31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Independent Component Analysis Using Clustering on Motor Imagery EEG

Hongzhi Qi, Yuhuan Zhu, Dong Ming*, Baikun Wan
Department of Biomedical Engineering,
Tianjin University
Tianjin, 300072, P. R. China,
e-mail address: richardming@tju.edu.cn

***Abstract*—Motor imagery is a popular paradigm in electrophysiology research and brain computer interface but the evoked EEG signals always contaminated significantly. In this paper we use the Independent Component Analysis to enhance the signal-to-noise ratio of multi trail EEG signals evoked by imaginary hand movement. Infomax algorithm was used to decompose multi channel EEG signals into independent components trail by trail, and then an automatic clustering method was used to group these components into several clusters. For the higher similarity between task relevant components, they can be assembled into one cluster that occupies the highest mean mutual information of pairwise components intra cluster. Furthermore, the reconstructed signals of task relevant cluster showed a high discrepancy features to left versus right hand task, which evaluated by Fisher criterion scores and served as the signal-to-noise ratio measurement.**

Keywords- Independent Component Analysis, clustering, Motor Imagery, Fisher Criterion Scores.

I. INTRODUCTION

Electroencephalogram (EEG) and Event Related Potentials (ERP) have been employed for research of brain functional activity for many decades. However, the signal-to-noise ratio of EEG and ERP are very below in most cases, it is necessary to use denoising methods enhance signal quality [1].Independent Component Analysis (ICA) is one of those techniques for signal boosting, especially for electrophysiological task stimulated brain fluctuations. In ICA processing, components corresponding different sources were decomposed and what produced by noise sources were removed while task relevant components were preserved [2, 3].

In multi trail conditions, which quite normally in physiological experiment, the linear decomposition ICA method is hard to trace the time variance functional fluctuations, then short time window analysis on each single trails are preferred [4]. However this single trail decomposition means large amount of Independent Components (ICs) recognitions, which always deal manually. The point is traditional ICA decomposition extract components without any specialized order and always inconsistent across different trails [2].

In this paper we use a clustering technique to automatically recognize ICs from different trails; we find that most task relevant components from different trail expressed a high similarity and can be assembled into one group. In this clustering method, mutual information (MI) is adopted to measure the similarity between pairwise ICs and served as the distance definition in clustering algorithm. For multi trail electroencephalogram data stimulation we use imaginary hand movement paradigm, which is broadly used in brain computer interface research.

This research was supported by National Natural Science Foundation of China (No.60501005), The National High Technology Research and Development Program of China (No.2007AA04Z236), Biomedical Engineering Key Program (No.07ZCKFSF01300) and Intentional Cooperation Key Program (No.08ZCGHHZ00300) of Tianjin Science Technology Support Plan.

II. METHODOLOGY

A. ICA model and Similarity measurement between ICs

ICA is a general-purpose statistical technique that attempts to recover a set of statistically independent sources from observed mixtures [2]. The data submitted to ICA is multi channel observed scalp EEG data $X(t) = [x_1(t), x_2(t), \cdots, x_n(t)]^T$, X was arranged in a matrix of n channels (rows) by t time points (columns). The task of ICA is to find a de-mixing matrix W that makes the components of $Y = W \times X$ mutually independent. Y is the estimate of sources $S(t) = [s_1(t), s_2(t), \cdots, s_n(t)]$, which are linearly mixed into observed EEG signals $X = V \times S$. The ideal W is the inverse of V, practically in this paper is obtained using Infomax ICA [5].

To clustering ICs from multi trail EEG decomposition a distance measurement is necessary. Under this measurement components from task stimulated should be more closely than other components. In this paper, mutual information based on Shannon theory was selected to evaluated how close between two components, $I(y_i, y_j) = H(y_i) + H(y_j) - H(y_i, y_j)$,where H is the Shannon entropy of ICA decomposed component, I is the mutual information between two components i and j [6].

