MUTUAL INFORMATION BASED RELEVANCE NETWORK ANALYSIS: A PARKINSON'S DISEASE STUDY

Pamela Wen-Hsin Lee, Z. Jane Wang *

Department of Electrical and Computer Engineering
University of British Columbia
Canada

Martin J. McKeown

Pacific Parkinson's Research Centre
Brain Research Centre
University of British Columbia
Canada

ABSTRACT

Monitoring the dynamics of networks in the brain is of central importance in normal and disease states. Current methods of detecting networks in the recorded EEG such as correlation and coherence only explore linear dependencies, which may be unsatisfactory. We propose applying mutual information as an alternative metric for assessing possible nonlinear statistical dependencies between EEG channels. However, EEG data are complicated by the fact that data are inherently non-stationary and also the brain may not work on the task continually. To address these concerns, we propose a novel EEG segmentation method based on the temporal dynamics of the cross-spectra of computed Independent Components. A real case study in Parkinson's disease and further group analysis employing ANOVA demonstrate different brain connectivity between tasks and between subject groups and also a plausible mechanism for the beneficial effects of medication used in this disease. The proposed method appears to be a promising approach for EEG analysis and warrants further study.

Index Terms— EEG, Cross-Spectrogram, Mutual Information, Parkinson's Disease, Relevance Network.

1. INTRODUCTION

Connectivity between brain regions is being increasingly recognized as important for normal brain functioning, and may be impaired in some neurological diseases such as Parkinson's disease (PD).A variety of electrophysiological techniques have been available for examining connectivity such as electroencephalogram (EEG). Both linear and non-linear measures have been proposed to infer the functional connectivity between and within brain hemispheres from the EEG.

Several methods such as coherence and correlation have been applied to the study of cortical connections [1, 2]. However, these methods consider only linear dependencies. In contrast, mutual information (MI) measures both the linear and non-linear statistical dependencies between time series and thus can be used to find the dynamical coupling or information transmission in the brain. Its value is maximized when the two time series are the same and is zero when they are completely independent.

The aim of this study is to investigate the functional connectivity of the brain of PD patients and normal subjects by estimating the MI in multi-channel EEG. In order to overcome the issues that only a small fraction of the EEG recording is task-related and the non-stationary nature of the EEG data, we propose a novel segmentation method of the EEG based on the temporal dynamics of the cross-spectrogram of the Independent Components (ICs). The network differences between tasks and between groups are analyzed statistically.

Previously, MI has been applied to Alzheimer's [3], Schizophrenic [4], and Odor Stimulation [5] data to assess information transmission between different parts of the brain and it's connectivity. However, to our knowledge, this is the first time that anyone has first segmented the data based on ICs or applied MI to PD data in order to gain insight into the difficulty that PD subjects face when performing simultaneous movements.

The paper is organized as follows: the methods are introduced in section 2. Section 3 presents the EEG experiment and summarizes the results in a real case study of PD. Section 4 concludes the paper.

2. METHODS

2.1. Preprocessing

Fig. 1 presents the flowchart of the steps of EEG preprocessing. The data are band-pass filtered to focus on the frequency ranges of clinical and physiological interests, typically 0.5-55Hz for clinical EEG [6]. EEG data are frequently contaminated with artifacts such as eye movements, eye blinks, cardiac signals, and muscle noise. Independent Component Analysis (ICA) can be used to isolate these artifacts because they are temporally independent from ongoing brain activity.

Since only a small percent of the entire EEG recording is task-related and the inherit non-stationary nature of EEG data [7], EEG segmentation based on the cross-spectrogram of the ICA components is adapted in order to address such concerns. The use of ICs requires further comment. If the derived components were truly independent, then the cross-spectrum would not be significant. However, in real data many of the assumptions of ICA are violated. The data are not stationary, and the time courses are not spatially white. By using infomax ICA, which does not incorporate time delays, the derived components will be maximally independent at zero lag. As such, it will deal with the problem of volume conduction – where a deep electrical source may impart common electrical activity to two or more electrodes. However, by examining the ICs within a short moving window, the non-stationary nature of the EEG will be explored, and significant dependencies between ICs become apparent [13].

The ICs are thus transformed into time-frequency domain and the cross-spectrogram is computed. A short (3s) time window shifted by 0.5s is applied to the frequency content to obtain the localized time information. The frequency content is computed

*Thanks to NSERC agency for funding.

