Brain Network Decomposition for Naturalistic Stimulus Paradigm

Yijun Liu[1], Jian Li[2,3,4], Jessica L. Wisnowski[5,6], Anand A. Joshi[1], Richard M. Leahy[1]

[1] Signal and Image Processing Institute, University of Southern California, Los Angeles, CA
[2] A. A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA
[3] Department of Radiology, Harvard Medical School, Boston, MA
[4] Department of Neurology, Massachusetts General Hospital, Boston, MA
[5] Radiology and Pediatrics, Division of Neonatology, Children's Hospital Los Angeles, Los Angeles, CA
[6] Keck School of Medicine, University of Southern California, Los Angeles, CA

Introduction

Functional MRI acquired using naturalistic stimulus paradigms, such as movie-watching, incorporates dynamic and diverse sensory information. It has been shown to be more ecologically valid than traditional task-based fMRI and produce brain responses with higher stability than resting-state fMRI[1]. Existing tools for analyzing this type of fMRI data often assume perfect temporal synchronization between subjects, which, however, may not be valid due to differing responses and/or latency to stimuli across different subjects. Popular methods such as intersubject correlation yield limited information about brain response patterns. Although the commonly used independent component analysis[5] can discover interesting structure, it imposes independence constraint on either the spatial or temporal domain, which may not be physiologically realistic. To address these issues, we applied a combination of a temporal synchronization technique (BrainSync Alignment[2,3]) and a tensor decomposition method (NASCAR[4]) to movie-watching data. The results showed that our method can provide rich information about the population's common brain responses to the naturalistic stimuli, yet with a parsimonious model.

Methods and Materials

The minimally preprocessed 7T movie-watching fMRI data of 110 subjects from the Human Connectome Project were used[7,8]. Continuous fMRI data were acquired during 4 sessions while subjects watched different audio-visual movies interleaved with rest periods. Each session ran ~15min (TR=1s). The fMRI data were resampled onto the cortical surface and co-registered to a common surface atlas. Each scan was represented as a $V \times T$ matrix ($V \approx 22K$ is the number of vertices across the two hemispheres, $T \approx 900$ is the number of time points). For each movie, we applied BrainSync Alignment (BSA) to jointly synchronize fMRI data across all subjects, Fig. 1a. For Movie 1, we formulated a tensor χ of size $V \times T \times S$ by concatenating the synchronized fMRI data along the

3rd dimension. NASCAR was then applied to approximate χ as a sum of 20 rank-1 tensors each representing a distinct brain network that is composed of a spatial activation map, a temporal dynamics component, and a subject participation level. We also extracted 20 components using the spatial ICA[7] (sICA) for comparison. We used the temporal mode of the auditory network (Fig 1b) from Movie 1 to predict soundtrack loudness of other movies by transferring the auditory temporal response from the group atlas G1 to each subjects' space in other sessions through a cascaded (invertible) BSA (Fig. 1a). The individual subjects' responses were then averaged together as the final prediction. Loudness measures, used as ground truths, were extracted using MIRtoolbox[9] and convolved with a double gamma HRF. For sICA, we applied a pairwise BrainSync between two sessions of each subject to predict the responses, which were also averaged across all subjects.

Results

Fig. 1b shows the spatial maps of two networks identified using BSA+NASCAR and sICA. Both methods capture activations in the auditory and visual cortex. Since the latter enforces spatial independences, spatial overlaps between the networks can be underestimated as shown in Fig. 1b. Fig 1c quantitatively confirms this observation by comparing the inter-network cross-correlation matrices between the two methods. Fig 2a shows the normalized sum of squared reconstruction error as a function of the number of identified networks. The errors are consistently lower using our method than that with sICA. Predictions of the loudness using the auditory temporal dynamics from BSA+NASCAR are also better correlated with loudness measures, for all 3 movies (Fig. 2b, c).

Conclusions

Our framework decomposes naturalistic fMRI data into brain networks under more physiologically valid assumptions, resulting in better stimulus neural encoding-decoding and better fit to the data than ICA.

