Inter-subject hyperalignment of neural representational space

J. Swaroop Guntupalli and James V. Haxby *Dept. of Psychological & Brain Sciences, Dartmouth College, Hanover, NH, USA*

262.20

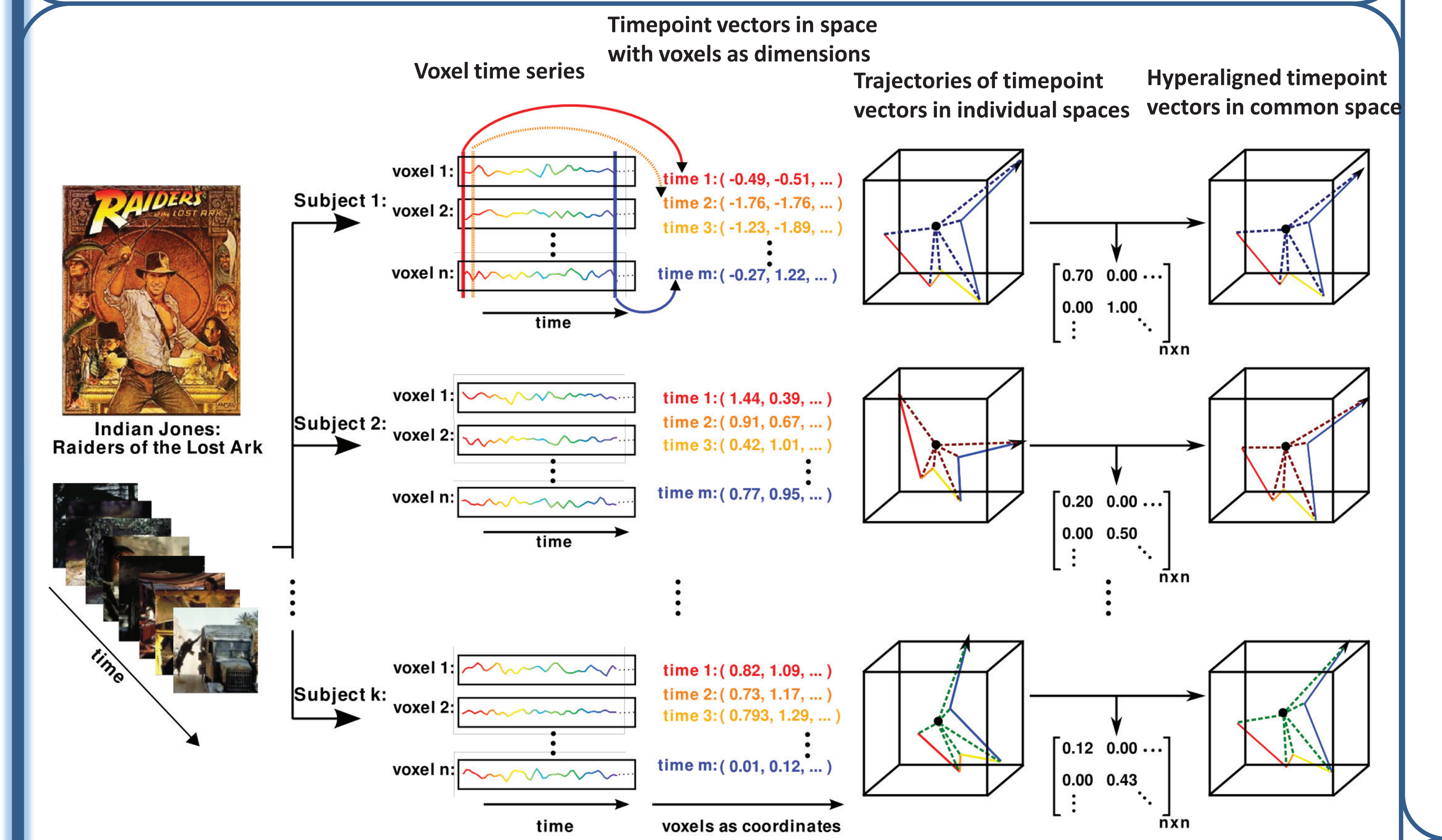
INTRODUCTION

Why

- The information content of brain activity patterns can be decoded using multivariate pattern (MVP) analysis[1].
- Since cross-subject registration of brain anatomy does not align fine-scale structure, MVP analysis is generally done by building a new classifier/model for each subject, limiting the cross-subject validity of the model.

What

- Hyperalignment solves this problem by aligning the neural representational spaces of different subjects into a common space.
- It generalizes across brains, across experimental designs, and across a wide range of stimuli, opening doors for building functional brain atlases.



