Published in final edited form as:
IEEE Trans Inf Technol Biomed. 2009 September ; 13(5): 711–720. doi:10.1109/TITB.2008.923773.

Feature-Based Fusion of Medical Imaging Data

Vince D. Calhoun [Senior Member] and **Tulay 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 multitask or multimodal 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 EEG information and spatial functional magnetic resonance imaging 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 provides the greatest value in discriminating groups. Finally, we summarize and describe future research topics.

Keywords

Data fusion; EEG; functional magnetic resonance imaging (fMRI); independent component analysis (ICA); multivariate data analysis

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 "crossinformation." In the former approach, any crossinformation 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.

V. D. Calhoun is with the Mind Research Network and Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM 87131 USA (e-mail: vcalhoun@unm.edu)
T. Adalı is with the Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD 21250 USA (e-mail: adalı@umbc.edu)

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

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 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 (GM) 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 “crossinformation” 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 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 6-D 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 aforementioned techniques includes 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 multitask 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 multitask 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, 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. 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—GM, white matter (WM), and cerebrospinal fluid (CSF). A different type of structural information is captured by DTI that 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

fMRI data provide a measure of brain function on a millimeter spatial scale and a subsecond 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 that 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

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

for R regressors, where $\mathbf{y}_m$, $\mathbf{x}_i$, and ε_m are $K \times 1$ for time points $k = 1,2, \cdots, K$ at brain locations $m = 1,2, \cdots, M$. 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., GM or WM). There are many methods for preprocessing sMRI data that may include bias field correction [intensity changes caused by RF or main magnetic field (Bo) inhomogeneities] [48], [49], spatial linear, or nonlinear [50] filtering normalization. MRIs are typically segmented using a tissue classifier producing images showing the spatial distribution of GM, WM, 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 × 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 ($\lambda 1$, $\lambda 2$, and $\lambda 3$) and the directions of these 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 nor aligned fibers. When diffusion is anisotropic, the direction of $\lambda 1$ 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]:

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

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

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

D. EEG

EEG 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 population although it is not clear if small populations also make a contribution. The recorded electrical signals are then amplified, digitized, and stored.

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 before, 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 subdataset extracted from one type of data, related to a selected brain activity or structure. A summary of some features is provided in Table I.

In the table, the first two modalities are fMRI, each of which has two stimuli that 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 DTI, for which a measure of fractional anisotropy is computed.

IV. Feature-Based Fusion Framework Using ICA

ICA 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 $\mathbf{x} = [x_1, \cdots, x_M]^T$ that is generated by the ICA model:

$$\mathbf{x}=\mathbf{A}\mathbf{s} \tag{3}$$

where $\mathbf{s} = [s_1, s_2, \cdots, s_N]^T$ is an N-dimensional vector whose elements are assumed to be 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 (4)] is a good approximation to the "true" sources $\mathbf{s}$:

$$\mathbf{y}=Wx. \tag{4}$$

ICA has been shown to be useful for fMRI analysis for several reasons. Spatial ICA finds systematically nonoverlapping, 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 intersubject and interevent 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.

ICA 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 4-D 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 (Fig. 1).

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

$$L(\mathbf{W}) = \prod_{n=1}^{N} \prod_{v=1}^{V} pJ, n(uJ, v) \tag{5}$$

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 $n = 1,\cdots, N$ 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

$$L(\mathbf{W}) = \prod_{n=1}^{N} \left(\prod_{v=1}^{V_1} pF, n(uF, v) \prod_{v=1}^{V_2} pG, n(uG, v) \right). \tag{6}$$

Depending on the data types in question, the previous 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], and [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 that do not require the same linear covariation for the different modalities [65], [66].

The underlying assumptions for the form given in (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 with 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 earlier formulation, 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 V to real data from multiple tasks/modalities using: 1) multitask fMRI data; 2) fMRI/sMRI; and 3) ERP/fMRI and 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 that 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 were then filtered with a 20 Hz 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. Multitask fMRI

We performed a joint analysis of fMRI data collected from a Sternberg (SB) task and an 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 Fig. 2(a) and (b) and the marginal histograms for the AOD and SB tasks are presented in Fig. 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.

B. fMRI-sMRI

It is also feasible to use jICA to combine structural and functional features. Our approach requires acceptance of the likelihood of GM 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 GM regions and functional regions that are spatially remote. Given the functional interconnectedness of widespread neural networks, we suggest that this is also a reasonable conception for the relationship between structural and functional changes.

The next example is from a jICA analysis of the fMRI data of AOD task and GM segmentation data [22]. AOD target activation maps and segmented GM maps were normalized to a study-specific template in order to control for intensity differences in MRIs based on scanner, template, and population variations [71].

Results are presented in Fig. 3. The AOD part of the joint source is shown in Fig. 3(a), the GM part of the joint source is shown in Fig. 3(b), and the ICA loading parameters separated by group and shown in Fig. 3(c). 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, AOD and GM 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 GM associated with bilateral temporal regions activated by the AOD target stimulus. This finding suggests that GM regions 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, GM concentrations were increased in patients versus controls, suggesting that these increases are somehow related to decreases in functional connectivity in the AOD fMRI task.

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 (Fig. 4, right), and in ERP at times during the N2/P3 complex (Fig. 4, left) that have been previously implicated in patients.

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

D. sMRI-DTI

We now present an example showing GM and FA maps in a jICA analysis in 11 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 Fig. 6 and demonstrates GM regions in occipital lobe, and bilateral lateral geniculate regions are associated with WM 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 Fig. 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 the loss of GM volume (and WM FA)with age.

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 earlier 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.

E. Parallel ICA

As discussed in Section V, 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 jICA algorithm [72], where a shared mixing matrix is used for both modalities. The fundamental difference is that paraICA assumes that the fMRI and ERP data are mixed in a similar pattern but not identically. The paraICA pays more attention to individually linked components and their connections, while the jICA studies intereffects 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 **f** is a multidimensional 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

$$D_{\mathrm{J}} = \frac{D_{\mathrm{KL}}(p,q) + D_{\mathrm{KL}}(q,p)}{2}.$$

.

