Spatially-Localized Compressed Sensing and Routing in Multi-Hop Sensor Networks*

Sungwon Lee, Sundeep Pattem, Maheswaran Sathiamoorthy, Bhaskar Krishnamachari, and Antonio Ortega

Dept. of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA,
{sungwonl,pattem,msathiam,bkrishna,antonio.ortega}@usc.edu

Abstract. We propose energy-efficient compressed sensing for wireless sensor networks using spatially-localized sparse projections. To keep the transmission cost for each measurement low, we obtain measurements from clusters of adjacent sensors. With localized projection, we show that joint reconstruction provides significantly better reconstruction than independent reconstruction. We also propose a metric of energy overlap between clusters and basis functions that allows us to characterize the gains of joint reconstruction for different basis functions. Compared with state of the art compressed sensing techniques for sensor network, our experimental results demonstrate significant gains in reconstruction accuracy and transmission cost.

1 Introduction

Joint routing and compression has been studied for efficient data gathering of locally correlated sensor network data. Most of the early works were theoretical in nature and, while providing important insights, ignored the practical details of how compression is to be achieved [1–3]. More recently, it has been shown how practical compression schemes such as distributed wavelets can be adapted to work efficiently with various routing strategies [4–6].

Existing transform-based techniques, including wavelet based approaches [4, 5, 7] and the distributed KLT [8], can reduce the number of bits to be transmitted to the sink thus achieving overall power savings. These transform techniques are essentially critically sampled approaches, so that their cost of gathering scales up with the number of sensors, which could be undesirable when large deployments are considered. Compressed sensing (CS) has been considered as a potential alternative in this context, as the number of samples required (i.e., number of sensors that need to transmit data), depends on the characteristics (sparseness)

* The work described here is supported in part by NSF through grants CNS-0347621, CNS-0627028, CCF-0430061, CNS-0325875, and by NASA through AIST-05-0081. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF or NASA.

of the signal [9–11]. In addition CS is also potentially attractive for wireless sensor networks because most computations take place at the decoder (sink), rather than encoder (sensors), and thus sensors with minimal computational power can efficiently encode data.

However, while the potential benefits of CS for sensor network applications have been recognized [12, 13], significant obstacles remain for it to become competitive with more established (e.g., transform-based) data gathering and compression techniques. A primary reason is that CS theoretical developments have focused on *minimizing the number of measurements* (i.e., the number of samples captured), rather than on *minimizing the cost of each measurement*. In many CS applications (e.g., [14] [15]), each measurement is a linear combination of many (or all) samples of the signal to be reconstructed. It is easy to see why this is not desirable in the context of a sensor network: the signal to be sampled is *spatially distributed* so that measuring a linear combination of all the samples would entail a significant transport cost to generate each aggregate measurement. To address this problem, *sparse measurement* approaches (where each measurement requires information from a few sensors) have been proposed for both single hop [16] and multi-hop [12, 13] sensor networks.

In this paper our goal is make explicit the trade-off between measurement cost and reconstruction quality. We note that lowering transport costs requires spatially localized gathering. Signals to be measured can be expressed in terms of elementary basis functions. We show that the performance of CS greatly depends on the nature of these bases (in particular, whether or not they are spatially localized). Thus, the specific data gathering strategy will depend in general on the signals to be measured. We propose a novel spatially-localized projection technique based on clustering groups of neighboring sensors. Data is gathered in each cluster and efficiently routed to the sink. We show that joint reconstruction of data across clusters leads to significant gains over independent reconstruction. Moreover, we show that reconstruction performance depends on the level of "overlap" between these data-gathering clusters and the elementary basis on which the signals are represented. We propose methods to quantify this spatial overlap, which allow us to design efficient clusters once the bases for the signal are known. Our experimental results demonstrate significant gains over state of the art CS techniques for sensor networks [16, 13].

The remainder of this paper is organized as follows. Section 2 presents CS basics and motivation. Section 3 introduces CS in a multi-hop sensor network. Section 4 presents spatially-localized CS and Section 5 provides experimental results, comparing proposed method and previously proposed CS techniques [16, 13]. Section 6 concludes the paper.

