Bayesian Hierarchical Modelling of Clustered Cerebral BOLD Images

OHRI IRSO
Ottawa Health Research Institute

C. Gómez-Laberge[1,2], A. Adler[2], I. Cameron[3], and M.J. Hogan[1]

1. Neuroscience Program, Ottawa Health Research Institute, Ottawa
2. Systems and Computer Engineering, Carleton University, Ottawa
3. Dept. Diagnostic Imaging, Ottawa Hospital, Ottawa

Carleton University

Introduction

- Cerebral BOLD fMRI may be possible without imposing a model for the hemodynamic function
- First, the image data set can be partitioned into a set of *clusters* of time-sequence correlated voxels
- Second, from each cluster, a set of *explanatory variables* can be measured
- Third, a *hierarchical model* can be fit to the data set by using all cluster measurements
- Once fit, the model can then be used to estimate the probability of chance-occurance of each observed cluster (Tablet 1)
- This poster describes a novel method for fMRI data analysis defining meaningful explanatory variables and applying a Bayesian hierarchical model for cluster analysis and selection

Methods

- 25 data sets from 6 healthy volunteers (28-55 yrs)
- Visually cued event-related hand motor task
- 1.5 T Siemens Magnetom SE-EPI pulse sequence
- Data sets analyzed with fuzzy c-means clustering[1]
- Clusters formed based on member-centroid correlation (Tablet 2)

Cluster Selection Criteria (Tablet 3)

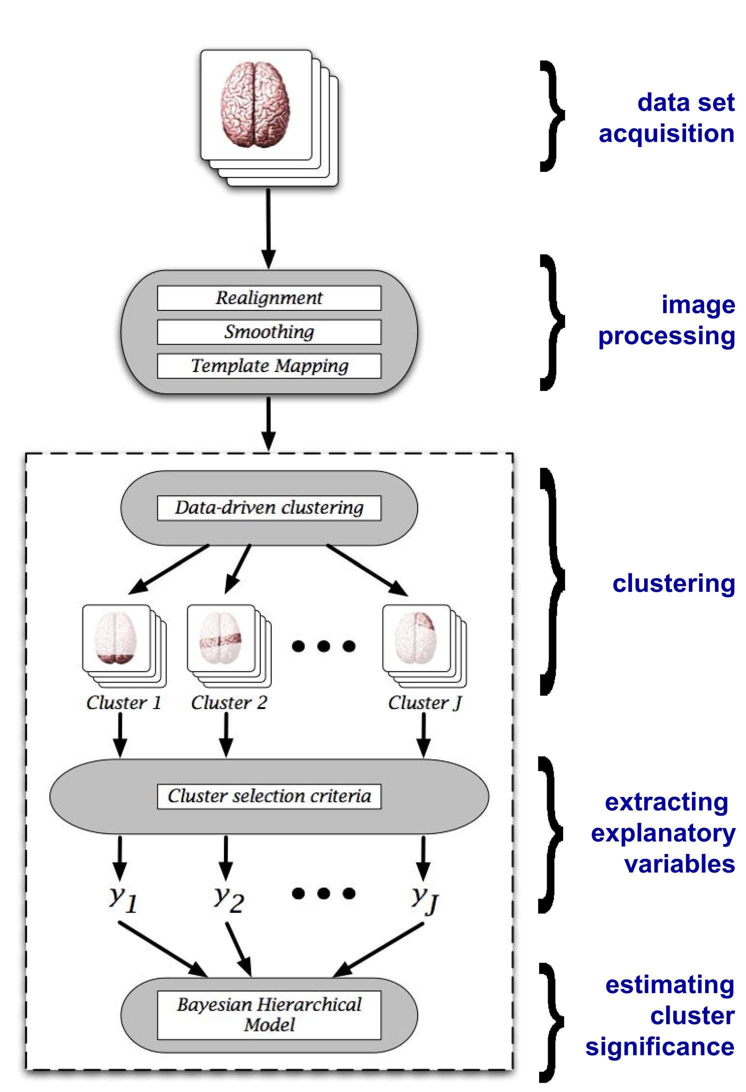
- 4 explanatory variables from selection criteria[2]:
- *Centroid-paradigm cross-correlation* provides the maximum Pearson correlation r of the cluster allowing for a time delay d
- *Cluster voxel-map contiguity* provides the relative "compactedness" c of cluster voxels as a function of a membership correlation threshold s

