Graph Theoretic Compressive Sensing Approach for Classification of Global Neurophysiological States from Electroencephalography (EEG) Signals

Mohammad Samie Tootooni[1], Miaolin Fan[3], Rajesh Sharma Sivasubramony[1], Chun-An Chou[3], Vladimir Miskovic[2], and Prahalada K. Rao[4(✉)]

[1] Systems Science and Industrial Engineering Department, State University of New York, Binghamton, NY, USA
[2] Psychology Department, State University of New York, Binghamton, NY, USA
[3] Mechanical and Industrial Engineering Department, Northeastern University, Boston, MA, USA
[4] Mechanical and Materials Engineering Department, University of Nebraska-Lincoln, Lincoln, NE, USA
Rao@unl.edu

Abstract. We present a data fusion framework integrating graph theoretic and compressive sensing (CS) techniques to detect global neurophysiological states using high-resolution electroencephalography (EEG) recordings. Acute stress induction (and control procedures) were used to elicit distinct states of neurophysiological arousal. We recorded EEG signals (128 channels) from 50 participants under two different states: hand immersion in room temperature water (control condition) or in chilled ($\sim$3 °C) water (stress condition). Thereafter, spectral graph theoretic Laplacian eigenvalues were extracted from these high-resolution EEG signals. Subsequently, the CS technique was applied for the classification of acute stress using the Laplacian eigenvalues as features. The proposed method was compared to a support vector machine (SVM) approach using conventional statistical features as inputs. Our results revealed that the proposed graph theoretic compressive sensing approach yielded better classification performance ($\sim$90 % F-score) compared to SVM with statistical features ($\sim$50 % F-Score). This finding indicates that the spectral graph theoretic compressive sensing approach presented in this work is capable of classifying global neurophysiological arousal with higher fidelity than conventional signal processing techniques.

Keywords: Graph theory · Compressive sensing · Laplacian eigenvalues · Electroencephalography · Stress · Classification

1 Introduction

Electroencephalography (EEG) is a neurophysiological method for non-invasively monitoring the large-scale electrical activity of the human brain. The objective of our work was to classify the global neurophysiological state of human subjects from

G.A. Ascoli et al. (Eds.): BIH 2016, LNAI 9919, pp. 42–51, 2016.
DOI: 10.1007/978-3-319-47103-7_5

multichannel EEG signals using a novel spectral graph theoretic compressive sensing approach.

Specifically, we wished to use the raw, time domain EEG signals in order to discriminate brain electrical signals collected during either an acute stress induction period or an appropriate, non-stressful comparison condition. Detecting stress using non-invasive physiological sensing can serve as a critical indicator of the onset of fatigue in mission-critical human activity, such as hazardous cargo trucking, air traffic control and railroad operation, to name a few examples. Thus, being able to detect stress from recordings of ongoing brain activity would have clear real-world applications.

In this context, the real-time monitoring of high-resolution EEG signals with compressive sensing (CS) has attracted considerable attention in recent years. CS is an ℓ_1-norm regularization-based signal compression and reconstruction approach that provides a sparse representation of the information in the original signal or image. Previous studies have shown the practical value of CS in EEG monitoring or brain computer interface systems for addressing problems, such as signal reconstruction and power consumption [1–3]. CS yields a more efficient representation, particularly with multi-channel EEG systems, of the original signal with a relatively smaller number of projected components for information reconstruction compared to signal reconstruction techniques. This feature of CS allows for a lower sampling rate than the Nyquist rate without losing information in the original signals.

In previous literature, CS was applied as a signal reconstruction technique to multichannel EEG signals based on various dictionaries, e.g. Gabor frame [1] or Slepian basis function [2]. Classification algorithms, i.e. Block sparse Bayesian learning, were performed on the reconstructed signals and showed that the CS-based compression was power-saving and effective compared to conventional transformation approaches, such as wavelets [4]. A review of CS applications to bioelectric signal processing is presented in [5].

The complexity of bio-sensor data arises from nonstationarity in the time domain [6], nonlinearity and quasi-periodicity in state-space [7], and intermittency [8]. Furthermore, the low signal to noise ratio (S/N), autocorrelation within and cross-correlation between sensor data, and the interactions across multiple neurological conditions [9] are other factors that impede the use of conventional statistical features for analysis of bio-sensor [10, 11].