2 Background and Motivation

Compressed Sensing (CS) builds on the observation that an n-sample signal ($\boldsymbol{x}$) having a sparse representation in one basis can be recovered from a small number of projections (smaller than n) onto a second basis that is incoherent with the first [9, 10]. If a signal, $\boldsymbol{x} \in \Re^n$, is sparse in a given basis $\boldsymbol{\Psi}$ (the sparsity inducing basis), $\boldsymbol{x} = \boldsymbol{\Psi a}, |\boldsymbol{a}|_0 = k$, where $k \ll n$, then we can reconstruct the original signal with $O(klogn)$ measurements by finding sparse solutions to under-determined, or ill-conditioned, linear systems of equations, $\boldsymbol{y} = \boldsymbol{\Phi x} = \boldsymbol{\Phi \Psi x} = \boldsymbol{Hx}$, where $\boldsymbol{H}$ is known as the holographic basis. Reconstruction is possible by solving the convex unconstrained optimization problem, $\min_x \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{Hx}\|_2^2 + \gamma\|\boldsymbol{x}\|_1$, if $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ are mutually incoherent [11]. The mutual coherence, $\mu(\boldsymbol{\Phi\Psi}) = \max_{k,j} |\langle \boldsymbol{\phi_k}, \boldsymbol{\psi_j} \rangle|$, serves as a rough characterization of the degree of similarity between the sparsity and measurement systems. For μ to be close to its minimum value, each of the measurement vectors must be spread out in the $\boldsymbol{\Psi}$ domain.

Measurements, y_i, are projections of the data onto the measurement vectors, $y_i = \langle \boldsymbol{\phi_i}, \boldsymbol{x} \rangle$, where $\boldsymbol{\phi_i}$ is the i^{th} row of $\boldsymbol{\Phi}$. Interestingly, independent and identically distributed (i.i.d.) Gaussian, Rademacher (random ± 1) or partial Fourier vectors provide useful universal measurement bases that are incoherent with any given $\boldsymbol{\Psi}$ with high probability. The measurement systems covered in traditional compressed sensing are typically based on these kinds of "dense" matrices, i.e., there are very few zero entries in $\boldsymbol{\Phi}$.

The dense projections of traditional compressed sensing are not suitable for sensor networks due to their high energy consumption. With a dense projection, every sensor is required to transmit its data once for each measurement, so the total cost can potentially be higher than that of a raw data gathering scheme. If the number of samples contributing to each measurement decreases, the cost is reduced by a factor that depends on the sparsity of the measurement matrix (see [12, 13] for an asymptotic analysis for different measurement matrices.) Note, however, that in addition to sparsity, the gathering cost also depends on the position of the sensors whose samples are aggregated in the measurements. If sensors contributing to a given measurement are far apart, the cost will still be significant even with a sparse measurement approach. This is our main motivation to develop spatially-localized sparse projections.

3 Spatially-localized Compressed Sensing

3.1 Low-cost sparse projection based on clustering

In order to design distributed measurements strategies that are both sparse *and* spatially localized, we propose dividing the network into clusters of adjacent nodes and forcing projections to be obtained only from nodes within a cluster. As an example, in this paper we consider two simple clustering approaches. For simplicity, we assume that all clusters contain the same number of nodes. When N_c clusters are used, each cluster will contain $\frac{N}{N_c}$ nodes. In "square clustering", the network is partitioned into a certain number of equal-size square regions. Alternatively, in "SPT-based clustering", we first construct shortest path tree (SPT) then, based on that, we iteratively construct clusters from leaf nodes to the sink. If incomplete clusters encounter nodes where multiple paths merge, we group them into a complete cluster under the assumption that nodes sharing a common parent node are likely to be close to each other.

Any chosen clustering scheme can be represented in CS terms by generating the corresponding measurement matrix, $\boldsymbol{\Phi}$, and using it to reconstruct the original signal. Each row of $\boldsymbol{\Phi}$ represents the aggregation corresponding to one measurement: we place non-zero (or random) coefficients in the positions corresponding to sensors that provide their data for a specific measurement and the other positions are set to zero. Thus, the sparsity of a particular measurement in $\boldsymbol{\Phi}$ depends on the number of nodes participating in this aggregation.

