

Feature-based Fusion of Medical Imaging Data

V. D. Calhoun, *Senior Member, IEEE*, T. Adalı, *Fellow, IEEE*

Abstract— The acquisition of multiple brain imaging types for a given study is a very common practice. There have been a number of approaches proposed for combining or fusing multi-task or multi-modal information. These can be roughly divided into those that attempt to study convergence of multimodal imaging, for example how function and structure are related in the same region of the brain, and those that attempt to study the complementary nature of modalities, for example utilizing temporal electroencephalography (EEG) information and spatial functional magnetic resonance imaging (fMRI) information. Within each of these categories, one can attempt data integration (the use of one imaging modality to improve the results of another) or true data fusion (in which multiple modalities are utilized to inform one another). We review both approaches and present a recent computational approach that first preprocesses the data to compute features of interest. The features are then analyzed in a multivariate manner using independent component analysis. We describe the approach in detail and provide examples of how it has been used for different fusion tasks. We also propose a method for selecting which combination of modalities provide the greatest value in discriminating groups. Finally, we summarize and describe future research topics.

Index Terms— fMRI, EEG, ICA, multivariate data analysis, data fusion

I. INTRODUCTION

Many studies are currently collecting multiple types of imaging data and information from the same participants. Each imaging method reports on a limited domain and is likely to provide some common information and some unique information. This motivates the need for a joint analysis of these data. Most commonly, each type of image is analyzed independently and then perhaps overlaid to demonstrate its relationship with other data types (e.g., structural and functional images). A second approach, called data fusion, utilizes multiple image-types together in order to take advantage of the ‘cross’-information. In the former approach, any cross-information is ‘thrown’ away, hence such an approach, for example, would not detect a change in

functional magnetic resonance imaging (fMRI) activation maps that are associated with a change in the brain structure while the second approach would be expected to detect such changes.

Many studies are currently collecting multiple types of imaging data from the same participants. Each imaging method reports on a limited domain and typically provides both common and unique information about the problem in question. Approaches for combining or fusing data in brain imaging can be conceptualized as having a place on an analytic spectrum with meta-analysis (highly distilled data) to examine convergent evidence at one end and large-scale computational modeling (highly detailed theoretical modeling) at the other end [1]. In between are methods that attempt to perform a direct data fusion [2].

Current approaches for combining different types of imaging information for the most part elect to constrain one type with another type of information—as in electroencephalography (EEG) [3, 4] or diffusion tensor imaging (DTI) [5, 6] being constrained by fMRI or structural MRI (sMRI) data. While these are powerful techniques, a limitation is that they impose potentially unrealistic assumptions upon the EEG or DTI data, which are fundamentally of a different nature than the fMRI data. An alternative approach—which we call data integration [7, 8]—is to analyze each data type separately and overlay them—thereby not allowing for any interaction between the data-types. For example, a data integration approach would not detect a change in fMRI activation maps that is related to a change in brain structure (in the example we provide the change is in gray matter concentration between patients and controls). One promising direction is to take an intermediate approach in which the processing of each image type is performed using features extracted from different modalities. These features are then examined for relationships among the data types at the group level (i.e., variations among individuals) and specifically, differences in these variations between patients and controls. This approach allows us to take advantage of the ‘cross’-information among data types [7, 8].

Methods such as structural equation modeling (SEM) or dynamic causal modeling (DCM) [9-11] can be used to examine the correlational structure between regions activated by different tasks [12] or between functional and structural variables [13, 14]. Such approaches are useful for model testing; however these approaches do not provide an examination of the full set of brain voxels, nor do they allow testing of unknown connections. Alternatively, one could choose to examine correlation (and potentially extend to

Manuscript received June 19, 2008. This research was supported in part by grants from the National Institutes of Health under grants 1 R01 EB 000840, R01 EB 005846, and R01 EB 006841 and the National Science Foundation under grant NSF-IIS 0612076.

V. D. Calhoun is with the Mind Research Network and the University of New Mexico, Dept. of ECE, Albuquerque, NM 87131 (phone: 505-272-1817; fax: 505-272-8248; e-mail: vcalhoun@unm.edu).

T. Adalı is with the University of Maryland Baltimore County, Dept. of CSEE, Baltimore, MD 21250 (e-mail: adali@umbc.edu).

nonlinear relationship through the use of other criteria such as mutual information) between all points of the data. This approach has been applied to examine functional connectivity in fMRI by computing a 6D matrix of correlations [15]. Such computations are straightforward; however the drawback is that they are high in dimensionality and hence potentially difficult to interpret. A natural set of tools for avoiding the disadvantages of the above techniques include those that transform data matrices into a smaller set of modes or components. Such approaches include those based upon singular value decomposition (SVD) [16, 17] as well as more recently, independent component analysis (ICA) [18]. An advantage of ICA over variance-based approaches like SVD or principal component analysis (PCA) is the use of higher-order statistics to reveal hidden structure [19, 20]. We have recently done work showing the value of combining multi-task fMRI data [21], fMRI and sMRI data [22], and fMRI and ERP data [23]. One important aspect of the approach is that it allows for the possibility that a change in a certain location in one modality is associated with a change in a different location in another modality (or, in the case of ERP one is associating time in ERP with space in fMRI) as we demonstrate with a number of examples in this paper.

In this paper, we first review the basic approaches for fusing information from multiple medical imaging data types. Next, we present a feature-based fusion approach that provides a general framework for fusing information from multiple data types, such as multi-task fMRI data, or fMRI and event-related potential (ERP) data. The extracted features for each data type are fused using a data-driven analysis technique, independent component analysis (ICA), which has proved quite fruitful for medical image analysis [23-27]. The fusion framework we present thus enables the discovery of relationships among data types for given samples, for example, at the group level, to study variations between patients and controls. In the following sections, after a background of some of the common fusion approaches, we introduce the feature-based fusion framework. In particular, we discuss the nature of features and their computation. Next, we present several examples showing the fusion of data from different modalities. The final section discusses the importance of selecting the important features for the joint analysis.

II. A BRIEF DESCRIPTION OF IMAGING MODALITIES AND FEATURE GENERATION

fMRI measures the hemodynamic response related to neural activity in the brain. sMRI provides information about the tissue type of the brain—gray matter (GM), white matter (WM), cerebrospinal fluid (CSF). A different type of structural information is captured by diffusion tensor imaging (DTI) which measures the diffusion of water in the brain and provides information about fiber direction. Another useful measure of brain function is EEG, which measures brain electrical activity with a higher temporal resolution than fMRI (and lower spatial resolution).

A. fMRI

Functional magnetic resonance imaging data provide a measure of brain function on a millimeter spatial scale and a sub-second temporal scale. There are a considerable number of available fMRI processing strategies [28, 29]. Two primary approaches include model-based approaches assuming certain hemodynamic properties and often utilizing the general linear model (GLM) [30], and data-driven approaches; one which has proven particularly fruitful is ICA [18, 31], which does not impose a priori constraints upon the temporal hemodynamic evolution.

A strength of GLM approaches is that they allow one to perform specific hypothesis tests, e.g., “where in the brain do these temporal patterns (i.e. the activation) occur?” In contrast, a strength of ICA is its ability to characterize fMRI activations without an a priori hemodynamic model in an exploratory manner, e.g., “what are the temporal and spatial patterns occurring in the brain?” Both approaches have obvious advantages. We next give a brief discussion of preprocessing and GLM analysis.

