Quantifying the performance of compressive sensing on scalp EEG signals

Amir M. Abdulghani, Alexander J. Casson and Esther Rodriguez-Villegas
Electrical and Electronic Engineering Department, Imperial College London, London, SW7 2AZ
Email: {amirm,acasson,e.rodriguez}@imperial.ac.uk

***Abstract*—Compressive sensing is a new data compression paradigm that has shown significant promise in fields such as MRI. However, the practical performance of the theory very much depends on the characteristics of the signal being sensed. As such the utility of the technique cannot be extrapolated from one application to another. Electroencephalography (EEG) is a fundamental tool for the investigation of many neurological disorders and is increasingly also used in many non-medical applications, such as Brain-Computer Interfaces. This paper characterises in detail the practical performance of different implementations of the compressive sensing theory when applied to scalp EEG signals for the first time. The results are of particular interest for wearable EEG communication systems requiring low power, real-time compression of the EEG data.**

I. Introduction

Electroencephalography (EEG) is the technique of measuring electrical signals generated within the brain by placing electrodes on the scalp. The EEG signal produced provides a non-invasive, high time resolution, interface to the brain, and as such the EEG is a key diagnosis tool for conditions such as epilepsy, and it is frequently used in Brain-Computer Interfaces [1]. In portable EEG systems the entire recording unit is battery powered, and the physical size of the batteries sets the overall device size and operational lifetime. The current technological trend is thus towards portable EEG systems that are as small and unobtrusive as possible, and that can record for very long periods of time [1].

It has been demonstrated that the use of low power, real-time data compression embedded in the portable EEG recorder itself is essential for such EEG systems to be realised [1]. Previous, offline, EEG compression schemes have achieved up to 65% data reduction with lossless compression [2], and up to 89% data reduction when lossy compression is employed [3]. However, to satisfy the constraints of real-time and low power operation it is essential that the computational complexity of the data compression algorithm to be embedded on the portable EEG system is kept low. This is not a requirement for many of the offline compression systems developed previously. In contrast, while the computational complexity of the algorithm on the portable EEG unit must be low, once the EEG data has been moved from the portable unit to a non-portable computer for storage or analysis there is no intrinsic need for low computational complexity algorithms as the power requirements of the fixed computer installation are much more relaxed.

Recently, a new compression technique named compressive sensing has been reported that can potentially satisfy these requirements for low complexity in the portable part of the system, instead utilising the unlimited power available in the fixed computer installation. It is thus potentially of significant interest for use in portable EEG systems [4]. Furthermore, compressive sensing has shown excellent performance in terms of compression ratio and reconstruction error in applications such as MRI [5], speech [6], and image/video coding [7]. The operation of compressive sensing, however, is based upon the assumptions that: the signal to be sensed is *sparse* in a particular domain (see Section II for definition) and that this domain is incoherent with a given measurement matrix. The validity of these assumptions differs from application to application, and so the performance of compressive sensing on a particular signal cannot be assumed *a priori*.

Very preliminary results have shown that compressive sensing may be suitable for use with scalp EEG signals [8]. However, until now, representative testing of the technique to assess its performance, merit and limitations using a large EEG test dataset has not been done. This paper provides this quantitative and comprehensive characterisation of compressive sensing performance when applied to scalp EEG signals by presenting performance results for 18 different implementations of the compressive sensing theory using a large, multi-channel, EEG data set. This provides essential information for guiding the choice of compressive sensing implementation for use in EEG systems, and in guiding future compressive sensing development.

The remainder of this paper is organised as follows. Section II summaries the core compressive sensing theory. Section III then describes the methods used to apply the theory to EEG signals with qualitative and quantitative reconstruction performance results presented in Section IV. Finally interpretations and conclusions are presented in Section V.

II. Compressive sensing theory

Compressive sensing is a lossy compression scheme based upon exploiting known information in the signal of interest to lower the effective sampling rate. The inherent redundancy in specific types of signals thus allows *compression while sampling*. A detailed introduction to the theory can be found in [9], [10]. Below is presented an overview of the signal compression and reconstruction methods to illustrate the procedure and to highlight the implementation choices that motivate the performance characterisation presented in this work.

