

Identification of Brain Image Biomarkers by Optimized Selection of Multimodal Independent Components

R. F. Silva^{1,2}, and V. D. Calhoun^{1,2}

¹Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM, United States, ²Image Analysis, The MIND Research Network, Albuquerque, NM, United States

MOTIVATION AND BACKGROUND

The acquisition of multiple imaging modalities on the same individual in brain imaging studies has become a common practice. In functional studies, often several different tasks are performed on the same person. It is likely and intuitive that such datasets contain some shared and some unique information across the different acquired modalities. This motivates an analysis method that directly looks for the joint (shared) information lying within multimodal datasets. However, the cross-correlational standard-of-practice for examination of different regions does not easily allow for the study of the ‘cross’-information of all brain voxels. A data fusion approach on the other hand has been shown to be beneficial in providing a unified framework for heterogenic datasets (Savopol et. al.). Hence the present work builds on the data fusion framework introduced by Calhoun et. al. (2006) called joint independent component analysis (jICA). As a ‘second level’ analysis this method utilizes summary maps of the patterns of brain activation generated for each individual in the dataset. Henceforth we refer to these summary maps as features derived from ‘first-level’ analysis of the original ‘raw’, 4D data using classic methods like the general linear model (GLM). Thus, with features generated for each modality and each individual of the dataset a jICA data fusion scheme as shown in Figure 1 is applied to yield joint (multimodal), maximally independent components (ICs) which capture the joint information from multiple modalities and enable identification of brain imaging biomarkers.

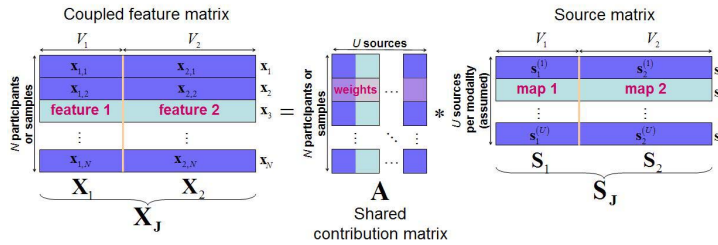


Figure 1 (left): Data fusion scheme for 2 modalities – Based upon the original ICA method nicely described by Hyvarinen et. al., popularized in the neuroscience community by McKeown et. al., and recently reviewed by Calhoun et. al. (2006a), jICA is characterized by a transformation in the dataset which (1) assumes the data are linear mixtures of statistically independent sources and therefore (2) attempts to decompose the data into maximally independent components and their mixing coefficients (or loading parameters). Hence the feature matrix is organized by placing the features from two modalities side by side yielding X_j . Next X_j is modeled as containing spatially independent joint sources S_j which share common mixing matrix parameters A . Each modality is allowed to differ in size ($V_1 \neq V_2$).

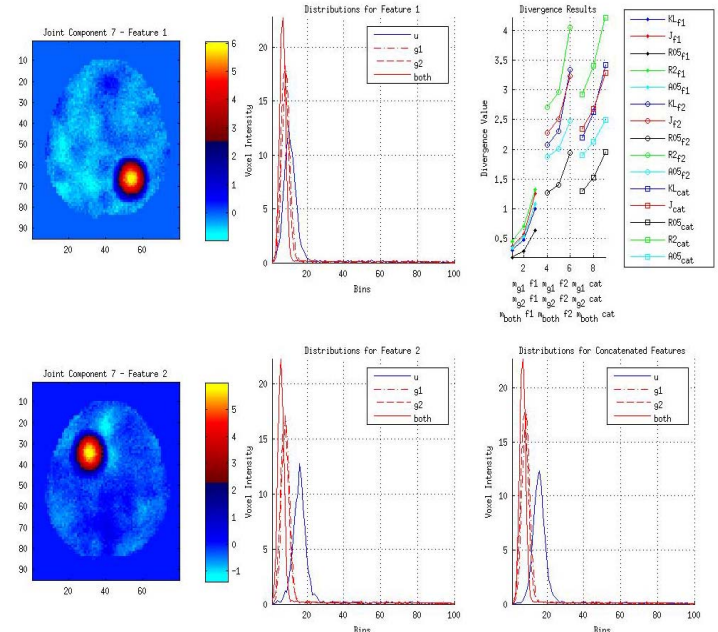
OBJECTIVES

In this study our goal is to improve an optimization step proposed in the study developed by Calhoun et. al. (2006). That study showed that in a general sense the loading parameters of a given independent component provide a measure of how a certain IC is representative of the overall activity within a given subject and therefore relevant information regarding the separation of groups can be extracted from a two-sample t-test on the loading parameters associated to a certain IC. The same study also suggested the use of such information as an optimization factor to discriminate which ICs display the most significant differences across groups of individuals. Going one step further from that we now propose an optimization factor based on the characterization of the differences in the across-group distribution functions for each modality individually and jointly as well. In order to accomplish that we (1) compute estimates of the (joint) distribution(s) of the different modalities for each group and thereafter (2) propose the use of a divergence metric on the estimated group distributions. Due to their metric property and optimality (Hero et. al.) special attention is being devoted to the behavior aspects of the J-divergence and Alpha divergence (with $\alpha = 0.5$), respectively. We also focus on the validation of results through the study of hybrid simulated datasets as well as by the analysis of real diagnostic datasets on schizophrenia studies.

PRIMARY RESULTS

Divergence measurements provide information on how different two distributions are from each other. In order to consistently validate the use of divergence measures a simulation framework has been developed as to simulate the selection of an IC. We achieve that by generating simulated activation in the form of ellipsoidal (known) sources that may be added to real diagnostic functional datasets on subjects of a given study. These sources are 21x21 half-cycle sinusoids added to the same section of each modality after multiplication by a random number drawn from a uniform distribution (the loading parameter) with half of the individuals (randomly chosen) having a mean shifted by some amount from the other half of the individuals. Extra variability is added to each group in terms of shape and displacement of the known source so that each realization actually differs from subject to subject. Finally joint ICs are estimated using jICA and then analyzed as depicted in Figure 2.

Figure 2 (right): Use of divergence measures for IC selection – The simulated sources generated for each individual undergo group-averaging across each feature yielding g_1 , g_2 , and **both** (the average of g_1 and g_2). These results are interpreted as ground (known) ‘truths’ for comparison to the estimated sources or ICs (u). Histograms are computed for each ground truth and each IC. Measures of divergence can be interpreted as the amount of information not contained in the IC as compared to the ground truth. The results show that when an IC correctly approximates the ground truth the distributions look similar and therefore divergence outputs very small values. These particular results explore the behavior of several divergence measures such as the Kullback-Leibler (KL), J, and Alpha (or Renyi) divergences. Particularly, computing the divergence on each feature individually yielded better results than on their concatenation.



REFERENCES

Calhoun, V. D. et. al. (2006): A method for multitask fMRI data fusion applied to schizophrenia. *Hum. Brain Map.* 27, 598-610.
 Calhoun VD, Adali T. (2006a): Unmixing fMRI with independent component analysis. *IEEE Eng Med Biol Mag* 25(2):79-90.
 Hyvarinen, A. et. al. (2001): Independent component analysis. John Wiley & Sons, New York, NY.
 McKeown, M. J. et. al. (2003): Independent component analysis of functional MRI: what is signal and what is noise? *Curr. Opin. Neurobiol.* 13, 620-629.
 Savopol F. et. al. (2002): Heterogeneous data for emergency mapping: data integration or data fusion? *Proc. ISPRS.*
 Hero, A. et. al. (2001): Divergence for classification, indexing and retrieval. 328.