Comparison of OMP and SOMP in the reconstruction of compressively sensed Hyperspectral Images

Aravind N V, Abhinandan K, Vineeth V Acharya, and Sumam David S.
Department of Electronics and Communication Engineering, National Institute of Technology Karnataka, Surathkal, India
e-mail: sumam@ieee.org

***Abstract*—In this paper, we present a novel method for the acquisition and compression of hyperspectral images based on two concepts – distributed source coding and compressive sensing. Compressive sensing (CS) is a signal acquisition method that samples at sub Nyquist rates which is possible for signals that are sparse in some transform domain. Distributed source coding (DSC) is a method to encode correlated sources separately and decode them together in an attempt to shift complexity from the encoder to the decoder. Distributed compressive sensing (DCS) is a new framework suggested for jointly sparse signals which we apply to the correlated bands of hyperspectral images. We compressively sense each band of the hyperspectral image individually and can then recover the bands separately or using a joint recovery method. We use the Orthogonal Matching Pursuit (OMP) for individual recovery and Simultaneous Orthogonal Matching Pursuit (SOMP) for joint decoding and compare the two methods. The latter is shown to perform consistently better showing that the Distributed Compressive Sensing method that exploits the joint sparsity of the hyperspectral image is much better than individual recovery.**

Keywords-Hyperspectral images, distributed source coding, compressive sensing, distributed compressive sensing, orthogonal matching pursuit, simultaneous orthogonal matching pursuit

I. INTRODUCTION

Hyperspectral imaging collects information across different bands in the electromagnetic spectrum that can be used for identification of different materials and features on the earth surface. However, the amount of data collected by hyperspectral sensors is very huge and hence compression is a must. Also, there is a need to keep the processing time onboard the aerial platform as low as possible so as to enable fast collection and transmission of data. The on board acquisition system and consequent encoder must be kept as simple as possible due to constraints on memory and size imposed by the nature of the aerial platform.

Conventional imaging consists of 2 stages – first huge amounts of data are collected by a large number of sensors and second, the data is compressed without substantially affecting the quality. Compressive sensing/sampling (CS) uses a different methodology by only acquiring some samples of the entire data thus removing the need for a separate compression step. This is possible due to the fact that the original data is usually sparse in some domain (e.g. DCT) and even in conventional compression most coefficients in that domain are insignificant. So for images (hyperspectral or otherwise), application of compressive sampling not only eliminates the separate compression step, it also helps in reducing the number of sensors needed. Thus, the bands of the hyperspectral image can be compressively sensed one at a time and one of many possible methods of reconstruction can be used to recover them individually.

The high correlation between the bands of hyperspectral images can also be exploited using the concept of distributed source coding (DSC). The bands can be encoded separately and then jointly decoded at a ground station. Thus if each band is compressively sensed and the bands are recovered together, the system thus developed is a form of distributed compressive sensing (DCS).

The paper is organized as follows. Section II gives a brief description of Distributed Source Coding (DSC). Section III deals with the theoretical aspects of Compressive Sensing (CS). Section IV describes the procedure of Orthogonal Matching Pursuit (OMP) which is a specific method of reconstruction used in CS. Section V briefly explains the paradigm of Distributed Compressive Sensing (DCS). Simultaneous Orthogonal Matching Pursuit (SOMP) which is an example of a reconstruction method in DCS is explained in Section VI. The experimental results obtained by comparing OMP and SOMP are given in Section VII.

II. DISTRIBUTED SOURCE CODING

Distributed source coding considers a situation in which two or more statistically dependent data sources must be encoded. In conventional compression, a single joint encoder exploits the statistical dependence of the source signals. However, efficient compression can also be achieved by exploiting source statistics at the decoder only. This way, the complexity is shifted from the encoder to the decoder.

Consider a communication system with two correlated signals that are encoded independently i.e. are distributed. Assume the receiver, on the other hand, can see both encoded signals and can perform joint decoding. Following the standard encoding paradigm, each source can be compressed lossless, with a total rate no less than the sum of the two source entropies. This is clearly less efficient than an encoder that jointly compresses the two sources where a bit rate equal to the joint entropy of the sources could be used. The surprising result of DSC theory as given by David Slepian and Jack Keil Wolf [1] is that, under certain assumptions, the same result can be achieved by using two separate encoders, provided that the two

978-1-4244-9799-7/11/$26.00 ©2011 IEEE

sources are decoded by a joint decoder. This bound was extended to cover more than two correlated sources by Thomas M. Cover [2]. With regard to lossy coding of joint Gaussian sources,similar results were obtained by Aaron D. Wyner and Jacob Ziv [3].