Schematic figure of analysis of movie data to derive hyperalignment parameters

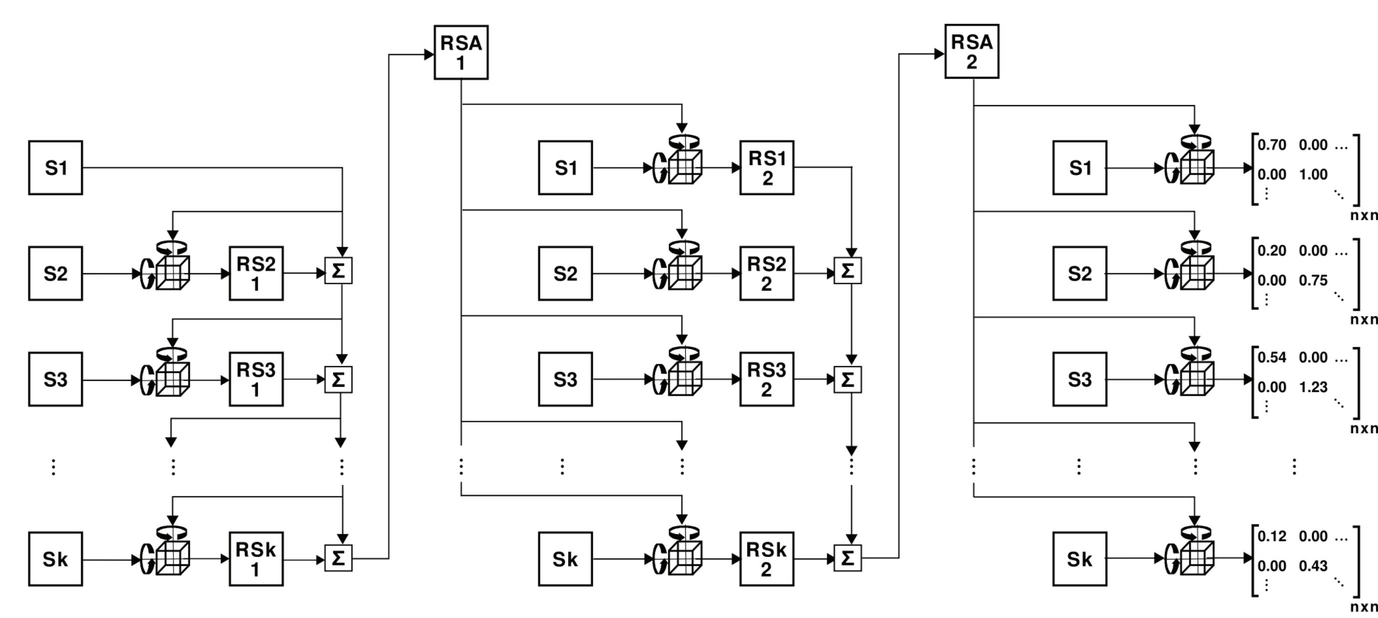
Methods – Imaging

- Ten healthy young subjects participated in two fMRI studies:
  - On one day, they watched the movie Raiders of the Lost Ark.
  - On a different day they viewed static pictures of four categories of faces (human females, human males, monkeys, and dogs) and three categories of objects (houses, chairs, and shoes).
- We used voxels from anatomically defined ventral temporal (VT) cortex and selected a fixed number of top-ranking VT voxels from each subject's left and right hemispheres based on their between-subject time series correlation.

Methods - Hyperalignment

- Response patterns of voxels during the movie form a trajectory of time-point vectors in a high-dimensional subject-specific space with voxels as dimensions.
- Our method assumes that these trajectories from different subjects reflect the same visual information[2], but the coordinate systems of their representational spaces are poorly aligned.
- Hyperalignment uses Procrustean transformation to align one subject's trajectory of time-point vectors to another subject's trajectory. It works by deriving an orthogonal matrix with global scaling factor that minimizes the sum of squared Euclidean distances between two sets of paired vectors.
- We derived a common representational space after aligning the coordinate systems of subjects and calculated Procrustean transformation parameters for each subject that align that subject's data to this common space. These are the hyperalignment parameters for that subject.