Graph theory is an approach whereby multi-dimensional signals can be fused. The Laplacian eigenvalues of a signal represented in graph space is used as input features for classification. We applied CS using these features as representations of high-resolution, continuous EEG signals to classify the signal patterns recorded during acute stress versus those recorded during an appropriate control condition. Figure 1 depicts a five second segment of the EEG time series (from a single electrode) recorded from two human subjects serving in different experimental conditions.

The rest of this paper is organized as follows; Sect. 2 explains the graph representation of EEG signal and the Laplacian eigen-spectra extraction. It also describes the acute stress experiments. Section 3 presents the results and compares it with both graph-based SVM and conventional methods using statistical features. Finally, Sect. 4 summarizes the conclusions and suggests avenues for future research.

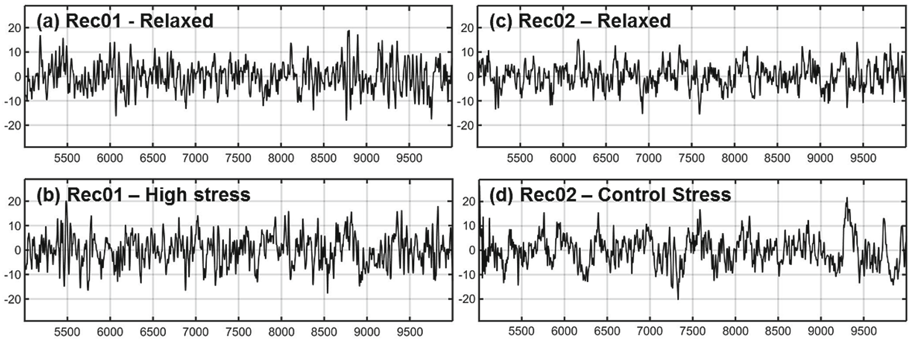

Fig. 1. Five second EEG fragment (single electrode) is depicted (a) when the first subject is relaxed; and (b) for the same subject during acute stress induction. Similarly, (c) and (d) depict EEG signals recorded from a different experimental subject, while relaxing, and during a control condition (hand immersion in room temperature water), respectively.

2 Methodology

In this section, we present our proposed approach which has two phases. First, the multi-channel EEG sensors will be fused using spectral graph theory and Laplacian eigenvalues will be extracted from the fused signals. In the second phase, the neuro-physiological states will be classified using CS with the Laplacian eigenvalues as inputs. The proposed methodology is schematically depicted in Fig. 2.

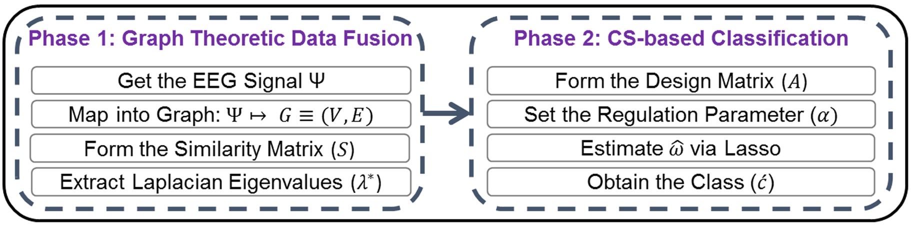

Fig. 2. The proposed graph theoretic compressive sensing classification approach

2.1 Phase 1 – Data Fusion with Graph Representation of EEG Signals

Let us consider matrix $\Psi \in \mathbb{R}^{q \times d}$ as a recorded EEG signal in which q is the length of the signal and d is number of EEG channels (in our practical case, $d = 128$ sensors and $q = 500$ data points (~ 1 s). See Sects. 2.3 and 2.4).

$$\Psi = \begin{bmatrix} \psi_1^1 & \cdots & \psi_1^d \\ \vdots & \ddots & \vdots \\ \psi_q^1 & \cdots & \psi_q^d \end{bmatrix}$$

In this signal *window*, each row represents voltage fluctuation recorded by all d sensors at a time instant. It is assumed that all sensors have equal recording rate (sampling rate). Choosing q (window length) is a heuristic choice; it should not be either too large, since it increases the computational time, or too short, else the window is not representative of the whole signal.

