NeuroImage 105 (2015) 286–299

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Identifying Sparse Connectivity Patterns in the brain using resting-state fMRI

Harini Eavani [a,*], Theodore D. Satterthwaite [a,b], Roman Filipovych [a], Raquel E. Gur [b], Ruben C. Gur [b], Christos Davatzikos [a]

[a] Center for Biomedical Image Computing and Analytics, University of Pennsylvania, USA
[b] Brain Behavior Laboratory, Department of Psychiatry, University of Pennsylvania, USA

ARTICLE INFO

Article history:
Accepted 25 September 2014
Available online 2 October 2014

Keywords:
Resting state fMRI
Functional connectivity
Sparsity

ABSTRACT

The human brain processes information via multiple distributed networks. An accurate model of the brain's functional connectome is critical for understanding both normal brain function as well as the dysfunction present in neuropsychiatric illnesses. Current methodologies that attempt to discover the organization of the functional connectome typically assume spatial or temporal separation of the underlying networks. This assumption deviates from an intuitive understanding of brain function, which is that of multiple, inter-dependent spatially overlapping brain networks that efficiently integrate information pertinent to diverse brain functions. It is now increasingly evident that neural systems use parsimonious formations and functional representations to efficiently process information while minimizing redundancy. Hence we exploit recent advances in the mathematics of sparse modeling to develop a methodological framework aiming to understand complex resting-state fMRI connectivity data. By favoring networks that explain the data via a relatively small number of participating brain regions, we obtain a parsimonious representation of brain function in terms of multiple "Sparse Connectivity Patterns" (SCPs), such that differential presence of these SCPs explains inter-subject variability. In this manner the sparsity-based framework can effectively capture the heterogeneity of functional activity patterns across individuals while potentially highlighting multiple sub-populations within the data that display similar patterns. Our results from simulated as well as real resting state fMRI data show that SCPs are accurate and reproducible between sub-samples as well as across datasets. These findings substantiate existing knowledge of intrinsic functional connectivity and provide novel insights into the functional organization of the human brain.

Introduction

The human brain is a complex system that consists of functionally specialized units working in unison to generate responses to internal and external stimuli. Resting-state fMRI (rs-fMRI) is a powerful tool for understanding the large-scale functional neuroanatomy of the brain through connectivity that is present independent of task performance. Functional connectivity is defined as correlations between the spontaneous fluctuations in the fMRI time-series among different regions. Prior research has shown that despite the absence of task performance, rs-fMRI connectivity can be used to delineate major functional brain systems as networks (Biswal et al., 1995; Fox et al., 2006; Vincent et al., 2008), often based on prior knowledge of a "seed" region of interest, and has demonstrated that network organization is altered in neuropsychiatric and neurological illnesses such as schizophrenia (Venkataraman et al., 2012) and Alzheimer's (Greicius et al., 2004). Thus, an accurate description of the brain's functional connectome is a critical prerequisite for understanding both normal brain function and its aberrations in disease.

Identifying these networks in a data-driven manner is a particularly challenging task due to the spatio-temporal complexity of rsfMRI. Robust identification requires the specification of the underlying common property that binds regions together to form a network. For example, graph partitioning approaches, such as InfoMap (Rosvall and Bergstrom, 2008) assume that any region of the brain can belong to only one brain network. This approach was applied to rsfMRI in Power et al. (2011). Retaining only high positive correlation values, the authors identified multiple spatially separated networks, or "subgraphs", whose regions consistently co-activate across subjects, and resemble functional systems discovered in task fMRI. However, up to date knowledge of the brain's functional organization seems to suggest that brain regions can participate in multiple functional networks. Graph partitioning approaches such as InfoMap do not allow for spatial overlap, and hence cannot identify such networks. Another disadvantage of such an approach is that it limits its analysis to strong positive correlations, while removing negative and weak edges from the graph that

* Corresponding author.
E-mail address: Harini.Eavani@uphs.upenn.edu (H. Eavani).

http://dx.doi.org/10.1016/j.neuroimage.2014.09.058
1053-8119/

could be informative, especially if considered collectively as a part of a distributed network (Fox et al., 2005; Keller et al., 2013).

Alternative approaches addressing some of these issues have been proposed in other fields. The hierarchical clustering algorithm proposed in Newman (2004) finds nested communities but does not allow for overlaps at each level in the hierarchy. The notion of "link communities" introduced in Ahn et al. (2010) is elegantly able to handle overlaps by assigning unique membership to edges rather than nodes, naturally resulting in multiple assignments per node. Approaches like correlation clustering (Bansal et al., 2004) and the Potts model based approach proposed in Traag and Bruggeman (2009) are partitioning approaches which allow negative values. Since most of these methods are used to analyze social networks, they interpret negative edge links as repulsion, and hence attempt to assign negatively connected groups to different communities. While this may be appropriate for social networks, in resting state fMRI, highly negative edges imply strong anti-correlation — meaning that despite opposing phase information, these nodes express the same information, since they are strongly statistically dependent. Allocating anti-correlated regions to the same network can provide interesting new insights into the functional organization of the brain. This leads to the formation of networks where topologically distinct partitions with similarly high values of modularity can be formed in a network. Most graph-theoretic approaches are ill-equipped to handle this scenario (Rubinov and Sporns, 2011).

