Research on the MEG of depression patients based on multivariate TE partial information decomposition

Zihan Chen¹, Yunjie Fang¹, Wei Yan²,Jun Wang¹,*, Jin Li³,b and Fengzhen Hou⁴,c

¹Smart Health Big Data Analysis and Location Services Engineering Lab of Jiangsu Province Nanjing University of Posts and Telecommunications, Nanjing, China
²Department of Psychiatry, The Affiliated Brain Hospital of Nanjing Medical University, Nanjing, China
³College of Physics and Information Technology Shaanxi Normal University, Xi’an, China
⁴Key Laboratory of Biomedical Functional Materials, China Pharmaceutical University, Nanjing, China

*Corresponding author e-mail: wangj@njupt.edu.cn
²yanwei@njmu.edu.cn, ³lijin1997@snnu.edu.cn, ⁴houfz@cpu.edu.cn

Abstract. Transfer entropy (TE) has been broadly used in the field of neurosciences. In this paper, the partial information decomposition algorithm is employed to decompose multivariate TE into synergistic, redundant and unique parts. In this work, the synergistic part is believed as more suitable as the computation method. We recorded the magnetoencephalogram (MEG) data of 6 subjects with depression and 13 normal subjects under different emotional stimulations, and studied the coupling between multiple symmetric channels in the frontal area in the brain of subject. The experimental results show that under different emotional stimulations, normal people present significant difference from the depression patients, especially in the right frontal area. Furthermore, under negative emotional stimulation, the difference in synergistic value between normal people and depression patients is smaller. The synergistic value of depression patients has become bigger, which indicates that the brain complexity of depression patients has grown, and their brain activities have increased.

1. Introduction

Depression is a disease of affective disorder. According to the report of the World Health Organization (WHO), approximately 800,000 people commit suicide due to depression every year, and it will become the major disease in the world in 2030[1]. The main clinical manifestations of depression include long-term low spirits, loss of interest in many things, insomnia and even self-harm[2]. Therefore, solution to depression has attracted broad concern worldwide.

In recent years, the MEG technology has been widely used in the clinical diagnosis of depression. The core of MEG is the superconducting quantum interference device (SQUID), which measures the brain magnetic field signal generated by the current within the nerve cells according to the superconductive Josephson effect, and then, by combining it with the MRI (magnetic resonance imaging) or other image information, it can directly reflect the activity of nerve cells[3]. Compared
with SPECT (single-photon emission computed tomography) and EEG (electroencephalogram), MEG has higher sensitivity and higher temporal and spatial resolutions, which is safe and reusable[4]. At present, MEG has broadly used in researches on complicated brain diseases, such as schizophrenia and epilepsy.

2. Multivariate TE Partial information decomposition

With the continuous progress of science and technology, research on nonlinear dynamics has achieved fast development. Many foreign and Chinese scholars have applied nonlinear dynamics into the research on MEG signals, and for example, both approximate entropy and sample entropy have been used in related researches on MEG[5]. Transfer entropy is generally used to study the binary connectivity of complicated nonlinear system, which has great performance. Therefore, many scholars have employed transfer entropy in researches on EEG and ECG (electrocardio). However, for multivariate connectivity, the traditional transfer entropy algorithm cannot provide good performance. Therefore, we adopt the partial information decomposition algorithm to improve the multivariate transfer entropy for research on MEG signals.

Mutual information is commonly used to quantify the correlation between two systems, and it does not have directivity nor include the dynamics features. The concept of transfer entropy proposed by Thomas Schreiber in 2000 is very suitable for nonlinear system, because it not only includes the characteristics of mutual information, but can also be used to study the mutual information between two systems that contains directivity and dynamics features[6]. Assume there are two independent random processes $I$ and $J$, the formula of transfer entropy can be defined as:

$$ T_{j\rightarrow i} = h_i - h_{ij} = \sum p(i_{n+1}, i_n^{(1)}, j_n^{(0)}) \log \frac{p(i_{n+1} | i_n^{(1)}, j_n^{(0)})}{p(i_{n+1} | i_n^{(1)})} $$

(1)

The transfer probability $p(i_{n+1}, i_n^{(1)}, j_n^{(0)})$ is from state $(i_n^{(1)}, j_n^{(0)})$ to state $i_{n+1}$, the conditional probability $p(i_{n+1} | i_n^{(1)}, j_n^{(0)})$ is from state $(i_n^{(1)}, j_n^{(0)})$ to state $i_{n+1}$, and the conditional probability $p(i_{n+1} | i_n^{(1)})$ is from state $i_n^{(1)}$ to state $i_{n+1}$.