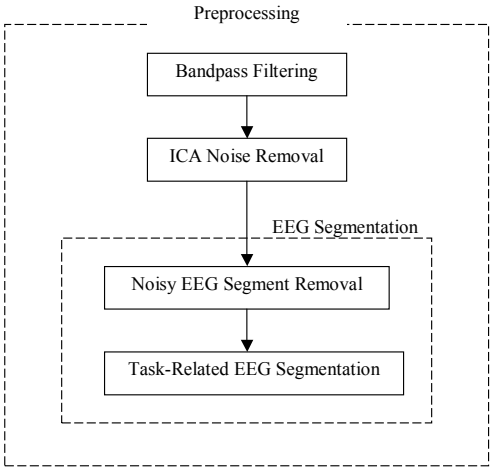


Fig. 1. Flowchart of EEG preprocessing

by cross power spectral density using the Welch's averaged, modified periodogram [8] method of spectral estimation. If x(t) and y(t) are signals normalized to zero mean and unit variance, their cross-correlation is:

$$R_{xy}(\tau) = \frac{1}{N-\tau} \sum_{k=1}^{N-\tau} x(k+\tau)\, y(k) \quad (1)$$

The cross spectral density function is a Fourier transform of the cross-correlation function that indicates the relationship between the two signals.

$$S(f) = \int_{-\infty}^{\infty} R_{xy}(t) e^{-j2\pi ft} dt \quad (2)$$

Based on the cross-spectrogram, we can identify the noisy and the task-related segments.

2.2. Mutual Information based Relevance Network

The relevance network takes large data sets of experimental data and and graphically depicts the result of pair-wise mutual information. Mutual information (MI) measures the mutual dependence or information gained about one signal from another. Given two random variables X and Y, the amount of uncertainty in X that is reduced by knowing Y is the mutual information,

$$I(X,Y) = H(X) - H(X|Y) = H(X) + H(Y) - H(X,Y), \quad (3)$$

which can also be written as

$$I(X,Y) = \sum_{x,y} P_{XY}(x,y) log_2 \frac{P_{XY}(x,y)}{P_X(x) P_Y(y)}, \quad (4)$$

where $H(X)$ means the entropy of X, $H(X|Y)$ means the conditional entropy of X when given Y, $H(X,Y)$ is the joint entropy that is the joint uncertainty in X and Y, and $P_{XY}(x,y)$ is the joint probability density for the measurements of X and Y that produce the values x and y.

MI is also the amount of information about X that Y contains. It is a symmetric function meaning $I(X,Y) = I(Y,X)$. If X is independent of Y, then $P_{XY}(x,y)$ factorizes to $P_X(x) P_Y(y)$ and the MI is zero. On the contrary, the higher the MI between two signals, the more information they contain about each other. Hence, the higher MI, the more likely that the two signals are biologically related. MI is estimated from a finite number of samples, the probability densities, $P_X(x)$ and $P_{XY}(x,y)$, are approximated by histogram. The detailed derivation and background of information theory can be found in [9].

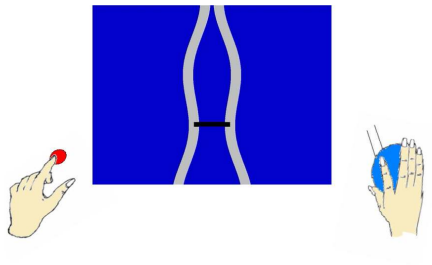


Fig. 2. Experiment design of the squeeze, button push, and simultaneous movement task

2.3. Statistical Analysis

The normality of the distribution of the MI value is tested by the Kolmogorov-Smirnov (KS) test. Within group differences of each MI value are analyzed using one way analysis of variance (ANOVA) with a task factor. Across group differences of each MI value are analyzed using ANOVA with a group/ subject factor. Because we are more interested in the significant connections with greater magnitude, a threshold is further applied. The threshold is chosen from the distribution of the permutation result [10] of the MI values.