There are a number of preprocessing steps important for fMRI. Phase correction is often applied because each slice is typically acquired sequentially, rather than acquiring all slices simultaneously [32, 33]. Registration is also required because of subject motion during an fMRI experiment. There are numerous algorithms for estimating and correcting for this motion including those based upon Fourier methods [34], Taylor approximations, Newton’s method [35], and others. The third preprocessing stage, normalization, is necessary to 1) compare brains across different individuals, and 2) use standardized atlases to identify particular brain regions. There are also many methods for applying spatial normalization including maximum likelihood and Newton’s methods [36] as well as localized methods.

The most common analysis approach for fMRI is based upon the general linear model, assuming a specific form for the hemodynamic response. In the simplest case the data are modeled as:

$$\mathbf{y}_m = \sum_{i=1}^R \mathbf{x}_i \beta_{mi} + \varepsilon_m \quad (1)$$

for R regressors, where \mathbf{y}_m , \mathbf{x}_i , and ε_m are $K \times 1$ for time points $k = 1, 2, \dots, K$ at brain locations $m = 1, 2, \dots, M$. The error is typically modeled as Gaussian, independent and identically distributed, zero-mean, with variance σ_v^2 .

B. sMRI

We define structural MRI analysis as the acquisition and processing of T1-, T2-, and/or proton density-weighted images. Multiple structural images are often collected to enable multispectral segmentation approaches. Both supervised and automated segmentation approaches have been developed for sMRI analysis [37-39]. The near-exponential pace of data collection [40] has stimulated the development of structural image analysis. Advanced methods include the

rapidly growing field of computational anatomy [41-43]. This field combines new approaches in computer vision, anatomical surface modeling [42, 44], differential geometry [41], and statistical field theory [45, 46] to capture anatomic variation, encode it, and detect group-specific patterns. Other approaches include voxel-based methods [47] and manual region-of-interest approaches. Each technique is optimized to detect specific features, and has its own strengths and limitations.

The primary outcome measure in a structural image may include a measure of a particular structure (e.g. volume or surface area) or a description of the tissue type (e.g. gray matter or white matter). There are many methods for preprocessing sMRI data which may include bias field correction [intensity changes caused by radio frequency (RF) or main magnetic field (B_0) inhomogeneities] [48, 49], spatial linear or non-linear [50] filtering normalization. MR images are typically segmented using a tissue classifier producing images showing the spatial distribution of gray matter, white matter, and CSF. Tissue classifiers may be supervised (where a user selects some points representing each tissue class to guide classification) or unsupervised (no user intervention). Bayesian segmentation methods [47, 51, 52] assign each image voxel to a specific class based on its intensity value as well as prior information on the likely spatial distribution of each tissue in the image. The classification step may be preceded by digital filtering to reduce intensity inhomogeneities due to fluctuations and susceptibility artifacts in the scanner magnetic field. In expectation-maximization (EM) techniques, RF correction and tissue classification steps are combined, using one to help estimate the other in an iterative sequence [37, 38].

C. DTI

Diffusion MRI is a technique that measures the extent of water diffusion along any desired direction in each voxel [53]. Such measurements have revealed that diffusion of brain water has strong directionality (anisotropy) attributed to the presence of axons and/or myelination [54]. Diffusion of brain water is often confined to a direction parallel to neuronal fibers. If there is a region where fibers align in a direction, diffusion of water may be restricted to a direction perpendicular to the fibers and tend to diffuse parallel to them. The properties of such water diffusion can be expressed mathematically as a “3 x 3 tensor” [55]. The tensor can be further conceptualized and visualized as an ellipsoid, the three main axes of which describe an orthogonal coordinate system. This ellipsoid can be characterized by six parameters; diffusion constants along the longest, middle, and shortest axes (λ_1 , λ_2 , and λ_3) and the directions of those axes. Once the diffusion ellipsoid is fully characterized at each pixel of the brain images, local fiber structure can be deduced. For example, if $\lambda_1 \gg \lambda_2 > \lambda_3$ (diffusion is anisotropic), it suggests the existence of dense and aligned fibers within each pixel, whereas isotropic diffusion ($\lambda_1 \sim \lambda_2 \sim \lambda_3$) suggests sparse or not aligned fibers. When diffusion is anisotropic, the

direction of λ_1 tells the direction of the fibers. The degree of anisotropy can be quantified using these parameters obtained from diffusion MRI measurements [56, 57].

Among the metrics for quantifying diffusion anisotropy, fractional anisotropy (FA) is considered to be the most robust [58].

$$FA = \frac{\sqrt{3}}{\sqrt{2}} \frac{\sqrt{(\lambda_1 - \lambda)^2 + (\lambda_2 - \lambda)^2 + (\lambda_3 - \lambda)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \quad (2)$$

$$\text{where } \lambda = (\lambda_1 + \lambda_2 + \lambda_3)/3$$

Within the constraints of in-plane resolution, some regions of white matter normally have very high FA, and this probably represents architectural differences in fiber tract organization at the intra-voxel level, *i.e.*, intact fibers crossing within a voxel. Many pathologic processes that cause changes at the micro-structural level, such as demyelination and corruption of microtubules, are likely to cause a significant measurable decrease in FA due to the diminished intra-voxel fiber incoherence.

D. EEG

Electroencephalography is a technique that measures brain function by recording and analyzing the scalp electrical activity generated by brain structures. Like MRI, it is a noninvasive procedure that can be applied repeatedly in patients, normal adults, and children with virtually no risks or limitations. Local current flows are produced when brain cells are activated. It is believed that contributions are made by large synchronous populations although it is not clear if small populations also make a contribution. The recorded electrical signals are then amplified, digitized, and stored.

Event-related potentials (ERPs) are small voltage fluctuations resulting from evoked neural activity and are one of many ways to process EEG data. These electrical changes are extracted from scalp recordings by computer averaging epochs (recording periods) of EEG time-locked to repeated occurrences of sensory, cognitive, or motor events. The spontaneous background EEG fluctuations, which are typically random relative to when the stimuli occurred, are averaged out, leaving the event-related brain potentials. These electrical signals reflect only that activity which is consistently associated with the stimulus processing in a time-locked way. The ERP thus reflects, with high temporal resolution, the patterns of neuronal activity evoked by a stimulus.

Due to their high temporal resolution, ERPs provide unique and important timing information about brain processing and are an ideal methodology for studying the timing aspects of both normal and abnormal cognitive processes. More recently, ICA has been used to take advantage of EEG activity that may be averaged out by computing an ERP [59]. Magnetoencephalography (MEG) is a complementary technique that senses the magnetic field produced by synchronously firing neurons. The MEG system is much more expensive, requiring superconducting sensors, but also has the advantage that the magnetic field is not attenuated by the scalp and skin.

III. BRAIN IMAGING FEATURE GENERATION

Often it is useful to use existing analysis approaches to derive a lower dimensional feature from the imaging data. These features can then be analyzed in order to integrate or fuse the information across multiple modalities. The data types on which we focus in this paper are: fMRI, sMRI [including T1- and T2-weighted scans], and EEG. Processing strategies for each of these data-types have been developed over a number of years. Each data type, after being preprocessed as above, is reduced into a feature, which contributes an input vector from each modality for each subject and each task to the joint ICA framework we introduce in this paper. A feature is a sub dataset extracted from one type of data, related to a selected brain activity or structure. A summary of some features is provided in Table 1.

Modality	Core-Feature
fMRI	Recognition related activity [77]
SB task	Encode-related activity [77]
fMRI	Target-related activity [78]
AOD task	Novel-related activity [78]
sMRI	GM concentration [79]
	WM concentration [79]
	CSF concentration [79]
EEG	Target-related ERP [80]
AOD task	Novel-related ERP [80]
DTI	Fractional anisotropy [58]

Table 1: Core features for fMRI (Sternberg working memory task [SB], auditory oddball task [AOD]), sMRI, and EEG (AOD task): The fMRI features are computed using the SPM software and represent fMRI activity in response to certain stimuli. The sMRI measures are the result of segmentation of T1-weighted brain images, and the EEG features are time-locked averages to the stimuli.

