Boosted Spatial and Temporal Precision in Functional Brain Imaging via Multimodal Analysis

Yaroslav O. Halchenko
yh42@njit.edu

Computer Science Department,
NJIT

Ph.D. Thesis Proposal

The Goal

General

Develop methods to achieve superior spatio-temporal resolution by combining signals from different brain imaging modalities that possess complementary temporal and spatial advantages.

Specific

Show that it is possible to obtain trustworthy estimate of neuronal activity at superior spatio-temporal resolution by combining EEG/MEG with fMRI data whenever forward models of the signals are appropriate to describe the data in terms of underlying neuronal processes.

Motivating Questions for Brain Scientists

Fundamental

How can we understand brain function?

Localization

Which areas of the brain are involved in the processing during a specific task?

Brain dynamics

What are the interactions among the areas during a specific task?

Motivating Questions for Engineers

Forward problem

How brain signals and stored information can be modeled to produce registered measurements?

Inverse problem

How viable estimates of the neuronal processes inside the brain can be obtained from a limited set of observations outside the brain?

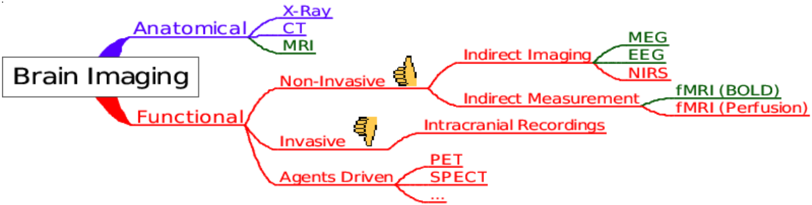
Signal processing

What characteristics (*e.g.* non-stationarity, statistical or frequency features, *etc.*) of the brain imaging data should be explored under heavy noise conditions?

Outline

Outline

Brain Imaging

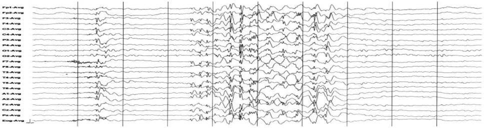


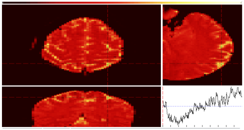
Non-Invasive Unimodal Brain Imaging

EEG

MEG

MRI





Electro- and Magnito- EncephaloGraphy

Common features

- Passive technique
- Post-synaptic ionic currents of synchronized pyramidal neurons generate the electro-magnetic field registered by E/MEG

Differences

EEG

- On the head surface
- Electric potential
- Reference electrode
- Silent to solenoidal currents

MEG

- Outside of the head
- Magnetic field
- Reference-free
- Silent to radially oriented currents

E/MEG Brain Imaging

Linear formulation: DECD

Both magnetic and electric fields linearly depend on the current strength at densely sampled fixed spatial locations

$$\mathbf{X} = \mathbf{G}\mathbf{Q}$$

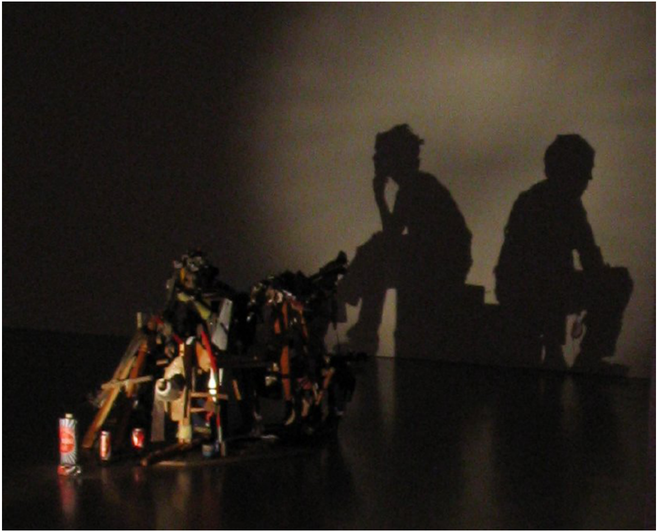
$\mathbf{X}$ $(M \times T)$ – E/MEG data;
$\mathbf{G}$ $(M \times N)$ – spatial filter (lead-field/gain matrix);
$\mathbf{Q}$ $(N \times T)$ – current strengths at each location

Easy!

For the linear case the solution is $\hat{\mathbf{Q}} = \mathbf{G}^{+}\mathbf{X}$

Not That Easy: Inverse Problem

Is That What You Had in Mind?