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 (α) divergence with α = 0.5 has been shown to be optimal for separating similar feature densities [74]; the limiting case of α = 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 II. 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 α = 0.5 as well as their symmetric counterparts.

The results shown in Fig. 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 earlier 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.

In summary, we have presented a data fusion approach based upon ICA and have shown multiple examples that 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 FIT.

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.

Acknowledgments

This work was supported in part by the National Institutes of Health under Grant 1 R01 EB 000840, Grant R01 EB 005846, and Grant R01 EB 006841, and in part by the National Science Foundation under Grant National Science Foundation (NSF)-Internet Information Services (IIS) 0612076.

Biographies

Vince D. Calhoun (S'xx-M'xx-SM'xx) received the Bachelor's degree in electrical engineering from the University of Kansas, Lawrence, in 1991, the Master's degree in biomedical engineering in 1993 and the Master's degree in information systems in 1996, both from Johns Hopkins University, Baltimore, MD, and the Ph.D. degree in electrical engineering from the University of Maryland Baltimore County, Baltimore, in 2002.

From 1993 to 2002, he was a Senior Research Engineer at the Psychiatric Neuro-Imaging Laboratory, Johns Hopkins University. He was the Director of Medical Image Analysis at the Olin Neuropsychiatry Research Center and an Associate Professor at Yale University. He is currently the Director of the Image Analysis and Magnetic Resonance (MR) Research at the Mind Research Network and also an Associate Professor in the Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque. He is the author or coauthor of more than 75 full journal papers, over 200 technical reports, abstracts, and conference proceedings. He has multiple R01 grants from the National Institutes of Health (NIH) on the incorporation of prior information into independent component analysis (ICA) for fMRI as well as two foundation grants. He is a Co-Investigator on several other grants. 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. His current research interests include the development of data driven analysis methods including data fusion approaches for MRI (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.

Dr. Calhoun is a member of the Organization for Human Brain Mapping and the International Society for Magnetic Resonance in Medicine. He was the recipient of the 2004 Early Career Investigator Award from the International Society for Neuroimaging in Psychiatry. He is currently serving on the IEEE Machine Learning for Signal Processing (MLSP) Technical Committee and has previous served as the General Chair of the 2005 meeting. He is the 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 International Journal of Computational Intelligence and Neuroscience.

Tulay Adalı (S'89-M'89-SM'98-F'XX) received the Ph.D. degree in electrical engineering from North Carolina State University, Raleigh, in 1992.

Since 1992, she has been with the University of Maryland Baltimore County (UMBC), Baltimore, where she is currently a Professor in the Department of Computer Science and Electrical Engineering. She has held visiting positions at the Technical University of Denmark, Lyngby, Denmark, Katholieke Universiteit, Leuven, Belgium, University of Camp-inas, Brazil, and Ecole Supérieure de Physique et de Chimie Industrielles, Paris, France. Her current research interests include statistical signal processing, machine learning for signal processing, biomedical data analysis [functional magnetic resonance imaging (MRI), MRI, positron emission tomography, CR, ECG, and EEG], bioinformatics, and signal processing for optical communications.

Prof. Adalı is a Fellow of the American Institute of Medical and Biological Engineering (AIMBE). She was a member of the Society of Physics Students (SPS) Conference Board (1998-2006) and an Associate Editor of the IEEE Transactions on Signal Processing (2003-2006). She is currently a member of the Machine Learning for Signal Processing (MLSP) Technical Committee and the Bio Imaging and Signal Processing Technical Committees, and an Associate Editor of the 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. She 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 MLSP. She was the General Co-Chair of NNSP (2001-2003), the Technical Chair of MLSP (2004-2006), and the Publicity Chair of ICASSP (2000 and 2005). She is currently the Technical Chair of the 2008 MLSP, a Publications Co-Chair of ICASSP 2008, and a Program Co-Chair of the 2008 Workshop on Cognitive Information Processing. She was also the Chair of the SPS MLSP Technical Committee (2003-2005). She was the recipient of the 1997 National Science Foundation (NSF) CAREER Award with more recent support from the National Institutes of Health, NSF, National Aeronautics and Space Administration (NASA), the US Army, and industry.

References

[1]. Husain FT, Nandipati G, Braun AR, Cohen LG, Tagamets MA, Horwitz B. Simulating transcranial magnetic stimulation during PET with a large-scale neural network model of the prefrontal cortex and the visual system. NeuroImage 2002;15:58–73. [PubMed: 11771974]

[2]. Horwitz B, Poeppel D. How can EEG/MEG and fMRI/PET data be combined? Hum. Brain Mapp 2002;17:1–3. [PubMed: 12203682]

[3]. Dale AM, Halgren E, Lewine JD, Buckner RL, Paulson K, Marinkovic K, Rosen BR. Spatio-temporal localization of cortical word repetition effects in a size judgement task using combined fMRI/MEG. NeuroImage 1997;5:S592–S592.

[4]. George JS, Aine CJ, Mosher JC, Schmidt DM, Ranken DM, Schlitt HA, Wood CC, Lewine JD, Sanders JA, Belliveau JW. Mapping function in the human brain with magnetoencephalography, anatomical magnetic resonance imaging, and functional magnetic resonance imaging. J.Clin. Neurophysiol 1995;12:406–431. [PubMed: 8576388]

[5]. Kim, DS.; Ronen, I.; Formisano, E.; Kim, KH.; Kim, M.; van Zijl, P.; Ugurbil, K.; Mori, S.; Goebel, R. Simultaneous mapping of functional maps and axonal connectivity in cat visual cortex. HBM; Sendai, Japan. 2003. presented at the

[6]. Ramnani N, Lee L, Mechelli A, Phillips C, Roebroeck A, Formisano E. Exploring brain connectivity: A new frontier in systems neuroscience. Functional Brain Connectivity, 4-6 April 2002, Dusseldorf, Germany. Trends Neurosci 2002;25:496–497. [PubMed: 12220872]

[7]. Savopol, F.; Armenakis, C. Mergine of heterogeneous data for emergency mapping: data integration or data fusion?. ISPRS; Ottawa, Canada. 2002. presented at the

[8]. Ardnt C. Information gained by data fusion. Proc. SPIE 1996;2784:32–40.