In the table, the first two modalities are fMRI, each of which has two stimuli which activate the brain differently. The first is a Sternberg task (SB) and the second is an auditory oddball task (AOD). The third modality is structural MRI for which three tissue segmentations can be computed. The fourth modality is EEG, also collected during an AOD task. The final modality is diffusion tensor imaging, for which a measure of fractional anisotropy is computed.

IV. FEATURE-BASED FUSION FRAMEWORK USING ICA

Independent component analysis is a statistical method used to discover hidden factors (sources or features) from a set of measurements or observed data such that the sources are maximally independent. Typically, it assumes a generative model where observations are assumed to be linear mixtures of independent sources, and unlike PCA that uncorrelates the data, ICA works with higher-order statistics to achieve independence. A typical ICA model assumes that the source signals are not observable, statistically independent and non-Gaussian, with an unknown, but linear, mixing process. Consider an observed M – dimensional random vector denoted by $\mathbf{x} = [x_1, \dots, x_M]^T$ which is generated by the ICA

model:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (3)$$

where $\mathbf{s} = [s_1, s_2, \dots, s_N]^T$ is an N -dimensional vector whose elements are assumed independent sources and $\mathbf{A}_{M \times N}$ is an unknown mixing matrix. Typically $M \geq N$, so that \mathbf{A} is usually of full rank. The goal of ICA is to estimate an unmixing matrix $\mathbf{W}_{N \times M}$ such that \mathbf{y} (defined in equation (4)) is a good approximation to the ‘true’ sources, \mathbf{s} :

$$\mathbf{y} = \mathbf{W}\mathbf{x} \quad (4)$$

Independent component analysis has shown to be useful for fMRI analysis for several reasons. Spatial ICA finds systematically non-overlapping, temporally coherent brain regions without a specific assumption about the shape of the temporal response. The temporal dynamics of many fMRI experiments are difficult to study with fMRI due to the lack of a well-understood brain-activation model. ICA can reveal inter-subject and inter-event differences in the temporal dynamics. A strength of ICA is its ability to reveal dynamics for which a temporal model is not available [60]. Spatial ICA also works well for fMRI as it is often the case that one is interested in spatially distributed brain networks.

Independent component analysis has demonstrated considerable promise for the analysis of fMRI [18], EEG [61], and sMRI [62] data. In this section, we present a data fusion framework utilizing ICA, which we call the joint ICA (jICA). Note that the ICA approach we described earlier for fMRI data is a first level analysis (*i.e.*, is applied directly to the 4D data without reduction into a feature, and though the basic algorithm is similar, with the same basic assumptions, the application details are different from the ICA we propose to utilize at the second level, on the generated features). An amplitude map generated by ICA at the first level would be considered a feature similar to an amplitude map generated by the GLM approach.

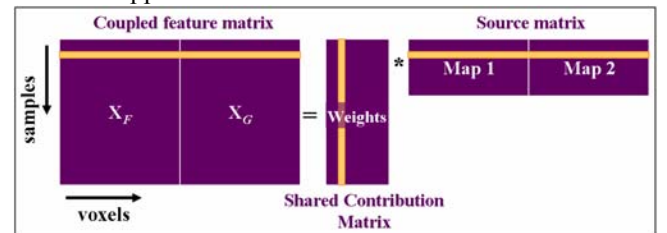


Figure 1: Illustration of model in which loading parameters are shared among features: The feature matrix is organized by placing the features (e.g. SPM map and GM map) from the two modalities side by side (with one row containing data collected from the same subject for both modalities). This matrix is then modeled as containing spatially independent joint source images which share common mixing matrix parameters.

Given two sets of data (can be more than two, for simplicity, we first consider two), \mathbf{X}_F and \mathbf{X}_G , we concatenate the two datasets side-by-side to form \mathbf{X}_J and write the likelihood as

$$L(\mathbf{W}) = \prod_{n=1}^N \prod_{v=1}^V p_{J,n}(u_{J,v}), \quad (5)$$

where $\mathbf{u}_j = \mathbf{W}\mathbf{x}_j$. Here, we use the notation in terms of random variables such that each entry in the vectors \mathbf{u}_j and \mathbf{x}_j correspond to a random variable, which is replaced by the observation for each sample $n = 1, \dots, N$ as rows of matrices \mathbf{U}_j and \mathbf{X}_j . When posed as a maximum likelihood problem, we estimate a *joint* demixing matrix \mathbf{W} such that the likelihood $L(\mathbf{W})$ is maximized.

Let the two datasets \mathbf{X}_F and \mathbf{X}_G have dimensionality $N \times V_1$ and $N \times V_2$, then we have

$$L(\mathbf{W}) = \prod_{n=1}^N \left(\prod_{v=1}^{V_1} p_{F,n}(u_{F,v}) \prod_{v=1}^{V_2} p_{G,n}(u_{G,v}) \right), \quad (6)$$

Depending on the data types in question, the above formula can be made more or less flexible.

This formulation assumes that the sources associated with the two data types (F and G) modulate the same way across N samples (usually subjects). This is a strong constraint; however it has a desirable regularization effect to the problem simplifying the estimation problem significantly, which is important especially when dealing with different data types. Also, the framework provides a natural link to two types of data by constraining the contributions to be similar. In addition, it is important to normalize the two data types independently so that they have similar contributions to the estimation and that $V_1 \approx V_2$. The normalization process is important and should be modality specific (see examples in [22, 23, 27, 63]). The assumption of the same linear covariation for both modalities is fairly strong, however we have demonstrated their utility in a variety of cases and they appear to provide meaningful results [22, 23, 27, 63, 64]. In addition, we are exploring other formulations which do not require the same linear covariation for the different modalities [65, 66].

The underlying assumptions for the form given in Equation (5) depend on the data types used for F and G . For example, when the two data types belong to the same data type but different tasks, the assumption of $p_J = p_F = p_G$ is more plausible than when dealing different data types. On the other hand, when little is known about the nature of the source distributions in a given problem, imposing a distribution of the same form provides significant advantages yielding meaningful results as we demonstrate with an fMRI-ERP fusion example. In addition, certain nonlinear functions such as the sigmoid function has been noted as providing a robust solution to the ICA problem providing a good match for a number of source distributions, especially when they assume super-Gaussian statistics.

Hence, there are different ways to relax the assumptions made in the formulation above, such as instead of constraining the two types of sources to share the same mixing coefficients, *i.e.*, to have the same modulation across N samples, we can require that the form of modulation across samples for the sources from two data types are correlated but not necessarily

the same. We have implemented such an approach, called parallel ICA [66, 67].

V. APPLICATION OF THE FUSION FRAMEWORK

In this section, we show examples of the application of jICA introduced in section 5 to real data from multiple tasks/modalities using: 1) multi-task fMRI data, 2) fMRI/sMRI, 3) ERP/fMRI, sMRI/DTI data. Furthermore, we address the selection of best input features for the jICA data fusion to achieve the best performance, in terms of identifying components which convey differences between two groups.

In the examples we present, fMRI data were preprocessed using the software package SPM2 [68]. SMRI data were segmented into GM, WM and CSF images using the same program. ICA was used to remove ocular artifacts from the EEG data [69]. The EEG data, then, were filtered with a 20Hz low pass filter. ERPs were constructed for trials in which participants correctly identified target stimuli, from the midline central position (Cz) because it appeared to be the best single channel to detect both anterior and posterior sources.