Inverse Problem

Why it is problematic

Ill-posed: the number of possible signal source locations (N) greatly exceeds the number of sensors (M) – infinite number of solutions

Ill-conditioned: instrumental and brain noise prevents from achieving stable solution by simply increasing number of sensors

E/MEG Inverse Regularization

Minimal 2-nd norm solution: pseudo-inverse

$$\mathbf{G}^{\dagger} = \mathbf{G}^{\top}(\mathbf{G}\mathbf{G}^{\top})^{-1}$$

Regularization: general formulation

$$\mathbf{G}^{+} = \mathbf{W_Q}\mathbf{G}^{\top}(\mathbf{G}\mathbf{W_Q}\mathbf{G}^{\top} + \lambda\mathbf{W_X})^{-1},$$

where $\mathbf{W_X}^{-1}$ and $\mathbf{W_Q}^{-1}$ are weighting matrices in sensor and source spaces correspondingly

E/MEG Pro et Contra

Pros: great temporal resolution

- Great for any event related design
- Epileptic spikes detection
- Coherence analysis
- Human brain interface

Cons: poor localization in space

- Non-linear optimization in the case of dipole modeling
- Inverse problem in the case of distributed dipole modeling

fMRI: Blood Oxygenation Level Dependent

Pros

Great spatial resolution: 1 mm and higher

Safe: does not require injections of radioactive isotopes

Cons

Indirect measurement: BOLD response reflects oxygenation

Low temporal resolution:

- Full volume can be acquired just every 2-4 seconds
- BOLD signal itself is of convolved nature

Noise:

- Inhomogeneities
- Blood vessels influence

Motivation for Multimodal Imaging

- Superior spatial resolution of fMRI
- Fine temporal resolution of E/MEG
- Reported agreement between E/MEG and BOLD signals

Existing Multimodal Techniques

- Correlative analysis
- Decomposition analysis
- Constrained equivalent current dipole (ECD) modeling
- FMRI-conditioned distributed ECD modeling
- Beamforming with fMRI-conditioned covariance
- Bayesian inference

Existing Multimodal Techniques

- Correlative analysis
- Decomposition analysis
- Constrained equivalent current dipole (ECD) modeling
- FMRI-conditioned distributed ECD modeling
- Beamforming with fMRI-conditioned covariance
- Bayesian inference

Problems

- Absent generative model of BOLD signal
- Variability of BOLD across subjects and within the brain
- True neural signal is not known
- Methods do not make use of temporal fMRI information

Problems

- Absent generative model of BOLD signal
- Variability of BOLD across subjects and within the brain
- True neural signal is not known
- Methods do not make use of temporal fMRI information

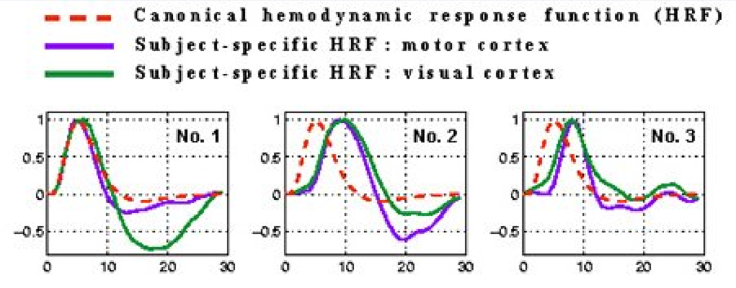
Outline

Major Obstacle: Absent Generative BOLD Model

Linear Time Invariant System

$$f(t) = (h * q)(t)$$

Hemodynamic Response Function



[Kalina Christoff, 2001]

BOLD Signal: LTIS (Convolutional) Model

Observation

Convolutional model is valid in many cases

Convolutional model

- provides good agreement between LFP and BOLD response
- permits the estimation of convolution kernel using simple stimulus
- has been used in most of the fMRI studies
- can be augmented with non-linearity to accommodate divergence from LTIS model

Forward Models

Temporally and spatially superior modality $\mathbf{Q}$ ($N \times T$) is used to reconstruct both $\mathbf{F}$ and $\mathbf{X}$ observed signals

Modality	Data Matrix	Size	Model	Description
E/MEG	$\mathbf{X}$	$M \times T$	$\hat{\mathbf{X}} = \mathbf{GQ}$	Spatial Filter
fMRI	$\mathbf{F}$	$N \times U$	$\hat{\mathbf{F}} = \tilde{\mathbf{Q}}\mathbf{B}$	Temporal Filter

