8-2016

Analysis of structural and functional brain networks

Jun Young Jeong
Purdue University

Recommended Citation

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Jun Young Jeong

Entitled
ANALYSIS OF STRUCTURAL AND FUNCTIONAL BRAIN NETWORKS

For the degree of   Master of Science in Electrical and Computer Engineering

Is approved by the final examining committee:

Zhongming Liu
Chair

Eugenio Culurciello

Xiaojun Lin

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Zhongming Liu

Approved by: Venkataramanan Balakrishnan   5/17/2016

Head of the Departmental Graduate Program   Date

ANALYSIS OF STRUCTURAL AND FUNCTIONAL BRAIN NETWORKS

A Thesis

Submitted to the Faculty

of

Purdue University

by

Jun Young Jeong

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science in Electrical and Computer Engineering

August 2016

Purdue University

West Lafayette, Indiana

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Jeong, Jun Young. M.S.E.C.E., Purdue University, August 2016. Analysis of Structural and Functional Brain Networks. Major Professor: Zhongming Liu.

The brain is a representative example of a network. It consists of numerous spatially distributed regions that continuously exchange information through structural connections. In the past decade, an increasing number of studies have explored the brain network in both structural and functional aspects; they have begun to decipher complex brain wirings, as well as elucidate how the rich functionality emerges from this architecture. Based upon previous studies, this thesis addresses three critical gaps in the field. (I) Although it is known that the community structures of brain network are spatially overlapping, conventional studies have focused on grouping brain regions into communities such that each region belongs to only one community. Therefore, a recent "link community" concept was employed to disentangle those overlapping architectures. (II) Spatial independent component analysis (sICA) and k-means clustering are two representative data-driven algorithms used to analyze functional networks. However, it is still unclear how these two methods compare to each other in terms of their theoretical basis and biological relevance. Hence, the relationship between these two methods were investigated. (III) Despite the multi-scale functional organization of the brain, previous studies have primarily examined the large-scale networks of the entire brain. Complex neural activity patterns in relatively smaller spatial scales have been poorly understood. Therefore, the fine-scale spatiotemporal patterns within visual cortex were explored. The distinguishing results obtained in this study may provide new insights regarding the brain's organization, as well as a better understanding of mathematical and statistical tools for functional and structural network analysis.

1. INTRODUCTION

The human brain is a network. It consists of a large number of spatially segregated neural regions that each performs a unique function, and continuously interact with each other to support the brain's rich functionality, including perception, cognition, and action. As such, they form a complex integrative and segregated network architecture in which neuronal information is consistently processed and transported between structurally- and functionally- linked brain regions. These are defined as structural and functional connectivity, respectively.

Structural connectivity indicates physical or structural (synaptic) connections linking sets of neural elements. The major pattern of structural connections is stable for longer time-scale, but slight change may occur due to plasticity. In animal models, invasive tracing methods have been widely used to directly obtain axonal connections [1-3]. In contrast, for humans, diffusion weighted imaging techniques, such as diffusion tensor imaging (DTI), are able to non-invasively characterize white matter projections linking cortical and subcortical regions [4-5].

Previously, an increasing number of studies have analyzed the structural connectivity datasets of both human and non-human primates by primarily using graph theory, which describes the structural brain network as a graph [6-9], composed of nodes denoting neural elements that are interconnected by edges indicating physical connections. As a result, a fundamental insight of the topological properties of the complex structural network have been revealed, including the small-world property [10], power-law degree distribution [6], modularity [11], hierarchy [12], rich-club distribution with a highly connected hub regions [13,14], etc.

In contrast, with structural connectivity, functional connectivity is more of a statistical concept. Since functional connectivity is highly time-dependent, often

continuously changing within short time-scale according to different sensory stimuli or task contexts, the main analysis of functional connectivity has primarily focused on identifying and characterizing representative functional networks from this dynamically changing neural information obtained with a variety of recording techniques, including electroencephalography (EEG), magnetoencephalography (MEG) and functional MRI (fMRI).

Specifically, the functional networks can be inferred by focusing on either the temporal or spatial characteristics of neural information obtained with fore-mentioned neuroimaging techniques. As an example in the former category, cross-correlations of neurophysiological timeseries can be calculated to measure the degree of functional interactions between different brain regions [15-17]. In contrast, methods with the latter aim to infer functional networks by examining the spatial distribution of the images of neural activation at different time points. To this end, broad spectra of multivariate statistical algorithms have been employed [18-21]. For instance, spatial independent component analysis (sICA) has been widely applied to fMRI and MEG datasets to produce a set of spatially independent network patterns [18, 22-24]. As another example, k-means clustering has been utilized to group instantaneous fMRI activity patterns into so-called co-activation patterns [20].

By combining various neuroimaging modalities with distinct mathematical and statistical algorithms, the network analysis in functional aspect has revealed many interesting findings about the functional connections between specific brain regions, as well as the overall organization of functional communication in the brain network. A number of functional networks in charge of distinct brain functions have been identified, including visual, attention, and motor network, etc. [15, 19]. These networks are highly reproducible among healthy individuals, different experimental sessions, and even in non-human primates [25, 26]. In addition, they are also commonly observed in the state with and without any explicit task or sensory stimulation [22, 27]. Moreover, these network patterns have also been applied to a clinical context, serving as an indicator of different neurodegenerative diseases such as Alzheimer's disease or Traumatic Brain Injury (TBI) [28, 29].

In the past decade, many prior theoretical and empirical studies have explored the brain network in both structural and functional perspectives. As a result, these long-term efforts have begun to decipher extremely complex brain wirings, as well as elucidate how rich brain's functionality emerges from this architecture. Based upon previous brain network studies, this thesis try to address three critical questions in the field. The brief summary of motivations and the problem-solving process of each question is explained below.

In chapter 2, the main emphasis was placed on brain network analysis with graph theory, especially regarding community structures. Detecting and characterizing these structures is particularly important, since this can reveal groups of structurally and functionally related regions, which support numerous brain's functions. [9, 14]. However, conventional network studies have primarily focused on grouping brain regions into communities such that each region belongs to only one community and thus cannot be used to reveal overlapping structures [8, 11, 12, 30]. In order to bypass this problem, a recent "link community" concept [31] was applied to various structural and functional network datasets of human and non-human primates.

In chapter 3, two representative multivariate statistical algorithms of analyzing resting-state fMRI: spatial independent component analysis (sICA) and k-means clustering were thoroughly explored. Spatial ICA has been widely used to produce spatially independent resting-state networks [18, 23, 24]. Recently, k-means clustering has been employed to group instantaneous fMRI activity patterns into the so-called co-activation patterns (CAPs) [20]. Despite their different naming and interpretation, it is still unclear how these two methods were compared to each other in terms of their theoretical basis and biological relevance. To address these questions, the resting-state networks extracted by either spatial ICA or k-means clustering were compared against the task-based networks obtained by applying these two methods to a large set of task activation maps.

In chapter 4, the complex spontaneous brain activity patterns occurred in visual cortex were explored. Despite the multi-scale functional organization of the brain [32, 33], most of previous network studies have primarily focused on the whole-brain scale for mapping large-scale functional networks [18, 20, 21, 22, 25]. Thus, the complex neural

activity patterns in relatively smaller spatial scales remains poorly known and rarely explored. Towards filling this gap, the spatiotemporal patterns of spontaneous activity within visual cortex were evaluated in three different aspects: spatial organization, inter-subject reproducibility, and functional relevance to naturalistic visual perception.

2. LINK COMMUNITIES REVEAL OVERLAPPING STRUCTURAL AND FUNCTIONAL NETWORKS

2.1 Motivation

In order to account for the rich variety of human brain functions, individual brain regions are engaged in multiple neural processes that emerge from extremely entangled structural or functional brain networks. The wiring diagram of complex neural architecture, defined as the Connectome [34], has been primarily investigated by employing a graph-theoretical approach. The graph theory approach adopts a formal network description in which neural regions are represented by nodes, and their anatomical connections or functional associations are represented by edges [7-9, 13, 35]. Consequently, this approach has provided deep insights into the large-scale topological features of the brain, including the small-world property [10], truncated power-law degree distribution [6], and high efficiency with low-wiring cost [13, 14].

An important feature of the brain connectome that has received a significant attention is the detection and characterization of community structure, which represents group of densely interconnected neural regions, with only sparser connections between different groups [9, 12, 14]. Examining this structure is particularly important, since it provides deeper understanding of brain architecture by identifying group of structurally and functionally interrelated neural regions that plays a distinct biological role to support brain's complex behaviors, including perception, action and cognition. According to previous studies, the presence of community structures was demonstrated in structural and functional brain networks of both human and non-human primate models [8, 11, 36]. In

addition, it was inferred that these structures may play a central role in supporting efficient information processing occurred in the brain [14].

By this time, broad spectrum of methodologies and algorithms have been applied to objectively detect community structures of brain network [8, 11, 12, 30]. However, conventional network studies have entirely focused on grouping brain regions into communities such a way that each region belongs to only one community and thus cannot be used to reveal overlapping networks. Compared with the conventional view on community as a group of nodes, the community structure of complex networks can also be revealed by clustering links into the so-called "link communities" [31]. This edge-centric perspective is potentially more powerful than the traditional node-centric perspective, since this can identify the nodes, which simultaneously belong to multiple communities. A number of previous studies using resting-state fMRI have already reported that brain network nodes are likely to belong to multiple networks [19-21]. Therefore, the link community concept might be a better choice to disentangle complex brain architecture.

In this study, the applicability of the link community concept to the brain network was explored by using different structural and functional network dataset of both human and mammalians. As a result, it was demonstrated that brain networks displayed rich and diverse link community structures, which were spatially overlapping at multiple brain regions. To provide possible biological interpretation of the identified results, the quality of link community was first evaluated by examining the functional similarity of brain regions with same community membership. Then, as a next step, whether the brain regions' involvement to different link communities was consistent with the function of those specific regions was verified. To this end, relevant literatures was thoroughly reviewed.

2.2 Materials and Methods

2.2.1 About the link community algorithm

The link community algorithm is summarized as follows:

1. First, calculate the link similarity between all connected pairs of links sharing a common node (also defined as keystone node in literature). The formula for calculating the

link similarity is defined differently for binary and weighted network datasets. In case of the binarized networks, link similarity S is defined based on the Jaccard index [37]:

For the link pairs e_{ik} and e_{jk} that share a common node k, the link similarity S is defined as:

$$S(e_{ik}, e_{jk}) = \frac{|n_+(i) \cap n_+(j)|}{|n_+(i) \cap n_+(j)|} \tag{2.1}$$

where $n_+(i)$ is the set of node i itself and its neighbors.

In case of the weighted, directed, or signed networks, the above Jaccard index formula is generalized to the Tanimoto coefficient [38] which is described as follows:

Consider a vector $a_i = (\tilde{A}_{i1}, \dots, \tilde{A}_{iN})$ with

$$\tilde{A}_{ij} = \frac{1}{k_i} \sum_{i' \in n(i)} w_{ii'} \delta_{ij} + w_{ij} \tag{2.2}$$

where w_{ij} is the connection strength on edge e_{ij}, $n(i) = \{j | w_{ij} > 0\}$ is the set of all neighbors of node i, $k_i = |n(i)|$ is the total number of neighbors, and $\delta_{ij} = 1$ if $i = j$ and zero otherwise. Then, the similarity between edges e_{ik} and e_{jk} becomes:

$$S(e_{ik}, e_{jk}) = \frac{a_i \cdot a_j}{|a_i|^2 + |a_j|^2 - a_i \cdot a_j} \tag{2.3}$$

2. After the link similarity is calculated between all pairs of links sharing a common node, the single-linkage (average-linkage or complete-linkage can be other options) hierarchical clustering algorithm is utilized to discover hierarchical link community structures. Initially, each link belongs to its own community; then at each step, the pair of links with the largest similarity are chosen, and their respective communities are merged into one. The following process is repeated until all the links become members of a single cluster. The process of merging steps stated previously is also simultaneously recorded in a link dendrogram, which contains all the information about the hierarchical link community organization.

3. In order to discover the most meaningful link communities from the link dendrogram structure obtained at previous step, it is necessary to determine the partition point. To this end, the link community algorithm employs an objective function called

"partition density", *D* which is defined as formula (2.5). The partition density essentially measures how "clique-ish" or "tree-ish" each link community is. For a network with *M* links, $\{P_1, \ldots, P_C\}$ is a partition of the links into C communities. Subset P_c has $m_c = |P_c|$ links and $n_c = \left| \bigcup_{e_{ij} \in P_c} \{i, j\} \right|$ nodes. Then, they can be used to define

$$D_c = \frac{m_c - (n_c - 1)}{n_c(n_c - 1)/2 - (n_c - 1)} \tag{2.4}$$

This is m_c normalized by the minimum and maximum numbers of links possible between n_c nodes. The partition density, *D*, is the average of D_c weighted by the fraction of present links:

$$D = \sum_c m_c \frac{m_c - (n_c - 1)}{(n_c - 2)(n_c - 1)} \tag{2.5}$$

MATLAB implementation for the above link community algorithm is provided in the Appendix.

2.2.2 Structural and functional network datasets information

This section will briefly discuss the structural and functional network datasets utilized throughout this work, including which imaging modalities were used to obtain structural or functional information and how the nodes and their interconnections were defined. In this study, four structural and one functional network datasets were used. The structural datasets include tracer studies of the visual cortex and the cerebral cortex of macaque monkey [1, 2], a tracer study of the whole mouse brain [3], and diffusion weighted imaging (DTI) of whole human brain. For the functional dataset, the PTN (Parcellation, Timeseries, Netmats) dataset was downloaded from the Human Connectome Project (HCP) website (http://www.humanconnectome.org/).

2.2.2.1 Macaque visual cortex

The structural network dataset of a macaque visual system was obtained from the Brain Connectivity Toolbox (BCT) website (https://sites.google.com/site/bctnet/Home). This network dataset consists of 32 visual and visual-association cortical regions interconnected with 194 binary, undirected links. In specific, different cortical regions were identified based on three methodologies: (1) input and output structural connection

characteristics for each cortical region, (2) cortical myeloarchitecture using myelin and different staining techniques and (3) topographic organization relying on an orderly mapping of the visual field in each region. The linkages between different visual regions were defined by summarizing all of the pre-existing visual pathway studies which used histological tracing technique. See [1] for detail information.

2.2.2.2 Macaque cerebral cortex

The structural network dataset of the macaque cerebral cortex was obtained from the Brain Connectivity Toolbox website. This network dataset consists of seventy-two cortical regions distributed through four different brain subsystems including visual, somatosensory-motor, auditory, and frontal-limbic systems. These different regions were connected by 438 binary undirected links. The division of different brain regions was based principally on the same categorization schemes used in the previously discussed macaque visual cortex study. Past neuroanatomical literature was reviewed for connections between the 72 regions, and the reported connections were collated together into the single connectivity matrix. See [2] for further information.

2.2.2.3 Whole mouse brain

For structural analysis of the whole mouse brain, the published dataset used at [3] was downloaded (www.nature.com/nature/journal/v508/n7495/full/nature13186.html). This dataset was reconstructed based on the Allen Mouse Brain Connectivity Atlas (http://www.brain-map.org/), the database containing the brain-wide, cellular-level, mesoscale connectome for the mouse. 213 non-overlapping brain regions were chosen from the Allen Reference Atlas to comprehensively cover the major structures. Then, the axonal connections between two regions were defined by summarizing 469 experiments which employed enhanced green fluorescent protein (EGFP)-expressing adeno-associated viral vectors. The best-fit model results from a bounded optimization followed by a linear regression to determine connection coefficients, assigning a statistical significance (p-values) to each connection in the matrix. See [3] for detail procedure and information.

Then, based on this weighted dataset, a binarized, undirected structural connectivity matrix was further generated. Two different regions were defined as "connected" if there

existed connection coefficients in either direction that were below the predefined thresholds. Since the optimized threshold value is unknown, four different *p*-value thresholds (0.001, 0.0001, 0.00001 and 0.000001) were applied further analyzed how the different threshold levels affect to the overall link community structures.

2.2.2.4 Human diffusion weighted imaging

For structural analysis of whole human brain, diffusion weighted imaging (DWI) dataset was acquired from each of three subjects. After defining 184 spatially non-overlapping brain regions based on the whole-brain parcellations by [39], the method used in [8] was employed to obtain white matter axons interconnecting those brain regions. See [8] for detail process of mapping the brain structural connectivity with diffusion MRI.

Then, from each of three subjects' DWI dataset, two percentage thresholds of 40 and 60 were respectively applied to binarize the specific percentage of top-weighted links. By investigating these three different subjects' structural datasets at different thresholds, the reproducible link community structures in the human can be demonstrated.

2.2.2.5 HCP parcellation, node time series and network matrices

The HCP-PTN dataset was the only dataset used in this work to examine the applicability of link community concept to functional network. This was not a classic voxel by timeseries functional MRI dataset, but was instead mainly preprocessed by spatial independent component analysis. The procedure for generating the HCP-PTN dataset is briefly summarized below.

1. Each of 468 subject's 15-minute resting-state fMRI data was preprocessed with the HCP processing pipelines [40], and different artifacts were eliminated by using "ICA-FIX" algorithm [28]. Then, each dataset was temporally standardized by subtracting its mean and dividing by its standard deviation. Based on these 468 subjects' preprocessed datasets, group principal components were extracted by using MELODIC's incremental group PCA (MIGP) [29]. Afterward, the following output was fed into MELODIC group-ICA tool, applying spatial ICA at several different dimensionalities (25, 50, 100, 200 and 300). A higher dimension typically indicates that the significantly modulated regions within the spatial component maps will be smaller.

2. Then, the set of group-ICA maps was projected back to each subject's resting-state fMRI dataset to derive one representative timeseries per ICA component. After concatenating all subjects' time series of each of the ICA components, the N by N group-level functional connectivity matrix was estimated by calculating the temporal correlation.

Before applying the link community algorithm to above functional connectivity matrix, various threshold levels ranging from 0.1 to 0.3 were filter this network. Then, further analyzed how the overall link community structures were reshaped.

The major weakness of spatial ICA is that this methodology cannot reveal spatially overlapping functional networks due to the constraint of statistical independence. However, this weakness can be bypassed by using link community algorithm, since it can reveal ICA maps, which belonging to multiple communities. Thus, the new pattern of resting-state functional networks were established by simply summing up all of the spatial maps with same link community memberships.

2.2.3 Force-directed algorithm

In this work, several number of link community results were displayed by using "Force-directed" algorithm which automatically arranges the two-dimensional locations of nodes, so that the crossing edges were as minimized as possible. MATLAB implementation of the following algorithm is available at MATLAB BGL toolbox written by David Gleich.