The inside of brain is a very complicated nonlinear system, and transfer entropy can be used to obtain the coupling strength and direction of two complicated systems. In order to study the information transfer among multiple channels of MEG, it is beyond the scope of traditional binary connectivity, so it is necessary to adopt multivariate measurement method. The partial information decomposition algorithm is a recently developed multivariate information measurement algorithm.

Figure 1 shows that the partial information decomposition algorithm decomposes the information transfer from source variable to target variable into the following parts: (a) Redundant information: all overlapped parts of source variable information are provided to the target variable; (b) Synergistic information: the synergistic effects of all source variables on the target variable; (c) Unique information: the information of each source variable to the target variable.

Assuming there are two source variables $X_{\tau-t}$ and $Y_{\tau-t}$, the target variable is $Z$, and PID is given as[6,7].

$$ I(X_{\tau-t}, Y_{\tau-t}; Z) = U_x + U_y + S + R $$

(2)

$$ I(X_{\tau-t}; Z) = U_x + R $$

(3)

$$ I(Y_{\tau-t}; Z) = U_y + R $$

(4)

![Figure 1. The partial information decomposition algorithm.](image)
Therefore, we employ partial information decomposition method to study the multivariate TE. The pid multivariate TE can include the four variables of multivariate TE measurement: the past states of three MEG signal and the future state of one MEG signal. Furthermore, it can decompose interaction into non-overlapped and non-negative parts. In this paper, we study the trivariate TE. Assuming the time series of 3 MEG signals is $I$, $J$, $K$, we can conduct the following decomposition:

$$
\text{TE}(J \rightarrow I) = \text{Synergy}(J, K \rightarrow I) + \text{Unique}(J; K \rightarrow I) + \text{Redundancy}(J, K \rightarrow I)
$$

(5)

$$
\text{Unique}(J; K \rightarrow I) = \text{Unique}(K; J \rightarrow I) + \text{Redundancy}(J, K \rightarrow I)
$$

(6)

$$
\text{Redundancy}(J, K \rightarrow I) = \text{Unique}(J; K \rightarrow I) + \text{Redundancy}(J, K \rightarrow I)
$$

(7)

In Formulas 5-7, $\{J, K\}$ represents the vector value of the combination of time series $J$ and time series $K$, and the TE part can be obtained from Formula 1; the synergistic part corresponds to the joint information effect of $J$ and $K$ on the future state of $I$; the unique part refers to the separate information effect of $J$ or $K$ on the future state of $I$; the redundant part represents the information effect of the overlapped part between $J$ and $K$ on the future state of $I$. The unique part can deduct the redundant part through the TE part; the synergistic part can deduct the redundant part and unique part in Formulas 6-7.

Among them, the most critical part is the redundant part. In 2010, Williams and Beer proposed the minimum information measurement method[7]. Here, the minimum information can be regarded as the weighted sum of the common information of each future state of time series $I$ held by time series $J$ and $K$ under the past state condition of given time series $I$, so we can calculate the redundant part via Formulas 8-10, and calculate the synergistic part and unique part according to the distributions in Formulas 5-7.

$$
I_{\text{min}}(I_p; J, K, P | I_p) = \sum_{i_p} p(i_p) \min_{K \in \{i_p, K\}} I_{\text{spec}}(I_p = i_p; R | I_p) = \sum_{i_p} p(i_p) \min_{K \in \{i_p, K\}} [I_{\text{spec}}(I_p = i_p; R, I_p) - I_{\text{spec}}(I_p = i_p; I_p)]
$$

(8)

$$
I_{\text{spec}}(I_p = i_p; R, I_p) = \sum_{r_i} p(r, i_p | i_p) \log \left( \frac{p(r, i_p | i_p)}{p(r, i_p) p(i_p)} \right)
$$

(9)

$$
I_{\text{spec}}(I_p = i_p; I_p) = \sum_{i_p} p(i_p | i_p) \log \left( \frac{p(i_p, i_p)}{p(i_p) p(i_p)} \right)
$$

(10)

3. Data analysis and result analysis

Compared to the traditional TE algorithm, the multivariate TE partial information algorithm separates the redundant part from the synergistic part, which can more effectively study the structure of multivariate information effect. For the MEG signals studied in this paper, the interaction between different channels can be more clearly obtained. Because this paper focuses on the MEG of trivariate symmetric channels, the input includes two source channels and one target channel, so this paper uses the synergy value as the measurement value of MEG signal.