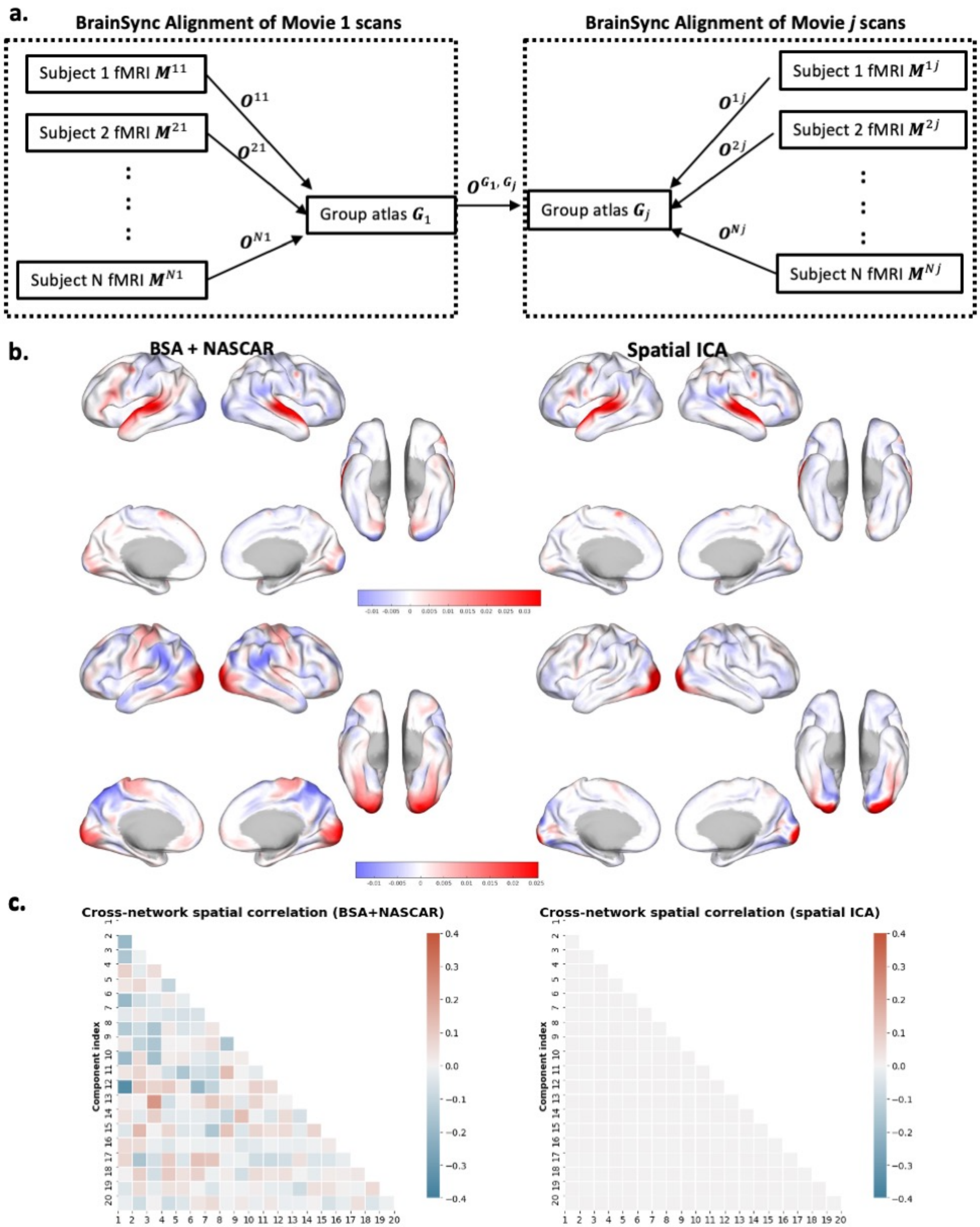
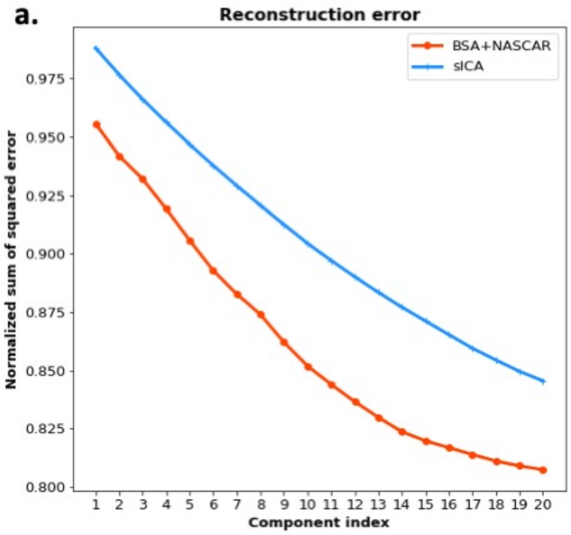


Figure 1 (a) Diagram of BrainSync Alignment (BSA) applied to the scans of two movies. For movie j, each subject i has a specific optimal transform $\boldsymbol{O}^{ij}$ which aligns the subject to the atlas. $\boldsymbol{O}^{G_1,G_j}$ is the orthogonal transform from $\boldsymbol{G}_1$ to $\boldsymbol{G}_j$ with BrainSync (b) Decomposed spatial maps containing activations in auditory cortex (row 1) and visual cortex (row 2) using BSA+NASCAR and spatial ICA (c) Correlation matrices of spatial maps from BSA+NASCAR and spatial ICA

a.



b.

