Wavelet-based Classification Applied to fMRI

Pedro A. Morettin

Instituto de Matemática e Estatística
Universidade de São Paulo
pam@ime.usp.br
http://www.ime.usp.br/∼ pam

1. Classification Procedures

- In this work we consider several wavelet-based procedures for clustering (classification) purposes. In some situations, the time domain approach may not lead to clear classification or discrimination. When we move to the wavelet domain, the multiresolution analysis leads to look at data in several levels of resolution (or scales) and then the separation may become better.
- Among the wavelet-based procedures, we mention:

  (a) Multifractal Spectra (MFS) and associated descriptors. Jeon et al. (2014).
  (b) DWT-CEM procedure: discrete wavelet transform combined with classification expectation maximization algorithm. Sato et al. (2007).
  (c) DWT-Schur measures: discrete wavelet transform followed by the use of some Schur monotone measure. Shi et al. (2006)
  (d) Wavelet-based Bayesian discriminant function. Chang et al. (2003).

2. fMRI

- Functional neuroimaging is making an increasingly important contribution to experimental neuroscience. One of the most popular imaging method is fMRI, which infers changes in brain function from the temporal evolution of the BOLD (blood oxygen-level dependent) signal (Ogawa et al., 1990), an indirect measure of brain activity.
- The BOLD signal primarily corresponds to the concentration of deoxyhaemoglobin. In simple terms, the magnetic resonance signal comes from exciting hydrogen nuclei with a radiofrequency pulse, and detecting the radio waves emitted as the nuclei return to a lower-energy configuration. Deoxyhaemoglobin has different magnetic properties than oxyhaemoglobin– it is paramagnetic, which means that it will make the local magnetic field over a microscopic domain inhomogenous. This has the effect of dephasing the signal emitted by the nuclei in this domain, causing destructive interference in the observed MR signal.

2. fMRI

- For the purposes of estimating the BOLD signal in an experimental paradigm, SPM (Statistical Parametric Maping) makes use of a canonical haemodynamic response function (HRF). This function is assumed to be the response of the system (as reflected by the MR signal) to a brief, intense period of neural stimulation. The HRF exhibits a rise peaking around 6 sec, followed by an undershoot that persists for a considerable period.

3. The Discrete Wavelet Transform

- Wavelet analysis has been widely applied with success in signal processing and image analysis. One application of the wavelet transform is to yield a multiresolution analysis of a time series in which the latter is decomposed into a representation at different temporal detail levels or resolutions (Daubechies, 1992).
- A wavelet basis is generated by dilations and translations of a mother wavelet ψ, , i.e., if $\mathbb{Z}$ denotes the set of integers,

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k), \;\; j,k \in \mathbb{Z}.$$

The discrete wavelet transform of data $(X_0, \ldots, X_{T-1})$ is defined by

$$d_{j,k} = \sum_{t=0}^{T-1} X_t \psi_{j,k}(t).$$

4. One special case: DWT-CEM Algorithm in fMRI

Considering fMRI time series analysis, we propose the following clustering algorithm (see Figure 1):

Step 1: Extract the wavelet coefficients of the detail level in the scale(s) of interest.

Step 2: Apply the CEM algorithm to the wavelets coefficients in the scale of interest.

Step 3: Extract the average or representative time series corresponding to each cluster. These time series describe the cluster BOLD signal and they can be used to identify clusters of interest, i.e those related to the experimental paradigm. This last step is similar to the strategy used in ICA to identify independent components of interest.

5. DWT-CEM Algorithm in fMRI

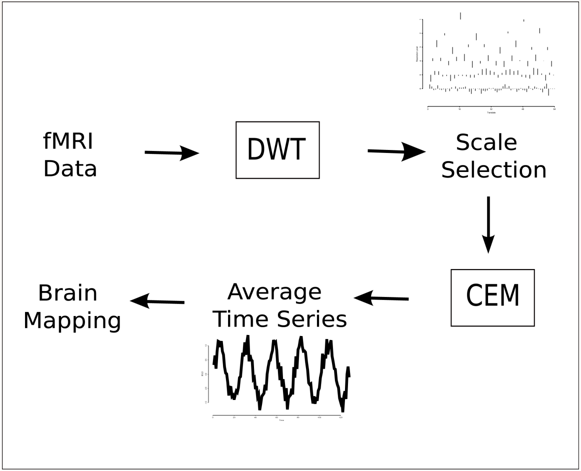

Figure 1: Wavelets clustering algorithm for fMRI.

