BrainSync: An Orthogonal Transformation for Synchronization of fMRI Data Across Subjects

*****

*****

Abstract. We describe a method that allows direct comparison of resting fMRI (rfMRI) time series across subjects. For this purpose, we exploit the geometry of the rfMRI signal space to conjecture the existence of an orthogonal transformation that synchronizes fMRI time series across sessions and subjects. The method is based on the observation that rfMRI data exhibit similar connectivity patterns across subjects, as reflected in the pairwise correlation between different brain regions. The orthogonal transformation that performs the synchronization is unique, invertible, efficient to compute, and preserves the connectivity structure of the original data for all subjects. Similarly to image registration, where we spatially align the anatomical brain images, this synchronization of brain signals across a population or within subject across sessions facilitates longitudinal and cross-sectional studies of rfMRI data. The utility of this transformation is illustrated through applications to quantification of fMRI variability across subjects and sessions, joint cortical clustering of a population and comparison of task-related and resting fMRI.

1 Introduction

Resting fMRI (rfMRI) is being increasingly used to study brain connectivity and functional organization [17]. It is also used for longitudinal studies of brain development and as a diagnostic biomarker in cross-sectional studies for various neurological and psychological diseases and conditions [13,12]. Large scale connectivity information derived from fMRI can be used to delineate functional regions [3]. By extension, identification of multiple contiguous areas, each of which exhibits distinct connectivity to the rest of the brain, can be used to define a functional parcellation of the entire cerebral cortex [19,17].

Since rfMRI data reflect spontaneous brain activity, it is not possible to directly compare signals across subjects [9]. Instead, comparisons make use of connectivity features, typically computed from pairwise correlations of the rfMRI time series between a point of interest and other locations in the brain [6]. For analysis of cerebral cortex, it is common to compute a feature vector at each location on a tessellated representation of the cortex as the correlation from that vertex to all other vertices. This is a very high dimensional feature and often requires dimensionality reduction or down-sampling for use in multi-subject comparisons. An alternative approach for inter-subjects comparisons is to use group independent component analysis (ICA) [5]. Group ICA concatenates

rfMRI data from multiple subjects and represents the data as a summation of independent spatial or temporal components. In this way common networks across subjects can be identified.

Here we describe a novel method for inter-subject comparison of fMRI signals in which a transformation is applied that allows direct comparison of time series across subjects. We use the geometry of normalized (zero mean, unit length) time series to represent the rfMRI data as a set of labeled points on the hypersphere. We then conjecture the existence of an orthogonal transformation, which we refer to as *BrainSync*, that makes the rfMRI from two subjects directly comparable. BrainSync retains the original signal geometry by preserving the pairwise geodesic distances between all pairs of points on the hypersphere while also temporally aligning or synchronizing the two scans. This synchronization results in an approximate matching of the time series at homologous locations across subjects. The synchronized data can then be directly pooled to facilitate large scale studies involving multiple subjects from cross-sectional as well as longitudinal studies. While the method is primarily designed for resting fMRI, we also show its application to task fMRI involving simple motor tasks.

2 Methods

We assume we have rfMRI and associated structural MRI data for two subjects. Our goal is to synchronize the rfMRI time series between these two subjects, although the method extends directly both to multiple sessions for a single subject or synchronization across multiple subjects. Our analysis below assumes that the rfMRI data has been mapped on to a tessellated representation of the midcortical layer of the cerebral cortex. The cortical surfaces for the two subjects must also be nonrigidly aligned and resampled onto a common mesh, as can be achieved using FreeSurfer [7] or BrainSuite [15].

Denote the cortically mapped rfMRI data for the subjects as matrices X and Y, each of size $T \times V$, where T represents the number of time points and V is the number of vertices in the cortical mesh. Corresponding columns in X and Y represent the time series at homologous locations in the two brains. The data vectors in each column in X and Y are normalized to have zero mean and unit length.

2.1 Geometry of the rfMRI signal space

Since the time series at each vertex are of unit length, we can represent each column of X and Y as a single point on the unit hypersphere $\mathbb{S}^{T-1}$ of dimension $T-1$, Fig. 1a, where T is the number of time samples .