Dipole projections: $q = \left[q_x q_y q_z\right]$

Dipole strength: $\tilde{q}_{jt} = \sqrt{q_{x\,it}^2 + q_{y\,it}^2 + q_{z\,it}^2}$

Dipole orientation: $\Theta_{jt} = q_{jt} / \tilde{q}_{it}$, where $i = j \bmod N$

Forward Models

Temporally and spatially superior modality $\mathbf{Q}$ $(N \times T)$ is used to reconstruct both $\mathbf{F}$ and $\mathbf{X}$ observed signals

Modality	Data Matrix	Size	Model	Description
E/MEG	$\mathbf{X}$	$M \times T$	$\hat{\mathbf{X}} = \mathbf{GQ}$	Spatial Filter
fMRI	$\mathbf{F}$	$N \times U$	$\hat{\mathbf{F}} = \tilde{\mathbf{Q}}\mathbf{B}$	Temporal Filter

Advantages

- Modeling both E/MEG and fMRI makes use of temporal and spatial information from both modalities
- Reconstruction of fMRI along with E/MEG provides regularization to the inverse E/MEG problem

The Unknown: Dipole Strength $\leftrightarrow$ BOLD

Scaling between dipole strength and BOLD signal is not known and can vary from location to location

Solutions

- Restrict range of applications to activations in small (thus approximately homogeneous) regions
- For the area of interest estimate scaling along with convolution kernel using simple experimental design
- Augment the model to include scaling parameter per each local region

The Unknown: Dipole Strength $\leftrightarrow$ BOLD

Scaling between dipole strength and BOLD signal is not known and can vary from location to location

Solutions

- Restrict range of applications to activations in small (thus approximately homogeneous) regions
- For the area of interest estimate scaling along with convolution kernel using simple experimental design
- Augment the model to include scaling parameter per each local region

Reconstruction Error

Residuals

$$\Delta_{\mathbf{X}}(\mathbf{Q}) = \frac{\hat{\mathbf{X}}(\mathbf{Q}) - \mathbf{X}}{\sqrt{\nu_{\mathbf{X}} MT}} \text{ and } \Delta_{\mathbf{F}}(\mathbf{Q}) = \frac{\hat{\mathbf{F}}(\mathbf{Q}) - \mathbf{F}}{\sqrt{\nu_{\mathbf{F}} NU}}$$

Quality of the reconstruction criterion:

$$\mathcal{E}_r(\mathbf{Q}) = \|\Delta_{\mathbf{X}}(\mathbf{Q})\|_l + \alpha\|\Delta_{\mathbf{F}}(\mathbf{Q})\|_l + \lambda\mathcal{C}(\mathbf{Q})$$

where

- $l \in \{1, 2\}$: the norm of error cost function
- $\mathcal{C}(\mathbf{Q})$: additional regularization term

$l = 2$: Gradient Descent Optimization

$$\frac{\partial \mathcal{E}_r(\mathbf{Q})}{\partial \mathbf{Q}} = \frac{\partial \Delta_{\mathbf{X}}(\mathbf{Q})}{\partial \mathbf{Q}} + \alpha \frac{\partial \Delta_{\mathbf{F}}(\mathbf{Q})}{\partial \mathbf{Q}} + \lambda \frac{\partial \mathcal{C}(\mathbf{Q})}{\partial \mathbf{Q}}$$

$$\frac{\partial \Delta_{\mathbf{X}}(\mathbf{Q})}{\partial \mathbf{Q}} = 2\mathbf{G}^T(\mathbf{X} - \mathbf{G}\mathbf{Q}), \quad \frac{\partial \Delta_{\mathbf{F}}(\mathbf{Q})}{\partial \mathbf{Q}} = 2\Theta \star \left((\mathbf{F} - \tilde{\mathbf{Q}}\mathbf{B})\mathbf{B}^T\right)$$

$l = 2$: Gradient Descent Optimization

$$\frac{\partial \mathcal{E}_r(\mathbf{Q})}{\partial \mathbf{Q}} = \frac{\partial \Delta_{\mathbf{X}}(\mathbf{Q})}{\partial \mathbf{Q}} + \alpha \frac{\partial \Delta_{\mathbf{F}}(\mathbf{Q})}{\partial \mathbf{Q}} + \lambda \frac{\partial \mathcal{C}(\mathbf{Q})}{\partial \mathbf{Q}}$$

$$\frac{\partial \Delta_{\mathbf{X}}(\mathbf{Q})}{\partial \mathbf{Q}} = 2\mathbf{G}^T(\mathbf{X} - \mathbf{G}\mathbf{Q}), \quad \frac{\partial \Delta_{\mathbf{F}}(\mathbf{Q})}{\partial \mathbf{Q}} = 2\Theta \star \left((\mathbf{F} - \tilde{\mathbf{Q}}\mathbf{B})\mathbf{B}^T\right)$$

