Hand Motor Imagery Classification Using Effective Connectivity and Hierarchical Machine Learning in EEG Signals

Maghsoudi A.¹, Shalbaf A.²*

ABSTRACT

Background: Motor Imagery (MI) Brain Computer Interface (BCI) directly links central nervous system to a computer or a device. Most MI-BCI structures rely on features of a single channel of Electroencephalogram (EEG). However, to provide more valuable features, the relationships among EEG channels in the form of effective brain connectivity analysis must be considered.

Objective: This study aims to identify a set of robust and discriminative effective connectivity features from EEG signals and to develop a hierarchical machine learning structure for discrimination of left and right hand MI task effectively.

Material and Methods: In this analytical study, we estimated effective connectivity using Granger Causality (GC) methods namely, Generalized Partial Directed Coherence (GPDC), Directed Transfer Function (DTF) and direct Directed Transfer Function (dDTF). These measures determine the transient causal relation between different brain areas. Then a feature subset selection method based on Kruskal–Wallis test was performed to choose most significant directed causal connection between channels. Moreover, the minimal-redundancy-maximal-relevance feature selection method is applied to discard non-significance features. Finally, support vector machine method is used for classification.

Results: The maximum value of the classification accuracies using GC methods over different frequency bands in 29 subjects in 60 trial is approximately 84% in Mu (8−12 Hz) - Beta1 (12−15 Hz) frequency band using GPDC method.

Conclusion: This new hierarchical automated BCI system could be applied for discrimination of left and right hand MI tasks from EEG signal, effectively.

Keywords
Electroencephalography; Motor Imagery; Effective Connectivity; Machine Learning; Brain-Computer Interfaces

Introduction

Brain Computer Interface (BCI) is a modern technology providing a communication between the brain and external environment. BCIs have been used for severely paralyzed patients as a communication option, an augmentative tool for rehabilitation and assistive technology [1, 2]. These systems offer effective assistance not only for those with motor disabilities, but also for healthy users such as computer game control [3] and mobile robots [4]. For practical applica-
tions in BCIs, electroencephalogram (EEG) is well accepted due to high temporal resolution information, and a noninvasive, inexpensive, portable method. A type of BCI known as motor imagery (MI) refers to the act of imagining a particular action without actual execution.

In the last decade, several signal processing techniques from the one-channel EEG have been proposed for developing BCIs [5-7] and specially MI-BCIs, namely, power spectral density using Fourier transform [8, 9], discrete wavelet transform [10], calculation of the autoregressive model coefficients [11], common spatial pattern [12, 13], sparse representation [14, 15], Hilbert–Huang transform [16], empirical mode decomposition [17, 18], and Hjorth parameters [19, 20]. Despite of significant achievements of the aforementioned methods, none of them have been proved to be adequately reliable in the practical settings because of using EEG features from individual channels and ignoring valuable information inherent between channels. Moreover, since EEG signals are non-stationary and very sensitive to noises, a single channel can hardly achieve a good accuracy and the extracted EEG patterns, based on multi-channel, must be considered to detect the dynamic characteristics of the EEG signals. MI task causes complicated neurophysiological changes and consequently, it is expected that for solving the aforementioned limitations in MI of different limbs, we consider the relationships among brain regions with different connectivity patterns.

The brain connectivity analysis has three general subdivisions, as follows [21]: (1) the structural connectivity (2), the functional connectivity, and (3) the effective connectivity. Among these fields, effective connectivity is a significant tool for the EEG analysis. A popular method for estimating effective connectivity is Granger Causality (GC) which is a data-driven approach [22]. Theoretically, GC characterizes the directed causal interaction among neural time-series as strong as causal mechanistic couplings. Actually, this method is used for analysis and visualization of time and frequency of multivariate dependent directed information flow and causality between localized EEG channels.

After the EEG feature extraction, the EEG signal is classified using machine learning methods. In the last decade, a wide variety of machine learning, including feature selection and classification methods have been used to EEG for developing MI-BCIs, such as, principal component analysis [23, 24], independent component analysis [25], sequential floating forward search [26], an evolutionary algorithm [27] and relative entropy criteria for feature selection [28], linear discriminant analysis [29, 30], multilayer perceptron network [31], radial basis function network [32], Support Vector Machine (SVM) [33-36], least squares classifier [37], Bayesian classifier [38], adaptive neuro-fuzzy classifier [39, 40], extreme learning machine [41], Naive bayes [42], sparse bayesian extreme learning machine [43], sparse group representation model [44] and deep learning approaches [45-47]. Despite the different machine learning algorithms, there hasn’t been a universally superior one for this application.

The aim of this study is to find a set of discriminative effective connectivity features from EEG signals and to design a hierarchical feature selection and classification methods for discrimination of left and right hand MI task. The ability of this novel system is evaluated with 29 subjects.

Material and Methods

A. Subjects and Data Acquisition

In this analytical study, 29 healthy subjects with no reported brain-related diseases participated [48]. EEG data was recorded at 1000 Hz sampling rate with thirty electrodes according to the international 10-5 system with Fz as the ground electrode. The subjects sat on an armchair in front of a 50-inch white screen with
distance of 1.6 m and without any movement of the body.