[9]. McIntosh AR, Gonzalez-Lima F. Structural equation modeling and its application to network analysis in functional brain imaging. Hum. Brain Mapp 1994;2:2–22.

[10]. Friston KJ, Harrison L, Penny W. Dynamic causal modelling. NeuroImage 2003;19:1273–1302. [PubMed: 12948688]

[11]. Mechelli A, Price CJ, Noppeney U, Friston KJ. A dynamic causal modeling study on category effects: Bottom-up or top-down mediation? J. Cogn. Neurosci 2003;15:925–934. [PubMed: 14628754]

[12]. Rajah MN, McIntosh AR. Overlap in the functional neural systems involved in semantic and episodic memory retrieval. J. Cogn. Neurosci 2005;17:470–482. [PubMed: 15814006]

[13]. Meyer-Lindenberg A, Kohn P, Mervis CB, Kippenhan JS, Olsen RK, Morris CA, Berman KF. Neural basis of genetically determined visuospatial construction deficit in Williams syndrome. Neuron 2004;43:623–631. [PubMed: 15339645]

[14]. Vitacco D, Brandeis D, Pascual-Marqui R, Martin E. Correspondence of event-related potential tomography and functional magnetic resonance imaging during language processing. Hum. Brain Mapp 2002;17:4–12. [PubMed: 12203683]

[15]. Worsley KJ, Cao J, Paus T, Petrides M, Evans AC. Applications of random field theory to functional connectivity. Hum. Brain Map 1998;6:364–367.

[16]. McIntosh AR, Bookstein FL, Haxby JV, Grady CL. Spatial pattern analysis of functional brain images using partial least squares. NeuroImage 1996;3:143–157. [PubMed: 9345485]

[17]. Friston K, Poline JP, Strother S, Holmes A, Frith CD, Frackowiak RS. A multivariate analysis of PET activation studies. Hum. Brain Mapp 1996;4:140–151.

[18]. McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ. Analysis of fMRI Data by blind separation into independent spatial components. Hum. Brain Mapp 1998;6:160–188. [PubMed: 9673671]

[19]. Cardoso, JF. Source separation using higher order moments; Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP); 1998. p. 2109-2112.

[20]. Hyvarinen A. Survey on independent component analysis. Neural Comput. Surv 1999;1:94–128.

[21]. Calhoun VD, Keihl KA, Pearlson GD. A method for multi-task fMRI data fusion applied to Schizophrenia. NeuroImage 2005;27(7):598–610.

[22]. Calhoun VD, Adalı T, Giuliani N, Pekar JJ, Pearlson GD, Kiehl KA. A method for multimodal analysis of independent source differences in Schizophrenia: Combining gray matter structural and auditory oddball functional data. Hum. Brain Mapp 2006;27:47–62. [PubMed: 16108017]

[23]. Calhoun VD, Pearlson GD, Kiehl KA. Neuronal chronometry of target detection: Fusion of hemodynamic and event-related potential data. NeuroImage 2006;30:544–553. [PubMed: 16246587]

[24]. McKeown MJ, Sejnowski TJ. Independent component analysis of fMRI data: Examining the assumptions. Hum. Brain Mapp 1998;6:368–372. [PubMed: 9788074]

[25]. Calhoun VD, Adalı T. Unmixing' functional magnetic resonance imaging with independent component analysis. IEEE Eng. Med. Biol Mar.-Apr.;2006 25(2):79–90.

[26]. McKeown MJ, Hansen LK, Sejnowsk TJ. Independent component analysis of functional MRI: What is signal and what is noise? Curr. Opin. Neurobiol 2003;13:620–629. [PubMed: 14630228]

[27]. Calhoun VD, Adalı T, Kiehl KA, Astur RS, Pekar JJ, Pearlson GD. A method for multi-task fMRI data fusion applied to Schizophrenia. Hum. Brain Mapp 2006;27:598–610. [PubMed: 16342150]

[28]. Lange, N.; Bandettini, P.; Moonen, C. Statistical methods in fMRI. Functional MRI; New York. 1999. presented at the

[29]. Lange N, Strother SC, Anderson JR, Nielsen FA, Holmes AP, Kolenda T, Savoy R, Hansen LK. Plurality and resemblance in fMRI data analysis. NeuroImage 1999;10:282–303. [PubMed: 10458943]

[30]. Friston K, Jezzard P, Turner R. Analysis of functional MRI time-series. Hum. Brain Mapp 1994;1:153–171.

[31]. Calhoun VD, Adalı T, Pearlson GD, Pekar JJ. Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. Hum. Brain Mapp 2001;13:43–53. [PubMed: 11284046]

[32]. Calhoun VD, Golay X, Pearlson GD. Improved fMRI slice timing correction: Interpolation errors and wrap around effects. Proc. ISMRM 2000:810.

[33]. Van de Moortele PF, Cerf B, Lobel E, Paradis AL, Faurion A, Le Bihan D. Latencies in fMRI time-series: Effect of slice acquisition order and perception. NMR Biomed 1997;10:230–236. [PubMed: 9430353]

[34]. Calhoun VD, Adalı T, Pearlson GD. A frequency-space approach for motion correction in fMRI. Proc. IMDSP 1998:229.

[35]. Friston K, Ashburner J, Frith CD, Poline JP, Heather JD, Frackowiak RS. Spatial registration and normalization of images. Hum. Brain Mapp 1995;2:165–189.

[36]. Friston, K. Spatial normalisation: A new approach. BrainMap'95 UTHSC; San-Antonio, TX. presented at the

[37]. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imag Jan;2001 20(1): 45–57.

[38]. Wells WM III, Grimson WEL, Kikinis R, Jolesz FA. Adaptive segmentation of MRI Data. IEEE Trans. Med. Imag Aug;1996 15(4):429–442.