A. Compression process

Compressive sensing theory starts from the assumption that a signal is *sparse* in a particular domain. A vector, of length N, is K sparse if it has K non-zero entries and the remaining $N-K$ entries are all zero. To illustrate this, consider a single-channel of digitised EEG data, **x**, which is an $N \times 1$ vector. Then assume that this signal can be represented by a projection onto a different basis set:

$$x = \sum_{i=1}^{N} s_i \Psi_i \text{ or } \mathbf{x} = \mathbf{\Psi s} \tag{1}$$

where **s** is an $N \times 1$ vector and $\mathbf{\Psi}$ is an $N \times N$ basis matrix. The vector **s** is given by the inner product of **x** and $\mathbf{\Psi}$, and the entries in $\mathbf{\Psi}$ are known as the dictionary functions. As an example, if $\mathbf{\Psi}$ is the Fourier dictionary of complex exponential functions, **s** is the Fourier transform of **x** and both **s** and **x** represent the signal equivalently, but in different domains. Compressive sensing assumes that a basis set $\mathbf{\Psi}$ is available in which **s** is sparse. Different choices for $\mathbf{\Psi}$ are available leading to one of the characterisation steps investigated here.

To actually compress the signal only a computationally simple operation is performed. In addition to the projection above, it is assumed that **x** can be related to another signal **y**:

$$\mathbf{y} = \mathbf{\Phi x} \tag{2}$$

where $\mathbf{\Phi}$ is a *measurement matrix* of dimensions $M \times N$ and **y** is the compressively sensed version of **x**. **y** has dimensions $M \times 1$ and if $M < N$ data compression is achieved. Provided that $\mathbf{\Phi}$ is correctly chosen, exact reconstruction of **x** from **y** is possible even though **y** has fewer samples than a signal sampled at the Nyquist rate. The *effective* sampling rate has thus been lowered. It can be shown that this technique is possible if $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are incoherent; that is if the elements of $\mathbf{\Phi}$ and $\mathbf{\Psi}$ have low correlation [11]. In general, to satisfy this condition $\mathbf{\Phi}$ is chosen as a random matrix following a given probability distribution. Again multiple choices for $\mathbf{\Phi}$ are available.

B. Signal reconstruction

The vector **y** is thus generated on the portable EEG unit and represents the compressively sensed signal **x**. To view and process the EEG signal at the non-portable computer the vector **x** must be recovered from the recorded signal **y**. This is done by solving the non-linear optimisation problem:

$$\min_{\mathbf{s} \in \Re^N} \|\mathbf{s}\|_{l_0} \text{ subject to } y_i = \langle \Phi_i, \mathbf{\Psi s} \rangle. \tag{3}$$

That is, find the vector **s** that is most sparse and best satisfies the observations made. This of course comes from the assumption that **s** is good sparse representation of the signal.

In practice the solution of (3) is a highly non-convex optimisation problem, and in general impractical even in the non-power constrained, non-portable part of the system. Instead, under certain conditions, solving the l_1 norm case of (3) gives the same solution, but can be computed in polynomial time [12]. In general, and in the remainder of this article, the l_1 norm and iterative algorithms which ensure strongly polynomial running times, are thus used. Again, multiple methods for carrying out this l_1 optimisation are possible.

III. Performance characterisation methods

This paper uses scalp EEG signals and assesses the practical reconstruction performance of compressive sensing using a number of different dictionary functions and signal reconstruction methods. This is done by using MATLAB to compress and reconstruct pre-recorded scalp EEG signals and then quantifying the amount of reconstruction error introduced.

A. Dictionary functions

Key to suitable choices for $\mathbf{\Psi}$ is that the resulting vector **s** must represent the EEG signal as sparsely as possible. Six different basis matrices $\mathbf{\Psi}$ are used in this work, each based on a different set of dictionary functions. Firstly the Gabor dictionary, as used in [8], [13], is used. Functions in this dictionary are defined by Gaussian envelope sinusoidal pulses:

$$\Psi_i(n, \omega, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(n-n_0)^2/\sigma^2} \cos(\omega n + \theta) \tag{4}$$

where n is the sample number, n_0 is the sample number of the centre of the envelope, $\omega \geq 0$ is the frequency of the sinusoid, $\sigma > 0$ is the spread of the envelope, and θ is the phase angle. Here the settings $\omega = 25$, $\theta = \{0, \pi/2\}$, and $\sigma = 0.015$ are used with these choices being based upon observations after running several preliminary simulations. A total of 2250 functions are thus present in the dictionary with this size being chosen to facilitate quasi-real time reconstruction.

