Finding Cortical Patches of Shared Representations

a comparison of clustering algorithms on representational geometries, and the effect of cross-validation to reduce physiological noise

Matteo Visconti di Oleggio Castello[1], Samuel A. Nastase[1], M. Ida Gobbini[1, 2], James V. Haxby[1, 3], Yaroslav O. Halchenko[1]

[1]Dept. of Psychological and Brain Sciences, Dartmouth College, Hanover (NH), USA; [2]Dept. of Medicina Specialistica, Diagnostica e Sperimentale (DIMES), Med. Sch., Univ. of Bologna, Italy; [3]Ctr. for Mind/Brain Sci., Univ. of Trento, Italy

SFN 2015 Poster #451.11

http://www.pymvpa.org http://Neuro.Debian.net

CCN Center for Cognitive Neuroscience at Dartmouth

Aims

- Apply clustering methods to searchlight RSA to functionally parcellate the cerebral cortex
- Investigate reproducibility for different clustering algorithms
- Compare experiment-specific parcellations to publicly available functional and anatomical parcellations
- Investigate the effect of different cross-validation schemes to reduce the effect of physiological noise [1]

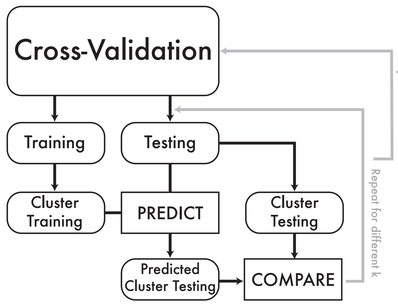
Methods

Clustering algorithms

- k-means
- Gaussian Mixture Models with diagonal covariance structure
- Ward with structural constraints

Metrics to compare cluster solutions

- Reproducibility procedure (inspired by [2] and [3], see figure below) using metrics
  - Adjusted Rand Index (ARI) [4]
  - Adjusted Mutual Information (AMI) [4]
  - Instability [2]
- Consistency of representational geometry quantified using a measure of homogeneity [5] based on pairwise correlation distance between al searchlight RDMs within a parcel



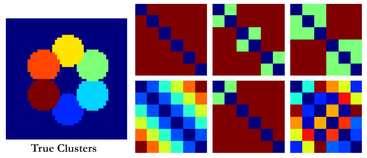
Reproducibility Procedure [2, 3]

Cross-validation schemes

- Across subjects with RDM estimated across runs
- Across subjects and runs with one RDM estimated within each run
- Across runs to compute RDMs (asymmetrical), then cross-validation across subjects

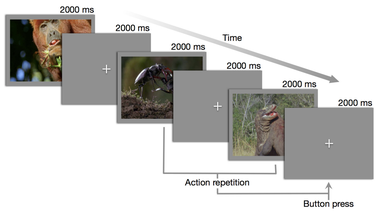
Datasets

Simulated



- 10 simulated subjects
- 1 or 8 runs
- 6 contiguous clusters with pattern information
- Different RDM in each cluster
- Random subject- and run-specific noise to simulate different noise sources

Real fMRI data



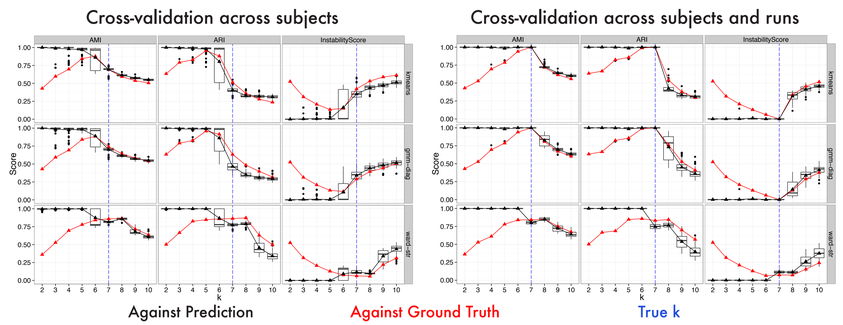
- 12 participants watching 2 s naturalistic video clips of behaving animals [6]
- 5 animal taxa, 4 actions (20 total conditions)
- 1-back task requiring attention to action categories
- 5 runs
- 20 x 20 RDM computed with surface-based searchlights (100 voxels) using correlation distance
- Anatomical alignment and whole-brain hyperalignment [7]

Code and contact information

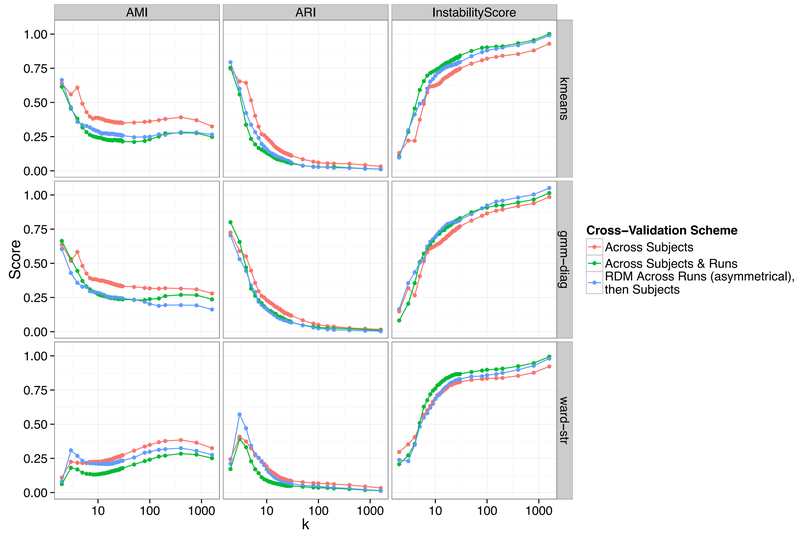
Code available at **www.github.com/mvdoc/reprclust** (work in progress)