2.3 Link Community Results

2.3.1 Macaque visual cortex

Fig. 2.1.A displays the link community representation of macaque visual cortex by using force-directed algorithm. A total of 28 link communities were revealed which are depicted in different colors. Among the 28 identified link communities, 21 link communities contained less than 5 links (termed as "small link communities").

The link community outcome clearly reflected a two-stream (dorsal and ventral) hypothesis which is the most influential model of neural information processing in visual system [43]. The brain regions in the green link community are mostly located in parietal lobe. They are known as to be functionally involved in the dorsal stream ("where" pathway),

which processes the object's spatial location relative to the viewer's location. The brain regions in red community are mostly located in the temporal lobe, and they are known to be functionally related to the ventral stream ("what" pathway), which is involved in the recognition, identification, and categorization of visual stimuli. Moreover, the blue link community contains the part of the brain regions in both green and red communities that may perform a role of exchanging neural information across the dorsal and ventral streams. Identification of blue link community is remarkable since the previous node-based algorithm could not capture this structure [9]. In addition, from Fig. 2.1.A, two significant overlapping architectures nearby green and red link communities were extracted (Fig. 2.1.B and C). Region 7a (Brodmann region 7) and TF (temporal region) respectively interconnects the dorsal and ventral streams through link communities depicted with yellow/light-green and dark/sky-blue. These two selective nodes may be involved in the integration and segregation of information obtained from dorsal and ventral streams.

As another means of validating the results, the link community result was overlaid in the circuit diagram of macaque visual cortex (Fig. 2.2.A). This circuit diagram represents the hierarchical processing of visual information, which implies that the complexity of neural representation increases as the information passes from low to high level of hierarchy [1]. For instance, neurons in low visual region such as V1 selectively respond to a very simple object such as line segment of a particular orientation. In contrast, neurons in relatively higher visual region interact together to process more complex objects such as human face. This hierarchy theory is generally consistent with the result shown in Fig. 2.2.B. Global and dense green-colored link community which may in charge of simple process dominated the low level of circuit diagram. However, as the level becomes higher, multiple link communities with their respective functions were emerged to process complex visual information.

Fig. 2.1 Link communities in the structural network of macaque visual cortex. (A) Link community representation by using Forced-directed algorithm. (B) Example of overlapping architectures around region 7a and TF

Fig. 2.2 Hierarchical structure of macaque visual cortex with link community results. (A) Original circuit diagram of macaque visual cortex. (B) Modified circuit diagram of macaque visual cortex overlaid with the link community results.

2.3.2 Macaque cerebral cortex

Fig. 2.3 shows the link community results of macaque primate cortex, which were overlaid on Fig. 3 extracted from [2]. In total, 106 link communities were identified, but 94 of those were small link communities, which were depicted in gray. Similar to the previous macaque visual cortex analysis, more emphasis was placed on the remaining 12 large link communities.

Fig. 2.4 displays all of the 12 large link communities, which each contain more than 5 links. Seven of the 12 communities are confined to specific sensory systems: Map 1, 2 and 3 contain the brain regions and their connections within the visual system. In specific, map 1 is distributed in low visual regions, and map 2 and 3 cover relatively higher visual regions. Map 4 and 5 contain link communities in the frontal/limbic system. Map 6 and map 7 correspond to the auditory and the somatosensory systems, respectively. The distribution of link communities in visual and frontal/limbic systems were dense and large, but were relatively sparse and small in the auditory and somatosensory systems. Three other communities spanned across different sensory systems: maps 8, 9, and 10 show the link communities spanning from visual to frontal/limbic, auditory to frontal/limbic, and auditory to somatosensory systems, respectively. Map 11 and 12 show the link communities that were formed around specific target regions, A7b and FEF. It was inferred that these link communities may be involved in information integration and segregation.

To provide more insight into the link community results, it would be of interest to examine the brain regions that belong to multiple communities. Therefore, the number of link community memberships was evaluated for each of brain regions. Consequently, it was revealed that a set of 7 top ranked regions (A7b, A6, LIP, A46, A24, A7a, FEF) were previously reported as "hub regions", which may play an important role of integrating and segregating neural signals across different functional modules.

As another means of validating the results, functionally well-analyzed brain regions were selected from above list of hub regions and further analyzed how these regions were interconnected with four different subsystems. Region LIP (Lateral Intraparietal Cortex) was connected to visual, somatosensory and limbic systems via multiple links, which belong to different link communities (Fig. 2.5.A). Distinct link clusters were generally

consistent with the functions of LIP, which is involved in eye movement and working memory of eye movement guidance. In a similar manner, region FEF (Frontal Eye Field) was connected to visual, auditory and limbic systems via links depicted with different colors (Fig. 2.5.B). The following architecture is also consistent with the function of FEF, which is in charge of eye movement and responding to auditory stimuli.
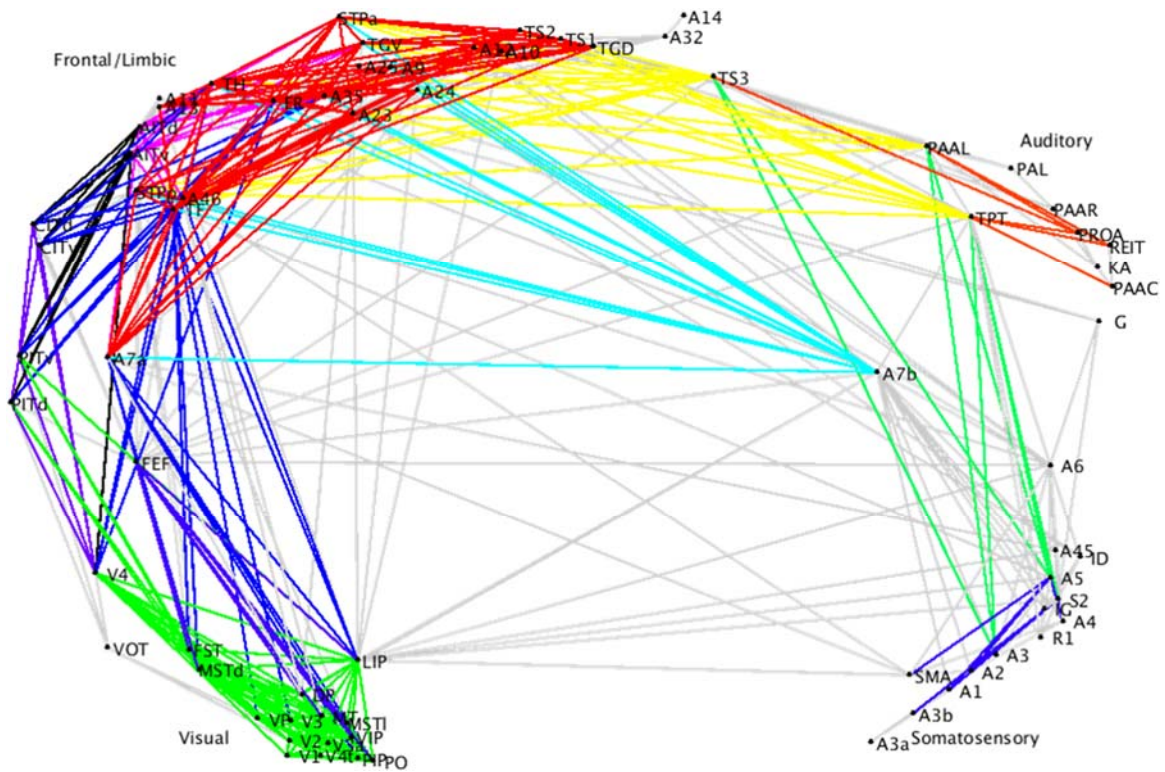
Fig. 2.3 Link communities in the macaque primate cortex
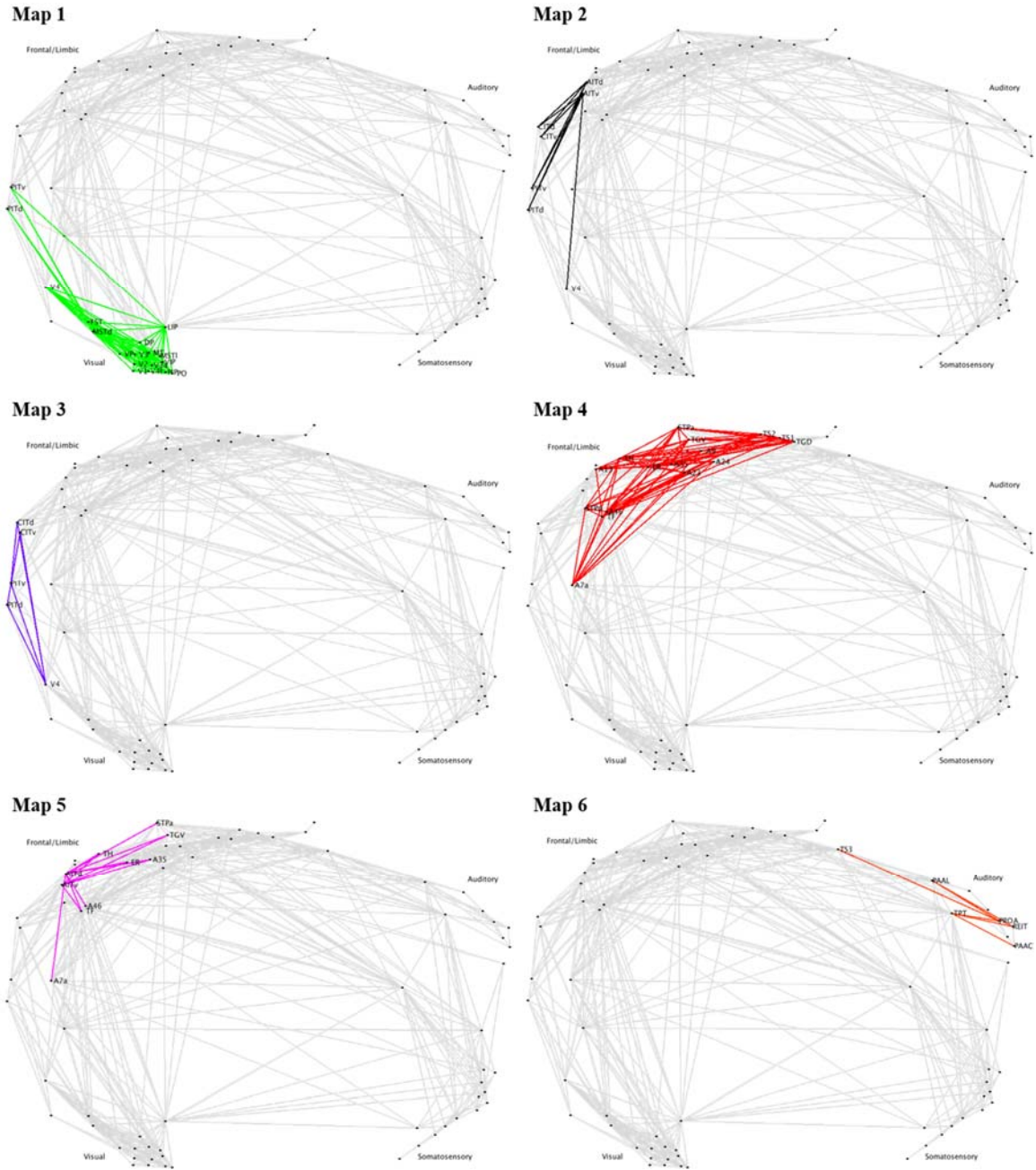
Fig. 2.4 Twelve selective link communities in the macaque primate cortex.
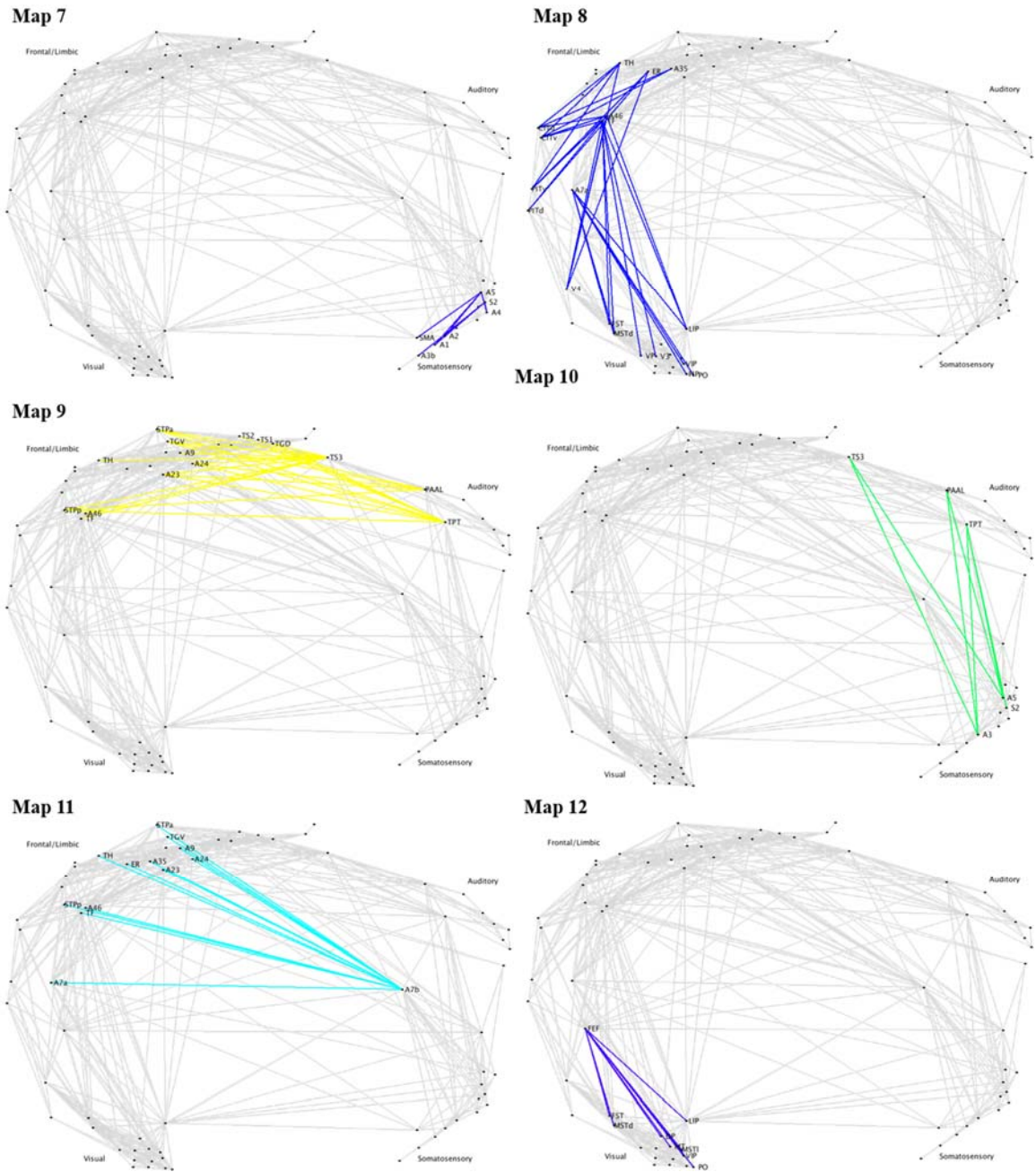
Fig. 2.4 Twelve selective link communities in the macaque primate cortex.
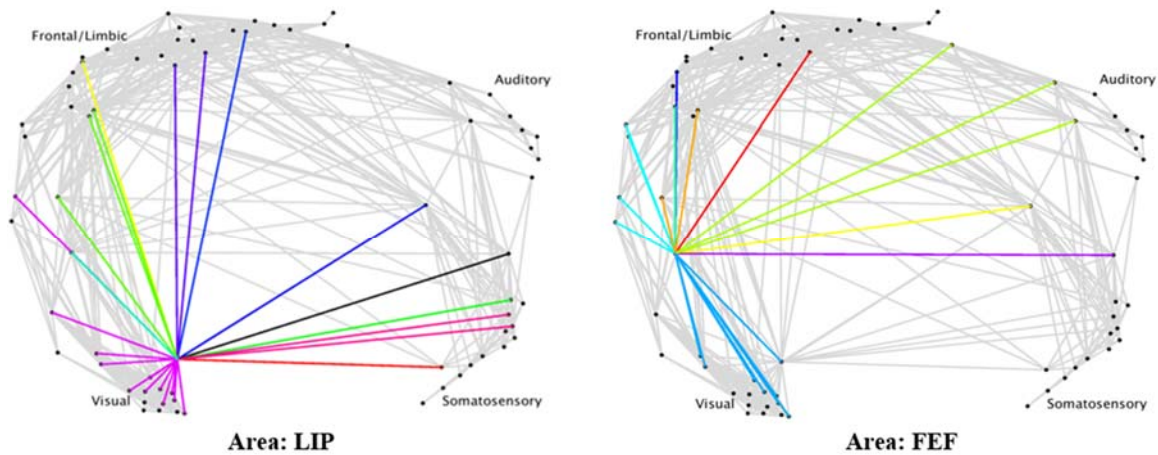
Fig. 2.5 Distinct connections from region LIP and FEF to different sensory systems

2.3.3 Whole mouse brain

Fig. 2.6 shows the link communities in the structural network of the whole mouse brain at four different thresholds (10^{-3}, 10^{-4}, 10^{-5}, and 10^{-6}). Regardless of the threshold used, 90% of the identified communities were small link communities, and were thereby excluded from this figure. By visual inspection, three reproducible overlapping architectures around the selected brain regions (CEA, PRNc and ENTI) were manually pulled out and separately displayed in Fig. 2.7.B, C and D.