A kernel function (Ω) is chosen to capture the distance between each pair of rows $\psi_i, \psi_j \in \mathbb{R}^{1\times d}$ of the matrix Ψ. In this paper, Gaussian kernel function is utilized to get the pairwise comparison matrix Υ (Eqs. (1) and (2)); where σ^2 is the total variance of the pairwise Euclidean distance matrix. A threshold function (Θ) is then applied on Υ (Eq. (3)). This threshold is set as the average of all element of matrix Υ. Rao *et al.* have discussed on setting the threshold value (r) [12]. A similarity matrix (S) is then acquired (Eq. (4)) to represent the corresponding unweighted and undirected network graph for matrix Ψ.

$$\mathcal{w}_{ij} = \Omega\left(\psi_i, \psi_j\right) = e^{-\left(\frac{\|\psi_i - \psi_j\|^2}{\sigma^2}\right)} \quad \forall\, i, j \in (1 \cdots q) \tag{1}$$

$$\Upsilon^{q\times q} = \left[\mathcal{w}_{ij}\right] \tag{2}$$

$$\Theta\left(\mathcal{w}_{ij}\right) = w_{ij} = \begin{cases} 1, & \mathcal{w}_{ij} \le r \\ 0. & \mathcal{w}_{ij} > r \end{cases} \tag{3}$$

$$S^{q\times q} = \left[w_{ij}\right] \tag{4}$$

The degree vector (deg_i) is formed then by row-wise summation of w_{ij} as shown in Eq. (5) and by Eq. (6), it transforms into a diagonal matrix called *Degree Matrix* (D). Finally Eqs. (7) and (8) denote the formation of Laplacian matrix (L) and the normalized Laplacian matrix ($\mathcal{L}$), respectively.

$$\deg_i = \sum_{j=1}^{j=q} w_{ij} \, \forall\, i, j \in (1 \cdots q) \tag{5}$$

$$D^{q\times q} = [d_{ij}] = \begin{cases} \deg_i. & i = j \\ 0. & i \neq j \end{cases} \tag{6}$$

$$L^{q\times q} = D - S \tag{7}$$

$$\mathcal{L}^{q\times \mathrm{q}} = D^{-\frac{1}{2}} \times L \times D^{-\frac{1}{2}} \tag{8}$$

$$\mathcal{L}v = \lambda^* v \tag{9}$$

In Eq. (9), $v \in \mathbb{R}^{q\times q}$ are the Laplacian eigenvectors; the Laplacian eigenvalues are indicated as $\lambda^* \in \mathbb{R}^{q\times q}$.

2.2 Phase 2 – Laplacian Eigen Compressive Sensing Classification

The aim of this phase is to classify the neurophysiological state of the subject using compressive sensing (CS) based on Laplacian eigenvalues λ^* from phase 1 (Eq. (9)) as input features. We note that CS is a supervised learning technique, in other words, we will define *a priori* classes from offline sensor signals.

In this context, consider $A\omega = y$ to be an underdetermined system of equation with N unknowns and m equations. Matrix A is referred as the training matrix which consists of Laplacian eigenvalues obtained from known-state EEG signal for each class. Where class refers to the neurophysiological state of the subject; in this case either relaxed vs. stressed. For instance, for a C-class classification problem, this matrix is designed as $A = [A_1, A_2, \ldots, A_C]$.

Zhan *et al.* [13] showed the first and last few eigenvalues have the highest variability among all Laplacian eigenvalues by analyzing their *relative deviation*. Accordingly in our paper, the first [Starting from the second eigenvalue since the first Laplacian eigenvalue is always zero $(\lambda_1 = 0)$] $m/2$ and the last $m/2$ of eigenvalues are chosen. We denote $\Lambda \in \mathbb{R}^{m\times 1}$ as the chosen eigenvalue vector for each window.