[39]. Bezdek JC, Hall LO, Clarke LP. Review of MR image segmentation techniques using pattern recognition [Review]. Med. Phys 1993;20:1033–1048. [PubMed: 8413011]

[40]. Fox PT. The growth of human brain mapping. Hum. Brain Mapp 1997;5:1–2.

[41]. Miller MI, Trouve A, Younes L. On the metrics and Euler-Lagrange equations of computational anatomy. Annu. Rev. Biomed. Eng 2002;4:375–405. [PubMed: 12117763]

[42]. Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. USA 2000;97:11050–11055. [PubMed: 10984517]

[43]. Ashburner J, Csernansky JG, Davatzikos C, Fox NC, Frisoni GB, Thompson PM. Computer-assisted imaging to assess brain structure in healthy and diseased brains. Lancet Neurol 2003;2:79–88. [PubMed: 12849264]

[44]. Gerig, G.; Styner, M.; Shenton, M.; Lieberman, JA. Improved understanding of the morphology of brain structures; Proc. MICCAI; The Netherlands: Ultrecht; 2001. p. 24-32.

[45]. Friston KJ. Statistical parametric mapping and other analyses of functional imaging data. Proc. Brain Mapp.: Methods 1996:363–386.

[46]. Worsley KJ, Andermann M, Koulis T, MacDonald D, Evans AC. Detecting changes in nonisotropic images. Hum. Brain Mapp 1999;8:98–101. [PubMed: 10524599]

[47]. Ashburner J, Friston KJ. Voxel-based morphometry-The methods. NeuroImage 2000;11:805–821. [PubMed: 10860804]

[48]. Cohen MS, DuBois RM, Zeineh MM. Rapid and effective correction of RF inhomogeneity for high field magnetic resonance imaging. Hum. Brain Mapp 2000;10:204–211. [PubMed: 10949058]

[49]. Tincher M, Meyer CR, Gupta R, Williams DM. Polynomial modeling and reduction of RF body coil spatial inhomogeneity in MRI. IEEE Trans. Med. Imag Jun;1993 12(2):361–365.

[50]. Gerig G, Martin J, Kikinis R, Kubler O, Shenton ME, Jolesz F. Unsupervised tissue type segmentation of 3-D dual-echo MR head data. Image Vis. Comput Jul./Aug;1992 10(6):349–360.

[51]. Shattuck DW, Sandor-Leahy SR, Schaper KA, Rottenberg DA, Leahy RM. Magnetic resonance image tissue classification using a partial volume model. NeuroImage 2001;13:856–876. [PubMed: 11304082]

[52]. Warfield SK, Kaus M, Jolesz FA, Kikinis R. Adaptive, template moderated, spatially varying statistical classification. Med. Image Anal 2000;4:43–55. [PubMed: 10972320]

[53]. Le Bihan D, Breton E, Lallemand D, Grenier P, Cabanis E, Laval-Jeantet M. MR imaging of intravoxel incoherent motions: Application to diffusion and perfusion in neurologic disorders. Radiology 1986;161:401–407. [PubMed: 3763909]

[54]. Moseley ME, Cohen Y, Kucharczyk J, Mintorovitch J, Asgari HS, Wendland MF, Tsuruda J, Norman D. Diffusion-weighted MR imaging of anisotropic water diffusion in cat central nervous system. Radiology 1990;176:439–445. [PubMed: 2367658]

[55]. Basser PJ, Mattiello J, LeBihan D. Estimation of the effective self-diffusion tensor from the NMR spin echo. J. Magn. Reson. B 1994;103:247–254. [PubMed: 8019776]

[56]. Basser PJ, Pierpaoli C. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. J. Magn. Reson. B 1996;111:209–219. [PubMed: 8661285]

[57]. Basser PJ. Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. NMR Biomed 1995;8:333–344. [PubMed: 8739270]

[58]. Pajevic S, Pierpaoli C. 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 1999;42:526–540. [PubMed: 10467297]

[59]. Jung TP, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ. Analysis and visualization of single-trial event-related potentials. Hum. Brain Mapp 2001;14:166–185. [PubMed: 11559961]

[60]. Calhoun VD, Pekar JJ, McGinty VB, Adalı T, Watson TD, Pearlson GD. Different activation dynamics in multiple neural systems during simulated driving. Hum. Brain Mapp 2002;16:158–167. [PubMed: 12112769]

[61]. Makeig S, Jung TP, Bell AJ, Ghahremani D, Sejnowski TJ. Blind separation of auditory event-related brain responses into independent components. Proc. Natl. Acad. Sci. USA 1997;94:10979–10984. [PubMed: 9380745]

[62]. Nakai T, Muraki S, Bagarinao E, Mikulis DJ, Takehara Y, Matsuo K, Kato C, Sakahara H, Isoda H. Application of independent component analysis to magnetic resonance imaging for enhancing the contrast of gray and white matter. NeuroImage 2004;12:251–260. [PubMed: 14741663]

[63]. Moosmann M, Eichele T, Nordby H, Hugdahl K, Calhoun VD. Joint independent component analysis for simultaneous EEG-fMRI: Principle and simulation. Int. J. Psych. to be published

[64]. Eichele T, Calhoun VD, Moosmann M, Specht K, Jongsma M, Quiroga R, Nordby H, Hugdahl K. Unmixing concurrent EEG-fMRI with parallel independent component analysis. Int. J. Psych. to be published

[65]. Liu J, Pearlson GD, Windemuth A, Ruano G, Perrone-Bizzozero NI, Calhoun VD. Combining fMRI and SNP data to investigate connections between brain function and genetics using parallel ICA. Hum.Brain Mapp. to be published

[66]. Liu J, Calhoun VD. Parallel independent component analysis for multimodal analysis: Application to fMRI and EEG Data. Proc. ISBI 2007:1028–1031.

[67]. Liu J, Calhoun VD. A novel approach to analyzing fMRI and SNP data via parallel independent component analysis. Proc. SPIE 2006:xx–xx.

