Statistical Analysis of fMRI Data using Wavelets in a Probabilistic Atlas Space

Ivo Dinov
Department of Statistics &
Laboratory of Neuro Imaging
UCLA School of Medicine
Los Angeles, CA 90095
dinov@loni.ucla.edu
http://www.loni.ucla.edu/~dinov

☺ IPAM BMI Summer School Quotes … ☺

I mean no disrespect with these quotes, on the contrary I find them impressive …
All quotes taken mostly out of context … Most quotes are paraphrased …

“Our Brain Mapping Approach is Fuzzy …”
“There are no nice flat-maps [preserving length/area]”
“All this is very simple and clear …”
 “[One] can use Bayesian MAP estimation to solve everything [tissue, warp, bias] …”
“This is a Brainless talk …”
 “[We] write this and this and apply this to get this …”
“I have no idea what this [slide] shows …”
“If Matlab can’t do it it’s not worth studying/trying …”
“We wanted to obtain an MRI and a CRYO volume of 1 postdoc …”
“Avoid any registration …”
 “[Our] challenge is to keep a baby quiet, it’d nothing to do with math/eng/neuroimaging …”
Work with:

- W. John Boscardin  
  Department of Biostatistics  
  UCLA School of Public Health  
  http://rem.ph.ucla.edu/~johnb/  

- Elizabeth Sowell, Paul Thompson,  
  Roger Woods & Arthur Toga  
  Department of Neurology, LONI, UCLA  
  http://www.loni.ucla.edu/  

- Michael Mega, Neural Net Research, Portland, OR  
  http://www.loni.ucla.edu/~mega  

Hemoglobin – a molecule to breathe with  

Sources: http://wsrv.clas.virginia.edu/~rjh9u/hemoglob.html, Jorge Jovicich

Hemoglobin (Hgb):
- four globin chains
- each globin chain contains a heme group
- at center of each heme group is an iron atom (Fe)
- to each heme group an oxygen atom (O₂) can attach
- oxy-Hgb (four O₂) is diamagnetic → no ∆B (net magnetization) effects
- deoxy-Hgb is paramagnetic → if [deoxy-Hgb] ↓ local ∆B ↓.
  
  like aluminum/platinum, deoxy-Hgb, has small positive magnetic susceptibility
**BOLD signal**

**Blood Oxygen Level Dependent signal**

↑neural activity ➔ ↑ blood flow ➔ ↑ oxyhemoglobin ➔ ↑ T2* ➔ ↑ MR signal

---

<table>
<thead>
<tr>
<th>Basal state</th>
<th>Activated state</th>
</tr>
</thead>
<tbody>
<tr>
<td>- normal flow</td>
<td>- increased flow</td>
</tr>
<tr>
<td>- basal level [Hbr]</td>
<td>- decreased [Hbr] (lower field gradients around vessels)</td>
</tr>
<tr>
<td>- normal MRI signal</td>
<td>- increased CBV</td>
</tr>
<tr>
<td>- increased [Hbr]</td>
<td>- increased MRI signal (from lower field gradients)</td>
</tr>
</tbody>
</table>

Source: [fMRI: Brief Introduction to fMRI](#)

---

**Hemodynamic Response Function**

% signal change = (point – baseline)/baseline 
usually 0.5-3%

Initial dip
- more focal and potentially a better measure
- somewhat elusive so far, not everyone can find it

Time to rise
- signal begins to rise soon after stimulus begins

Time to peak
- signal peaks 4-6 sec after stimulus begins

Post stimulus undershoot
- signal suppressed after stimulation ends

Source: [Jorge Jovicich](#)
Haemoglobin acts as an endogenous intravascular contrast agent. As the level of oxygenation changes, so too does the contrast in the images.
Problems with BOLD

- How localized is the BOLD response to the site of neuronal activity?
  - Is the signal from draining veins rather than the tissue itself?
- How is the signal change coupled to neuronal activity?
  - Do changes in timing of BOLD responses reliably tell us about changes in timing of neural activity?