Advantages

- Simple formulation
- Efficient modifications of gradient descent can be used
- Can easily incorporate other regularization terms

$l = 2$: Gradient Descent Optimization

$$\frac{\partial \mathcal{E}_r(\mathbf{Q})}{\partial \mathbf{Q}} = \frac{\partial \Delta_{\mathbf{X}}(\mathbf{Q})}{\partial \mathbf{Q}} + \alpha \frac{\partial \Delta_{\mathbf{F}}(\mathbf{Q})}{\partial \mathbf{Q}} + \lambda \frac{\partial \mathcal{C}(\mathbf{Q})}{\partial \mathbf{Q}}$$

$$\frac{\partial \Delta_{\mathbf{X}}(\mathbf{Q})}{\partial \mathbf{Q}} = 2\mathbf{G}^T(\mathbf{X} - \mathbf{G}\mathbf{Q}), \quad \frac{\partial \Delta_{\mathbf{F}}(\mathbf{Q})}{\partial \mathbf{Q}} = 2\ominus \star \left((\mathbf{F} - \tilde{\mathbf{Q}}\mathbf{B})\mathbf{B}^T\right)$$

Problems

- Optimization can fall into local minima

Linear Programming Formulation

Minimization task can be formulated as an LP problem

$$\begin{aligned} \hat{\mathbf{X}} + \Delta_{\mathbf{X}} &= \mathbf{X} && \textit{Constraints} \\ \hat{\mathbf{F}} + \Delta_{\mathbf{F}} &= \mathbf{F} && \\ \tilde{q}_{ij} &\geq 0 && \textit{Region} \\ \mathcal{E} = \|\Delta_{\mathbf{X}}\|_1 + & \alpha\|\Delta_{\mathbf{F}}\|_1 && \textit{Objective} \end{aligned}$$

Linear Programming Formulation

Minimization task can be formulated as an LP problem

$$
\begin{aligned}
\hat{\mathbf{X}} + \Delta_{\mathbf{X}} &= \mathbf{X} && \textit{Constraints} \\
\hat{\mathbf{F}} + \Delta_{\mathbf{F}} &= \mathbf{F} && \\
\tilde{q}_{ij} &\geq 0 && \textit{Region} \\
\mathcal{E} = \|\Delta_{\mathbf{X}}\|_1 + &\alpha \|\Delta_{\mathbf{F}}\|_1 && \textit{Objective}
\end{aligned}
$$

Advantages

- Sum of absolute errors found to be a much better criterion in the case of present outliers
- Side effect of LP formulation is the minimization of $\|\mathbf{Q}\|_1$

Linear Programming Formulation

Minimization task can be formulated as an LP problem

$$
\begin{aligned}
\hat{\mathbf{X}} + \Delta_{\mathbf{X}} &= \mathbf{X} && \textit{Constraints} \\
\hat{\mathbf{F}} + \Delta_{\mathbf{F}} &= \mathbf{F} && \\
\tilde{q}_{ij} &\geq 0 && \textit{Region} \\
\mathcal{E} = \|\Delta_{\mathbf{X}}\|_1 + &\ \alpha \|\Delta_{\mathbf{F}}\|_1 && \textit{Objective}
\end{aligned}
$$

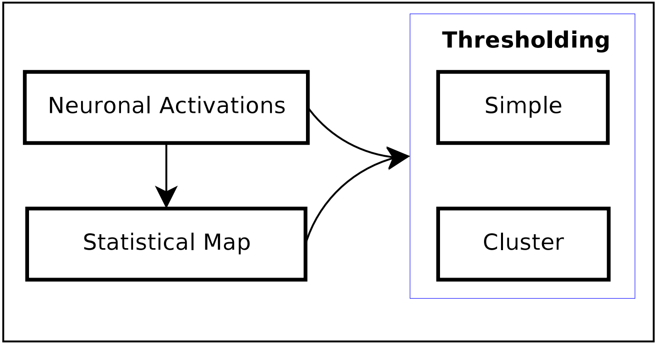
Problems

- Efficient LP solvers are necessary due to the large size of LP problem (MOSEK)
- Possibly poor performance if noise is indeed Gaussian