3. EXPERIMENT AND RESULTS

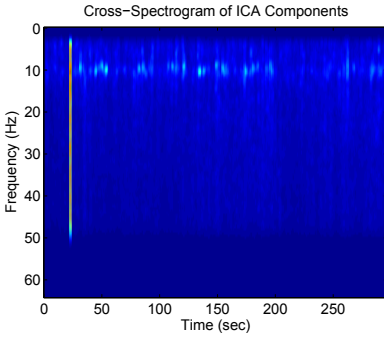
3.1. Subjects and Experiment Design

Seven PD and six age-matched control subjects volunteered to participate in the study. Subjects were asked to hold a custom-built rubber squeeze bulb in their right hand. They were instructed to control an "inflatable" ring as shown as the horizontal bar in Fig. 2 by squeezing the bulb. The ring must move through an undulating tunnel without touching the sides. Three five-minute trials were performed by normal subjects (N), PD before medication (Ppre), and PD 1hr after L-dopa medication (Ppost). During one trial, subjects were asked to squeeze the bulb (SQ) with right hand alone. In another trial, they were asked to press a mouse button (BU) with their left hand alone when they observed a color changes in the ring. Lastly, the subjects were required to do both movements simultaneously (BO) as seen in Fig. 2.

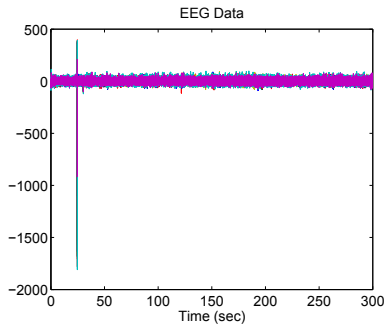
3.2. EEG Data Preprocessing

The 19-channel EEG data are sampled at 128Hz and bandpass filter between 0.5∼55Hz. The artifactual ICA components are removed from the data by visual inspection. Two operations are done based on the cross-spectrogram of the ICA components: noisy EEG segment removal and task-related EEG segmentation.

Noisy EEG Segment Removal: The data in the frequency range of 45∼55Hz do not contain information of clinical and physiological interest. Thus, activity in the cross-spectrogram of the ICA components between 45∼55Hz was a good marker of transient broadband artifacts that were not eliminated by ICA Noise Removal step (Fig.

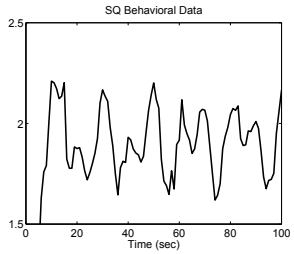


(a) Cross-Spectrogram of a pair of ICA components

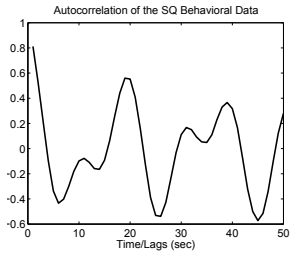


(b) Corresponding EEG Data

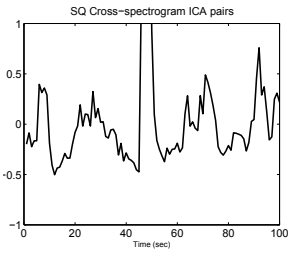
Fig. 3. Cross-spectrogram and EEG



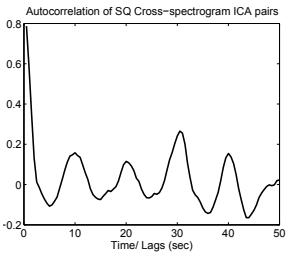
(a) Squeeze task: the behavioral data



(b) Squeeze task: the autocorrelation of the behavioral data



(c) Squeeze task: the sum of all behaviorally relevant cross-spectra between ICA components pairs



(d) Squeeze task: the autocorrelation of the sum of cross-spectra between ICA component pairs

Fig. 4. Example of the original and autocorrelation of the squeeze behavioral data and the summ of all cross-spectrogram ICA component pairs integrated between 8∼12Hz

3). Thresholding is applied to eliminate EEG segments with greater power in this frequency range and exclude them from further analysis.

Task-Related EEG Segmentation: The segmentation and network analysis are based on three frequency bands: 5∼8Hz (Delta band), 8∼12Hz (Theta band), 12∼30Hz (Beta band) as these have in the past been implicated with different neurological processes [6]. Since not all the cross-spectrogram of the ICA component pairs are task-related, only the ones that are modulated with the behavioral data are used. The squeezing task movement is composed of 2 sinusoids with period of 10 and 18 seconds. The button was pressed every 20 seconds. The simultaneous movement thus contains both the squeeze and button push information.

