

Hierarchical nonlinear embedding reveals brain states and performance differences during working memory tasks

Siyuan Gao (siyuan.gao@yale.edu)

Department of Biomedical Engineering, Yale University,
10 Hillhouse Avenue, New Haven, CT

Gal Mishne (gal.mishne@yale.edu)

Department of Mathematics, New Haven, CT
51 Prospect Street, New Haven, CT

Dustin Scheinost (dustin.scheinost@yale.edu)

Department of Radiology and Biomedical Imaging, Yale School of Medicine, New Haven, CT
300 Cedar Street, New Haven, CT

Abstract

Investigation of the moment-to-moment changes in brain activity using functional magnetic resonance imaging (fMRI) is an emerging field. However, one of the major problems is how to represent and evaluate these temporal relationships from the high-dimensional fMRI data. While many linear approaches have been proposed, nonlinear dimensionality reduction approaches may offer better solutions for these high-dimensional data. In this work, we propose a hierarchical, dimensionality reduction framework for time-synchronized fMRI data based on diffusion maps—a type of nonlinear embedding—labeled 2-step diffusion maps (2sDM). For evaluation, we apply the framework to task fMRI during a working memory task for two large, independent datasets. By applying the embedding on the time domain, we show that our framework can detect brain states as defined by task blocks. By applying the embedding on the subjects domain, we show that subjects can be separated by their working memory performance. Together, these results show the promise of 2sDM as a nonlinear embedding framework for fMRI data.

Keywords: fMRI; diffusion maps; brain states; clustering

Introduction

Recent studies of functional magnetic resonance imaging (fMRI) are beginning to quantify moment-to-moment changes in brain activation or connectivity (Allen et al., 2014; Monti et al., 2017; Lindquist, Xu, Nebel, & Caffo, 2014). A main goal of these works is to find representative brain states—or distinct, repeatable patterns of brain activity or connectivity—as a way of quantifying these brain dynamics. Focusing on a few specific states operationalizes the characterization of brain dynamics into computational tractable problems. However, due to the high-dimensional nature of these brain patterns, assigning time points to specific brain states or even estimating the number of brain states remain unsolved problems.

While previous works have used a range of supervised or unsupervised methods to define brain states, linear dimensionality reduction approaches, like principal component analysis (PCA), are the most widely used (Allen et al., 2014;

Monti et al., 2017). For example in (Monti et al., 2017), the brain states estimated from PCA were moderately correlated with the underlying task, providing evidence that the observed brain states represented some underlying neurobiology. However, these patterns are insufficient to classify between several task states. Classifying multiple states associated with different, yet related tasks, would further provide evidence that any observed brain state represents a neurobiological process rather than simply artifacts or confounds in the data.

In this paper, we propose a hierarchical, nonlinear dimensionality reduction framework based on diffusion maps for time-synchronized fMRI data. In contrast to linear embedding methods (e.g. PCA), nonlinear methods (e.g. diffusion maps) focus on discovering the underlying manifold structure of the data by integrating local similarities at different scales. By focusing on local similarities rather than global similarities, diffusion maps can better capture complex information in high-dimensional data. We initially validate our approach on task fMRI in two large open-source datasets. Similar to (Monti et al., 2017), task fMRI is used instead of resting-state in order to have different task blocks as explicit brain states in the data. Overall, our new framework provides a novel way to detect underlying brain states in order to quantify moment-to-moment changes in the fMRI signal.

Methods

Diffusion maps

Diffusion maps (Coifman & Lafon, 2006) is part of a broad class of nonlinear dimensionality reduction algorithms; specifically, diffusion maps gives a global description of the data by considering only local similarities and is robust to noise perturbation. One of the capabilities of diffusion maps is to reveal underlying intrinsic states governing the data. As such, these approaches should be suitable to detect repeatable brain states in fMRI data. The diffusion maps algorithm is as follows. The input is the similarity matrix S between vectors of data, which can be computed using various kernels like the Gaussian kernel or the cosine similarity. From the similarity matrix, the rows are normalized by $P = D^{-1}S$, where $D_{ii} = \sum_j S_{ij}$. Eigendecomposition is then performed on