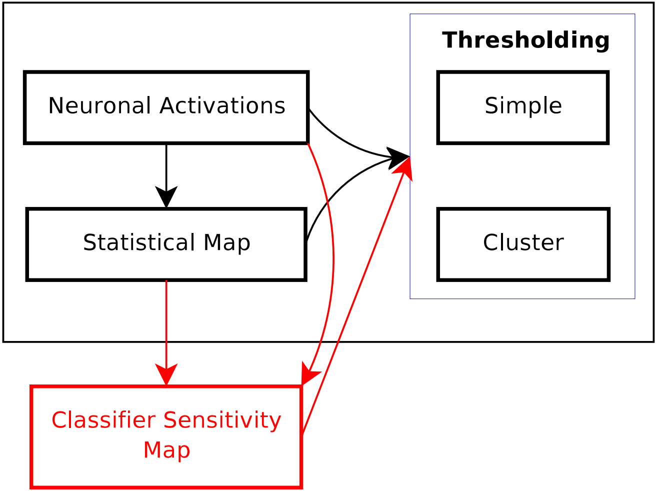
Outline

1. The “State of Art”
2. Research Issues
   - Multiple Modalities Data Integration
   - Activity Localization
3. Problem Area
4. Simulations
5. Research Plan and Timeline

Localization Workflow



Localization Workflow



Classifier as a Localizer

Localization using classifiers

Temporal: trained classifier

Spatial: sensitivity map of the classifier

Advantages

- Notion of generalization
- Fast classification after the classifier has been trained

Disadvantages

- Training can be lengthy
- Might not generalize
- Sensitivity map might reflect just a subset of activations

Classifier as a Localizer

Localization using classifiers

Temporal: trained classifier

Spatial: sensitivity map of the classifier

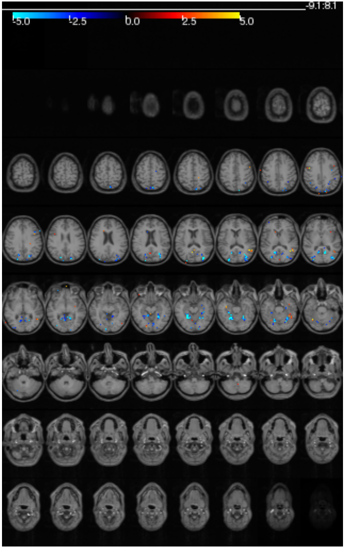
Advantages

- Notion of generalization
- Fast classification after the classifier has been trained

Disadvantages

- Training can be lengthy
- Might not generalize
- Sensitivity map might reflect just a subset of activations

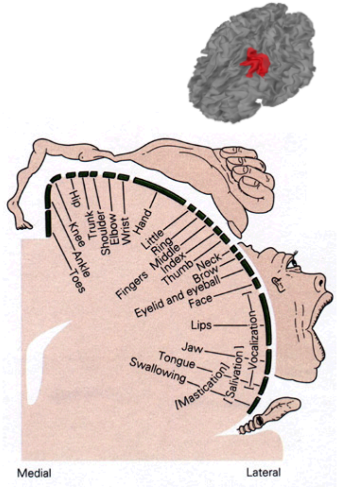
Localization Using SVM



- Great ability to generalize
- Fast to train (constrained quadratic problem)
- Can easily work with data of huge dimensionality
- Sensitivity map of linear SVM is given by the decision hyper-plane normal
- Results are consistent with conventional analysis

Outline

Somatotopy: Mapping of the Primary Motor (M1)



- Simple motor response
- Experiment is easily reproducible
- Coarse information about spatial organization is available
- Temporal separation between events is easily controllable

M1 Mapping

Possible problems

- Convolutional model might not be valid
- Activations in other areas (PMA, SMA and PI) can interfere with registration of the signal of interest
- Suggested multimodal analysis methods may not produce good estimates of neuronal activity

Solutions

- Carry out a pilot experiment to verify applicability of the convolutional model
- Augment the model with non-linearity if necessary
- Preprocess the data to extract signal components of interest (ICA?, SOBI?)