6. Simulation

- For all simulations we used the D16 wavelet in the DWT. Different types of noise (white, AR(1) and long memory) were added to a simulated HRF (T=128 timepoints) consisting of a linear combination of two Poisson functions with peaks at 4 and 8 seconds replicated 6 times in a block design (see Figure 2, top).
- The AR(1) parameter is 0.8 and for the long memory time series fractional integrated white noise (d=0.4) was considered. The histogram, kernel estimates and Gaussian theoretical density show that the Gaussiam assumption is reasonable.

6. Simulation

- Furthermore, the performance of the combination between wavelet transform and CEM was evaluated considering three simulated situations. These were (a) two out of phase responses (denoted HRF-1 and HRF-2, T=128 timepoints), (b) a linear combination of two Poisson functions with peak in 4 and 8 replicated 6 times in a block design fashion (see Figure 3 for illustrative examples) and (c) no haemodynamic response.
- Considering that only a small percentage of voxels in fMRI datasets normally respond to any given stimulus, the simulated data was composed by 4096 time series, 128 (3.125%) using HRF-1, 128 (3.125%) using HRF-2 and 3840 (93.75%) of no response. Gaussian white noise was than added to these curves (SNR=0.5 and 1, SNR here indicating the ratio of signal to noise).

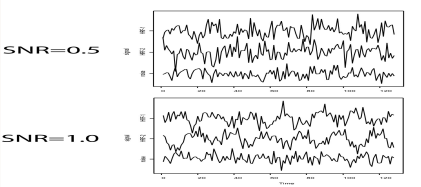

Figure 2: Illustrative simulated time series for HRF-1, HRF-2 and white noise.

- The DWT-CEM algorithm was then applied to the simulated data (200 simulations).
- In order to compare the CEM algorithm with other clustering methods, we also applied k-means and Fuzzy C-Means (FCM) (Jahanian, 2004, Baumgartner, 2001) to the wavelet-transformed simulated data. For each simulation, the maximum number of clusters was selected using the BIC.
- The results for classification mean accuracy are presented in Figure 4. The simulations suggest a good performance of DWT-CEM algorithm, and that it has advantages over k-means and FCM particularly in cases of low SNR.

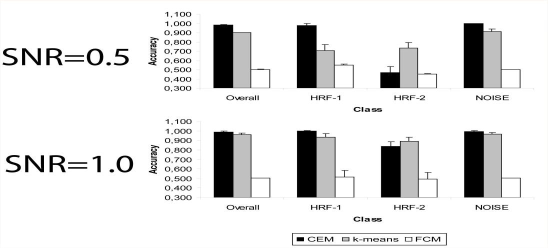

Figure 3: Classification mean accuracy for simulated data (200 simulations). The error bars describe two standard errors.

- **Visual and Auditory Experiment**: DWT-CEM was applied to a simple visual-auditory stimulation fMRI data, for which HRF model and expected brain activated regions are well established.
- Four voluntary normal healthy subjects were scanned under visual and auditory stimulation. All the images were collected at Hospital das Clínicas, University of São Paulo using a 1.5 Tesla GE Signa scanner (TR=2s, TE=40ms, 24 slices oriented to AC/PC line), 128 volumes acquired.

- The visual stimulus consisted of an AB periodic block design (block duration= 24s), alternating an 8Hz flashing checkerboard stimulus with a fixation cross in the centre of an average gray level background, with 6 cycles.
- The auditory stimulus was delivered in the same run, based in AB periodic block design (block duration= 36s, 4 cycles), alternating between silence and passive listening to words via MR compatible headphones (background scanner noise was present in both conditions).
- In summary, the subject had visual and auditory stimulation out of phase, both presented in block design with different cycle durations.