this row-stochastic matrix, resulting in the eigenvalues λ_i and eigenvectors ψ_i . The corresponding diffusion maps is then $\Psi_t(x) = (\lambda_1^t \psi_1(x), \lambda_2^t \psi_2(x), \dots, \lambda_k^t \psi_k(x))$, where t is the diffusion time. In practice, eigenvalues of P typically exhibit a spectral gap such that the first few eigenvalues are close to one with all additional eigenvalues much smaller than one. In this case, the diffusion distance, which is the Euclidean distance between points in the embedding space, can then be well approximated by only these first few eigenvectors (Nadler, Lafon, Coifman, & Kevrekidis, 2006). Thus, we obtain a low-dimensional representation of the data by considering only the first few eigenvectors of the diffusion maps. Intuitively, diffusion maps embeds the data points closer when it is hard for the data to escape the local region within time t .

Diffusion maps is similar to the normalized cuts algorithm (Shi & Malik, 2000) which has previously been used in fMRI analysis (Shen, Tokoglu, Papademetris, & Constable, 2013). Normalized cuts aims to find the eigendecomposition of $D^{-1}L$ where L is the Laplacian matrix $L = D - S$. The eigendecomposition of $D^{-1}L$ yields the same eigenvectors ψ as for diffusion maps, with corresponding eigenvalues $1 - \lambda$. Thus, performing k -means clustering on diffusion maps coordinates as we do below is mathematically similar to spectral clustering. The key difference is that in diffusion maps the coordinates are weighted by the corresponding eigenvalues.

2-step diffusion maps

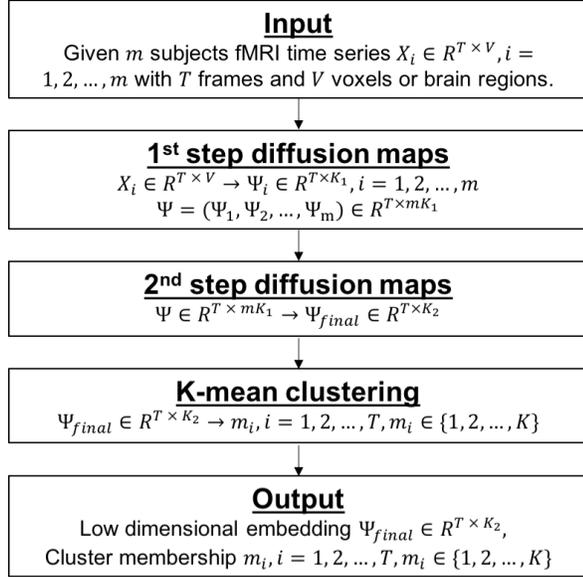


Figure 1: Framework of 2sDM applied to the time domain.

Based on diffusion maps, we design a hierarchical dimensionality reduction framework for multi-subject fMRI BOLD time series. Under the assumption that subjects' fMRI response is time-synchronized, we represent fMRI time series data as $X_i \in \mathbb{R}^{T \times V}$, $i = 1, \dots, m$. Here T is the number of frames in the scan, V is the number of voxels or brain regions, and m is the total number of subjects we have. We

label this framework 2-step diffusion maps (2sDM). Note that 2sDM can be applied to either domains of the data, resulting in a lower-dimensional representation of either time, subjects or brain regions. Here we illustrate the framework by applying the reduction on the time domain. Reducing the other two domains just requires trivial adaptation.

To perform 2sDM to the time domain, first, we apply diffusion maps dimensionality reduction on every single subject X_i in every time frame, reducing each subject's V voxels or brain regions to a K_1 -dimensional Euclidean space. Then, we concatenate all the subjects' new representation to a single matrix $X_c \in \mathbb{R}^{T \times (mK_1)}$. Next, we perform a second-step diffusion maps to further reduce the dimensions of this matrix to a vector of length K_2 . As a result, the final time frame representation matrix is $X_r \in \mathbb{R}^{T \times K_2}$.