M1 Mapping

Possible problems

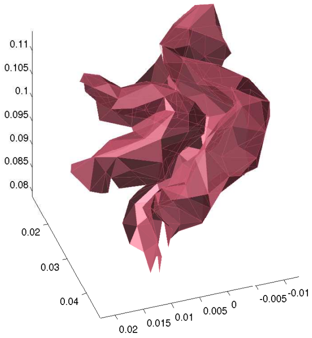
- Convolutional model might not be valid
- Activations in other areas (PMA, SMA and PI) can interfere with registration of the signal of interest
- Suggested multimodal analysis methods may not produce good estimates of neuronal activity

Solutions

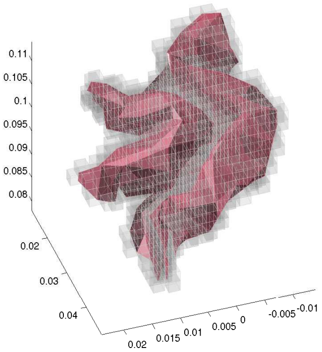
- Carry out a pilot experiment to verify applicability of the convolutional model
- Augment the model with non-linearity if necessary
- Preprocess the data to extract signal components of interest (ICA?, SOBI?)

Outline

Region of Interest: M1 “hand area”



(a) Cortical Mesh



(b) 895 Surrounding 2 mm Voxels

Datasets

E/MEG sensors: 30 sensors (895 voxels)

Sampling rate: Sources (and E/MEG): 16 [Hz], fMRI: 1 [Hz]

Duration: Sources (and E/MEG): 1 [sec], fMRI: 10 [sec]

Noise: (1) Gaussian white and (2) empirical

Noise levels: $\varepsilon = \sigma_\epsilon / \max(s) \in [\,0,\ 0.1,\ 0.2,\ 0.4,\ 0.6\,]$

An activation: Modeled as a Gaussian (σ=50 [ms])

Trials: 30 trials

Arrangement: 5 datasets

- Spatially non-overlapping: [1, 10, 100, 895] active
- Spatially overlapping: 10 randomly activated locations followed by 2nd activation within next 100–300 [ms]

FMRI Conditioned E/MEG Inverse (FMRI-DECD)

$$\hat{\mathbf{Q}} = \mathbf{G}^{+}\mathbf{X}, \text{ where } \mathbf{G}^{+} = \mathbf{W_Q}\mathbf{G}^{\top}(\mathbf{G}\mathbf{W_Q}\mathbf{G}^{\top})^{-1}$$

Conditioning of the inverse :
Truncated SVD of $(\mathbf{G}\mathbf{W_Q}\mathbf{G}^{\top})$

Gain matrix normalization :
$\mathbf{W_Q} = \mathbf{W}_\mathrm{n} = \left(\mathrm{diag}\,(\mathbf{G}^{\top}\mathbf{G})\right)^{-1}$

Relative fMRI weighting :
$(\mathbf{W}_{\mathrm{fMRI}'})_{ii} = \nu_0 + (1 - \nu_0)\Delta_i/\Delta_{\mathrm{max}}$.
$\nu_0 \in [\,1.0,\ 0.5,\ 0.1\,]$ which corresponds to 0, 50, and 90% of relative fMRI weighting

Dipole orientations :
Variable and Fixed

L_2 -Fusion

$$\hat{\mathbf{Q}} = \arg\min_{\mathbf{Q}} \|\Delta_{\mathbf{X}}(\mathbf{Q})\|_2 + \alpha\|\Delta_{\mathbf{F}}(\mathbf{Q})\|_2$$

Trade-off Parameter :

$\alpha = [0.5, 1, 10]$ for a tradeoff between E/MEG and FMRI fit was used

Reconstruction Quality Criterion

Relative noise energy brought into the source signal estimation

$$E = \left(\frac{||\hat{\mathbf{Q}} - \mathbf{Q}||_2}{||\mathbf{Q}||_2} \right)^2$$

Minimal value $E = 0$ corresponds to the perfect restoration of the sources time course.

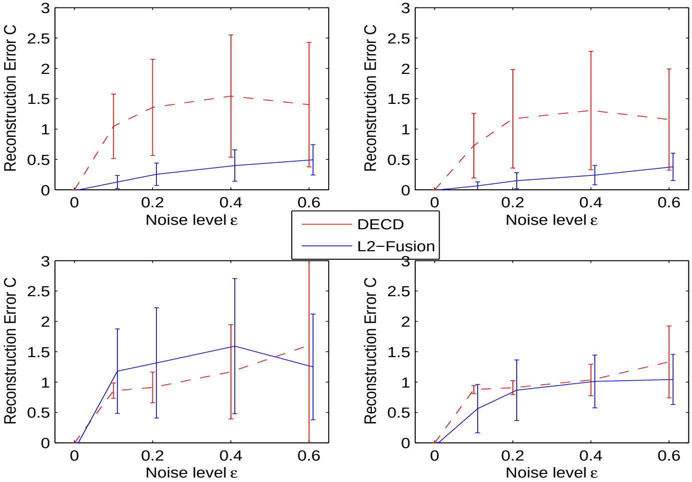
The best result obtained with fMRI conditioned E/MEG inverse was chosen to be compared against L_2 -Fusion results.