The concept has a strong potential for remote sensing image compression as we can exploit the correlation between two (or more) bands of a multispectral or hyperspectral image to achieve lower encoder complexity by avoiding explicit decorrelation of the bands. Current work on the application of DSC to hyperspectral image compression has been presented clearly by Magli et al.[4].

III. COMPRESSIVE SENSING

Shannon-Nyquist's sampling theorem states that the sampling frequency of a signal must be at least two times the highest frequency present in the signal to prevent information loss through aliasing. CS is a new sampling theory that states that compressible signals can be reconstructed using far fewer samples than Shannon suggests. CS was first put forth by Emmanuel J.Candes [5] in 2004 while working on a problem in magnetic resonance imaging who discovered that a test image could be reconstructed exactly even with data deemed insufficient by the Shannon-Nyquist criterion.

Compressive Sampling can be explained as given in [6] as follows. Consider a real valued, discrete signal $\boldsymbol{x}$, which can be viewed as an $N \times 1$ column vector in R^N with elements $x[n]$, $n = 1, \ldots, N$. Suppose any signal in R^N can be represented in terms of a basis of $N \times 1$ vectors $\{\psi_i\}_{i=1}^{N}$. The signal $\boldsymbol{x}$ can be expressed as

$$\boldsymbol{x} = \sum_{i=1}^{N} s_i \boldsymbol{\psi}_i \text{ or } \boldsymbol{x} = \boldsymbol{\psi s} \qquad (1)$$

where $\boldsymbol{s}$ is the $N \times 1$ column vector of weighting coefficients. Clearly, $\boldsymbol{x}$ and $\boldsymbol{s}$ are equivalent representations of the signal, with $\boldsymbol{x}$ in the time or space domain and $\boldsymbol{s}$ in the ψ domain. $\boldsymbol{x}$ is *K-sparse* if it is a linear combination of only K basis vectors; in other words only K of the s_i coefficients in equation (1) are nonzero and the remaining are zero or small enough to be approximated to zero. $\boldsymbol{x}$ is said to be compressible if $K << N$.

Consider a linear measurement process that computes $M<N$ inner products between $\boldsymbol{x}$ and a collection of vectors $\{\boldsymbol{\Phi}_j\}_{j=1}^{M}$. Arrange the measurements y_i in an $M \times 1$ vector $\boldsymbol{y}$ and the measurement vectors $\boldsymbol{\Phi}_j^T$ as rows in an $M \times N$ matrix $\boldsymbol{\Phi}$. Then, by substituting $\boldsymbol{x}$ from equation (1), $\boldsymbol{y}$ can be written as

$$\boldsymbol{y} = \boldsymbol{\Phi x} = \boldsymbol{\Phi \psi s} = \boldsymbol{\Theta s} \qquad (2)$$

where $\boldsymbol{\Theta} = \boldsymbol{\Phi} \boldsymbol{\psi}$ is an $M \times N$ matrix.

The measurement process is not adaptive, meaning that $\boldsymbol{\Phi}$ is fixed and does not depend on the signal $\boldsymbol{x}$. The problem consists of designing a stable measurement matrix $\boldsymbol{\Phi}$ so that no important information is damaged by the dimensionality reduction and a reconstruction algorithm to recover $\boldsymbol{x}$ from only M measurements of $\boldsymbol{y}$ is possible. The choice of $\boldsymbol{\Phi}$ and $\boldsymbol{\psi}$ is critical for CS. In general, we can design a stable measurement matrix based on certain properties (like Restricted Isometry Property) as stated in [6].

The reconstruction of $\boldsymbol{x}$ or equivalently, $\boldsymbol{s}$ from vector $\boldsymbol{y}$ of M samples is not trivial. The exact solution is NP-hard and consists of finding the minimum L_0 norm (the number of nonzero elements). However, excellent approximation can be obtained via the L_1 norm minimization given by:

$$\tilde{\boldsymbol{s}} = argmin\|\boldsymbol{s}'\|_1 \text{ } such \text{ } that \text{ } \boldsymbol{\Phi\psi s}' = \boldsymbol{y} \qquad (3)$$

where $\tilde{\boldsymbol{s}}$ is the reconstructed $\boldsymbol{s}$.

