Sparse Representation of Higher-Order Functional Interaction Patterns in Task-Based FMRI Data

Shu Zhang[1,2], Xiang Li[1], Jinglei Lv[2], Xi Jiang[1], Dajiang Zhu[1], Hanbo Chen[1], Tuo Zhang[2], Lei Guo[2], and Tianming Liu[1]

[1] Department of Computer Science and Bioimaging Research Center, The University of Georgia, Athens, GA, USA
[2] School of Automation, Northwestern Polytechnical University, Xi'an, China

Abstract. Traditional task-based fMRI activation detection methods, e.g., the widely used general linear model (GLM), assume that the brain's hemodynamic responses follow the block-based or event-related stimulus paradigm. Typically, these activation detections are performed voxel-wise independently, and then are usually followed by statistical corrections. Despite remarkable successes and wide adoption of these methods, it remains largely unknown how functional brain regions interact with each other within specific networks during task performance blocks and in the baseline. In this paper, we present a novel algorithmic pipeline to statistically infer and sparsely represent higher-order functional interaction patterns within the working memory network during task performance and in the baseline. Specifically, a collection of higher-order interactions are inferred via the greedy equivalence search (GES) algorithm for both task and baseline blocks. In the next stage, an effective online dictionary learning algorithm is utilized for sparse representation of the inferred higher-order interaction patterns. Application of this framework on a working memory task-based fMRI data reveals interesting and meaningful distributions of the learned sparse dictionary atoms in task and baseline blocks. In comparison with traditional voxel-wise activation detection and recent pair-wise functional connectivity analysis, our framework offers a new methodology for representation and exploration of higher-order functional activities in the brain.

Keywords: task-based fMRI, GES, sparse coding, dictionary learning, higher-order interaction.

1 Introduction

Voxel-based fMRI activation detection has been widely adopted in the functional brain mapping field. For instance, the general linear model (GLM) [1] is often used to detect activated voxels in task-based fMRI data, and followed by statistical corrections of the detected foci. However, voxel-wise activation detection methods have their limitations in terms of revealing the complex functional interaction patterns, since the brain often functions a network behavior. In recognition of this limitation, recently, several new studies have examined the functional connectivities during task performance in task fMRI data. The authors in [2] proposed a fiber-centered

K. Mori et al. (Eds.): MICCAI 2013, Part III, LNCS 8151, pp. 626–634, 2013.

activation detection method to find the activated connectivity patterns and the results demonstrated activated fiber-connected regions covered substantially wider brain areas than the traditional voxel-based activation methods. Another recent literature study in [3] examined the temporal dynamics of functional connectivity during task performance and found that the whole-brain's functional connectivity pattern well correlated with the block-based stimulus curve [3]. The results in [2, 3] suggested the feasibility and promise of examining functional connectivity patterns in task-based fMRI.

However, all the above-mentioned method [1, 2, 3] are still constrained in terms of the lack of quantitative representation of higher-order functional brain responses, and multivariate functional interaction patterns within brain networks are omitted in these methods. Essentially, both basic neuroscience research and computational modeling of neuroimaging data have proved that brain functions are typically realized via higher-order functional interactions among specific networks [4, 5]. In the literature, there are several other methods published to deal with higher-order interactions among multiple ROIs (regions of interests) such as independent component analysis [6], Granger causality modeling [7], dynamic causal modeling [8] and Bayesian graphical models [4, 5]. In particular, the Bayesian graphical causal models are proven to exhibit superior performance on estimating the network structure [4, 5] in both simulated and real data. Conceptually, Bayesian models are based on marginal and conditional probabilistic dependencies, which determines this method more suitable in estimating the network structure and less sensitive to the noises in fMRI signals [4]. However, only Bayesian models is not enough, the inferred interaction patterns in both baseline and task blocks among the whole-brain structural connectomes could be potentially overlapping with each other, their temporal transitions could be gradual.

