Simultaneous EEG-fMRI Analysis with Application to Detection of Seizure Signal Sources

Min Jing[1] & Saeid Sanei[2]
[1]*Intelligent Systems Research Centre*
University of Ulster
[2]*Centre of Digital Signal Processing*
Cardiff University
United Kingdom

1. Introduction

1.1 Brain Function Monitoring Modalities

Introduced by Hans Berger in 1929, Electroencephalography (EEG) is a recording of the electrical current potentials spontaneously generated by cortical nerve cell inhibitory and excitatory postsynaptic potentials. These postsynaptic potentials summate in the cortex and extend to the scalp surface where they are recorded as the EEG signals (EEGs). EEG provides a noninvasive means of monitoring brain activity and investigating brain function disorders. Generally, the EEG reflects the changes of cerebral function directly and reliably, especially if the structural lesions are localized near the surface of the hemispheres. EEG plays a very important role in the diagnosis of specific neurological diseases such as epilepsy and it is a very useful tool in clinical applications.

Functional Magnetic Resonance Imaging (fMRI) is an advanced imaging technique which delineates the brain activated areas responding to the designed stimuli such as sound, light or finger's movement. The principles of MRI are based on nuclear magnetic resonance (NMR). The NMR signal originates from the hydrogen nucleus which has a single proton. When the proton is placed in an external magnetic field, transitions of energy occur as the proton absorbs or emits a photon. The NMR signals generated by this energy transition can be detected and presented in an anatomical image. fMRI is able to show the blood flow within the brain activated areas in the image as the blood oxygen level-dependence (BOLD) response, which provides valuable spatial information on the brain. However, because it relies on blood flow response rather than electrical activity, it has a relatively slow response to temporal changes.

Many unanswered questions about the relationship between the cerebral haemodynamic changes (measured by fMRI) and the underlying neural electrical activity (revealed by EEG) are of interest to many researchers. Although the spatiotemporal relationship between fMRI and EEG is still far from straightforward, there are promising perspectives presented in the

literature. Logothetis et al. compared local field potentials (LFPs) with the fMRI responses from the visual cortex of monkeys (Logothetis et al, 2001). The largest magnitude changes observed in LFPs at recording sites characterized by transient responses were the signals that highly correlated with the haemodynamic response. The relationship between fMRI and event-related potential (ERP) was also examined during an auditory odd-ball paradigm (Horovitz et al, 2002). The results have shown for the first time that for auditory stimuli, the amplitude of the hemodynamic response in the region of interest (ROI) follows the amplitude of the ERP changes, and BOLD signals from the source location of P300 have high correlation with the amplitude of the P300. In a more recent simultaneous EEG-fMRI study of painful electric stimulation (Christman et al, 2007), it has been shown that the BOLD changes in ROI were correlated with the dipole strength of the EEG source. The results revealed a close relationship of BOLD signal and possible underlying neural electrical activity in ROI. The results provided evidence that there were underlying connections between fMRI and EEG.

1.2 Aspects of Application in Fusion of EEG and fMRI

It is obvious that each modality has its own advantages. For example, EEG reflects the brain changes on a timescale of milliseconds (1kHz or more), which is capable of capturing the dynamic changes of the brain very well. However, it has poor spatial resolution due to being recorded from a limited number of electrodes on the scalp, and the problem of source localization from the EEG still remains a research challenge. On the other hand, fMRI is extremely powerful in investigating the brain function, but it is slow to follow the brain activities because it relies on the brain blood flow response rather than electrical activities. How to integrate these two modalities technically and maximize their advantages in real applications has started to attract more attention in recent years.

Numerous efforts have been directed towards combining high spatial information provided by fMRI, with the high quality temporal data generated by EEG or MEG. These approaches mainly focus on three aspects (Horwitz & Poeppel, 2002). The first aspect is referred to as direct data fusion, which is normally applied in EEG source localization. The geometrical information on the source activation region obtained from fMRI can be used as the constraint for localizing the EEG dipole sources (Ahlfors et al, 1999; Phillips et al, 2002; Babiloni et al, 2003; Christman et al, 2007), because there is no unique solution for the ill-posed inverse problem for EEG source localization in the absence of constraints. For example, in (Babiloni et al, 2003), the geometrical information from fMRI provided the realistic head model as a volume conductor medium to help reduce the EEG localisation solution space. In (Ahlfors et al, 1999), fMRI was not used as a rigid constraint but helped in selecting more likely inverse solution among the possible solutions. The second aspect is based on the use of computational neural models, in which the relation of EEG and fMRI is modelled on the basis of some hypotheses of certain neural activities (Schilling et al, 2001; Babajani & Soltanian-Zadeh, 2006). In these approaches, the EEG and fMRI data are not linked directly, but are compared inside a simulated neural model. The main challenge for this approach is to construct a recurrent neural model for simulation of the complex neural physiological activities, which is still questionable from the physiological aspect because the simulated neural model simplifies the complexity of neural activities (Horwitz & Poeppel, 2002). The third type of these fusion approaches is more commonly used in the clinical or neurological field in which temporal information from the EEG helps to time-lock the events

in the fMRI. In some studies (Lemieux et al, 2001; Diehl et al, 2003), the epileptic EEG and fMRI data were simultaneously recorded during seizure onset, and the spikes within the seizure EEG were modelled as the stimuli which can be used in the statistical parametric mapping in fMRI. The provided brain map of seizure can be used for help in surgical planning. In this study, the work is related to the third aspect, which aims at incorporating information from EEG into fMRI analysis for the detection of the "epileptogenic zone".

1.3 Analysis for Epileptic Seizure fMRI

Epilepsy covers a group of related disorders characterized by a tendency for recurrent seizures. The seizures are due to a sudden development of synchronous neuronal firing in the cerebral cortex. The cause of epilepsy still remains unknown and different types of epilepsy have been discovered. For example, based on whether the whole brain or just part of it is involved, the seizure can be classified as generalized, focal or partial. Generalized type means that the seizure involves the whole brain at once. Focal or partial seizure means that the seizure originates from one area of the brain. In the present study, the work is related to focal seizure, and the "epileptogenic zone" is known as priori.

