EXAMINING ASSOCIATIONS BETWEEN FMRI AND EEG DATA USING CANONICAL CORRELATION ANALYSIS*

Nicolle Correa, Yi-Ou Li, and Tülay Adalı

University of Maryland, Baltimore County
Department of CSEE
Baltimore, MD 21250

Vince D. Calhoun

The MIND Institute and University of
New Mexico, Department of ECE
Albuquerque, NM 87131

ABSTRACT

Functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) provide complementary information about the brain function. We propose a novel scheme to examine associations between these two modalities using canonical correlation analysis (CCA). Our multimodal canonical correlation analysis (mCCA) scheme utilizes inter-subject covariations to quantitatively link the two modalities and to estimate the spatio-temporal areas of association. We evaluate the performance of mCCA using simulated fMRI- and EEG-like data and note its ability to effectively identify associations across modalities. Also, our experiments on actual data from an auditory oddball task reveals associations of the temporal and motor areas with the N2 and P3 peaks, a finding that is consistent with previous studies. Additionally, we compare the performance of mCCA to the recently introduced joint-ICA technique for estimating spatio-temporal connections from multimodal data and discuss the advantages and limitations of each.

Index Terms— biomedical signal analysis, magnetic resonance, electroencephalography, multimodal analysis, canonical correlation analysis

1. INTRODUCTION

An increasing number of studies are collecting multiple measurements e.g. functional magnetic resonance imaging (fMRI) data, structural MRI data, electroencephalography (EEG) data, genetic data, and others from the same participants to identify differences across treatment groups. Each modality has its advantages as well as limitations and the purpose of analyzing multiple modalities is to utilize the common as well as unique information from complementary modalities to better understand neurological activity. Recently developed techniques for combined analysis perform either data integration (based on simple co-registration or using one modality to constrain the other) or data fusion (incorporating interactions among imaging types).

In this paper, we examine the connections between fMRI and EEG data through inter-subject covariations. FMRI provides information about dynamic blood flow changes in response to a stimulus with millimeter spatial and second temporal resolution, whereas EEG provides information about electrical activity of the brain with centimeter spatial and millisecond temporal resolution. Since these two modalities are rich in information at different spatial and temporal scales, a multimodal analysis could provide novel understanding about the brain function. At the same time, the fundamentally different nature of these modalities requires careful examination of the assumptions made to the multimodal analysis model.

We propose a scheme to examine associations between multiple modalities using canonical correlation analysis. The multimodal canonical correlation analysis (mCCA) scheme, initially preprocesses the two modalities to obtain lower dimensional features of interest. Our data model enables the decomposition of the two feature sets, into a set of components (i.e., spatial areas for fMRI or temporal segments for EEG), each component associated with a single modulation profile across subjects (i.e., profile of intersubject variations). Multimodal CCA seeks the spatial areas of the brain that are associated with the temporal portions of the bioelectric potentials on the basis of the correlation between the corresponding inter-subject variations. This is a novel application of CCA to the data fusion model since previous applications of CCA to fMRI data [1, 2] utilize spatial correlation rather than inter-subject co-variances to perform data decomposition.

Independent component analysis (ICA) has been successfully utilized for source separation of biomedical data [3] and more recently, it has been extended to jointly analyze data from different modalities. Joint-ICA [4] has been used to fuse information from fMRI and event related potential (ERP) data. It uses statistical independence among the components, which are formed by combining the spatial fMRI locations and the temporal ERP segments, as the criteria for decomposition. Para-ICA [5] on the other hand, performs separate ICA analysis on each modality and enhances the inter-subject co-variations by constraining the correlations between modulation profiles. Like mCCA, para-ICA is flexible in that it

does not assume identical inter-subject variations, however, para-ICA uses an adaptive algorithm whose performance depends on the choice of the learning rates and the threshold above which the correlations are to be enhanced.

The foremost difference between mCCA and the ICA-based techniques is that the associations across modalities in the former method are solely based on inter-subject co-variations whereas the associations in the latter methods are based on the co-variations as well as statistical independence among the corresponding spatial and temporal associations. While statistical independence may be a reasonable assumption to detect a network of areas linked by a particular function, the ICA algorithms may not be able to satisfy both these constraints and may end up achieving a tradeoff solution between the two constraints. Relaxing the independence criterion may enable multimodal-CCA to explore the connectivity between different networks of functional areas associated with both hemodynamic response and electric potential.