Activation Imaging using BOLD

- **Resting** state versus **Active** state
  - e.g. Finger tapping, word recognition.
- Whole brain scanned in ~3-5 seconds using a high speed imaging technique (EPI).
- Perform analysis to detect regions which show a signal increase in response to the stimulus.
fMRI Data Analysis Tools (non exclusive!)

- **Brain Voyager** [http://www.brainvoyager.de/](http://www.brainvoyager.de/)
- **CARET** [http://brainmap.wustl.edu/resources/caretnew.html](http://brainmap.wustl.edu/resources/caretnew.html)
- **FSL** [http://www.fmrib.ox.ac.uk/fsl/](http://www.fmrib.ox.ac.uk/fsl/)
- **SCIrun** [http://software.sci.utah.edu/scirun.html](http://software.sci.utah.edu/scirun.html)
- **LONI Pipeline** [http://www.loni.ucla.edu](http://www.loni.ucla.edu)
- **SPM** [http://www.fil.ion.ucl.ac.uk/spm/](http://www.fil.ion.ucl.ac.uk/spm/)

---

**Hypotheses vs. Data driven Approaches**

**Hypothesis-driven**
- Examples: t-tests, correlations, general linear model (GLM)
  - a priori model of activation is suggested
  - data is checked to see how closely it matches components of the model
  - most commonly used approach (e.g., SPM)

**Data-driven** — Independent Component Analysis (ICA)
- No prior hypotheses are necessary
- Multivariate techniques determine the patterns in the data that account for the most variance across all voxels
- Can be used to validate a model (see if the math comes up with the components you would've predicted)
- Can be inspected to see if there are things happening in your data that you didn't predict
- Can be used to identify confounds (e.g., head motion)
- Need a way to organize the many possible components
- New and upcoming
A. Region-Of-Interest (ROI) approach (e.g., LONI’s Sub-Volume Thresholding)
1. A localizer run(s) to find a region (e.g., show moving rings to find middle temporal MT area)
2. Extract time course information from that region in separate independent runs
3. See if the trends in that region are statistically significant
   The runs that are used to generate the area are independent from those used to test the hypothesis.

Example study: Tootell et al, 1995, Motion Aftereffect

Localize “motion area” MT (V5) in a run comparing moving vs. stationary rings

B. Whole volume statistical approach
1. Make predictions about what differences you should see if your hypotheses are correct
2. Decide on statistical measures to test for predicted differences (e.g., t-tests, correlations, GLMs)
3. Determine appropriate statistical threshold
4. See if statistical estimates are significant

Statistics available
1. T-test
2. Correlation
3. Frequency-Based (Fourier/Wavelet/Fractal) modeling
4. General Linear Model
   - overarching statistical model that lets you perform many types of statistical analyses (incl. correlation/regression, ANOVA)
Why do we need statistics?

- fMRI intensities essentially represent a random field
  - variation caused by scanner (magnet, coil, time)
  - variation in neurophysiology (area, task, subject)

- The rest-state scans used against activation tasks in subtraction paradigm setting to correct for some of these effects within the same run.

- Statistical analyses help us confirm whether the eyeball tests of significance based on visual thresholding are real/correct.

- Because we do so many comparisons (~10^6), we need a way to compensate (increase of Type I error).

---

**Paradigm:** Event-related design to assess between-population differences of amplitude and variance of hemodynamic response to visual stimulus (14 young, 14 old nonAD and 13 AD subjects)

- 16 coronal slices
- 64x64 in-plane voxels
- volume time = 2 sec
- 3 x 3 x 5 mm^3 voxels

- 4 runs per Subject
- 128 volumes per run
- 2 (randomly chosen) conditions (always 7+8)
- 1 Trial = 21.44 sec (8 vol’s)
- Total 60 Trials/subject

**Stimuli required a right hand index-finger click**

**blank**

**1-trial condition—1.5 second 8Hz B/W checkerboard flicker**

**2-trial condition—2 visual stimuli 5.36 sec apart**

One Approach – Voxel-Based T-tests

Look for differences b/w One-Trial & Two-Trial Activation, for a given voxel:

Measure average MR signal and SD for each volume in which 1-Trial stimuli were presented (7-8 epochs x 8 volumes/epoch = 56-64 volumes).