Because the pairs that are modulated with the behavioral data all have a common task-related feature, they are integrated to enhance the signal-to-noise for segmentation. As an example, both the autocorrelation of the behavioral data and the autocorrelation of the integrated cross-spectrogram ICA component pairs contain peaks around every 10 and 18 seconds as seen in Fig. 4(b) and 4(d), yet there are some discrepancies between their actual time courses in Fig. 4(a) and 4(c). In general the behavioral data cannot reflect exactly how the brain functionally relates to the task, although in this case, we can use the smooth behavioral performance as a rough check on the validity of the proposed segmentation approach. The discrepancies include the actual magnitude of the correlation and delays in information transmission. EEG segmentation is thus based on the actual cross-spectrogram of the ICA pairs directly from the brain for more accurate segmentation. Depending on the features of each dataset, approximately five pairs are chosen for each task. The EEG segments that have higher power indicate that they are highly correlated for the particular task they are performing. We therefore threshold the EEG data every 20 second to obtain the task-related segments. Only segments that are above the mean plus the mean absolute deviation are selected for further analysis.

3.3. Mutual Information based Relevance Network Analysis

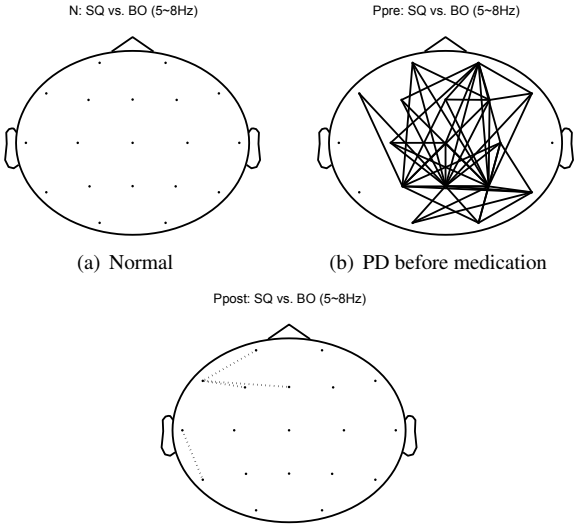
The data are broken into 4 second epochs for MI computation in order to increase the sample size, and enhance the stationarity and consistency of the MI estimates. A two-tailed P value from the ANOVA ($p < 0.001$) is considered significant for the within group analysis (SQ vs. BO) and the same goes for the across group analysis (N vs. Ppre vs. Ppost). The permutation test is used in conjunction with ANOVA test to select relevant features for the MI based relevance network. We have taken the maximum value of the permutation distribution as our threshold to select the significant connections for the interpretation.

We first examine the results for the within group analysis. As an illustrative example, the graphical results for the within group at the frequency range of 5 to 8 Hz are presented in Fig. 5. The details of the comparison for the PD group after medication are summarized in Table 1. The results suggest that PD subjects are unable to independently recruit different areas while performing simultaneous tasks; they attempt to recruit a single coherent widespread network. Additionally, medication appears to partially normalize this deficiency, consistent with prior studies investigating corticomuscular coherence in PD [12].

We also investigate the results for across group analysis. Due to the space limitation, we only report an illustrative example here. The results for across group analysis for BO are shown in Fig. 6. We observe higher MI values in the frontal region of the brain for lower

Table 1. Ppost 5∼8 Hz: SQ vs. BO

Node #	Node #	P value	SQ mean	SQ STD	BO mean	BO STD
1	3	0.00181	0.18402	0.08481	0.16174	0.06261
3	4	0.00307	0.18680	0.07635	0.16781	0.05615
3	5	0.00210	0.14559	0.07767	0.12510	0.06041
8	13	0.00486	0.12596	0.04397	0.11383	0.04610



(a) Normal (b) PD before medication

(c) PD after medication

Fig. 5. Mutual information based relevance network for within group analysis: squeeze (SQ) vs. both (BO). The solid line denotes that MIs of BO are significantly greater than the ones of SQ. The dotted line means MIs of SQ are significantly greater than the ones of BO.

and medium frequencies and motor cortex for higher frequencies in PD.

