SEIZURE SIGNALS SEPARATION USING CONSTRAINED TOPOGRAPHIC BLIND SOURCE SEPARATION

Min Jing and Saeid Sanei

Centre of Digital Signal Processing, Cardiff University
Cardiff, CF24 3AA, South Wales, UK
Email: {jingm, saneis}@cf.ac.uk

ABSTRACT

Topographic independent component analysis (TICA) is an effective tool to group the geometrically nearby source signals. The TICA algorithm further improves the results if the desired signal sources have particular properties which can be exploited in the separation process as constraints. Here the spatial-frequency information of the seizure signals is used to design a constrained TICA (CTICA) for the separation of epileptic seizure signal sources from the multichannel electroencephalogram (EEG). The results are compared with those from the TICA and the superiority of the new CTICA has been demonstrated. [1]

1. INTRODUCTION

Epilepsy affects more than fifty million people in the world. Many studies have been carried out to explore the mechanisms of epileptogenesis and the possible solutions for anticipation and therapia [1]- [5]. Prediction of seizure has been investigated for decades. The most widely applied approach is based on the nonlinear analysis methods for the intracranial EEG recordings. The nonlinear analysis methods consider the epileptic brain as a dynamical system and the prediction problems can be solved by means of analysis of the corresponding nonlinear features. Although the dynamic properties have been shown in the scalp EEGs [6] [7], there are two main challenges in applying the traditional nonlinear methods to the scalp EEGs: one is that the scalp signals are more subject to the environmental noise and artifacts; second, the meaningful signals are attenuated and mixed when passing through the brain, skull, and scalp. A novel approach proposed in [8] [9] has already shown very promising results of the predictability from the scalp EEGs by incorporating the popular blind source separation (BSS) algorithm. The main concept of this approach is to consider the seizures as the independent components which are linearly and instantaneously combined together with the noise and artifacts over the scalp. All independent components can be separated by BSS algorithms, and the seizure sources can be selected by postprocessing. The traditional nonlinear analysis methods can be applied to these seizure components. This approach can be improved if the separation can achieve better performance. The objective of this work is to develop a method that can provide more plausible estimation of the seizure sources and eventually pave the way for the prediction of epileptic seizure from the scalp EEGs.

Independent component analysis (ICA) has been increasingly applied to brain signal analysis for decomposition of multivariate EEGs to extract the desired sources. It has found a fruitful application in the analysis of multichannel EEGs [10] including epileptic seizure signals. The conventional noise-free ICA model can be expressed as:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (1)$$

where $\mathbf{x}(t) = [x_1(t), x_2(t), ..., x_n(t)]^T$, $\mathbf{x} \in \mathfrak{R}^n$ is the vector of observed signals at time t, $(\cdot)^T$ denotes transpose operation. $\mathbf{s}(t) = [s_1(t), s_2(t), ..., s_m(t)]^T$ is the unknown independent source, $\mathbf{s} \in \mathfrak{R}^m$, $m \leq n$ for the over-determined mixture and $\mathbf{A} \in \mathfrak{R}^{n \times m}$ is the mixing matrix. The estimated sources $\mathbf{y}(t) = [y_1(t), y_2(t), ..., y_m(t)]^T$ can be obtained by a separation matrix $\mathbf{W}$ through inversion of the above data model,

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t) \quad (2)$$

where $\mathbf{W} = \mathbf{A}^\dagger$ is the pseudo-inverse of the mixing matrix and $\mathbf{W}\mathbf{A} = \mathbf{I}$. There are certain assumptions and restrictions for this model: 1) the source signals are statistically independent; 2) the independent components must have non-Gaussian distributions; and 3) the number of independent components are less or equal to the number of input channels [11]. The ICA algorithm has its own limitations which are the scale ambiguity and the permutation problem. Sometime it cannot take all the prior physiological information into account and the results of separation cannot be interpreted physiologically.