Analysis of epileptic seizure fMRI remains a challenging problem in the related research field. In comparison with the common fMRI analysis, for which the functional data are acquired from the designed experiments, the fMRI data of epileptic seizures are very different. As the spontaneous brain activity caused by brain functional disorders, the response of epileptic seizure is very difficult to be modelled as the normal hemadynamic response, which means the commonly used method as General Linear Model (GLM) for fMRI analysis faces the challenge for epileptic seizure fMRI. There has been limited literature (Lemieux et al, 2001; Salek-Haddadi et al, 2003; Diehl et al, 2003) that investigated the statistical parametric mapping of epileptic seizure spikes. Those results were limited by carefully choosing functional data that have distinguishable periodic seizure spikes, in which the spikes were used as the stimulus to construct the design matrix in GLM. Although these approaches partly solved the problem of model specification in GLM, the clinical expertise is required to time-lock the spikes, which makes those methods practically difficult to be applied to the unpredictable brain events, such as seizures. However, the seizure active area can be detected without concerning the model specification by using a data-based method such as Independent Component Analysis (ICA). If the information available in EEG and fMRI can be incorporated in the separation process of ICA, then the performance of ICA can be improved, which is the main objective of the present work.

This chapter is arranged as follows. First, the most commonly used fMRI analysis method GLM is introduced. Second, the spatial ICA model is described in detail and a comparison between the spatial and temporal ICA models is discussed, with the development of the constrained spatial ICA algorithm given afterwards. In the final section, the experimental results are given and discussed.

2. Methods of fMRI Analysis

2.1 Model Based Method

The most commonly used model-based approach is the general linear model (GLM) (Friston et al. 1995). According to GLM, fMRI data from each voxel is considered as a linear combination of the hemodynamic responses of stimuli and their corresponding weighted

parameters. The stimuli are modelled as delta functions and the response is the convolution of the stimulus and the predefined hemodynamic response functions (HRF). The response is referred to as design matrix in GLM, and the HRF can be chosen from some well defined functions such as the Gamma function and Fourier set (windowed sines and cosines). Therefore, based on GLM, in order to specify the model, prior knowledge or specific assumptions about the time courses contributing to the signal changes are required. After the model specification, the weight parameters can be estimated by using estimation techniques such as maximum likelihood estimation (MLE) or Bayesian estimation. The active areas are then detected by evaluating the statistical significance of the whole brain voxels (Friston et al, 1995). Several software toolboxes for fMRI analysis have been developed based on GLM, which can be used for fMRI data preprocessing, model specification, statistical parameter estimation and parameter mapping, such as statistical parametric mapping (SPM) and FMRIB Software Library (FSL). The general linear model is a linear model which can be presented as,

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E} \tag{1}$$

where **Y** is an N × V matrix representing the fMRI time series in each voxel, N is the number of scans, and V is the number of voxels involved in the analysis. **X** is an N × F matrix referred to as the design matrix, which is the predicted event response by convolving the stimulus with the predefined hemadynamic response function, and F is the number of stimuli (events). B is an F × V matrix of unknown parameters which are to be estimated, and **E** represents the errors which are assumed to be independently and identically distributed normal random variables (Friston et al, 1995), For the fMRI data from the jth voxel, the elements of the above matrix equation can be represented as:

$$\begin{pmatrix} y_{1j} \\ \vdots \\ y_{ij} \\ \vdots \\ y_{Nj} \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1j} & \cdots & x_{1F} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i1} & \cdots & x_{ij} & \cdots & x_{iF} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{Nj} & \cdots & x_{NF} \end{pmatrix} \begin{pmatrix} b_{1j} \\ \vdots \\ b_{ij} \\ \vdots \\ b_{Fj} \end{pmatrix} + \begin{pmatrix} \epsilon_{1j} \\ \vdots \\ \epsilon_{ij} \\ \vdots \\ \epsilon_{Nj} \end{pmatrix} \tag{2}$$

From the above equation, one can see that the column vector $\mathbf{b}_j = [b_{1j},\ldots,b_{Fj}]^T$ is actually the weight factor of each event response at the *jth* voxel. An estimation of the parameter **B** which represents the weight for all voxels, denoted as $\hat{\mathbf{B}}$, can be obtained as,

$$\hat{\mathbf{B}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} \tag{3}$$

The conventional approach for fMRI analysis is based on evaluation of the statistical significance in each voxel, for example by means of the t-statistic (Friston et al, 1995). The activated areas are detected by selecting the voxels in which the statistical significance is higher than a certain threshold value.

2.2 Data Based Method

In contrast to the model-based GLM, the data-based model relies on the data instead of prior information on stimuli or predefined brain function. One of the data-based approaches,

proposed by McKeown et al as the first application of ICA to fMRI data analysis (McKeown et al, 1998), has attracted more attention (McKeown et al, 1998; Biswal & Ulmer, 1999; Suzuki et al, 2002; Duann et al, 2002; Beckmann & Smith, 2004; Calhoun et al, 2006; Reidl et al, 2007). In the ICA model, the fMRI data are considered as a linear combination of a number of temporally or spatially independent components. Comparing with GLM, the data-driven model is more suitable for analysis of brain signals because no assumptions regarding the stimulus response are required. The brain function and its hemadynamic response are so complicated that it is still questionable to simply choose certain predefined HRF and to assume that the shape of the HRF remains constant during the events for each brain voxel. In the following section, the details of the ICA model are discussed. The experimental results show that the ICA approach can be used to analyze the fMRI data in those cases that the GLM is difficult to work.

Apart from ICA, some other model-free approaches have been applied to fMRI analysis, such as support vector machine (SVM) (LaConte et al, 2005; Fan et al, 2006; Wang et al, 2006). The core concept of using SVM or any other pattern classification method is based on the fact that the voxels within the active areas certainly contain some special patterns which can distinguish the active part from the rest of the brain. These features can be image features such as image intensity, probability density, statistical information from each voxel, or the shape of the temporal sequence of the designed event. By means of feature extraction, the region of activation can be detected.

Information theory has also been applied to fMRI analysis. The main idea is to estimate the information across the whole brain voxels under the designed experiments and then to decide the active areas based on the mutual information (MI) criterion. As in (Galit et al, 2007), Galit et al. proposed a model-free method based on measuring the entropy and MI, which detected the location of the event-related activity by evaluation of the temporal information across different brain regions. The stimulus was used as a reference, the MI between the fMRI signals in each voxel and the stimulus was measured. The active areas were then detected by selecting those voxels which had higher MI according to the task conditions.