The second dictionary basis is the Mexican hat, which is often used for time-frequency analysis of EEG signals. Here the dictionary functions are defined by the second derivative of Gaussian functions:

$$\Psi_i(n) = \frac{2}{3} \pi^{-1/4} (1 - n^2) e^{-n^2/2} \tag{5}$$

where n is the interval over which the Mexican hat is defined. Here the Mexican hat dictionary has been defined with two intervals of $\{-5, 5\}$ and one interval of $\{-1, 1\}$ giving a total of 2250 functions in the dictionary by shifting the centre of the generated functions.

The last four dictionaries considered are spline based. These have also been used previously for EEG feature extraction and moreover are known to offer very compact support. Based upon [14] multi-resolution-like spline dictionaries, both linear and cubic varieties, are used. Based upon [15] a B-spline dictionary is also used, again in linear and cubic variants. Both dictionaries are constructed on the interval $\{1, 7\}$ with a dilation factor of 2 and translation factor of 1. Note that the suitability of B-spline dictionaries for use with sparse problems has been established previously in [16].

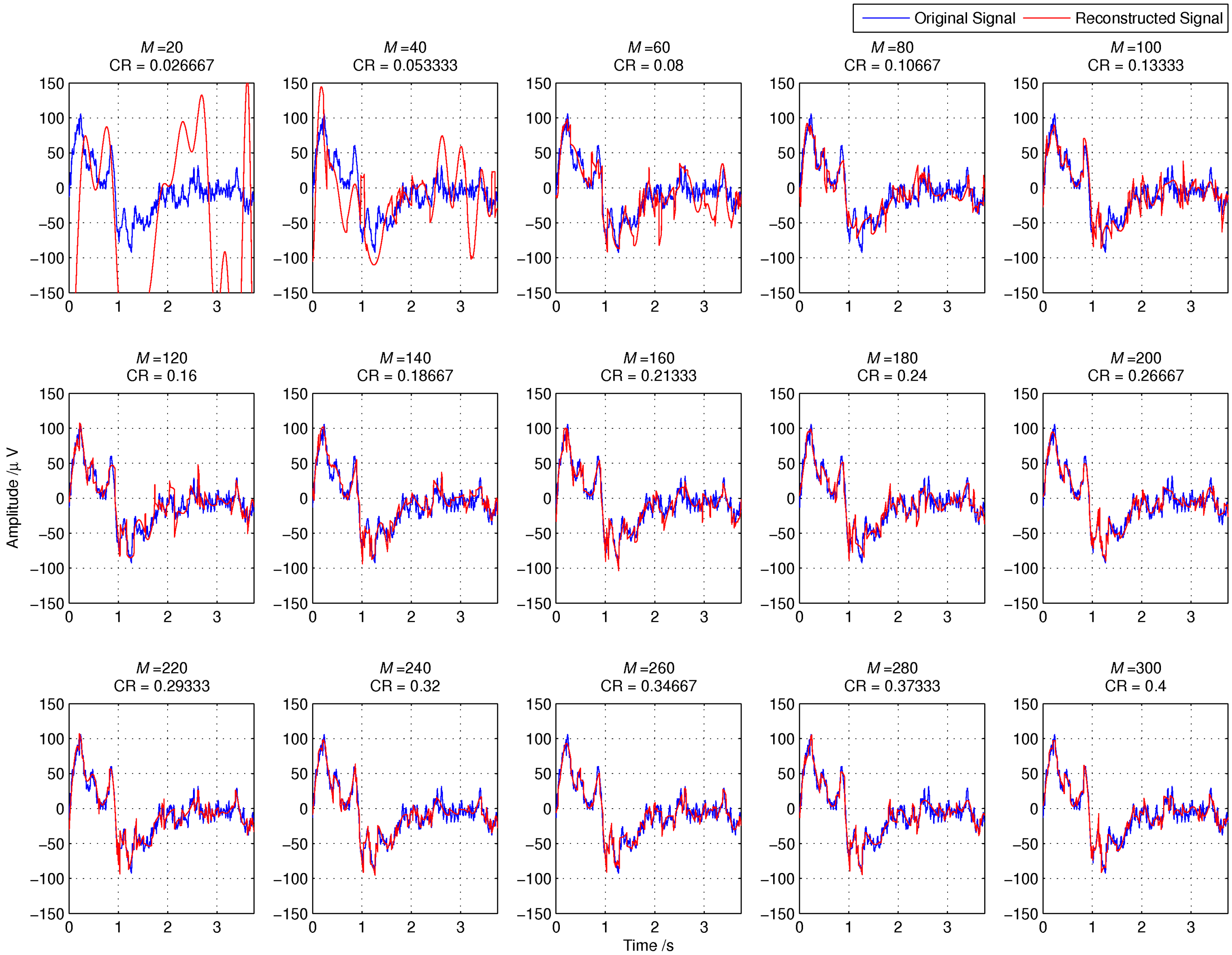

Fig. 1. Qualitative illustration of the reconstruction performance of compressive sensing applied to scalp EEG signals at a range of different Compression Ratios (CR). The EEG section is from channel F7 and is selected as a random background section rather attempting to be representative of the entire EEG.

