Single Trial BOLD functional MRI

César Caballero Gaudes

Marie Curie Research Fellow
PhD Student, School of Computer Science

The University of Nottingham
• Introduction to MRI and fMRI

• Brief introduction to existing fMRI data analysis techniques

• Research project: Detection of Single Trial BOLD responses

• Experiments and results

• Conclusions
Magnetic Resonance Imaging (MRI)
Functional Magnetic Resonance Imaging (fMRI)
Blood oxgenation level dependent (BOLD) imaging

FMRIB, Oxford

Baseline State

Active State

Oxyhemoglobin
Diamagnetic

Deoxyhemoglobin
Paramagnetic
Blood oxigenation level dependent (BOLD) imaging

BOLD fMRI is not a direct measure of neuronal activity. It is related to blood hemodynamics (blood flow, blood volume and oxygen concentration).
Model-Based Methods

- Traditional model-based approaches rely on an accurate definition of the paradigm, e.g. onset and duration of the stimuli.

- In their simplest form, Model based methods assume identical hemodynamic (BOLD) response across different brain regions.

- Model comparison and statistical inference (interpretation) is simple.

- Best example: GLM. Implemented in SPM, FSL or AFNI. They differ on how to deal with the noise.
Exploratory Methods

• Data-driven approaches
• No a-priori knowledge of the shape of the BOLD response
• No paradigm specification
• Identify components (or clusters) which best fit the data according to some statistical measure.
• Multivariate (spatio-temporal) analysis: Functional connectivity & Resting States
• Post-processing to decide which (or how many) components are relevant.
• Best Example: ICA, but also PCA, CCA, TCA or clustering methods.

<table>
<thead>
<tr>
<th>Introduction</th>
<th>Methods</th>
<th>Experiments</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Our goal</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Map in space and time the brain’s response to <strong>single trial</strong> stimulus <strong>without prior information of the stimulus timing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Detection of true single-events</strong> can facilitate the characterization of higher cognitive processes, such as learning or adaptation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Challenging task</strong> in the presence of physiological and systematic fluctuations which obscure the detection of the BOLD event</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our methodology aims to identify individual trials and <strong>track the spatial/temporal evolution</strong> of the BOLD response</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Our proposal: Paradigm-Free Mapping

Spatial and Temporal Preprocessing

Ridge Regression deconvolution → Spatio-temporal T-statistic → Multiple Hypothesis Testing: FDR → Temporal T-maps
Our main assumption: fMRI signal as a linear model

BOLD response modelled as the sum of two gamma functions

\[ h(t) = g(t; \tau_1, l_1) - \frac{1}{6} g(t; \tau_2, l_2) \]

\[ g(t; \tau, l) = \frac{l^\tau t^{\tau-1} e^{-lt}}{\Gamma(\tau)} \]

fMRI signal = convolution of stimulus function with BOLD response

\[ x(t) = h(t) \otimes s(t) + e(t) \]

G.M. Boynton et al., J. Neuroscience, 1996, 16(13):4207-4221
### Linear model

\[ y = Hs + e \]

\( H = \) canonical HRF toeplitz matrix

### Least Squares:

\[ \min \| y - Hs \|^2 \]

\[ \hat{S}_{\text{LS}} = \left( H^T \Sigma^{-1} H \right)^{-1} H^T \Sigma^{-1} x \]

### Regularized Least Squares:

(\textbf{Ridge Regression})

\[ \min \| y - Hs \|^2 + \lambda \| s \|^2 \]

\[ \hat{S}_{\text{RR}} = \left( H^T \Sigma^{-1} H + \lambda I \right)^{-1} H^T \Sigma^{-1} x \]

\[ \lambda = \frac{N \hat{\sigma}^2}{s_{\text{LS}}^T H^T \Sigma^{-1} H s_{\text{LS}}} \]
Noise variance \( (\sigma_k^2) \) is estimated voxelwise from the voxel \( k \) baseline images.

Noise covariance matrix \( \Sigma_k \) is estimated voxel-wise after pooling the baseline images of a group of voxels surrounding the current voxel.

\[
V_C = \sum_{k \in C} x_k^B \left( x_k^B \right)^T
\]

Levinson Durbin Recursion + finite sample MDL criteria

\[
\Sigma_k
\]
Autoregressive model for the noise

\[ e(t) - \sum_{i=1}^{n} a_i e(t-i) = w(t) \]

\[ r(0) - \sum_{i=1}^{n} a_i r(i) = \sigma_n^2 \]

Yule-Walker Equations

\[
\begin{bmatrix}
    r(0) & r(-1) & \cdots & r(-N) \\
    r(1) & r(0) & \cdots & r(-2) \\
    \vdots & \ddots & \ddots & \vdots \\
    r(N) & \cdots & r(1) & r(0)
\end{bmatrix}
\begin{bmatrix}
    1 \\
    a_1 \\
    \vdots \\
    a_n
\end{bmatrix}
=
\begin{bmatrix}
    \sigma_n^2 \\
    0 \\
    \vdots \\
    0
\end{bmatrix}
\]

Baselines are input to Levinson-Durbin Recursion to estimate AR coefficients and residuals for each candidate order \( n=\{1,\ldots,N\} \)

Model selection criteria: Which is the best order \( N \) for the noise model?
Model selection criteria: Which is the best order $N$ for the noise model?

- Selection is based on the residual power for each candidate order.
- Traditional criteria, such as AIC, GIC or MDL, are asymptotic criteria, i.e., developed for infinite number of observations.
- Finite sample versions apply correction factors to account for few number of observations.