Measure average MR signal and SD for each volume in which 2-Trial stimuli were presented (56-64 volumes).

Stat-Significant Mean difference?

Calculate $T_0$ value. Look up $p$ value for that number of degrees of freedom (df $\geq 56 \times 2 = 112$).

e.g., For $\sim 112$ df $T_0 > 1.98 \Rightarrow p < 0.05$

$T_0 > 3.39 \Rightarrow p < 0.001$

Repeat this process 65,536 more times (64x64x16), once for each voxel.

To look for OneTrial – TwoTrial Activation, look at the negative tail of the comparison.

If $p=0.001$ and 65,536 voxels, 65,536*0.001 $\Rightarrow$ 66 voxels could be significant purely by chance.

Activation Statistics

Statistical Map superimposed on anatomical MRI image

Functional images ~2s

ROI Time Course

Time Condition

Region of interest (ROI)

4 ~ 5 min
T-test: Maps

For each voxel in the brain, we can now color code a map based on the computed t and p values:

We can do this for the positive tail (1 Trial – 2 Trial)
Orange = low significance
Yellow = high significance
Clear False-Positives!

And we can also do this for the negative tail (2 Trial – 1 Trial)
Blue = low significance
Green = high significance

Correlation: Incorporating the HRF

We can model the expected curve of the data by convolving our predictor with the hemodynamic response function.

To find a 1-Trial responsive area, we can correlate the convolved face predictor with each voxel time course

56 Model points
56 data points
Data 0
1 Model
Value of Predictor
Value of fMRI Signal

m(x)
Model

Predictor
m(x)

g(x)
Kernel

f(x)
Data

(1, 21, 41, 61, 81, 101, 121, 141)

(0, 21, 41, 61, 81, 101, 121, 141)

(0, 21, 41, 61, 81, 101, 121, 141)
Two Main fMRI Designs

**Block design**
- Compare long periods (e.g., 16 sec) of one condition with long periods of another
- Traditional approach
- Most statistically powerful approach
- Less dependent on how well you model the HRF (hemo)
- Habituation effects?!!

**Event-related design**
- Compare brief trials (1 sec) of one condition with brief of another
- Newer approach (1997+)
- Less statistically powerful but has some advantages
- Trials can be well-spaced to allow the MR signal to return to baseline b/w trials (e.g., 21+ sec apart) or closely spaced (e.g., every 2 sec)

Variability of HRF: Evidence

Aguirre, Zarahn & D'Esposito, 1998
- HRF shows considerable variability between subjects
- Within subjects, responses are more consistent, although there is still some variability between sessions

- [Graph](#)
  - different subjects
  - same subject, same session
  - same subject, different sessions

Parse out variance in the voxel’s time course to the contributions of six predictors plus residual noise (what predictors can’t account for).

\[ fMRI \text{ signal} = \beta_1 \times + \beta_2 \times + \beta_3 \times + \beta_4 \times + \beta_5 \times + \beta_6 \times \]

**Examples:**
- Stimulus
- Subject
- Run
- Trial
- Group (AD/Young)
- ROI

**Advantages of General Linear Model (GLM):**
- Can perform data analysis within and between subjects without the need to average the data itself
- Allows you to counterbalance random stimuli orders
- Allows you to exclude segments of runs with artifacts
- Can perform more sophisticated analyses (e.g., 2 factor ANOVA with interactions)
- Easier to work with (do one GLM vs. many T-tests and/or correlations)
GLM Parameter Estimates

- realignment & motion correction
- smoothing
- GLM model fitting statistic image
- smoothing kernel
- design matrix
- Brain Atlas – anatomical reference
- normalization
- Statistical Parametric Map
- corrected p-values random field theory

General Linear Model Approach

\[ Y = X \times \beta + \varepsilon \]

- Voxel timeseries data vector
- GLM design matrix
- parameters
- error vector

Example:
- Stimulus
- Subject
- Run
- Trial
- Group
- ROI
- Hand
- Hemi
- Tissue
Completeness and Efficiency of Signal Representation – Wavelet functions 1D

A 2D Daubechies wavelet illustrating the compact support, fast decaying and oscillatory properties of wavelets.