This convex optimization problem, namely, basis pursuit can be solved using a linear program algorithm of $O(N^3)$ complexity [7]. Due to complexity and low speed of linear programming algorithms, faster solutions were proposed at the expense of slightly more measurements, such as matching pursuit, tree matching pursuit, orthogonal matching pursuit (OMP) [8], and TwIST algorithm [9].

There has been very limited research on the applications of compressive sensing theories for remote sensing and multi-spectral images. Jianwei Ma [10] proposed two possible systems, SPMT (single-pixel but multi-time) and MPST (multi-pixel but single-time) for different applications of CS in aerospace remote sensing. As suggested in [10] and [11], we also have employed a noiselet transform [12] as the measurement matrix due to the fact that a fast transform is available allowing low computational cost. The sparse domain chosen is the DCT domain.

IV. ORTHOGONAL MATCHING PURSUIT

The OMP Algorithm [8] is a sparse approximation algorithm. From equation (2) we have $\boldsymbol{y} = \boldsymbol{\Theta s}$. $\boldsymbol{y}$ is a linear combination of $\mathbf{m}$ columns from $\boldsymbol{\Theta}$. Denote the columns of $\boldsymbol{\Theta}$ by $\boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_N$.

To identify $\boldsymbol{s}$, we need to determine which columns of $\boldsymbol{\Theta}$ participate in the measurement vector $\boldsymbol{y}$. The idea behind the algorithm is to pick columns in a greedy fashion. In each iteration, we choose the column of $\boldsymbol{\Theta}$ that is most strongly correlated with the remaining part of $\boldsymbol{y}$. Then we subtract off its contribution to $\boldsymbol{y}$ and iterate on the residual. The basic algorithm is described below as given in [8]. The outputs are $\tilde{\boldsymbol{s}}$ which is the estimate for the ideal signal and set $\boldsymbol{\Lambda}_M$ containing M elements from 1,.,N.

Procedure:

1. Initialize the residual $\mathbf{r}_0 = \boldsymbol{y}$, the index set $\boldsymbol{\Lambda}_0 = \emptyset$, and the iteration counter t = 1.
2. Find index λ_t that solves the optimization problem
$$\lambda_t = \arg max_{j=1,2\ldots,N} |\langle \boldsymbol{r}_{t-1}, \boldsymbol{\varphi}_j \rangle| \qquad (4)$$
3. If the maximum occurs for multiple indices, the tie is broken deterministically.

4. Augment the index set $\boldsymbol{\Lambda}_t = \boldsymbol{\Lambda}_{t-1} \cup \{\lambda_t\}$ and the matrix of chosen atoms
$$\boldsymbol{\theta}_t = [\boldsymbol{\theta}_{t-1}, \boldsymbol{\varphi}_{j_t}] \quad (5)$$
By convention, $\boldsymbol{\theta}_0$ is taken to be an empty matrix.
5. Solve a least-squares problem to obtain a new signal estimate
$$\boldsymbol{x}_t = \arg min_x \|\boldsymbol{\theta}_t \boldsymbol{x} - \boldsymbol{y}\|_2 \quad (6)$$
6. Calculate the new approximation of the data and the new residual
$$\boldsymbol{a}_t = \boldsymbol{\theta}_t \boldsymbol{x}_t \quad (7)$$
$$\boldsymbol{r}_t = \boldsymbol{y} - \boldsymbol{a}_t \quad (8)$$
7. Increment t, and return to Step 2 if t < m.
8. The estimate $\tilde{s}$ for the ideal signal has nonzero indices at the components listed in Λ_m. The value of the estimate $\tilde{s}$ in component λ_j equals the j[th] component of $\boldsymbol{x}_t$.

V. DISTRIBUTED COMPRESSIVE SENSING

Distributed Compressive Sensing (DCS) is a combination of DSC and CS and was put forward by Baron et.al. [13]. In a typical DCS scenario, a number of sensors measure signals that are each individually sparse in some basis and also correlated from sensor to sensor. The theory rests on a concept termed joint sparsity– the sparsity of the entire signal space. There are 3 models of joint sparsity suggested in [13] – sparse common component with innovations, common sparse support and non-sparse common component with sparse innovations. An example of its implementation is given in [14] where the concept of DCS is applied to colour images.