For simplicity, we consider non-overlapped clusters with the same size. This leads to a block-diagonal structure for $\boldsymbol{\Phi}$. Note that recent work [17] [18], seeking to achieve fast CS computation, has also proposed measurement matrices with a block-diagonal structure, with results comparable to those of dense random projections. Our work, however, is motivated by achieving spatially localized projections so that our choice of block-diagonal structure will be constrained by the relative positions of the sensors (each block corresponds to a cluster).

3.2 Sparsity-inducing basis and cluster selection

While it is clear that localized gathering leads to lower costs, it is not obvious how it may impact reconstruction quality. Thus, an important goal of this paper is to study the interaction between localized gathering and reconstruction. A key observation is that in order to achieve both efficient routing and adequate reconstruction accuracy, the structure of the sparsity-inducing basis should be considered. To see this, consider the case where signals captured by the sensor network can be represented by a "global" basis, e.g., DCT, where each basis spans all the sensors in the network. Then the optimally incoherent measurement matrix will be the identity matrix, $\boldsymbol{I}$, thus a good measurement strategy is simply to sample $k \log n$ randomly chosen sensors and then forward each measurement directly to the sink (no aggregation). Alternatively, for a completely localized basis, e.g., $\boldsymbol{\Psi} = \boldsymbol{I}$, a dense projection may be best. However, once the transport costs have been taken into account, it is better to just have sensors transmit data to the sink via the SPT whenever they sense something "new" (e.g., when measurements exceed a threshold). In other words, even if CS theory suggests a given type of measurements (e.g., dense projection for the $\boldsymbol{\Psi} = \boldsymbol{I}$ case), applying these directly may not lead to an efficient routing and therefore efficient distributed CS may not be achievable.

In this paper we consider intermediate cases, in particular those where localized bases with different spatial resolutions are considered (e.g., wavelets). Candes *et al.* [11] have shown that a partial Fourier measurement matrix is incoherent with wavelet bases at fine scales. However, such a dense projection is not suitable for low-cost data gathering for the reasons discussed above. Next we explore appropriate spatially-localized gathering for data that can be represented in localized bases such as wavelets.

4 Efficient clustering for spatially-localized CS

4.1 Independent vs. Joint reconstruction

To study what clustering scheme is appropriate for CS, we first compare two types of reconstruction: independent reconstruction and joint reconstruction. Suppose that we construct a set of clusters of nodes and collect a certain number of local measurements from each cluster. With a given clustering/localized projection, joint reconstruction is performed with the basis where sparseness of signal is originally defined while independent reconstruction is performed with truncated basis functions corresponding to each cluster.

Equation (1) describes an example for two clusters with the same size. $\boldsymbol{\psi_1}$ and $\boldsymbol{\psi_1}$ correspond to localized projections in each cluster. For joint reconstruction, the original sparsity inducing basis, $\boldsymbol{\Psi}$, is employed. But, for independent reconstruction, data in the first cluster are reconstructed with partial basis functions, $\boldsymbol{\psi_1}$ and $\boldsymbol{\psi_2}$, and those in the second cluster are with $\boldsymbol{\psi_3}$ and $\boldsymbol{\psi_4}$ thus, when N_c clusters are involved, independent reconstruction should be performed N_c times, once for each cluster.

$$\boldsymbol{H} = \boldsymbol{\Phi\Psi} = \begin{bmatrix} \boldsymbol{\phi_1} & 0 \\ 0 & \boldsymbol{\phi_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{\psi_1} & \boldsymbol{\psi_2} \\ \boldsymbol{\psi_3} & \boldsymbol{\psi_4} \end{bmatrix} \Rightarrow \begin{Bmatrix} \boldsymbol{H_1} = \left[\boldsymbol{\phi_1\psi_1},\, \boldsymbol{\phi_1\psi_2}\right] \\ \\ \boldsymbol{H_2} = \left[\boldsymbol{\phi_2\psi_3},\, \boldsymbol{\phi_2\psi_4}\right] \end{Bmatrix} \tag{1}$$

Joint reconstruction is expected to outperform independent reconstruction because the sparseness of data is determined by the original basis, $\boldsymbol{\Psi}$. Also, with joint reconstruction, measurements taken from a cluster can also convey information about data in other clusters because basis functions overlapped with more than one clusters can be identified with measurements from those clusters.