The experiment included three sessions of left and right hand MI. Each session comprised a 1 min pre and post-experiment resting period, and 20 repetitions of the given task (10 trials for each left and right hand MI). The task started with 2 s of a visual introduction of the task, followed by 10 s of a task period and 15 to 17 s resting period randomly. Figure 1 shows the diagram of the experimental paradigm. In the task period, subjects imagine the opening and closing their hands with a speed of 1 Hz. Therefore, for each subject in the whole three sessions with 30 trials for left and also 30 trials for right hand MI were performed.

Figure 1: Schematic diagram of the data acquisition [48].

B. Effective connectivity

Effective connectivity is an important aspect of modern neuroscience due to its potential to describe interactions between brain areas and in other words representation of the direction and strength of the information flow between different brain areas [49]. A popular statistical method for estimating effective connectivity is Granger causality (GC) which is a data-driven approach [50].

GC in the frequency domain results in the analysis of EEG frequencies bands [51]. This method is based on estimation of parameters of Multi-Variable Auto-Regressive (MVAR) model for an individual signal data. Quantitative spectral measures are: Generalized Partial Directed Coherence (GPDC) [52, 53], Directed Transfer Function (DTF) [54], and Direct DTF (dDTF) [55]. These techniques characterize the direction and spectral properties between any pair of channels in EEG signals simultaneously. By these measures, it’s possible to quantify the value of each possible electrode combination per specific frequency range. We extract each frequency range for each measure by averaging theta (4 – 7 Hz), alpha or mu (8 – 12 Hz), beta1 (12 – 15 Hz), beta2 (15 – 22 Hz), beta3 (22 – 30 Hz) and gamma (30–45 Hz). All calculations were done using MATLAB (The Mathworks, Inc., Natick, MA, USA) via the open-source SIFT toolbox [56].

C. Statistical Analysis

Due to the limited dataset, k-fold cross validation was used. This method partitions the dataset into k equal sized subsamples (in this paper, k is set to 10). In each trial, the classification structure is constructed with 90 percent of data and evaluated with the remaining data as testing data. The process is repeated 10 times, each subsample used exactly once as the testing data, until all the dataset has been used for testing. Evaluation performance is reported by averaging the 10 results from each subsample. Moreover, the 10-fold cross validation is done for 100 consecutive runs and the average of the results is calculated. The advantage of this method compared with repeated random sub-sampling is the use of all
dataset for both training and testing.

D. Preprocessing

First, the EEG data was re-referenced to common average reference, then a fourth order of Chebyshev type II with a passband of 2-45 Hz was applied to the signal. Finally, electrooculogram (EOG) artifact was rejected using independent component analysis (ICA) method via the automatic artifact rejection toolbox in EEGLAB [57].

E. Feature selection method

The significance of the extracted features from effective connectivity methods between right and left hand MI groups is tested by the Kruskal-Wallis method [58]. Then, the non-significant features with \( p > 0.01 \) are discarded. After that, the Minimum Redundancy Maximum Relevance (mRMR) method [59, 60] is used to select the best features. In this method, features can be selected far away from each other while still having the highest relevance to the target classes. This selection method is more powerful than the only maximum relevance selection methods.

F. SVM classification

SVM is one of the powerful tools for classification issues. In this method, the parameters of the separator function should be designed so that the margin between the hyperplane becomes maximum [61]. SVM maps the input space to a higher dimension in order to unfold the complexity existing in the dataset using a proper kernel function and after that, a linear decision surface is designed to identify the labels with fewer complications. Appropriate function is Radial Basis Function (RBF) and the parameters are determined in the optimization process.

Results

Several measures of effective connectivity using GC methods (GPDC, DTF, dDTF) and parameters of calculated MVAR model were estimated based on each signal frequency band ranges \([\text{theta} (4 – 7 \text{ Hz}), \text{alpha or mu} (8 – 12 \text{ Hz}), \text{beta1} (12 – 15 \text{ Hz}), \text{beta2} (15 – 22 \text{ Hz}), \text{beta3} (22 – 30 \text{ Hz}) \) and \( \text{gamma} (30–45 \text{ Hz}) \]) in 29 subjects (30 left hand MI and 30 right MI task for each subject). Best parameters for MVAR model fitting were selected according to autocorrelation function and portmanteau tests. The optimized parameters are: 5 s of window length, and model order of 60. Having 30-channel EEG, 900 \((30 \times 30)\) directed causal connection between channels as effective connectivity features are extracted for each GC method in each frequency band range that make computation complex. As a result, a feature subset selection based on Kruskal-Wallis statistical test is done to choose most significant features for discrimination of right and left hand MI tasks. Using this test, the nonsignificant features, which result in \( p > 0.01 \), are discarded. Then, using the mRMR algorithm via 5-fold cross validation, the best features remained after the Kruskal-Wallis test are selected. This method chooses the features with the minimum redundancy among the selected features and the highest relevance to the target classes. Finally, the best selected features are fed to SVM classification structure. The classification procedure aims to accurately classify EEG data into left and right hand MI tasks in 29 subjects. A 10-fold cross-validation was performed to evaluate the classification performance. RBF is selected as the kernel function of SVM with the optimal sigma value \( (\sigma) \) of 0.90 by try and error methodology. The schematic diagram of the proposed MI-BCI system is shown in Figure 2.