[68]. Worsley KJ, Friston KJ. Analysis of fMRI time-series revisited—Again. NeuroImage 1995;2:173–181. [PubMed: 9343600]

[69]. Jung TP, Makeig S, Humphries C, Lee TW, McKeown MJ, Iragui V, Sejnowski TJ. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 2000;37:163–178. [PubMed: 10731767]

[70]. Thomsen T, Specht K, Hammar A, Nyttingnes J, Ersland L, Hugdahl K. Brain localization of attentional control in different age groups by combining functional and structural MRI. Neuroimage 2004;22:912–919. [PubMed: 15193622]

[71]. Clark L, Iversen SD, Goodwin GM. A neuropsychological investigation of prefrontal cortex involvement in acute mania. Amer. J. Psychiatry 2001;158:1605–1611. [PubMed: 11578991]

[72]. Calhoun VD, Adalı T, Pearlson GD, Kiehl KA. Neuronal chronometry of target detection: Fusion of hemodynamic and event-related potential data. Neuroimage Apr 1;2006 30:544–553. [PubMed: 16246587]

[73]. Cover, T.; Thomas, KM. Elements of Information Theory. Wiley; New York: 1991.

[74]. Hero, A.; Ma, B.; Michel, O.; Gorman, J. Alpha-Divergence For Classification, Indexing And Retrieval. Dept. EECS, Univ. Michigan; Ann Arbor, MI: 2001. p. 328

[75]. Demirci O, Calhoun VD. Detection of Schizophrenia using fMRI Data via Projection Pursuit. Proc. MLSP 2007:199–204.

[76]. Demirci O, Clark VP, Calhoun VD. A projection pursuit algorithm to classify individuals using fMRI data: Application to schizophrenia. NeuroImage. to be published

[77]. Johnson MR, Morris N, Astur RS, Calhoun VD, Kiehl KA, Pearlson GD. Schizophrenia and working memory: A closer look at fMRI of the dorsolateral prefrontal cortex during a working memory task. Proc.CNS 2005:xx–xx.

[78]. Kiehl KA, Stevens M, Laurens KR, Pearlson GD, Calhoun VD, Liddle PF. An adaptive reflexive processing model of neurocognitive function: Supporting evidence from a large scale (n = 100) fMRI study of an auditory oddball task. NeuroImage 2005;25:899–915. [PubMed: 15808990]

[79]. Giuliani N, Calhoun VD, Pearlson GD, Francis A, Buchanan RW. Voxel-based morphometry versus regions of interest: A comparison of two methods for analyzing gray matter disturbances in Schizophrenia. Schizophr. Res 2005;74:135–147. [PubMed: 15721994]

[80]. Kiehl KA, Smith AM, Hare RD, Liddle PF. An event-related potential investigation of response inhibition in Schizophrenia and psychopathy. Biol. Psychiatry 2000;48:210–221. [PubMed: 10924664]

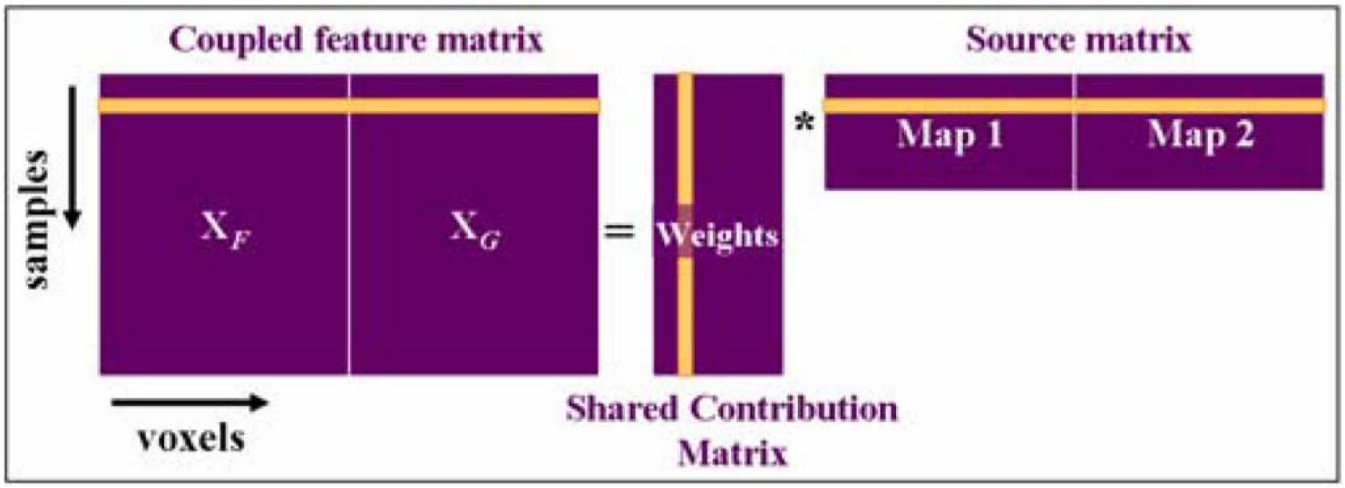

Fig. 1.

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 that share common mixing matrix parameters.