A *sample* vector $\bar{\Lambda} \in \mathbb{R}^{m\times 1}$ is then defined as average of k randomly chosen eigenvalue vectors (Λ) from each class. Although this averaging reduces the number of available samples for each class, it helps to increase the reliability of the training matrix (A). Suppose we use n sample vector to train the classification algorithm in each class. Therefore, a sample vectors is denoted by $\bar{\Lambda}_{j,c}$ where $c \in \{1, \ldots, C\}$ and $j \in \{1, \ldots, n\}$ are the class and sample indices, respectively. Ergo, the training matrix ($A \in \mathbb{R}^{m\times N}$) is designed as Eq. (10) where $N = n * C$.

$$A = \left[\left[\bar{\Lambda}_{1,1} \quad \ldots \quad \bar{\Lambda}_{n,1} \right] \left[\bar{\Lambda}_{1,2} \quad \ldots \quad \bar{\Lambda}_{n,2} \right] \ldots \left[\bar{\Lambda}_{1,C} \quad \ldots \quad \bar{\Lambda}_{n,C} \right] \right] \tag{10}$$

Also, measurement vector $Y \in \mathbb{R}^{m\times 1}$ represents the testing (new arrived information) set. This set is basically the average Laplacian eigenvalue vector ($\bar{\Lambda}$) extracted from the incoming EEG signal. Our aim is to find out unknown vector $\omega \in \mathbb{R}^{N\times 1}$ using compressive sensing to solve the linear system of equations mentioned $A\omega = y$ and eventually, to determine the class label of the incoming signal. Thereafter, an ℓ_0-minimization problem should be formulated as Eq. (11). Equation (12) replaces it with its corresponding ℓ_1-minimization problem [14]. To approximate a sparse solution, LASSO (Least Absolute Shrinkage and Selection Operator) algorithm is applied as shown in Eq. (13); where α is the regulation parameter for the LASSO algorithm. These concepts are clarified in detail in [15–18].

$$\textit{minimize}\ \|\omega\|_0 \quad \textit{subject to}\ A\omega = y \tag{11}$$

$$\textit{minimize}\ \|\omega\|_1 \quad \textit{subject to}\ A\omega = y \tag{12}$$

$$\hat{\omega} = \underset{\omega}{\operatorname{argmin}}\ \alpha\|\omega\|_1 + \|Y - A\omega\|_2 \tag{13}$$

$$\acute{c} = \underset{c}{\operatorname{argmin}} A\delta_c(\hat{\omega}) - Y \tag{14}$$

vector $\hat{\omega} = [\hat{\omega}_1, \hat{\omega}_2, \ldots, \hat{\omega}_C]^T$ is obtained from Eq. (13) and eventually, Eq. (14) determines the class index of the new-arriving data where $\delta_c(\hat{\omega}) = [0^T, \ldots, \hat{\omega}_c^T, \ldots, 0^T]^T$.

2.3 Data Acquisition and Processing

We used a 128-channel EEG sensor network by Electrical Geodesics, Inc. (EGI), (HydroCel Geodesic Sensor Net) to collect the resting-state EEG data with subjects keeping their eyes open in two experiments with a sampling rate of 1 kHz and Net Station 4.5.6 software. In the first experiment (dataset 1), the resting-state EEG was recorded respectively from a *stress* condition of two male subjects who were instructed to place their hand into ice water (0–3 °C) and a pre-stimulus phase (*relaxed* condition). The length of EEG recording for each state was 2 min. In the second experiment (Dataset 2), we extended the study to 49 participants who were randomly assigned to either an acute stress or a comparison condition, where the ice water was replaced by lukewarm water. However, the subjects were not informed beforehand about which treatment they were assigned. Some participants in acute stress conditions do not have full length (2 min) recordings since they were unable to maintain their hand in the cold water. Furthermore, 1-min EEG recordings from *relaxed* condition was also collected.

The EEG recordings were down-sampled to 500 Hz for reducing the computational cost in the data analysis. After removing the facial sensors, a spatial principal component analysis (sPCA) was applied for the artifact correction, with 98 % of total variance explained. Furthermore, a reduced-rank independent component analysis (ICA) was performed to extract the same number of components as in sPCA. Finally, a binary classification on recorded EEG signals is performed to detect whether the participant is under stress in a within-participant manner.