7. Application to a fMRI Dataset

- The images were pre-processed by realignment to minimize the effects of subject motion, slice-time correction and spatial smoothing. GLM activation maps in individual native space were obtained using the software XBAM and they are shown in Figures 6 and 7.
- Activation maps are built by a regression, for each voxel, of the form

$$S_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + e_t,$$

where the regressors X_{1t} (auditory) and X_{2t} (visual) are vectors containing zeros (no stimulation) or ones (stimulation), convolved with a HRF. The maps simply show if the β_i are significant or not.

7. Application to a fMRI Dataset

- For the cluster analysis, the time series in all voxels were normalized to zero mean and unit standard deviation. This procedure is important in order to prevent clustering based on structural image features (reflecting the average image intensity at each voxel) rather than the intended clustering of BOLD responses.
- The DWT of the experiment design was computed in order to find the decomposition level with the largest mean absolute value of the wavelet coefficients (scale of interest). The detail levels identified were the third and fourth, for visual and auditory experiments respectively (the zero level is the finest scale), resulting in 24 wavelet coefficients (predictors) to be considered in CEM analysis.

7. Application to a fMRI Dataset

- Considering the Bayesian Information Criterion, 12 clusters were obtained in the whole volume. Thus, these clusters are based on temporal similarity and removing noise influences. The clusters in individual native space are presented in Figure 8. The results of DWT-CEM analysis evidence a clear pattern of clusters in both auditory and visual primary areas, which is consistent across subjects and similar to GLM results presented in Figure 5 and 6.
- Furthermore, some clusters are also identified in areas where artifacts are commonly described, but this pattern is not recurrent across subjects. These results are obtained without any prior specification about the HRF, only the wavelet scales of stimulation are informed. The time series of visual and auditory cortical clusters are presented in Figures 7 and 8..

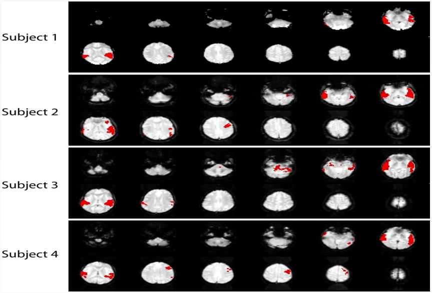

Figure 4: Auditory task brain activation maps obtained using the GLM.

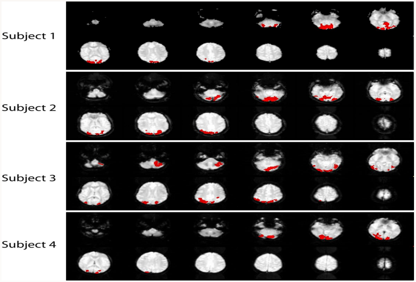

Figure 5 : Visual task brain activation maps obtained using the GLM.

7. Application to a fMRI Dataset

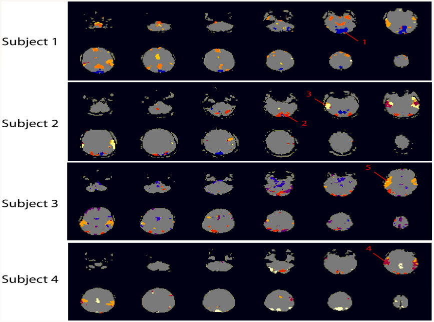

Figure 6: : DWT-CEM analysis. The time series of clusters indicated by the numbers are presented in Figure 9.

7. Application to a fMRI Dataset

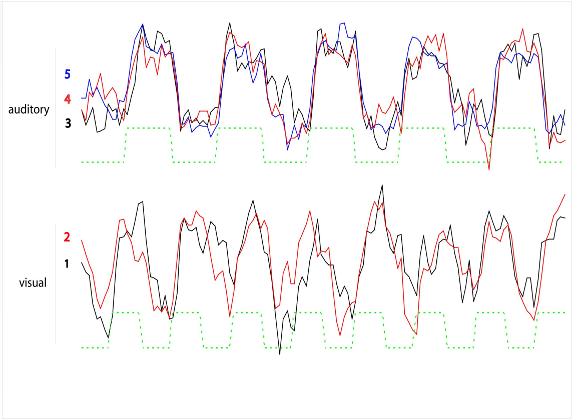

Figure 9: BOLD signals of clusters at visual and auditory cortex. The green dotted lines describe the respective stimulation.