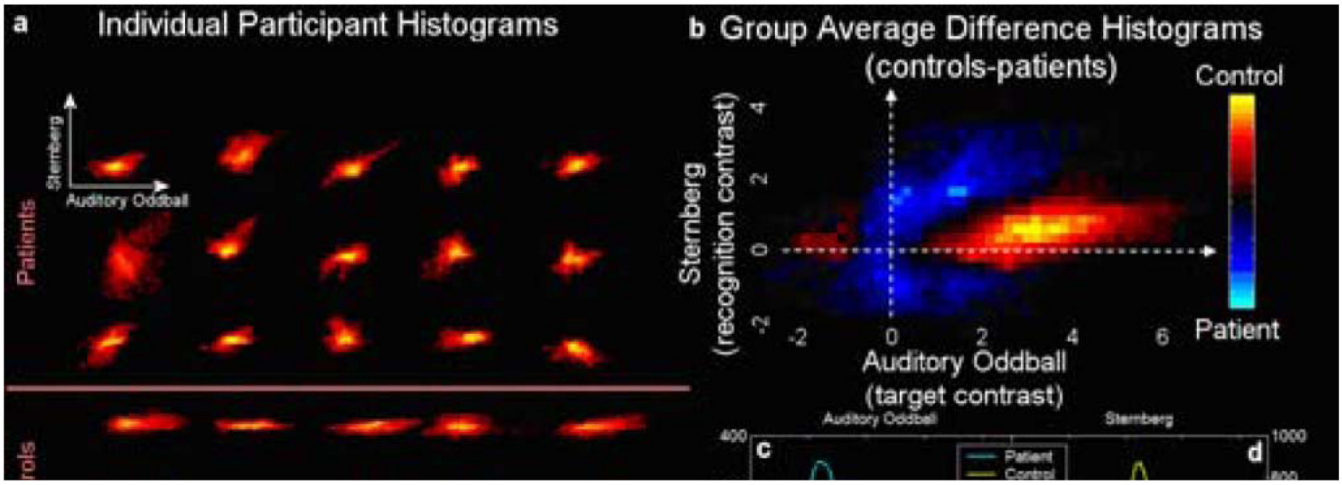

Fig. 2.

Cross-task 2-D histograms for AOD versus SB fMRI activation: joint 2-D 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 AOD (SPM contrast image for “targets”) (c) and Sternberg (SPM contrast image for “recall”) (d) data.

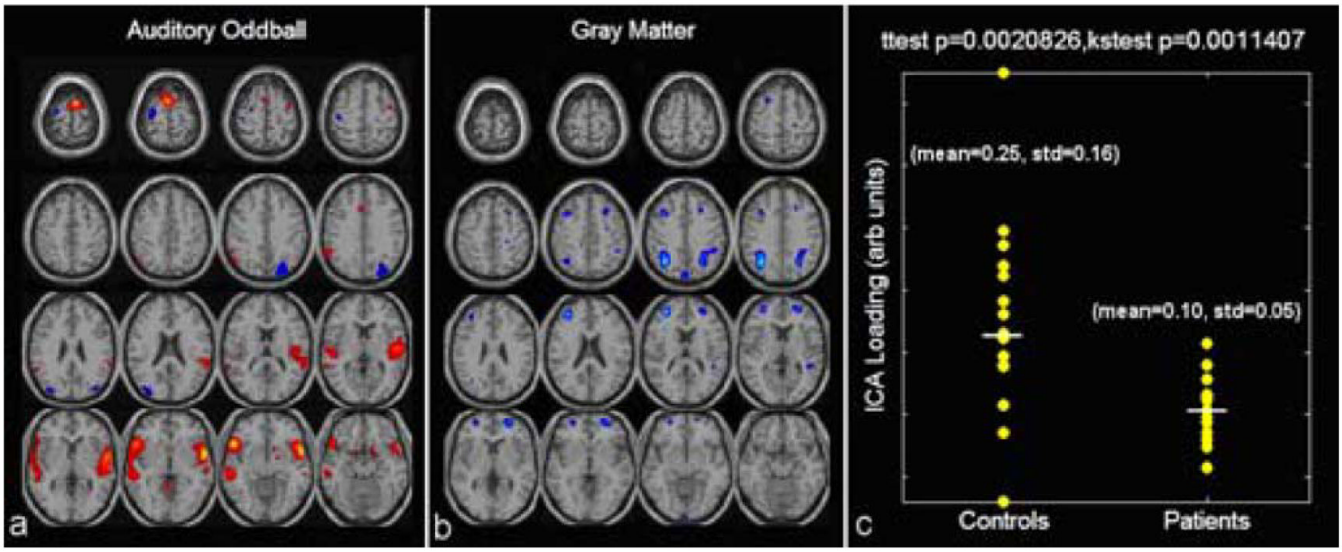

Fig. 3.

AOD/GM jICA analysis: only one component demonstrated a significant difference between patients and controls. The joint source map for the AOD (left) and GM (middle) data is presented along with the loading parameters for patients and controls (far right).

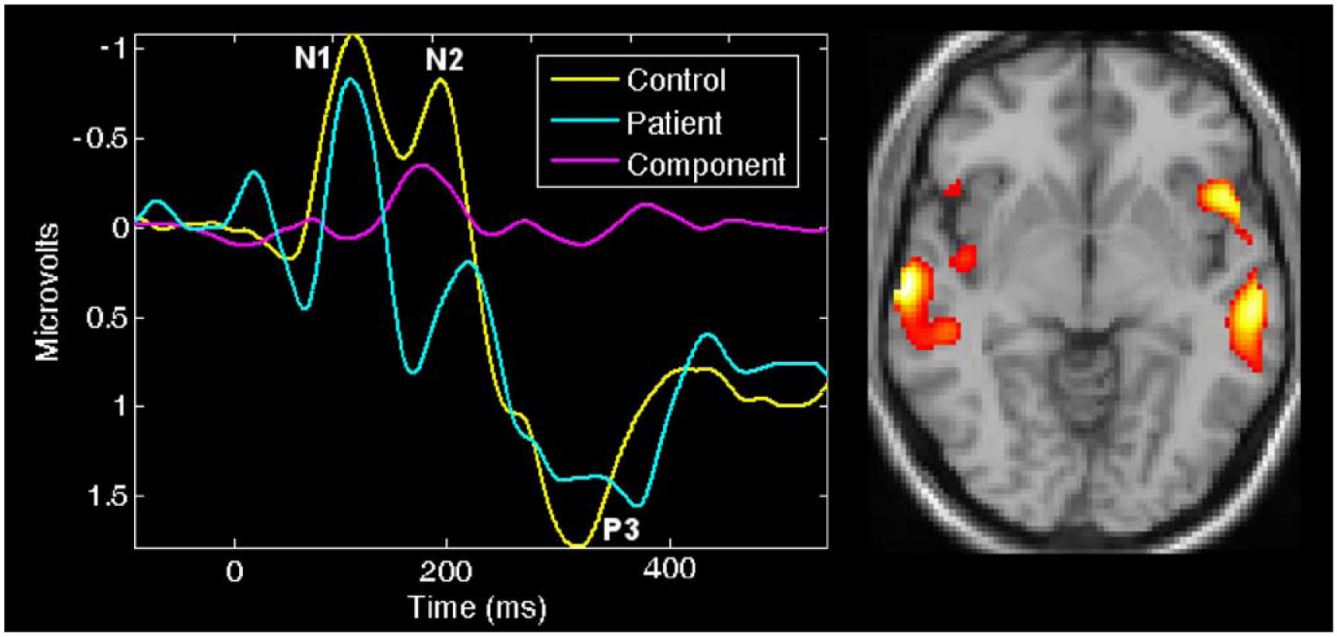

Fig. 4.

