Outline

- fMRI/EEG data
- Three Approaches to integration/fusion
  - Prediction
  - Constraints
  - 2nd Level Fusion
- Conclusions

Approaches to EEG-MRI data integration

Data Integration through:

(i) Prediction
- some features of EEG to predict fMRI responses.

(ii) Constraints
- Spatial information from fMRI as Priors for Source Reconstruction

(iii) Fusion
- Common forward or generative model to explain EEG and fMRI data

Data integration/fusion: Previous Work

Data Fusion vs. Data Integration

- Definitions [Savopol ISPRS 2004]
  - Data Integration: The use of another data-type to improve the geometry or other information reconstructed from the first data-type (e.g. fMRI constrained EEG).
  - Data Fusion: Images analyzed jointly, hence allowing them to influence one another, to reveal inter-relationships between data-types.

Possible approaches for joint analyses

- Point-based
  - Correlation [Rosenstiel 2002]
  - Straightforward, but difficult to visualize

- Region-based
  - Interregional correlation [Thulborn et al. 1992, Mathalon et al. 2007]
  - Structural-equlotion modeling or dynamic causal modeling [McIntosh et al., 2003; Friston et al., 2003]
  - Useful for model testing, does not take into account all brain regions/times

- Transformation-based
  - A natural set of tools for this problem include those that transform data matrices into a smaller set of modes or components
    - Singular value decomposition [Friston et al., 1993; Friston et al., 1996]
    - Partial Least Squares [McIntosh, Bookstein et al., 1996]
    - Canonical Variates Analysis [Strother et al, 1995]

- Bayesian Hierarchical Models
  - Reverend Thomas Bayes (1702-1761)
  - Karl Friston et al. (now)

- Multiway Partial Least Squares
  - Pedro Valdes-Sosa

- ICA or jICA

\[ S = Ax \]
Separate vs. Joint Estimation

Linear mixtures with shared mixing parameter:

\[ x(t) = \sum_{i} a_i x_i(t) \quad \text{and} \quad x(t) = \sum_{i} a_i x_i(t) \]

In a non-joint analysis, we maximize the likelihood functions for each modality separately:

\[ p(x(t)|w) = \prod_{i} p(x_i(t)|w_i) \quad w = \text{somehow be fused together} \]

Resulting in two unmixing parameters, that then have to be somehow be fused together.

In contrast, for a joint analysis we maximize the joint likelihood function, resulting in a single fused unmixing parameter.

Joint ICA Approach

Mixing Model: \[ X(t) = AS(t) \]

For two sources:

\[ X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \]

Shared mixing matrix: \[ A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \]

The mixing equations are:

\[ x_1(t) = a_{11} s_1(t) + a_{12} s_2(t) \]

\[ x_2(t) = a_{21} s_1(t) + a_{22} s_2(t) \]

Update equation:

\[ \Delta W = -2y^T (y W) \]

where \( y(t) = g(u(t)) \) and \( g(x) = 1(1 + e^{-x}) \)

Overview

Brain 1 Brain 2 Brain N

\[ c \quad 0 \text{ ms} \quad 200 \text{ ms} \quad 300 \text{ ms} \]

ERP

fMRI

Auditory Oddball Task

Target Novel Standard

1 kHz tone, sweep, whistle

0.5 kHz Standard

Group Averaged fMRI/ERP Features

Preprocess

Feature Extraction

jICA

Noise Removal

Visualization
Joint ERP/fMRI Components

Visualization Snapshots…

ERP (temporal) Components: \( \mathbf{T} = [t_1 \ldots t_n] \)
FMRI (spatial) Components: \( \mathbf{S} = [s_1 \ldots s_n] \)

FMRI Image Snapshot: \( \mathbf{M}_f(t) = \mathbf{T} \cdot \mathbf{S}^T(t) \)

ERP Snapshot: \( \mathbf{M}_e(s) = \mathbf{T} \cdot \mathbf{S}^T(s) \)

FMRI Snapshots (movie)

FMRI Snapshots (stills)