Let $\boldsymbol{x}$ and $\boldsymbol{y}$ represent time series from two points in the brain. Then the dot product of $\boldsymbol{x}$ and $\boldsymbol{y}$ yields the Pearson correlation $\rho_{\boldsymbol{xy}}$ between them. The inverse cosine of $\rho_{\boldsymbol{xy}}$ gives the geodesic distance between the points on the hypersphere. The squared Euclidean distance between them is given by $\|\boldsymbol{x}-\boldsymbol{y}\|^2 = 2-2\rho_{\boldsymbol{xy}}$ and so is also solely a function of $\rho_{\boldsymbol{xy}}$. Since distance on the hypersphere depends

only on correlation, we expect clusters of points on the hypersphere, with each cluster corresponding to vertices within a functionally homogeneous region. It therefore follows that if two subjects have similar connectivity patterns to each other, then the clusters and the distance between these clusters will be similar for both subjects. With this picture in mind, we conjecture the existence of an orthogonal transformation (rotation and/or reflection) that will map the data from one subject onto that of the other based on the following well known result [4]:

Proposition 1. *Let* $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_V$ *and* $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_V$ *be points in* $\mathbb{R}^T$. *If* $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| = \|\boldsymbol{y}_i - \boldsymbol{y}_j\|, \forall i, j \in \{1, \cdots, V\}$, *then there exists a rigid motion* (O, t) *such that* $\boldsymbol{x}_i = O\boldsymbol{y}_i + t, \forall i \in \{1, \cdots, V\}$.

Since in our case the points are on a hypersphere $\mathbb{S}^{T-1}$, we can exclude the translation and apply a strict orthogonal transformation. In order to illustrate this concept, we performed the following illustrative experiment using in vivo rfMRI data. We took cortically constrained data from a pair of subjects consisting of $T = 1200$ time samples per vertex. For illustrative purposes we need to reduce dimensionality to $\mathbb{R}^3$ so that we can plot the data on the $\mathbb{S}^2$ sphere. We consider data from only three locations: cingulate, motor cortex and visual cortex. We projected this data onto the $\mathbb{R}^3$ subspace corresponding to the three largest singular values in the data matrix and renormalized to unit length. This data is of sufficiently low rank that we can indeed see the clustering of points on the sphere (Fig. 1b, c). Fig. 1d shows the result of applying the BrainSync orthogonal transformation described below to the original data and then applying the same dimensionality reduction as previously applied to subject 1. The data for subject 2 is now very similar to that of subject 1, consistent with our earlier conjecture.

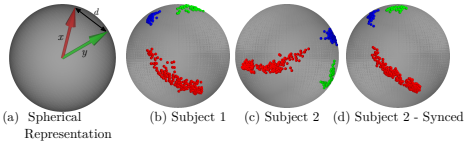


Fig. 1. Illustration of the BrainSync concept: (a) the geodesic and Euclidean distances between points on the hypersphere; (b),(c): data from cingulate (red), motor (green) and visual (blue) cortex for two subjects after dimensionality reduction to 3D; (d) data from subject 2 after application of BrainSync to subject 1 followed by dimensionality reduction identical to that for subject 1 in (b).

2.2 A method for temporal synchronization

The orthogonal transform O^s to synchronize the two data sets, X and Y, is chosen to minimize the overall squared error: $O^s = \arg\min_{O \in O(T)} \|X - OY\|^2$ where $O(T)$ represents the group of $T \times T$ orthogonal matrices. Given the high dimensionality of the surface vertices ($V \simeq 32,000$) relative to number of time samples ($T \simeq 1,200$) in the data analyzed below, the problem is well-posed and can be solved using the Kabsch algorithm [10]. Following the derivation in [18],

we first form the $T \times T$ cross-correlation matrix XY^t and compute its SVD: $XY^t = U\Sigma V^t$. The optimal orthogonal matrix is then given by $O^s = UV^T$.

To illustrate the behavior of BrainSync we applied this orthogonal transformation to data from a pair of rfMRI data sets from the HCP database described and explored more extensively below. Fig. 2 shows an example of the time series before and after BrainSync for the same vertex for the two subjects.

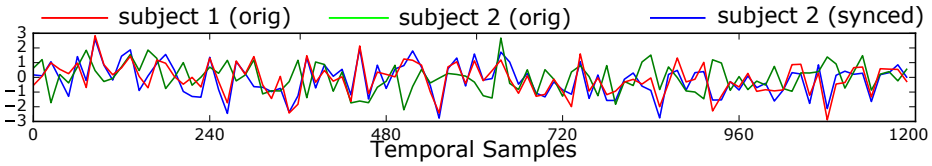


Fig. 2. Representative time series for two subjects for a single cortical location before and after synchronization.