Completeness and Efficiency of Signal Representation – Wavelet functions 2D
Visual Wavelets

If Time Permits:

C:\Ivo.dir\Research\movies\WaveletMovie.mpg

Online At:
http://www.loni.ucla.edu/Software/
http://www.loni.ucla.edu/~dinov/WAIR.html

Completeness and Efficiency of Signal Representation – Wavelet functions 3D

Signal & it’s 3D WT
Completeness and Efficiency of Signal Representation – Wavelet functions 3D
Completeness and Efficiency of Signal Representation – Wavelet functions 3D

Wavelets

http://socr.stat.ucla.edu/

The three curves on this graph represent the original signal (*HeavySine* function, thin red curve), its wavelet transform (thin blue curve) and the reconstructed function estimator using only the largest 2% of the wavelet parameters. Note the space-frequency decorrelation of the original data in the wavelet-space (blue curve).
Wavelet-space Shrinkage

- Wavelet Denoising: \( y_i = f(x_i) + e_i, \ 0 \leq i \leq N-1, \ e_i \sim N(0, \sigma^2) \) IID. If \( W \) is the WT operator and \( d \) is a subset of \( \{ w_{j,k} \} \), \( w = Wy = Wf + We \) and \( \hat{f} = T_{SW}(y, d) = \sum_{(j,k) \in d} w_{j,k}W_{j,k}(x) \), Selected Wavelet reconstruction.

**Ideal Risk:** For \( f \) an L-piece-wise poly of degree \( D \):

\[
R_{N,\sigma}(SW, f) = \inf_d \{ R_{N,\sigma}(T_{SW}(y, d), f) \} = \\
\inf_d \left\{ \frac{1}{N} E \left( \sum_{i=0}^{N-1} (T_{SW}(y, d)(x_i) - f(x_i))^2 \right) \right\} = \\
\begin{cases} 
\frac{L(D+1)\sigma^2}{N}, & \text{known - breaks} \\
O\left( \frac{\ln N \sigma^2}{N} \right), & \text{unknown - breaks}
\end{cases}
\]

Is attainable?

---

Wavelet-space Shrinkage

- Spatial- vs. Frequency-Adaptive wavelet shrinkage

\( \eta_{j,k}(w_{j,k}) = sgn(w_{j,k})|w_{j,k} - \lambda_{j,k}| \), where \( \lambda_{j,k} = \begin{cases} \lambda_{j,k}^{DJ} = \frac{\sigma \sqrt{2 \ln N}}{N} & \sigma \sqrt{2 \ln (2n_j - 4)} \\ \lambda_{j,k}^{DS} = \sigma \sqrt{2 \ln (2n_j - 4)} & \end{cases} \)

If \( \hat{f}_{j,k}^{DJ} = \eta_{j,k}^{DJ}(w_{j,k}) \), \( \hat{f}_{j,k}^{DS} = \eta_{j,k}^{DS}(w_{j,k}) \) and \( \hat{f} = W^{-1}(\hat{\theta}(W(y))) \) then

\[
R(\hat{f}_{DJ}, f) = R(\hat{\theta}_{DJ}, \theta) \leq (1 + 2 \ln N) \left( \frac{\sigma^2}{N} + R_{N,\sigma}(SW, f) \right) \\
R(\hat{f}_{DS}, f) = R(\hat{\theta}_{DS}, \theta) \leq (1 + 2 \ln(N + 4)) \left( \frac{\ln N \sigma^2}{2N} + R_{N,\sigma}(SW, f) \right)
\]

in essence an optimal estimator.