ERP Snapshots

Parallel ICA
• Impose additional inter-modality correlation to improve ability to identify connections between EEG and fMRI
neuronal sources $s = (s_1, s_2, …, s_n)^T$ transformed by an unknown mixing matrix $A$ to the signal mixture $x = As$; features of the sources and their mixture observed as scalp EEG/ERP waveforms $E$ at timepoints $t$ and channels $c$ as $AE = [xE_1(tc), xE_2(tc), …, xEn(tc)]^T$; hemodynamic correlates of the detected in fMRI volumes $F$ at the respective locations $v$ as $AF = [xF_1(v), xF_2(v), …, xFn(v)]^T$; the mixture is entered into a common 2D space from which fused signals $y_{EF} = [yEF_1(tcv), yEF_2(tcv), …, yEFn(tcv)]^T$ are extracted by adjusting the unmixing matrix $W$ (the inverse of $A$), such that $y = Wx$ optimally represents $s$. The update equation for the algorithm to compute the common unmixing matrix $W$ and the fused EEG/ERP and fMRI sources, $u_{EF}$ is 

$$
\Delta W = \eta \{I - 2 y_{EF}(u_{EF})^T\}W, \quad \text{where} \quad y_{EF} = g(u_{EF}), \quad \text{and} \quad g(x) = \frac{1}{1+e^{-x}} \text{is the nonlinearity in the neural network}.
$$

Eichele, Calhoun, Moosmann et al., in progress

Fusion ICA Toolbox (FIT)

Outline

- Motivation
- Data integration/fusion
- ICA for data fusion
- Study 1: Multi-task fMRI in Sz
- Study 2: fMRI and GM in Sz
- Other Projects
- Summary
Motivation: we collect many data types

Multi-modal data collected

Brain Function (spatiotemporal, task-related)
EEG
fMRI

Brain Structure (spatial)
T1/T2
DTI

Covariates (Age, etc.)
Other*

Benefit to a coupled/joint analysis?

Data integration/fusion: Previous Work

Data Fusion vs. Data Integration

Definitions [Savopol ISPRS 2004]

Data Integration: The use of another data-type to improve the geometry or other information reconstructed from the first data-type (e.g. fMRI constrained EEG/DTI).

Data Fusion: Images analyzed jointly, hence allowing them to influence one another, to reveal inter-relationships between data-types.

No Inter-relationship
One-Way Inter-Relationship

- EEG
- fMRI
- DTI
- T1/T2
- Other*

Temporal Resolution:
- low
- high

Spatial Resolution:
- low
- high

Two-Way Inter-Relationship

- EEG
- fMRI
- DTI
- T1/T2
- Other*

Temporal Resolution:
- low
- high

Spatial Resolution:
- low
- high

Previous approaches

- Data integration models
  - fMRI constrained EEG [Dale, 1997; George, 1995]
  - fMRI constrained DTI [Kim, 2003]
- Data fusion models
  - Multi-task PET using SVD [Friston, 1994]
  - PET and fMRI using PLS [Chen, 2004]
  - EEG and fMRI using multiway-PLS [Martinez-Moreno, 2004]

Possible approaches for joint analyses

- Voxel-based
  - Correlation (0 wealthy)
  - Straightforward, but difficult to visualize
- Region-based
  - Intergroup correlation [Marin, et al., 1984]
  - Structural equation modeling (McIntosh and Gonzalez-Lima, 1994; Friston, et al., 2003)
  - Structural equation modeling (McIntosh and Gonzalez-Lima, 1994; Beckel & Friston, 1997)
  - Multiple regression and extensions (e.g., McIntosh, Beckel & Friston, 1993)
  - Bayesian networks (Dynamic Causal Modeling, Friston, Penny, et al., 2003)
- Transformation-based
  - A natural set of tools for this problem include those that transform data matrices into a smaller set of modes or components
  - Singular value decomposition [Friston, et al., 1993; Friston, et al., 1996]
  - Partial Least Squares [McIntosh, Bookstein, et al., 1993]
  - Canonical Variates Analysis [Strother, et al., 1995]