<table>
<thead>
<tr>
<th>Asymptotic</th>
<th>Finite sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AIC(n)</strong></td>
<td>$\ln\left(\sigma_n^2\right) + 2 \frac{n}{K}$</td>
</tr>
<tr>
<td><strong>MDL(n)</strong></td>
<td>$\ln\left(\sigma_n^2\right) + \ln(K) \frac{n}{K}$</td>
</tr>
</tbody>
</table>

Our proposal: Paradigm-Free Mapping

Spatial and Temporal Preprocessing

Ridge Regression deconvolution → Spatio-temporal T-statistic → Multiple Hypothesis Testing: FDR → Temporal T-maps

Introduction
Methods
Experiments
Conclusions
To statistically test for the presence of activation, a t-statistic is defined from the Ridge Regression estimate, comparing for each voxels the signal at each time point to the mean of the baseline.

\[ s(n) = \frac{1}{L} \sum_{k=1}^{L} s_k(n) \]

\[ s(n) \sim N(\mu_0, \sigma_L^2) \quad i \leq 1 \leq B \]

\[ s(n) \sim N(\mu_1, \sigma_L^2) \quad B \leq i \]

Hypothesis Test

\[ H_0 : \mu_0 = \mu_1 \]

\[ H_1 : \mu_0 \neq \mu_1 \]

\[ t(n) = \frac{s_k(n) - \hat{\mu}_L}{\hat{\sigma}_L \sqrt{1 + \frac{1}{B}}} \sim t_{B-1} \quad n \geq B \]
Our proposal: Paradigm-Free Mapping

Spatial and Temporal Preprocessing

Ridge Regression deconvolution → Spatio-temporal T-statistic → Multiple Hypothesis Testing: FDR → Temporal T-maps
<table>
<thead>
<tr>
<th></th>
<th>Negatives</th>
<th>Positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$ true</td>
<td>$V_{0N}$</td>
<td>$V_{0P}$</td>
</tr>
<tr>
<td>$H_0$ false</td>
<td>$V_{1N}$</td>
<td>$V_{1P}$</td>
</tr>
<tr>
<td></td>
<td>$V_N$</td>
<td>$V_P$</td>
</tr>
</tbody>
</table>

**Family-wise error rate (Bonferroni):**
Probability of any false positives $\leq \alpha$

$$P(V_{0P} \geq 1) \leq \alpha_{FWE}$$

**False Discovery Rate:**
Ratio of false positives $\leq \alpha$

$$E \left\{ \frac{V_{0P}}{V_P} \right\} \leq \alpha_{FDR}$$

Too strict
Low power

Data-dependent
More power
1- Sort uncorrected p-values for $V$ tests

\[ p_{(1)} \leq \cdots \leq p_{(i)} \leq \cdots \leq p_{(V)} \]

2- Reject those hypothesis where

\[ p_{(i)} \leq \alpha_{FDR} \frac{i}{V} \]

Or compute & reject corrected FDR p-values

\[ Q(p_{(i)}) = \min_{k \geq i} \left\{ p_{(k)} \frac{V}{k} \right\} \leq \alpha_{FDR} \]

5 subjects were scanned (one subject twice) on a Philips 7T MR system. Two data sets were acquired per subject at TR 2s and 0.4s. Experiment was a motor paradigm where subjects performed unilateral visually-cued (VC) and self-paced (SP) finger tapping.

Electromyography (EMG) signals were recorded in the left extensor, right flexor and right extensor to detect hand-related movements.
Datasets were corrected for motion (AFNI), low frequency drifts (DCT order 4 and linear trend) and physiological noise (RETROICOR).

CSF voxels were excluded from analysis after classification based on the mean and standard deviation of the baseline volumes.

Spatial clustering (AFNI, NIMH/NIH) was applied on the thresholded maps in order to reduce false positives.

Glover et al., MRM, 2000, 44:162-167
No information about the paradigm (except # of baselines) was employed for data analysis
Results: T-dynamics

Brain activity is detected in:

- Supplementary motor area and cingulate gyrus: initiation and self-control of motor movements.
- Primary motor cortex: motor execution
- Primary somatosensory cortex: touch and proprioception.
- Primary and secondary visual cortex
Results: Delay maps/Time to peak

Graphs showing fMRI change over peristimulus time for Visual cue (VCT) and Self-paced (VCT).

Maps showing regional brain activity for VCT and SPT conditions.
Results: Spontaneous brain responses at rest

Figure 1. Temporal T-statistic computed for one subject, and the corresponding EMG. Activity corresponding to instructed tapping is marked on the graphs.

Figure 2. A, B, C show example activation maps for three slices corresponding to the to the A, B, C markings on the graphs in Figure 1. The slice locations are shown on the T2* image. Panels D: TAP2 and E illustrate maps during tapping and when no activity was detected (marked in Figure 1).
Conclusions

BOLD fMRI allows to map in space and time the cortical response associated to a given cognitive process.

A novel fMRI data analysis method has been developed based on the Ridge Regression deconvolution and an FDR corrected temporal T-statistic.

Using this method it is feasible to detect single trial events with NO information about stimuli timing and without averaging: Paradigm-Free Mapping.

Future work will address signal processing techniques which consider the sparse nature of the brain: LASSO, Dantzig Selector.
Acknowledgements

Work in collaboration with

Natalia Petridou
Penny Gowland
Susan Francis
Ian Dryden