A. Multi-task fMRI

We performed a joint analysis of fMRI data collected from a Sternberg (SB) task and an auditory oddball (AOD) task. Data in each task were collected from 15 controls and 15 patients with schizophrenia. Additional details of the tasks and subjects are provided in [27]. A single joint component was found to discriminate schizophrenia patients and healthy controls. A joint histogram was computed by ranking voxels surviving the threshold for the AOD and SB parts of the joint source in descending order and pairing these two voxel sets. Single subject and group-averaged joint histograms are presented in Figure 2 (a) and (b) and the marginal histograms for the AOD and SB tasks are presented in Figure 2 (c) and (d).

In general, more AOD task voxels were active in the controls and the SB task showed a slight increase standard deviation for the patients. Results also revealed significantly more correlation between the two tasks in the patients ($p < 0.000085$). A possible synthesis of the findings is that patients are activating less, but also activating with a less unique set of regions for these very different tasks, consistent with a generalized cognitive deficit.

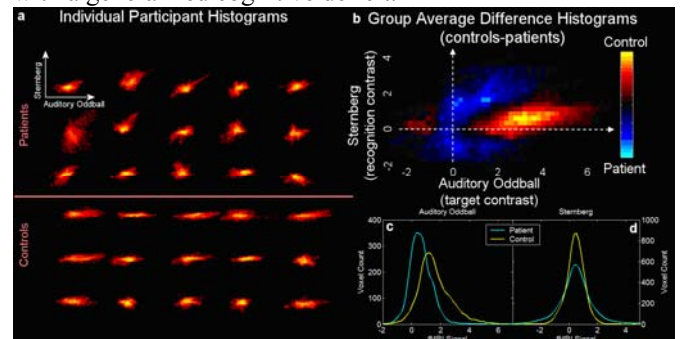


Figure 2: Cross-Task 2D Histograms for AOD versus SB fMRI Activation: Joint 2D histograms for voxels identified in the analysis. Individual (a) and

group average difference (b) histograms [with orange areas larger in controls and blue areas larger in patients] are provided along with the marginal histograms for the auditory oddball (SPM contrast image for ‘targets’) (c) and Sternberg (SPM contrast image for ‘recall’) (d) data.

B. fMRI-sMRI

It is also feasible to use jICA to combine structural and functional features. Our approach requires acceptance of the likelihood of gray matter changes being related to functional activation. This is not an unreasonable premise when considering the same set of voxels [70], or even adjacent voxels [13] but as the current study shows, it also requires the acceptance of related gray matter regions and functional regions which are spatially remote. Given the functional interconnectedness of widespread neural networks, we suggest that this, also, is a reasonable conception for the relationship between structural and functional changes.

The next example is from a joint-ICA analysis of fMRI data of auditory oddball task and gray matter segmentation data [22]. Auditory oddball target activation maps and segmented gray matter maps were normalized to a study specific template in order to control for intensity differences in MR images based on scanner, template and population variations [71].

Results are presented in Figure 3. The AOD part of the joint source is shown in Figure 3a, the GM part of the joint source is shown in Figure 3b, and the ICA loading parameters separated by group and shown in Figure 3c. Only one component demonstrated significantly different loadings ($p \sim 0.0012$) in patients and controls (loading for controls was higher than that for patients). Different regions were identified for the fMRI and sMRI data. For display, auditory oddball and gray matter sources were converted to Z-values and thresholded at $|Z| > 3.5$.

The main finding was that the jICA results identified group differences in bilateral parietal and frontal as well as right temporal regions in gray matter associated with bilateral temporal regions activated by the auditory oddball target stimulus. This finding suggests gray matter regions that may serve as a morphological substrate for changes in (functional) connectivity. An unexpected corollary to this finding was that, in the regions showing the largest group differences, gray matter concentrations were increased in patients versus controls, suggesting that these increases are somehow related to decreases in functional connectivity in the AOD fMRI task.

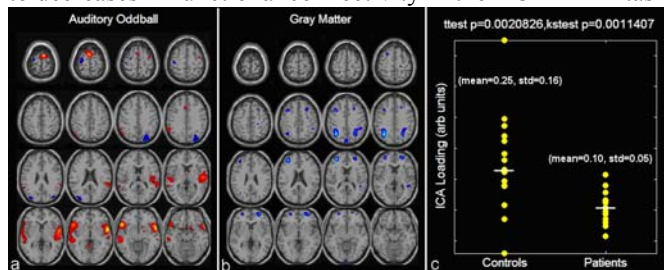


Figure 3: Auditory Oddball/Gray Matter jICA Analysis: Only one component demonstrated a significant difference between patients and controls. The joint source map for the auditory oddball (left) and gray matter (middle) data is presented along with the loading parameters for patients and controls (far right).

C. fMRI-ERP

The feature-based jICA framework was used for ERP and fMRI data collected from 23 healthy controls and 18 chronic schizophrenia patients, during the performance of the AOD task. Fifteen joint components were estimated from the target-related ERP time courses and fMRI activation maps via the jICA. One joint component was found to distinguish patients and controls using a two-sample t-test ($p < 0.0001$) on patient and control loading parameters. This identified component shows a clear difference in fMRI at bilateral frontotemporal regions implicated in schizophrenia (Figure 4 right) and in ERP at times during the N2/P3 complex (Figure 4 left) which have been previously implicated in patients.

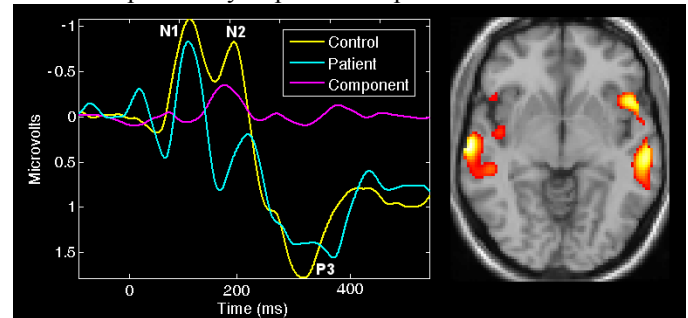


Figure 4: ERP/fMRI jICA: Joint component which showed significantly different loading parameters ($p < 0.0001$) for patients versus controls.: Left: control (yellow) and patient (blue) average ERP plots along with the ERP part of the identified joint component (pink). Right: Thresholded fMRI part of the joint component showing bilateral temporal and frontal lobe regions.

In the same way as for Figure 2 significant voxels/time points were used to generate an ERP vs fMRI histogram for controls (orange) and patients (blue), shown in Figure 5. The controls are clearly showing increases in both fMRI and ERP data.

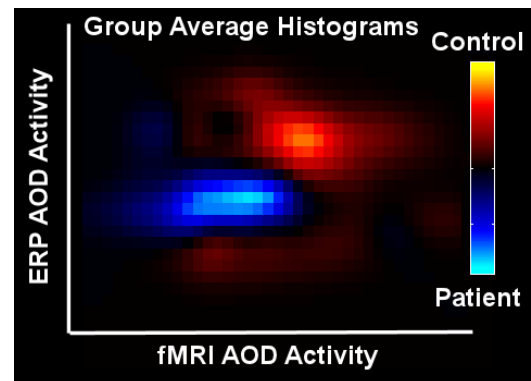


Figure 5: ERP/fMRI Histograms: (Left) Joint histograms for patients (blue) and controls (orange). (Right) Simulated data from two Gaussians (a) showing a case in which marginal histograms (b, c) are less able to detect differences between groups whereas the histogram in the direction of maximal separation (d) clearly shows the two distributions from patients and controls

6.4. sMRI-DTI

We now present an example showing gray matter and fractional anisotropy maps in a joint ICA analysis in eleven participants. Seven components were estimated and a template of the occipital lobe generated from a previous study was used

to select the FA and joint GM map. A picture of the resulting joint source (for each modality, the slice corresponding to the maximal coupling is displayed) is shown in Figure 6 and demonstrates GM regions in occipital lobe and bilateral lateral geniculate regions are associated with white matter occipital lobe regions consistent with the optic radiations (that is, higher FA in optic radiations is associated with lower GM values in lateral geniculate and visual cortex).