2.2.1 Spatial ICA

The first application of the spatial ICA model for fMRI analysis was proposed by McKeown et al. According to the authors, the physiological foundation for the ICA model is based on the two complementary principles of brain function, namely localization and connectionism. Localization implies that each psychomotor function is performed in a small region of the brain area; the principle of connectionism reveals that the active brain area involved in certain functions may be widely distributed in the multiple distinct brain systems (McKeown et al, 1998). The spatial ICA model was introduced based on these two principles. In this model, the brain areas executing different tasks are assumed to be spatially independent. Each of these areas can be considered as an independent component associated with a time course. According to the statistical definition of independency, the spatial independence can be defined as:

$$p(\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_n) = \prod_{i=1}^{n} p_i(\boldsymbol{c}_i) \tag{4}$$

where $\mathbf{c}_i$ is the ith independent spatial component, and the joint pdf $p(.)$ is the multiplication of the marginal pdfs of the components. As for the conventional ICA model, the spatial ICA model is formed as:

$$\mathbf{X} = \mathbf{MC} \tag{5}$$

where $\mathbf{X}$ is a T × V matrix of the mixtures, T is the length of the fMRI scan, V is the number of brain voxels involved in the analysis. $\mathbf{C}$ is an N × V matrix of unknown sources. $\mathbf{M}$ is a T × N mixing matrix, and N is the number of unknown spatially independent sources. Each column of $\mathbf{M}$ represents the time course of the corresponding independent component. Based on this model, fMRI signals can be decomposed into a number of spatially independent components $\mathbf{C}$ and their associated time course of activation $\mathbf{M}$. The spatial components can be estimated from:

$$\mathbf{C} = \mathbf{WX} \tag{6}$$

where $\mathbf{W}$ is an N × T unmixing matrix to be estimated, and $\mathbf{W}$ is the pseudoinverse of $\mathbf{M}$, i.e. $W= M^{\dagger}$.

In contrast to conventional ICA, which is based on temporal independence, the spatial ICA is based on the assumption of spatial independence. Although some research has exploited both the spatial and temporal independence (Stone et al, 2002), most approaches are still based on the assumption of spatial independence due to lack of good understanding of the unknown brain activities for the temporal dynamics of fMRI. Another reason that spatial ICA is more favourable for decomposing fMRI data is that it is computationally less expensive. Temporal and spatial ICA analysis of fMRI data has been compared in the previous work (Calhoun et al, 2001). Figure 1 illustrates the difference between the two models. It is noticed that, for the mixture, $\mathbf{X}$ (fMRI data matrix) in the spatial model is the transpose of $\mathbf{X}$ in the temporal model, and the spatial dimension is much higher than the temporal one because the number of brain voxels is larger than the number of time points of the scans. For estimation of the unmixing matrix in temporal ICA, a covariance matrix on the order of V^2 must be calculated, which is more computationally expensive than that for spatial ICA. For these reasons, spatial ICA is selected in this work and is denoted as SICA in the following section.

2.2.2 Constrained ICA for fMRI Analysis

The constrained ICA has been applied to fMRI signal analysis in order to incorporate prior information since the SICA model does not take the fMRI time course into account. Although very limited, recent work has shown that the performance of the application of ICA to fMRI analysis is improved if some prior information is incorporated into the estimation process (Calhoun et al, 2006; Lu & Rajapakse, 2005). Lu et al applied a predefined stimulus as the reference signal in the temporal ICA model. By minimizing the distance between the output and reference signal, the source component closest to the reference can be obtained. Calhoun et al. developed a semi-blind spatial ICA (Calhoun et al, 2005), in which a constraint is introduced by incorporating the GLM design matrix which contains information on the fMRI time course. The column vectors of the mixing matrix are constrained according to their closeness to the time course estimated from the design matrix.

In each iteration, apart from updating **W** based on the Infomax algorithm, an additional updating rule is set by measuring the correlation between the columns of the mixing matrix $[\mathbf{W}]^{-1}$ and the time course estimated from the design matrix. The additional updating rule is only performed if the correlation is lower than a certain threshold.
Although the semi-blind spatial ICA developed by Calhoun provides some promising results, their constraint still relies on GLM, therefore it is only suitable for the case that the stimuli of the fMRI data are pre-specified. For an epileptic fMRI, this approach cannot work because it is very difficult to model the epileptic seizures. In this work, the idea is to incorporate the information from the simultaneously recorded EEG as the constraint into the spatial ICA. As one can see from the spatial ICA model (shown in Figure 1), the fMRI data can be decomposed into spatially independent components and the associated time courses, for which each column of the mixing matrix represents the time course of one component activation. Therefore, the temporal constraint can be introduced by linking the EEG with the columns of $[\mathbf{W}]^{-1}$ in the separation process. In the following section, constrained spatial ICA is denoted CSICA.

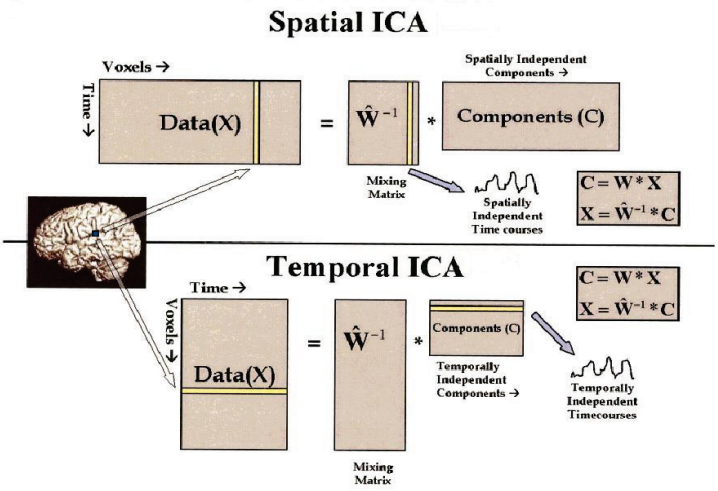

Fig. 1: Comparison of the key stages of processing within spatial ICA and temporal ICA (Calhoun et al, 2001).

2.3 Development of Algorithm

The performance of different ICA algorithms for fMRI data separation has been studied (Calhoun et al, 2004; Correa et al, 2005). Basically, selection of the separation algorithm depends on the assumption about the distribution of the sources. For example, fMRI data are commonly assumed to have super-Gaussian distribution (Calhoun et al, 2006), from which one can expect that the algorithm which is more suitable for super-Gaussian signals will achieve a better performance. As shown in (Correa et al, 2005), Infomax consistently yields reliable results for separation of fMRI followed by JADE and FastICA. Therefore, Infomax was selected in this work. In the following sections, the principle of Infomax is explained. Then the constraint is introduced to the Infomax learning rule and the development of the constrained algorithm is given.