As basis functions are overlapped with more clusters, joint reconstruction has potentially higher chance to reconstruct signal correctly. To augment gains from localized CS with clustering scheme, how to choose the clustering should be based on the basis, so that overlap is encouraged. The degree of overlapping between basis functions and clusters can be measured in many different ways. One possible approach is to measure energy of basis functions captured by each cluster.

4.2 Energy overlap analysis

To characterize the distribution of energy of basis functions with respect to the clusters, we present a metric, E_{oa}, and an analysis of the worst-case scenario. We assume that signal is sparse in $\boldsymbol{\Psi} \in \Re^{n \times n}$. And, $\psi(i,j)$ corresponds to the j^{th} entry in the i^{th} column of $\boldsymbol{\Psi}$ and each column is normalized to one. Suppose that N_c is the number of clusters and C_i is a set of nodes contained in the i^{th} cluster. The energy overlap between the i^{th} cluster and the j^{th} basis vector, $E_o(i,j)$, is

$$E_o(i,j) = \sum_{k \in C_i} \psi(j,k)^2 \tag{2}$$

Energy overlap per overlapped basis, E_{oa} Average energy overlap per overlapped basis is a good indicator of distribution of energy of basis functions. For each cluster, $E_{oa}(i)$ is computed as

$$E_{oa}(i) = \frac{1}{N_o(i)} \sum_{j=1}^{N} E_o(i,j),\ \forall i \in \{1, 2, \cdots, N_c\}, \tag{3}$$

where $N_o(i)$ is the number of basis functions overlapped with the i^{th} cluster. Then, we compute E_{oa} by taking average of $E_{oa}(i)$ over all clusters, $E_{oa} = \frac{1}{N_c} \sum_{i=1}^{N_c} E_{oa}(i)$. Intuitively, this metric shows how much energy of basis functions are captured by each cluster. Thus, as energy of basis functions are more evenly distributed over overlapped clusters, E_{oa} decreases, which leads to better reconstruction performance with joint reconstruction. If the specific basis contributing a lot to the cluster is not in the data support, the measurements from this cluster does not notably increase the reconstruction performance.

Worst case analysis It would be useful to have a metric to determine the number of projections required from each local cluster in order to achieve a certain level of reconstruction performance. We first define what the worst-case is, then try to characterize the 'worst-case scenario' performance.

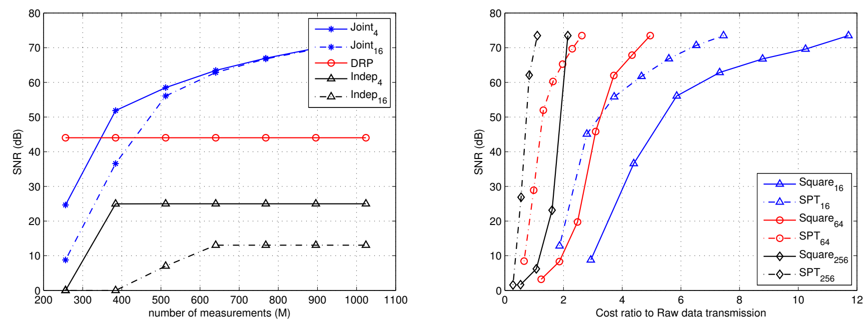
With the global sparsity of K, the worst case scenario is when all K basis vectors supporting data are completely contained in a single cluster. Since the identity of this cluster is not known *a priori* and projections from other clusters not overlapped with those basis vectors do not contribute to reconstruction performance as much as projections from that cluster, $O(K)$ projections would be required from each cluster. But, note that, in general, the coarsest basis vector representing DC component is likely to be overlapped with more than one cluster and chosen as data support for real signal. Thus, in practice, performance could be better than our estimation.