In this paper, our novel algorithmic framework offers a new methodology to explore and represent higher-order functional activities in task-based fMRI data. Specifically, the greedy equivalence search (GES) algorithm [9] is employed to infer multivariate functional interactions for both task performance and baseline blocks. Then, an effective online dictionary learning algorithm [10] is utilized for sparse representation of the inferred higher-order functional interaction patterns obtained by GES. Both of the GES algorithm [9] and online dictionary learning algorithm [10] are not new to the world, but the novelty of this paper lies in its integration of both methods into an effective framework for modeling task-based fMRI data. Importantly, applying the proposed method on a working memory task-based fMRI data [11] has revealed distinguished and interesting of the learned sparse dictionary atoms in task and baseline blocks.

2 Materials and Methods

2.1 Overview

Based on the recently developed and publicly released DICCCOL (Dense Individualized and Common Connectivity-based Cortical Landmarks) system [12], structural connectome is constructed (Fig.1a) and the higher-order functional interactions are

modeled on the relevant sub-networks (Fig.1b) of the connectome. Then, the fMRI signals from each subject split into blocks by following the task and baseline paradigm (Fig.1c). All these segmented blocks of fMRI signals then generate one functional interaction pattern inferred by GES (Fig.1d). Afterwards, all of these interaction patterns are arranged into one matrix, as the bases of online dictionary learning algorithm used for sparse coding (Fig.1e). Finally, the proposed methods are applied on an operational span (OSPAN) working memory task-based fMRI dataset [11]. The flowchart is summarized in Fig. 1.

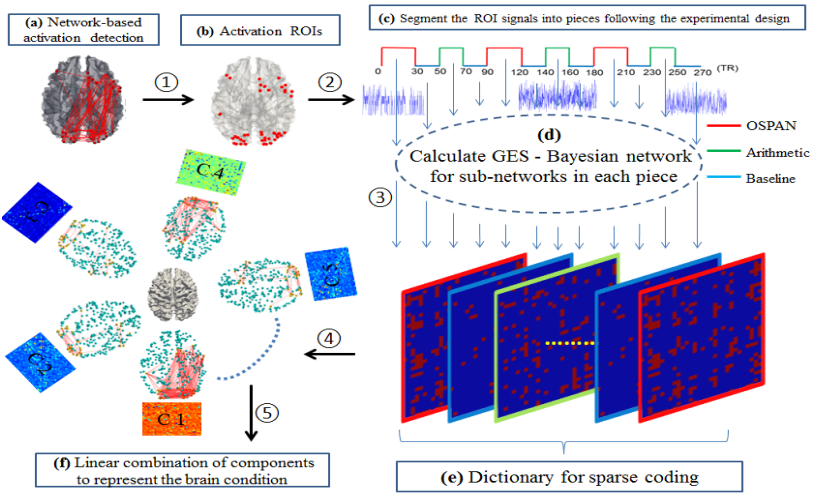

Fig. 1. The overview of sparse representation of higher-order functional interaction patterns in task-based fMRI data. The framework includes five main steps.

2.2 Data Acquisition and Pre-processing

In an operational span (OSPAN) working memory task-based fMRI experiment under IRB approval [11], 19 healthy young adult subjects were scanned and fMRI images were acquired on a 3T GE Sigma scanner. Briefly, acquisition parameters were taken as following: fMRI: 64×64 matrix, 4mm slice thickness, 220mm FOV, 30 slices, TR=1.5s, TE=25ms, ASSET=2. Each subject was performed by a modified version of the OSPAN task (3 block types: OSPAN, Arithmetic, and Baseline) while fMRI data were acquired [11]. DTI data was acquired with dimensionality 128×128×60, spatial resolution 2mm×2mm×2mm; parameters were TR 15.5s and TE 89.5ms, with 30 DWI gradient directions and 3 B0 volumes acquired. More details about pre-processing can be referred to [11].

2.3 Bayesian Network Modeling and The GES Algorithm

Bayesian network is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a direct acyclic graph (DAG).