2.3.1 Infomax

Infomax (Bell & Sejnowski, 1995) is based on information theory by maximizing the output entropy or information flow of a neural network with nonlinear outputs, hence the name Infomax. Assume that the neural network with output y and input x is of the form:

$$\mathbf{y}_i = \varphi_i\left(\mathbf{w}_i^T\mathbf{x}\right) + \mathbf{e} \tag{7}$$

where $\varphi_i(\cdot)$ are some nonlinear scalar functions, the $\mathbf{w}_i$ are the weight vectors of the neurons, and $\boldsymbol{e}$ is the additive Gaussian white noise. The entropy of the output is:

$$H(\mathbf{y}) = H\left(\varphi_1\left(\mathbf{w}_1^T\mathbf{x}\right), \ldots \varphi_n\left(\mathbf{w}_n^T\mathbf{x}\right)\right) \tag{8}$$

For an invertible transformation of the random vector $\mathbf{x}$, $\mathbf{y} = f(\mathbf{x})$, the relation between the entropies of $\mathbf{y}$ and $\mathbf{x}$ can be expressed as (Hyvarinen et al, 2001),

$$H(\mathbf{y}) = H(\mathbf{x}) + E\{\log|\det Jf(\mathbf{x})|\} \tag{9}$$

where $Jf(\cdot)$ is the Jacobian matrix of function $f(\cdot)$, and $E\{\cdot\}$ denotes the expectation operation. According to Eq. (9), the transformation of the entropy in Eq. (8) can be obtained as

$$H(\mathbf{y}) = H(\mathbf{x}) + E\left\{\log\left|\det\frac{\partial F}{\partial \mathbf{W}}(\mathbf{x})\right|\right\} \tag{10}$$

where $F(\mathbf{x}) = \varphi_1\left(\mathbf{w}_1^T\mathbf{x}\right), \ldots \varphi_n\left(\mathbf{w}_n^T\mathbf{x}\right)$ denotes the nonlinear function defined by the neural network. As $H(\mathbf{x})$ is independent of $\mathbf{W}$, the entropy of output can be expressed as

$$H(\mathbf{y}) = \sum_i E\left\{\log\varphi_i'(\mathbf{w}_i^T\mathbf{x})\right\} + \log|\det\mathbf{W}| \tag{11}$$

It is noticed that maximization of the output entropy is very closely related to the maximum likelihood (ML) estimation. The ML estimation yields (Bell & Sejnowski, 1995),

$$\Delta\mathbf{W} \propto [(\mathbf{W}^T)^{-1}\, g(\mathbf{y})\mathbf{x}^T] \tag{12}$$

where the nonlinear function $g(\mathbf{y})$ is the column vector whose i-th component is

$$g_i(y_i) = -\frac{\frac{\partial p(y_i)}{\partial y_i}}{p(y_i)} \tag{13}$$

where $p(y_i)$ is an approximate model of the pdf of the i-th source signal. In practice, $g_i(y_i) = 2\tanh(y_i)$ is selected as it is suitable for super-Gaussian signals (Hyvarinen et al, 2001). By defining $g(\cdot)$ in a different way, the Infomax can also work for sub-Gaussian signals, which is referred to as extended-Infomax (Lee et al, 1999). An efficient way to maximize the log-likelihood is to follow the natural gradient learning method

$$\Delta \mathbf{W} \propto \frac{\partial \mathrm{H}(\mathbf{y}, \mathbf{W})}{\partial \mathbf{W}} \mathbf{W}^{\mathrm{T}} \mathbf{W} = [\mathbf{I} - \varphi(\mathbf{y})\mathbf{y}^{\mathbf{T}}]\mathbf{W} \tag{14}$$

Then, the updating rule can be written as

$$\mathbf{W}\,(\mathrm{k}+1) = \mathbf{W}(\mathrm{k}) - \eta(\mathrm{k})\big[\mathbf{I} - \varphi\big(\mathbf{y}(\mathrm{k})\big)\mathbf{y}(\mathrm{k})^{\mathbf{T}}\big]\mathbf{W}(\mathrm{k}) \tag{15}$$

where k is the iteration number, **I** is the identity matrix and $\eta(k)$ is the learning rate. The optimal solution for **W** is obtained when the estimated sources y_i and y_j are independent. As shown in (Cichochi & Amari 2002), the stability condition of the learning rule in Eq. (14) converges to an equilibrium point corresponding to the optimal solution which can be expressed as $\mathbf{E}\{g_i(y_i)y_i\} = 1$. This is not only the condition for local stability of the algorithm, but also determines the scaling of the estimated sources. This condition is invariant with respect to the sign of y_i as $g_i(y_i)$ is selected as an invertible (monotonic) function $\tanh(y_i)$. As Infomax is very similar as ML estimation as shown in Eq. (11), the unmixing matrix must be constrained to be orthogonal such that the determinant of **W** is one and the second term in Eq. (11) can be ignored.

2.3.2 Constrained Algorithm

In this work, the objective is to incorporate the EEG signal as the constraint into the fMRI data separation process since EEG contains valuable temporal information about the brain activity. As the columns of the mixing matrix represent the time courses of the estimated components, intuitively the temporal constraint can be added to the columns of the mixing matrix such that the EEG information can be taken into account. However, incorporating this information into the separation process is a problem to be resolved.

The relationship between fMRI and EEG is far from straightforward due to the complexity of the brain mechanism and very limited understanding of it at present. Although neural networks have been exploited to model the relationship between fMRI and EEG (Schilling et al, 2001; Babajani & Soltanian-Zadeh, 2006), it is still hard to be established in the neuro-physiological and clinical fields, because theoretically the complexity between the hemadynamic changes and neural activities cannot be fully represented by a simple mathematical model (Horwitz & Poeppel, 2002). Practically, correlation measurement has been widely used in the existing studies for investigating the relationship between fMRI and EEG (Logothetis et al, 2001; Horovitz et al, 2002; Mathalon et al, 2003; Calhoun et al, 2006; Christman et al, 2007). Therefore, in this work, correlation is used to connect the time course of fMRI components and the corresponding EEG signals.