8. Conclusions

- GLM analysis is a very popular approach in brain activation mapping. However, temporal clustering may be also an interesting tool, providing and alternative way to identify activation foci. Temporal clustering, unlike most GLM approaches is also naturally multivariate.
- In this work, we propose a wavelet cluster analysis method (DWT-CEM), which automatically selects the optimum number of clusters and works well in low SNR. The clusters are based on temporal similarity of signals, indicating possible neural modules or networks. Furthermore, the analysis suggests functional regions of interest, which could be objective candidates for subsequent ROI - based analysis.

8. Conclusions

- DWT-CEM can identify auditory and visual responses in real fMRI data but without the necessity for a detailed prior specification of an "expected" HRF or the total number of clusters. Providing the response is contained within the chosen scale(s) of interest, DWT-CEM will also be insensitive to phase shifts of the response and to variations of response amplitude between blocks. In addition to "real" experimental responses, clustered artifactual components can also be identified in the chosen scale.
- Our main aim of applying DWT-CEM to fMRI, however, is not the simple identification of activated areas, a replacement for standard GLM methodology, but the unsupervised grouping voxels according to their temporal similarities at scales of interest. Additionally, the property to separate different clusters in the same functionally identified region represents an avenue of possible use for wavelet clustering, based on discriminating different wave shapes.

9. References

Celeux, G. and Govaert, G. (1992). A classication EM algorithm for clustering and two stochastic versions. *Computational Statististics and Data Analysis*, **14**, 315–332.

Chang, W., Kim, S.-H. and Vidakovic, B. (2003). Wavelet-based estimation of a discriminant function. *Applied Stochastic Models in Business and Industry*, **19**, 185–198.

Daubechies, I. (1992). *Ten Lectures on Wavelets*. SIAM, Philadelphia.

Fujita, A., Sato, J.R., Garay-Malpartida, H.M., Morettin, P.A., Sogayar, M.C. and Ferreira, C.E. (2007). Time-varying modelling of gene expression regulatory networks using the wavelet dynamic vector autoregressive method. *Bioinformatics*, **23**, 1623–1630.

9. References

Harris, K.D., Henze, D.A., Csicsvari, J., Hirase, H. and Buzsaki, G. (2000) Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *Journal of Neurophysiology*, **84**, 401–414.

Hastie, T., Tibshirani, R. and Friedman, J. (2001). *The Elements of Statistical Learning*. Springer.

Jahanian, H., Hossein-Zadeh, G.A., Soltanian-Zadeh, H. and Ardekani, B.A. (2004). Controlling the false positive rate in fuzzy clustering using randominzation: application to fMRI activation detection. *Magnetic Resonance Imaging*, **22**, 631–638.

Jeon, S., Nicolis, O. and Vidakovic, B. (2014). Mammogram diagnostics via 2-D complex wavelet-based self-similarities measures. *São Paulo Journal of Mathematical Sciences*, **8**, 265–284.

Kohonen T (1995). Self-Organizing Maps. Berlim:Springer.

McQueen J (1967). Some methods for classification and analysis of multivariate observations. Proc. Fifth Berkeley Symposium on Math. Stat. and Prob. 1:281-296.

Ogawa, S. et al. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings National Academy Sciences USA*, **87**, 9868–72.

Sato, J.R., Fujita, A., Amaro Jr, E., Mourão Miranda, J., Morettin, P.A. and Brammer, M.J. (2007). DWT-CEM: An algorithm for scale-temporal clustering in fMRI. *Biological Cybernetics*, **97**, 33–45.

Shi, B., Moloney, K.P., Pan, Y., Leonard, V.K., Vidakovic, B., Jacko, J.A. and Sainfort, F. (2006). Wavelet classification of high-frequency pupillary responses. *Journal of Statistical Computation and Simulation*, **76**, 431–446.

Vidakovic, B. (1999). *Statistical Modeling by Wavelets*. Wiley.