2.4 Classification and Verification

When applying the Laplacian CS classification algorithm to the two datasets, the window size is chosen to be 1 s ($q = 500$), the number of features are $m = 20$, and the sample size is set to be $k = 5$ for the first dataset and $k = 3$ for the second dataset[1]. Among all the samples, we randomly allocate 60 % to training set which forms the design matrix (A). 30 % is randomly specified to validation set, which is used to find obtain the LASSO regularization parameter. An enumerative heuristic approach is applied to find a value which minimizes the overall classification error of the validation set. Finally, we use the remaining 10 % to evaluate the classification performance. Beside the proposed algorithm, to verify the capability of compressive sensing

[1] Due to lower length of available recorded signal for *Relaxed* class.

approach, the Laplacian eigen-spectra extracted from EEG signals are used in a Support Vector Machine with linear kernel function (LSVM).

To further verify the applicability of graph theoretic Laplacian eigenvalues in classification of EEG signals, we applied the CS and the SVM with the conventional statistical features, including 6 main chosen features: mean, median, standard deviation, kurtosis, skewness, and interquartile range of each signal *window*. However, having several channels in recorded EEG signals result in a large number of features (in our practical case $d = 128$, which results in $128 * 6 = 768$ features for each window) which intensely increases the computational time. Therefore, to make it comparable to other algorithms, we applied the ICA method on the acquired statistical features and chose the first m independent components.

Moreover, to avoid bias due to random partition of training and test sets in classification, each Laplacian-based algorithm is run 20 times; and each statistical feature-based algorithm is run 10 times.

3 Results

In this section, we present the results of applying the proposed binary classification algorithm. To assess classification performance, we use a confusion matrix with F-score as the evaluation criterion to compare classification performances of selected algorithms $(F - score = 2 * (Precision.Recall)/(Precision + Recall))$ [19]. It should be noted that if either one of *precision* and *recall* does not exist, consequently, the F-score cannot be calculated which is shown as *NaN*.

3.1 Dataset 1

Table 1 shows the result confusion matrix for all discussed classification approaches. As shown in the table, the proposed graph theoretic CS approach has significantly higher F-score than graph theoretic SVM. Besides that, both Graph theoretic based approaches dominate the approaches based on conventional statistics. This result

Table 1. Confusion matrices for classification of the EEG signals. All numbers are reported as percentage. The two classes are *Relaxed* (Rel.) and *Stressed* (Str.) conditions.

Confusion matrix		Classifier	Laplacian eigenvalues			Conventional statistics		
			Predicted		Recall	Predicted		Recall
			Rel.	Str.		Rel.	Str.	
Actual	Rel.	CS (Proposed)	85.0	15.0	**85.0**	73.3	26.7	**73.3**
	Str.		5.0	95.0	**95.0**	70.0	30.0	**30.0**
Precision			**94.4**	**86.4**	***F* = 90.20**	51.2	52.9	***F* = 51.86**
Actual	Rel.	SVM	90.0	10.0	**90.0**	80.0	20.0	**80.0**
	Str.		33.3	66.7	**66.7**	80.0	20.0	**20.0**
Precision			**73.0**	**87.0**	***F* = 79.14**	**50.0**	**50.0**	***F* = 50.00**

indicates that the proposed graph theoretic compressive sensing approach has higher fidelity compare to other conventional methods.

3.2 Dataset 2

In this dataset, there were many runs that F-score was not acquirable (the precision did not exist). Therefore, we introduced a metric, *Number of Success,* as number of the Runs that the F-score is estimable for each subject. This metric provides an appropriate criterion for evaluating the algorithms' performance. Indeed before comparing the average F-score, the *Success Rate* should be considered to compare the feasibility of the classification algorithms $(Success\,Rate = (Number\ of\ Success)/(Number\ of\ Runs))$.