For instance, $D = (V, E)$ is a directed acyclic graph satisfying the causal condition, and thus Markov factorization of the distribution can be utilized. Here, V is a finite set of DAG nodes and E is a finite set of directed edges between the DAG nodes. As a measure of how well a DAG is, D represents the conditional independencies between the random variables, we can use the relative probability:

$$S(D) = p(D, d) = p(d|D)p(D) \quad (1)$$

where *S(D)* refers to a network score [13], *p(d|D)* calculates the probability for observing a given dataset *d* under a given graphical model *D* with un-weighted edges, where an "improved" model will have larger *p(d|D)*. Thus the direction obtained by GES is based on the fact that two graphical models with the same undirected connectivity but different directionality would have different *p(d|D)*, resulting in the different scores for *S(D)*. In this work, we adopt the GES algorithm [10] to infer the D from fMRI signals extracted for the relevant sub-networks of DICCCOL-based structural connectomes. Briefly, GES begins with an empty graph, and then each time it searches for one edge to add to the graph over the space of Markov equivalence, and it only stops when the graph is not further improved by adding any more edges. Then, GES starts to search backwards, and each time it removes one edge until no improvement occurs by removing any edge. Thus the converged network graph is obtained.

In this work, each fMRI signal is split into 12 temporal segments by following the task paradigm for the 19 subjects. It leads to 228 pieces of fMRI segments. Then, for each segment, GES is used to generate a functional interaction graph that can represent the current higher-order brain activities. The GES toolbox we used is the publicly available TETRAD system [14], which is an effective tool for the GES algorithm. Until now, there are 228 GES-derived graphs obtained, among which 57 are for OSPAN blocks, 57 graphs for the Arithmetic blocks, and the last 114 graphs for the Baseline blocks. Two examples with different pattern of the GES graphs are shown in Fig. 2. Accordingly, higher-order functional interactions, instead of pair-wise connectivities, on the structural connectomes can be clearly appreciated.

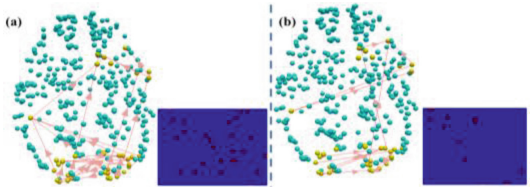

Fig. 2. Two exemplar patterns of the GES-derived graphs. The original data matrix is shown on the right bottom for each case.

2.4 Dictionary Learning and Sparse Coding

A large amount of recent studies in the machine learning field have demonstrated that sparse coding is superior in representing features and patterns. In this paper, we adopt the effective online dictionary learning algorithm [10] for sparse representation of the higher-order functional interaction patterns obtained in section 2.3. Specifically, for the problem interested, each GES graph matrix derived from section 2.3 with m_V rows and m_V columns is first reorganized into a vector x_i with the size

of $m = m_V \times m_V$, where m_V is the number of vertex in the graph. Then, given the intrinsically-established correspondences of DICCCOL-based connectomes across individual brains [12], we pool all the GES graphs from 19 subjects together to form a set of training samples $X = [x_1, x_2, \ldots x_i \ldots x_n], x_i \epsilon \mathbb{R}^m$ for dictionary learning. Thus, by using a learned dictionary $D \epsilon \mathbb{R}^{m \times k}$, the aim is to represent each sample with a weight vector $\alpha_i \epsilon \mathbb{R}^k$ that sparsely and linearly combine dictionary atoms, i.e., $x_i = D \times \alpha_i$. Here, the empirical cost function is defined:

$$f_n(D) \triangleq {}^1\!/_n \sum_{i=1}^{n} \ell(x_i, D) \tag{2}$$

in which the loss function to minimize the regression error of x_i with D is defined with a ℓ_1 regulation that seeks a sparse solution of α_i in Eq.(3). The parameter λ is a weight to control sparsity level. In our experiments, λ is all empirically set as 0.1 to minimize regression error residual.

$$\ell(x_i, D) \triangleq \min_{\alpha_i \epsilon \mathbb{R}^k} \frac{1}{2} ||x_i - D\alpha_i||_2^2 + \lambda ||\alpha_i||_1 \tag{3}$$