Contact me at **matteo.visconti.gr@dartmouth.edu - www.mvdoc.me**

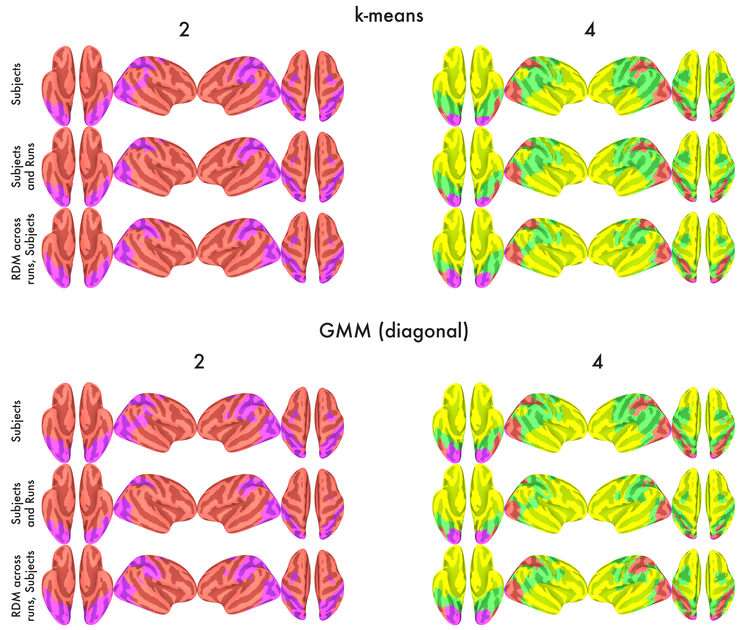
Results on simulated data



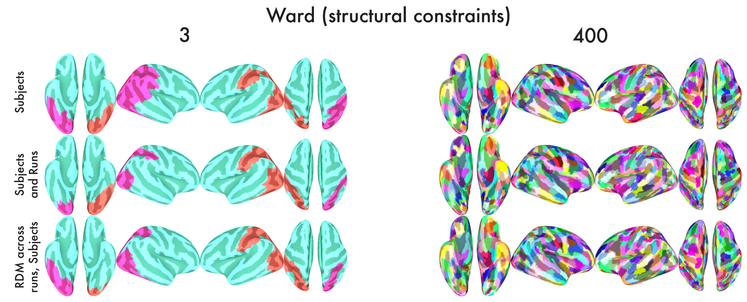
Results on real fMRI data



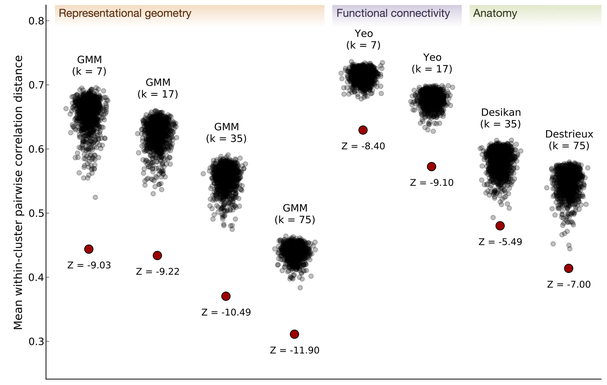
Cluster solutions on full datasets according to different RDM estimations



Results on real fMRI data (continued)



Homogeneity [5]



- Within each parcel, average pairwise correlation distance between all searchlight RDMs
- Null distribution of homogeneities estimated by randomly rotating the spherical projection of the cortical surface
- Comparison with parcellations derived from anatomy (FreeSurfer) and resting-state functional connectivity [2]

Conclusions

- Meaningful parcellations can be obtained by clustering shared representational geometries
- None of the methods tested provided "the ultimate solution" for whole brain parcellation according to reproducibility, and different metrics can provide contrasting results
- Experiment-specific parcellations exhibit higher homogeneity compared to resting-state and anatomical parcellations
- Reproducibility estimates and cluster solutions remained stable with different cross-validation schemes in the dataset we analyzed

Future Directions

- Remove uninformative and unstable parcels
  - Clustering is performed on task-related data, so some nodes might lack useful representational geometry information
  - This would yield higher instability at greater values of k when the nodes are clustered across folds
- Compare to thresholding data before running stability procedure to keep only informative nodes

References

1. Henriksson, L et al (2015). Visual representations are dominated by intrinsic fluctuations correlated between areas. NeuroImage.
2. Lange, T et al. (2004). Stability-based validation of clustering solutions. Neural Computation.
3. Yeo, B et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of Neurophysiology.
4. Thirion, B et al. (2014). Which fMRI clustering gives good brain parcellations? Frontiers in Neuroscience.
5. Gordon, E et al. (2014). Generation and evaluation of a cortical area parcellation from resting-state correlations. Cerebral Cortex.
6. Nastase, SA et al. (2015). Attention reshapes distributed population codes for semantic representation. Under review.
7. Haxby, JV et al. (2011). A common, high-dimensional model of the representational space in human ventral temporal cortex. Neuron.