3.1. Data collection

240 affective pictures were selected from the International Affective Picture System (IAPS) database, of which, 80 are negative affective pictures, 80 are positive affective pictures, and 80 are neutral affective pictures[9]. In order to ensure the preciseness and reliability of experiment, all selected pictures are set with the same size and consistent contrast and brightness. For all experimental subjects, the data was collected in a magnetically shielded room to reduce the interference to the data collection process by other external factors. The experimenter was outside of the magnetically shielded room, and used a projector to successively project the affective pictures on the big screen. The eyes of experimental subjects were approximately 60cm from the screen, and they are asked to relax and try not to blink.
The 275 channels in the data collected by the MEG system are divided based on area. They can be divided into three main areas: the left, middle and right areas; each main area can be further divided into 5 areas: the frontal, central, parietal, occipital and temporal areas, which are represented as F, C, P, O and T, respectively. Of them, the left area includes 132 channels, the middle area includes 11 channels, and the right area consists of 132 channels.

The name of each channel consists of 3 letters and 2 numbers: the first letter M stands for MEG; the second letter represents the main area; the third letter refers to the small area under this main area; the two numbers represent the line number and column number of the channel.

3.2. Data processing and result analysis

All data collected by the MEG system are files ended with .meg4. We use the SPM8 software to extract required data, adjust the trigger delay, and intercept data from -200ms to 600ms. The data is processed into the format that can be recognized by Matlab, and the 275×161×80 3D data is obtained through baseline correction and frequency reduction of data. After the MEG data collected from an experimental subject under the stimulation of a picture is processed, Fig. 2 shows the oscillogram of processed data on channel MRF36.

![Figure 2. MEG oscillogram on channel MRF63](image)

The frontal lobe is the lobe close to forehead, which is at the front of each cerebral hemisphere (before the parietal lobe). The prefrontal cortex of lobe is closely related to the advanced cognitive ability, and for example, the executive capability, memory and concentration are all controlled and adjusted by the prefrontal cortex [10].

Therefore, the frontal area is used the area of MEG studied in the experiment. The left (right) frontal area corresponding to prefrontal lobe is the target object, the central frontal area is one of source variables, and the other variable is the right (left) frontal area. They are combined in the channel distribution order from middle to two sides, and from up down, and each combination is numbered with 1,2,3⋯⋯99. As a result, when the left frontal area is the target object, there are 99 combinations in total, and there are also 99 combinations when the right frontal area is the target object. The time series of each channel has the threshold value of zero, which is processed into the binary time series.

The multivariate TE partial information decomposition algorithm is employed to calculate the synergy value of each combination. There are 6 depression patients and 13 healthy people in total. After obtaining the synergy value of each combination, the SPSS software is used to conduct T-test and statistical analysis of each independent sample. When \( P < 0.05 \), it indicates the difference between depression patients and healthy people has statistical significance, and that the normal sample can be distinguished from the patient sample in this combination. Finally, under different emotional stimulations, select the combinations that satisfy T-test, obtain the statistical mean of normal people and depression cases, and use Matlab to draw diagram and analyze the results.

Fig. 3 shows all the results that satisfy T-test under different stimulations when the target object is the channel in left frontal area. Fig. 3(a) shows the results of depression patient samples and healthy samples that satisfy T-test among the symmetric channel combinations in the frontal area under neutral stimulation. In these symmetric channel combinations, the normal people can be effectively distinguished from the depression patients, and the combination number is 9. When the target object is
channel MLF25, the source object 1 is MZF01, and the source object 2 is MRF25, the distinction is the best, and the difference between the mean value of healthy samples and that of depression patient samples is also the biggest. Similarly, Fig.3(b) shows the results under positive stimulation, and the combination number is 3. The distinction is the best when the target object is channel MLF13, the source object 1 is MZF01, and the source object 2 is MLF13. In Fig.3(c), the combination number is 39. When the target object is channel MLF22, the source object 1 is MZF02, and the source object 2 is MRF13, the difference between the mean values of healthy samples and depression patients is the biggest.