In this paper, we investigate the performance of mCCA on both simulated as well as actual data and compare it to that of joint-ICA (jICA). The results demonstrate the promising potential of application of mCCA to multimodal fusion.

2. MULTIMODAL-CCA FOR DATA FUSION

In this section, we explain the generic data model, assumptions, and steps involved in mCCA while the specifics for the fMRI and EEG data sets used in our experiments are given in Section 3. As mentioned in the previous section, in order
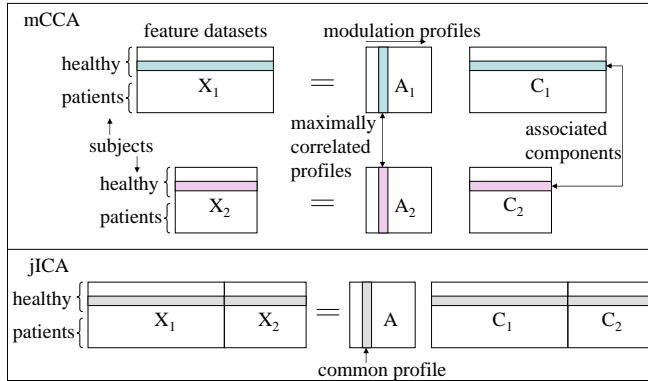

Fig. 1. Data model for mCCA and jICA

to employ an additional dimension of coherence, our scheme utilizes lower dimensional features of interest. We develop the following generative model for data fusion.

Given two feature subdatasets $\mathbf{X}_1$ and $\mathbf{X}_2$, we seek to decompose them into two sets of components, $\mathbf{C}_1$ and $\mathbf{C}_2$, and corresponding modulation profiles (inter-subject variations), $\mathbf{A}_1$ and $\mathbf{A}_2$. The connection across the two modalities can be evaluated based on correlations of modulation profiles of one modality with those of the other. If the modulation profiles

are uncorrelated within modalities, each component can be associated with only one component across modalities. This one to one correspondence aids examination of associations across modalities. The generative model is then given by

$$\mathbf{X}_k = \mathbf{A}_k \mathbf{C}_k, \quad \text{for } k = 1, 2$$

where $\mathbf{X}_k \in \mathbb{R}^{s \times v_k}$, $\mathbf{A}_k \in \mathbb{R}^{s \times d}$, and $\mathbf{C}_k \in \mathbb{R}^{d \times v_k}$, v_k is the number of variables and s is the number of observations in $\mathbf{X}_k$ and d is the min(rank($\mathbf{X}_1,\mathbf{X}_2$)). As per the model, the modulation profiles, $\mathbf{A}_{1n}$ and $\mathbf{A}_{2n}$ ($n = 1, \ldots, d$) are assumed to have the following constraints:

$$E[\mathbf{A}_{1i}, \mathbf{A}_{2i}^T] \geq 0 \tag{1}$$
$$E[\mathbf{A}_{1i}, \mathbf{A}_{2j}^T] = 0, \quad \text{for } i \neq j \tag{2}$$
$$E[\mathbf{A}_{ki}, \mathbf{A}_{kj}^T] = 0, \quad \text{for } i \neq j, k = 1, 2 \tag{3}$$

We propose to use CCA, which is a statistical tool for identifying linear relationships between two sets of variables [6], to determine the inter-subject co-variances. CCA seeks two sets of transformed variates such that the transformed variates assume maximum correlation across the two datasets, while the transformed variates within each data set are uncorrelated. CCA solves the following maximization problem:

$$\max_{\mathbf{P},\mathbf{Q}} \text{corr}(\mathbf{P}\mathbf{X}_1^T, \mathbf{Q}\mathbf{X}_2^T) = \frac{\mathbf{P}\mathbf{R}_{X_1,X_2}\mathbf{Q}^T}{\sqrt{(\mathbf{P}\mathbf{R}_{X_1}\mathbf{P}^T)(\mathbf{Q}R_{X_2}\mathbf{Q}^T)}}$$