Topographic ICA (TICA) proposed by Hyvärinen et al. [12] is a modified ICA model, which relaxes the assumption of statistical independency of the components, considering the components topographically closed to each other are not completely independent but have certain dependencies. The dependencies are used to define a topographic order between these components. This provides an efficient approach for separation of the multichannel EEG source signals, because the dependencies between such sources cannot be simply cancelled out by some statistical assumptions. In this work, we show how TICA can be used for separation of the seizure signals from EEGs, and how the performance can be improved by introducing the novel spatial and frequency constraints. (The constrained TICA is denoted as CTICA).

The paper is organized as follows. Section 2 describes the algorithm development. The basic TICA model and principles are explained and the CTICA model is developed. Section 3 gives the experimental results obtained by applying the proposed methods to the epileptic seizure EEGs, the performance of CTICA and TICA is compared and the superiority of CTICA is demonstrated. The final section concludes the paper.

---

[1] An extended and detailed version of this manuscript is to be published in The Journal of Computational Intelligence and Neuroscience 2007.

2. ALGORITHM DEVELOPMENT

2.1 Topographic ICA

In the conventional ICA, the sources are assumed to be completely statistically independent and the estimated signals have no particular order. But in most real applications, some sources maybe more or less dependent on each other, such as the EEG sources which are fired from close locations within the cortex. In order to estimate the dependency of the independent components, Hyvärinen et al. proposed the TICA [12]. In TICA, the independency of the components has been relaxed, which means that the sources geometrically far from each other in topography are considered approximately independent and the sources close to each other are assumed to have certain dependencies. The dependency is defined as the higher-order correlation between the estimated sources, such as the correlation of the energies. In the TICA model, the variances of estimated components are not constant, instead, they are generated by some high-order independent variables. These variables are mixed linearly in the topographic neighborhood, which are defined by a neighborhood function $h(i, j)$. Based on this model, the estimated components in the same neighborhood are energy correlated. The approximation of the density of source s_i is given as:

$$\tilde{p}(s) = \prod_k exp(G(\sum_i h(i,k)s_i^2)), \quad (3)$$

where k is the index of the components within the same neighborhood. $G(\cdot)$ is the scalar function defined by incorporating certain nonlinearity as those defined in [12]. The approximation of the log-likelihood of this model is given as:

$$log\tilde{L}(\mathbf{W}) = \sum_{t=1}^{N} \sum_{j=1}^{n} G(\sum_{i=1}^{n} h(i,j)(\mathbf{w}_i^T \mathbf{x}(t))^2) + Nlog(|det\mathbf{W}|)$$

$$(4)$$

where, $\mathbf{w}_i$ is the column vector of the unmixing matrix, N is the length of the data, and $h(i, j)$ is the neighborhood function, which can be defined as a monotonically decreasing function of some distance. The second term of the above equation can be ignored because the unmixing matrix is constrained to be orthogonal and the determinant of an orthogonal matrix is one. Therefore, the estimation of the TICA model changes to choosing the optimal matrix $\mathbf{W}_{opt}$ that maximizes the above log likelihood function. The estimation of maximization of the log-likelihood can be found by:

$$\frac{\partial}{\partial \mathbf{W}} log\tilde{L}(\mathbf{W})|_{\mathbf{W}=\mathbf{W}_{opt}} = 0 \quad (5)$$

The computation of the gradient can be found in [12]:

$$\nabla_{\mathbf{W}_k} = 2 \sum_{t=1}^{N} \mathbf{x}(t)(\mathbf{w}_k^T \mathbf{x}(t)) \sum_{j=1}^{n} h(k,j)g(\sum_{i=1}^{n} h(i,j)(\mathbf{w}_i^T \mathbf{x}(t))^2)$$

$$(6)$$

where $g(\cdot)$ is the derivative of the scalar function $G(\cdot)$.

2.2 Constrained Topographic ICA

The estimated components from the TICA maybe dependent if they fall into the same neighborhood, i.e, the sources coming from the nearby location will be grouped together. However, the performance of TICA algorithm has certain limits.