With the effective online dictionary learning method in [10], this problem can be solved using the publicly available SPAM-toolbox [10, 15]. Here, the learned dictionary D consists of k columns of components ($d_j \epsilon \mathbb{R}^m$) are believed to be bases of GES graph patterns in the form of vectors. As our training dataset is not large with n=228, we empirically choose a relatively smaller dictionary size k=10. Further, in order to identify the most relative GES graph pattern atom for each sample x_i, a sparse coding approach is implemented via the fast Orthogonal Matching Pursuit algorithm [3] in the SPAM-toolbox. Specifically, the purpose is to minimize the representation error with a limited number of dictionary atoms, i.e., $||\alpha_i||_0 \leq L$, see Eq.(4), which used for sparse coding of the input GES graph with the constraint on the number of atoms used

$$min_{\alpha_i \epsilon \mathbb{R}^k} ||x_i - D\alpha_i||_2^2, \ \ s.t. \ ||\alpha_i||_0 \leq L \tag{4}$$

In our experiments, the sparsity constrain L is set as 5, and in this way, each input GES graph pattern can be sparsely represented by 5 or less than 5 atoms in the learned dictionary. The sparse coding representation of the group-wise GES graph patterns can be used to discriminate different states of the brain during working memory task.

3 Experimental Results

3.1 Dictionary Learning Results

According to the algorithmic pipeline in Fig.1, first, group-wise activation detection is performed on 358 DICCCOLs via traditional GLM method, and 37 identified ROIs are most consistently activated in order to narrow down the number of ROIs concerned. Then, the 228 segments of fMRI signals within this sub-network of 37 ROIs are obtained via the online dictionary learning methods [10]. The size of the dictionary is empirically set as 10 in our application and the learned dictionary is visualized in Fig.3. It is interesting that the dictionary atom #2, #4 and #7 exhibit different levels of intense functional interactions within this sub-network, they are very important as

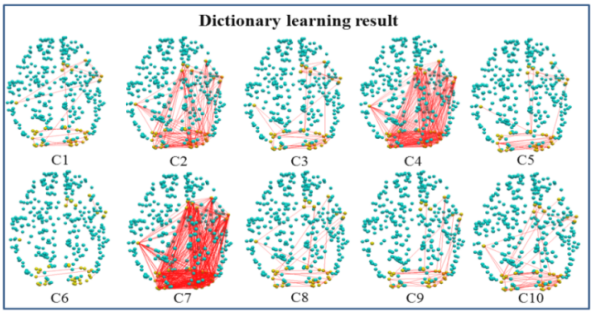

Fig. 3. The 10 learned dictionary atom components (C1-C10, respectively) are shown

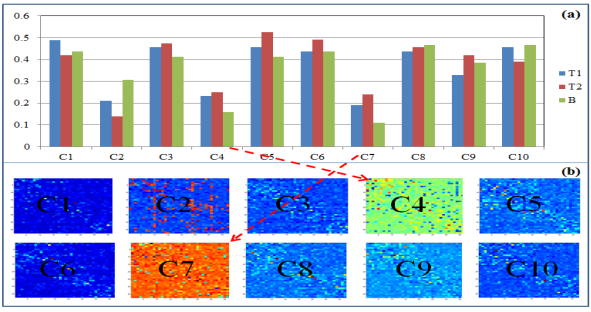

Fig. 4. (a) Frequency distributions of the 10 dictionary atoms in the three types of blocks. Here, T1 stands for OSPAN task, T2 stands for Arithmetic task, and B for Baseline. (b) Visualization of the 10 atoms (C1-C10) in the functional interaction matrices. The red color in the matrix represents higher functional interaction and the blue indicates lower functional interaction.

they have similar interaction patterns with the result from group-wise GLM, while other atoms show much less but variable functional interaction patterns. Moreover, they have similar patterns because the learning process needs more atoms to regress the dense (thus higher-weighted) interaction patterns. Though the frequencies of individual atoms are not high, their combined presence in task blocks is nearly 50%, which is really high in Fig.4a.