3 Applications, Experiments and Results

3.1 Data

We used the minimally preprocessed (ICA-FIX denoised) resting fMRI data from 40 unrelated subjects, which are publicly available from the Human Connectome Project (HCP) [17,8]. In addition to this processing, we also applied the temporal non-local means (tNLM) filter [2] to improve SNR. Finally, we normalized the filtered resting fMRI time series at each vertex to zero mean and unit length.

3.2 Application 1: Quantifying variability of rfMRI across population

To compute within subject variability, we computed the correlation at each vertex between two sessions in the same subject after synchronization and averaged the result over all 40 subjects, Fig. 3a. To compute between subject variability, we chose one reference brain and synchronized the data from the other 39 brains to it. We then computed the between-subject correlations at each vertex after synchronization, averaged over all subjects and plotted the result as the map shown in Fig. 3. Within-subject variability across sessions reveals that most of the brain exhibits repeatable patterns of correlation, which lead to accurate syncing. Exceptions of areas with relative less correlation including limbic cortex, anterior temporal pole, insula and medial frontal cortex, but much of this is probably due to low signal artifacts. Across subject correlations are lower than within subject correlation, as expected the brain was less synchronized across subjects, although still very highly correlated in the similar regions.

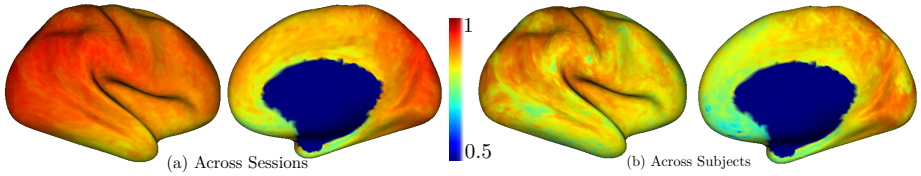


Fig. 3. Correlation of Resting fMRI and BrainSync: (a) across two sessions for the same subject, averaged over 40 subjects; (b) between subjects averaged over all pairs and two sessions.

3.3 Application 2: Cortical parcellation

Parcellations of the human cerebral cortex representing cyto-, myelo- or chemo-architecture are helpful in understanding the structural and function organization of the brain [20,1]. In vivo imaging, specifically, rfMRI has been used for identification of contiguous areas of cortex that exhibit similar functional connectivity to define a functional parcellation [19,17]. One major problem in using rfMRI for single subject parcellation is that the amount of data from a single subject is usually not sufficient to reliably parcellate the cerebral cortex brain into a large number of regions [14,16]. Since BrainSync makes data across subjects directly comparable, this synchronized data can be easily pooled and a joint parcellation of a large number of subjects is possible.

We synchronized all the subject data to a single reference brain and pooled the data from 40 subjects $\times$ 2 sesssions. Let $B_{i,j}$ represent the $T \times V$ data matrix for the i^{th} subject and j^{th} scan, all synchronized to the first subject's first scan. The concatenated data matrix is then $B = [B_{1,1}, B_{2,1}, \cdots, B_{40,1}, B_{1,2}, \cdots, B_{40,2}]$. The k-means algorithm was then applied to this data to generate simultaneous parcellation into k=100 regions of all the 40×2 brain scans. Note that we do not enforce any spatial prior or topological constraint on the data. Sample parcellations for two subjects, two sessions each, are shown in Fig. 4 for individual clustering and joint (BrainSync) clustering. For visualization of the individual clustering results, we used the Hungarian algorithm [11] for label matching across subjects. For joint clustering, corresponding regions are automatically identified through k-means clustering and no re-labelling is required.

To quantify performance, we computed the Adjusted Rand Index (ARI) between all pairs of subjects and scans and report both within-subject and across-subject similarity by averaging across subjects and sessions respectively. ARIs were computed for both the individual and group parcellation.

As expected, individual clustering results for $k = 100$ parcels are very variable across both sessions and subjects, Fig. 4, because of the limited information in a single 15min rfMRI scan. After synchronization results appear far more consistent across sessions. They also appear more coherent across subjects, although, unsurprisingly less so than the within-subject comparisons. Table 1 shows that the ARI is substantially higher for synchronized joint clustering vs. individual

	Orig (k=30)	Sync (k=30)	Orig (k=100)	Sync (k=100)	Orig (k=200)	Sync (k=200)
Across Subject	0.90(0.10)	0.97(0.07)	0.32(0.16)	0.89(0.07)	0.16(0.05)	0.64(0.04)
Across Sessions	0.94(0.08)	0.99(0.06)	0.42(0.12)	0.94(0.03)	0.32(0.02)	0.83(0.07)

Table 1. Adjusted Rand Indices: mean(std) for different number of classes (k) for individual (Orig) and group (Sync) parcellation.