The group was then split into an old (mean age 63 ± 10) and young (mean age 44 ± 10) cohort. The weight parameter calculated from the earlier ICA estimation is plotted in Figure 6 (right) as a function of cohort membership. A highly significant difference was observed with young participants showing higher FA and GM (see) and older participants showing lower FA and GM. This is consistent with loss of gray matter volume (and white matter FA) with age.

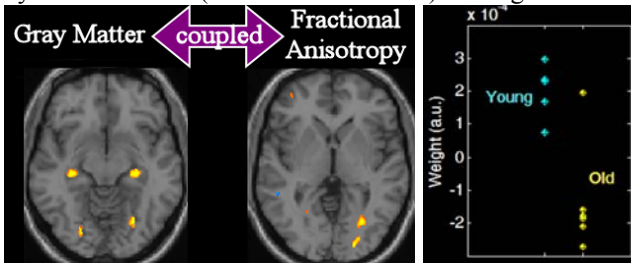


Figure 6: GM and FA Spatial Correspondence (left) and corresponding weights (right) comparing older and younger participants ($p < 0.002$)

Analyzing GM maps and FA maps together can be challenging as the FA images are warped due to the various gradient directions used. In the above analysis, GM and FA are coupled together at an image level, but not at a voxel level. Thus misregistration between image-types will not directly affect the results.

D. 6.5. Parallel ICA

As discussed in Section 5, the strong regularization imposed by the jICA framework can be relaxed in a number of ways to allow for more flexibility in the estimation. One such approach we investigated is called parallel independent component analysis (paraICA). As a framework to investigate the integration of data from two imaging modalities, this method is dedicated to identify components of both modalities and connections between them through enhancing intrinsic interrelationships. We have applied this approach to link fMRI/ERP data and also fMRI and genetic data (single nucleotide polymorphism arrays) [65-67]. Results show that paraICA provides stable results and can identify the linked components with a relatively high accuracy.

The result for fMRI/ERP data is consistent with that found by the joint ICA algorithm [72], where a shared mixing matrix is used for both modalities. The fundamental difference is that paraICA assumes the fMRI and ERP data are mixed in a similar pattern but not identically. ParaICA pays more attention to individually linked components and their connections, while the joint ICA studies inter-effects between EEG and fMRI as a whole [72]. It provides a promising way to analyze the detail coupling between hemodynamics and neural activation.

The methods in this section along with some example data are available in a new Matlab toolbox called Fusion ICA Toolbox or FIT (<http://icatb.sourceforge.net>).

VI. SELECTION OF JOINT COMPONENTS

In some cases, it is important to define criteria for selecting among joint components. For example, when studying two groups, *e.g.*, patients and controls, we may want to determine which combination of joint features is optimal in some sense. We apply our approach to the problem of identifying image-based biomarkers in schizophrenia. A *biomarker* is a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention. In our proposal, a biomarker would be the joint component resulting from the optimal combination of fMRI, EEG and sMRI features. We have used two criteria for this, separation of the mixing parameters and separation of the source distributions.

The estimated source distributions for these components are computed separately for each group and the Kullback-Leibler (KL) divergence is computed between the patient (sz) and control (hc) distributions (*i.e.*, $D_{KL}(p_{sz}(\mathbf{f}) \| p_{hc}(\mathbf{f}))$, where \mathbf{f} is a multi-dimensional feature/modality vector [73]). There are several possible divergence measures including J, KL, and Renyi. J divergence is simply a symmetric KL divergence, such that:

$$D_J = (D_{KL}(p, q) + D_{KL}(q, p)) / 2.$$

Typical reasons for selecting one divergence over another are the ability to express the solution analytically given a particular assumed distributional form or to sensitize the divergence to particular distributional features [*e.g.*, minimizing the Renyi (alpha) divergence with $\alpha = 0.5$ has been shown to be optimal for separating similar feature densities [74]; the limiting case of $\alpha = 1$ results in the KL divergence].

We compute features for 1) AOD target-related fMRI activity (AOD_T), 2) AOD novel-related fMRI activity (AOD_N), 3) SB recognition fMRI activity (SB), and 4) sMRI GM values (GM). All four of these features were collected on each of 15 controls and 15 patients with schizophrenia.

In order to evaluate the impact of the choice of divergence measures, we have computed results for several divergence measures. An example of the results is shown in Table 2. The number in parenthesis indicates the relative gain (ratio of current row/next row) where larger numbers indicate better patient/control separation. We compare the KL and Renyi divergences with alpha = 0.5, as well as their symmetric counterparts.

	KL	J	Renyi (0.5)
Oddball & Sternberg	0.497 (7.4)	0.480 (5.9)	0.2874 (7.4)
Oddball	0.067 (1.3)	0.081 (1.6)	0.039 (1.5)
Sternberg	0.050	0.050	0.0254

Table 2: Several divergence measures for three combinations of fMRI

features

The results shown in Figure 7 are ranked according to a divergence measure, in order to determine which combination of features/modalities provides better separation. For example, the result shows us that combining the SB feature with the AOD_N or AOD_T features provides increased separation beyond SB or AOD_T alone. It also suggests that the incorporation of GM tends to decrease the separation. Note that though we have demonstrated above a comparison of patients and controls, this method is a general one, useful for studying a variety of questions. For example, instead of patient versus controls, we could have examined age-related activity only in healthy controls.

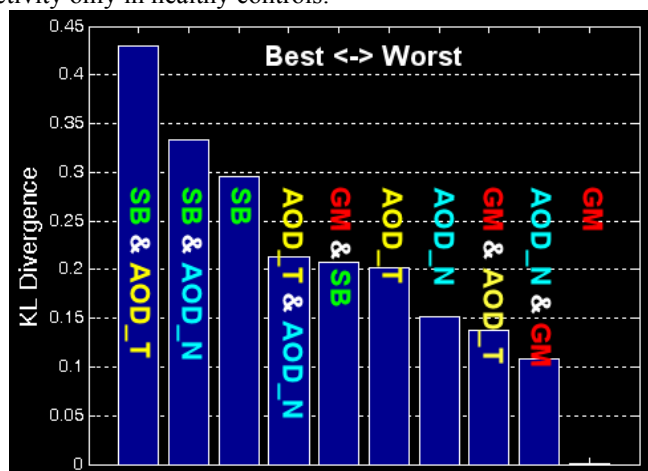


Figure 7: Example of evaluation of KL Divergence for combinations of features comparing 2-way fusion with non-fusion. Larger KL Divergence values represent better patient versus control separation.

In summary, we have presented a data fusion approach based upon ICA and have shown multiple examples which demonstrate the feasibility of combining multimodal data. The results we show appear to be clinically plausible and may be useful in improving our understanding of schizophrenia. There is also some possibility to use multimodal approaches to develop new tools for classification and treatment prediction in schizophrenia. We have made some progress in this regard [75, 76] however much more work is needed including validation and comparison with currently used approaches in the clinic. The methods and much of the data shown in the paper are provided as part of our Matlab toolbox called Fusion ICA Toolbox (FIT) downloadable at <http://icatb.sourceforge.net>.