All data processing were calculated separately over segments of 5 s. The classification accuracies obtained from the proposed method using GC (GPDC, DTF, dDTF) connectivity measures over all subjects for 0-5 s and 5-10 s for each frequency band range are given in Table 1, separately. As seen, GPDC yields the best results with high accuracy in classifying right and left hand MI task within Mu (83.87)
and beta1 (83.05) frequency bands of EEG for 0-5 s. However, other GC methods (DTF, dDTF) and other frequency bands of EEG (theta, beta2, beta3, and gamma) are not able to categorize right and left hand MI tasks with significant accuracy. It should be noted that at a time interval of 0 - 5s due to its inherent fast responses of brain electrical activity, the EEG is the best separable.

Raw 900 (30×30) connectivity features for
the best results of GC method with high accuracy (GPDC) for Mu, and Beta1 frequency band over all subjects in 0-5 s for left and right hand MI task are shown in Figure 3. A higher absolute value of connectivity feature is shown in warm colors. Moreover, the scalp topographies of the averages of normalized log (p-value) obtained from the best selected of GPDC connectivity features (68 connectivity) in feature selection procedure for 0-5 s in Mu and Beta1 frequency band over all subjects are shown in Figure 4. A higher absolute value of log (p-value) is shown in warm colors i.e. a better separability with higher significance between left and right hand MI task. As seen, during the MI task, EEG generally show good separation around the motor cortex and high separation around frontal and parietal areas (Figure 4).

**Discussion**

In this paper, we proposed a new automated algorithm for discrimination of left and right hand MI tasks from EEG signal, based on a set of discriminative features using GPDC method and two hierarchical feature selection methods and finally SVM classification structure. This algorithm could classify the EEG data in 29 subjects in 60 trials with an overall accuracy of 84% during Mu-Beta1 frequency band, effectively.

Results of directional connectivity metrics

**Figure 3:** Raw 900 (30×30) Generalized Partial Directed Coherence (GPDC) connectivity features for Mu, and Beta1 frequency band over all subjects for 0-5 seconds for left and right hand Motor Imagery (MI) task. A higher absolute value of connectivity feature shows with warm colors. Thirty electrodes are as follow: F7, AFF5h, F3, AFp1, AFp2, AFF6h, F4, F8, AFF1h, AFF2h, Cz, Pz, FCC5h, FCC3h, CCP5h, CCP3h, T7, P7, P3, PPO1h, POO1, POO2, PPO2h, P4, FCC4h, FCC6h, CCP4h, CCP6h, P8, T8.
Connectivity Based Motor Imagery Classification

Figure 4: Normalized-log (p-value) obtained from the best selected of Generalized Partial Directed Coherence (GPDC) connectivity features (68 connectivity) in feature selection procedure for Mu and Beta1 frequency band over all subjects for 0-5 seconds. A higher absolute value of log (p-value) shows with warm colors and means a better separability between left and right hand Motor Imagery (MI) task. Directional connectivity represented by the arrows.

(GPDC features) in multichannel EEG in the present study infer that information flow from different parts of the brain to the others with different direct paths plays an important role in differentiation of right and left hand MI tasks. As results shown in Figure 4, differential patterns of connectivity in GPDC method is around the motor areas and frontal and parietal areas. Moreover, the accuracies achieved with our approach (84%) is higher than other method that uses EEG features from individual channels (70%) [48] in the same database.

Previous studies have shown that when a person is performing imagining the left or right hand movements, the mu and beta1 rhythms of EEG signals as a neural oscillation of brain electrophysiological activity at the sensorimotor area are (de)synchronized [8-10, 13, 48, 62, 63]. Thus, this study on MI hand movement pattern discrimination is in accordance with previous results.

To mention a limitation of our study, we believe that the performance of a multi-modal BCI system based on EEG and near-infrared spectroscopy (NIRS) might improve accuracy. Complementary, information measured by these methods is capable to improve the performance of either method and finally the performance of the system of discrimination of right and left hand MI tasks might be improving.

Conclusion

This study addresses a new method based on effective connectivity quantified with GPDC method and a hierarchical machine learning structure methods for discrimination of left and right hand MI tasks from EEG signals. Results indicate that exploring causal dependencies between brain regions of subjects using directed information flow plays an important role and has potential discriminative value. This new system could reach the accuracy of 84% on the MI EEG data of 29 subjects within Mu-beta1 frequency band.

Conflict of Interest

None

References

2. Wolpaw JR, Birbaumer N, McFarland DJ,
Maghsoudi A. and Shalbaf A.


Connectivity Based Motor Imagery Classification


Maghsoudi A. and Shalbaf A.