Entirely 49 subjects participated in the second experiment, 17 of which did not have enough recorded signal to be considered in the analysis. Therefore, in this dataset there are 20 participants treated with warm water and 12 participants underwent cold water immersion. Figure 3 shows the performance (F-score and *Number of Successes*) of the graph theoretic features (Fig. 3(a) and (b)) as well as the statistical features (Fig. 3(c) and (d)). In this figure, the line charts represent the *number of success* and the bars show the average of available F-scores for each classification technique. As Fig. 3 (c) and (d) show, SVM with statistical features (SVM-ST) were unable to classify the state of the incoming signal almost in all runs; and the CS classifier with statistical (CS-ST) features had poor classification result as well as low success rate compare with CS based on graph theoretic features (CS-GT). This shows that using graph theoretic features for signal classification purposes is preferred over conventional statistics. We refer to the complex structure of EEG data, discussed in the Sect. 1, as one reason to make the statistical feature-based algorithms unable to capture the dynamics.

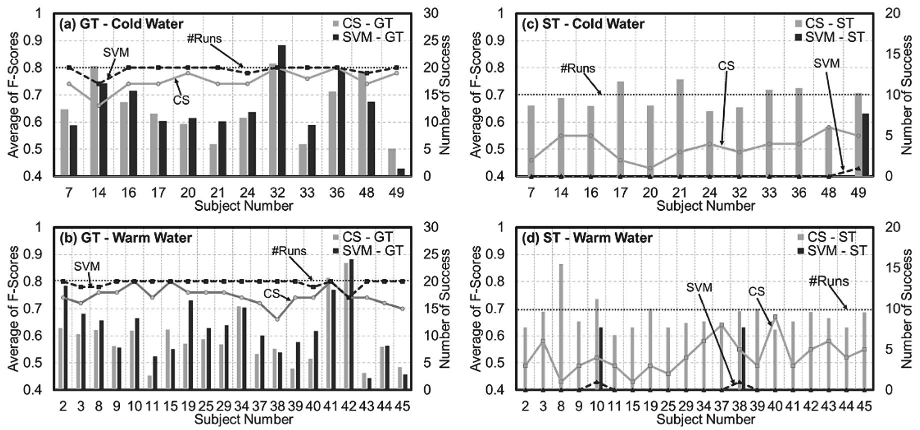

Fig. 3. Performance of the classifiers to detect stress for different participant. Bars show the average F-score in primary vertical axes; and lines represent the number of success in the secondary vertical axes. CS and SVM stand for Compressive Sensing and Support Vector Machine classifiers, respectively. Also GT and ST are representors of features: Graph Theoretic and Statistical features, correspondingly.

Also, Fig. 3(a) and (b) show the performance of the two graph theoretic classifiers for each participant, separately. The CS has a relatively lower F-score in warm water classification (Fig. 3(b)) compared to the cold water classification (acute stress, Fig. 3 (a)). This means that the proposed CS-GT was able to distinguish the acute stress easier than the control stress from the *relaxed* condition. Although the SVM has higher success rate in both groups of participants, the GT-CS has comparable F-score and success rate for detection of acute stress states (Fig. 3(a)). However, the SVM is highly sensitive to the size of the training and testing sets, and its performance is dependent on choosing the right kernel function and tuning parameters [20, 21]. In contrast, there is only LASSO regularization parameter to be set in the proposed CS-GT algorithm. It can thus be used to classify the neurophysiological signals in real-time with low computational load. Nonetheless, both CS and SVM with graph theoretic quantifiers outperformed the statistical features based approaches.

4 Conclusion

In this paper we applied a graph-based data fusion compressive sensing approach for high-dimensional signal classification. Two continuous EEG datasets we collected in an acute stress experiment to test the proposed graph-based compressive sensing (CS) approach. The validation procedure has two stages; first with graph-based SVM, and then, with other conventional-based methods. It was found that graph-based classifiers features were able to demarcate distinct states of neurophysiological arousal with higher fidelity compared to conventional statistical methods. The authors suggest two avenues for future research; using graph-theoretic features in other multi-class classification algorithms, and applying these features for prediction of high-stress conditions.

Acknowledgements. This research is made possible due to funding by the National Science Foundation under the Service, Manufacturing, and Operation Research (SMOR) program (grant number CMMI – 1538059), and the Transdisciplinary Area of Excellence (TAE) exploratory research grant by Binghamton University.