Bayesian Hierarchical Model (Tablet 4)

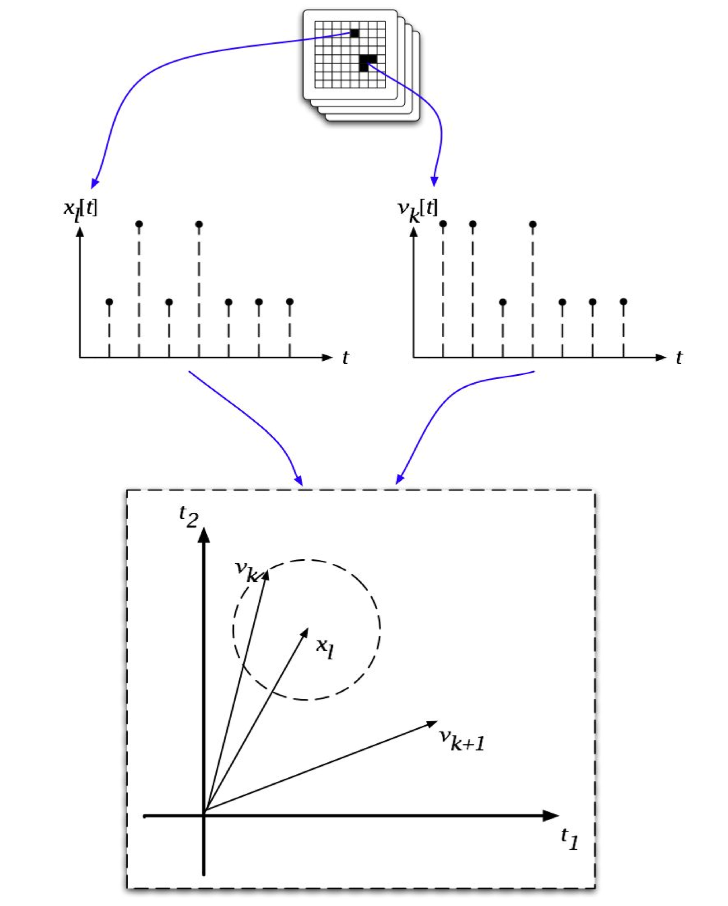
- Model variables from cluster j as $y_j = (r_j, d_j, c_j, s_j)^T$ a realization from a stochastic system of 3 levels:
  1) observable cluster data y, drawn from
  2) underlying processes β, affected in turn by
  3) an overall effect μ of subject & instrument
- Hierarchical model[3] has normal likelihood $y_j|\beta_j$, conjugate normal population $\beta_j|\mu,\tau$ and non-informative μ,τ prior densities, leading to the *conditional posterior density*