The constrained term reflects the closeness between the *i*-th column vector of the inverse of unmixing matrix $[\mathbf{W}]^{-1}$ and the processed EEG time series **u**. The seizure signal **u** has to be selected carefully either by applying temporal ICA to EEG data to obtain the seizure component, or based on prior clinical information about the seizure. In this work, **u** is formed on the basis of prior clinical information, since the epileptogenic zone is known as a *priori*.

The constraint is imposed in the Infomax update rule in Eq. (15). A similar method to add a constraint in the natural gradient rule can be found in the nonholonomic learning rules (Cichochi & Amari 2002; Amari et al, 2002). The basic natural gradient learning equation is then extended as:

$$\mathbf{W}(\mathrm{k}+1) = \mathbf{W}(\mathbf{k}) - \eta(\mathrm{k})\left[\mathbf{I} - \varphi\left(\mathbf{y}(\mathrm{k})\right)\mathbf{y}(\mathrm{k})^{\mathrm{T}} - \alpha\boldsymbol{\Lambda}(\mathrm{k})\right]\mathbf{W}(\mathrm{k}) \tag{16}$$

where α is the factor adjusted based on the stability of the algorithm. $\Lambda = \mathrm{diag}\{\Lambda_{ii}\}$, $i=1,..,N$ is a diagonal weight matrix containing the information from the EEG, which is updated as

$$\boldsymbol{\Lambda}(k) = diag\left(cor\left([\mathbf{W}]_i^{-1}(k), \mathbf{u}\right)\right) \tag{17}$$

where *cor(·)* denotes correlation. (Here, correlation is defined as a normalised version of covariance, i.e., $cor(y_i, y_j) = E\{y_i y_j\}$). According to the adaptive learning rule in Eq. (16), **W** is updated based on the Infomax principle, also the column vectors of its inverse are forced to be close to the corresponding processed EEG signal **u**. As shown in Eq. (17), Λ is updated iteratively according to the closeness between and **u**. Here, the entries of Λ are bounded since the absolute value of correlation coefficient is less than 1. Due to the additional constraint, the new algorithm can converge to the lower minimum of the cost function than the one before imposing a constraint, whereby the performance of algorithm can be improved.

3 Implementation of the Algorithm

The experiments comprise two parts. In the first part, the format of fMRI data and the basic preprocess procedure are introduced. In the second experiment, SICA and CSICA are applied to the epileptic fMRI data. GLM cannot work in this case because it is very difficult to model the epileptic seizure signals. The simultaneously recorded EEG was introduced as the constraint in CSICA and the process of constructing the constraint is described in detail. The performances of SICA and CSICA are compared in terms of algorithm convergence, the detected seizure area, and the closeness between the seizure component and EEG signal.

3.1 Preprocessing of fMRI Data

3.1.1 fMRI Data Format

The primary functional image data format used in this work is Analyze 7.5. An Analyze 7.5 data format consists of two files, an image file and a header file, with extensions ".img" and ".hdr" respectively. The .img file contains the image data information. The .hdr file contains the volume information of the .img file, such as voxel size, and the number of pixels in the x, y and z directions (dimensions). Also, a MATLAB file .mat is added to the .hdr and .img pair. The .mat file includes some information on the orientation of the image, which are generated by the realignment and coregistration processes. In the coordinate system for Analyze data format, the x-direction is from left to right, the y-direction is from back to front, and the z-direction is from bottom to top. The image obtained from one scan is referred to as one volume. Each volume consists of a number of slices through the brain, and each slice has a certain thickness and is composed of a number of 3D unit elements called voxels. The volume of a voxel is approximately 3 mm^3. In general, the analysis of fMRI is executed based on each voxel.

3.1.2 Preprocessing

In order to apply ICA to the functional data, certain preprocessing must be performed before the data are ready for further analysis. The preprocessing includes not only the basic temporal and spatial preprocessing as needed in GLM, but also the data dimension reduction and data structure conversion which are required before applying ICA.

The first basic step is temporal and spatial preprocessing for the raw fMRI data by means of slice timing and realignment in order to remove the motion artifacts. In the experiments, this preprocessing is carried out by using the established tool in the SPM. The second step is data dimension reduction. The analysis of fMRI is always computationally expensive due to a large number of brain voxels involved. Therefore, a very important step of preprocessing before performing any fMRI analysis is to reduce the number of voxels involved in the analysis, and thereby to reduce the data dimension for further analysis. This can be executed by either removing the off-brain voxels (i.e. the voxels which fall outside the brain boundary), or extracting the voxels within the area of interest. Both processes can be carried out by applying various techniques.

For example, in SPM, the standard images of grey and white matter of the brain are provided to be used as mask images, so that the user can extract the voxels of interest. In another medical image viewing and processing toolbox MRIcro (Rorden, C. & Brett, M., 2000), the user can choose some masking images as in SPM. Also, in MRIcro, extraction of the region of interest (ROI) can be performed directly by drawing the ROI on the original images manually. Then the ROI can be converted to the Analyze 7.5 format which is readable in SPM. In the Matlab based software FMRLAB (http://sccn.ucsd.edu/fmrlab), the off-brain voxels are excluded by manually setting a threshold value. The user can visually check the changes in brain images during the process of removing off-brain voxels in the graphic interface, and decide on the threshold value. In the following experiment, the toolbox in FMRLAB was used to extract the brain voxels.

In the case of applying GLM to fMRI analysis, the above two steps of preprocessing are required. But for applying ICA to fMRI, one also needs to construct the input (data mixture) in order to perform the ICA. As seen from the ICA model (Eq. 5), the input data mixture has dimensions of $T \times V$, where T is the number of scans and V is the number of brain voxels involved in the analysis, which are the voxels after excluding the off-brain voxels. For each scan (one time point of fMRI data), which is referred to as one volume, the data are in 3D form (i.e., with x, y and z directions). In order to construct the 2D mixture with size $T \times V$, it is necessary to first compress all scans into a 4D dataset, then reshape it into the 2D mixture with dimension $T \times V$.

In the following experiment, the data were first centred to have zero mean. Then preliminary whitening was performed to make the data have unit variance before further separation processing. The estimation of number of spatial components in fMRI has been attempted recently (Li et al, 2007), in which the information-theoretic criteria was applied on the simulated data based on the minimization of the Kullback-Leibler divergence between the true model and the fitted model. However, it is still a relatively new research topic and many questions remain. In this study, the number of source was selected as the same as the number of input channels, which is the time points of fMRI data.