B. Multi trails clustering

In clustering, K-means algorithm was selected for assemble components from multi trails into different groups. With presumed group number K, K-means method firstly select k initial cluster centers and then iteratively refine them until the total inter-class distance to minimum

[7]. For clustering the whole multi trail decomposed component set, assume that there are N components y_i, $i = 1,2,\cdots,N$, K classes C_i, $i = 1,2,\cdots,K$, and N labels, $\lambda_{i,k} = \{1, y_i \in c_k; 0, y_i \notin c_k\}$, which describe whether component y_i belong to kth class. Then class center can be defined as $\overline{y_k}$ and inter-class distance of kth class is $D_k = \sum_{i=1}^{n} \lambda_{i,k} I(y_i, \overline{y_k})^2$. The objective function in clustering is the whole inter-class distance $\widehat{D} = \sum_{k=1}^{K} D_k = \sum_{k=1}^{K} \sum_{i=1}^{n} \lambda_{i,k} I(y_i, \overline{y_k})^2$. The partition of multi-trail ICA component set can be achieved by using gradient descent algorithm to minimize whole inter-class distance $\widehat{D}$, the optimized $\lambda_{i,k}$ give the class label of each component.

Multi channels task evoked signals of each trail were decomposed by Infomax algorithm individually and all components were grouped into one unit set [8]. Mutual information distance of each pairwise component were Calculated in this unit set, then K-means algorithm was used to Separate the whole component set into K clusters.

The imaginary hand movement task evoked a stolid and time locked fluctuation characteristics, which mean the components correlated to the stimulus of different trails should have a high similarity, whereas the stimulus is left or right hand imagining. Compared to task-relevant inter trail similarity, the task-irrelevant components from different trails should have a low similarity because spontaneous EEG and noise component hardly be affected by these imaginary stimulus. Therefore using mutual information based distance measurement the task-relevant and task-irrelevant will clustered into different groups, which could be recognized by the intra cluster mean mutual information (MMI).

III. EXPERIMENT

Motor Imagery evoked EEG is popular in current Brian computer interface study and also draws a broad interest in neurophysiology research [9]. A traditional imaginary hand movement experiment has a preconfigured task rhythm and a visual cue for task start-up.

In experiment the subjects were seated in an armchair at 1-m distance in front of a computer screen and were asked to imagine left versus right hand movements during each trial. The experimental paradigm is described as follows: after trial begin, the first 2s were quite; at t=2s a cross "+" is displayed for 1s; then from t=3s an arrow to the left or right was displayed, at the same time the subject was asked to imagine a left hand or right hand movement respectively, until the arrow disappeared at t=8s. Each of the 2 cues was displayed 32 times within each run in a randomized order; the experiment consists of 2 runs with a 5 minutes break between them.

We recorded EEG signals from eleven untrained right handed subjects using 19 electrodes, whose recording position follows standard 10-20 system. The reference electrodes were positioned at ears and two electrodes were used to record possible EOG artifacts and eye blinks. The EEG was sampled with 512 Hz and it was filtered between 1 and 50Hz with Notch filter on.

IV. PREPROCESSING AND RESULT

A. SNR evaluation with Fisher criterion (FC)

As Pfurtscheller and da Silva have reported that movement related desynchronization of the μ-rhythm (8–13 Hz) on sensorimotor region is the primary EEG features of this hand motor imagery task [10]. Restricted to the C3 and C4 channels that were located over sensorimotor cortex we calculated the time variance power spectral density (PSD) of μ-rhythm as the representation of task evoked signal. In this process a 1s removing Hanning window was used on the continuous EEG data, then PSD was computed and changed to db units by log transformation.

To evaluate the signal-to-noise ratio, a FC score factor was used to describe how strongly the power spectral density features correlated to the task [11]. FC score R defined as: $R(X) = (U(X^+) - U(X^-))^{\wedge}2/(V(X^{\mp}) + V(X^-))$, where U(X) is the mean value and V(X) is the variance value of features within one class, +/- represent two classes (here is left/right hand task).

In left hand versus right hand imaginary movement task, μ rhythm on sensorimotor cortex expressed a label related PSD feature; consequently, we can use FC scores to evaluate the signal qualification instead of SNR, which is hardly to measure in scalp electroencephalogram directly.