A Single Active Source

Empirical **Gaussian**

EEG

MEG

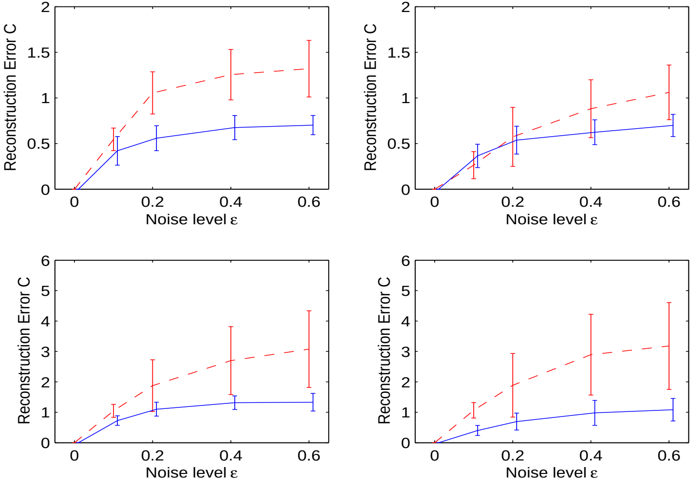


10 Active Sources

Empirical **Gaussian**

EEG

MEG

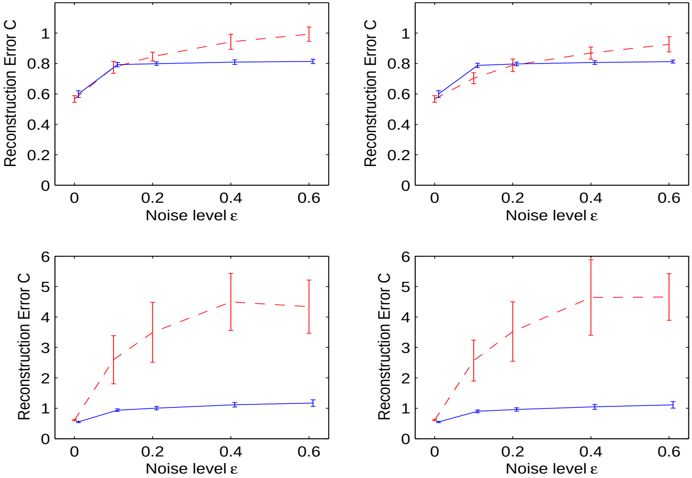


100 Active Sources

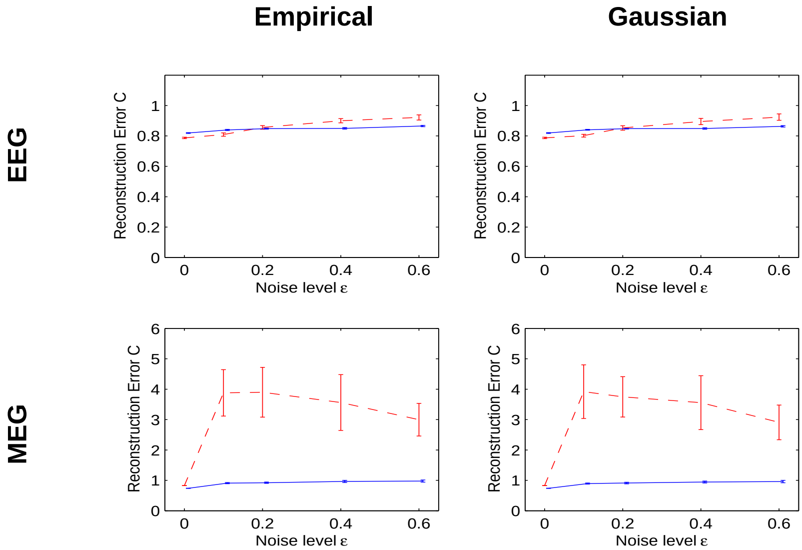
Empirical | Gaussian

EEG

MEG



895 Active Sources

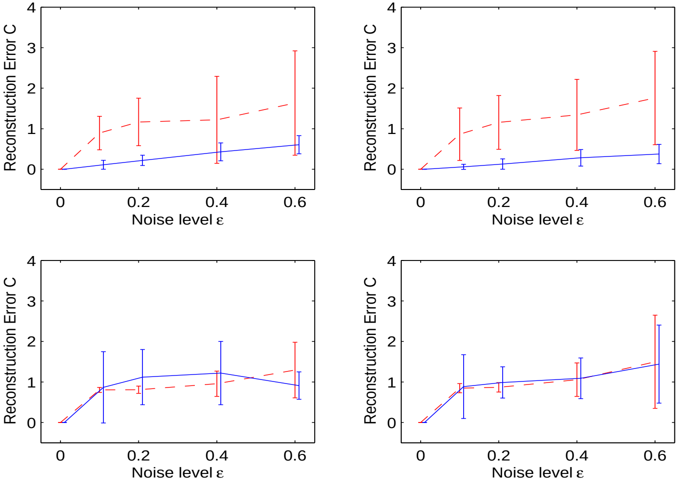


10 Spatially Overlapping Active Sources

Empirical | Gaussian

EEG

MEG



L_2 -Fusion Outperforms FMRI-DECD

- L_2 -Fusion is more noise-robust than FMRI-DECD
- L_2 -Fusion constantly outperforms FMRI-DECD on the large number of non-overlapping sources
- L_2 -Fusion performs as well as FMRI-DECD on overlapping sources in case of MEG and outperforms it with EEG
- FMRI-DECD on MEG data fails with increased number of sources
- Gaussian noise model is well suited for modeling of E/MEG instrumental noise

Outline

Summary: Completed Work

- An overview of the existing multimodal imaging approaches revealed advantages, drawbacks and difficulties associated with any particular method
- Two novel methods (L_1 - and L_2 -Fusion) of multimodal analysis were suggested
- Neuroimaging problem to be tackled with multimodal methods was chosen
- The simulation environment for a somatotopic experiment was created to facilitate comparative performance analysis of different methods
- Simulated data was used to compare L_2 -Fusion with the conventional methods under different noise conditions and source arrangements

Proposed Work Timeline

Sep – Oct 2005

- Evaluate the quality of reconstruction achieved using L_1 -Fusion on the simulated dataset
- Apply proposed localization method to the simulated data to assess its performance
- Carry out a pilot fMRI/EEG experiment to verify applicability of the convolutional model for fMRI

Proposed Work Timeline: Continued

Nov – Dec 2005

- Analyze the trade-off between spatial and temporal resolution achieved by the proposed methods on simulated data
- Setup fMRI acquisition protocol to achieve reliable sub-mm spatial resolution over the region of interest
- Design somatotopic experiment based on resolution limits of the methods revealed by simulation studies