References

1. Aviyente, S.: Compressed sensing framework for EEG compression (2007)
2. Senay, S., Chaparro, L.F., Sun, M., Sclabassi, R.J.: Compressive sensing and random filtering of EEG signals using Slepian basis. In: 16th European 2008 Signal Processing Conference, pp. 1–5. IEEE (2008)
3. Abdulghani, A.M., Casson, A.J., Rodriguez-Villegas, E.: Compressive sensing scalp EEG signals: implementations and practical performance. Med. Biol. Eng. Comput. **50**(11), 1137–1145 (2012)
4. Liu, B., Zhang, Z., Xu, G., Fan, H., Fu, Q.: Energy efficient telemonitoring of physiological signals via compressed sensing: a fast algorithm and power consumption evaluation. Biomed. Sig. Process. Control **11**, 80–88 (2014)

5. Craven, D., McGinley, B., Kilmartin, L., Glavin, M., Jones, E.: Compressed sensing for bioelectric signals: a review. IEEE J. Biomed. Health Inform. **19**(2), 529–540 (2015)
6. Kaplan, A.Y., Fingelkurts, A.A., Fingelkurts, A.A., Borisov, S.V., Darkhovsky, B.S.: Nonstationary nature of the brain activity as revealed by EEG/MEG: methodological, practical and conceptual challenges. Sig. Process. **85**(11), 2190–2212 (2005)
7. Pereda, E., Quiroga, R.Q., Bhattacharya, J.: Nonlinear multivariate analysis of neurophysiological signals. Progress Neurobiol. **77**(1), 1–37 (2005)
8. Siegel, M., Donner, T.H., Engel, A.K.: Spectral fingerprints of large-scale neuronal interactions. Nat. Rev. Neurosci. **13**(2), 121–134 (2012)
9. Hall, D.L., Llinas, J.: An introduction to multisensor data fusion. Proc. IEEE **85**(1), 6–23 (1997)
10. Stam, C.J.: Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. Clin. Neurophysiol. **116**(10), 2266–2301 (2005)
11. Stam, C.J., Breakspear, M., van Walsum, A.M.V.C., van Dijk, B.W.: Nonlinear synchronization in EEG and whole-head MEG recordings of healthy subjects. Hum. Brain Mapp. **19**(2), 63–78 (2003)
12. Rao, P.K., Kong, Z., Duty, C.E., Smith, R.J., Kunc, V., Love, L.J.: Assessment of dimensional integrity and spatial defect localization in additive manufacturing using spectral graph theory. J. Manuf. Sci. Eng. **138**(5), 051007 (2016)
13. Zhan, C., Chen, G., Yeung, L.F.: On the distributions of Laplacian eigenvalues versus node degrees in complex networks. Phys. A: Stat. Mech. Appl. **389**(8), 1779–1788 (2010)
14. Chen, S.S., Donoho, D.L., Saunders, M.A.: Atomic decomposition by basis pursuit. SIAM Rev. **43**(1), 129–159 (2001)
15. Boche, H., Calderbank, R., Kutyniok, G., Vybíral, J.: Compressed Sensing and its Applications. Springer, Berlin (2015)
16. Foucart, S., Rauhut, H.: A mathematical introduction to compressive sensing, vol. 1, 3. Springer, Berlin (2013)
17. Qaisar, S., Bilal, R.M., Iqbal, W., Naureen, M., Lee, S.: Compressive sensing: from theory to applications, a survey. J. Commun. Netw. **15**(5), 443–456 (2013)
18. Bastani, K., Rao, P.K., Kong, Z.: An online sparse estimation-based classification approach for real-time monitoring in advanced manufacturing processes from heterogeneous sensor data. IIE Trans. **48**(7), 579–598 (2016)
19. Van Rijsbergen, C.: Information retrieval. Department of Computer Science, University of Glasgow (1979). citeseer.ist.psu.edu/vanrijsbergen79information.html
20. Suykens, J.A.: Advances in Learning Theory: Methods, Models, and Applications, vol. 190. IOS Press, Amsterdam (2003)
21. Burges, C.J.: A tutorial on support vector machines for pattern recognition. Data Min. Knowl. Disc. **2**(2), 121–167 (1998)