The connectivity dataset used in this study also contained additional information about the corresponding major brain subdivisions of each of 213 brain regions. Initially, it was assumed that the brain regions with same community membership might correspond to similar major subdivisions. However, this assumption was not applicable to those subnetworks; in fact, the biological significance was difficult to interpret. For example, in Fig. 2.7.A, the brain regions in community 1 correspond to various major subdivisions such as cerebellar cortex, cerebellar nuclei, and medulla. In a similar fashion, the regions in community 5 correspond to the medulla, midbrain, pons, and hypothalamus.
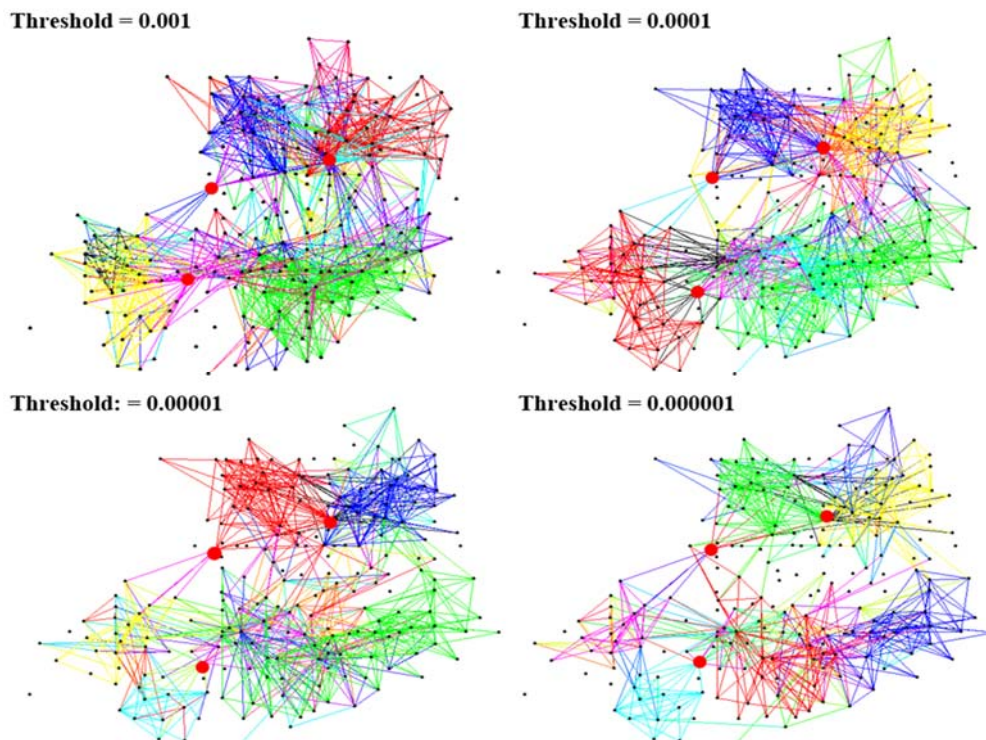
Fig. 2.6 Link communities in the whole mouse brain at four different thresholds.
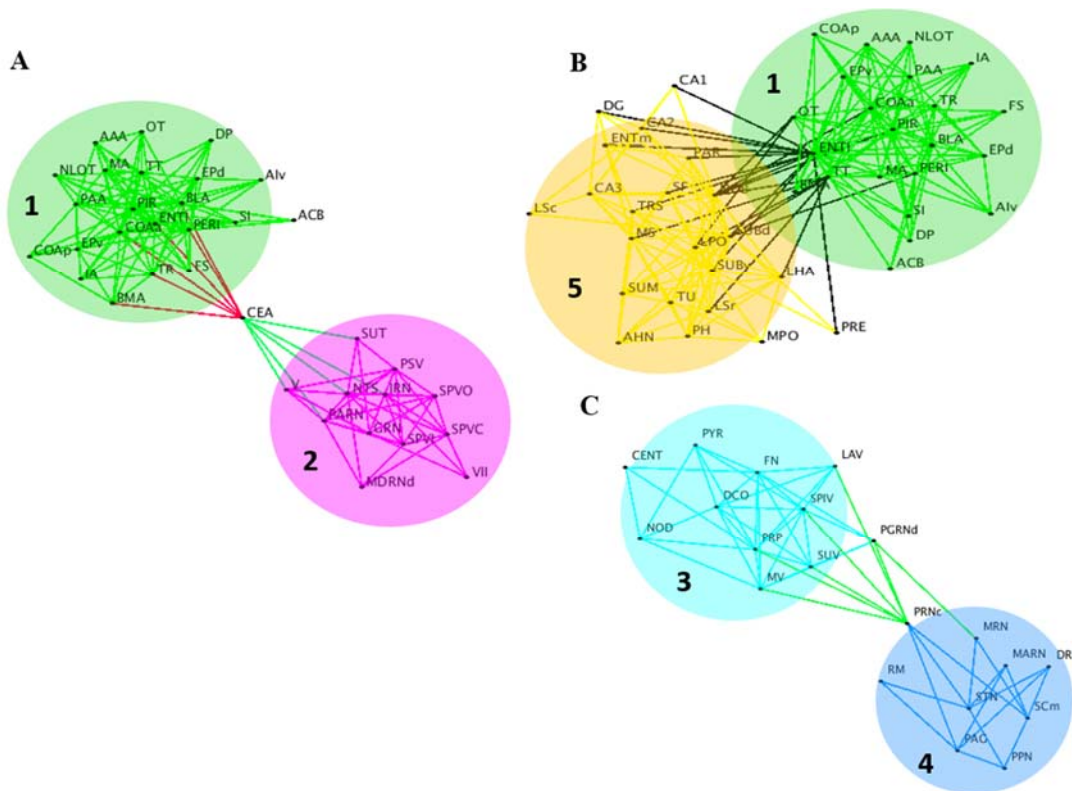
Fig. 2.7 Examples of overlapping community structures in the mouse brain with threshold equals 000001. (B), (C) and (D) Subnetworks around region CEA, PRNC and CEA respectively.

2.3.4 Human diffusion weighted imaging

Fig 2.9 and Fig 2.10 show the link community results respectively obtained from binarized three subjects' DWI datasets with 40 and 60 percentage thresholds. The total number of link communities were significantly different across three subjects. At 40 percent threshold, subjects 1, 2, and 3 contain 298, 305, and 210 link communities, respectively, and, at 60 percent threshold, subjects 1, 2 and 3 contain 317, 369 and 325 link communities, respectively. Except for the extensive green-colored link community at occipital lobe, no reproducible structures was observed.

Similar to the previous analysis, after determining the ranking of brain regions according to their link community memberships, it was identified that L.BA 6, L.BA 39 and R.BA 45 are commonly included in 10 top-ranked brain regions across three different subjects. Region BA 6 contains premotor and supplementary motor regions, which are involved in motor sequencing and movement planning. BA 39 corresponds to angular gyrus, the cortical region coordinates the information from visual, somatosensory and auditory system. BA 45 is also known as pars-triangularis, which is engaged in complex verbal functions. Based on previous studies, these three nodes were already known as hub regions [44].

Threshold = 40%
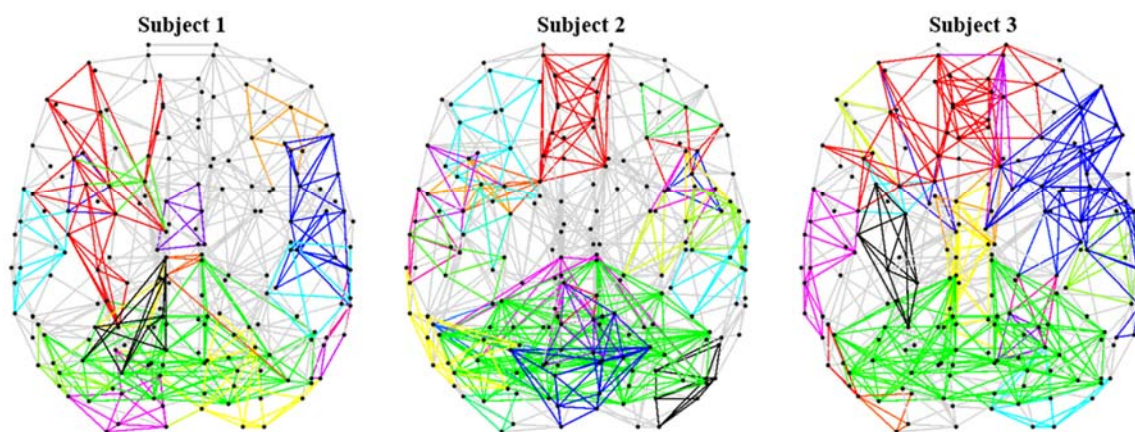
Fig. 2.8 Link communities in the three human subjects with threshold equals 60%.

Threshold = 60%
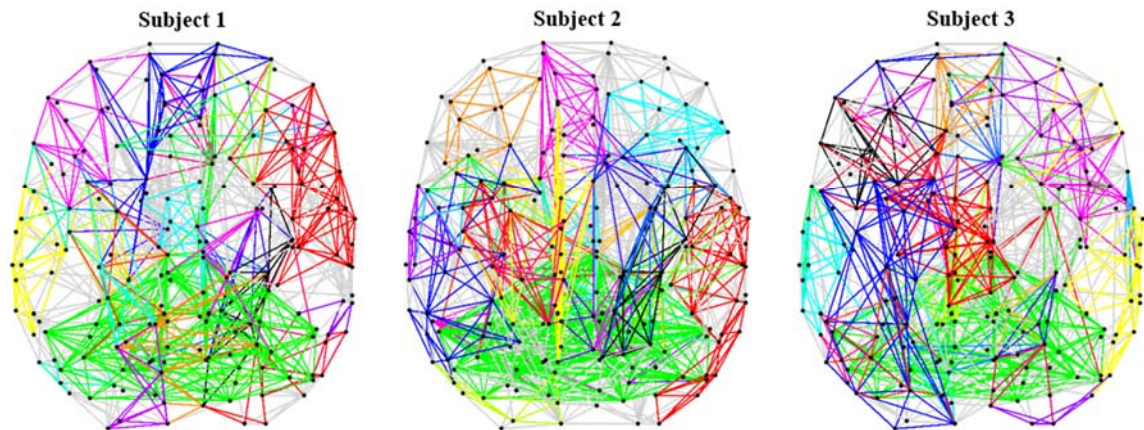
Fig. 2.9 Link communities in the three human subjects with threshold equals 40%.

2.3.5 HCP parcellation, node time series and network matrices

Fig. 2.10 displays the link communities derived from group functional connectivity matrix with threshold equals 0.25. As mentioned above, node and link on this graph respectively represents ICA map and functional connectivity strength. In total, 78 link communities were revealed, but 60 contained less than 5 links and they were not shown in this figure.

The ICA maps at same link community cover the brain regions that are functionally associated with each other. Maps in community 39 primarily contain lateral occipital cortex, lingual gyrus and pericalcarine sulcus, which are regions recognized as part of the low visual network. Maps in community 61 contain large regions of the default mode network (DMN) as described by, including angular gyrus, anterior cingulate cortex (ACC), posterior cingulate cortex (PCC), precuneus and superior frontal cortex, along with some of non-DMN regions such as calcarine sulcus and temporal gyrus. Maps in community 67 can be largely classified as visual-related, motor-related and auditory-related networks. Map 2, 5, 6, 7, 10, 13, 14, 16, 38 and 44 involve several visual regions at occipital lobe, including calcarine sulcus, cuneus gyrus, lingual gyrus, superior occipital sulcus and parieto-occipital sulcus. Map 8, 30, 29, 34 and 46 encompass central sulcus, precentral sulcus/gyrus, postcentral gyrus and subcentral sulcus/gyrus, where the primary motor

cortex and somatosensory cortex are located. Map 45 contains superior temporal gyrus and superior temporal sulcus, representing the auditory network.

Fig. 2.11 displays the selective maps of new functional networks that were based on the link community outcomes. Compared with canonical functional networks extracted with ICA [18, 25], these maps were more spatially extended and overlapping to each other. Map 1 contains occipital pole, lingual gyrus and cuneus gyrus, representing the low visual network. Map 2 encompasses the regions of map 1, along with more extended to superior parietal gyrus, inferior and superior part of occipital gyrus, which are regions part of the high visual network. Map 3 contains the regions of map 2 except part of occipital pole, with further extended to interior temporal gyrus and intraparietal sulcus. Map 4 includes the brain regions of map 3, with inferior and superior occipital gyrus and postcentral gyrus. Map 5 contains large regions related to the DMN, including PCC/precuneus, superior frontal gyrus and angular gyrus, along with part of the middle temporal gyrus and superior frontal gyrus/sulcus. Map 6 encompasses similar regions as Map 5, but more activation is observed at superior and middle frontal gyrus/sulcus. Map 7 involves widespread motor, somatosensory regions as well as auditory regions and Wernicke's region. Map 8 includes middle frontal gyrus, inferior frontal sulcus, orbital gyrus, suparmarginal gyrus and intraparietal sulcus.
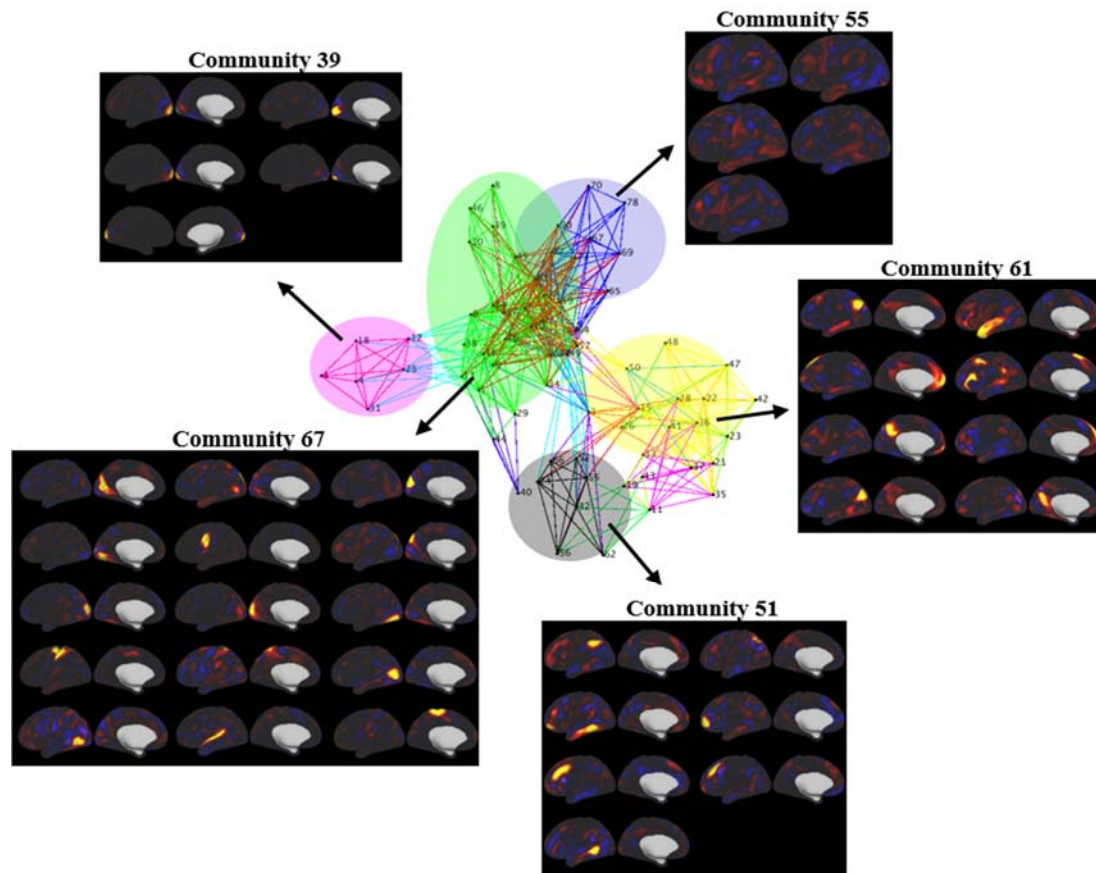
Fig. 2.10 Link communities in the PTN dataset with threshold equals 0.25.
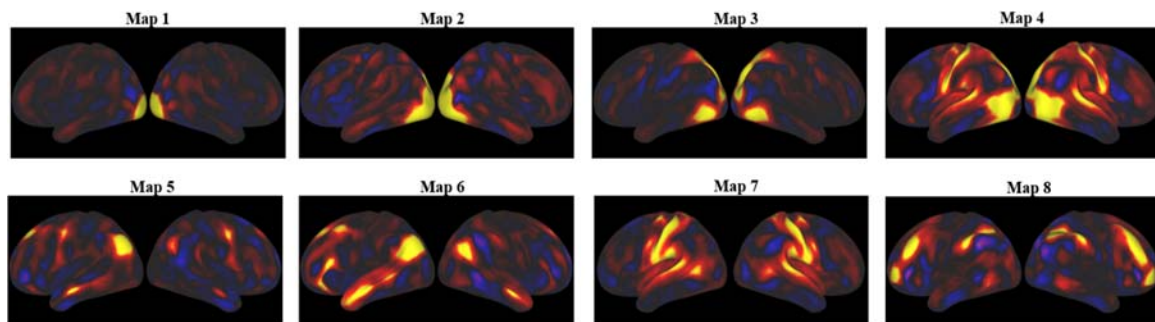
Fig. 2.11 Selective functional networks based on the link community results.

2.4 Discussion

This work demonstrates that structural and functional brain networks of both human and non-human primates contain rich and diverse link community structures. At structural network of macaque visual cortex, the link communities correspond to dorsal and ventral visual streams as well as the overlapping architectures interconnecting those two streams were identified. Furthermore, the distribution of link communities reflected the biological property of hierarchical organization of the visual system. With structural network of macaque cerebral cortex, distinct link connections formed between selected regions and different sensory systems were generally consistent with the functions of those target regions. Moreover, at HCP-PTN functional connectivity dataset, the functional networks derived based on the link communities showed distinct differences in various brain regions with respect to the canonical networks extracted with spatial ICA.

Complete biological interpretation of all link communities and their pervasive overlapping relationships is not immediately clear and be likely opened to scientific debate. At many of previous network studies, the relevance of community detection algorithms has been primarily evaluated by revealing priori known subsystems or building blocks of the networks [7, 9, 11, 12]. In contrast, the full understanding of brain's architecture is far from being completed. From the different link community results, several interesting observations were made. For instance, at overlapping architectures around region TF and 7a of macaque visual cortex, it was inferred that these two regions may be involved in integrating and segregating of neuronal information from dorsal and ventral streams. Similar inference was also made to region CEA and surrounding modules of whole mouse brain. Nevertheless, no previous studies were found to elucidate above relationships. Further researches are needed by focusing on exploring more detailed functions of specific brain regions and their individual connections.

A distinguishing feature of link community is that links from multiple communities converged at a single node, revealing a brain 'hub'. Hub regions are densely connected brain regions in the network and have been known as involved in the integration and segregation of information between different functional modules [9, 10, 13]. According to above results, the regions that previously known as hubs tended to participated in more

number of communities. Though the interpretation of the resulting link communities is not straightforward, this observation may partly support the potential application of link community concept to the brain network.

One apparent limitation of link community algorithm is that the link similarity measurement given by Jaccard index or Tanimoto coefficient only considered connected pair of links, which share a common node. However, by considering real-world network examples, it is easy to realize that the following link similarity formula is oversimplified. For instance, it has been known that visual regions V1, V2, V3 and V4 are both structurally and functionally connected to process complex visual information. Though, with the original formula, the similarity of link V1-V2 and V3-V4 results in 0, which is highly likely not to be the case. Quantifying the functional similarity of neuronal connections solely based on topological structures must be a challenging question. However, employing more biologically and mathematically reasonable formulas may identify more rich and diverse community structures in the brain.