Proposed Work Timeline: Continued

31 Dec 2005 – 02 Jan 2006

- Celebrate New Year

Jan – Mar 2006

- Collect fMRI and EEG data
- Perform the described analysis and draw conclusions
- Complete the dissertation

Do Not Forget to Shut Down the Lights

Thank you

Somatotopy

Definition

Somatotopy The topographic association of positional relationships of receptors in the body via respective nerve fibres to their terminal distribution in specific functional areas of the cerebral cortex.

Requirements for a Benchmark Study

- BOLD signal should be well described by convolutional model
- Experimental design has to be non-parametric
- Activations have to be reproducible and stationary in time
- There must be a possibility to control the spatial and temporal distance between the activations

Outline

6 Experimental Design and Data Preprocessing

The Structure of a Brain Imaging Study

- Choose a brain imaging problem
- Design and setup an experiment
- Acquire the data
- Preprocess the data
- Fusion: integrate imaging data from multiple modalities
- Localize neuronal activity of interest

The Structure of a Brain Imaging Study

- Choose a brain imaging problem
- Design and setup an experiment
- Acquire the data
- Preprocess the data
- Fusion: integrate imaging data from multiple modalities
- Localize neuronal activity of interest

The Structure of a Brain Imaging Study

- Choose a brain imaging problem
- Design and setup an experiment
- Acquire the data
- Preprocess the data
- Fusion: integrate imaging data from multiple modalities
- Localize neuronal activity of interest

Specifics of the Experimental Design

"In Concert" multimodal data :
collect EEG and fMRI data in separate sessions on the same subject and using identical experimental design. Subject responses has to be recorded

Controlled spatio-temporal tradeoff :
explore different spatial and temporal distances between the sources

Data corregistration :
3D digitization of fiducial points and their alignment across the sessions

EEG data preprocessing :
ICA (or SOBI) analysis to extract components of interest