B. Reconstruction methods

Three methods for carrying out the l_1 optimisation in (3) are used in this work. Basis Pursuit (BP) [17], Matching Pursuit (MP) [18] and Orthogonal Matching Pursuit (OMP) [19] are investigated. These are commonly used numerical techniques, each achieving different performance in terms of computational complexity and reconstruction accuracy.

C. Measurement matrix

In order to keep the number of results generated and their presentation practical only one choice for the measurement matrix $\mathbf{\Phi}$ is used here. This is selected as a Gaussian random matrix and is generated using the MATLAB `randn` function. The same matrix was used in [8] and it is generally a popular choice to ensure incoherence. The impact of other choices for $\mathbf{\Phi}$ on the compressive sensing performance is left to future work.

D. Analysis methods

Quantitative testing of the compressive sensing performance is carried out using a set of scalp EEG data provided by the National Society for Epilepsy in the UK. One hour recordings from three subjects are used with each recording having 19 referential channels (giving a total of 57 hours of EEG data). All data uses an FCz reference and a 200 Hz sampling frequency. The channels present are: F7, F8, F3, F4, Fz, C3, C4, Cz, Fp1, Fp2, T3, T4, T5, T6, P3, P4, Pz, O1, O2. For analysis each channel is broken down into non-overlapping frames of 750 samples ($N = 750$) which are compressed and reconstructed separately.[1] The reconstructed frames are then concatenated and performance metrics derived by averaging the performance across channels.

A total of six performance metrics are presented here, simply because there is no uniformity in the literature as to the metrics used to quantify other compression techniques, and so the aim is to provide the reader with comprehensive information about compressive sensing to allow comparison. The performance metrics used include the conventional Signal-to-Noise Ratio (SNR), Peak Signal-to-Noise Ratio (PSNR), Root

[1]This frame size matches that used in [8] to allow direct comparison of results.

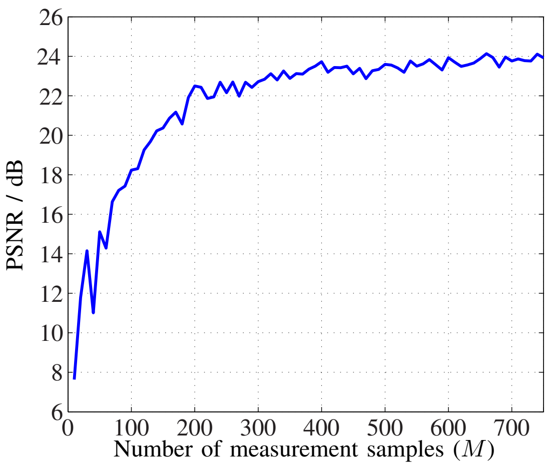

Fig. 2. Illustration of how reconstruction accuracy assessed via the PSNR varies with the compression ratio for a background EEG section. As long as 250 or more measurement samples are taken the reconstruction accuracy is somewhat constant.

Mean Square (RMS), Percent of Root-mean-square Difference (PRD), and Cross-Correlation (CC). The time required for the reconstruction of a 750 sample frame (corresponding to 3.75 s of data) using a Quad core Xeon processor with 4 GB of RAM is also presented. Note that this last metric is not intended as an absolute measure, but as a factor for comparison between the complexity of the different reconstruction methods.

IV. Results

A. *Qualitative performance*

Fig. 1 illustrates the typical reconstruction performance of a single 750 sample frame of scalp EEG. This is based upon the use of the Gabor dictionary, a Gaussian random measurement matrix and the OMP reconstruction method as the number of measurement samples (M) is varied. This is equivalent to changing the Compression Ratio (CR): CR $= M/N \times 100\%$; where lower compression ratios represent better performance.