In addition, the 10 dictionary atoms are visualized in the matrix formats of functional interactions in Fig.4b and the statistical analysis of the distributions of these atoms is performed in three types of blocks as shown in Fig.4a. It is interesting that the two most active atoms #4 and #7 are substantially more frequent in both OSPAN and Arithmetic task blocks than in the Baseline block, as illustrated by the two red lines in Fig.4. This result is proven to be reasonable since task blocks exhibit more active interaction patterns. Also, for

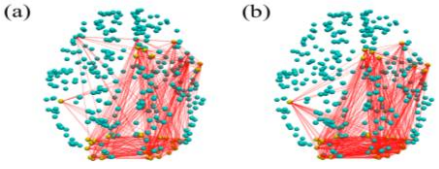

Fig. 5. (a) Activated networks of 37 ROIs detected from group-wise GLM [16]. (b) The most active atom by the method in this paper with 37 ROIs.

comparison purpose, the activated networks are visualized within the same set of 37 ROIs via our recently developed methods in [16] via group-wise GLM in Fig.5a. It is evident that the functional interaction pattern in atom #7 in Fig.5b exhibits similar activated pair-wise connections as those in Fig.5a, though there are more higher-order functional interactions in Fig.5b. This result not only suggests the validity of higher-order functional interaction modeling in this paper, but also demonstrates the superiority of GES-based inference of multivariate interaction patterns.

3.2 Reproducibility Study

As a reproducibility study of our methods, the z-value threshold was decreased when detecting the consistently activated DICCCOL landmarks via group-wise GLM, and thus 60 and 74 activated ROIs were obtained, respectively, to re-perform the experiments in section 3.1. Then we applied the methods in section 2 on these two experiments with different sizes of ROIs, and the results were shown in Fig.6. In both experiments, there were two active dictionary atoms, atoms #5 and #7 in Fig.6a and atoms #7 and #9 in Fig.6b. Again, these active atoms were substantially more frequent in both OSPAN and Arithmetic blocks than in the Baseline block. This result replicated the conclusion in section 3.1, suggesting the good reproducibility of our methods and results. Notably, for some inactive interaction patterns such as atom #10 in Fig.6a and atom #8 in Fig.6b, their frequencies were substantially higher in the Baseline block than in OSPAN and Arithmetic blocks, as illustrated by the green lines in Fig.6. This result further manifested the validity of our methods. What I want to emphasize is the atoms with same number in different experiment are independent.

For another comparison, we visualized the activated connections in the networks obtained by group-wise GLM methods and the interaction pattern of the most active atoms by our purposed methods with different numbers of selected ROIs in Fig.7. It was inspiring to observe more and more active connections from the left to the right in Fig.7, and the connections on the left side figure tended to be a sub-set of those on the right side figure. This result triggered us to re-think the limitations of traditional voxel-based [1] and connection-based [2, 3] activation detection methods. Hence, a large portion of the whole brain might be responsive to external tasks, such as the working memory in this paper, instead of a very small number of activated foci as detected by traditional voxel-based and connection-based methods [1-3].

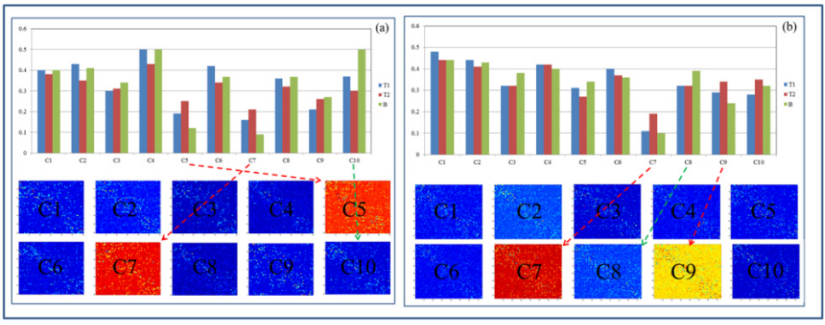

Fig. 6. Reproducibility study. In the charts, blue means OSPAN, red means Arithmetic and green means Baseline. (a) Results for 60 ROIs. (b) Results for 74 ROIs.