B. ICA processing and clustering

Infomax algorithm of ICA was selected to decompose the single trail EEG data, and then 19 components (same to 19 EEG electrodes) were achieved for each left versus right hand imaginary movement task. Each component set was grouped from 64 trails, i.e. one run, so there is totally 1216 components involved in the clustering process. For each trail ICA processing, a decomposition matrix $W_i, i = 1,2,\cdots,64$, and an independent components set $Y_{i,j}(t), i = 1,2,\cdots,64, j = 1,2,\cdots,19$, were achieved, here i corresponding to trails and j corresponding to channels.

The distance between each pairwise component $(Y_m, Y_n), m, n = 1,2,\cdots,64 \times 19$ were computed follow mutual information definition described in section II, and then this 1216*1216 size distance matrix was clustering by K-means algorithm, which operated on Matlab statistics toolbox. Clustering process gives each IC a label to represent which cluster it belongs to, $Y_{i,j}^{c}(t), c \in \{1,2,\cdots,K\}$, here c is cluster index.

To evaluate the FC scores of each component cluster a reconstruction procedure was processed at first, $S_i^c = W_i^{-1} \times Y_i^c(t)$, W_i^{-1} is inverse of decomposition matrix,. Then FC scores R_j were computed by the multi trail reconstruction signal $\{S_{i,j}^c\}$, here j is the channel index. In this paper we use left hand task specified channel, C4, to study what happened on the signal-to-noise ratio of task relevant signals after this single cluster reconstruction. The FC scores and reconstruction signal were R_{C4} and $S_{i,C4}^c$ respectively.

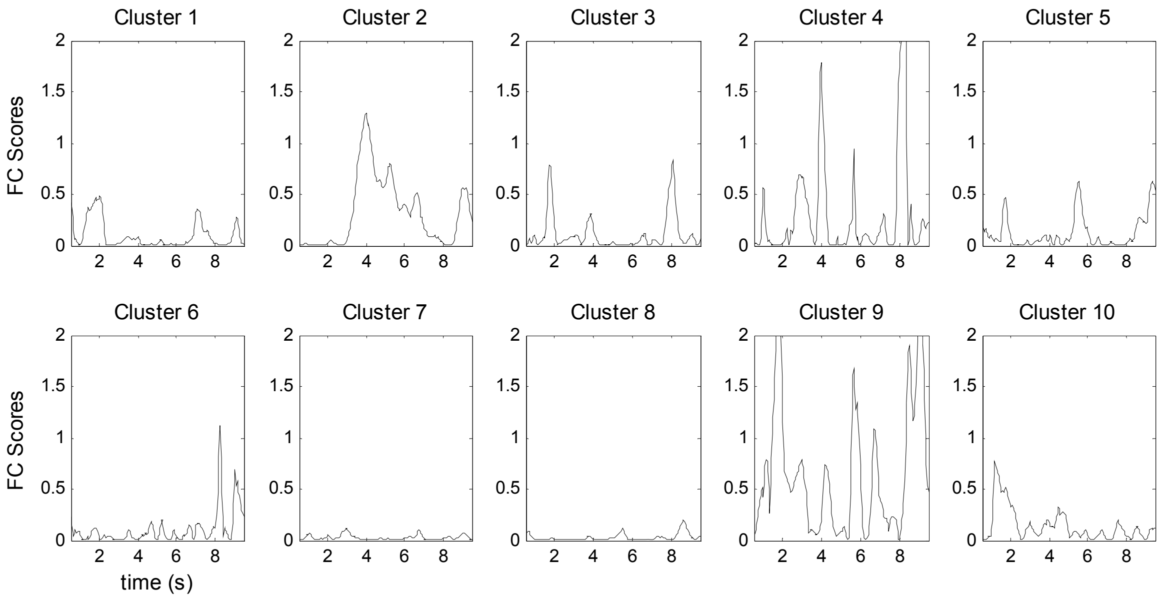

Fig 1. Time variance FC Scores curve of reconstructed signals on C4 using components of each cluster only.