It maybe not easy to obtain the sources grouped together unless the nearby sources are active at the same time. In [12], in order to obtain better visualization results, the experiment was set to generate some typical high energy sources such as biting teeth for 20 seconds. But in the most cases of real applications, the source signals may not be so significant, or there maybe only one or two of active sources. This could be the factor that affects the performance of TICA. Another factor is the number of input channels. It is obvious that the more input channels, the more information one can have and the better results can be achieved. This can be another limitation for the practical applications. However, the performance can be improved by introducing certain constraints in the algorithm.

Application of the constrained ICA for EEG analysis has been previously reported [8] [13]- [17]. The classic ICA dose not exploit the dependency of the sources therefore not always provides the desired outputs. For EEGs, there are valuable prior knowledge which can help to separate the desired sources. In this study, we consider two constraints which are based on spatial and frequency information. Firstly, in the focal epileptic seizures, the location of the seizure sources, "epileptogenic zone", is often known as the prior information. Secondly, seizure signals manifest themselves within certain frequency band. Based on the research findings from the clinicians and the neurologists, although the dominant frequency may vary for different types of seizures, the frequency band of the epileptic seizure onset is normally from 2.5 to 15.5 Hz. (Frequencies below 2.5 Hz are considered to be mainly due to eye-blinking artifacts) [18] [19] [20]. Therefore, the constraint can be determined based on both spatial and frequency domain information. The model of the constrained TICA problem can be presented as:

$$\max J_m(\mathbf{W}) \text{ s.t. } \min J_c(\mathbf{W}) \quad (7)$$

where $J_m(\mathbf{W})$ is the main cost function, here the main cost function is given in Eq. (4). $J_c(\mathbf{W})$ is the constrained cost function, which can be defined as minimizing the distance between the output and a reference signal:

$$J_c(\mathbf{W}) = \arg \min_{\mathbf{w}} \sum_{t=1}^{N} \|\mathbf{w}_i^T \mathbf{x}(t) - \mathbf{y}_r(t)\|_2^2 \quad (8)$$

where $\mathbf{y}_r$ is the reference signal defined based on the spatial and frequency constraints. $\| \cdot \|_2$ is the Euclidean distance. The CTICA is then changed to an unconstrained function by using a Lagrange multiplier. Therefore, the overall cost function is written as:

$$J(\mathbf{W}, \Lambda) = J_m(\mathbf{W}) - \Lambda J_c(\mathbf{W}) \quad (9)$$

where $\Lambda = diag\{\Lambda_{ii}\}, i = 1, ..., m$, is a diagonal weight matrix formed by

$$\Lambda = p \cdot diag(cor(\mathbf{y}_r, \mathbf{y}_i)) \quad (10)$$

where p is an adjust constant, $cor(\cdot)$ is the correlation measurement, and $\mathbf{y}_i$ is the ith estimated source. Then the update equation is obtained as:

$$\mathbf{W}(k+1) = \mathbf{W}(k) + \mu(k)\{\frac{\partial J_m(\mathbf{W})}{\partial \mathbf{W}} + \Lambda(\mathbf{X}(\mathbf{W}\mathbf{X} - \mathbf{Y}_r)^T)\}$$

$$(11)$$

where μ is the step size which is updated iteratively. $\mathbf{Y}_r$ is

the matrix with the reference signal $\mathbf{y}_r$ in each row .