3.2 Experiment: Analysis of Epileptic EEG-fMRI

3.2.1 Data Details

The simultaneously recorded EEG and fMRI data were provided by the National Society for Epilepsy, University College London (UCL). The functional data were acquired on a modified 3T GE Horizon system and EEG data were recorded by the Brain Product system. The length of EEG-fMRI data is approximately 5 minutes before and during the seizure onset. The functional data were acquired from the 16*th* scan. In this experiment, the functional data were truncated from acquisition 20 to acquisition 107, which is the scan just before seizure onset. The first four scans were discarded in order to remove the initial magnetic gradient effect in the fMRI recording. Each acquisition comprised 47 contiguous slices, with image dimension of 64 × 64 × 47, and the volume size was 3.75*mm* × 3.75*mm* × 2.5*mm*. The interval between each scan was 3 sec. The simultaneous 64 channel EEGs were sampled at 250 Hz. Before applying the proposed algorithm over the simultaneous EEG-fMRI data, the scanner artifacts were removed from EEGs by the data provider. The epileptic seizure area was known as the right temporal area.

3.2.2 Experiment Setup

In this experiment, SICA and CSICA were applied to the epileptic EEG-fMRI data and the performances of these two algorithms were compared. After realignment to remove the motion artifacts, the raw functional data were first compressed into 4D format, then the off-brain voxels were removed and the 2D data were constructed for ICA.

As described in the above section, the constraint was formulated as the closeness between the EEG information and the column vectors of the mixing matrix, in which the closeness was measured by correlation. The special electrodes F8 and P8, which contain the most significant seizure information, were selected as the reference signal as suggested by the clinical consultant.

For measuring the correlation, the difference in resolution between the EEG and fMRI must be resolved first, because the temporal resolution of EEG is much higher than that of fMRI time series. In order to solve this problem, the following process was performed: (1) selecting electrode P8 or F8 as the EEG reference signal; (2) filtering the reference signal by lowpass filter with cut-off frequency 15 Hz, thus ensuring that the important seizure information is kept because the dominant frequency of seizure is in the range of 2.5 Hz to 15 Hz (Lantz et al, 1999; Blanke et al, 2000); (3) down-sampling the reference EEG and up-sampling each column of the mixing matrix to ensure that they have the same data length.

3.2.3 Results and Discussion

SICA and the proposed CSICA were applied to the processed fMRI data. In each iteration of CSICA, the correlation between the column vectors of $[\mathbf{W}]^{-1}$ and the EEG reference vector was measured. Then, the unmixing matrix $\mathbf{W}$ was updated according to Eq. (16). The performances of the two algorithms were compared in terms of algorithm convergence, the correlation between the seizure EEG and the corresponding columns of the mixing matrix, and the mapping of the selected component. The selection of component was based on the prior information about the seizure area, which was the right temporal area. The epileptic seizure area was known as the right temporal area.

Figure 2 gives the algorithm convergence curve. It clearly shows that the proposed CSICA algorithm converges to the minimum of the cost function, which is less than that for unconstrained SICA. It means that CSICA attained less estimation error than SICA. Figure 3 illustrates the region of activation obtained from both algorithms. The level of activity is represented by the standard deviation (z-value). The activated area is the brain area in which the voxels have a higher z-value than the threshold level (1.5 in this experimental result). The mapping of the component (activated area) is then displayed by overlaying the area on top of the high resolution structural images. Based on the clinical expertise, the highlighted part in the left frontal area is introduced as the MRI scan noise, and the right temporal area is verified to be within the epileptic seizure area. This means that the detected BOLD is in line with the clinical findings. Also based on clinical investigations, the small patch at the right temporal region is more focused in the result obtained from CSICA.

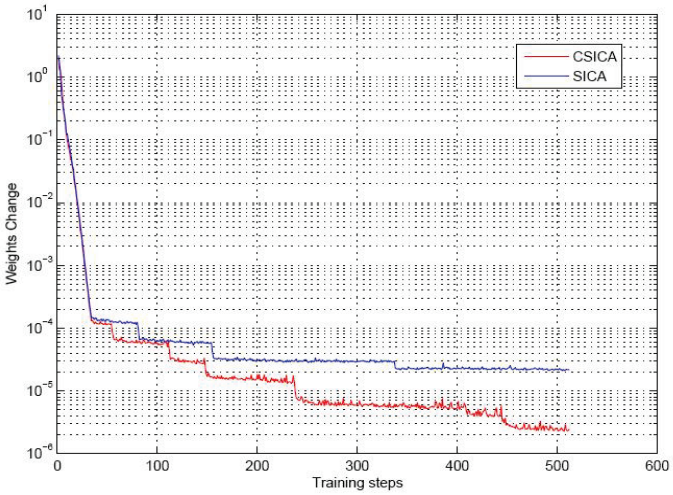

Fig. 2: Comparison of algorithm estimation error for SICA and CSICA.

The maximum correlation coefficient between the column vectors of the mixing matrix and the EEG reference vector **u** were obtained by averaging five trials for SICA and CSICA, which are 0.181 and 0.195 respectively. It shows that the results from CSICA provide a higher correlation between the seizure signal and the corresponding column vectors of $[\mathbf{W}]^{-1}$ than that from SICA.

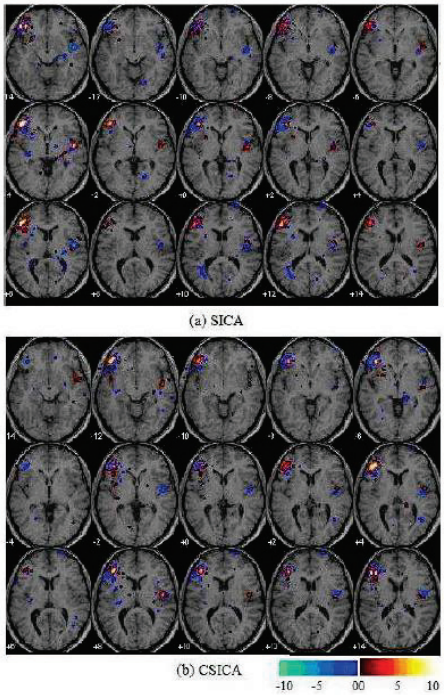

Fig. 3: The BOLD obtained from separation of fMRI data by using (a) SICA and (b) the proposed CSICA, which incorporates the EEG signals as the constraint into the update equation.