As shown in previous results, link community algorithm tend to generate many number of small-sized communities. For instance, among the 106 link communities revealed from macaque primate cortex, 91 communities contained less than 5 links. This result is strongly contrast with the result obtained by node-based algorithm such as modularity, which only extracted 5 node communities from the identical dataset. The same phenomenon was also commonly observed at the different datasets used in this study. Since the basic principle of link community algorithm is grouping links instead of nodes, it is very natural that more number of communities should be extracted. Moreover, as a brain network exhibits a highly complex structure, it is likely that many number of communities with executing distinct functions should exist. It is difficult to validate whether this phenomenon is primarily caused by the link community algorithm itself or really reflects the biological characteristics of brain network structure. Further studies are needed to examine the potential biological relevance of such an arrangement of structural and functional connections.

Finally, it should be noted that the current work only focused on the link communities obtained at the cutoff point corresponding to maximum partition density. Given the previous studies, it has been suggested that both structural and functional brain networks exhibit hierarchical organization on multiple topological scales. Investigating the brain network as hierarchical organization has revealed several interesting features, including greater robustness, adaptivity and evolvability of network function. Until now, several methods were employed to explore this network property [33], however, none can handle overlap since hierarchical structure almost assumes disjoint community partition. Therefore, exploring the link dendrogram at multiple levels will provide more deep insights about the organization of brain network by allowing to analyze in both hierarchical and overlapping features.

3. FUNCTIONAL RELEVANCE OF SPATIAL ICA AND K-MEANS CLUSTERING

3.1 Motivation

In the absence of any overt task, spontaneous brain activity observed with functional MRI (fMRI) exhibits complex, but dynamic and systematic patterns [15, 45, 46]. It is increasingly recognized that spatiotemporal analysis of resting-state fMRI data holds great potential to uncover the full repository of large-scale intrinsic neural networks underlying various brain functions or diseases [17, 22, 28, 47].

Toward this end, existing network analysis methods are generally classified into two categories that focus on either the temporal or spatial characteristics of resting-state fMRI. As an example in the former category, cross-correlations of fMRI blood oxygenation level-dependent (BOLD) time series can be calculated to measure the degree of neural interaction between different brain locations, or the so-called "functional connectivity" [15, 17, 46]. This terminology aligns with and extends from the conventional concept that localized functional information is encoded into the temporal fluctuation in neural activity indirectly measured with fMRI. It is thus reasonable to further take the cross-correlation in fMRI time series as a putative measure of inter-regional functional relationship that defines temporally coherent networks. In contrast, methods in the second category aim to infer intrinsic functional networks by analyzing the fMRI spatial distribution at every time point. For example, spatial independent component analysis (ICA) has been widely used to decompose resting-state fMRI data into a number of spatially independent activity components [18, 23]. Despite their methodological and conceptual differences, the spatially independent maps obtained with ICA are highly consistent with the temporally synchronized networks obtained with time-series correlation analysis. In short, these two

popular methods arrive at converging findings about a common set of intrinsic functional networks.

For both temporal correlation and spatial ICA, it is implicitly assumed that functional networks are stationary over time. This assumption is perhaps questionable given recent findings suggesting that different regions may interact in a dynamic manner [48, 49]. In an attempt to capture connectivity dynamics, a common practice is to assess the time-dependent correlation by using much shorter time series extracted from a sliding window [48-50]. However, the functional interpretation of such dynamic correlational patterns is often confounded by the trivial statistical instability due to limited time samples in the presence of various noises.

Bypassing this limitation is a recently proposed notion that spontaneous brain activity is dominated by distinct co-activation patterns (CAPs) arising from discrete neural events. Applying the k-means clustering to instantaneous, whole-brain fMRI volumes serves to extract these CAPs in a data-driven manner [20]. Without assuming temporal stationarity or spatial independence, this method is conceptually meritorious and potentially allows for the discovery of spatially overlapping and temporally varying network patterns. In fact, resting-state CAPs obtained with this method have been shown to involve many novel spatial configurations notably different from those obtained with conventional methods [18, 22, 25, 51]. These CAPs form an alternative set of intrinsic networks that are potentially more informative and indicative of the brain's functional architecture.

However, it remains unclear whether and what functionally meaningful features accounts for the apparently distinct network patterns identified by these different analysis methods. It is challenging to address this question, because the precise nature, origin, and role of spontaneous brain activity still remain poorly understood. The task-free resting state is inherently lack of a specific behavioral or cognitive context. As a result, the naming and interpretation of intrinsic networks have been based mostly on their qualitative similarity to well-established functional systems, or existing activation or deactivation patterns obtained with various task paradigms that are designed to address specific aspects of brain functioning. Therefore, an objective and systematic evaluation of any conventional or

emerging resting-state fMRI analysis method should benefit greatly from directly comparing its resulting network patterns with task activation maps, for which the functional interpretation is readily available. Without such comparison, it is difficult, and perhaps speculative, to determine whether a resting-state fMRI analysis method would lead to insightful network features that truly predict and support a wide range of brain functions.

To evaluate the functional relevance is of particular importance for those analysis methods based on resting-state fMRI spatial characteristics, such as the aforementioned spatial k-means clustering and ICA methods. Note that the fMRI signal is indirectly related to regional neural activity through the neurovascular coupling, which is widely recognized and modeled as a time-domain transfer function. However, there has been no established theoretical or empirical relationship between instantaneous fMRI volumes and underlying whole-brain neural activity patterns. When the functional relationship between regions is defined merely based on the instantaneous image intensity distribution in the whole brain, it is worth being cautious in interpreting the so-defined intrinsic networks that support spontaneous hemodynamic patterns as those that support neuronal activities and interactions. In addition, the k-means clustering and ICA methods are very similar in that they both treat instantaneous fMRI volumes as independent high-dimensional input data, from which a given number of multivariate features are learned in order to efficiently represent (or reconstruct) the input data [52]. As two off-the-shelf learning algorithms, these two methods use similarly constrained optimization schemes that both encourage the learned features to be sparse. Given these considerations, the theoretical essence of these two methods may not be as distinct as is implied by the underlying notion that motivates their development and application to resting-state fMRI network analysis.

In this work, functional relevance of the resting state network patterns revealed by using the spatial k-means clustering and spatial ICA was explored. For this purpose, either the k-means clustering or spatial ICA was employed to 15 subjects' resting-state fMRI data, and a large and comprehensive set of neuroimaging-based task activation data that have been previously published and stored in the BrainMap database. It was attempted to match the resulting resting-state CAPs to the corresponding task-based CAPs, and also to match the resulting resting-state ICA components to their task-based counterparts in a similar way

as in previously published studies. The correspondence of the best paired network patterns during rest and task conditions served as a quantitative measure of the functional relevance of the different sets of intrinsic network patterns, providing objective evaluation and proper interpretation of the methods that gave rise to such networks.

3.2 Materials and Methods

3.2.1 Resting-state fMRI

10-minute resting-state fMRI data were acquired from each of 16 subjects. The pre-processed resting-state fMRI dataset was resampled from $2 \times 2 \times 2$ mm^3 to $3 \times 3 \times 3$ mm^3 spatial resolution which was consistent with the Montreal Neurological Institute (MNI) standard brain. After the data was spatially smoothed with a Gaussian kernel (FWHM = 6), each voxel's time series was temporally normalized by subtracting its mean and dividing by its standard deviation. Global signal regression was not applied in this work since it may force artificial anti-correlation between different brain regions [53]. Finally, all 16 subjects' datasets were reshaped in 2-D (space by time) and temporally concatenated for the subsequent multi-subject analysis.

3.2.2 Task activation maps

As a means of comparison with the resting-state fMRI, Brainmap database (http://www.brainmap.org) was employed, which publicly shares the results of a large number of task activation studies. At the time of current analysis, the Brainmap database included the results from ~1900 published functional neuroimaging articles, which resulted in ~9100 individual activation images. Multiple experimental conditions were employed in each study; the spatial distributions of the following activation results were stored as 3-D foci in Talairach space depicting the statistically-significant local maxima. Then, from each of activation images, a set of "pseudoactivation" images were recreated by filling an empty brain image with points corresponding to the 3-D foci and spatially smoothed with a FWHM 12mm Gaussian kernel. Although the actual spatial information of the original activation has not been preserved, this smoothing extent is a reasonably close match to that applied as data preprocessing in most FMRI activation studies [54]. Then, the resulting

~9100 activation images were concatenated to generate a 2-D dataset so that the first dimension is space the second is experiment ID. In case of spatial ICA analysis, the above dataset was temporally demeaned and variance normalized since this step was necessary to perform principal component analysis [55]. However, in case of k-means clustering, the above-mentioned steps were not applied to preserve as much of the original information as possible of the dataset itself.

3.2.3 Extraction of co-activation patterns

The rest and task co-activation patterns were derived by using the k-means clustering algorithm, which had already been implemented at previous work [20]. After the above-mentioned preprocessing steps were applied to each of rest and task dataset, k-means clustering technique was subsequently applied to classify fMRI time points (experiment ID in case of Brainmap) into k clusters based on their similarity of spatial activity distributions. Then, the maps assigned to the same cluster were simply averaged, resulting in k maps which are called co-activation patterns (CAPs) referred to earlier. The number of rest and task CAPs, k, was set to 25 in this work.

3.2.4 Extraction of spatially independent components

As a means of comparison with spatial CAPs, the procedure of group-level spatial ICA with temporal concatenation was implemented [18]. The following analysis was applied independently for the identical rest and task datasets utilized in CAPs analysis. After reducing the concatenated dataset into n strongest spatial eigenvectors by using PCA, the resulting eigenvectors were fed into Infomax ICA to extract the n most representative functional networks. At the end, the resulting ICA component maps were spatially normalized by dividing its magnitude. The number of ICs, n, was also set as 25 to match with the number of CAPs.

3.2.5 Extraction of temporally constrained resting-state CAPs

Temporal sparsity was constrained to resting-state fMRI dataset by following procedure: after temporally normalizing each subject's resting-state fMRI dataset, a level of threshold was defined based on the standard deviation (STD) and identically applied to each voxel's

time series such that the original signal intensities were preserved at the time points in which the absolute values exceeded this threshold, while the remaining points were simply padded with 0 (Fig. 3.1). Then, same as previous analysis, after temporally concatenated all subjects' datasets, the identical procedure of k-means clustering was employed again to the following grouped dataset, which resulted in another set of 25 temporally constrained resting-state CAPs (termed as rest tcCAPs). In this analysis, threshold level was progressively increased from 0.5 to 2.2 SD, then the resulting rest tcCAPs were compared with respect to the rest ICs, task ICs and task CAPs obtained at above procedures.
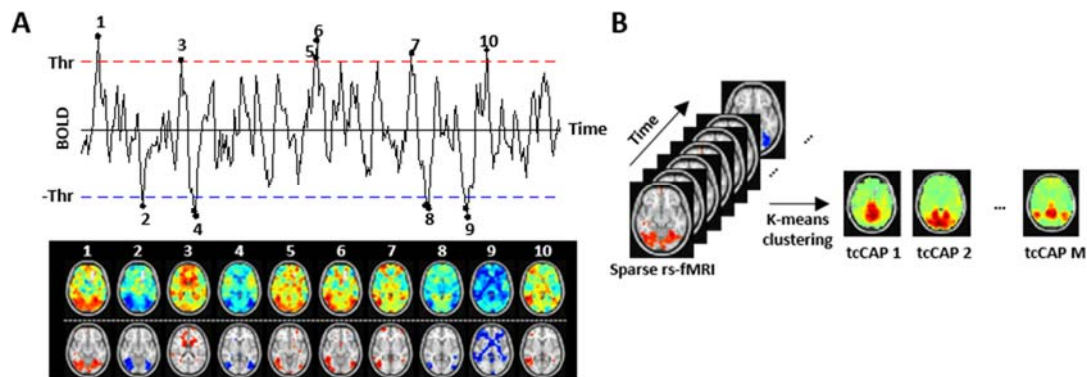
Fig. 3.1 Procedure of obtaining temporally-constrained CAPs

3.2.6 Spatial similarity of different spatial maps

Based on the above procedures, 5 different sets of spatial maps each with 25 components were extracted; rest & task ICs, rest & task CAPs and rest tcCAPs. The spatial maps from these five different groups were matched with each other based on their spatial similarity using Pearson's correlation of the unthresholded maps. Since all components of one set were paired with the only one best match from a second set, there arose a multiple pairings problem, in which some of the same spatial maps from the first set may highly correlated with two or more spatial maps in the second set. In order to overcome with this multiple pairings problem, greedy algorithm was used to discover the optimal pairs between different groups. In summary, after randomly permuting the indices of spatial maps of the first set, the maps from the first set were sequentially paired with the highest correlated map of the second set. In each iteration, the most strongly paired map of the second set is removed from consideration for future pairings with the first set. After 300

iterations, one-to-one pairs between any two sets was obtained. Then, the spatial correlations between the pairs were summed for each iteration; the set of pairs which maximized the sum of spatial correlations were selected as the optimal pairs.

3.3 Results

3.3.1 ICs and CAPs from resting-state fMRI and Brainmap

The sets of rest and task functional networks which were respectively extracted by spatial ICA and k-means clustering were shown in Fig. 3.2.A and C. With a naked eye, several ICs from rest and task datasets were unambiguously matched to each other, showing strong spatial cross-correlations (Fig. 3.2.B), as previously reported [22, 56]. In addition, the following rest and task ICs strongly resembled the canonical functional networks shown in previous studies [18, 25, 51]. Given several examples, map A contains striate and parastriate area, which are areas recognized as part of the visual cortex. In comparison with map A, map B covers more lateralized visual areas, including peristriate area and lateral and superior occipital gyrus. Map D contains the superior temporal gyrus, representing the auditory network. Map E encompasses the thalamus region, a small symmetrical structure situated between the cerebral cortex and the midbrain. Map F includes prefrontal, anterior cingulate and posterior cingulate cortex (ACC/PCC), inferior temporal gyrus, and the superior parietal region, known as the default mode network (DMN). Map G contains superior parietal cortex, occipito-temporal and precentral regions, representing the visuospatial network.

In comparison with rest and task ICs, only coarse correspondences were observed between rest and task CAPs (Fig. 3.2.C). The resulting rest CAPs showed significant differences with the conventional networks extracted with spatial ICA by covering more extended brain regions and spatially overlapping to each other (Fig. 3.2.D). In addition, these rest CAPs contained both strong co-activation and co-deactivation in one map. In contrast, the resulting task CAPs were more confined to specific sets of regions. Interestingly, both spatial ICA and k-means clustering produced nearly identical spatial patterns for task activation data, showing significant spatial cross-correlations (Fig. 3.3.A and B).

Brief descriptions of the selective rest CAPs are following: Map A contains low and high visual regions with further extended to motor and auditory networks. In specific, lateral occipital cortex, pericalcarine sulcus and cuneus, fusiform, precentral and postcentral gyrus were included. Map B covers large areas of the DMN, including PCC, ACC and angular gyrus, with anti-correlation at regions of low visual, visuospatial and ventral networks. Map D contains medial and lateral visual regions, including lateral occipital cortex, cuneus and lingual gyrus and pericalcarine sulcus. Map E shows particularly global spatial distribution, containing insular cortex, PCC, precentral, postcentral, superior frontal and temporal gyrus and paracentral sulcus. Map J contains large areas of the motor network, including precentral, paracentral and postcentral gyrus, along with several frontal regions.

3.3.2 Rest tcCAPs at different temporal sparsity levels

The list of correlation matrices in Fig. 3.4 denoted how the network patterns of rest tcCAPs were altered with respect to itself as well as rest ICs and task CAPs at four different selective STD thresholds (0, 0.6, 1.2 and 1.8). The correlation matrices at three different rows sequentially represent spatial autocorrelations of rest tcCAPs, spatial cross-correlations of rest tcCAPs and rest ICs, and spatial cross-correlations of rest tcCAPs and task CAPs. All of cross-modality correlation matrices were rearranged based on correlation strengths.