3. EXPERIMENT

3.1 Data acquisition and the experiment setup

The multichannel EEGs with the frontal focal epileptic seizure were recorded from the standard silver cup electrodes applied according to the 'Maudsley' electrode placement system, which is a modification of the extended 10-20 system [21]. The 16 channels EEGs were sampled at 200 Hz and bandpass filtered in the frequency range of 0.3 - 70Hz. The system input range was 2 mV and data were digitized with a 12-bit analog-to-digital converter [8]. The EEGs were filtered by a 10th order Butterworth digital filter with a cut frequency of 45Hz in order to eliminate the 50Hz frequency component. The data used in the following experiments were truncated from the original recordings to include the duration of 10 seconds with seizure onset as shown in Fig.1.
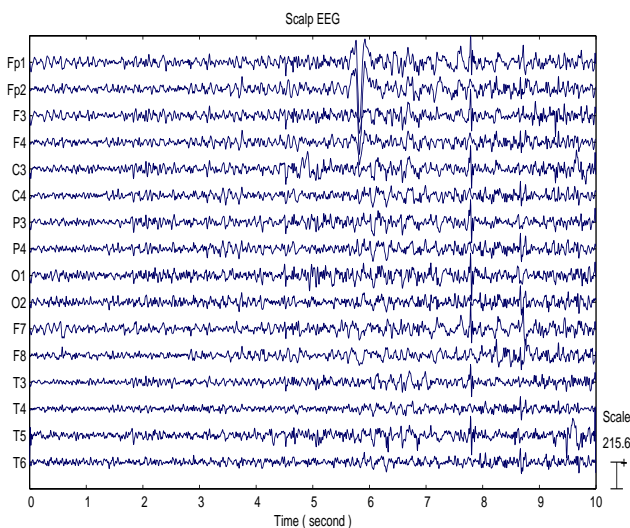

Figure 1: Multichannel EEGs from an epileptic patient with the seizure onset.

The reference was obtained by first averaging the special channels closed to the epileptogenic zone. In the experiment, F3, F4, F7, F8, C3 and C4 were selected. Then 3-15Hz bandpass filtering was undertaken to extract the information within the seizure frequency band. The final reference is a vector bounded within the designed spatial and frequency information of the seizure.

The neighborhood function indicates how the estimated sources are energy correlated with each other, which can be defined as a function of the width of the neighborhood. In this study, because of the limited number of input channels, the function was chosen as the simple one-dimensional form, such as $h(i,j) = 1$, if $|i - j| \leq m$, otherwise, $h(i,j) = 0$. m is the width of the neighborhood. It can be noticed that the neighborhood function is symmetric as $h(i,j) = h(j,i)$.

3.2 Results

The separation results of TICA and CTICA are given in Fig. 2 and Fig. 3, both with the width of neighborhood $m = 1$. A

simple detection rule based on the dominant frequency and respective estimated spectrum is applied to select the sources which have the significant ictal activities. The source with a maximum spectrum amplitude higher than a threshold and also with the dominant frequency in the seizure band, is taken as a seizure source. These sources are, IC7, IC8, IC9 and IC10 in Fig. 2, IC5, IC6, IC7 and IC8 in Fig. 3. One can see that the high amplitude spike signal are separated from the other sources. Another distinct source related to the eye blink can be seen from both of the outputs, which is IC12 in Fig. 2 and IC4 in Fig. 3.
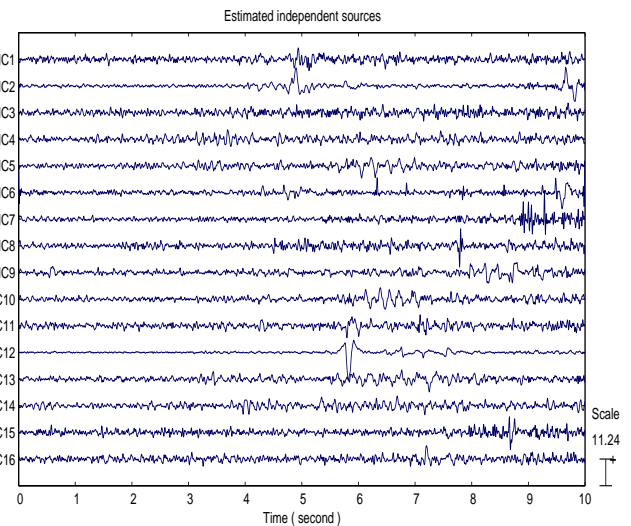

Figure 2: EEG source signals estimated by the TICA.

