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

R. F. Silva1,2, and V. D. Calhoun1,2
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.

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.) these divergence measures 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 known source so that each realization actually differs from subject to subject. Finally joint ICs are estimated using jICA and then are analyzed as depicted in Figure 2.

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 are analyzed as depicted in Figure 2.

REFERENCES
Savopol F. et. al. (2002): Heterogeneous data for emergency mapping: data integration or data fusion? Proc. ISPRS.