By comparing the three situations of neutral stimulation, positive stimulation and negative stimulation as shown in Fig.3, it is found that in these combinations where the depression patients can be effectively distinguished from healthy people, the synergy value of depression patients is generally lower than that of healthy people, i.e., the connectivity between the left and right frontal areas in healthy people is generally higher than that in depression patients. This also verifies the low spirits and low complexity of brain activities of depression patients introduced above. Under different emotional stimulations, healthy people do not present significant difference in various factors. However, under negative stimulation, the mean synergy value of depression patients is higher than that under neutral and positive stimulations, which is close to and even higher than that of healthy people. This means that the complexity of brain activities of depression patients increases under negative stimulation.

Fig.4 shows all the results that satisfy T-test under different stimulations when the target object is the channel in right frontal area. Fig.4(a) shows all the satisfying channel combinations under neutral stimulation, Fig.4(b) presents all the satisfying channel combinations under positive stimulation, and Fig.4(c) shows all the satisfying channel combinations under negative stimulation. Obviously, compared to the results of left frontal area as shown in Fig.3, the results are better under different stimulations when the right frontal area is the target object, and more combinations in this area satisfy T-test. Similarly, just like Fig.3, the synergy value of healthy people is higher than that of depression patients. Under negative stimulation, the synergy value of depression patients increases significantly, which means that the complexity of brain activities of depression patients increases under negative stimulation.

Figure 3. Synergy value of multivariate TE partial information decomposition under three emotional stimulations when the target object is three-channel combination in the left frontal area

Figure 4. Synergy value of multivariate TE partial information decomposition under three emotional stimulations when the target object is three-channel combination in the right frontal area
4. Conclusion

The analysis results show that the multivariate TE partial information decomposition algorithm proposed in this paper can well distinguish normal people from depression patients in various symmetric channel combinations in the frontal area. The synergy value of multivariate TE of the symmetric channel in the frontal area of healthy people is significantly higher than that of depression patients, especially when the target object is the right frontal area, which indicates that healthy people have better connectivity in brain activities than the depression patients. However, under negative stimulation, the complexity of depression patients’ brain activities increases greatly. We wish the experimental method provided in this paper can help further study the distinction between depression patients and healthy people, which can provide certain assistance to the clinical diagnosis and treatment of depression.

5. Acknowledgments

The project is supported by the National Natural Science Foundation of China (Grant Nos. 31671006, 61771251, 11974231), Jiangsu Provincial Key R&D Program (Social Development)(Grant No.BE2015700,BE2016773), Natural Science Research Major Program in Universities of Jiangsu Province (Grant No.16KJA310002).

References

[1] R.J. Davidson, D. Pizzagalli, J.B. Nitschke, Depression: Perspectives from affective neuroscience, Annu. Rev. Psychol., 53 (2002) 545-574.
[2] G.C. Medeiros, L. Seger, J.E. Grant, Major depressive disorder and depressive symptoms in intermittent explosive disorder, Psychiat. Res., 262 (2018) 209-212.
[3] P.J. Basser, C. Pierpaoli, Microstructural and Physiological Features of Tissues Elucidated by Quantitative-Diffusion-Tensor MRI, J. Magn. Reson. Imaging., 213 (2011) 2560-570.
[4] R. Oostenveld, P. Fries, E. Maris, FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data, Comput. Intel. Neurosc., 2011 (2011) 3 156869.
[5] S. Baillet, J.C. Mosher, R.M. Leahy, Electromagnetic brain mapping, IEEE. Signal. Proc. Magazine., 18 (2001) 6 14-30.
[6] T. Schreiber, Measuring information transfer, Phys. Rev. Lett., 85 (2000) 2 461-464.
[7] P.L. Williams, R.D. Beer, Nonnegative decomposition of multivariate information, arXiv:1004.2515, (2010).
[8] P.L. Williams, R.D. Beer, Generalized measures of information transfer, arXiv:1102.1507, (2011).
[9] A.C Constantinescu, M. Wolters, A. Moore, A cluster-based approach to selecting representative stimuli from the International Affective Picture System (IAPS) database, Behav. Res. Methods., 49 (2017) 3 896-912.
[10] Y. Moriguchi, K. Hiraki, Longitudinal development of prefrontal function during early childhood, Dev. Cogn. Neuros-neth., 1 (2011) 2 153-162.