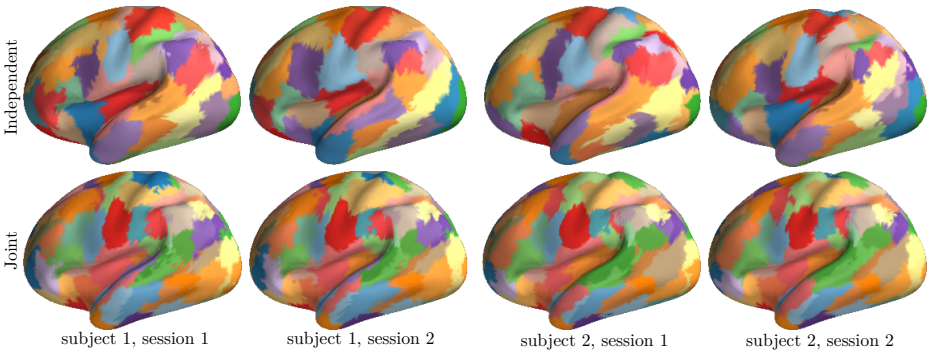


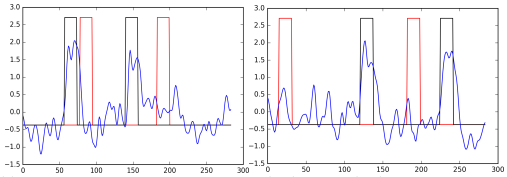
Fig. 4. Representative individual parcellation results ($k = 100$) for two subjects, two sessions each.

clustering. Table 1 also shows significantly higher across session similarity than across subjects.

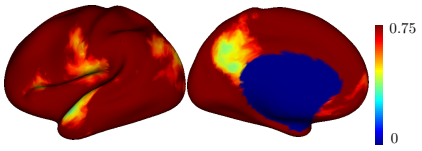
3.4 Applications to task fMRI

Predicting timing information To further analyze the performance of BrainSync, we considered two sessions of a block motor task for a single subject. These two sessions involved identical motor tasks but the timing blocks were different. The first session was synchronized to the second and the resulting orthogonal transformation O^s was applied to the timing blocks of the first session. As shown in Fig. 5(a), O^s allows us to predict the task timing of the second session from the rfMRI data from the first session.

Contrasting task vs rest We also use BrainSync to directly compare resting and motor activity. For this purpose we considered motor activity (self-paced tongue movement) and resting data from the HCP database for a single subject. The resting and task data were synchronized using BrainSync. At each point on the brain, the correlation between synced task and resting data was computed (Fig. 5 (b)). Results shown in Fig 5 (b) indicate that despite the fact that we are comparing task and resting data, much of the brain can still be synchronized. Exceptions include the facial primary motor cortex area and portions of the default mode network. This observation is consistent with the fact that we would expect increased motor activity and decreased default mode activity during the



(a) Synchronization of timing blocks for (left) tongue (left) right hand tasks



(b) Motor (tongue) to Resting correlation

Fig. 5. Task data (a) red: timing blocks for session 1, black: timing blocks for session 2 and blue: timing for session 1 after applying the orthogonal transformation predicted by BrainSync to timing blocks for (left) tongue and (right) right hand motor tasks. (b) Correlation between resting and synchronized motor tongue task time series. Strong correlation between task and resting fMRI can be seen throughout the brain other than in primary motor and default mode regions where we expect maximum dissimilarity between brain activity during motor and rest states.

motor task. This result shows promise for the use of BrainSync to compare brain activity between different states in a manner that accounts for ongoing (spontaneous) brain activity in both conditions.