3.4. Conclusion

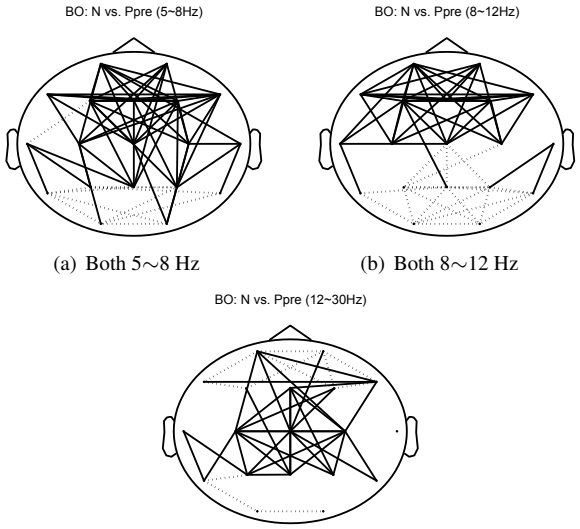
This paper proposed a mutual information based relevance network approach for brain connectivity in EEG data analysis. Since mutual information measures both the linear and nonlinear dependencies, it is considered a suitable metric for identifying brain connectivity between EEG times series. EEG segmentation was applied to the data in order to extract the task-related EEG segments. A real case study of Parkinson's disease yielded promising results, as it is able to identify different brain network patterns between tasks and between groups characterized by different factors such as disease and medication.

4. REFERENCES

[1] Shen B., Nadkarni M., and Zappulla R.A., "Spectral modulation of cortical connections measured by EEG coherence in human.," *Clinical Neurophysiology*, vol. 110, pp. 115-125, 1999.

[2] Weiss S., and Rappelsberger P., "Long-range EEG synchronisation during word encoding correlates with successful memory performance," *Cognitive Brain Ressearch*, vol. 9, pp. 299-312, 2000.

[3] Jaeseung Jeong, John C. Gore, and Bradley S. Peterson, "Mutual information analysis of the EEG in patients with Alzheimers disease," *Clinical Neurophysiology*, vol. 112, pp. 827-835, 2001.



(a) Both 5∼8 Hz (b) Both 8∼12 Hz

(c) Both 12∼30 Hz

Fig. 6. Mutual information based relevance network for across group analysis: normal (N) vs. PD before medication (Ppre). The solid line denotes that MIs of Ppre are significantly greater than the ones of N. The dotted line means MIs of N are significantly greater than the ones of Ppre.

[4] Sun Hee Na, Seung-Hyun Jin, Soo Yong Kim, and Byung-Joo Ham, "EEG in schizophrenic patients: mutual information analysis," *Clinical Neurophysiology*, vol. 113, pp. 1954-1960, 2002.

[5] B.C. Min, S.H. Jin, I.H. Kang, D.H. Lee, J.K. Kang, S.T. Lee, and K. Sakamoto, "Analysis of Mutual Information Content for EEG Responses to Odor Stimulation for Subjects Classified by Occupation," *Chem. Senses*, vol. 28, pp. 741-749, 2003.

[6] G Pfurtscheller and FH Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clinical Neurophysiology*, vol 110, pp. 1842-1857, 1999.

[7] Alexander Ya. Kaplan, Andrew A. Fingelkurts, Alexander A. Fingelkurts, Sergei V. Borisov, and Boris S. Darkhovsky,"Nonstationary nature of the brain activity as revealed by EEG/MEG: Metholdological, practical and conceptual challenges" *Signal Processing*, vol. 85, pp. 2190-2212, 2005.

[8] Alan V. Oppenheim and Ronald W. Schafer, *Discrete-Time Signal Processing*, Upper Saddle River, NJ: Prentice-Hall, pp. 737, 1999.

[9] Cover TM and Thomas JA, *Elements of information theory*, New York: John Willey & Sons, 1991.

[10] D. Francois, V. Wertz, and M. Verleysen, "The permutation test for feature selection by mutual information," *ESANN Proceedings*, 26-28 April 2006.

[11] Hoehn M and Yahr M,"Parkinsonism: onset, progression and mortality," *Neurology*, vol 17, pp. 427-442, 1967.

[12] S. Salenius, S. Avikainen, S. Kaakkola, R. Hari, and P. Brown, "Defective cortical drive to muscle in Parkinsons disease and its improvement with levodopa," *Brain*, vol 125, pp. 491-500, 2002.

[13] D. Gupta and C. J. James,"Narrowband vs. Broadband Phase Synchronization Analysis Applied to Independent Components of Ictal and Interictal EE", *IEEE-EMBS Proceedings*, Aug 2007.