$$\text{where} \quad \mathbf{P} \in \mathbb{R}^{d \times v_1} \quad \text{and} \quad \mathbf{Q} \in \mathbb{R}^{d \times v_2}$$

to obtain canonical variates given by

$$\mathbf{A}_1^T = \mathbf{P}\mathbf{X}_1^T \quad \text{and} \quad \mathbf{A}_2^T = \mathbf{Q}\mathbf{X}_2^T$$

which satisfy the constraints given in Eqs. (1)–(3). Hence, $\mathbf{P}$ and $\mathbf{Q}$ are the eigenvectors of the two matrices

$$(\mathbf{R}_{X_2}^{-1}\mathbf{R}_{X_2,X_1}\mathbf{R}_{X_1}^{-1}\mathbf{R}_{X_1,X_2} - \mathbf{rI})\mathbf{P} = \mathbf{0}$$
$$(\mathbf{R}_{X_1}^{-1}\mathbf{R}_{X_1,X_2}\mathbf{R}_{X_2}^{-1}\mathbf{R}_{X_2,X_1} - \mathbf{rI})\mathbf{Q} = \mathbf{0}$$

where $\mathbf{r}$ is vector of eigenvalues or squared canonical correlations. $\mathbf{R}_{X_1,X_2}$ is the cross-correlation matrix of $\mathbf{X}_1$ and $\mathbf{X}_2$ ($\mathbf{R}_{X_2,X_1} = \mathbf{R}_{X_1,X_2}^T$), and $\mathbf{R}_{X_1}$ and $\mathbf{R}_{X_2}$ are the autocorrelation matrices of $\mathbf{X}_1$ and $\mathbf{X}_2$ respectively. Thus, mCCA models the inter-subject covariations as the canonical covariates obtained by CCA and the least squares approximations of the components are given by

$$\hat{\mathbf{C}}_k = (\mathbf{A}_k^T\mathbf{A}_k)^{-1}\mathbf{A}_k^T\mathbf{X}_k, \quad \text{for } k = 1, 2.$$

In Fig. 1, we compare the model of mCCA to that of jICA [4]. The main difference between the two models is that jICA assumes a common set of modulation profiles for both modalities and imposes an independence constraint on the components. Additionally, jICA does not impose any uncorrelation constraint among the modulation profiles. Hence, jICA examines the common connection between independent networks in both modalities while mCCA allows for common as well as distinct components and describes the level of connection between the two modalities.

3. EXPERIMENTS

We perform experiments on both simulated data as well as actual fMRI and EEG data to evaluate the performance of mCCA and compare it to jICA. In this section, we describe the procedure, data sets and results.

3.1. Dimension reduction

For both fMRI and ERP data, the number of variables in the feature datasets are much larger than the number of observations. Transforming each set of features to a subspace with smaller number of variables helps reduce any redundancy in the analysis. Dimension reduction is performed on the feature dataset using singular value decomposition (SVD). SVD of $\mathbf{X}_1$ and $\mathbf{X}_2$ is given by

$$\mathbf{X}_k = \mathbf{E}_k \mathbf{D}_k \mathbf{F}_k = [\mathbf{E}_k^{'} \mathbf{E}_k^{''}] \mathbf{D}_k \mathbf{F}_k, \quad \text{for} \ \ k = 1, 2.$$

where $\mathbf{E}_1^{'}$ and $\mathbf{E}_2^{'}$ contain the eigenvectors corresponding to the significant eigenvalues in $\mathbf{D}_1$ and $\mathbf{D}_2$ respectively. We perform CCA on the dimension-reduced datasets given by

$$\mathbf{Y}_k = \mathbf{X}_k \mathbf{E}_k^{'}, \quad \text{for} \ \ k = 1, 2.$$

3.2. Simulated data

For brain imaging data, there is no *a priori* knowledge about the ground truth of the underlying components and their modulation profiles across subjects. Hence, in order to test the performance of mCCA, we generate a simulated fMRI-like set of components and an ERP-like set of components and mix each set with a different set of modulation profiles to obtain two sets of mixtures. For the ERP data set, we use the peaks of the actual ERP time course from the data described in the section 3.3. The modulation profiles are chosen from a random normal distribution. The profiles are kept orthogonal within each set. Connections between the two modalities are simulated by generating correlation between profile pairs, formed across modalities.