S-> Subject's original data; RS-> Subject's Rotated data; RSA-> Average of all subjects' rotated data



Block diagram of the algorithm for deriving a common space and parameters using pairwise Procrustean transformations

Analyses & Results

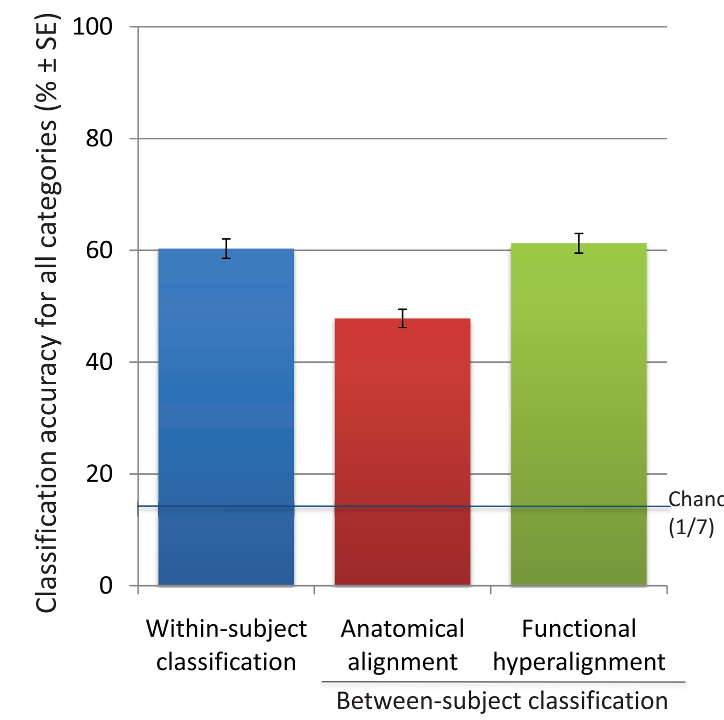
In our analysis we tested the following characteristics of hyperalignment:

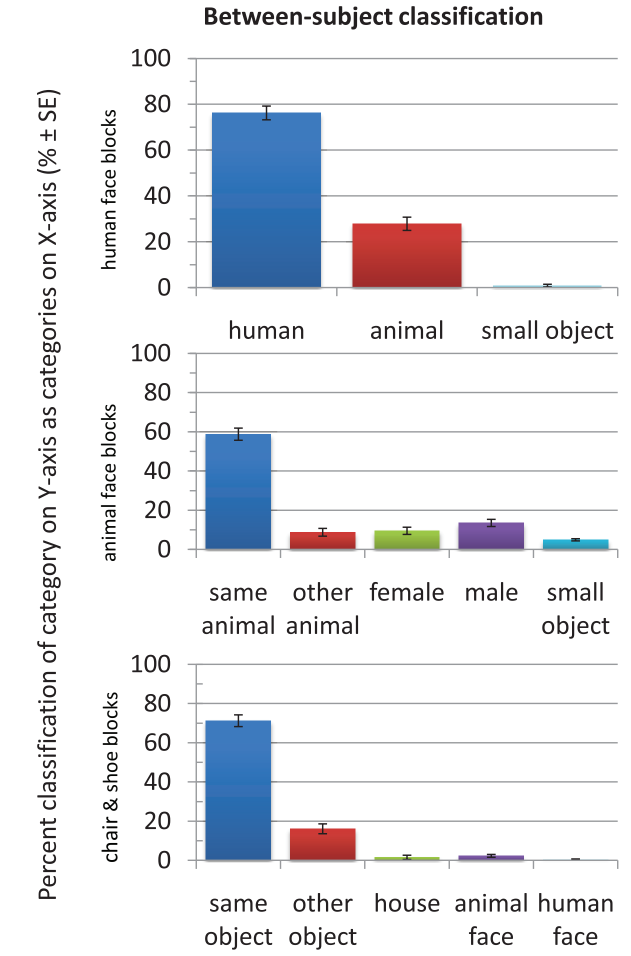
I. Generalization across experiments
We derived the hyperalignment parameters using full movie data and applied them to the faces and objects study data for all subjects. We classified the categories of blocks from each subject based on models from other subjects' data.

II. Significance of rich stimulus set with variety of conditions
We derived the hyperalignment parameters from one session of the movie study and applied the parameters to data from the other session. We classified 30 second time segments in the other half of the movie for each subject using other subjects' data. We repeated the classification analysis using hyperalignment parameters derived from faces and objects study time series data.