VII. CONCLUSION

We present a general framework for combining different types of brain imaging data at the group level via features computed from each data type. We also show that by combining modalities in certain ways performance is improved. This approach enables us to take advantages of the strengths and limitations of various modalities in a unified analytic framework and demonstrates that data fusion techniques can be successfully applied to joint brain imaging data to reveal unique information that cannot be evaluated in

any one modality.

REFERENCES

- [1] F. T. Husain, G. Nandipati, A. R. Braun, L. G. Cohen, M. A. Tagamets, and B. Horwitz, "Simulating transcranial magnetic stimulation during PET with a large-scale neural network model of the prefrontal cortex and the visual system," *NeuroImage*, vol. 15, pp. 58-73, 2002.
- [2] B. Horwitz and D. Poeppel, "How can EEG/MEG and fMRI/PET data be combined?," *Hum.Brain Mapp.*, vol. 17, pp. 1-3, 2002.
- [3] A. M. Dale, E. Halgren, J. D. Lewine, R. L. Buckner, K. Paulson, K. Marinkovic, and B. R. Rosen, "Spatio-temporal localization of cortical word repetition effects in a size judgement task using combined fMRI/MEG," *NeuroImage*, vol. 5, pp. S592-S592, 1997.
- [4] J. S. George, C. J. Aine, J. C. Mosher, D. M. Schmidt, D. M. Ranken, H. A. Schlitt, C. C. Wood, J. D. Lewine, J. A. Sanders, and J. W. Belliveau, "Mapping function in the human brain with magnetoencephalography, anatomical magnetic resonance imaging, and functional magnetic resonance imaging," *J.Clin.Neurophysiol.*, vol. 12, pp. 406-431, 1995.
- [5] D. S. Kim, I. Ronen, E. Formisano, K. H. Kim, M. Kim, P. van Zijl, K. Ugurbil, S. Mori, and R. Goebel, "Simultaneous mapping of functional maps and axonal connectivity in cat visual cortex," in *Proc.HBM, Sendai, Japan, 2003*.
- [6] N. Ramnani, L. Lee, A. Mechelli, C. Phillips, A. Roebroeck, and E. Formisano, "Exploring brain connectivity: a new frontier in systems neuroscience. Functional Brain Connectivity, 4-6 April 2002, Dusseldorf, Germany," *Trends Neurosci.*, vol. 25, pp. 496-497, 2002.
- [7] F. Savopoul and C. Armenakis, "Merge of Heterogeneous Data for Emergency Mapping: Data integration or Data Fusion?," *Proc.ISPRS*, 2002.
- [8] C. Arndt, "Information gained by data fusion," in *SPIE Proc.*, 1996.
- [9] A. R. McIntosh and F. Gonzalez-Lima, "Structural Equation Modeling and Its Application to Network Analysis in Functional Brain Imaging," *Hum.Brain Map.*, vol. 2, pp. 2-22, 1994.
- [10] K. J. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *NeuroImage*, vol. 19, pp. 1273-1302, 2003.
- [11] A. Mechelli, C. J. Price, U. Noppeney, and K. J. Friston, "A dynamic causal modeling study on category effects: bottom-up or top-down mediation?," *J Cogn Neurosci.*, vol. 15, pp. 925-934, 2003.
- [12] M. N. Rajah and A. R. McIntosh, "Overlap in the functional neural systems involved in semantic and episodic memory retrieval," *J Cogn Neurosci.*, vol. 17, pp. 470-482, 2005.
- [13] A. Meyer-Lindenberg, P. Kohn, C. B. Mervis, J. S. Kippenhan, R. K. Olsen, C. A. Morris, and K. F. Berman, "Neural basis of genetically determined visuospatial construction deficit in williams syndrome," *Neuron*, vol. 43, pp. 623-631, 2004.
- [14] D. Vitacco, D. Brandeis, R. Pascual-Marqui, and E. Martin, "Correspondence of event-related potential tomography and functional magnetic resonance imaging during language processing," *Hum.Brain Mapp.*, vol. 17, pp. 4-12, 2002.
- [15] K. J. Worsley, J. Cao, T. Paus, M. Petrides, and A. C. Evans, "Applications of random field theory to functional connectivity," *Hum.Brain Map.*, vol. 6, pp. 364-367, 1998.
- [16] A. R. McIntosh, F. L. Bookstein, J. V. Haxby, and C. L. Grady, "Spatial pattern analysis of functional brain images using partial least squares," *NeuroImage*, vol. 3, pp. 143-157, 1996.
- [17] K. Friston, J. P. Poline, S. Strother, A. Holmes, C. D. Frith, and R. S. Frackowiak, "A Multivariate analysis of PET activation studies," *Hum.Brain Map.*, vol. 4, pp. 140-151, 1996.
- [18] M. J. McKeown, S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, "Analysis of fMRI Data by Blind Separation Into Independent Spatial Components," *Hum.Brain Map.*, vol. 6, pp. 160-188, 1998.
- [19] J. F. Cardoso, "Source separation using higher order moments," in *Proc.IEEE Int.Conf.Acoustics, Speech, Signal Processing (ICASSP)*, 1998, pp. 2109-2112.
- [20] A. Hyvarinen, "Survey on Independent Component Analysis," *Neural Computing Surveys*, vol. 1, pp. 94-128, 1999.
- [21] V. D. Calhoun, K. A. Kiehl, and G. D. Pearlson, "A Method for Multi-task fMRI Data Fusion Applied to Schizophrenia," in *NeuroImage*, 2005.
- [22] V. D. Calhoun, T. Adali, N. Giuliani, J. J. Pekar, G. D. Pearlson, and K. A. Kiehl, "A Method for Multimodal Analysis of Independent Source Differences in Schizophrenia: Combining Gray Matter Structural and