Alternately, continuous matrix factorization approaches like Principal Component Analysis (PCA), Independent Component Analysis (ICA) or Non-negative Matrix Factorization (NMF) are applied directly to the time-series to obtain a set of basis, where each vector is a set of weights, one for each node. In some cases, matrix factorization can be interpreted as *soft*-clustering, or a continuous relaxation of the discrete clustering problem. For example, it has been shown that components obtained using Principal Component Analysis (PCA) are a continuous relaxation to the discrete clusters obtained using K-means (Ding and He, 2004). The symmetric Non-Negative Matrix Factorization (NMF) model is considered to be the continuous equivalent to kernel K-means and spectral clustering approaches (Ding et al., 2005). While PCA also exploits the second-order moment (correlations) to perform clustering, ICA incorporates higher-order moments to reveal sub-networks that are maximally independent. Such continuous approaches do not suffer from issues of non-overlap and negative values, but their main drawback is the lack of interpretability of the resulting components. The resulting basis vectors are dense, i.e., the weight of every node is typically non-zero, making clustering inference difficult. Approaches such as Independent Component Analysis (ICA) (Hyvarinen, 1999) are driven by the assumption of maximal spatial or temporal independence between networks. Spatial ICA is widely applied to rsfMRI data to obtain spatially independent components, commonly referred to as "Intrinsic Connectivity Networks (ICNs)" (Calhoun et al., 2003). In practice, ICNs found using Spatial ICA are usually non-overlapping. To address this issue of non-overlap, the study by Smith et al. (2012) applied Temporal ICA to rsfMRI data and found multiple functional brain networks, or "Temporal Functional Modes (TFMs)". Although this is a significant advancement, these networks have been identified on the basis of independent temporal behavior, i.e., lacking between-network interactions, which is contrary to the notion that brain systems often act in concert during complex cognitive functioning, for instance, for executive functioning (Dosenbach et al., 2006).

A major disadvantage of connectivity based approaches is their inability to directly quantify inter-subject variability in functional connectivity, requiring additional post-processing and analysis. An important source of variation across subjects is the average strength of networks. In this scenario, we assume that the inter-subject variability is introduced due the variation in the strength of each network across subjects. This could possibly be due to the extent to which (how much and for how long) that functional unit is recruited in each subject, or as an indicator of functional development or abnormality. There are studies that have found strong relationships between the clinical variable of interest and the strength of such intrinsic rsfMRI networks (von dem Hagen et al., 2013; Mayer et al., 2011; He et al., 2007; Satterthwaite et al., 2010). Another scenario that introduces inter-subject variability is in the membership of nodes to networks; this was modeled in Ng et al. (2012). In these prior clinical studies, inter-subject variability did not play a role in network identification; rather, average functional connectivity (strength) was computed *after* network identification. Hence quantifying inter-subject variability in connectivity in an automated, data-driven manner is crucial.

In this paper we propose a method that addresses these limitations. Motivated by models of neuronal activity (Vinje and Gallant, 2000), we propose the use of spatial sparsity to drive network identification. In a neuronal sparse coding system, information is encoded by a small number of synchronous neurons that are selective to a particular property of the stimulus (e.g. edges of a particular orientation within a visual stimulus). Multiple such spatial patterns of neurons constitute a sparse neural basis which acts in concert in response to the stimulus. A nearly infinite number of stimuli can be parsimoniously encoded by varying the proportion in which these patterns are combined.

Extending this idea to rs-fMRI, we assume that the observed spontaneous activity arises from the concerted activity of multiple "Sparse Connectivity Patterns (SCPs)" that encode system-level function, similar to sparse codes that are present at the level of neurons. Each SCP consists of a small set of spatially distributed, functionally synchronous brain regions, forming a basic pattern of co-activation. These SCPs capture the range of resting functional connectivity patterns in the brain, although they do not necessarily need to be present in each individual or subsets of individuals. Using spatial sparsity as a constraint, we learn the identity of these SCPs and the strength of their presence in each individual, revealing the heterogeneity in the population. Sparsity-based approaches bridge the gap between discrete clustering techniques and continuous dimensionality reduction approaches. The proposed method is not limited by problems related to negative correlations, overlapping sub-networks or modular degeneracy. The proposed approach is motivated by methods proposed for computer vision and machine learning applications in Sra and Cherian (2011) and Sivalingam et al. (2011). A preliminary version of this method was used in Eavani et al. (2013, 2014) but for different objectives, which was to find networks that characterize temporal variations and two-group differences in connectivity respectively. In this paper, the proposed method focuses on finding common networks that characterize average whole-brain functional connectivity in a group of subjects, while capturing inter-subject variations. The performance of the method is evaluated using simulated data and multiple resting-state fMRI datasets.

In the following sections, we describe the SCPs obtained in a rsfMRI dataset of young healthy adults, and how they compare to existing knowledge of functional organization of the brain. We investigate the accuracy and reproducibility of SCPs vis-a-vis sub-graphs, ICNs and TFMs using simulated data as well as real rsfMRI data. Furthermore, we provide evidence of inter-subject variability in the presence of SCPs, which is a valuable measurement that can facilitate inter-group comparisons in clinical studies.

Identification of Sparse Connectivity Patterns

The objective of our method is to find SCPs consisting of functionally synchronous regions, and are smaller than the whole-brain network. The information content within any one of these SCPs is also relatively low, since all the nodes within an SCP are correlated, and express the same information. Hence, if a correlation matrix were constructed for each of these SCPs, it would show two properties — (1) large number of edges with zero weights, or sparsity and (2) low information content — or rank deficiency. Our method takes as input correlation matrices and finds SCPs that satisfy both properties. We assume that if a set of ROIs act as a functional system, then, in a set of normal subjects,