4 Discussion and Conclusion

We have described a novel method for synchronization of rfMRI data across subjects and scans. By exploiting similarity in correlation structure across subjects we are able to transform the time series so that they become highly correlated across subjects. This synchronization process bears some similarity to image registration, in the sense that after synchronization comparisons can be made directly with respect to these time series across subjects. Importantly, since the transformation is orthogonal, correlations in the original data are preserved and the transform is invertible. One of the implicit assumptions in this work is that the rfMRI signal is stationary in the sense that correlation patterns are preserved over time. Our results show good correspondence of signals over the 15min windows used in this analysis. However, even within a 15min period we would expect to see variations in the activity of different networks, and it would be interesting to explore whether BrainSync is able to enhance our ability to identify and characterize these dynamic changes in network activity. While the proposed method is primarily designed for resting fMRI, we also show preliminary results indicating the possibility of synchronizing brain activity during different forms of activity in a manner that accounts for ongoing spontaneous brain activity.

References

1. K. Amunts, A. Schleicher, and K. Zilles. Cytoarchitecture of the cerebral cortex-more than localization. *Neuroimage*, 37(4):1061–1065, 2007.
2. C. Bhushan, M. Chong, S. Choi, A. A. Joshi, J. P. Haldar, H. Damasio, and R. M. Leahy. Temporal non-local means filtering reveals real-time whole-brain cortical interactions in resting fmri. *PloS one*, 11(7):e0158504, 2016.

3. B. Biswal, Zerrin Y. F., V. M. Haughton, and J. S. Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic resonance in medicine*, 34(4):537–541, 1995.
4. M. Boutin and G. Kemper. On reconstructing n-point configurations from the distribution of distances or areas. *Advances in Applied Mathematics*, 32(4):709–735, 2004.
5. V. D Calhoun, J. Liu, and T. Adalı. A review of group ica for fmri data and ica for joint inference of imaging, genetic, and erp data. *Neuroimage*, 45(1):S163–S172, 2009.
6. Y. Fan, L. D. Nickerson, H. Li, Y. Ma, B. Lyu, X. Miao, Y. Zhuo, J. Ge, Q. Zou, and J. Gao. Functional connectivity-based parcellation of the thalamus: an unsupervised clustering method and its validity investigation. *Brain connectivity*, 5(10):620–630, 2015.
7. B. Fischl. Freesurfer. *Neuroimage*, 62(2):774–781, 2012.
8. M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, Jesper L. A., J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D.C. Van Essen, and Jenkinson M. The minimal preprocessing pipelines for the human connectome project. *Neuroimage*, 80:105–124, 2013.
9. A. Iraji, V. D. Calhoun, N. M. Wiseman, E. Davoodi-Bojd, M. R.N. Avanaki, E. M. Haacke, and Z. Kou. The connectivity domain: Analyzing resting state fmri data using feature-based data-driven and model-based methods. *NeuroImage*, 134:494–507, 2016.
10. W. Kabsch. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, 32(5):922–923, 1976.
11. H. W. Kuhn. The hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2):83–97, 1955.
12. J. Posner, C. Park, and Z. Wang. Connecting the dots: a review of resting connectivity mri studies in attention-deficit/hyperactivity disorder. *Neuropsychology review*, 24(1):3–15, 2014.
13. E. Redcay, J. M. Moran, P. L. Mavros, H. Tager-Flusberg, J. D.E. Gabrieli, and S. Whitfield-Gabrieli. Intrinsic functional network organization in high-functioning adolescents with autism spectrum disorder. *Frontiers in human neuroscience*, 7:573, 2013.
14. M. Rubinov and O. Sporns. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, 52(3):1059–1069, 2010.
15. D. W. Shattuck and R. M. Leahy. Brainsuite: an automated cortical surface identification tool. *Medical image analysis*, 6(2):129–142, 2002.
16. S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey, and M. W. Woolrich. Network modelling methods for fmri. *Neuroimage*, 54(2):875–891, 2011.
17. S. M. Smith, D. Vidaurre, C. F. Beckmann, M. F. Glasser, M. Jenkinson, K. L. Miller, T. E. Nichols, E. C. Robinson, G. Salimi-Khorshidi, M. W. Woolrich, M. W. Barch, K. Ugurbil, and D.C. Van Essen. Functional connectomics from resting-state fmri. *Trends in cognitive sciences*, 17(12):666–682, 2013.
18. O. Sorkine. Least-squares rigid motion using svd. *Technical notes*, 120(3), 2009.
19. O. Sporns, G. Tononi, and R. Kötter. The human connectome: a structural description of the human brain. *PLoS Comput Biol*, 1(4):e42, 2005.
20. C.B. Von Economo and G.N. Koskinas. The cytoarchitectonics of the adult human cortex. *Vienna and Berlin: Julius Springer Verlag*, 1925.