$$p(\beta|\,\mu,\tau,\,y) \propto p(\beta|\,\mu,\tau)\;p(y|\,\beta)$$

- Simulate parameters using MCMC Gibbs sampler



TABLET 1: method stages

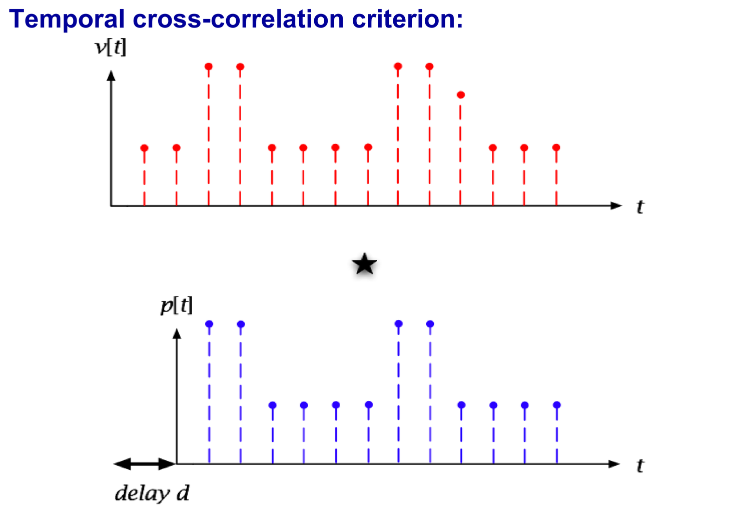


Distance metric based on Pearson product-moment correlation (PPMC)

$$r(x,v) = \frac{\sum_{t=1}^{T}(x[t]-\bar{x})(v[t]-\bar{v})}{(T-1)s_x s_v}$$

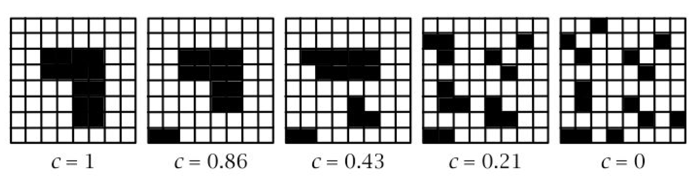
TABLET 2: data-driven clustering

Temporal cross-correlation criterion:



$$(v \star p)[d] = \sum_{t=1}^{T} v[t+d-T]p[t], \quad d = 1,\ldots,2T-1$$

Spatial contiguity criterion:

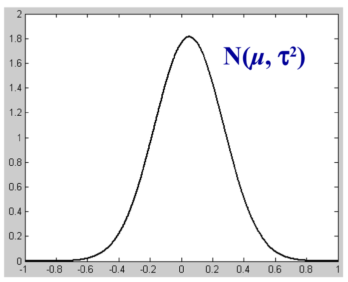


$$c = \frac{1}{GL}\sum_{l=m}^{M} l g_l$$

For example, the third image has contiguity:

$$c = (1\times3+1\times9)/(2\times14) = 0.43$$

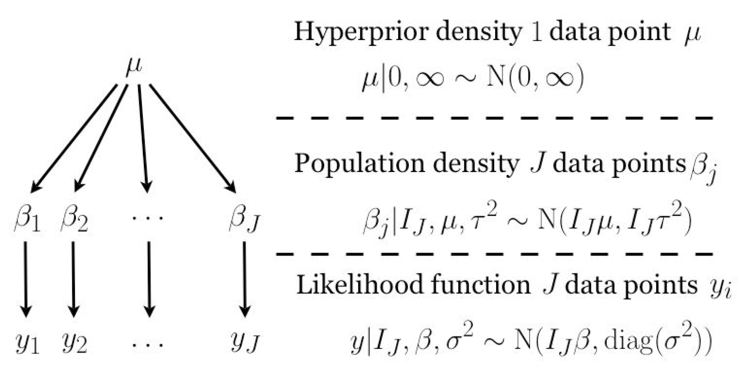
TABLET 3: explanatory variables



Null hypothesis: All cluster variables y were drawn from the some underlying normal distribution with unknown parameters μ, τ^2

Given y, can one calculate the probability of generating similar outcomes?

Hierarchical layers modelling y as realizations of an stochastic process—each stage is a random variable:

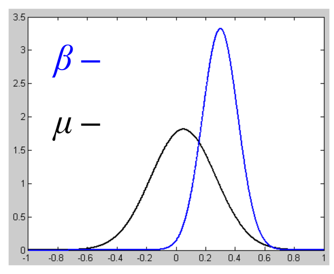


The joint posterior distribution for this model:

$$p(\beta,\mu,\tau|y) \propto p(\mu,\tau)p(\beta|\mu,\tau)p(y|\beta)$$

$$= p(\mu,\tau)\prod_{j=1}^{J}\mathrm{N}(\beta_j|\mu,\tau^2)\prod_{j=1}^{J}\mathrm{N}(y_j|\beta_j,\sigma_j^2)$$