- Optimal Function Estimators

If \( \hat{f}_{j,k}^{DJ} = \eta_{j,k}^{DJ}(w_{j,k}), \) \( \inf \sup_{\hat{\theta} \in \Theta} \frac{1}{N} E(\hat{\theta} - \theta)^2 \leq \frac{2 \ln N}{N} \)

If \( \hat{f}_{j,k}^{DS} = \eta_{j,k}^{DS}(w_{j,k}), \) \( \inf \sup_{\hat{\theta} \in \Theta} \frac{1}{N} E(\hat{\theta} - \theta)^2 \leq \frac{4 \ln N}{3 N} \)
Wavelet-space Shrinkage -
What function spaces are these estimates optimal for?

**MiniMax** attaining estimators over scale of function classes, \( \mathcal{F} \), Risk \( \text{Risk}(\mathcal{N}, \mathcal{F}) = \inf_{\mathcal{F}} \sup_{f \in \mathcal{F}} R(f^\wedge, f) \).

E.g., \( L^2 \) Sobolev spaces

\[
W^m_2(C) = \left\{ f : \| f \|_2^2 + \left\| \frac{d^m f}{dt^m} \right\|_2^2 \leq C^2 \right\}
\]

(parameter\(=m\)). More generally, for \( L^p \to W^m_p(C) \).

Virtually all of our data is represented by signals belonging to \( L^2 \) Sobolev spaces with \( m<4 \).

Wavelets & PDE’s

- **Procedure**: Use heat equation to numerically (de)blur an image. Iteratively, convolve the image with a smoothing kernel, then subsample (\( \downarrow 2 \)) and compute the difference, a Laplacian. The result is a faded version of the original signal. This pyramidal representation is **sparse but not contrast invariant** (due to the multiscale subsampling.)

Ref. Gabor, 1960, Morel, 2004
**Wavelets & PDE’s**

- Marr’s *edge detector* (1980’s) is to use **second derivative** to *locate* the maxima of the first derivative (which the edge contours pass through).

- *Haar Basis* (1909) encodes the edges into image representation via the first derivative operator (i.e. moving average/difference):

  $$(x_{2n},x_{2n+1},...) \xrightarrow{\text{Haar}} (a_n = \frac{x_{2n} + x_{2n+1}}{2}, d_n = \frac{x_{2n} - x_{2n+1}}{2},...)$$

**Wavelets & Engineering**

- $H$ is a low-pass filter
- $G$ is a high-pass filter
- $\downarrow$ is the down-sampling operator: $(1,3,4,6,5,8,7) \rightarrow (1,4,5,7)$
- $\uparrow$ is the up-sampling operator: $(1,4,5,7) \rightarrow (1,0,4,0,5,0,7)$

Orthogonal filter bank is biorthogonal, $\leftrightarrow$ both *analysis* filters $H’$ and $G’$ are the *time reversals* of the *synthesis* filters $H$ & $G$: e.g., $H=(1,2,3) \rightarrow H’=(3,2,1)$

Biorthogonal (perfect) filter bank: if $y = x$ for all inputs $x$. 

Wavelets & Engineering ➔ Pyramidal Algorithm

Alternating quadrature mirror – Smooth (S) and Detail (D) filters for Daub4 filterbank.

Data(y) Apply Elements Matrix: C\textsubscript{xy} Apply Elements Forward DWT
\[ \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{bmatrix} \rightarrow \begin{bmatrix} s_1 \\ d_1 \\ s_2 \\ d_2 \\ s_3 \\ d_3 \\ s_4 \\ d_4 \end{bmatrix} \rightarrow \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix} \rightarrow \begin{bmatrix} s_1 \\ d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix} \]

Smooth Components, Mother function coefficients
Details, wavelet coefficients

Wavelets & Fractals

Fractal quadtree-based transformation is equivalent to Haar-filter wavelet transform.