Table. Pearson correlation of predicted loudness time-series with the ground truth

Movie	Ours	Spatial ICA	Gaussian noise
2	0.20	0.183	<0.01
3	0.498	0.487	
4	0.497	0.402	

c.

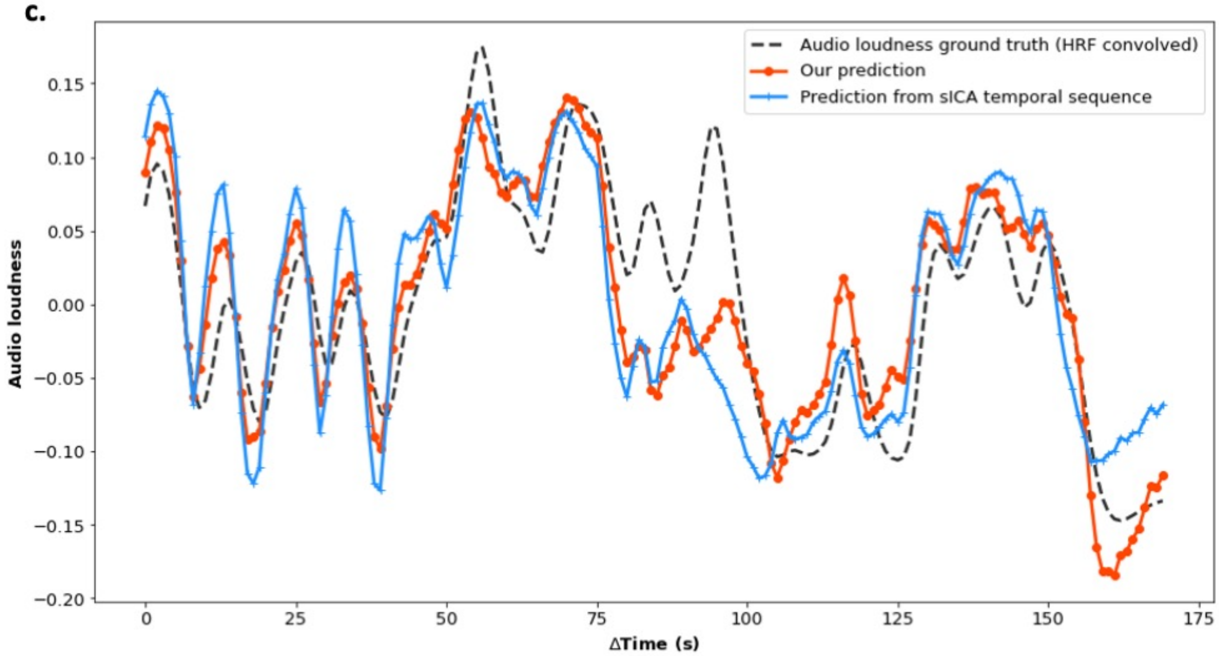


Figure 2 (a) Reconstruction errors using our method and spatial ICA (b) Pearson correlation of predicted loudness time-series with the ground truth using the temporal representations from our approach, sICA and Gaussian noise (c) Segments of ground truth audio loudness, our prediction, and prediction from sICA for Movie 4. Sequences are standardized for visualization purposes.

References

1. Sonkusare, S., et al. (2019). Naturalistic stimuli in neuroscience: critically acclaimed. *Trends in cognitive sciences*, *23*(8), 699-714. doi: 10.1016/j.tics.2019.05.004

2. Akrami, H. et al. (2019). Group-wise alignment of resting fMRI in space and time. In *Medical Imaging 2019: Image Processing* (Vol. 10949, p. 109492W. doi: 10.1117/12.2512564

3. Joshi, A. A., et al. (2018). Are you thinking what I'm thinking? Synchronization of resting fMRI time-series across subjects. *NeuroImage*, *172*, 740-752. doi: 10.1016/j.neuroimage.2018.01.058

4. Li, J., et al. (2021). Robust brain network identification from multi-subject asynchronous fMRI data. *NeuroImage*, *227*, 117615. doi: 10.1016/j.neuroimage.2020.117615

5. Varoquaux, G., et al. (2010). A group model for stable multi-subject ICA on fMRI datasets. *Neuroimage*, *51*(1), 288-299. doi: 10.1016/j.neuroimage.2010.02.010

6. https://nilearn.github.io/

7. Van Essen, D. C., et al. (2012). The Human Connectome Project: a data acquisition perspective. *Neuroimage*, *62*(4), 2222-2231. doi: 10.1016/j.neuroimage.2012.02.018

8. Glasser, M. F., et al. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage*, *80*, 105-124. doi: 10.1016/j.neuroimage.2013.04.127

9. Lartillot O., et al. (2008) "A Matlab Toolbox for Music Information Retrieval", *Data Analysis, Machine Learning and Applications*, Springer, doi: 10.1007/978-3-540-78246-9_31