The differences between the source candidates may not be easily noticed by visual inspection of the time course of the sources, hence the topography is applied for this purpose. Fig.4 and Fig.5 show the topography of the sources estimated by TICA and CTICA respectively. Topography can be achieved by backprojecting the estimated source onto the original signal space, i.e. multiplying each column vector of the inverse of unmixing matrix by the corresponding estimated source. Topography reveals how each source signal contributes to the electrode signals. For example, one can notice that in both sets of results, the distribution of eye blink (IC12 in Fig. 4 and IC4 in Fig. 5) appears on the area near the electrodes Fp1 and Fp2. It can be found that, the four selected ICs are grouped together. The difference is, in Fig. 5, the selected ICs (IC5, IC6, IC7 and IC8) from the CTICA are localized in the frontal region, but in Fig. 4, the distribution of the corresponding sources (IC7, IC8, IC9 and IC10) by the TICA are rather dispersed. For instance, for IC10, the spatial distribution is highlighted in both frontal and temporal areas. A similar result can be noticed for IC11.

The differences can also be evaluated by comparing the performance in terms of the signal-to-interference ratio (SIR). Here, we define SIR to be the averaged signal energy for the estimated source $y(t)$ from the direct source divided
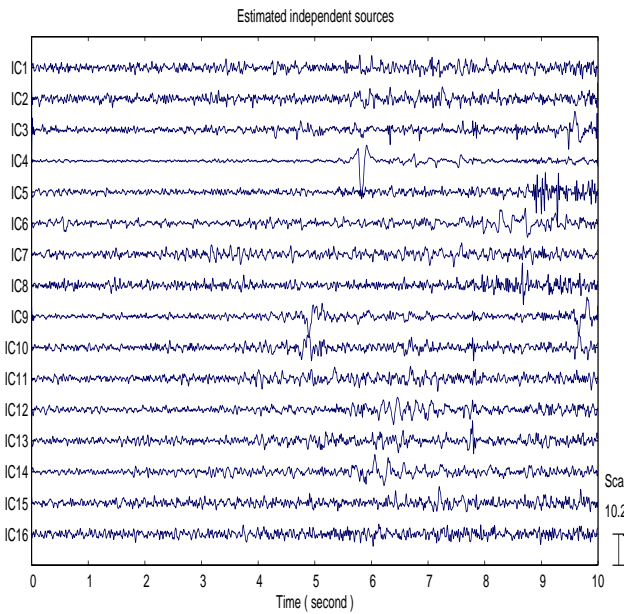
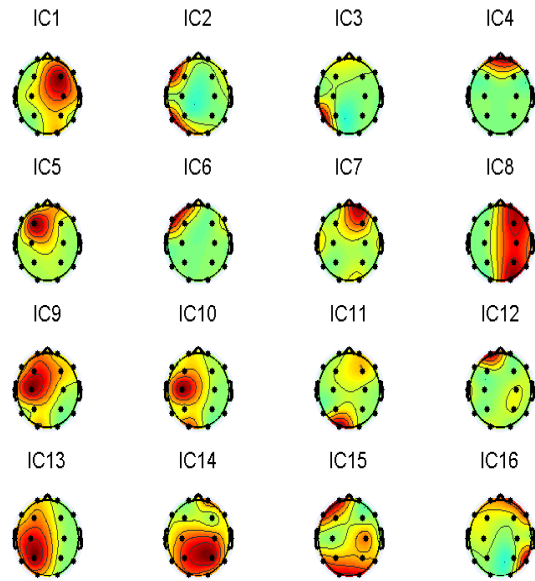Figure 3: EEG source signals estimated by the CTICA.

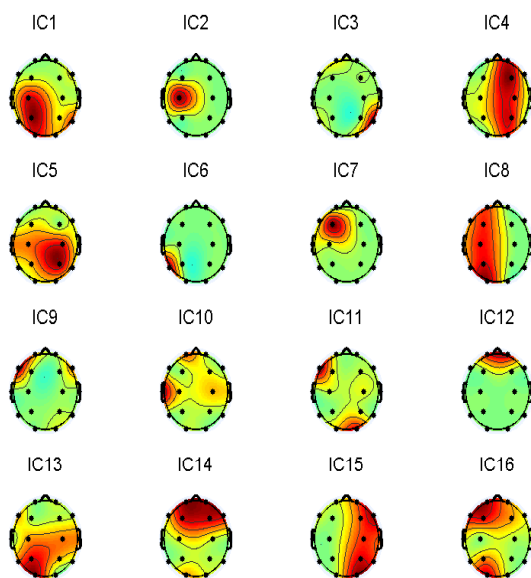Figure 4: Topography of the estimated EEG sources from TICA.