4 Concluding Remarks

In this chapter, various techniques for fMRI analysis have been reviewed and the established techniques such as GLM and SICA have been discussed in terms of their mathematical frameworks. More importantly, a novel constrained spatial ICA algorithm has been proposed which incorporates the simultaneously recorded seizure EEG into the fMRI separation process. The experimental results have shown that the BOLD region, as the result of seizure onset, has been detected using the proposed constrained SICA. This algorithm outperforms the existing unconstrained ICA algorithm in terms of estimation error and closeness between the component time course and the seizure EEG signals.

The relationship between fMRI and EEG has indeed been a challenging problem to date, therefore in this study, this relationship was simply chosen by closeness between time course of fMRI component and processed EEG signal. Further improvements to the proposed method can be achieved if a better mathematical modelling of the relationship between EEG and fMRI can be developed. Another limitation of the proposed method is that the EEG signal used in this work was the scalp EEG, which is the signal that is mixed with noise and artifacts, and is therefore not a perfect choice for seizure reference. This may be the reason why the results for the correlation measurement are not sufficiently

significant. Further investigations can be carried out by applying a proper ICA algorithm to extract the EEG seizure component. These can be an agenda for future research in this area. Nevertheless, the results presented here are still very promising, and have shown a new direction for fusion of fMRI and EEG. The idea can be further exploited in both separation and localization of seizure signals in joint EEG-fMRI signal processing.

5. Acknowledgement

The authors would like to thank Dr. Khalid Hamandi in University Hospital of Wales for his valuable advices and providing data.

6. References

Ahlfors, S. P.; Simpson, G. V.; Dale, A. M.; Belliveau, J. W.; Liu, A. K.; Korvenoja, A.; Virtanen, J.; Huotilainen, M.; Tootell, R.; Aronen, H. J. & Iimoniemi, R. J. (1999). Spatiotemporal activity of a cortical network for processing visual motion revealed by MEG and fMRI. *J. Neurophysiol.* Vol. 82, pp. 2545-2555.

Amari, S.; Chen, T. P. & Cichocki, A. (2002). Nonholonomic Orthogonal Learning Algorithms for Blind Source Separation. *Neural Computation.* Vol. 12, pp. 1463-1484.

Babajani, A. & Soltanian-Zadeh, H. (2006). Integrated MEG/EEG and fMRI Model Based on Neural Masses. *IEEE Tran on Biomedical Eng.* Vol. 53, No. 7, pp. 1794-1801.

Babiloni, F.; Babiloni, C.; Carducci, F.; Romani, G. L.; Rossini,P. M.; Angelone, L. M. & Cincottia, F. (2003). Multimodal integration of high-resolution EEG and functional magnetic resonance imaging data: a simulation study. *NeuroImage.* Vol. 19, pp. 1-5.

Beckmann, C. and Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans. Med. Imaging.* Vol. 23, pp. 137-142.

Bell, A. J. and Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation.* Vol. 7, pp. 1129 - 1159.

Blanke, O.; Lantz, G.; Seeck, M.; Spinelli, L.; Peralta, R. G. D.; Thut, G.; Landis, T. & Michel, C. M. (2000). Temporal and Spatial Determination of EEG-Seizure Onset in the Frequency Domain. *Clin. Neurophysiol.* Vol. 111, pp. 763-772.

Biswal, B. B. & Ulmer, J. L. (1999). Blind source separation of multiple signal sources of fMRI data sets using independent component analysis. *J. Comput. Assist. Tomogr.* Vol. 23, No. 2, pp. 265-271.

Calhoun, V. D.; Adali, T.; Pearlson, G. D. & Pekar, J. J. (2001). Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. *Human Brain Mapping.* Vol.13, pp. 43-53.

Calhoun, V. D.; Adali, T. & Pearlson, G. D. (2004). Independent component analysis applied to fMRI data: A generative model for validating results. *J. VLSI Signal Process. Syst.* Vol. 37, pp. 281 - 291.

Calhoun, V. D.; Stevens, M.; Pearlson, G. D. & Kiehl, K. A. (2004). FMRI analysis with the general linear model: removal of latency-induced amplitude bias by incorporation of hemodynamic derivative terms. *NeuroImage.* Vol. 22, pp. 252 - 257.

Calhoun, V. D.; Adali, T.; Stevens, M.; Kiehl, K. A. and Pekar, J. (2005). Semi-blind ICA of fMRI: A method for utilizing hypothesis-derived time courses in a spatial ICA

analysis. *NeuroImage.* Vol. 25, No. 2, pp. 527-538.

Calhoun, V. D. and Adali, T. (2006). Unmixing fMRI with independent component analysis. *IEEE Eng. in Medicine and Biology Magazine*, pp. 79-90, March/April, 2006.

Calhoun, V. and Adali, T. (2006). Fusion of Multisubject Hemodynamic and Event-Related Potential Data Using Independent Component Analysis. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) Proceedings,* Vol. 5, pp. 14-19, May, 2006.

Christman, C.; Koeppe, C.; Braus, D. F. & Flora, H. (2007). A simultaneous EEG and fMRI study of painful electric stimulation. *NeuroImage.* Vol. 34, pp. 1428-1437.

Cichochi, A. & Amari, S. (2002). *Adaptive blind signal and image processing, learning algorithm and application.* John Wiley and Sons.

Correa, N.; Adali, T.; Li, Y. & Calhoun, V. D. (2005). Comparison of blind source separation algorithms for fMRI using a new matlab toolbox: GIFT. *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, pp. 401 - 404, Philadelphia, PA, 2005.

Diehl, B.; Salek-Haddadi, A.; Fish, D. R. & Lemieux, L. (2003). Mapping of spikes, slow waves, and motor tasks in a patient with malformation of cortical development using simultaneous EEG and fMRI. *Magn. Reson. Imaging*, Vol. 21, No. 10, pp. 1167-1173.

Duann, J. R.; Jung, T. P.; Kuo, W. J.; Makeig, S.; Hsieh, J. C. & Sejnowski, T. J. (2002). Single-trial variability in event-related bold signals. *NeuroImage,* Vol. 15, No. 4, pp. 823-835.

Fan, Y.; Shen, D. & Davatzikos, C. (2006). Detecting Cognitive States from fMRI Images by Machine Learning and Multivariate Classiffication. *Computer Vision and Pattern Recognition Workshop*, pp. 89 - 89, June 2006.

Friston, K.; Holmes, A.; Worsley, K.; Poline, J. B.; Frith, C. & Frackowiak, R. (1995). Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping*. Vol. 2, pp.189 - 210.