Fig.1 shows FC scores time variance curve of each cluster reconstruction signal of one subject. There are only few clusters have components more than 64 (trail number), the index of these cluster is 2, 7, 8, and components amount is 579,106,309 respectively. Because task relevant components should be find in all trails, those clusters without full trails contribution could be regarded as task irrelevant and then can be removed.

The left versus right hand imaginary movement task will draw a μ rhythm desynchronization on contra lateral sensorimotor area. Consequently, compared to right hand task, there is a higher depression of μ rhythm PSD on C4 when left hand task was executed. Using FC scores to characterize this difference between two tasks on each time coordinate, a time variance separability curve can be achieved. Fig 1 shows FC score curves of reconstructed signals of all clusters. In this figure, only cluster 2 expressed a significant increase in the active period (3~8s) and a very low magnitude in quiet (0~2s) and rest period (8~10). Therefore, cluster 2 shows an obvious task related specialty and denotes that components belonging this cluster contains enough task evoked information.

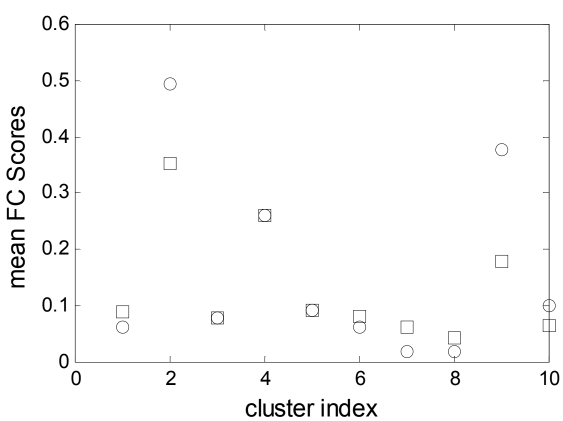

Fig 2. Mean mutual information of pairwise components in each cluster.

Clustering algorithm used mutual information distance to group ICs close with each other together. To identify the compactness of ICs in each cluster the distance between pairwise ICs was calculated and averaged within cluster. As Fig 2 depicted, Mean MI of each cluster varied greatly and cluster 2 occupied an obvious higher MMI than other clusters. Consequently this MMI difference could be used to recognize which cluster is task relevant.

ICs in each cluster consist of two classes, intra trail and inter trail components. For intra trail components the mutual information can be regarded as the residue dependence after ICA, which may be caused by the linear warp of ICA model. Another possible reason of this residue MI existence is the mismatch between source and recording channels, when recording channels occupies a higher number than sources ICA decomposed more components than sources fact, which means there are several ICs come from a same source.

For inter trail components, the mutual information can be regarded as to describe long range similarity in different time. Under this condition, inter trail task relevant components should expressed a higher similarity because the task iteratively occurrence provide different IC sets the same evoked sources. Whereas sources corresponding to the spontaneous Electroencephalogram background and other noises such as EOG, MEG and ECG, are all lack of this stable rhythm iteratively emergence, hence the similarity between these inter trail task irrelevant components would be lower.

To evaluate the signal to noise ratio of motor imagery task evoked signals we reconstructed signals using each individual cluster. The SNR of each cluster was evaluated by FC scores on left versus right hand class and compared to original signals before ICA and clustering process. Fig 3 depicted the mean FC scores during active time range, i.e. 3~8s, of reconstructed signals on C4 and C3 channel. It is very evident in this fig of that the task relevant cluster expressed a higher FC scores than other clusters no matter what on C4 or on C3. Note that cluster 4 and 9 also showed a high mean FC scores because there are only a few components in these clusters (<< 64), which means the FC estimation in these clusters were hardly accurate.

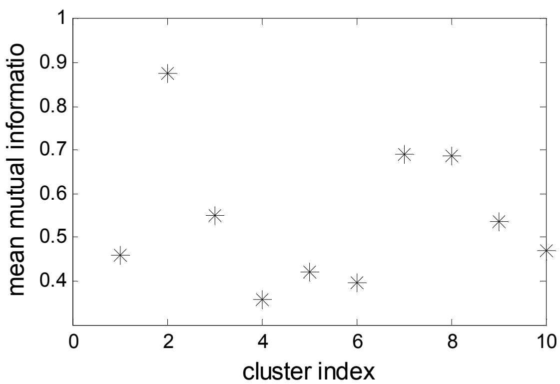

Fig 3. Mean FC scores during active time on C4 and C3 constructed signals using components of each cluster.