The common sparse support model seems to be the easiest that can be applied in our case since the model assumes a zero common component (and hence we need not find a common component across the bands of the hyperspectral image). In this model, the innovations remain sparse but the common component is equal to zero. Also, the innovations share the same sparse support which means that the index of the coefficients are the same for all signals but the value of the coefficient themselves vary.

VI. SIMULTANEOUS ORTHOGONAL MATCHING PURSUIT

To recover correlated signals, a greedy pursuit method, Simultaneous Orthogonal Matching Pursuit(SOMP), has been proposed in [15]. The algorithm is very similar to the OMP algorithm described before but with a few minor changes.

Let there be B compressively sampled signals $\mathbf{y}_1, \mathbf{y}_2, \ldots \mathbf{y}_B$. The second step of the OMP algorithm is modified so as to find the index λ_t that solves the easy optimization problem as
$$\max_{\omega\in\Omega} \sum_{k=1}^{B} |\langle \boldsymbol{R}_{k,t-1}, \boldsymbol{\varphi}_j \rangle| \quad (9)$$
where $\boldsymbol{R}_{k,t-1}$ is the residual of the k[th] compressively sampled signal. The rest of the procedure remains the same with each of $\mathbf{s}_1, \mathbf{s}_2, \ldots \mathbf{s}_B$ being reconstructed. This procedure reduces to standard Orthogonal Matching Pursuit [8] when B = 1. The idea behind maximizing the sum of absolute correlations is that we wish to find that column that contributes the most energy to as many of the input signals as possible.

The implementation of the above is done over a small group of consecutive bands and not all the bands taken together (since the latter would require large memory and take a lot of time).

VII. EXPERIMENTAL RESULTS

Experiments were conducted to compare the relative performance of compressive sensing using OMP and distributed compressive sensing using SOMP. All experiments were performed on Airborne Visual/InfraRed Imaging Spectrometer (AVIRIS) images which contain data of 224 contiguous spectral channels (bands) with wavelengths from 400 to 2500 nanometres. Three standard data products available for download from NASA's AVIRIS website – Cuprite, Lunar lake and Jasper Ridge have been used [16]. All the proposed algorithms have been applied on the radiance data of the three data sets.

The AVIRIS images were cropped to size 256x256. All compressive sensing was done using 12.5% measurements (i.e. 1/8[th] of the original number of pixels = 8196 or a compression ratio of 0.125). Reconstructions are on compressively sensed images taking DCT as the sparse domain and random noiselet transform as the measurement matrix. Initially, the programs were run on the first 5 bands of the 3 AVIRIS images and this was extended to the first 10 bands and results were found to be consistent. The results in terms of average MSE and SNR over the 10 bands of the images reconstructed by OMP and SOMP are given in Table I. Figures 1, 2 and 3 shows the SNR comparison for the 3 images by the two methods and Figure 4 shows the time for reconstruction using SOMP and OMP for the three images.

Algorithms were implemented in MATLAB version 7.0.1 on a workstation with Intel Dual Core 1.73 Ghz processor using the l1-magic toolbox [7] and the Sparsify toolbox [17], [18]. The SOMP algorithm was developed from the OMP algorithm in the Sparsify toolbox which is based on QR factorisation [17], [19]. The SNR was computed using the following formula
$$SNR = 10\log\left(\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\{I(i,j)\}^2}{M\cdot N\cdot MSE}\right) \quad (9)$$
where the image is of size MxN and I(i,j) is the brightness value of the pixel in i[th] row and j[th]column. Reconstruction time was estimated using the etime function available in MATLAB.

As is evident from Table I and Figures 1, 2 and 3, performance of SOMP is much better than OMP in terms of MSE and SNR with an improvement of 2723.89 in MSE and 0.961dB in SNR on average. Also the time for reconstruction of the 10 bands together using SOMP is much lesser than that required to reconstruct all 10 bands individually by OMP as shown in Figure 4. Thus the usage of Distributed Source Coding in tandem with Compressive Sensing methods does improve performance as well as reconstruction time.