To analyze the worst-case scenario, we assume that we know the basis functions, clustering scheme and the value of K *a priori*. For each cluster, we first choose K basis functions with highest energy overlap with the cluster; in general, finer (more localized) basis functions in spatial domain are likely to be chosen. Then, we compute the sum of the energy overlap of the chosen basis functions. To simplify analysis, we take the average over all clusters.

Minimum number of measurements for each cluster indirectly depends on overlap energy in the worst-case scenario. For example, with DCT basis and four clusters with the same size, overlap energy for each cluster is equal to $\frac{K}{4}$ in the worst-case thus the total number of measurements will be $O(K)$.

5 Experimental Results

For our experiments, we used 500 data generated with 55 random coefficients in different basis. In the network, 1024 nodes are deployed on the square grid and error free communication is assumed. To compare what clustering is apposite for CS, two types of clustering are considered: square-clustering and SPT-based clustering. And we do not assume any priority to clusters for measurements; we collect the same number of localized measurements for each cluster. With localized projection in each cluster, data is reconstructed jointly or independently with Gradient Pursuit for Sparse Reconstruction (GPSR) [19] provided in its package [20]. To evaluate performance, SNR is used to evaluate reconstruction accuracy. For cost evaluation, transmission cost is computed by $\sum(bit) \times (distance)^2$ and the cost ratio is the ratio to the cost for raw data gathering: a simplest scheme that every sensor independently transmits its data along the shortest path to the sink.



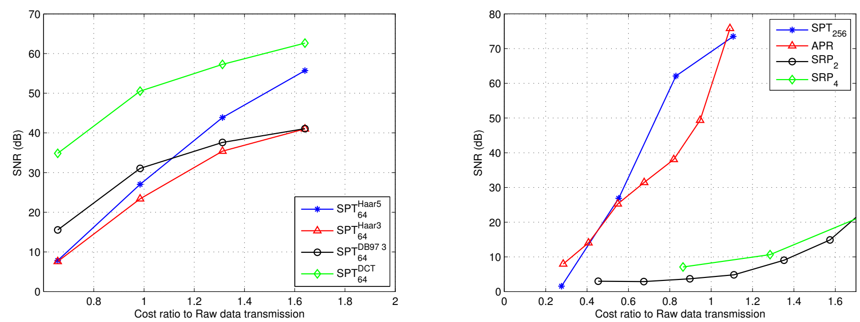
(a) (b)

Fig. 1. Comparison of different types of reconstruction and clustering scheme with Haar basis with decomposition level of 5. (a) Comparison of independent reconstruction and joint reconstruction. (b) Cost ratio to raw data gathering vs. SNR with different number of clusters and clustering schemes

To compare independent reconstruction with joint reconstruction, we used square-clustering scheme with two different number of clusters and Haar basis with decomposition level of 5. In Fig. 1(a), DRP corresponds to the case that takes 256 global measurements from all the nodes in the network then reconstructs data with joint reconstruction. Other curves are generated from localized measurements in each cluster and the two types of reconstruction are applied respectively.

Fig. 1 (a) shows that joint reconstruction outperforms independent reconstruction. As discussed in Section 4.1, joint reconstruction can alleviate the worst situation by taking measurements from other clusters overlapped with basis functions in the data support. In following experiments, all the data was jointly reconstructed.

With joint reconstruction and Haar basis, Fig. 1 (b) shows that SPT-based clustering outperforms square clustering for different number of clusters (N_c). As N_c increases, reconstruction accuracy decreases because measurement matrix becomes sparser as network is separated into more equal-size clusters. However, once the transport costs have been taken into account, more clusters show better performance because cost per each measurement decreases. Since we also observed this trend for different bases, we will focus on 64 SPT-based clusters in following experiments.