From Fig. 1 it can be seen how reconstruction of the signal is possible, and how the quality of this reconstruction improves as more measurement samples are taken. This is a key trade-off for anybody interested in implementing a compressive sensing scheme. Fig. 2 illustrates this, showing how the PSNR between the original and reconstructed signals varies with the number of measurement samples (M) taken. It can be seen that as long as M is greater than around a third of the total number of samples, which in this case is 250, the reconstruction PSNR is somewhat constant. Little improvement in reconstruction accuracy is then achieved for increasing the compression ratio. From these results it is clear that compressive sensing can be successfully applied to scalp EEG signals. However, the impact of the dictionary and reconstruction method has yet to be evaluated.

B. *Quantitative performance*

Table I presents detailed results for the reconstruction performance of the 18 different compressive sensing implementations (six dictionaries each used with three reconstruction methods) used here. To keep Table I practical, only one compression ratio is considered. This is selected as $M = 300$, giving a compression ratio of 40% as no substantial gain in performance was witnessed with higher values of M in Fig. 2. It is assumed that this provides results representative of other compression ratios.

The first conclusion that can be extracted from Table I is that, although potentially interesting, system designers should be aware of the limitations of compressive sensing theory when applied to EEG signals. For example, at the same compression ratio the reconstruction accuracy can vary significantly depending on the settings used with the compressive sensing. It is clear that if reconstruction accuracy is the most important consideration the Basis Pursuit reconstruction method works considerably better than either Matching Pursuit or Orthogonal Matching Pursuit. However, this comes at the cost of computational complexity and reconstruction time. Thus Basis Pursuit implementations may not be suitable if real or quasi-real time reconstruction implementations are aimed for. If time and complexity are issues and the reconstruction error can somehow be compromised, the Orthogonal Matching Pursuit method offers a better option. Also, it is apparent that B-spline dictionaries are particularly suitable for use with EEG signals. Independent of the reconstruction method used they lead to the lowest reconstruction errors for a similar level of complexity.

V. Conclusions

This paper has characterised the performance of compressive sensing theory when applied to scalp EEG signals. The characterisation has been done by taking a total of 57 hours of EEG and quantifying the errors after signal reconstruction in terms of the CC, SNR, PSNR, RMS, PRD and reconstruction time. This has been done for 18 different implementations of the theory using six dictionaries and three reconstruction methods. We have thus presented performance results that can aid the EEG system designer to decide whether the technique is worth using or not, and if so, which one of the different implementations to opt for given the particular application aims and constraints.

The results show that at present compressive sensing, applied to a single EEG channel at a time, has limited applicability as a compression technique for EEG signals, depending mostly on the application requirements and more specifically on the reconstruction error that is acceptable. This acceptable error may vary significantly depending on the specific application and use of the EEG system. Overall, Basis Pursuit as a reconstruction technique works considerably better than Matching Pursuit or Orthogonal Matching Pursuit, but this comes at the expense of increased computational complexity. Similarly B-spline dictionaries are the most promising in terms of reconstruction error. However, again, there are other factors to take into account before considering a certain kind of function for a practical system design. In particular, the power requirements of the specific chosen hardware platform and whether the chosen dictionary functions are realisable in analogue or digital hardware, and continuous or discrete time,

TABLE I

DETAILED PERFORMANCE OF 18 DIFFERENT VERSIONS OF COMPRESSIVE SENSING THEORY APPLIED TO 57 HOURS OF SCALP EEG DATA.

Dictionary	CC	SNR / dB	PSNR / dB	RMS / μV	PRD / %	Reconstruction time / s
Basis Pursuit (BP) reconstruction method						
Gabor	0.97	13.49	50.42	9.03	23.14	4.87
Mexican hat	0.97	12.51	49.48	10.29	25.11	5.20
Linear Spline	0.97	13.35	50.28	9.04	23.04	3.34
Cubic Spline	0.97	12.70	49.68	9.71	24.96	3.42
Linear B-Spline	0.98	14.59	51.43	7.91	20.39	3.25
Cubic B-Spline	0.98	15.28	52.11	7.38	18.61	3.25
Matching Pursuit (MP) reconstruction method						
Gabor	0.85	4.84	42.08	26.47	52.67	1.13
Mexican hat	0.62	-3.71	34.89	75.91	184.77	1.13
Linear Spline	0.95	11.46	48.38	11.10	28.68	1.19
Cubic Spline	MP method failed to reconstruct for this dictionary					
Linear B-Spline	0.96	12.54	49.43	9.87	25.29	1.17
Cubic B-Spline	0.95	11.25	48.44	10.96	28.71	1.17
Orthogonal Matching Pursuit (OMP) reconstruction method						
Gabor	0.94	9.82	46.66	13.51	34.84	2.00
Mexican hat	0.94	9.98	46.93	13.33	34.15	1.78
Linear Spline	0.95	11.24	48.05	11.41	29.60	0.93
Cubic Spline	0.94	10.54	47.42	12.37	31.96	0.94
Linear B-Spline	0.96	11.71	48.52	10.93	27.91	0.89
Cubic B-Spline	0.96	12.17	49.03	10.39	26.47	0.94