The first-step diffusion maps produces a cleaner representation of the fMRI data, as diffusion maps reduces noise. The reasoning of performing an embedding based on the results of the first-step embedding is that the Euclidean distance in the diffusion coordinates approximates the diffusion distance. As such, if two concatenated vectors have relatively small Euclidean distance, it suggests that for all of the subjects there is small diffusion distance between the two time frames. Our framework is similar to related work on integrating information from multiple sensors by applying multiple embeddings (Rabin & Averbuch, 2010). However, our framework keeps the magnitude of diffusion coordinates while the previous approach discards this magnitude by normalizing the coordinates. As the magnitude carries information about the diffusion process, retaining this information should increase the utility of the second diffusion map.

Experimental setup

We assess the performance of 2sDM using working memory tasks on two independent fMRI datasets from the Human Connectome Project (Van Essen et al., 2013) and Philadelphia Neurodevelopmental Cohort (PNC) (Satterthwaite et al., 2016). From the HCP dataset, subjects executed interleaved blocks of 0-back and 2-back working memory tasks. Each subject had two scans corresponding to the left-right (LR) and right-left (RL) phase encoding direction where the task block orders were different for the two scans. 515 subjects from the HCP dataset were retained after removing subjects for high motion or incomplete data. For the PNC dataset, subjects executed interleaved blocks of 0-back, 1-back and 2-back working memory tasks. 571 subjects were retained after removing subjects for high motion or incomplete data. For both datasets, 268 timecourses of fMRI data were extracted using a whole-brain, functional atlas (Shen et al., 2013).

Results

Temporal embedding: HCP dataset

As our framework does not rely on the temporal structure, we first concatenate time series of LR and RL acquisition. While this increases the number of time points to be used in the em-

bedding, concatenation requires our algorithm to be robust to noise differences across the different acquisitions. We first examine the first three coordinates of the embedding in 3D space (Fig. 2). Four directions are clearly revealed in the embedding. Fixation, 0-back and 2-back task each takes one direction. The last direction consists of points from both 0-back and 2-back task blocks. To further assess the embedding, we perform k -means clustering using only the first three embedding coordinates. Fixation, 0-back and 2-back form individual clusters (Fig. 3) as expected from the embedding shape (Fig. 2). Further, the embedding direction consisting of 0-back and 2-back task blocks can be explained from the clustering results. Cluster 3 forms after the task cue occurs, which corresponds to when the subjects are viewing the task cues. The task block order is visualized as the background color of Figure 3. Timing of the task block is delayed by 5 seconds to account for the lagging in BOLD signals.

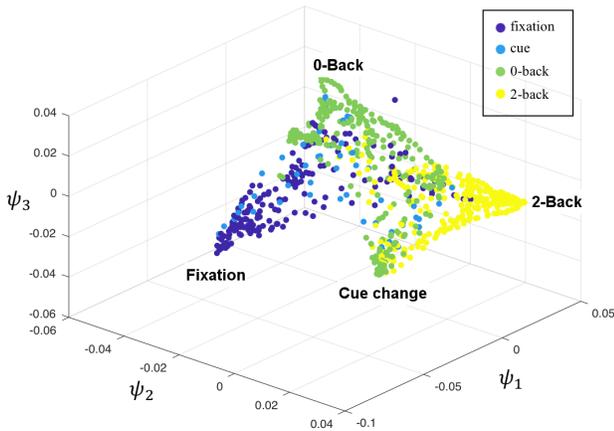


Figure 2: 2sDM embedding of the HCP dataset. Points are embedded in the 3D space by the first three non-trivial coordinates of diffusion maps. Each point represents a time frame and the point is colored by the task block type.

Although our 2sDM can reveal overall task patterns by incorporating data from all subjects, individual differences in brain states during a working memory task likely exist. For example, some subjects may exhibit larger shifts of their brain states during task blocks, and some subjects may exhibit a longer delay in entering the state associated with specific task blocks. To investigate these individual differences, we selected a subsample from the HCP subjects consisting of the 50 subjects with the lowest working memory accuracy and 50 subjects with the highest working memory accuracy. Using only these subjects, we recalculated our 2sDM. As shown in Fig. 3b, the patterns of the states for the best performers are more consistent with the task whereas for the worst performers they are more perturbed, suggesting distinct patterns of brain states for the best and the worst performing subjects.