At relatively low threshold levels (0 and 0.6 STD), the rest tcCAPs showed global autocorrelation with itself, since their network patterns were still spatially extended and overlapping to each other. In contrast, at relatively higher threshold levels (1.2 and 1.8 STD), they were decorrelated with itself, since their extents of co-activation and co-deactivation were more confined to specific sets of regions. Strong anti-correlations were observed at all threshold levels due to the similar rest tcCAPs with reverse polarity, which might be occurred due to the methodological property of k-means clustering.
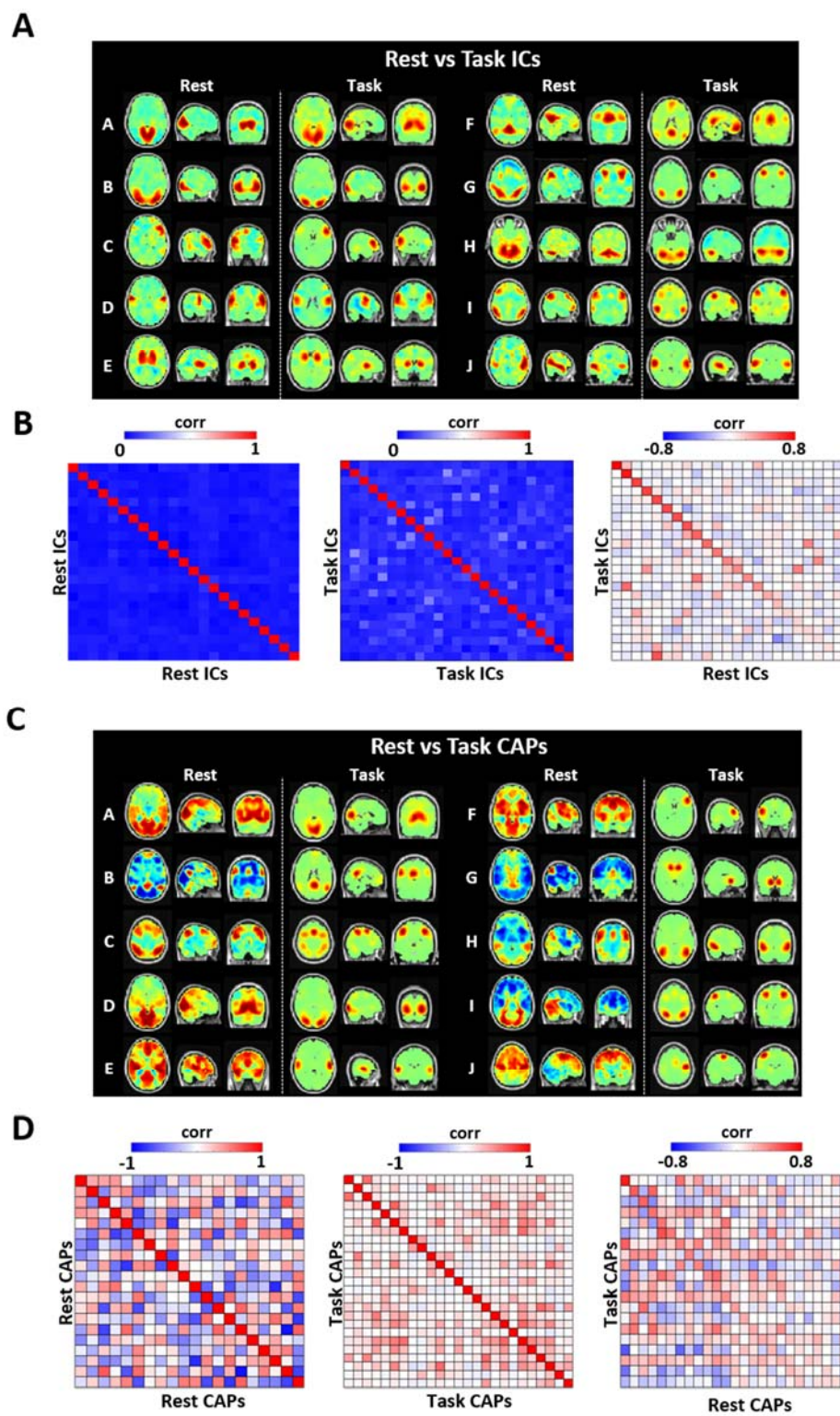
Fig. 3.2 Comparison of ICs and CAPs from resting-state fMRI and Brainmap.
(A) Network patterns of rest and task ICs. (B) Spatial similarity between rest and task ICs.
(C) Network patterns of rest and task CAPs. (D) Spatial similarity between rest and task CAPs.
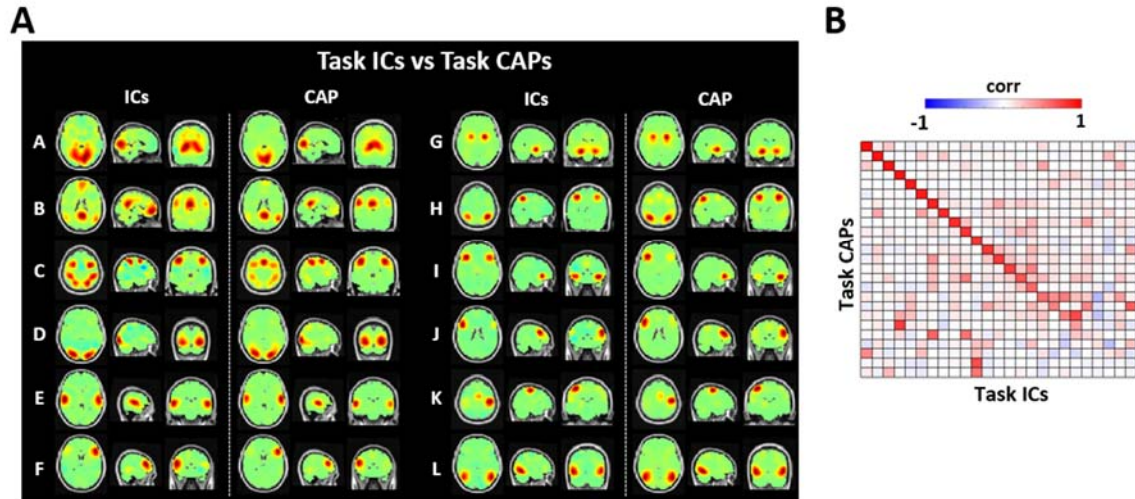
Fig. 3.3 Comparison of task ICs and task CAPs from Brainmap. (A) Network patterns of task ICs and task CAPs. (B) Spatial similarity between task ICs and task CAPs.

In comparison of rest ICs and task CAPs, at low threshold levels (0 and 0.6 STD), no strong correspondence was observed with rest tcCAPs. Instead, multiple rest tcCAPs showed global spatial cross-correlations with several rest ICs as well as task CAPs, and vice versa. However, as threshold levels became higher (1.2 and 1.8 STD), the rest tcCAPs progressively achieved strong one-to-one or two-to-one correspondences with both rest ICs and task CAPs. As shown in the correlation matrices, the diagonal elements became stronger and the off-diagonal became sparser. Two-to-one pairs were achieved, since similar co-activation (red) and co-deactivation (blue) patterns were matched with only one ICs or task CAPs.

In this analysis, the temporal sparsity was set to 2 STD after comparing the outcomes of ranging from 0 to 2.2 STD. At the range between 0 to 1.2 STD, several rest tcCAPs still showed extensive co-activation and co-deactivation patterns which merely matched with both rest ICs and task CAPs. In contrast, at range between 1.3 to 2.2 STD, the rest tcCAPs were mostly similar to each other, but, threshold with 2 STD led to the highest number of matched network patterns with rest ICs as well as task CAPs.
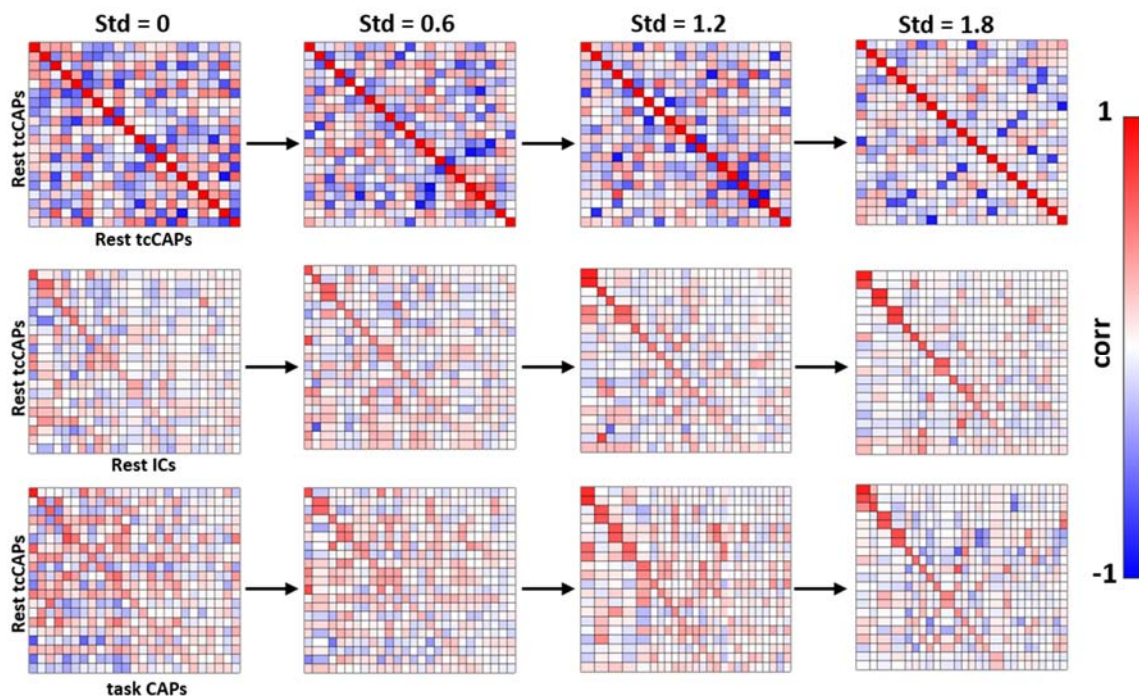
Fig. 3.4. Shape alteration of rest tcCAPs at different threshold levels.

3.3.3 CAPs from temporally constrained resting-state fMRI

When the temporal sparsity was constrained to resting-state fMRI datasets with STD = 2, k-means clustering generated remarkably different CAPs compared to those yielded without any constraint conditions (Fig. 3.5.A). The resulting rest tcCAPs were more spatially confined to specific sets of regions and the strong co-activation (red) and co-deactivation patterns (blue) were disassociated into individual maps. In addition, a number of rest tcCAPs showed similar co-activation and co-deactivation patterns with reverse polarity. These rest tcCAPs also showed a fair degree of correspondence with both rest ICs and task CAPs obtained at previous steps. Between rest tcCAPs and rest ICs, 16 rest tcCAPs were closely matched with 12 rest ICs, and, between rest tcCAPs and task ICs, 13 rest tcCAPs showed correspondence with 10 task CAPs. As previously mentioned, the similar co-activation and co-deactivation patterns were paired with only one rest IC or task CAP.

Set A to H display network patterns that were commonly shown in rest tcCAPs, rest ICs and task CAPs. Set A, C and E are primarily associated with visual processing, which correspond to primary visual, visuospatial and high visual network respectively. Set

B contains PCC, ACC and angular gyrus, known as default mode. Set D contains the superior temporal gyrus, where the primary auditory cortex is located. Rest tcCAPs also contained a network pattern with fine-scale structure such as thalamus, which is shown in set F. Set G encompasses medial-frontal areas, which mainly includes anterior cingulate and paracingulate cortex. The network patterns in set I to L were commonly shown in rest tcCAPs and rest ICs. Set I shows a network pattern that predominantly involves dorsal parietal and lateral prefrontal cortex. Set J contains precentral, paracentral and postcentral gyrus, where primary somatosensory and motor cortex were located. Set L shows a cerebellum, which plays a major role in motor control. Set M shows a network pattern that was observed in rest tcCAPs and task CAPs. This map contains rostral and caudal middle frontal regions, which are part of the executive control network.

Fig. 3.5 Comparison of rest tcCAPs, rest ICs and task CAPs. (A) Network patterns of rest tcCAPs, rest ICs and task CAPs. (B) Spatial similarity among rest tcCAPs, rest ICs and task CAPs.
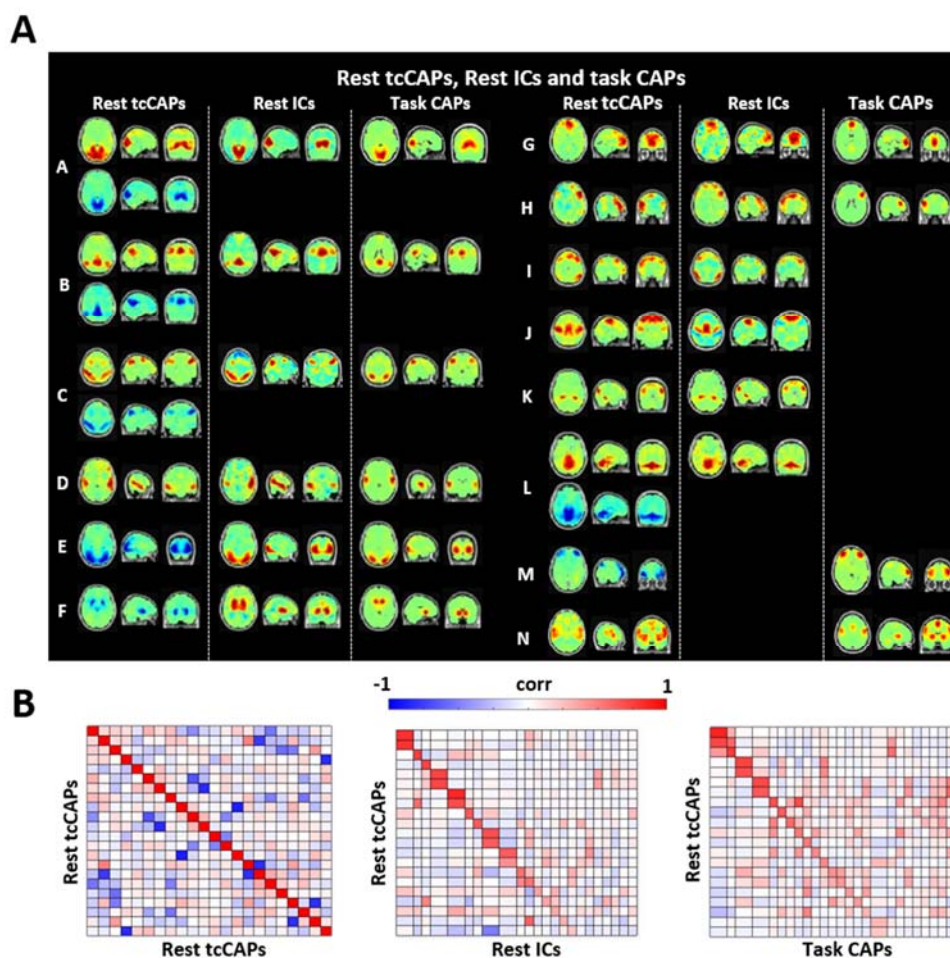
3.4 Discussion

In this study, two representative multivariate data-driven analysis of resting-state fMRI: spatial independent component analysis (ICA) and k-means clustering were thoroughly investigated in order to elucidate their functional relevance and similarity of theoretical basis. To this end, these two methods were respectively applied to 15 subjects' resting-state fMRI and a large set of task activation maps published in the BrainMap database. Then the sets of resting-state and task-based independent components (ICs) and co-activation patterns (CAPs) were compared against to each other. Initially, k-means clustering yielded rest CAPs, which showed distinct differences with respect to both rest ICs and task CAPs. However, after constraining the temporal sparsity of resting-state fMRI, the network patterns of temporally-constraint CAPs (tcCAPs) closely resembled those of rest ICs as well as task CAPs. Based on these observations, it was inferred that these two algorithms may not distinct from each other, but converge in the discovery of similar resting-state networks.

As mentioned previously, in the field of machine learning, k-means clustering and ICA have been considered as conceptually and theoretically closely related methodologies [52]. These two methods explicitly or implicitly impose very similar "sparsity" constraint to the input dataset, so that the basis function learned by k-means clustering closely resemble those obtained with ICA. Consequently, it was not so surprised that both k-means clustering and sICA extracted similar functional networks from task activation dataset. In contrast, the differences observed in the resting-state fMRI were conflict with the above claim.

In order to investigate the reason why both k-means clustering and spatial ICA yielded almost identical functional network patterns from task activation datasets, but not from resting-state fMRI, a careful attention was initially paid to the characteristic of task activation data itself. As mentioned earlier, this dataset is a "pseudo-fMRI data", which was created based on the 3-D activation foci in a few brain locations with applying a Gaussian kernel [54]. Consequently, the data itself implicitly had a strong sparsity constraint in a way that only specific regions around the activation points contain strong signal intensities. In contrast, resting-state fMRI datasets used in this study showed a

certain level of BOLD intensities around all of the regions. Based on this observation, a specific level of temporal sparsity was applied to resting-state fMRI by thresholding voxel timeseries. This approach was inspired by the notion of point process analysis (PPA) [57], which indicates that the co-activation and co-deactivation of functionally related brain regions is dominated by instantaneous, discrete neuronal events rather than the result of a continuous process. Therefore, the brain activity patterns during resting-state are contributed not only by peak activations or deactivations, but also may be contributed by the signals occurred during intermediate phases in transition to either activation or deactivations. Therefore, removing these intermediate phase signals by thresholding voxel time series may be a more reasonable approach for the subsequent k-means clustering analysis.

One obvious difficulty of this work is the selection of a suitable threshold level, so that the "unnecessary" signals occurred during the intermediate transition are efficiently removed. The current study was initialized from the claim that both spatial ICA and k-means clustering might yield similar functional networks from the resting-state fMRI [52]. Accordingly, the optimized threshold was chosen by comparing the network patterns of tcCAPs with those of rest ICs. However, it is hard to decide whether this threshold level eliminated the important signals solely caused by the sporadic neuronal events. Focusing only on meaningful neural signals may increase the specificity and detailedness of resting-state fMRI analysis. The enhancement of the proposed method can be achieved through the differentiation of those critical signals.

A main advantage of the presented method is that it requires only few assumptions and transformation of the data. The method presented here resembles spatial ICA in the sense that both are spatial-domain methodologies, which regard fMRI volumes as the basic units of analysis. However, spatial ICA requires additional processing step of the fMRI data itself such as whitening, and further assume that component maps are statistically independent to each other [55]. In contrast, the classification procedure executed by k-means clustering does not need any transformation of the input data, and the resulting CAPs are just the simple average of fMRI volumes with same cluster memberships. Accordingly, the characterization and interpretation of resulting CAPs are more straightforward.

An interesting observation made from rest tcCAPs is the similar co-activation (red) and co-deactivation maps (blue) with reverse polarity (Fig. 3.5). In case of spatial ICA, polarity of component maps does not bear any significant meaning, since the fMRI volumes are represented as the weighted combinations of a set of ICs [23, 55]. It is fairly obvious that these reverse-polarity maps were primarily caused by the methodological property of k-means clustering, since its' optimization criteria always regarded the instantaneous activity patterns with opposite polarity as "very dissimilar" and assigned those patterns to different clusters. As shown in Fig. 3.5, some of rest tcCAPs such as default mode network (map B) showed both strong co-activation and co-deactivation. In comparison, the others such as thalamus network (map F) and motor network (map J) only showed either one of them. However, the biological origin of this observation remains unclear, and may need further exploration.

Finally, an obvious difficulty of k-means clustering analysis is a selection of k, a number of maps to be extracted from the input fMRI dataset. In this study, the value was set arbitrarily as 25 after comparing the tcCAPs obtained with k equals 20, 30, 35 and 40. In general, a higher k value means that the significant co-activated and co-deactivated regions will be smaller. In order to set k in more objective way, different methods, including the Silhouette coefficient and Elbow method were employed, but their suggested optimum values were fairly inconsistent with each other. This may be due to the dissimilarity among the CAPs is likely to have a skew distribution, with a portion of the CAPs being much closer to one another than to others, which increases the difficulty in finding a clear division.

4. INTRINSIC FUNCTIONAL NETWORKS WITHIN VISUAL CORTEX SUPPORTS NATURALISTIC VISUAL PERCEPTION

4.1 Motivation

In the absence of sensory stimulation or any overt task, spontaneous brain activity observed with functional MRI (fMRI) exhibits a rich and diverse spatiotemporal patterns. Specifically, low-frequency (around 0.01 - 0.1 Hz) blood-oxygenated-level-dependent (BOLD) signals obtained during resting-state fMRI are temporally coherent across spatially distributed brain regions, which are also termed as resting-state networks (RSNs) [15, 46]. Although being called as RSNs, these networks show similar patterns as those observed during various task activations, including visual, auditory and motor networks [22, 56]. In addition, these patterns are reproducible across different experimental sessions, healthy individuals and even in non-human primates [25, 26].

Most previous analyses of resting-state connectivity have primarily focused on the large-scale organization of whole brain [13, 18, 19, 20, 21, 22]. However, the growing body of evidences have suggested that the spontaneous brain activity can exhibit systematic spatiotemporal organizations within the specific sensory systems, and even within individual cortical regions [32, 33]. Characterizing this fine-scale connectivity patterns is especially important, since this information can provide more rich insights into the relationship of the spontaneous activity and the underlying functional architectures that support diverse brain's functions. In consequence, many recent studies have comprehensively investigated such patterns of different sensory systems. For instance, from experiment using micro-electrocorticography (μEcoG), the correspondence of sensory maps and spontaneous field-potential signals was demonstrated in the macaque auditory cortex [58]. In addition, a high field 9.4T fMRI study identified the

correspondence of spontaneous activity patterns and the somatotopic arrangement of the monkeys' somatosensory cortex [59] Moreover, the study with single unit electrophysiology reported a high level of correlation in spontaneous spiking activity between neurons with similar tuning properties [60].