will be key. Ultimately, opting for compressive sensing as a data reduction technique for EEG signals will be beneficial depending on the overall system design trade-offs. The results presented here have quantified the compressive sensing trade-offs for a set of 18 different compressive sensing arrangements.

Finally, whilst outside the scope of this work it is necessary to note that other compressive sensing implementations are possible, not least through other dictionaries and measurement matrices, and so the results here are not exhaustive. Also, it is known that EEG signals have high inter-channel correlation, or in other words can be jointly sparse. Potentially this joint sparsity could be exploited to improve the reconstruction performance. This is especially relevant for EEG systems which are customised for medical use where variations in the degree of inter-channel correlation can also be related to the specific nature of the disease being investigated. The work presented here provides an analysis framework and quantification of baseline performance essential for establishing the utility of any such future compressive sensing implementations.

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Research Council under the European Community's 7$^{\text{th}}$ Framework Programme (FP7/2007-2013) / ERC grant agreement no. 239749.

REFERENCES

[1] A. J. Casson, D. C. Yates, S. J. Smith, J. S. Duncan, and E. Rodriguez-Villegas, "Wearable electroencephalography," *IEEE Eng. Med. Biol. Mag.*, vol. 29, no. 3, pp. 44–56, 2010.

[2] Y. Wongsawat, S. Oraintara, T. Tanaka, and K. R. Rao, "Lossless multi-channel EEG compression," in *IEEE ISCAS*, Kos, May 2006.

[3] J. Cardenas-Barrera, J. Lorenzo-Ginori, and E. Rodriguez-Valdivia, "A wavelet-packets based algorithm for EEG signal compression," *Med. Informatic. and Internet in Med.*, vol. 29, no. 1, pp. 15–27, 2004.

[4] A. M. Abdulghani, A. J. Casson, and E. Rodriguez-Villegas, "Quantifying the feasibility of compressive sensing in portable electroencephalography systems," in *HCI international*, San Diego, July 2009.

[5] U. Gamper, P. Boesiger, and S. Kozerke, "Compressed sensing in dynamic MRI," *Magn. Reson. Med.*, vol. 59, no. 2, pp. 365–373, 2008.

[6] T. V. Sreenivas and W. B. Kleijn, "Compressive sensing for sparsely excited speech signals," in *IEEE ICASSP*, Taipei, April 2009.

[7] Y. Zhang, S. Mei, Q. Chen, and Z. Chen, "A novel image/video coding method based on compressive sensing theory," in *IEEE ICASSP*, Las Vagas, April 2008.

[8] S. Aviyente, "Compressed sensing framework for EEG compression," in *IEEE/SP SSP*, Madison, August 2007.

[9] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[10] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 21–30, 2008.

[11] E. J. Candes, "Compressive sampling," in *Proc. Int. Congr. Math.*, Madrid, August 2006.

[12] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[13] C. Sieluycki, R. Konig, A. Matysiak, R. Kus, D. Ircha, and P. J. Durka, "Single-trial evoked brain responses modeled by multivariate matching pursuit," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 1, pp. 74–82, 2009.

[14] M. Andrle and L. Rebollo-Neira, "Spline wavelet dictionaries for non-linear signal approximation," in *Proc. Int. Conf. Interactions between Wavelets and Splines*, Athens, May 2005.

[15] ——, "Cardinal B-spline dictionaries on a compact interval," *Appl. Comput. Harmon. Anal.*, vol. 18, no. 3, pp. 336–346, 2005.

[16] ——, "From cardinal spline wavelet bases to highly coherent dictionaries," *J. Phys. A: Math. Theory*, vol. 41, no. 17, p. 172001, 2008.

[17] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001.

[18] S. Mallat and Z. Zhang, "Matching pursuit in a time frequency dictionary," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[19] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *ACSSC*, Pacific Grove, November 1993.