This is a J+2 dimensional function that is best sampled numerically using Markov Chain Monte Carlo methods



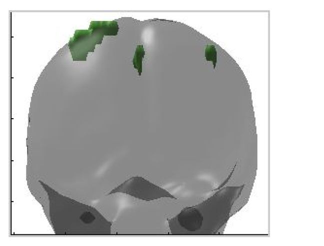
Now with the generated samples for β, μ, τ conditioned on y, one can estimate the probability $\Pr(\beta = \mu \mid y)$

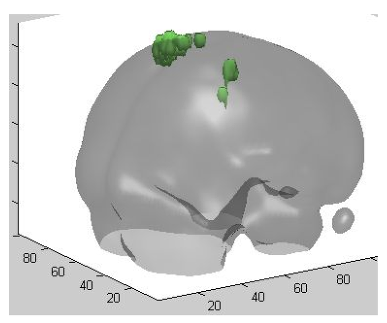
TABLET 4: hierarchical model

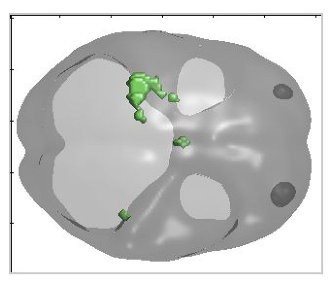
BAYESIAN CLUSTER ANALYSIS: 002-01-GIST01 20070521 Fourth swr

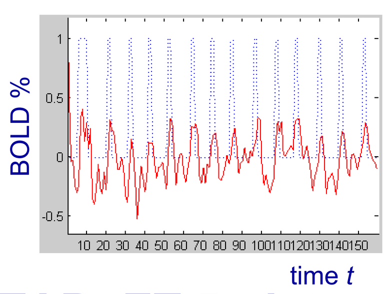
Cluster	Posterior Quantiles					Significance		Convergence
j	2.5%	25%	50%	75%	97.5%	Pr(ß=μ|y)	Pr(ß=y)	Rhat
1	-0.37	-0.15	-0.03	+0.08	+0.28	0.274	0.73	1.00
2	-0.25	-0.12	-0.06	+0.01	+0.13	0.138	0.61	1.00
3	-0.07	+0.03	+0.08	+0.13	+0.23	0.460	0.50	1.00
4	-0.27	-0.19	-0.15	-0.11	-0.04	0.013	0.61	1.00
5	-0.29	-0.20	-0.15	-0.10	-0.02	0.015	0.65	1.00
6	-0.32	-0.17	-0.09	-0.01	+0.12	0.101	0.66	1.00
7	+0.35	+0.42	+0.46	+0.50	+0.58	0.000	0.68	1.00
8	+0.00	+0.09	+0.14	+0.19	+0.29	0.240	0.55	1.00
9	+0.04	+0.21	+0.30	+0.39	+0.58	0.049	0.78	1.00
10	-0.02	+0.08	+0.14	+0.19	+0.30	0.278	0.56	1.00
11	-0.06	+0.07	+0.14	+0.21	+0.33	0.295	0.57	1.00
mu	-0.10	+0.02	+0.07	+0.12	+0.23	N/A	N/A	1.00
tau	0.14	0.20	0.23	0.28	0.43	N/A	N/A	1.00

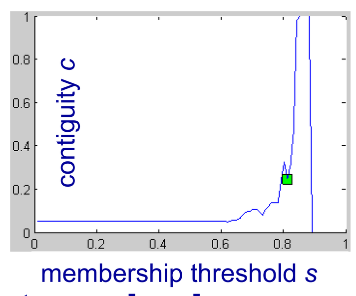
Cluster 7 has 260 voxels
Explanatory variables:
r = 0.490; d = 2.0 s
c = 0.249; s = 0.81
Pr($\beta = \mu \mid y$) < 0.0002