Figure 5: Topography of the estimated EEG sources from CTICA.

by the energy stemming from the other sources as:

$$SIR = \frac{\frac{1}{m}\sum_i^m |\mathbf{W}_{ii}^{-1}|^2 < |\mathbf{y}_i|^2 >}{\frac{1}{m(m-1)}\sum_{i\neq j}^m \sum_j^m |\mathbf{W}_{ij}^{-1}|^2 < |\mathbf{y}_j|^2 >} \qquad (12)$$

where $\mathbf{W}_{ii}^{-1}$ includes the diagonal elements in the inverse of unmixing matrix, i.e. the weights from source $\mathbf{y}_i$ to sensor $\mathbf{x}_i$. The off-diagonal elements $\mathbf{W}_{ij}^{-1}$ provide the weights from the source $\mathbf{y}_j$ to the sensor $\mathbf{x}_i$. It shows how the source $\mathbf{y}_j$ interferes the source $\mathbf{y}_i$, since each column of the inverse of unmixing matrix indicates the distribution of each source in the mixture, and a higher value indicates a better performance.

The performance of the algorithm is evaluated by the average of five trials for both the TICA and the CTICA. Fig. 6 shows the separation performance via the changes of the width of the neighborhood. It can be noticed that, the SIR of TICA decreases with increasing the neighborhood width. This is because the wider the neighborhood is, the more the sources will be separated based on the energy correlation. However, for the CTICA, because of the spatial and frequency constraints, the SIR slightly decreases at the beginning, then stays approximately at certain level. It shows that generally, the CTICA has a better performance than the TICA. It also works better than the TICA when the width of the neighborhood increases.

4. CONCLUSION

A new constrained topographic BSS algorithm has been developed to separate the epileptic seizure sources from the EEG signals. The incorporated constraints are based on the prior information about spatial and frequency of the focal epileptic seizures. The experimental results show that seizure
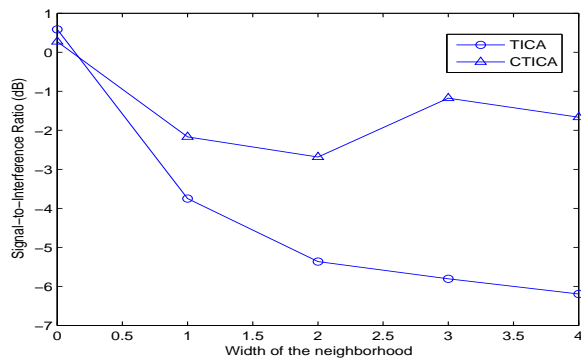
Figure 6: Performance comparison of the TICA and the CTICA.

sources coming from the epileptogenic zone are grouped together. The CTICA algorithm achieves better performance in terms of seizure source separation, and also works better when the width of the neighborhood increases.

REFERENCES

[1] L. D. Iasemidis, "Epileptic seizure prediction and control," IEEE Trans. on Biomedical Engineering, vol.50, pp. 549 - 558, 2003.

[2] F. H. L da Silva, W. Blanes, S. N. Kalitzin, J. Parra, P. Suffczynski, and D. N. Velis, "Dynamical diseases of brain systems: different routes to epileptic seizures," IEEE Trans. on Biomedical Engineering, vol. 50, pp. 540-548, 2003.

[3] M. L. V. Quyen, J. Martinerie, V. Navarro, M. Baulac and F. J. Varela, "Characterizing neurodynamic changes before seizures," J. Clin. Neurophysiol., vol. 18, no. 3, pp, 191-208, 2001.

[4] L. D. Iasemidis, D. S. Shiau, W. Chaovalitwongse and J. C. Sackellares,"Adaptive epileptic seizure prediction system," IEEE Trans. on Biomedical Engineering, vol. 50, pp. 616-626, 2003.

