2005, Submitted.

[22] W. Bajwa, A. Sayeed, and R. Nowak, "Matched source-channel communication for field estimation in wireless sensor networks," in *Proc. Information Processing in Sensor Networks (IPSN)*, Los Angeles, CA, Apr. 2005.

[23] J. Tropp, A. C. Gilbert, and M. J. Strauss, "Simulataneous sparse approximation via greedy pursuit," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Philadelphia, PA, Mar. 2005.

[24] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure and Applied Mathematics*, 2005, Submitted.

[25] J. A. Tropp, M. B. Wakin, M. F. Duarte, D. Baron, and R. G. Baraniuk, "Random filters for compressive sampling and reconstruction," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, May 2006.

[26] D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, S. Sarvotham, D. Baron, K. F. Kelly, and R. G. Baraniuk, "A compressed sensing camera: New theory and an implementation using digital micromirrors," in *Proc. Computational Imaging IV at SPIE Electronic Imaging*, San Jose, CA, January 2006.

[27] S. PalChaudhuri, S. Du, A. K. Saha, and D. B. Johnson, "TreeCast: A stateless addressing and routing architecture for sensor networks," in *Proc. International Parallel and Distributed Processing Symposium (IPDPS)*, Santa Fe, NM, Apr. 2004, pp. 221–228.

[28] Jerry Zhao and Ramesh Govindan, "Understanding packet delivery performance in dense wireless sensor networks," in *Proc. Int. Conf. Embedded Networked Sensor Systems (SENSYS)*, 2003, pp. 1–13.

[29] S. D. Servetto, K. Ramchandran, V. A. Vaishampayan, and K. Nahrstedt, "Multiple Description Wavelet Based Image Coding," *IEEE Trans. Image Processing*, vol. 9, no. 5, pp. 813–826, 2000.

[30] Y. Wang, M. T. Orchard, and A. Reibman, "Multiple description image coding for noisy channels by pairing transform coefficients," in *Proc. Workshop Multimedia Signal Processing (MMSP)*, Princeton, NJ, 1997.

[31] S. N. Diggavi and V. A. Vaishampayan, "On multiple description source coding with decoder side information," in *Proc. Information Theory Workshop (ITW)*, Oct. 2004, pp. 88–93.

[32] S. D. Rane, A. Aaron, and B. Girod, "Systematic lossy forward error protection for error-resilient digital video broadcasting," in *Proc. Security, Steganography, and Watermarking of Multimedia Contents VI at SPIE Electronic Imaging*, Jan. 2004, pp. 588–595.

[33] R. Grivonbal, R. M. Figueras i Ventura, and P. Vandergheynst, "A simple test to check the optimality of sparse signal approximations," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Philadelphia, PA, Mar. 2005.

[34] M. F. Duarte, M. Davenport, M. B. Wakin, and R. G. Baraniuk, "Sparse signal detection from incoherent projections," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Toulouse, France, May 2006.

[35] F. Koushanfar, N. Taft, and M. Potkonjak, "Sleeping coordination for comprehensive sensing using isotonic regressiond and domatic partitions," in *Proc. IEEE INFOCOM*, Barcelona, Spain, Apr. 2006.