Design protocol for the wavelet-based construction and utilization of the fMRI functional atlas (F&A SVPA). The construction of the anatomical probabilistic atlas is augmented by including the regional distributions of the wavelet coefficients of the functional signals for one population. Statistical analysis of functional variability between a new fMRI volume and the F&A SVPA atlas is assessed following wavelet-space shrinkage.
Atlas–Based Wavelet–Space Stat Analysis – Demo

C:\Ivo.dir\LONI_Viz\LONI_Viz_VR_driver.bat (Sub-Sampling 2x2x2)
File ➔ Load Data
(E:\Ivo.dir\Research\Data.dir\FA_SVPA_WaveletBasedAtlas\stat_S1_HR VR_2 ADSVPA Multi10.img.gz)
File ➔ Load Mask
(E:\Ivo.dir\Research\Data.dir\FA_SVPA_WaveletBasedAtlas\WT_SVPA_20 lin h lobes30.img.gz)
Voxel (40,20,33)

Distribution of the wavelet coefficients

Displayed is the frequency histogram of the wavelet coefficients at 1,000 randomly selected locations (random indices of \( w_{jk} \)), averaged across all 578 MRI volumes part of the ICBM database [Mazziotta et al., 1995]. Notably, most of the wavelet coefficients are near the origin, with some having sporadic, but large magnitudes. Heavy-tail distributions models will be appropriate for these data.
Distr’n of the wavelet coefficients – 3 volumes

Shown here are the frequency distributions of the wavelet coefficients for three separate MRI volumes (randomly selected from the pool of 578). There is little variation between these three individuals, however the overall shape of the distribution of the magnitudes of the wavelet coefficients of each individual across the 1,000 random locations is more regular, still heavy-tailed, than the averaged distribution across subjects depicted in the previous Figure.

Wavelet Coefficient distributions MRI data

This image illustrates a portion of all of the Individual distributions of the wavelet parameters for the 578 ICBM MRI volumes. The horizontal axis represents the magnitude of the wavelet coefficients in the range [-4; 4], vertical axis labels the subject index and the row color map indicates the frequencies of occurrence of a wavelet coefficient of certain magnitude across all 1,000 voxels. Bright colors indicate high, and dark colors represent low, frequencies.
This diagram depicts the frequency histogram of the average magnitude of a wavelet coefficient, across all 128 fMRI 3D time-volumes (part of the fMRI study of Buckner and colleagues [Buckner et al., 2000]) at 1,000 randomly selected locations (random indices of $w_{j,k}$). Again, we observe the heavy-tailness of the data.

A 2D image displaying all of the individual distributions for each 128 3D time points (1 run, 15 trials) of an fMRI timeseries. On the horizontal axis is the magnitude of the wavelet and the vertical axis are the voxel locations.

Colors indicate the frequencies of occurrence of a wavelet coefficient of certain magnitude across all 1,000 voxel locations. The across row average is effectively shown on previous Figure.
Distribution of 4D wavelet coefficients

This graph shows the frequency distribution of the wavelet coefficients of the complete 4D WT of the fMRI timeseries. Note the slow asymptotic decay of the tails of this distribution (data-size: 64x64y16z128t, floating point). Here the x-axis represents the magnitude of the wavelet coefficients and the vertical axis indicates the frequency of occurrence of a wavelet coefficient of the given magnitude within the 4D dataset. Sporadic behavior of the large-magnitude wavelet coefficients is clearly visible at the tails of this empirical distribution.