Galit, F. A.; Sun, F. T.; Daniel, H. Desposito, M. & Knight, R. T. (2007). Spatio-temporal information analysis of event-related BOLD responses. *NeuroImage.* Vol. 34, pp. 1545 - 1561.

Horovitz, G.; Skudlarski, P. & Gore, J. C. (2002). Correlations and Dissociations Between BOLD Signal and P300 Amplitude in an Auditory Oddball Task: a Parametric Approach to Combining FMRI and ERP. *Magn. Reson. Imaging,* Vol. 20, pp. 319-325.

Horwitz, B. & Poeppel, D. (2002). How can EEG/MEG and fMRI/PET data be combined? *Hum. Brain Mapp.,* vol. 17, pp. 1-3.

Hu, D.; Yan, L.; Liu, Y.; Zhou, Z.; Friston, K. J.; Tan, C. & Wu, D. (2005). Unified SPM-ICA for fMRI analysis. *NeuroImage.* Vol. 25, No. 3, pp. 746 - 755.

Hyvarinen, A.; Karhunen, J. & Oja, E. (2001). *Independent Component Analysis.* Wiley-Interscince Publication.

Ji, Y.; Liu, H. B.; Wang, X. K. & Tang, Y. Y.(2004). Cognitive states classification from fMRI data using support vector machines. *Proc. Machine Learning and Cybernetics*, Vol 5, pp. 2919 - 2926, Aug., 2004.

Josephs, O.; Turner, R. & Friston, K. J. (1997) Event-related fMRI. *Hum. Brain Mapp.,* Vol. 5, pp. 243-248.

LaConte, S.; Strother, S.; Cherkassky, V.; Anderson, J. & Hu, X. (2005). Support vector machines for temporal classification of block design fMRI data. *NeuroImage.* Vol. 26,

pp. 317 - 329.

Lantz, G.; Michel, C. M.; Seeck, M.; Blanke, O., Landis, T. & Rose, I. (1999). 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.

Lee, T. W.; Girolami, M. and Sejnowski, T. J. (1999). Independent Component Analysis Using an Extended Infomax Algorithm for Mixed Subgaussian and Supergaussian Sources. *Neural Computation.* Vol. 11, pp. 417-441.

Lemieux, L.; Salek-Haddadi, A.; Josephs, O.; Allen, P.; Toms, N. and Scott, C. (2001). Event-related fMRI with simultaneous and continuous EEG: description of the method and initial case report. *Neuroimage,* Vol. 14, No. 3, pp. 780-787.

Li, Y.; Adal, T. and Calhoun, V. D. (2007). Estimating the number of independent components for functional magnetic resonance imaging data. *Human Brain Mapping*. Vol. 28, Issue 11, pp. 1251-1266.

Logothetis, K.; Pauls, J.; Augath, M.; Trinath, T. & Oeltermann, A. (2001). Neurophysiological Investigation of the Basis of the FMRI Signal. *Nature,* Vol. 412, pp. 150-157.

Lu, W. & Rajapakse, J. C. (2005). Approach and Applications of Constrained ICA. *IEEE Tran. on Neural Network.* Vol. 16, No. 1, pp. 203 - 212.

Mathalon, D. H.; Whitfield, S. L. & Ford, J. M. (2003). Anatomy of an Error: ERP and FMRI. *Biol. Psychol.,* Vol. 64, pp. 119-141.

McKeown, M. J.; Makeig, S.; Brown, G. G.; Jung, T. P.; Kindermann, S. S.; Bell, A. J. & Sejnowski, T. J. (1998). Analysis of fMRI data by blind separation into independent spatial components. *Hum. Brain Map.* Vol. 6, No. 3, pp. 160 - 188.

McKeown, M. J. & Sejnowski, T. J. (1998). Independent component analysis of fMRI data: Examining the assumptions. *Hum. Brain Map.* Vol. 6, No. 5, pp. 368 - 372.

McKeown, M. J.; Hansen, L. K. & Sejnowsk, T. J. (2003). Independent component analysis of functional MRI: What is signal and what is noise? *Curr. Opin. Neurobiol.,* Vol. 13, No. 5, pp. 620-629.

Phillips, C.; Rugg, M. D. & Friston, K. J.(2002). Anatomically Informed Basis Functions for EEG Source Localization: Combining Functional and Anatomical Constraints. *NeuroImage,* Vol. 16, pp. 678-695.

Price, C. J. & Friston, K. J. (1997). Cognitive conjunction: a new approach to brain activation experiment. *NeuroImage.* Vol. 5, pp. 261-270.

Reidl, J.; Starke, J.; Omer, D. B.; Grinvald, A. & Sporsa, H.(2007). Independent component analysis of high-resolution imaging data identifies distinct functional domains. *NeuroImage.* Vol. 34, pp. 94 - 108.

Rorden, C. & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology*, Vol. 12, pp. 191-200.

Salek-Haddadi, A.; Merschhemke, M.; Lemieux, L. & Fish, D. (2002). Simultaneous EEG-Correlated Ictal fMRI. *Neuroimage,* Vol. 16(1), pp. 32-40.

Salek-Haddadi, A.; Friston, K. J.; Lemieux, L. & Fish, D. R. (2003). Studying spontaneous EEG activity with fMRI. *Brain Research Reviews*. Vol. 43, pp. 111-133.

Schilling, R. J.; Carroll, J. J. & Al-Ajlouni, A. F.(2001). Approximation of nonlinear systems with radial basis function neural networks" *IEEE Tran. Neural Network,* Vol. 12, pp. 1-15.

Stone, J. V.; Porrill, J.; Porter, N. R. & Wilkinson, I. D. (2002). Spatiotemporal independent

component analysis of event-related FMRI data using skewed probability density functions. *NeuroImage.* Vol. 15, No. 2, pp. 407- 421.

Suzuki, K.; Kiryu, T. & Nakada, T. (2002). Fast and precise independent component analysis for high field fMRI time series tailored using prior information on spatiotemporal structure. *Hum. Brain Map.* Vol. 15, No. 2, pp. 407-421.

Wang, Z.; Childress, A. R. and Detre, J. A. (2006). Boost up the detection sensitivity of ASL perfusion fMRI through support vector machine. *The 28th IEEE Engineering in Medicine and Biology Society (EMBS)*, pp. 1006-1009, Aug., 2006.