A particularly convenient brain sensory system to investigate such fine-scale connectivity patterns is the human visual cortex. This is because its functional architecture has been widely understood and solidly established with retinotopic mapping [61, 62]. After the finding of robust functional connectivity in human visual cortex without sensory stimuli [63], recent studies have explored the relationship between spatiotemporal patterns of spontaneous activity and those evoked by a standard phase-encoded retinotopic mapping paradigm. Accordingly, several studies reported that the spontaneous signals recorded in similar receptive fields tend to fluctuate in a correlated manner [64, 65]. In addition, the connectivity patterns formed nearby low visual region (V1-V3) followed the eccentricity organization [64-66]. Recently, a similar analysis was conducted by using naturalistic visual stimuli [66]. Then, it was demonstrated that the connectivity patterns that appear during rest better reflected the naturalistic activations than artificially controlled phase-encoding paradigm.

However, most of these studies have primarily focused on the spontaneous connectivity patterns in low visual region, so that the similar information in high visual region is largely missing. Toward filling this gap, the current study characterized the complex neural activity patterns within whole visual cortex in three different aspects: spatial organization, inter-subject reproducibility and functional relevance to naturalistic vision. To this end, two different 3T functional MRI datasets were employed. Dataset 1 included 45 subjects' resting-state fMRI from the Human Connectome Project (HCP). Dataset 2 contained 3 subjects' fMRI obtained under the naturalistic movie stimuli. The fine-scale connectivity patterns during resting-state and movie conditions were mainly obtained by using spatial independent component analysis (sICA) instead of seed-based correlation. This multivariate statistic may potentially provide more in-depth connectivity information than univariate analysis by simultaneously identifying the co-activations of multiple brain regions.

4.2 Materials and Methods

In this study, two different 3-T fMRI datasets were used. For the dataset 1, the high quality, high-resolution resting-state fMRI was downloaded from the Human Connectome Project website (HCP; http://www.humanconnectome.org). For the dataset 2, functional MRI was acquired from each of three subjects under natural movie viewing. Detailed information about above two datasets, including subjects' information, MRI parameters, experimental paradigm and data pre-processing steps will be discussed below.

4.2.1 HCP resting-state fMRI

Subjects

From the HCP website, 45 human subjects' resting-state fMRI (26-35 years of age; 19 males, 26 females) were randomly selected and utilized in the current study [67]. These selected participants were unrelated to each other, healthy individuals who doesn't experience any significant neurological or psychological diagnosis. All subjects gave informed consent as approved by the Washington University in St. Louis institutional review board.

MRI Parameters

All experiments were conducted in a 3T MRI system (Skyra, Siemens, Germany) using a Siemen's standard 32-channel head coil. T1-weighted structural images were acquired with a magnetization-prepared rapid gradient-echo (MP-RAGE) sequence (TR = 2400 ms, TI = 1000 ms, TE = 2.14 ms, FA: 8°, $0.7 \times 0.7 \times 0.7$ mm voxels and FOV = 208 $\times$ 180 mm). FMRI data were acquired using a gradient-echo EPI sequence (TR = 720 ms, TE = 33.1 ms, FA = 52°, $2 \times 2 \times 2$ mm voxels and FOV = 208 $\times$ 180 mm).

Resting-state fMRI acquisition

For each of a participant, resting-state fMRI were acquired in four runs of 14 minutes and 33 seconds (1200 time frames) each, two runs in one session and two in another session, with eyes open and fixated on a cross-hair presented on a dark background. Within each session, oblique axial acquisitions alternated between phase encoding in a right-to-left direction in one run and in a left-to-right direction in the other run. In the

current study, only the data acquired during the first session with a left-to-right phase encoding direction was used.

Pre-processing

The resting-state fMRI datasets published by HCP already contained basic, but necessary preprocessing steps [40], so these preprocessed datasets were directly used in this study. FMRI preprocessing steps included 1) gradient nonlinearity distortion correction, 2) FSL's FLIRT motion correction, 3) FSL's top up distortion correction, 4) registration to the T1-weighted structural MRI scan, 5) spline resampling from the original EPI frames to FSL MNI152 2 mm space using FSL FNIRT and vi) image intensity normalization to mean of 10000 and bias field correction. Then, the following preprocessed fMRI dataset were projected from FSL MNI152 space onto the FreeSurfer surface space (1 mm mesh), spatially smoothed using a 6 mm FWHM Gaussian kernel and downsampled to a 4 mm mesh. Along with the above steps, the surface-based fMRI dataset was temporally de-trended by using a fourth-order polynomial function, band-pass filtered at 0.01-0.2 Hz, and temporally standardized by subtracting its mean and dividing by its standard deviation.

4.2.2 Natural movie experiment

Natural Movie Stimuli

A color natural movie (8-minutes in length) was constructed by temporally concatenating multiple video clips (8-12 seconds in length), which were downloaded from the Videoblocks website (https://www.videoblocks.com/). The clips used in this experiment contain everyday objects, including airplane, bird, car, face, flower, fruit, insect, animal, people, ship and various natural scenes. The sequence of stimulus was created by randomly drawing up clips from the entire set. For the purpose of this study, no sound was inserted into the movie.

MRI Parameters

The experimental protocol was approved by the Institutional Review Board at Purdue University. Functional scans were conducted in a 3T MRI system (Signa HDx, General Electric, Milwaukee). A 16-channel receive-only surface phase-array coil (NOVA Medical, Wilmington) was used throughout every experiment. T1-weighted structural images were acquired with a spoiled gradient recalled acquisition (SPGR) sequence (256 sagittal slices with 1 mm thickness and 1×1 mm^2 in-plane resolution, TR/TE = 5.7/2ms, flip angle: 12°). FMRI data were acquired using a standard single-shot, gradient-recalled (GRE) EPI sequence (TR = 2s, TE = 35ms, FA = 78°, $3.5 \times 3.5 \times 4$ mm voxels and FOV = 220×220 mm).

FMRI acquisition

Three human subjects (23-26 years of age; 3 females) who were healthy and had normal vision participated in this study. Each subject underwent four repeated sessions of 8-minutes natural movie presentation per day, in total of 12 sessions were acquired across three different days. Every session started with a blank gray screen presented for 12 seconds, followed by the movie presented for 8 minutes, ended with the blank screen again for 10 seconds. The movie was presented using the MATLAB-based Psychophysics Toolbox [68, 69], and it was delivered to subjects through a binocular goggle system (NordicNeuroLab, Norway) mounted on the head coil. During the movie presentation, all subjects were instructed to fixate at a cross-hair (24 pixels in width and height) presented at the screen center.

Pre-processing

MRI and fMRI data were preprocessed by using FSL [70], AFNI [71] and MATLAB functions developed in house. In summary, 1) T1-weighted anatomical images were non-linearly registered to the MNI brain template, 2) T2*-weighted functional image series were corrected for slice timing, registered to the first volume within each series to account for head motion, 3) masked out non-brain tissues, 4) aligned to the T1-weighted structural MRI, 5) registered to the MNI template and resampled into 3×3×3 mm3 voxels, 6) the fMRI data were temporally de-trended by using a third-order polynomial function to

model the slow signal drift, and 7) spatially smoothed by using a 3-D Gaussian filter with 6 mm full width at half maximum (FWHM). Then, fMRI time series were projected from FSL MNI152 space onto the FreeSurfer surface space (1 mm mesh) and downsampled to a 4 mm mesh. Finally, these projected fMRI time series were temporally standardized by subtracting its mean and dividing by its standard deviation.

4.2.3 Removal the effect of spontaneous activity from movie fMRI

FMRI BOLD signals acquired under natural movie condition might contain the neural information not only evoked by the movie itself but also be further contributed by the endogenous activity. Therefore, in order to eliminate the effect of this endogenous signals for subsequent analysis, each subject's fMRI signals were simply averaged across twelve different sessions.

4.2.4 Spatial independent component analysis

A cortical mask defined by [72] was used to select the surface region corresponding to visual cortex (Fig. 4.1 .A). Then, the fMRI signals within those selected regions were temporally concatenated and the following multi-subject dataset was analyzed by using spatial independent component analysis (ICA) code developed in house. Following analysis was applied independently for three different grouped datasets each with 15 subjects' resting-state fMRI and the other dataset with three subjects' averaged fMRI under natural movie condition. After reducing the concatenated dataset into 70 strongest spatial eigenvectors by using principal component analysis (PCA), the resulting spatial eigenvectors were inserted into Infomax ICA to obtain the 70 most representative functional networks. At the end, all of the ICA component maps were spatially normalized by dividing its magnitude.

4.2.5 Spatial similarity of ICA maps

To examine the reproducibility of spatial ICA component maps across three different datasets of resting-state fMRI and between resting-state and natural movie condition, their spatial similarity was first calculated by using simple Pearson correlation of the unthresholded, normalized spatial maps. However, in occasion, a few number of

pairs showed a fair degree of correlation values even though they were only "partially" similar to each other. To avoid this case, the maps matched based on the spatial cross-correlations were further inspected again by visual.

4.2.6 Functional parcellation of visual cortex

After identifying the set of reproducible ICA maps from three different resting-state fMRI datasets, k-means clustering was subsequently applied to the resulting ICA weights to group cortical locations into 15 parcels. Then, the following functional parcellations were visually compared with each other to verify their stability. Moreover, the biological relevance of this functional parcellation was evaluated with respect to the classical visual areas defined by Destrieux (2010) and Van Essen (2012) [73, 74]. In summary, Destrieux (2010) is the anatomical parcellation of cortical sulci and gyri based on the technique that automatically allocates a neuroanatomical label to each location on a cortical surface according to probabilistic information estimated from a manually labeled training set. Van Essen (2012) is a composite cortical parcellation based on architectonic or retinotopic fMRI maps that covers only part of the cortical surface. See [73] and [74] for further information about above two parcellations.

4.3 Results

4.3.1 Reproducible cortical visual networks

Spontaneous BOLD signals within the human visual cortex (Fig. 4.1.A) were decomposed into 70 spatially independent components for each of the three resting-state fMRI datasets. A number of components were reproducible across datasets, showing significant spatial cross-correlations between the matched components extracted from different datasets (Fig. 4.1.B). Fig. 4.1.C displays selective examples of unthresholded maps to demonstrate their reproducibility across the different datasets. Based on the spatial cross-correlations and visual inspection, total of 19 reliable ICA components were identified, as shown in Fig. 4.2 and their spatial information was briefly summarized in Table. 4.1. Sixteen components are focal with well-defined borders and the remaining three (map 5, 6 and 16) are globally distributed over the visual cortical area. Eight components

are bilaterally distributed, whereas eleven are strongly lateralized to either left or right surface.

Among the 19 reliable ICA components, a number of maps with focal spatial distribution agree with existing visual regions or visual evoked activations previously reported in different literatures. To provide more detailed views of these selective maps, they were projected onto the inflated surface and shown in Fig. 4.3. Map 2 shows the fovea representation of the classically retinotopic region, where low visual regions V1, V2 and V3 converge around the occipital pole. Discovery of this map also matches with the result of previous studies, which observed topographic connectivity between V1 and V3 in absence of visual input [64, 65]. Map 3 and 7 showed almost identical spatial patterns, but distributed at left and right hemisphere respectively. These two maps contain V1, V2 and V4v regions at the periphery of fovea region. Moreover, some of maps even showed interesting network patterns at high visual region. Map 9 shows strong coherence with middle temporal (MT) region, which is known as playing a major role in perception and processing of visual motion. Map 13 agrees with V7 region, but the activation region on the right hemisphere is more global than those on the left.

Unlike other component maps, map 5, 6 and 16 show widespread spatial distribution across the visual cortical region (Fig. 4.4). They are well-organized and show bilaterally symmetric distribution. Since they were consistently observed across different datasets, they might contain the signals occurred from neurophysiological origin. However, it is difficult to distinguish whether there could be some non-neural physiological contributions, including cardiac pulsation or breathing effects. These three maps contain different brain regions in parietal, medial and temporal sections of occipital lobe. Map 5 is largely divided into three sections, which correspond to parietal occipital, medial occipital and temporal occipital region respectively. Map 6 shows clear boundary with V1 and fovea region. Map 16 is roughly divided into top and bottom section around the fovea region.
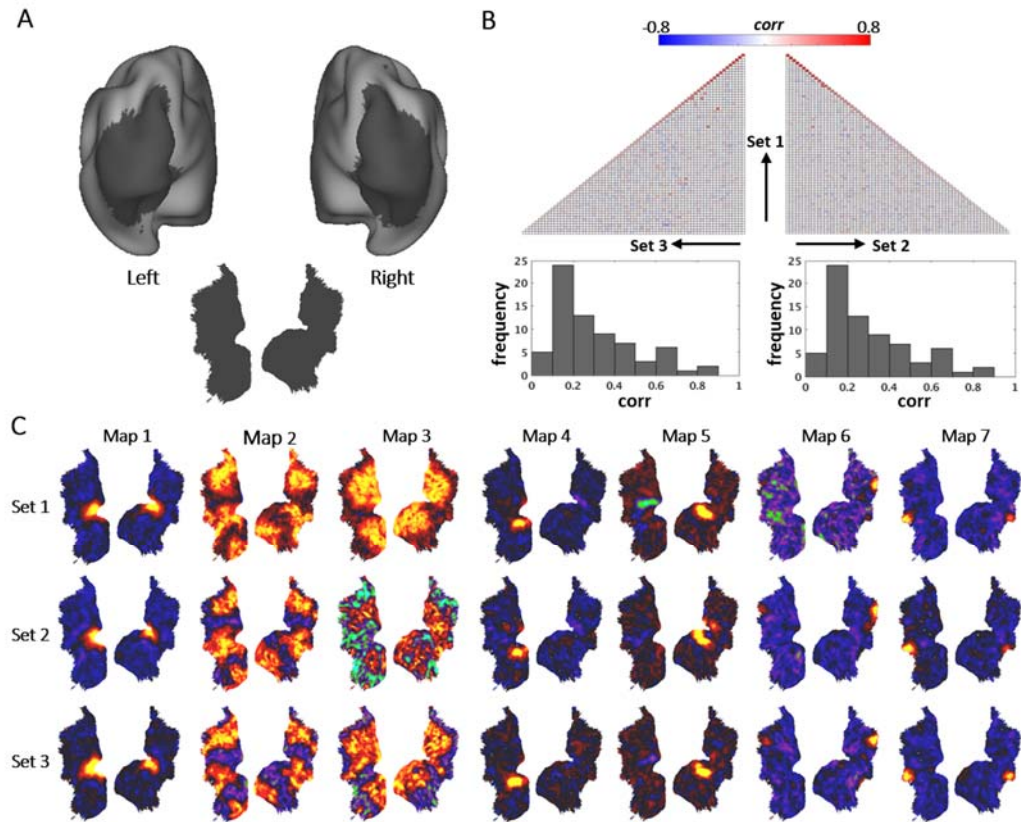
Fig. 4.1 Reproducibility of ICA components within visual cortex. (A) Cortical parcellation of visual cortex. (B) Spatial correlations between 70 components of different datasets. (C) Seven selective examples of reproducible components.
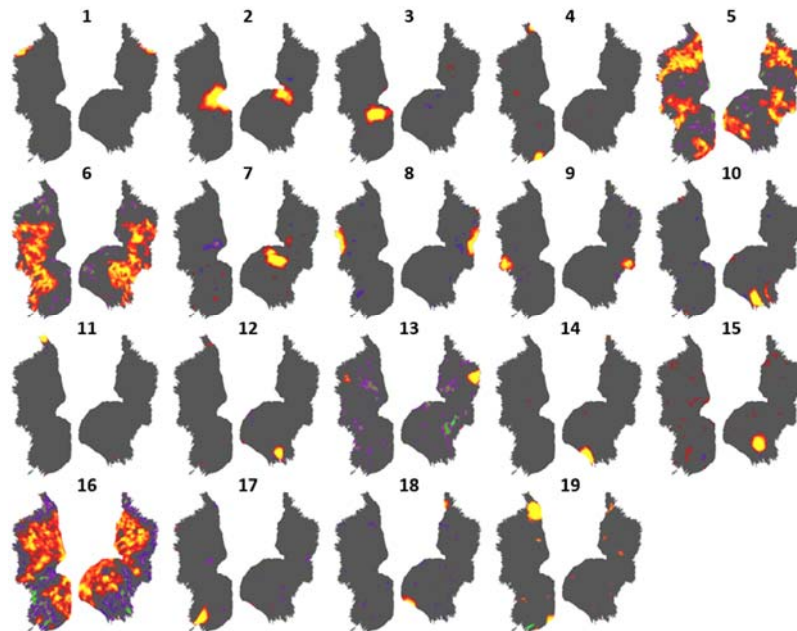
Fig. 4.2 Nineteen reproducible components across three different datasets.
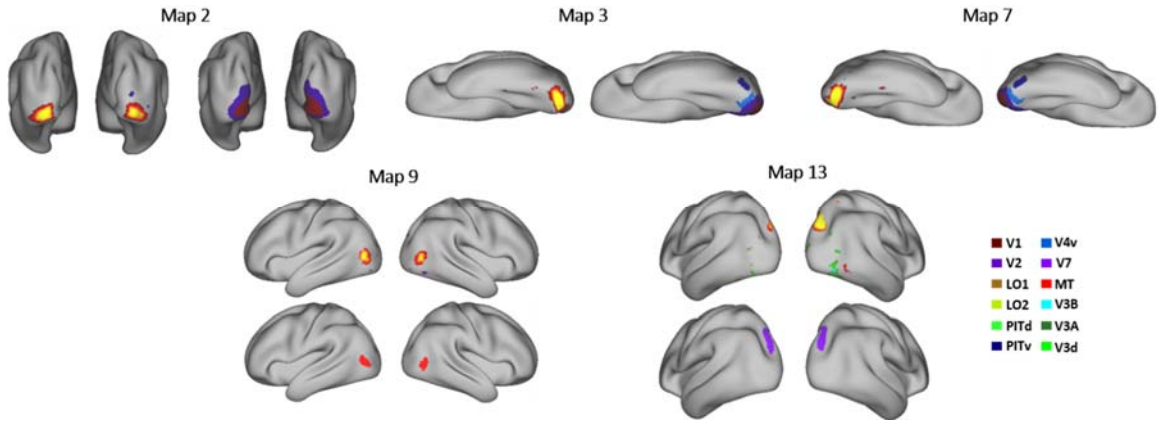
Fig. 4.3 Inflated surface representation of five selective components.