ICA for data fusion

- In contrast to a first-level ICA approach, we instead introduce the idea of a second level, feature-based analysis of the fMRI activation maps (the features) generated from a first-level analysis
- We propose a method, called joint ICA, that enables the decomposition of activation maps generated from two tasks into joint, maximally spatially independent components

Feature-based

- What is a feature?
  - Lower dimensional data containing information of interest
  - Examples: An image of activation amplitudes, A gray matter segmentation image, fractional anisotropy image
- Advantages
  - Less-computationally complex/easier to model
  - Takes advantages of existing analytic approaches
  - Examines covariance of multiple data types at the subject level

<table>
<thead>
<tr>
<th>Feature</th>
<th>EEG</th>
<th>fMRI</th>
<th>DTI</th>
<th>Other*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compression ratio</td>
<td>1:56.45MB</td>
<td>1:1.628 MB</td>
<td>1:2010 MB</td>
<td>1:2001.3 MB</td>
</tr>
<tr>
<td>Feature size</td>
<td>25MB</td>
<td>46 MB</td>
<td>200 MB</td>
<td>260 MB</td>
</tr>
</tbody>
</table>
Data preprocessing and feature extraction are crucial steps in the analysis of multimodal brain imaging data. These steps involve several stages:

1. **Preprocessing**: This stage includes operations such as spatial normalization, smoothing, and normalization of the data. The goal is to align and standardize the data from different modalities to facilitate further analysis.

2. **Feature Extraction**: Features are extracted from the preprocessed data. This can involve methods such as spatial ICA decomposition, which can separate the data into independent components. Each component is associated with a specific brain region or functional activity.

3. **Data Mining**: Techniques like joint feature matrix composition are used to combine the features from different modalities. This results in a single fused unmixing parameter, which can be used to model the shared mixing parameters across different tasks.

4. **Component Selection**: The selected components are then used for further analysis, such as image reconstruction or task-specific modeling.

5. **Algorithm for jICA Analysis**: The joint ICA (jICA) algorithm is used to analyze the data. This involves maximizing the joint likelihood function, which results in a single fused unmixing parameter. In contrast to a non-joint analysis, where the likelihood functions for each parameter have to be fused together, a joint analysis can model the shared mixing parameters across tasks.

6. **Simulation: Hybrid-Experiment**: The simulation demonstrates the application of the jICA algorithm to real-world data, showing how the algorithm can be used to generate component images for different tasks and groups.

The framework for this analysis is depicted in the diagram, showing the flow of data from raw data to the final component selection. The algorithm for jICA analysis is described more formally in the text, with equations detailing the process. The simulation shows a practical application of these concepts.
jICA of multiple-tasks

Study 1: Multi-task fMRI

Subject Information
- Fifteen healthy controls and fifteen chronic schizophrenia outpatients in complete or partial remission
- Patients met criteria for schizophrenia in the DSM-IV on the basis of a structured clinical interview and review of the case file (mean duration of illness=11.2±6.3 years)
- All but one patient with schizophrenia were stabilized on atypical antipsychotic medications.
- Equal numbers of males (N=12/12) and females (N=3/3) and all but two participants in each group were right handed.
- All participants had normal hearing (assessed by self-report) and were able to perform both tasks successfully during practice prior to the scanning session.

Scan Parameters
- 3.0 Tesla Siemens Allegra
- TR=1.50s [AOD]/1.86s [SB]
- TE=27ms
- Field of view=24cm
- Acquisition matrix=64x64
- Flip angle=70 deg
- Slice thickness=4 [AOD]/3 [SB] mm
- Gap=1mm, 29 [AOD]/36 [SB] slices
- Ascending acquisition
- Six “dummy” scans

Schizophrenia and Disconnection
- Schizophrenia is a complex condition with a diverse and heterogeneous clinical presentation, which makes it unlikely that the neural mechanism underlying the disorder is limited to a circumscribed brain dysfunction.
- Schizophrenia has been associated with both structural and functional abnormalities in neocortical networks including frontal, parietal, and temporal regions.
- Schizophrenia is thought to involve a disturbance of coupling between large-scale cortical systems [Breakspear, 2003; Friston, 1998; Pearlson, 1999].
- There have been a number of models proposed with many studies implicating regions in temporal lobe, cerebellum, thalamus, basal ganglia, and lateral frontal regions [Weinberger 1987; McCarley et al., 1991; Andreasen et al., 1998; Braver et al., 1999; Friston 1999].
- Some suggested models
  - Dis-coordination models [Andreasen et al., 1996]
  - Fronto-temporal disconnection models [Liddle et al., 1992]