- Auditory Oddball Functional Data," *Hum.Brain Map.*, vol. 27, pp. 47-62, 2006.
- [23] V. D. Calhoun, G. D. Pearlson, and K. A. Kiehl, "Neuronal Chronometry of Target Detection: Fusion of Hemodynamic and Event-related Potential Data," *NeuroImage*, vol. 30, pp. 544-553, 2006.
- [24] M. J. McKeown and T. J. Sejnowski, "Independent component analysis of fMRI data: examining the assumptions," *Hum.Brain Map.*, vol. 6, pp. 368-372, 1998.
- [25] V. D. Calhoun and T. Adali, "'Unmixing' Functional Magnetic Resonance Imaging with Independent Component Analysis," *IEEE Eng.in Medicine and Biology*, vol. 25, pp. 79-90, 2006.
- [26] M. J. McKeown, L. K. Hansen, and T. J. Sejnowski, "Independent component analysis of functional MRI: what is signal and what is noise?," *Curr.Opin.Neurobiol.*, vol. 13, pp. 620-629, 2003.
- [27] V. D. Calhoun, T. Adali, K. A. Kiehl, R. S. Astur, J. J. Pekar, and G. D. Pearlson, "A Method for Multi-task fMRI Data Fusion Applied to Schizophrenia," *Hum.Brain Map.*, vol. 27, pp. 598-610, 2006.
- [28] N. Lange, P. Bandettini, and C. Moonen, "Statistical Methods in fMRI," in *Functional MRI* New York, 1999.
- [29] N. Lange, S. C. Strother, J. R. Anderson, F. A. Nielsen, A. P. Holmes, T. Kolenda, R. Savoy, and L. K. Hansen, "Plurality and resemblance in fMRI data analysis," *NeuroImage*, vol. 10, pp. 282-303, 1999.
- [30] K. Friston, P. Jezzard, and R. Turner, "Analysis of functional MRI time-series," *Hum.Brain Map.*, vol. 1, pp. 153-171, 1994.
- [31] V. D. Calhoun, T. Adali, G. D. Pearlson, and J. J. Pekar, "Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms," *Hum.Brain Map.*, vol. 13, pp. 43-53, 2001.
- [32] V. D. Calhoun, X. Golay, and G. D. Pearlson, "Improved fMRI Slice Timing Correction: Interpolation Errors and Wrap Around Effects," in *Proc.ISMRM*, 2000, p. 810.
- [33] P. F. Van de Moortele, B. Cerf, E. Lobel, A. L. Paradis, A. Faurion, and D. Le Bihan, "Latencies in fMRI time-series: effect of slice acquisition order and perception," *NMR Biomed.*, vol. 10, pp. 230-236, 1997.
- [34] V. D. Calhoun, T. Adali, and G. D. Pearlson, "A frequency-space approach for motion correction in fMRI," in *Proc.IMDSP*, 1998, p. 229.
- [35] K. Friston, J. Ashburner, C. D. Frith, J. P. Poline, J. D. Heather, and R. S. Frackowiak, "Spatial Registration and normalization of images," *Hum.Brain Map.*, vol. 2, pp. 165-189, 1995.
- [36] K. Friston, "Spatial Normalisation: A new approach," 1995.
- [37] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Trans.Med.Imaging*, vol. 20, pp. 45-57, 2001.
- [38] W. M. Wells, III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz, "Adaptive Segmentation of MRI Data," in *IEEE Trans.Med.Imaging*, vol. 15(4), 1996, pp. 429-442.
- [39] J. C. Bezdek, L. O. Hall, and L. P. Clarke, "Review of MR image segmentation techniques using pattern recognition. [Review]," *Med.Phys.*, vol. 20, pp. 1033-1048, 1993.
- [40] P. T. Fox, "The growth of human brain mapping," *Hum.Brain Map.*, vol. 5 pp. 1-2, 1997.
- [41] M. I. Miller, A. Trouve, and L. Younes, "On the metrics and euler-lagrange equations of computational anatomy," *Annu.Rev.Biomed.Eng.*, vol. 4, pp. 375-405, 2002.
- [42] B. Fischl and A. M. Dale, "Measuring the thickness of the human cerebral cortex from magnetic resonance images," *Proc.Natl.Acad.Sci.U.S.A.*, vol. 97, pp. 11050-11055, 2000.
- [43] J. Ashburner, J. G. Csernansky, C. Davatzikos, N. C. Fox, G. B. Frisoni, and P. M. Thompson, "Computer-assisted imaging to assess brain structure in healthy and diseased brains," *Lancet Neurol.*, vol. 2, pp. 79-88, 2003.
- [44] G. Gerig, M. Styner, M. Shenton, and J. A. Lieberman, "Improved understanding of the morphology of brain structures," in *Proc.MICCAI*, Utrecht, The Netherlands 2001, pp. 24-32.
- [45] K. J. Friston, "Statistical parametric Mapping and Other Analyses of Functional Imaging Data," in *Brain Mapping: The Methods*, 1996, pp. 363-386.
- [46] K. J. Worsley, M. Andermann, T. Koulis, D. MacDonald, and A. C. Evans, "Detecting changes in nonisotropic images," *Hum.Brain Mapp.*, vol. 8, pp. 98-101, 1999.
- [47] J. Ashburner and K. J. Friston, "Voxel-based morphometry--the methods," *NeuroImage*, vol. 11, pp. 805-821, 2000.
- [48] M. S. Cohen, R. M. DuBois, and M. M. Zeineh, "Rapid and effective correction of RF inhomogeneity for high field magnetic resonance imaging," *Hum.Brain Mapp.*, vol. 10, pp. 204-211, 2000.
- [49] M. Tincher, C. R. Meyer, R. Gupta, and D. M. Williams, "Polynomial Modeling and reduction of RF body coil spatial inhomogeneity in MRI," in *IEEE Trans.Med.Imaging*, vol. 12(2), 1993, pp. 361-365.
- [50] G. Gerig, J. Martin, R. Kikinis, O. Kubler, M. E. Shenton, and F. Jolesz, "Unsupervised tissue type segmentation of 3D dual-echo MR head data," 1992.
- [51] D. W. Shattuck, S. R. Sandor-Leahy, K. A. Schaper, D. A. Rottenberg, and R. M. Leahy, "Magnetic resonance image tissue classification using a partial volume model," *NeuroImage*, vol. 13, pp. 856-876, 2001.
- [52] S. K. Warfield, M. Kaus, F. A. Jolesz, and R. Kikinis, "Adaptive, template moderated, spatially varying statistical classification," *Med.Image Anal.*, vol. 4, pp. 43-55, 2000.
- [53] D. Le Bihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanis, and M. Laval-Jeantet, "MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders," *Radiology*, vol. 161, pp. 401-407, 1986.
- [54] M. E. Moseley, Y. Cohen, J. Kucharczyk, J. Mintorovitch, H. S. Asgari, M. F. Wendland, J. Tsuruda, and D. Norman, "Diffusion-weighted MR imaging of anisotropic water diffusion in cat central nervous system," *Radiology*, vol. 176, pp. 439-445, 1990.
- [55] P. J. Basser, J. Mattiello, and D. LeBihan, "Estimation of the effective self-diffusion tensor from the NMR spin echo," *J.Magn Reson.B*, vol. 103, pp. 247-254, 1994.
- [56] P. J. Basser and C. Pierpaoli, "Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI," *J.Magn Reson.B*, vol. 111, pp. 209-219, 1996.
- [57] P. J. Basser, "Inferring microstructural features and the physiological state of tissues from diffusion-weighted images," *NMR Biomed.*, vol. 8, pp. 333-344, 1995.
- [58] S. Pajevic and C. Pierpaoli, "Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data: application to white matter fiber tract mapping in the human brain," *Magn Reson.Med.*, vol. 42, pp. 526-540, 1999.
- [59] T. P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Analysis and visualization of single-trial event-related potentials," *Hum.Brain Mapp.*, vol. 14, pp. 166-185, 2001.
- [60] V. D. Calhoun, J. J. Pekar, V. B. McGinty, T. Adali, T. D. Watson, and G. D. Pearlson, "Different activation dynamics in multiple neural systems during simulated driving," *Hum.Brain Map.*, vol. 16, pp. 158-167, 2002.
- [61] S. Makeig, T. P. Jung, A. J. Bell, D. Ghahremani, and T. J. Sejnowski, "Blind separation of auditory event-related brain responses into independent components," *Proc.Natl.Acad.Sci.*, vol. 94, pp. 10979-10984, 1997.
- [62] T. Nakai, S. Muraki, E. Bagarinao, D. J. Mikulis, Y. Takehara, K. Matsuo, C. Kato, H. Sakahara, and H. Isoda, "Application of independent component analysis to magnetic resonance imaging for enhancing the contrast of gray and white matter," *NeuroImage*, vol. 12, pp. 251-260, 2004.
- [63] M. Moosmann, T. Eichele, H. Nordby, K. Hugdahl, and V. D. Calhoun, "Joint Independent Component Analysis for Simultaneous EEG-fMRI: Principle and Simulation," *Int. J. Psych.*, In Press.
- [64] T. Eichele, V. D. Calhoun, M. Moosmann, K. Specht, M. Jongsma, R. Quiroga, H. Nordby, and K. Hugdahl, "Unmixing concurrent EEG-fMRI with parallel independent component analysis," *Int. J. Psych.*, In Press.
- [65] J. Liu, G. D. Pearlson, A. Windemuth, G. Ruano, N. I. Perrone-Bizzozero, and V. D. Calhoun, "Combining fMRI and SNP data to investigate connections between brain function and genetics using parallel ICA," *Hum.Brain Map.*, In Press.
- [66] J. Liu and V. D. Calhoun, "Parallel independent component analysis for multimodal analysis: Application to fMRI and EEG Data," in *Proc.ISBI*, 2007.
- [67] J. Liu and V. D. Calhoun, "A novel approach to analyzing fMRI and SNP data via parallel independent component analysis," in *Proc.SPIE*, 2006.
- [68] K. J. Worsley and K. J. Friston, "Analysis of fMRI time-series revisited--again," *NeuroImage*, vol. 2, pp. 173-181, 1995.
- [69] T. P. Jung, S. Makeig, C. Humphries, T. W. Lee, M. J. McKeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, pp. 163-178, 2000.
- [70] T. Thomsen, K. Specht, A. Hammar, J. Nytingnes, L. Erslund, and K. Hugdahl, "Brain localization of attentional control in different age groups by combining functional and structural MRI," *Neuroimage.*, vol. 22, pp. 912-919, 2004.
- [71] L. Clark, S. D. Iversen, and G. M. Goodwin, "A neuropsychological investigation of prefrontal cortex involvement in acute mania," *Am.J.Psychiatry*, vol. 158, pp. 1605-1611, 2001.