TABLE I. MSE AND SNR FOR RECONSTRUCTION BY OMP AND SOMP

AVIRIS image name	Reconstruction by OMP (average over all bands)		Reconstruction by SOMP (average over all bands)	
	mse	SNR (in db)	**mse**	**SNR (in dB)**
Cuprite	27055.85	26.8364	**25416.45**	**27.7181**
Jasper	7635.71	26.6065	**6218.68**	**27.8054**
Lunar Lake	80208.29	24.5171	**75093.06**	**25.3197**
Average	38299.95	25.987	**35576.06**	**26.948**

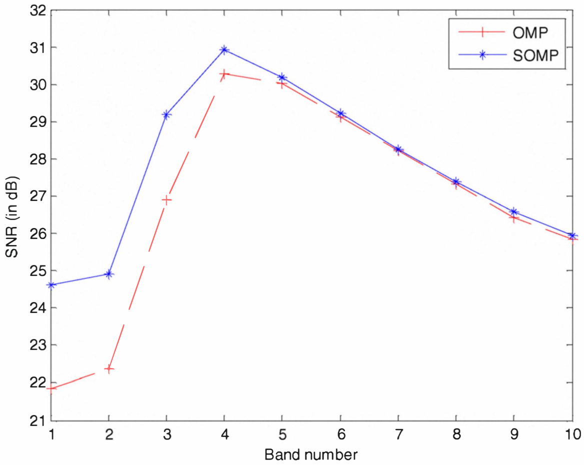

Figure 1. Comparison of SNR of OMP reconstructed and SOMP reconstructed bands of the Cuprite image for the first 10 bands

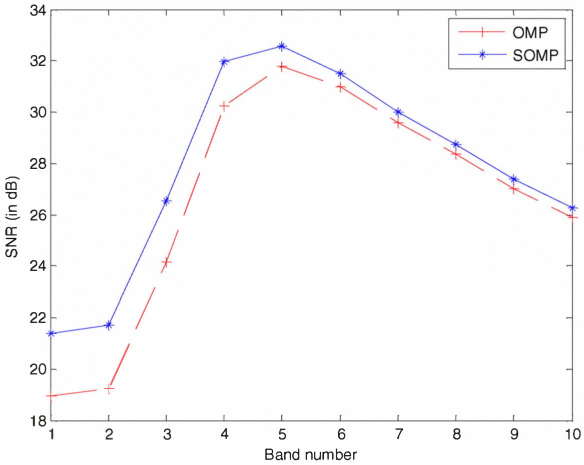

Figure 2. Comparison of SNR of OMP reconstructed and SOMP reconstructed bands of the Jasper image for the first 10 bands

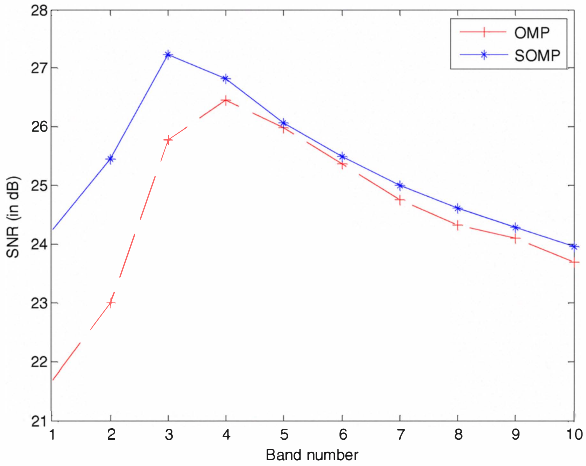

Figure 3. Comparison of SNR of OMP reconstructed and SOMP reconstructed bands of the Lunar Lake image for the first 10 bands

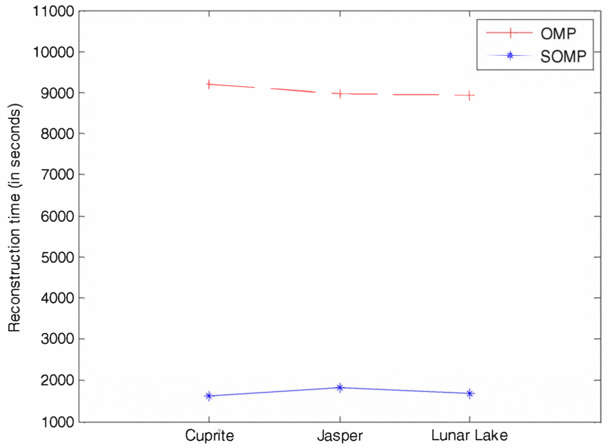

Figure 4. Comparison of the reconstruction time for first 10 bands using OMP and SOMP for the 3 hyperspectral images – Cuprite, Jasper and Lunar lake