[5] L. D. Iasemidis, J. C. Principle, and J. C. Sackellares, "Measurement and quantification of spatio-temporal dynamics of humna eplptic seizures," Nonlinear Signal Processing in Medicine, IEEE Press, 1999.

[6] D. S. Shiau and L. D. Iasemidis "Detection of the preictal period by dynamical analysis of scalp EEG," Epilepsia, vol.44(S.9), pp. 233-234, 2003.

[7] J. C. Sackellares, L. D. Iasemidis, D. S. Shiau, R. L. Gilmore and S. N. Roper, "Detection of the preictal transition from scalp EEG recordings," Epilepsia, vol.40(S7), pp.176, 1999.

[8] J. Corsini, L. Shoker, S. Sanei, and G. Alarcon, "Epileptic seizure predictability from scalp EEG incorporating BSS," IEEE Trans. on Biomedical Engineering, vol. 53, no. 5, pp. 1654 - 1669, 2006.

[9] M. Jing, S. Sanei, J. Corsini, and G. Alarcon, "Incorporating BSS to Epileptic Seizure Predictability Measure from Scalp EEG," 27th Annual International Conference of Engineering in Medicine and Biology Society (IEEE-EMBS 2005), pp.5950 - 5953, 2005.

[10] S. Sanei and J. Chambers, "EEG Signal Processing," John Wiley - Sons, 2007.

[11] A. Hyvärinen, J. Karhunen and E. Oja,"Independent Component Analysis", Wiley-Interscince Publication, 2001.

[12] A. Hyvärinen, P. O. Hoyer and M. Inki," Topographic Independent Component Analysis," Neural Computation, vol. 13, pp. 1527-1558, 2001.

[13] W. Wang, M. G. Jafari, S. Sanei and J. Chambers,"Blind sepration of convolutive mixtures of cyclostationary signals," Int. J. Adapt. Control Signal process, vol. 18, pp. 279-298, 2004.

[14] M. A. Latif, S. Sanei, J. Chambers, and L. Shoker, "Localization of Abnormal EEG Sources Using Blind Source Separation Partially Constrained by the Locations of Known Sources" IEEE Signal Processing Letters, vol. 13, no. 3, pp. 117 - 120, 2006.

[15] L. Spyrou, M. Jing, S. Sanei, and A. Sumich "Separation and Localisation of P300 Sources and Their Subcomponents Using Constrained Blind Source Separation," EURASIP J. on Advances in Signal Processing, vol. 2007, pp. 1-10, 2007.

[16] W. Wang, S. Sanei and J. Chambers, "Penalty Function-Based Joint Diagonalization Approach for Convolutive Blind Separation of Nonstationary Sources" IEEE Tran on singal processing, vol. 53, no. 5, pp. 1654 - 1669, 2005.

[17] N. Ille, R. Beucker, and M. Scherg,"Spatially constrained independent component analysis for artifact correction in EEG and MEG," Neuroimage, vol. 13, S159, 2001.

[18] G. Lantz, C. M. Michel, M. Seeck, O. Blanke, T. Landis, and I. Rose,"Frequency domain EEG source localization of ictal epileptiform activity in patients with partial complex epilepsy of temporal lobe origin," Clin. Neurophysiol., vol. 110, pp. 176 - 184, 1999.

[19] O. Blanke G. Lantz, M. Seeck, L. Spinelli, R. Grave de Peralta, G. Thut, T. Landis, and C. M. Michel,"Temporal and Spatial Determination of EEG-Seizure Onset in the Frequency Domain," Clin. Neurophysiol., vol. 111, pp. 763-772, 2000.

[20] P. Wahlberg and G. Lantz,"Approximate time-variable coherence analysis of multichannel signals", Multidimensional Systems and Signal Processing, vol. 13, pp. 237-264, 2002.

[21] J. L. Fernndez, G. Alarcn, C. D. Binnie, and C. E. Polkey,"Comparison of sphenoidal, foramen ovale and anterior temporal placements for detecting interictal epileptiform discharges in presurgical assessment for temporal lobe epilepsy," Clin. Neurophysiol., vol. 110, pp. 895-904, 1999.