(a) (b)

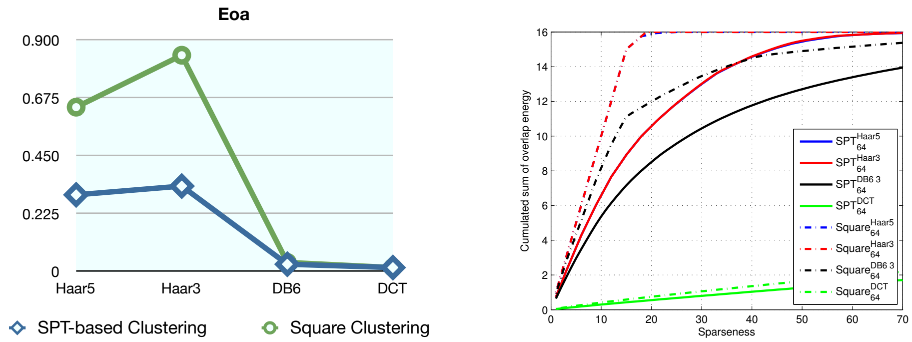
Fig. 2. Performance comparison in terms of cost ratio to raw data gathering vs. SNR (a) for different basis functions and 64 SPT-based clusters. (b) for 256 SPT-based clusters with other CS approaches with Haar basis with level of decomposition of 5

To investigate effects of different bases, we consider the joint reconstruction performance with different basis functions with 64 SPT-based clusters: 1) DCT basis, where each basis vectors have high overlaps in energy which distributed throughout the network 2) Haar basis, where the basis vectors have less overlap and the energy distribution varies from being very localized to global for different basis vectors and 3) Daubechies (DB6) basis, where the overlaps and distribution are intermediate to DCT and Haar. The result in Fig. 2 (a) confirms our intuition. Thus, for the same clustering scheme, the gains from joint reconstruction depend on how"well-spread" the energy in the basis vectors is.

As an indicator of distribution of energy of basis functions, E_{oa} is computed with different clustering schemes for different basis functions. Fig 3 (a) shows that E_{oa} accurately distinguishes performance between two different clusters; SPT-based clusters capture more energy of basis functions than square clusters then lead to better reconstruction. For different basis functions, E_{ov} shows lower overlap energy as basis functions are more spread over in spatial domain.

The result for worst-case analysis is shown in Fig. 3 (b). The results show that, for the same basis function, SPT-based clustering reduces more energy than square clustering thus requires fewer measurements for each cluster. For DCT basis, the energy increases slower than other basis because energy of DCT basis is evenly spread over all clusters. For Haar basis, overlap energy increases very sharply with a few basis functions then it is saturated because energy of some basis functions is concentrated in small regions. With Daubechies basis, the energy is somewhat between two previous bases as we expected.

Fig. 2(b) shows that our approach outperforms other CS approaches [16, 13]. APR corresponds to a scheme that aggregation occurs along the shortest path to the sink and all the sensors on the paths provide their data for measurements.



(a) (b)

Fig. 3. E_{oa} and worst-case analysis with different basis and clustering schemes. Smaller values of metrics indicate more even distribution of overlap energy thus better reconstruction. (a) E_{oa}; values are average over clusters and variances are ignored because they are relatively small. (b) Worst-case analysis; average of cumulated overlap energy with increasing number of basis

SRP with different parameter, $s' = s \times n$, represents a scheme that randomly chooses s' nodes without considering routing then transmit data to the sink via SPT with opportunistic aggregation.

In the comparison, SRP performs worse than the others because, as we expected, taking samples from random nodes for each measurement significantly increases total transmission cost. Our approach and APR are comparable in terms of transmission cost but our approach shows better reconstruction. The performance gap is well explained by N_{oa} for APR and our approach: 0.247 and 0.171 respectively. Lower N_{oa} indicates that energy of basis functions are more evenly distributed over overlapped clusters thus those functions are more likely to be identified with joint reconstruction.