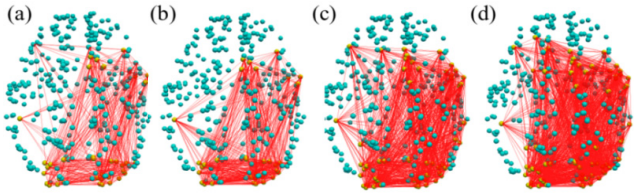

Fig. 7. Visualization of interaction patterns in (a) activated networks via group-wise GLM method, (b) most active atom with the ROI number of 37, (c) most active atom with the ROI number of 60, and (d) most active atom with the ROI number of 74

4 Conclusion

In this paper, in comparison with traditional voxel-wise activation detection and recent pair-wise functional connectivity analysis, a novel algorithmic framework was presented, which statistically inferred and sparsely represented higher-order functional interaction patterns within the working memory network during task performance and in the baseline. Experimental results on a working memory task-based fMRI data demonstrated meaningful, reproducible and interesting results. In the future, further evaluations and applying this framework on larger scale task-based fMRI datasets are to be planed.

References

1. Friston, K.J., et al.: Statistical parametric maps in functional imaging: a general linear approach. Human Brain Mapping 2(4), 189–210 (1994)
2. Lv, J., Guo, L., Li, K., Hu, X., Zhu, D., Han, J., Liu, T.: Activated Fibers: Fiber-centered Activation Detection in Task-based FMRI. In: Székely, G., Hahn, H.K. (eds.) IPMI 2011. LNCS, vol. 6801, pp. 574–587. Springer, Heidelberg (2011)
3. Li, X., Lim, C., Li, K., Guo, L., Liu, T.: Detecting Brain State Changes via Fiber-Centered Functional Connectivity Analysis. Neuroinformatics (2012)
4. Sun, J., et al.: Inferring Consistent Functional Interaction Patterns from Natural Stimulus FMRI Data. NeuroImage (2012)
5. Ramsey, J., et al.: Six problems for causal inference from fMRI. Neuroimage 49(2), 1545–1558 (2010)
6. Calhoun, V.D., et al.: A method for making group inferences from functional MRI data using independent component analysis. Human Brain Mapping 14(3), 140–151 (2001)
7. Roebroeck, A., Formisano, E., Goebel, R.: Mapping directed influence over the brain using Granger Causality and fMRI. NeuroImage 25, 230–242 (2005)
8. Friston, K.J., et al.: Dynamic causal modelling. NeuroImage 19(3), 1273–1302 (2003)
9. Meek, C.: Graphical Models: Selecting Causal and Statistical Models. Carnegie Mellon University, Pittsburgh (1997)
10. Mairal, J., Bach, F., Ponce, J., Sapiro, G.: Online dictionary learning for sparse coding. In: Proceedings of the International Conference on Machine Learning, ICML (2009)
11. Faraco, C.C., et al.: Complex span tasks and hippocampal recruitment during working memory. NeuroImage 55(2), 773–787 (2011)

12. Zhu, D., Li, K., Guo, L., Jiang, et al.: DICCCOL: Dense Individualized and Common Connectivity-Based Cortical Landmarks. Cerebral Cortex (2012)
13. Susanne, G., et al.: DEAL: A Package for Learning Bayesian Networks (2003)
14. Eberhardt, F., Hoyer, P.O., Scheines, R.: Combining Experiments to Discover Linear Cyclic Models with Latent Variables. Journal of Machine Learning, Workshop and Conference Proceedings (AISTATS 2010) 9, 185–192 (2010)
15. Mairal, J., Bach, F., Ponce, J., Sapiro, G.: Online learning for matrix factorization and sparse coding. Journal of Machine Learning Research 11, 19–60 (2010)
16. Shu, Z., et al.: Activated Cliques: Network-based Activation Detection in Task-based FMRI. In: ISBI 2013 (2013)