We simulate five fMRI and five ERP components. Each simulated fMRI component is a 60×60 pixel image with a 200-subject modulation profile. The image is then reshaped to form a vector by concatenating columns of the image. Each ERP component is 451 time point segment with a 200-subject modulation profile. We evaluated the performance of mCCA for a number of different connections strengths and chose an example that demonstrates its representative performance. In the example presented in this paper, the connection between the profiles of the five fMRI components and the five ERP components is 0.9, 0.7, 0.4, 0.25, and 0, respectively. The results are an average of twenty runs, each with independent realizations of the modulation profiles. The criteria we use for performance evaluation are the correlations between the true and estimated components and profiles as well as the differences between the actual and estimated connection strengths.
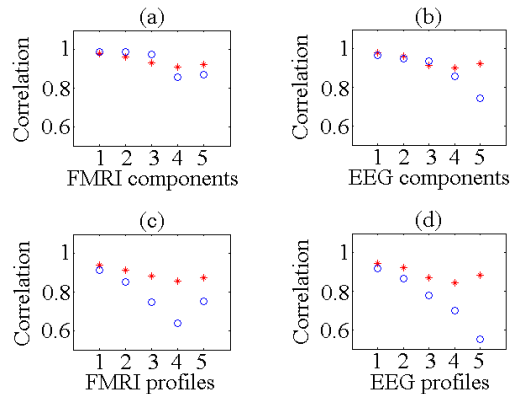

Fig. 2. Correlation with ground truth of components and profiles as estimated by mCCA ($*$) and jICA ($\circ$)

Figure 2 (c) and (d), shows that the profile correlation values are higher for mCCA than for jICA. Joint-ICA estimates a common set of modulation profiles instead of a different set for each modality. In Fig. 2 (a) and (b), the component correlation values of both mCCA and jICA are quite similar for the first three components. These three components have a considerable amount of correlation between profiles of the two modalities (0.9, 0.7, and 0.4) and thus, a common profile estimated by jICA, that forms a tradeoff between the actual separate profiles, suffices to efficiently estimate these components. However, when the connection between the two modalities is weak, the common profile departs further from the ground truth as seen in Fig. 2 (c) and (d), in the case of components 4 and 5, which have connection strengths of 0.25 and 0, respectively. Consequently, in Fig. 2 (a) and (b) we see that components 4 and 5 as estimated by mCCA are closer to the ground truth than the ones estimated by jICA. The connection strengths estimated by mCCA are found to be exact upto one decimal place.

In another simulation case, when the connection strengths are selected to be very close to each other the performance of mCCA degraded and was similar to that of jICA (results not displayed due to space restrictions). For eigen-analysis based solutions, identical eigenvalues are known to pose problems and further investigation needs to be performed to analyze this problem. Next we discuss the experiment on actual data.

3.3. Actual data

FMRI and EEG data are acquired from 39 subjects (23 healthy and 16 schizophrenic subjects) performing an auditory oddball (AOD) task that requires the subjects to press a button when they detect a particular infrequent sound among three kinds of auditory stimuli. Details of the task design and the participants are given in [7]. Both the fMRI and EEG data are processed into lower dimensional features using existing analysis techniques. We use the Statistical Parametric Map-
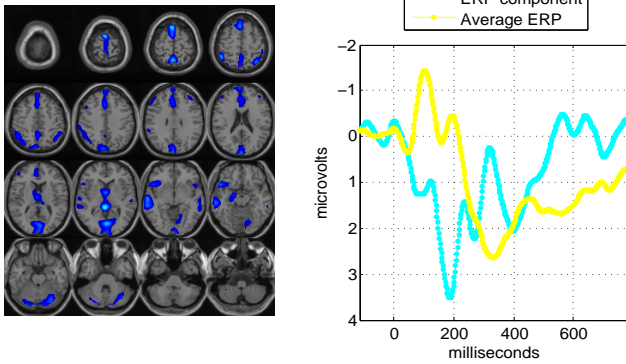
Fig. 3. Pair of associated components estimated by mCCA

ping software package (SPM2) [8] to obtain target related features for fMRI data. For EEG features, ERPs are calculated by averaging epochs of EEG, time-locked to the event of correct target detection.