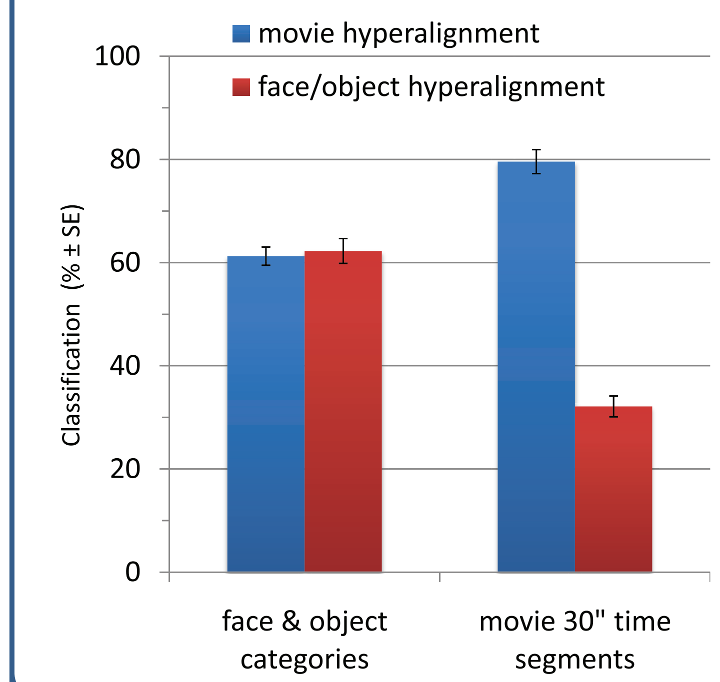
III. Alignment of fine-scale features
We repeated analysis (I) on voxels from FFA alone and PPA alone.

Analysis I: Classification of faces and objects categories in VT cortex

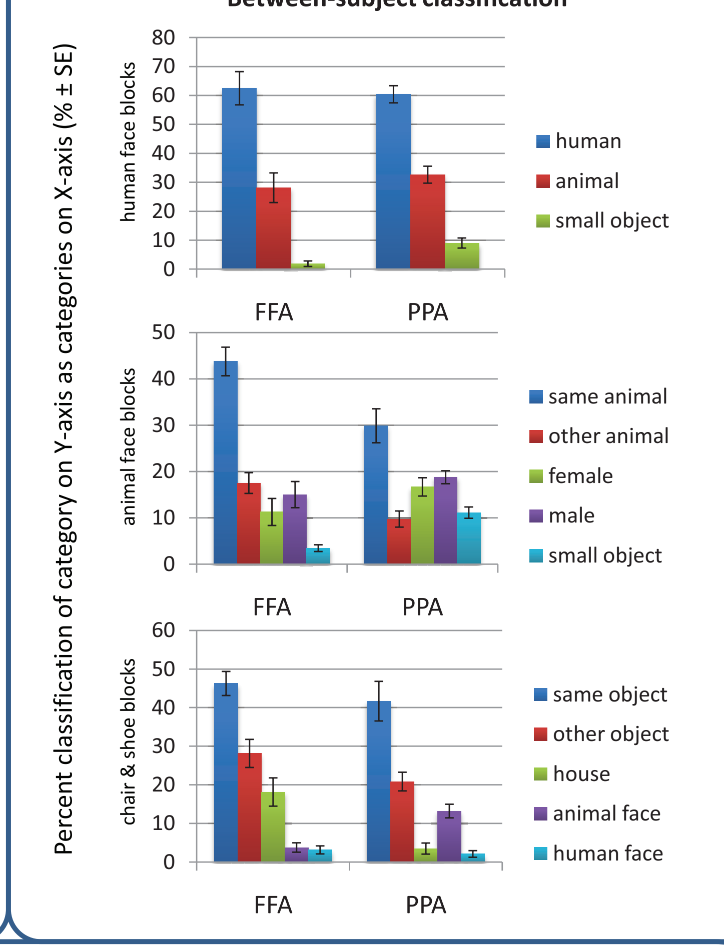




Analysis II: Movie time-segment classification in VT cortex



Analysis III: Analysis I in FFA & PPA



CONCLUSIONS

- Hyperalignment derives a common high-dimensional space for fine-scale neural representations and parameters for transforming subject-specific neural representational spaces into this space.
- Hyperalignment allows the brain states of one individual to be classified based on a model derived from other subjects' data.
- The common representational space models neural representations in individual brains at least as good as individually-tailored models, indicating that neural representations for complex stimuli in different brains have a higher level of commonality than anticipated.

References

1. Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* **293**, 2425-2430 (2001).
2. Sabuncu, M.R., Singer, B.D., Conroy, B., Bryan, R.E., Ramadge, P.J., Haxby J.V. Function-based intersubject alignment of human cortical anatomy. *Cereb. Cortex*, in press.