- [72] V. D. Calhoun, T. Adali, G. D. Pearlson, and K. A. Kiehl, "Neuronal chronometry of target detection: fusion of hemodynamic and event-related potential data," *Neuroimage*, vol. 30, pp. 544-553, Apr 1 2006.
- [73] T. Cover and K. M. Thomas, *Elements of Information Theory*. New York: Johns Wiley and Sons, 1991.
- [74] A. Hero, B. Ma, O. Michel, and J. Gorman, "Alpha-Divergence for classification, indexing and retrieval," Dept. EECS, University of Michigan, Ann Arbor 328, 2001.
- [75] O. Demirci and V. D. Calhoun, "Detection of Schizophrenia using fMRI Data via Projection Pursuit," in *Proc.MLSP*, 2007.
- [76] O. Demirci, V. P. Clark, and V. D. Calhoun, "A Projection Pursuit Algorithm to Classify Individuals Using fMRI Data: Application to Schizophrenia," *NeuroImage*, In Press.
- [77] M. R. Johnson, N. Morris, R. S. Astur, V. D. Calhoun, K. A. Kiehl, and G. D. Pearlson, "Schizophrenia and Working Memory: A Closer Look at fMRI of the Dorsolateral Prefrontal Cortex During a Working Memory Task," in *Proc.CNS*, 2005.
- [78] 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.
- [79] N. Giuliani, V. D. Calhoun, G. D. Pearlson, A. Francis, and R. W. Buchanan, "Voxel-Based Morphometry versus Regions of Interest: A Comparison of Two Methods for Analyzing Gray Matter Disturbances in Schizophrenia," *Schizophr.Res.*, vol. 74, pp. 135-147, 2005.
- [80] K. A. Kiehl, A. M. Smith, R. D. Hare, and P. F. Liddle, "An event-related potential investigation of response inhibition in schizophrenia and psychopathy," *Biological Psychiatry*, vol. 48, pp. 210-221, 2000.



Vince D. Calhoun VINCE CALHOUN received a bachelor's degree in Electrical Engineering from the University of Kansas, Lawrence, Kansas, in 1991, master's degrees in Biomedical Engineering and Information Systems from Johns Hopkins University, Baltimore, in 1993 and 1996, respectively, and the Ph.D. degree in electrical engineering from the University of Maryland Baltimore County, Baltimore, in 2002. He worked as a Senior Research Engineer at the Psychiatric Neuro-Imaging Laboratory at Johns Hopkins from 1993 until 2002. Then as the director of medical image analysis at the Olin Neuropsychiatry Research Center and an associate professor at Yale University.

Dr. Calhoun is currently the Director of the Image Analysis and MR Research at the Mind Research Network and is an associate professor in the Dept of ECE at the University of New Mexico. He is the author of more than 75 full journal articles, over 200 technical reports, abstracts and conference proceedings. Much of his career has been spent on the development of data driven approaches for the analysis of functional magnetic resonance imaging (fMRI) data. He has multiple R01 grants from the National Institutes of Health on the incorporation of prior information into independent component analysis (ICA) for fMRI as well as two foundation grants, including one from the NARSAD foundation to apply the methods he is developing to differentiate healthy controls from patients with schizophrenia and bipolars. He is a co-investigator on several other grants. Dr. Calhoun was the recipient of the 2004 Early Career Investigator Award from the International Society for Neuroimaging in Psychiatry.

Dr. Calhoun is a senior member of the IEEE, the Organization for Human Brain Mapping, and the International Society for Magnetic Resonance in Medicine. He has participated in multiple NIH study sections. He has worked in the organization of workshops at conferences including the society of biological psychiatry (SOBP) and the international conference of independent component analysis and blind source separation (ICA). He is currently serving on the IEEE Machine Learning for Signal Processing (MLSP) Technical Committee and has previously served as the general chair of the 2005 meeting. He is a reviewer for a number of international journals and is on the editorial board of the *Human Brain Mapping* journal and an associate editor for the *IEEE Signal Processing Letters* and the *International Journal of Computational Intelligence and Neuroscience*.

His interests include the development of data driven analysis methods including data fusion approaches for magnetic resonance imaging (functional imaging, diffusion tensor imaging, and structural imaging), electroencephalography, and genetics in the areas of image processing, adaptive signal processing, estimation theory, neural networks, statistical signal processing, and pattern recognition.



Tulay Adali TULAY ADALI received the Ph.D. degree in electrical engineering from North Carolina State University, Raleigh, in 1992 and joined the faculty at the University of Maryland Baltimore County (UMBC), Baltimore, the same year. She is currently a Professor in the Department of Computer Science and Electrical Engineering at UMBC. She has held visiting positions at Technical University of Denmark, Lyngby, Denmark, Katholieke Universiteit, Leuven, Belgium, University of Campinas, Brazil, and Ecole Supérieure de Physique et de Chimie Industrielles, Paris, France.

Prof. Adali assisted in the organization of a number of international conferences and workshops including the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), the IEEE International Workshop on Neural Networks for Signal Processing (NNSP), and the IEEE International Workshops on Machine Learning for Signal Processing (MLSP). She was the General Co-Chair, NNSP (2001-2003); Technical Chair, MLSP (2004-2006); and Publicity Chair, ICASSP (2000 and 2005). She is currently the Technical Chair, 2008 MLSP, Publications Co-Chair, ICASSP 2008, and Program Co-Chair, 2008 Workshop on Cognitive Information Processing.

Prof. Adali chaired the SPS Machine Learning for Signal Processing Technical Committee (2003-2005); Member, SPS Conference Board (1998-2006); and Associate Editor, *IEEE Transactions on Signal Processing* (2003-2006). She is currently a Member, Machine Learning for Signal Processing Technical Committee and the Bio Imaging and Signal Processing Technical Committees and an Associate Editor, *IEEE Transactions on Biomedical Engineering*, *Signal Processing Journal*, *Research Letters in Signal Processing*, and *Journal of Signal Processing Systems for Signal, Image, and Video Technology*.

Prof. Adali is a Fellow of the AIMBE and a Senior Member of the IEEE. Her research interests are in the areas of statistical signal processing, machine learning for signal processing, biomedical data analysis (functional MRI, MRI, PET, CR, ECG, and EEG), bioinformatics, and signal processing for optical communications. She is the recipient of a 1997 National Science Foundation (NSF) CAREER Award with more recent support from the National Institutes of Health, NSF, NASA, the US Army, and industry.