A number of strong connections are identified by mCCA, three of which have modulation profiles that are significantly different ($\alpha \leq 0.05$) between healthy and schizophrenic treatment groups. We report the results for these three pairs of components. One of these pairs of components, having a connection strength given by 0.85, is displayed in Fig. 3. The fMRI component shows activations in the temporal lobe and the middle anterior cingulate region and the EEG component shows a maximum peak at around 300 msec (N2) after the stimulus onset. This is similar to the result obtained in [9] using jICA on a similar dataset. Another pair of components, having a connection strength of 0.66, also shows activation in the motor areas and the bilateral temporal lobe to be associated with the N2 peak. The final pair of components of interest shows significant differences in treatment groups for only the EEG component and not the fMRI component. This finding is not possible with jICA since it uses a common profile and it is quite plausible that for this particular component pair, fMRI is not sensitive to the differences in controls and patients. The association for this pair is obtained as 0.58. The fMRI component shows temporal lobe areas and the EEG component shows the P3A peak (early part of P300).

4. DISCUSSION

In this paper, we have introduced mCCA for fusing two modalities and demonstrated its usefulness on fMRI and EEG data. Multimodal CCA seeks spatial areas of the brain associated with temporal portions of the bioelectric potentials and utilizes the correlation between the corresponding inter-subject variations to identify these connections.

Our experiments show that mCCA correctly identifies and quantifies associations between two modalities in a simulated environment. Multimodal CCA yields similar performance

as jICA when the connections are strong and yields better performance for weak connections. This is because mCCA identifies separate inter-subject variations as opposed to jICA which assumes identical modulation profiles for both modalities. The assumption of separate modulation profiles is less stringent and also facilitates the quantification of the strength of connections between two modalities. Another difference between the two methods is that jICA constrains the components to be independent while mCCA does not impose any constraints on them. Hence, mCCA identifies different networks of areas that are similarly modulated across subjects while jICA identifies independent networks. On the other hand, mCCA requires the modulation profiles to be uncorrelated within each modality. This makes it easier to examine associations since each component from one modality can be linked to only one component in the other modality. Joint-ICA does not impose any constraints on the profiles. Thus, the two methods are complementary, though it is advantageous not to assume a shared profile for two modalities.

Our experiments demonstrated the usefulness of mCCA to examine linear associations between fMRI and EEG data. An interesting direction for future research would be to examine higher order connections through non-linear CCA and also, to investigate the performance of mCCA to other modalities of complementary nature.

5. REFERENCES

[1] Y-O. Li, T. Adalı, and V. D. Calhoun, "A multivariate model for comparison of two datasets and its application to fMRI analysis," *Proc. MLSP 2007*, Thessaloniki, Greece.

[2] O. Friman, M. Borga, P. Lundberg, and H. Knutsson, "Exploratory fMRI analysis by autocorrelation maximization," *Neuroimage*, vol. 16, pp. 454–464, 2002.

[3] M. J. Mckeown, S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann and T. J. Sejnowski, "Analysis of fMRI by blind separation into independent spatial components," *Human Brain Mapping*, vol. 6, no. 3, pp 160–188, 1998.

[4] V. D. Calhoun, T. Adalı, G. D. Pearlson, and K. A. Kiehl, "Neuronal chronometry of target detection: Fusion of hemodynamic and event-related potential data," *Neuroimage*, vol. 30, no. 2, pp. 544–553, 2006.

[5] J. Liu and V. D. Calhoun, "Parallel independent component analysis for multimodal analysis: an application to fMRI and EEG data," in *Proc. ISBI 2007*, Arlington, Virginia.

[6] A. C. Rencher, *Methods of Multivariate Analysis (Second Edition)*, Wiley-Interscience, 2002.

[7] K. A. Kiehl, M .Stevens, K. R. Laurens, G. D. Pearlson, V. D. Calhoun, and P. F. Liddle, "An adaptive reflexive processing model of neurocognitive function: Supporting evidence from a large scale (n=100) fMRI study of an auditory oddball task," *NeuroImage*, vol. 25, pp. 899–915, 2005.

[8] SPM: http://www.fil.ion.ucl.ac.uk/spm2.

[9] V. D. Calhoun and T. Adalı, "ICA for fusion of brain imaging data," submitted to *IEEE Trans. Info. Technology in Biomedicine*.