ERP/fMRI jICA: joint component that showed significantly different loading parameters ($p < 0.0001$) for patients versus controls. Control (yellow) and patient (blue) average ERP plots along with the ERP part of the identified joint component (pink) (left). Thresholded fMRI part of the joint component showing bilateral temporal and frontal lobe regions (right).

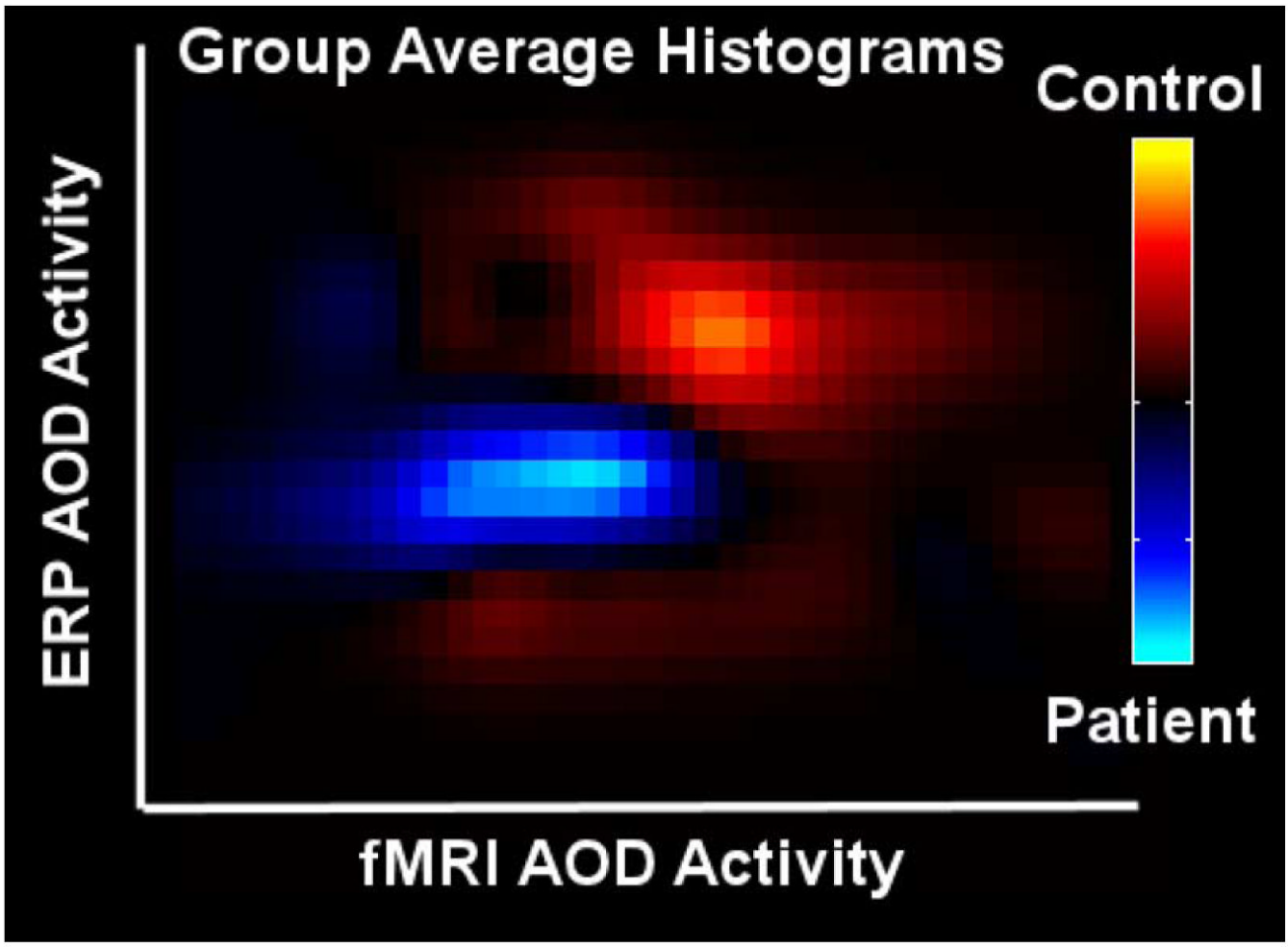

Fig. 5.
ERP/fMRI histograms: joint histograms for patients (blue) and controls (orange) (left). Simulated data from two Gaussians (a) showing a case in which marginal histograms (b) and (c) are less able to detect differences between groups whereas the histogram in the direction of maximal separation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