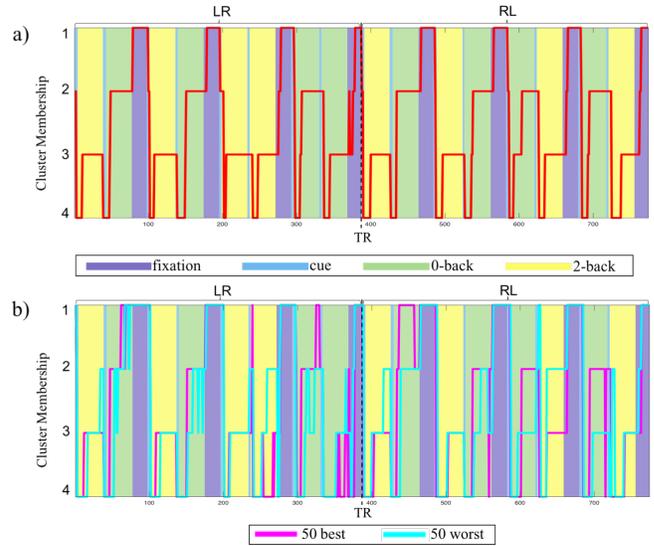


Figure 3: K -means clustering result on embedded coordinates of HCP time frames combining LR and RL acquisition. Cluster 1 corresponds to fixation, cluster 2 corresponds to 0-back, cluster 3 corresponds to 2-back, cluster 4 corresponds to cue change. *a)* shows the clustering result using all the subjects' information. *b)* shows the clustering result using only the 50 subjects with the best or worst working memory accuracy.

Subjects embedding: HCP dataset

Using the same subset of subjects (i.e. the 50 subjects with the worst working memory accuracy and the 50 subjects with the best working memory accuracy), we performed embedding across subjects, rather than across time. In contrast to the temporal embedding where information across different subjects is gathered together, subject embedding gathers temporal information to project subjects with similar temporal patterns of brain states closer together in the embedding. In the embedded 2D space (Fig. 4), subjects with the best working memory performance (yellow points) are clearly projected together; subjects with the worst performance, especially those with less than 70% accuracy, are mainly projected on the other side of the space. From the embedding, it can be observed that the best performers share similar and steady brain states. In contrast, the worst performers exhibit inconsistent brain states patterns, as displayed to more largely spread across subjects in the embedding.

Subjects embedding: PNC dataset

Similar to the HCP dataset, we selected the 50 best and 50 worst performing subjects from the PNC dataset for subject embedding. In contrast to the HCP dataset, for the PNC dataset, we used out-of-scanner performance on a working memory task as accuracy during the actual fMRI task was not available. While the exact grouping of subjects (e.g. the 50 best performing subjects) may be different if in-scanner accuracy was used, we expect that the groupings would be similar as working memory performances are generally highly repeat-

able. The 2D embedding (Fig. 4b) shows similar patterns as the HCP dataset. The best performing subjects are clustered together; while, the worst performing subjects are scattered as outliers around the major cluster.