Table 4.1 Spatial information of 19 reproducible components.

Map	MNI coordinates	Symmetry	Distribution
1	(15.5,-71.1,40.9)	bilateral	focal
2	(20.9,-98.3,-9)	bilateral	focal
3	(-18.3,-89.1,-10.3)	left	focal
4	(-16.7,-46.4,-2.7)	left	focal
7	(14.7,-88.5,-8.9)	right	focal
8	(38.7,83.7,21)	bilateral	focal
9	(44.2,73.6,8.1)	bilateral	focal
10	(14.7,-38.9,-6.8)	right	focal
11	(-9.1,-4.3,1.5)	left	focal
12	(34.3,-37.5,-11.2)	right	focal
13	(28.6,-76.3,35.6)	bilateral	focal
14	(22.1,-43.1,-17.8)	right	focal
15	(27.9,-58.6,-15.3)	right	focal
17	(-28.5,-48,-18.9)	left	focal
18	(19.3,-61.4,-0.2)	right	focal
19	(-16.7,-70.3,5.1)	left	focal
5		bilateral	distributed
6		bilateral	distributed
16		Bilateral	distributed
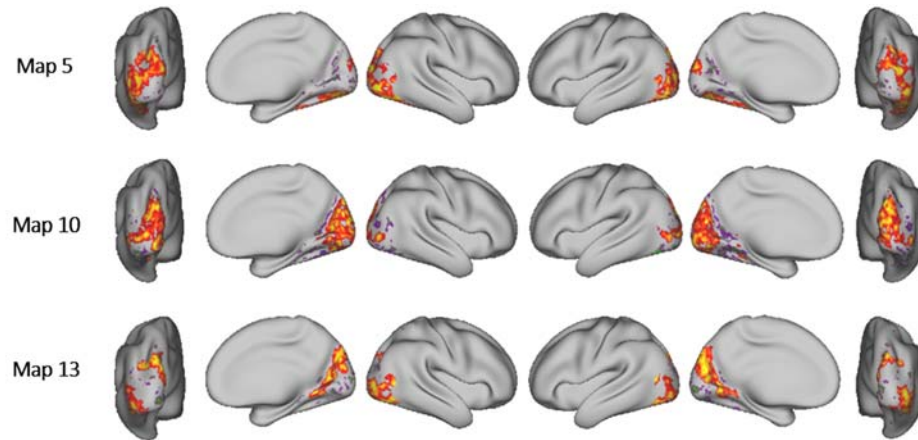
Fig. 4.4 Inflated surface representation of three global components.

4.3.2 Functional parcellation of visual cortex

To generate the functional parcellation of visual cortex, k-means clustering was employed to the weights of 19 reproducible maps and grouped the cortical vertices into 15 different parcels (Fig. 4.5.B). The resulting functional parcellation is highly consistent across three different datasets, indicating its robustness. Fifteen parcels marked with different colors are well-defined with showing clear borders to each other. Among 15 parcels, 8 show bilateral symmetric, the others are strongly left or right lateralized and some even show top-down symmetric. Furthermore, the following parcellation reflects the biological properties of visual system by showing that large parcels exist at low-visual region, but relatively minute parcels emerged as the level becomes higher.

To objectively compare this functional parcellation with Destrieux 2009 (Fig. 4.5.A) and Van Essen 2012 (Fig. 4.5.C), their borders are overlaid in white color and shown in Fig. 4.5.D. With a naked eye, only coarse correspondence was observed to each other. In comparison with Van Essen 2012, parcel 11 shows correspondence with MT region. Parcel 3 roughly aligns with the border between V1 and V2. The combined area of parcel 1, 2 and 3 formed confluence region of V1, V2 and V3. Parcel 9 roughly aligns with V7 region. In comparison with Destrieux 2009, the combined area of parcel 1, 2 and 3 corresponds to occipital pole. Parcel 8 roughly aligns with medial occipito-temporal sulcus and lingual sulcus.
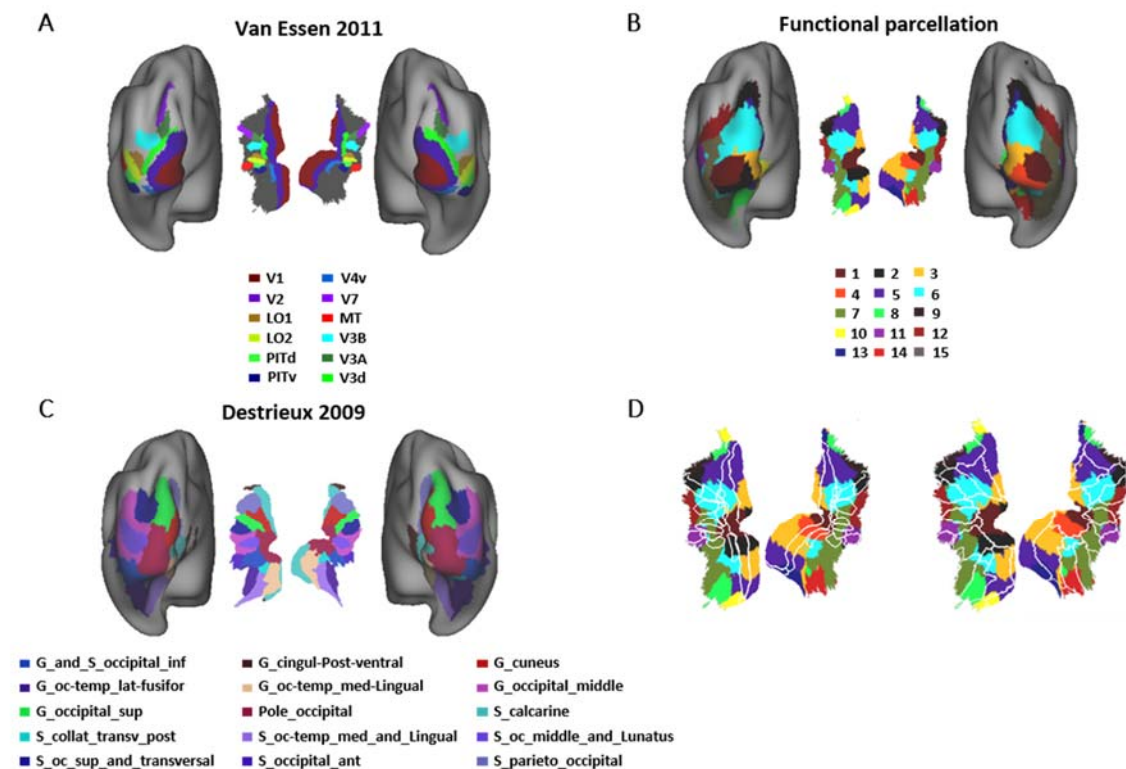
Fig. 4.5 Comparison of functional parcellation with visual regions defined with retinotopic mapping and cytoarchitecture. (A) Parcellation of Van Essen 2011. (B) Functional parcellation. (C) Parcellation of Destrieux 2009. (D) Functional parcellation overlaid with the borders of Van Essen 2011 and Destrieux 2009.

4.3.3 Visual networks under resting-state and natural visual perception

As a comparison with intrinsic functional connectivity patterns, three subjects' fMRI datasets obtained under the repeated natural movie sessions were simply averaged and also decomposed into 70 independent components. Of the 70 components individually generated from rest and movie datasets, 8 maps from each set were unambiguously matched to each other, showing significant spatial cross-correlations (Fig. 4.6.A). In general, the spatial distribution of movie components is generally more confined than those of rest.

Brief descriptions of 8 reproducible maps are provided below. Map 1 includes the medial occipito-temporal gyrus which may be involved in the information processing occurred at dorsal stream. As previously mentioned, Map 2 and 3 display identical network patterns by containing V1, V2 and V4v at the periphery of fovea region. However, the rest component of Map 2 is a little bit further extended to the ventral region. Map 4 represents the foveal representation of retinotopic mapping, but the rest component is more extended to V1, V2 and V3 regions around fovea. Revealing of this pattern during both resting-state and naturalistic visual perception also in-line with the results of previous study [64-66]. Map 5 and Map 6 are bilaterally symmetric and corresponds to medial temporal (MT) and V7 region respectively. Map 7 and Map 8 are strongly left lateralized and distributed at the end of V1 and V2 region.
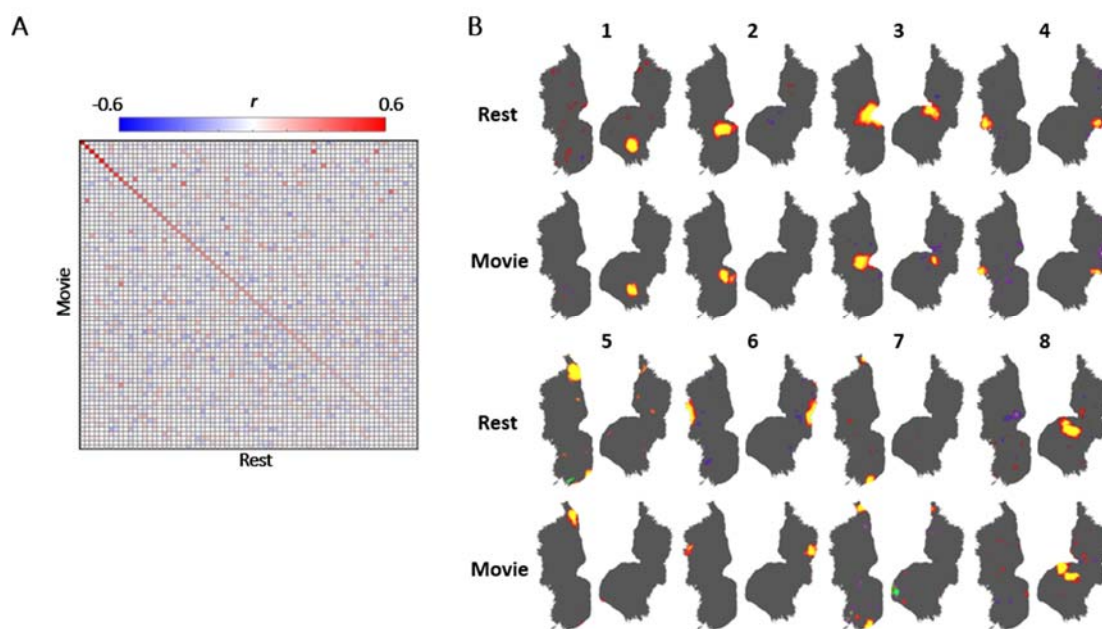

Fig. 4.6 Comparison of fine-scale networks under resting-state and natural movie condition. (A) Spatial similarity between rest and movie component maps. (B) Eight reproducible components between rest and movie dataset.

4.4 Discussion

In the current work, it was demonstrated that the spontaneous brain activity at rest exhibits reliable network patterns not only in the whole-brain scale, but also in much finer scales. Such intrinsic networks in the visual system were reproducible across different subjects and showed correspondence with formerly defined visual areas as well as visual-evoked activations previously reported in different literatures. In addition, these networks provide an information basis to divide the visual cortex into small parcels, which showed distinct differences compared to the classical visual regions defined with retinotopic mapping or cytoarchitecture. Moreover, a number of intrinsic networks at low and high visual regions are preserved even during naturalistic visual stimulation, suggesting their roles in supporting brain function for vision.

A number of results obtained from this study were consistent with previous studies, which demonstrated that the spontaneous brain activity showed the eccentricity organization near the fovea region [64-66, 72]. In addition, this study further extended the previous results by demonstrating that the fine-scale connectivity patterns in higher visual regions also exhibited exquisite patterns and several of them were even preserved during naturalistic visual perception. However, what kind of underlying neural processes trigger these specific organizing principles are still unclear. One possibility is that "rest" participants were engaged in some kind of visual imagery, which may activate visual regions in a retinotopic fashion [75, 76]. Another explanation maybe that retinal ganglion cells exhibit spontaneous bursts of action potentials that were spread to the visual system [77]. The other is that the connectivity patterns that appear during rest may actually reflect the record of habitual past cortical coactivation [66, 78].

Until now, many of previous studies have spent lots of efforts to establish different visual regions of human brain. In specific, those regions were predominantly revealed by using functional MRI recorded under specific task condition called phase-encoding [62, 74]. In comparison, the current study identified visual regions MT and V7 by simply applying spatial ICA to resting-state fMRI signals within the visual cortex. In this sense, the different ICA component maps (Fig. 4.2) and the functional parcellation (Fig. 4.3.B)

shown in this study may provide a future framework to discover unique visual regions which have not been explored before.

Most previous fine-scale connectivity analyses have primarily used seed-based correlation analysis to measure the degree of functional interactions between different brain regions. Although this bivariate statistic is a reliable method to measure coactivation level of two regions, this provides limited information about high-level correlation, such as coactivation of more than two regions. To bypass this problem, in this study, spatial independent component analysis (ICA) was employed and the applicability of this multivariate statistic to fine-scale connectivity analysis was confirmed. In fact, a number of results obtained in this study converged into similar findings as those obtained by using temporal correlation [63-66]. Applying various multivariate statistics algorithms to different subsystems, including auditory and somatosensory systems may potentially reveal richer fine-scale connectivity patterns.

When examining the reproducibility of component maps with resting-state fMRI datasets, the proportion of reproducible maps to the total number of maps was not significant. This phenomenon may be occurred, since only three datasets were used to test their reproducibility. In fact, several maps were commonly included only at two datasets, but these maps were simply excluded from reproducible candidates. Increasing the statistical sensitivity by using bootstrap method may identify more number of reliable and meaningful components. In biological aspect, this observation may imply that the spontaneous activity occurred in visual cortex may utilize the set of reproducible fine-scale networks as well as those that were uniquely exist at different subjects.

Similar as above, the number of pairs achieved between rest and movie component maps was also not high. As explained previously, the movie fMRI signals were further processed by averaging across different sessions. This step might cause unpredictable results to subsequent ICA analysis and those resulting components. Moreover, the current work only used three subjects' dataset. Performing the same analysis with more number of subjects may increase the reliability of movie component maps. In biological aspect, it could be inferred that the functional dynamics occurred in visual cortex during natural perception may employ a set of networks exhibited during resting-state as well as those

that are uniquely shown in task condition. Further investigating the unmatched components across different datasets of resting-state fMRI and between resting-state and natural visual perception may elucidate the underlying neuronal bases that trigger the difference during behavioral and cognitive states.

One potential drawback of this work is the visual cortex region (Fig. 4.1.A) used in this analysis does not cover enough regions of the temporal and parietal lobes, which are also known to be deeply involved in the high-level visual information processing [43]. It was inferred that this could be the reason why some component maps such as map 1 and 18 shown in Fig. 4.2 were somewhat spatially "cropped" at the boundary of analyzed region. Performing the similar analysis with further extended visual cortex parcellation may potentially provide more valuable fine-scale information.

Lastly, although spatial ICA is a powerful data-driven method of extracting functional networks, this algorithm cannot capture the dynamic changes of resting-state fMRI due to the assumption of temporal stationarity. In order to capture these connectivity dynamics, different methodologies have been suggested and applied to the resting-state fMRI. First, by using a shorter-time sliding window, dynamic functional connectivity can be obtained by calculating temporal correlation of the data points within that specific window [48-50]. Second, by applying k-means clustering technique to instantaneous fMRI volumes, spatially overlapping co-activation patterns (CAPs) was identified [20]. Third, by combining the signal processing technique called TA (Total Activation) regularization and clustering technique, spatially and temporally overlapping iCAPs (innovation-driven co-activation patterns) were extracted [21]. Investigating the fine-scale connectivity pattern in dynamic aspects with above-mentioned methodologies may reveal more rich spatiotemporal information of intrinsic activity.

5. CONCLUSION

By this time, the analysis of brain network in both structural and functional perspectives has provided deep insights into the organization of brain, as well as how the brain's rich functionality emerge from this complex architecture. These findings are feasible by the technological developments in various non-invasive neuroimaging techniques as well as new powerful tools from graph theory and dynamical systems. Based upon previous studies, this thesis tried to answer three critical questions in the field by analyzing different structural and functional datasets with distinct mathematical and statistical methodologies. The brief summary of results and conclusion of each question is explained below.

In chapter 2, a recent "link community" algorithm was applied to different structural and functional network datasets to disentangle complex brain architecture. As a result, it was demonstrated that the brain network contains rich and diverse link communities, which were spatially overlapping, but biologically significant. However, the full interpretation of all link communities is not immediately clear and open to scientific debate. This is because the complete understanding of the functional roles of individual brain regions as well as their interconnections is far from being clear. Further studies must be needed to elucidate the potential biological relevance of such an arrangement of structural or functional connections and biological significance of those link communities.

In chapter 3, two representative multivariate analysis algorithms of resting-state fMRI: spatial independent component analysis (sICA) and k-means clustering were investigated to elucidate their relationship in terms of functional relevance and theoretical basis. This study was originated from the idea of machine learning field that ICA and k-means clustering are conceptually and theoretically closely related methodologies. Initially,

the resting-state network patterns extracted with those two methods showed significant differences. However, after removing the signals occurred during intermediate phases by constraining the temporal sparsity of resting-state fMRI, k-means clustering yielded similar network patterns with rest ICs and task CAPs. Based on this observation, it was inferred that these two methods might converge into the similar findings of resting-state networks.

In chapter 4, the complex spatiotemporal patterns of spontaneous activity within the visual cortex were explored by evaluating their spatial organization, inter-subject reproducibility and functional relevance to naturalistic visual perception. As a result, it was revealed that fine-scale intrinsic connectivity patterns in visual cortex exhibits robust and reliable network patterns. In addition, the parcels obtained based on these network patterns showed novel configurations compared to those defined with retinotopic mapping and cytoarchitecture. Moreover, a number of these patterns were also even preserved during naturalistic visual perception, suggesting their functions in supporting visual task.

Perhaps one of the greatest scientific challenges is to understand the human brain. One key aspect to tackle this challenge is to explore the structural and functional networks that maintain numerous brain's functions and behaviors. By complementing the previous network studies, the distinguishing results obtained in this thesis may provide new insights regarding the brain's organization, as well as a better understanding of mathematical and statistical tools for functional and structural network analysis.

LIST OF REFERENCES

LIST OF REFERENCES

[1] Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex*, *1*(1), 1-47.

[2] Young, M. P. (1993). The organization of neural systems in the primate cerebral cortex. *Proceedings of the Royal Society of London B: Biological Sciences*, *252*(1333), 13-18.

[3] Oh, S. W., Harris, J. A., Ng, L., Winslow, B., Cain, N., Mihalas, S., ... & Mortrud, M. T. (2014). A mesoscale connectome of the mouse brain. Nature, 508(7495), 207-214.