6 Conclusion

We have proposed a framework for efficient data gathering in wireless sensor network by using spatially-localized compressed sensing. With localized projection in each cluster, joint reconstruction has shown better performance than independent reconstruction because joint reconstruction can exploit measurements in multiple clusters, corresponding to energy in a given basis function that overlaps those clusters. Our proposed approach outperforms over state of the art CS techniques for sensor networks [16, 13] because our method achieves power savings with localized aggregation and captures more evenly distributed energy of basis functions. Moreover, we proposed methods to quantify the level of "energy overlap" between the data gathering clusters and the elementary basis on which the signals are represented, which allows us to to design efficient clusters once the bases for the signal are known. Based on the metric, we hope to design an optimal clustering scheme in near future.

References

1. Cristescu, R., Beferull-Lozano, B., Vetterli, M.: On network correlated data gathering. In: INFOCOM 2004. (March 2004)
2. Pattem, S., Krishnamachari, B., Govindan, R.: The impact of spatial correlation on routing with compression in wireless sensor networks. In: IPSN. (April 2004)
3. von Rickenbach, P., Wattenhofer, R.: Gathering correlated data in sensor networks. In: DIALM-POMC, ACM (October 2004)
4. Ciancio, A., Pattem, S., Ortega, A., Krishnamachari, B.: Energy-efficient data representation and routing for wireless sensor networks based on a distributed wavelet compression algorithm. In: IPSN. (April 2006)
5. Shen, G., Ortega, A.: Joint routing and 2d transform optimization for irregular sensor network grids using wavelet lifting. In: IPSN. (April 2008)
6. Pattem, S., Shen, G., Chen, Y., Krishnamachari, B., Ortega, A.: Senzip: An architecture for distributed en-route compression in wireless sensor networks. In: ESSA. (April 2009)
7. Wagner, R., Choi, H., Baraniuk, R., Delouille, V.: Distributed wavelet transform for irregular sensor network grids. In: SSP. (July 2005)
8. Gastpar, M., Dragotti, P., Vetterli, M.: The distributed karhunen-loeve transform. In: MMSP. (December 2002)
9. Donoho, D.L.: Compressed sensing. In: IEEE Transactions on Information Theory, IEEE (April 2006)
10. Candes, E., Romberg, J., Tao, T.: Robust uncertainity principles : exact signal reconstruction from highly incomplete frequency information. In: IEEE Transactions on Information Theory, IEEE (February 2006) 489–509
11. Candes, E., Romberg, J.: Sparsity and incoherence in compressive sampling. In: Inverse Problems. (June 2007)
12. Lee, S., Pattem, S., Sathiamoorthy, M., Krishnamachari, B., Ortega, A.: Compressed sensing and routing in sensor networks. In: USC CENG Technical Report. (April 2009)
13. Quer, G., Masierto, R., Munaretto, D., Rossi, M., Widmer, J., Zorzi, M.: On the interplay between routing and signal representation for compressive sensing in wireless sensor network. In: ITA. (February 2009)
14. Lustig, M., Donoho, D., Pauly, J.M.: Sparse mri: The application of compressed sensing for rapid mr imaging. In: Magnetic Resonance in Medicine. (December 2007)
15. Duarte, M.F., Davenport, M.A., Takhar, D., Laska, J.N., Sun, T., Kelly, K.F., Baraniuk, R.G.: Single pixel imaging via compressive sampling. In: IEEE Signal Processing Magazine. (March 2008)
16. Wang, W., Garofalakis, M., Ramchandran, K.: Distributed sparse random projections for refinable approximation. In: IPSN, Springer Verlag (April 2007)
17. Gan, L., Do, T.T., Tran, T.D.: fast compressive imaging using scrambled block hadamard ensemble. In: preprint. (2008)
18. Do, T., Tran, T., Gan, L.: Fast compressive sampling with structurally random matrices. In: ICASSP. (April 2008)
19. Figueiredo, M., Nowak, R., Wright, S.: Gradient projection for sparse reconstruction:application to compressed sensing and other inverse problems. In: IEEE Journal of Selected Topics in Signal Processing. (2007)
20. Figueiredo, M., Nowak, R., Wright, S.: Gradient projection for sparse reconstruction. http://www.lx.it.pt/ mtf/GPSR/ (January 2009)