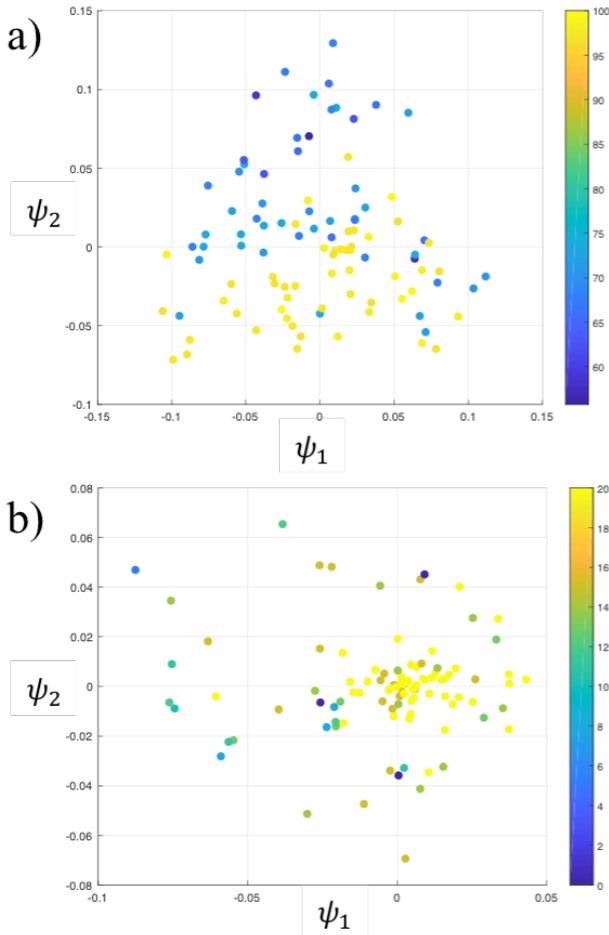


Figure 4: 2sDM embedding of the HCP and PNC dataset. Points are embedded in the 2D space by the first two non-trivial coordinates of diffusion map. Each point represents a subject. a) shows the embedding of HCP subjects, colored by in-scanner working memory task accuracy. b) shows the embedding of PNC subjects, colored by out-of-scanner working memory task correct responses.

Conclusion

In this paper, we propose a hierarchical diffusion maps-based fMRI data embedding framework. In two independent task-fMRI datasets, we showed that 2sDM sufficiently separates between different conditions in task and subjects, suggesting our framework can extract meaningful brain states from fMRI data. While this framework is designed for time-synchronized task fMRI data, recent methods have been developed to create time-synchronized resting-state fMRI data (Joshi, Chong, Li, Choi, & Leahy, 2018). Thus, for future work, we will adapt our framework for resting-state fMRI to investigate brain states

when a subject is not explicitly performing a task.

Acknowledgments

Data were provided in part by the Human Connectome Project, WU-MinnConsortium (Principal Investigators: David Van Essen and Kamil Ugurbil;1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The authors would also like to thank Ronald Coifman and R. Todd Constable for their support. GM is supported by the United States-Israel Binational Science Foundation and by the NSF (2015582) and the NIH (1R01HG008383-01A1).

References

- Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., & Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cerebral cortex*, *24*(3), 663–676.
- Coifman, R. R., & Lafon, S. (2006). Diffusion maps. *Applied and computational harmonic analysis*, *21*(1), 5–30.
- Joshi, A. A., Chong, M., Li, J., Choi, S., & Leahy, R. M. (2018). Are you thinking what i'm thinking? synchronization of resting fmri time-series across subjects. *NeuroImage*, *172*, 740–752.
- Lindquist, M. A., Xu, Y., Nebel, M. B., & Caffo, B. S. (2014). Evaluating dynamic bivariate correlations in resting-state fmri: a comparison study and a new approach. *NeuroImage*, *101*, 531–546.
- Monti, R. P., Lorenz, R., Hellyer, P., Leech, R., Anagnostopoulos, C., & Montana, G. (2017). Decoding time-varying functional connectivity networks via linear graph embedding methods. *Frontiers in computational neuroscience*, *11*, 14.
- Nadler, B., Lafon, S., Coifman, R. R., & Kevrekidis, I. G. (2006). Diffusion maps, spectral clustering and reaction coordinates of dynamical systems. *Applied and Computational Harmonic Analysis*, *21*(1), 113–127.
- Rabin, N., & Averbuch, A. (2010). Detection of anomaly trends in dynamically evolving systems. In *AAAI fall symposium: Manifold learning and its applications*.
- Satterthwaite, T. D., Connolly, J. J., Ruparel, K., Calkins, M. E., Jackson, C., Elliott, M. A., ... others (2016). The Philadelphia Neurodevelopmental Cohort: a publicly available resource for the study of normal and abnormal brain development in youth. *NeuroImage*, *124*, 1115–1119.
- Shen, X., Tokoglu, F., Papademetris, X., & Constable, R. T. (2013). Groupwise whole-brain parcellation from resting-state fmri data for network node identification. *NeuroImage*, *82*, 403–415.
- Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence*, *22*(8), 888–905.
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., ... others (2013). The WU-Minn human connectome project: an overview. *NeuroImage*, *80*, 62–79.