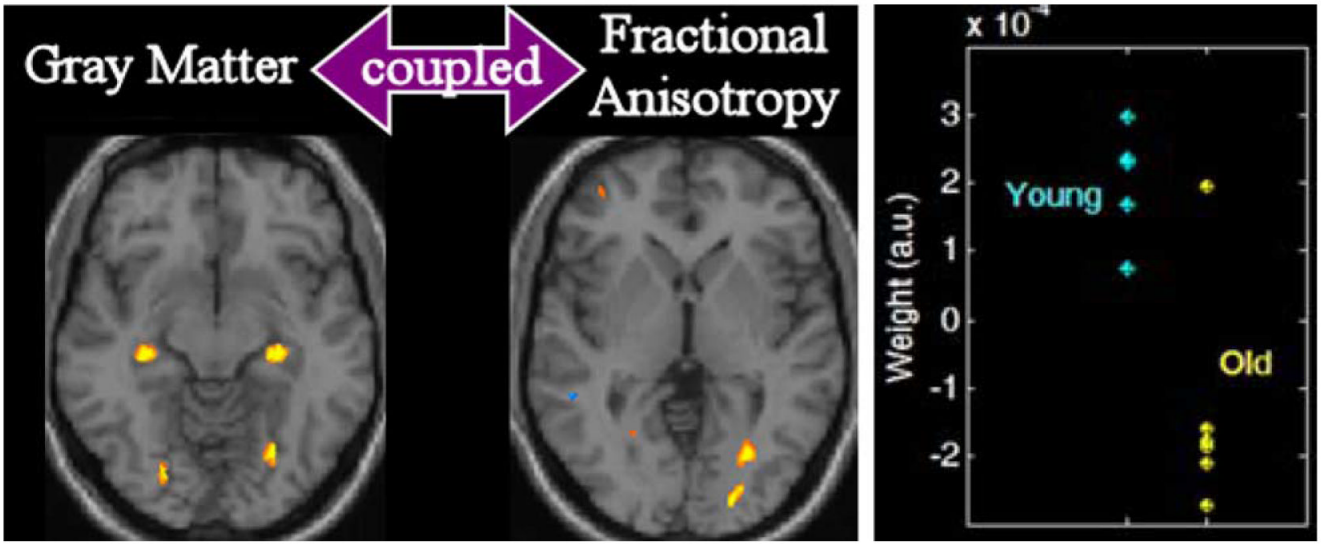

Fig. 6.
GM and FA spatial correspondence (left) and corresponding weights (right) comparing older and younger participants ($p < 0.002$).

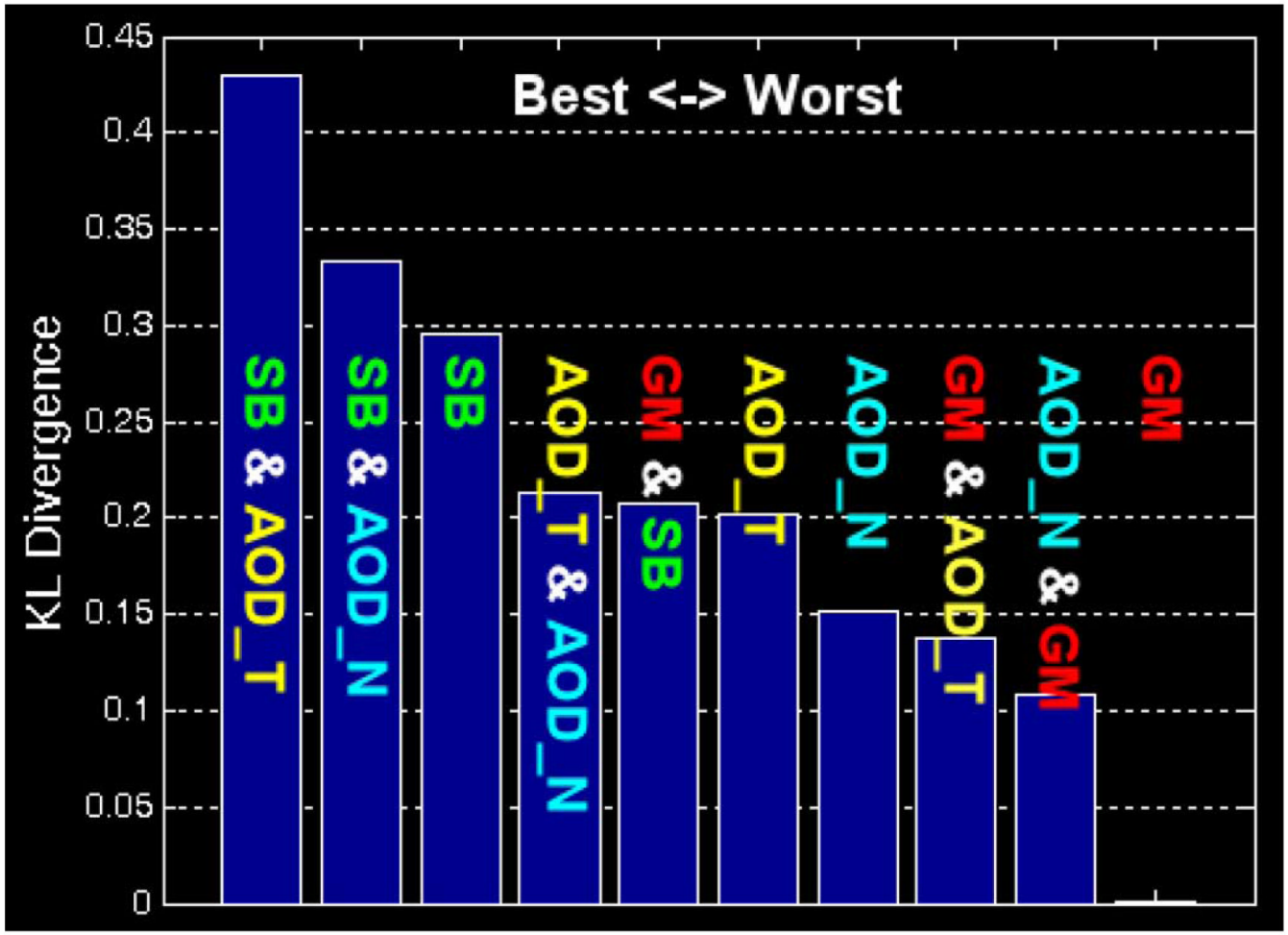

Fig. 7.
Example of evaluation of KL divergence for combinations of features comparing two-way fusion with nonfusion. Larger KL divergence values represent better patient versus control separation.

TABLE I

Core Features for FMRI (Sternberg Working Memory Task [SB], AOD), sMRI, and EEG (AOD task)

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

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.

TABLE II

Several Divergence Measures for Three Combinations of fMRI Features

	KL	J	Renyi (0.5)
Oddball & Stemberg	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)
Steinberg	0.050	0.050	0.0254