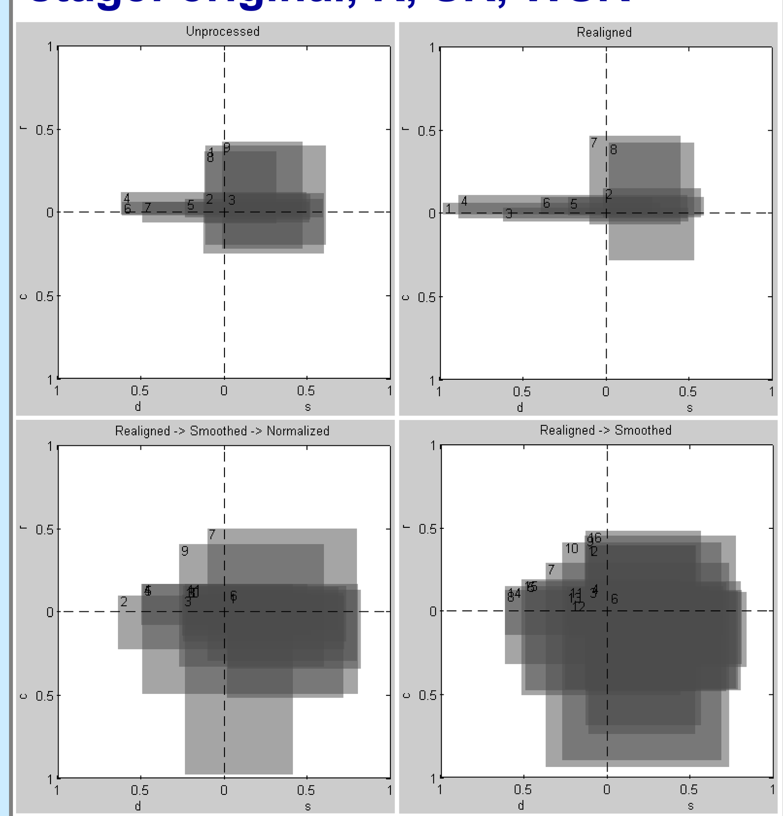






TABLET 5: data set analysis

Cluster footprints from one data set for each processing stage: original, R, SR, WSR



TABLET 6: cluster footprints

Results

- We computed conditional posterior probability of cluster realization $\Pr(\beta = \mu \mid y) < 0.05$ and compared results against selection with threshold $|r| > 0.30$

Efficacy of Hierarchical Model (Tablet 5)

- Model identified relevant cerebral territories of sensorimotor cortex and cerebellum in all data sets in agreement with threshold selection
- Clusters selected by model have mean $|r| = 0.30 \pm 0.12$, while rejected have $|r| = 0.05 \pm 0.22$
- Model rejects highly correlated clusters caused by movement because of additional requirement of measured sample variance σ_j^2 misfit with τ^2

Effect of Data Preprocessing (Tablet 6)

- Exploratory results show largest performance benefit obtained from (in decreasing order) normalization, realignment, and smoothing
- Effects can be seen by examining magnitude of cluster vector $y' \in [0,1]^4$ onto planar projections referred to as cluster "footprint"
- The clustering algorithm used was less effective on unprocessed data, producing few clusters from several sources; sensorimotor cortex was identified in roughly 40% of the data sets

Conclusion

- The proposed model fits cerebral BOLD data from healthy subjects and agrees with both threshold-based cluster selection and standard ANOVA

Research in progress

- Multivariate hierarchical regression analysis
- Gaussian random field analysis of contiguity

Brief References:

- Jarmasz, M *et al.* (2002). Exploring regions of interest with cluster analysis. *Artif. Intell. Med.*, 25, 45-67.
- Gómez-Laberge, C *et al.* (2008). Selection criteria for the analysis of data-driven clusters in cerebral fMRI. *IEEE T. Biomed. Eng.*, in press: TBME-00102-2008.
- Gelman, A *et al.* (2004). *Bayesian Data Analysis*, 2nd ed., Boca Raton, USA: Chapman & Hall/CRC press.

Acknowledgements: This research was supported by the Behavioural Research and Imaging Network in partnership with the Ontario Research Fund, and by the Heart and Stroke Foundation Centre for Stroke Recovery