Heavy-Tail wavelet coefficient distribution models …

**Leptokurtic Distribution Candidates:**

1. **Double-exponential** distribution
   \[ f(x) = \frac{1}{2\beta} \exp \left( -\frac{|x-\mu|}{\beta} \right) \]

2. **Double-Pareto** distribution
   \[ f(x) = \frac{\alpha \mu}{|x-\mu|^{\alpha+1}}, \quad |x-\mu| > c \]

3. **Cauchy** distribution
   \[ f(x) = \frac{\gamma}{\pi} \frac{1}{\gamma^2 + (x-\mu)^2} \]

4. **T** distribution
   \[ f(x) = \frac{\Gamma\left(\frac{1}{2}(n+1)\right)}{\sqrt{n\pi\Gamma\left(\frac{n}{2}\right)}} \times \left(1+x^2/n\right)^{-\frac{1}{2}(n+1)} \]

5. [http://socr.stat.ucla.edu/](http://socr.stat.ucla.edu/)
Distribution of (and Models for) the Wavelet Coefficients of the 4D fMRI volume

Several heavy-tail distribution models are fitted to the frequency histogram of 4D wavelet coefficients. These include double-exponential, Cauchy, T, double-Pareto and Bessel K form models. Because our statistical tests will be applied on the coefficients that survive wavelet shrinkage it is important to utilize a distribution model that provides accurate asymptotic approximation to the real data in the tail regions.

Right Tail of Distribution of (and Models for) the Wavelet Coefficients of the 4D fMRI volume

Shows the extreme left trail of the data distribution (data is symmetric). The double-exponential and the T distribution models underestimate the tails of the data asymptotically, but provide good fits around the mean. Cauchy, Bessel K forms and double-Pareto distribution models, in that order, provide increasingly heavier tails with Cauchy being the most likely candidate for the best fit to the observed wavelet coefficients across the entire range. The double-Pareto and the Bessel K form densities provide the heaviest tails, however, they are inadequate in the central range [-3 : 3] and undefined near the mean of zero.
Wavelet-domain Atlas-based fMRI Analysis

- Anatomical Sub-Volume Probabilistic Atlas (SVPA)

- Analysis using conventional time-domain SVT (sub-volume thresholding)

Subtraction paradigm: AD-ElderlyNC

Wavelet-domain Atlas-based fMRI Analysis

- Raw results following analysis in the wavelet-domain (has ringing effects outside of the true ROI where the activation supposedly occurred).
- Post-processed wavelet domain statistics [final uniform thresholding (top 1%), reconstruction of all wavelet-space stats into a single volume – regional calculations]:

Bayesian Mixture Models for fMRI Analysis

• We use two component mixture prior distributions on the wavelet coefficients $\theta_{j,k}$ with

$$\theta_{j,k} \mid \pi_j \tau_j \sim \pi_j N\left(0, \tau_j^2\right) + (1 - \pi_j) \delta(0)$$

where $\pi_j$ is a proportion between 0 and 1, $\delta(0)$ is the Dirac point mass at zero and $\tau_j > 0$. In other words, there is a level-dependent positive probability $\pi_j$ a priori that each wavelet coefficient will be exactly zero. If not, the coefficient will be normally distributed with mean zero and a level-specific standard deviation $\tau_j$.

Bayesian Mixture Models for fMRI Analysis

• Given the observed data over an ROI $y = f + \varepsilon$, the corresponding wavelet representation $w = \theta + z$, where $W$ is the discrete WT, $w = Wy$, $\theta = Wf$ and $z = W\varepsilon$, and the above prior distribution for the true wavelet coefficients, the posterior distributions of $\theta_{j,k}$ are again independent two-component mixtures:

$$p(\theta_{j,k} \mid w_{j,k}) \sim \lambda_{j,k} N\left(\frac{w_{j,k} \tau_j^2}{\sigma^2 + \tau_j^2}, \frac{\sigma^2 \tau_j^2}{\sigma^2 + \tau_j^2}\right) + (1 - \lambda_{j,k}) \delta(0)$$

Where the $\lambda_{j,k} = 1/(1 + \rho_{j,k})$ are the posterior odds that $\theta_{j,k}$ is exactly zero are:

$$\rho_{j,k} = \frac{1 - \pi_j}{\pi_j} \sqrt{\frac{\tau_j^2 + \sigma^2}{\sigma^2}} \exp\left(-\tau_j^2 w_{j,k} \frac{\sigma^2}{2 \sigma^2 (\tau_j^2 + \sigma^2)}\right)$$
LONI Resource Collaborators