Auditory Oddball & Sternberg Tasks
Preprocessing
Timing/motion correction
Spatial normalization
Spatial smoothing
Feature Extraction
GLM Analysis

AOD
SBS--WM

jICA
10 components

Joint Feature Matrix
Component Selection

15 Patients
15 Controls

jICA Results
One component
Significant at p<0.01

Joint Histograms

Inter-task correlation

Brief Summary of Study 1

• Finding 1:
  • Schizophrenia patients demonstrate “decreased” connectivity in the joint network identified using the jICA approach.
  • This network includes regions in temporal lobe, cerebellum, thalamus, basal ganglia, visual cortex, and lateral frontal regions, and these findings are consistent with both the cognitive dysmetria and fronto-temporal disconnection.
  • Our findings thus argue against new, less efficient, areas being recruited by patients due to impaired connectivity between regions which are normally utilized.

• Finding 2:
  • A second finding is that for the voxels identified by the jICA analysis, the correlation between the two tasks was significantly higher in patients than controls.
  • This finding suggests that schizophrenia patients activate “more similarly” for both tasks than controls. The degree to which a brain activation map is different from that of another task may reflect the degree to which performance on a task is “specialized” to a certain set of regions.
  • A possible synthesis of both findings is that patients are activating less, but also activating with a less unique set of regions for these very different tasks. This suggests both a global attenuation of activity as well as a breakdown of specialized wiring between cognitive domains. Alternative explanations?
Study 2: fMRI and GM Fusion in Sz

- All participants had normal hearing (assessed by self-report) and were able to perform both tasks successfully during practice prior to the scanning session.
- Scan Parameters (3.0 Tesla Allegra)
  - fMRI EPI Scan
    - TR=1.50s, TE=27ms
    - field of view=24cm, 64x64
    - flip angle=70 deg
    - st=4mm, gap=1mm, 29 slices
    - ascending acquisition
  - sMRI T1 MP-RAGE
    - TR/TE/TI = 2300/2.74/900 ms
    - flip angle = 8°
    - FOV = 176x256 mm
    - Slab thickness = 176 mm
    - Matrix = 176x256x176
    - Voxel size =1x1x1 mm
    - Number of echoes = 2
    - Pixel bandwidth =190 Hz
    - Total scan time =10:09 min

jICA Results

- One component
- Significant at p>0.01
We have used a joint ICA model to examine linearly related fMRI auditory oddball target activation data and gray matter segmentation data.

A single component which was significantly different between patients with schizophrenia and healthy controls were examined.

Finding 1:
- Gray matter regions in bilateral parietal lobe and frontal lobe as well as right temporal lobe were found to be associated with auditory oddball activations in bilateral temporal lobe.
- Previous findings of diminished activity to AOD and decreased gray matter in temporal lobe regions were replicated.

Finding 2:
- In the regions showing the largest group differences, gray matter concentrations were larger in patients versus controls, suggesting that more gray matter may be related to less functional connectivity in the auditory oddball fMRI task.
- As a whole, these findings suggest a possible morphological substrate for auditory oddball connectivity changes in schizophrenia (although this cannot be proven in the current study).
- In general, we suggest studying interactions between gray matter data and fMRI data provides a useful way to examine structure and function.

Much more data fusion is needed to make sense of the vast amounts of data we are collecting!

Caution: causality cannot be established with correlational or related methods!

Data fusion approaches enable us to ask novel questions about fMRI data and provides a way to potentially help us gain a more complete understanding of brain structure and function.

Our choice of modeling the shared dependence between the tasks with the mixing parameters should be examined in more studies before the true utility of this assumption will be known.

Future work includes advancing the statistical models used, especially incorporating prior information about each of the data types.