In reconstructed signals, compared to original signals, the mean FC scores were increased on 31 percent (from 0.39 to 0.51) and 20 percent (from 0.28 to 0.35) on C4 and C3 respectively. This SNR increment proved that the irrelevant task components would be disturbing factors in left versus right hand recognition and the removal of these components would enhance this motor imagery evoked signal quality.

V. CONCLUSION AND DISCUSSION

In this paper we use the Infomax ICA algorithm on imaginary hand movement stimulated EEG signals one trail by one trail and then use the mutual information based distance to clustering components into task relevant and task irrelevant clusters. By using the FC scores as the signal noise ratio measurement on reconstructed signals on task relevant cluster, this method present a high quality to increase signals SNR. Another advantage of this method is the convenience to recognize task relevant ICA decomposed components, because this method change the single component recognition to component cluster recognition, which can be easily and automatically completed by using mean mutual information within each cluster.

Otherwise, using K-means clustering algorithm to group samples required presume the number of clusters, it is necessary to optimize factor K to pursue the highest SNR of reconstructed signals. In this optimization the mean FC could be used as the cost function, then the best K would provide the most evidently discrepancy on sensorimotor area.

Compared to others task relevant cluster has a prominent higher number of components, the reason is each trail provide several task-relevant components and these components have larger mutual information than task-irrelevant ones. Although stimulus correlated FC scores of reconstructed signals from this cluster demonstrated that it maintains enough task evoked features, they can't guarantee that all components in this cluster are task-relevant. Actually, whether the ICA decomposition has the capability to divide the multi channel EEG signals into task relevant and irrelevant components still have a lot of argument.

REFERENCES

[1] Todd C Handy. Event-Related Potentials: A Methods Handbook. Cambridge, MIT Press, 2004

[2] M Girolami. Self-organizing Neural Networks: Independent Component Analysis and Blind Source Separation, London: SpringerVerlag, 1999

[3] T. P. Jung, S Makeig, C. C. Humphries, A. T. Lee, M. J. Mckeown, A V. Iragui, et al. "Removing electroencephalogramphic artifacts by blind source separation," Psychophysiology, vol. 37, pp. 163-178, 2000.

[4] J. Onton, M. Westerfield, J. Townsend, S. Makeig, "Imaging human EEG dynamics using independent component analysis", Neuroscience and Biobehavioral Reviews, vol. 30, pp.808-822, 2006

[5] P. COMON, "Independent Component Analysis, a new concept ?," Signal Processing, vol. 36, pp. 287--314, 1994,

[6] H. Stögbauer, A. Kraskov, S. A. Astakhov, P. Grassberger, "Least-dependent-component analysis based on mutual information," Phys Rev E Stat Nonlin Soft Matter Phys. vol. 70, pp. 066-123, 2004

[7] A. K. Jain, M. N. Murthy, P. J. Flynn, "Data Clustering: A Review," ACM Computing Reviews, vol. 31, pp. 264-323, 1999.

[8] T. W. Lee, M. Girolami, T. J. Sejnowski, "Independent Component Analysis Using an Extended Infomax Algorithm for Sub-Gaussian and Super-Gaussian Sources," Neural Computation,vol. 11, pp. 409-433, 1999

[9] G. Pfurtscheller, C. Neuper, G. R. Müller, B. Obermaier, G. Krausz, A. Schlogl, et al., "Graz-BCI: State of the Art and Clinical Applications," IEEE Trans. Neural. Syst. Rehabil. Eng., vol. 11, pp. 1-4, 2003

[10] G. Pfurtscheller, L. d. Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," Clin. Neurophysiol. Vol. 110, pp. 1842-1857, 1999.

[11] S. Theodoridis, K. Koutroumbas, Pattern Recognition. Second Edition, New York, Academic Press, 2003.