VIII. CONCLUSIONS AND FUTURE WORK

Performance of SOMP is much better than OMP both in terms of quality of reconstructed image and time for reconstruction. This shows that distributed compressive sensing techniques can indeed improve on compressive sensing methods. Thus, jointly reconstructing small groups of consecutive bands using some recovery method like SOMP is better than individual recovery.

As far as compressive sensing is concerned, there has been little emphasis on hyperspectral images thus far, showing the need for more research work on them. While we implemented one method of compressive sensing and reconstruction on hyperspectral images, remaining methods found in literature can also be implemented to choose the most appropriate one. This could involve variation of the measurement matrix, the sparse domain and the reconstruction method to find which would give better performance. A hardware framework of how the sensors could be adapted for compressive sensing of hyperspectral images will also need to be designed and specified.

REFERENCES

[1] D. Slepian and J.K. Wolf, "Noiseless coding of correlated information sources", IEEE Trans. Inform. Theory, pp. 471-480, July 1973

[2] T.M. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources", IEEE Trans. Inform. Theory,pp. 226-228, March 1975

[3] A.D. Wyner and J. Ziv, "The rate distortion function for source coding with side information at the decoder", IEEE Trans. Information Theory, vol. 22, no. 1, pp. 1–10, Jan. 1976

[4] E. Magli, M. Barni, A. Abrardo, and M. Grangetto, "Distributed source coding techniques for lossless compression of hyperspectral images", EURASIP Journal on Advances in Signal Processing, vol. 2007 Issue 1, 2007

[5] E. J. Candès, "Compressive sampling", Proceedings of the International Congress of Mathematicians, Madrid, Spain, vol. 3, pp. 1433–1452, 2006.

[6] R. G. Baraniuk, "A Lecture on Compressive Sensing." IEEE Signal Processing Magazine, v. 24, iss. 4, pp. 118-121, July 2007

[7] E. J. Candes and J. Romberg, "l1-magic : Recovery of Sparse Signals", Caltech, 2005.

[8] J. A. Tropp and A. C. Gilbert: "Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit", IEEE Transactions on Information Theory 53 (12), pp. 4655-4666, 2007

[9] J. M. Bioucas-Dias and M. A. T. Figueiredo, "A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration," IEEE Trans. Image Process., vol. 16, no. 12, pp. 2992-3004, Dec. 2007

[10] J. Ma, "Single-Pixel Remote Sensing", IEEE Geoscience and Remote Sensing Letters, vol. 6, issue 2, pp. 199-203, 2009

[11] J. Bobin, J. Starck, and R. Ottensamer, "Compressed Sensing in Astronomy," IEEE J. Sel. Topics Signal Process., vol. 2, no. 5, pp. 718–726, Oct. 2008

[12] R. Coifman, F. Geshwind, and Y. Meyer, "Noiselets," Appl. Comput. Harmon. Anal., vol. 10, no. 1, pp. 27–44, Jan. 2001

[13] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, "Distributed compressive sensing", Preprint, 2005

[14] Pradeep Nagesh and Baoxin Li, "Compressive imaging of color images", Proc. ICASSP Taipei, pp. 1261-1264, 2009

[15] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Simultaneous sparse approximation via greedy pursuit", Proc. ICASSP, Philadelphia , 2005

[16] AVIRIS Homepage, JPL, NASA. [Online]. Available:http://aviris.jpl.nasa.gov/

[17] T T. Blumensath and M. E. Davies; "Gradient Pursuits", IEEE Transactions on Signal Processing, vol. 56, no. 6, pp. 2370-2382, June 2008.

[18] Thomas Blumensath Software, Sparsify toolbox.[Online].Available: http://www.personal.soton.ac.uk/tb1m08/sparsify/sparsify.html

[19] G. Davis, S. Mallat and M. Avellaneda, "Greedy adaptive approximation," Journal of Constructive Approximation, vol. 12, pp. 57–98, 1997.