[4] Mori, S., & Zhang, J. (2006). Principles of diffusion tensor imaging and its applications to basic neuroscience research. *Neuron*, *51*(5), 527-539.

[5] Hagmann, P., Thiran, J. P., Jonasson, L., Vandergheynst, P., Clarke, S., Maeder, P., & Meuli, R. (2003). DTI mapping of human brain connectivity: statistical fibre tracking and virtual dissection. Neuroimage, 19(3), 545-554.

[6] Sporns, O., Chialvo, D. R., Kaiser, M., & Hilgetag, C. C. (2004). Organization, development and function of complex brain networks. Trends in cognitive sciences, 8(9), 418-425.

[7] Iturria-Medina, Y., Sotero, R. C., Canales-Rodríguez, E. J., Alemán-Gómez, Y., & Melie-García, L. (2008). Studying the human brain anatomical network via diffusion-weighted MRI and Graph Theory. *Neuroimage*, *40*(3), 1064-1076.

[8] Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biol*, *6*(7), e159

[9] Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*(3), 186-198.

[10] Bassett, D. S., & Bullmore, E. D. (2006). Small-world brain networks. The neuroscientist, 12(6), 512-523.

[11] Chen, Z. J., He, Y., Rosa-Neto, P., Germann, J., & Evans, A. C. (2008). Revealing modular architecture of human brain structural networks by using cortical thickness from MRI. Cerebral cortex, 18(10), 2374-2381.

[12] Meunier, D., Lambiotte, R., Fornito, A., Ersche, K. D., & Bullmore, E. T. (2010). Hierarchical modularity in human brain functional networks. Hierarchy and dynamics in neural networks, 1, 2.

[13] Van den Heuvel, M. P., & Sporns, O. (2011). Rich-club organization of the human connectome. The Journal of neuroscience, 31(44), 15775-15786.

[14] Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. Nature Reviews Neuroscience, 13(5), 336-349.

[15] Biswal, B., Zerrin Yetkin, F., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic resonance in medicine*, *34*(4), 537-541.

[16] Cordes, D., Haughton, V. M., Arfanakis, K., Wendt, G. J., Turski, P. A., Moritz, C. H., ... & Meyerand, M. E. (2000). Mapping functionally related regions of brain with functional connectivity MR imaging. American Journal of Neuroradiology, 21(9), 1636-1644.

[17] Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, *100*(1), 253-258.

[18] Beckmann, C. F., DeLuca, M., Devlin, J. T., & Smith, S. M. (2005). Investigations into resting-state connectivity using independent component analysis. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 360(1457), 1001-1013.

[19] Smith, S. M., Miller, K. L., Moeller, S., Xu, J., Auerbach, E. J., Woolrich, M. W., ... & Van Essen, D. C. (2012). Temporally-independent functional modes of spontaneous brain activity. Proceedings of the National Academy of Sciences, 109(8), 3131-3136.

[20] Liu, X., Chang, C., & Duyn, J. H. (2013). Decomposition of spontaneous brain activity into distinct fMRI co-activation patterns. Frontiers in systems neuroscience, 7.

[21] Karahanoğlu, F. I., & Van De Ville, D. (2015). Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks. *Nature communications*, *6*.

[22] Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., ... & Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. Proceedings of the National Academy of Sciences, 106(31), 13040-13045.

[23] McKeown, M. J., Makeig, S., Brown, G. G., Jung, T. P., Kindermann, S. S., Bell, A. J., & Sejnowski, T. J. (1997). *Analysis of fMRI data by blind separation into independent spatial components* (No. NHRC-REPT-97-42). NAVAL HEALTH RESEARCH CENTER SAN DIEGO CA.

[24] Brookes, M. J., Woolrich, M., Luckhoo, H., Price, D., Hale, J. R., Stephenson, M. C., ... & Morris, P. G. (2011). Investigating the electrophysiological basis of resting state networks using magnetoencephalography. Proceedings of the National Academy of Sciences, 108(40), 16783-16788.

[25] Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., & Beckmann, C. F. (2006). Consistent resting-state networks across healthy subjects. *Proceedings of the national academy of sciences*, *103*(37), 13848-13853.

[26] Zhang, N., Rane, P., Huang, W., Liang, Z., Kennedy, D., Frazier, J. A., & King, J. (2010). Mapping resting-state brain networks in conscious animals. *Journal of neuroscience methods*, *189*(2), 186-196.

[27] Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., & Petersen, S. E. (2014). Intrinsic and task-evoked network architectures of the human brain. *Neuron*, *83*(1), 238-251.

[28] Stam, C. J., Jones, B. F., Nolte, G., Breakspear, M., & Scheltens, P. (2007). Small-world networks and functional connectivity in Alzheimer's disease. *Cerebral cortex*, *17*(1), 92-99.

[29] Braun, U., Muldoon, S. F., & Bassett, D. S. (2009). On human brain networks in health and disease. *eLS*.

[30] Crossley, N. A., Mechelli, A., Vértes, P. E., Winton-Brown, T. T., Patel, A. X., Ginestet, C. E., ... & Bullmore, E. T. (2013). Cognitive relevance of the community structure of the human brain functional coactivation network. *Proceedings of the National Academy of Sciences*, *110*(28), 11583-11588.

[31] Ahn, Y. Y., Bagrow, J. P., & Lehmann, S. (2010). Link communities reveal multiscale complexity in networks. *Nature*, *466*(7307), 761-764.

[32] Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A., & Arieli, A. (2003). Spontaneously emerging cortical representations of visual attributes. *Nature*, *425*(6961), 954-956.

[33] Ohiorhenuan, I. E., Mechler, F., Purpura, K. P., Schmid, A. M., Hu, Q., & Victor, J. D. (2010). Sparse coding and high-order correlations in fine-scale cortical networks. Nature, 466(7306), 617-621.

[34] Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS Comput Biol*, *1*(4), e42.

[35] Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, *52*(3), 1059-1069.

[36] Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., ... & Petersen, S. E. (2011). Functional network organization of the human brain. Neuron, 72(4), 665-678.

[37] Jaccard, P. (1901). Etude comparative de la distribution florale dans une portion des Alpes et du Jura. Impr. Corbaz.

[38] Tanimoto, T. T. (1958). elementary mathematical theory of classification and prediction.

[39] 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.

[40] Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., ... & Van Essen, D. C. (2013). The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage, 80, 105-124.

[41] Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., & Smith, S. M. (2014). Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. Neuroimage, 90, 449-468.

[42] Smith, S. M., Hyvärinen, A., Varoquaux, G., Miller, K. L., & Beckmann, C. F. (2014). Group-PCA for very large fMRI datasets. NeuroImage, 101, 738-749.

[43] Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, *15*(1), 20-25.

[44] Hwang, K., Hallquist, M. N., & Luna, B. (2013). The development of hub architecture in the human functional brain network. *Cerebral Cortex*, *23*(10), 2380-2393.

[45] Arieli, A., Sterkin, A., Grinvald, A., & Aertsen, A. D. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science*, *273*(5283), 1868-1871.

[46] Cordes, D., Haughton, V. M., Arfanakis, K., Wendt, G. J., Turski, P. A., Moritz, C. H., ... & Meyerand, M. E. (2000). Mapping functionally related regions of brain with functional connectivity MR imaging. *American Journal of Neuroradiology*, *21*(9), 1636-1644.

[47] Fox, M. D., & Greicius, M. (2010). Clinical applications of resting state functional connectivity. *Frontiers in systems neuroscience*, *4*, 19.

[48] Chang, C., & Glover, G. H. (2010). Time–frequency dynamics of resting-state brain connectivity measured with fMRI. *Neuroimage*, *50*(1), 81-98.

[49] Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., ... & Handwerker, D. A. (2013). Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage*, *80*, 360-378.

[50] Hutchison, R. M., Womelsdorf, T., Gati, J. S., Everling, S., & Menon, R. S. (2013). Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Human brain mapping*, *34*(9), 2154-2177.

[51] Biswal, B. B., Mennes, M., Zuo, X. N., Gohel, S., Kelly, C., Smith, S. M., ... & Dogonowski, A. M. (2010). Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences*, *107*(10), 4734-4739.

[52] Coates, A., & Ng, A. Y. (2012). Learning feature representations with k-means. In *Neural Networks: Tricks of the Trade* (pp. 561-580). Springer Berlin Heidelberg.

[53] Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., & Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: are anti-correlated networks introduced?. *Neuroimage*, *44*(3), 893-905.

[54] Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage*, *16*(3), 765-780.

[55] Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural networks*, *13*(4), 411-430.

[56] Laird, A. R., Fox, P. M., Eickhoff, S. B., Turner, J. A., Ray, K. L., McKay, D. R., ... & Fox, P. T. (2011). Behavioral interpretations of intrinsic connectivity networks. Journal of cognitive neuroscience, 23(12), 4022-4037.

[57] Barbieri, R., & Shimono, M. (2012). Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis. *Networking of Psychophysics, Psychology and Neurophysiology*, 61.

[58] Fukushima, M., Saunders, R. C., Leopold, D. A., Mishkin, M., & Averbeck, B. B. (2012). Spontaneous high-gamma band activity reflects functional organization of auditory cortex in the awake macaque. *Neuron*, *74*(5), 899-910.

[59] Chen, L., Mishra, A., Newton, A. T., Morgan, V. L., Stringer, E. A., Rogers, B. P., & Gore, J. C. (2011). Fine-scale functional connectivity in somatosensory cortex revealed by high-resolution fMRI. *Magnetic resonance imaging*, *29*(10), 1330-1337.

[60] Crowe, D. A., Averbeck, B. B., & Chafee, M. V. (2010). Rapid sequences of population activity patterns dynamically encode task-critical spatial information in parietal cortex. *The Journal of Neuroscience*, *30*(35), 11640-11653.

[61] Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral cortex*, *7*(2), 181-192.

[62] Warnking, J., Dojat, M., Guérin-Dugué, A., Delon-Martin, C., Olympieff, S., Richard, N., ... & Segebarth, C. (2002). fMRI retinotopic mapping—step by step. *NeuroImage*, *17*(4), 1665-1683.

[63] Nir, Y., Hasson, U., Levy, I., Yeshurun, Y., & Malach, R. (2006). Widespread functional connectivity and fMRI fluctuations in human visual cortex in the absence of visual stimulation. *Neuroimage*, *30*(4), 1313-1324.

[64] Heinzle, J., Kahnt, T., & Haynes, J. D. (2011). Topographically specific functional connectivity between visual field maps in the human brain. *Neuroimage*, *56*(3), 1426-1436.

[65] Raemaekers, M., Schellekens, W., van Wezel, R. J., Petridou, N., Kristo, G., & Ramsey, N. F. (2014). Patterns of resting state connectivity in human primary visual cortical areas: a 7T fMRI study. *Neuroimage*, *84*, 911-921.

[66] Wilf, M., Strappini, F., Golan, T., Hahamy, A., Harel, M., & Malach, R. (2015). Spontaneously Emerging Patterns in Human Visual Cortex Reflect Responses to Naturalistic Sensory Stimuli. *Cerebral Cortex*, bhv275.

[67] Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., & WU-Minn HCP Consortium. (2013). The WU-Minn human connectome project: an overview. *Neuroimage*, *80*, 62-79.

[68] Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, *10*, 433-436.

[69] Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*(4), 437-442.

[70] Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., ... & Niazy, R. K. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, *23*, S208-S219.

[71] Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research*, *29*(3), 162-173.

[72] Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., ... & Fischl, B. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*, *106*(3), 1125-1165.

[73] Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage*, *53*(1), 1-15.

[74] Van Essen, D. C., Glasser, M. F., Dierker, D. L., Harwell, J., & Coalson, T. (2012). Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. *Cerebral Cortex*, *22*(10), 2241-2262.

[75] Ishai, A., & Sagi, D. (1995). Common mechanisms of visual imagery and perception. *Science*, *268*(5218), 1772-1774.

[76] Slotnick, S. D., Thompson, W. L., & Kosslyn, S. M. (2005). Visual mental imagery induces retinotopically organized activation of early visual areas. *Cerebral cortex*, *15*(10), 1570-1583.

[77] Blankenship, A. G., & Feller, M. B. (2010). Mechanisms underlying spontaneous patterned activity in developing neural circuits. *Nature Reviews Neuroscience*, *11*(1), 18-29.

[78] Fiser, J., Berkes, P., Orbán, G., & Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. *Trends in cognitive sciences*, *14*(3), 119-130.

APPENDIX

APPENDIX

```
%%% link_community
% link_community() - apply link community algorithm to network connectivity
%               matrix
%
% Usage
%   [edge_cid,linkage,cutoff,max_D,Thr_D] =...
%                       link_community(net_mat,'keyword','value')
%
% Inputs
%   net_mat: connection matrix (undirected binary or undirected weighted)
%
% Outputs
%   edge_cid: edge with its' assigned link community ID
%    linkage: record hierarchical structure of link clustering.
%     cutoff: the level of cutting dendrogram
%      max_D: value of a maximum partition density
%      Thr_D: cutoff level with respect to partition density
%
% Keywords
%      thr: in case of the weighted network, remove out
%          edges below a threshold {default: 0}
%
% Version:
%   1.02
%
% Reference:
%   Ahn, Yong-Yeol, James P. Bagrow, and Sune Lehmann.
%   "Link communities reveal multiscale complexity in networks"
%   Nature 466.7307 (2010)


% CREATED:
%    2/02/2014
%    Author: Jun Young Jeong, Haiguang Wen and Zhongming Liu

%%% History
% 1.00 - 2/02/2014: create the file
% 1.01 - 2/27/2014: add Keywords 'cutoff' and 'threshold'
% 1.02 - 3/05/2014: change Tanimoto coefficient formula




function [edge_cid,linkage,cutoff,max_D,Thr_D] = link_community(W,varargin)
% if the number of input arguments is less than 1,
if nargin<1
    eval('help link_community');
    return
end
```

```matlab
%%% check for errors in input
% if input is not a matrix format,
if ~ismatrix(W)
    error('link_community(): must input a matrix');
else
    % if input is a not square matrix format,
    if  ~issquare(W)
        error('link_community(): must input a square matrix');
    end
end

%%% default
flag_threshold = 0; % 0(not given)|1(given)
flag_directed = ~issymmetric(single(W)); % 0(undirected)|1(directed)
flag_weighted = isweight(W); % 0(unweighted)|1(weighted)

%%% keywords
for ik = 1:2:size(varargin,2)
    Keyword = varargin{ik};
    Value   = varargin{ik+1};
    if strcmpi(Keyword,'thr')
        flag_threshold = 1;
        thr = Value;
    else
        warning(['link_community: unknown keyword ' Keyword]);
    end
end

%%% reorganize the node-connectivity matrix as list format
% let "link" be the connections between every pair of nodes (i,j)
% link(:,1): ni (starting node)
% link(:,2): nj (ending node)
% link(:,3): connection strength between node ni and nj

% if the input matrix is undirected,
if flag_directed == 0;
    % in case of the weighted network, filter edges whose strength is below
    % a threshold
    if flag_weighted && flag_threshold
        W(abs(W)<thr)=0;
    end
    % starting and ending node
    [ni,nj] = find(triu(W)~=0);
    % total number of links
    nlink = size([ni,nj],1);
    % [ni,nj,connection strength]
    link = zeros(nlink,3);
    link(:,1) = ni; link(:,2) = nj;
    link(:,3) = W(triu(W)~=0);

% if the input matrix is directed,
elseif flag_directed == 1;
    % to be continue
```

```matlab
end

%%% calculate the similarity of links sharing a common (keystone) node
disp('computing link similarities...');
% extract nodes
N = unique(link(:,1:2));
% [link similarity,link1(node1,node2),link2(node3,node4)]
S_link = zeros(length(N)*(length(N)-1)/2,5);

cnt = 1; % counter
% if the network is undirected,
if flag_directed == 0;
    % knode is a "keystone" node
    for iN = 1:length(N);
        knode = N(iN);
        % inclusive neighbors of a knode
        inc_ngh_k = unique(link((link(:,1)==knode)|(link(:,2)==knode),1:2));
        % non-inclusive neightbors of a knode
        ngh_k = inc_ngh_k(inc_ngh_k~=knode);
        if length(ngh_k) > 1;
            % pair node i and node j using combnk function
            pair_ij = combnk(ngh_k,2);
            for jpair = 1:size(pair_ij,1)
                inode = pair_ij(jpair,1); jnode = pair_ij(jpair,2);
                % inclusive neighbors of inode
                inc_ngh_i = unique(link((link(:,1)==inode)|(link(:,2)==inode),1:2));
                % inclusive neighbors of jnode
                inc_ngh_j = unique(link((link(:,1)==jnode)|(link(:,2)==jnode),1:2));

                % if the network is binary,
                if flag_weighted == 0;
                    % calculate the link similarity using Jaccard index
                    S = length(intersect(inc_ngh_i,inc_ngh_j))/...
                                length(union(inc_ngh_i,inc_ngh_j));


                % if the network is weighted,
                elseif flag_weighted == 1;
                    % calculate the link similarity using Tanimoto coefficient
                    ai = W(inode,:);
                    ai(inode) = sum(W(inode,(inc_ngh_i(inc_ngh_i~=inode))))/...
                                sum(inc_ngh_i~=inode);
                    aj = W(jnode,:);
                    aj(jnode) = sum(W(jnode,(inc_ngh_j(inc_ngh_j~=jnode))))/...
                                sum(inc_ngh_j~=jnode);
                    % Tanimoto coefficient formula
                    S = (ai*aj')/(norm(ai)^2+norm(aj)^2-ai*aj');
                end

                S_link(cnt,1:5) = [S,min(knode,inode),max(knode,inode),...
                                min(knode,jnode),max(knode,jnode)];
                % update cnt
                cnt=cnt+1;
            end
```

```
        end
    end

% if network is directed,
elseif flag_directed == 1;
    % continue
end
S_link = S_link(any(S_link,2),:); % remove unfilled rows
% rearrange 'S_link' according to the descend order of link similarity
[~,I] = sort(S_link(:,1),'descend');
S_link = S_link(I,:);

%% sequentially merge links in the order of similarity & calculate partition density
disp('hierarchical clustering...');
% initially, every link is assigned to its' own community
E = [link(:,1:2),(1:length(link(:,1)))'];
% record the dendrogram structure
linkage = zeros(nlink-1,3);
% at each merging step, record the value of partition density
D = zeros(nlink-1,1);
% at each merging step, record the partition density and community id
D_cid = zeros(nlink-1,nlink+1);

% updated partition density
tmp_D = 0;
% updated community id
tmp_cid = nlink;
% counter
cnt = 1;

for kpair = 1:length(S_link(:,1));
    link1 = S_link(kpair,2:3); link2 = S_link(kpair,4:5);
    cid1 = E((E(:,1)==link1(1))&(E(:,2)==link1(2)),3); % cid of link1
    cid2 = E((E(:,1)==link2(1))&(E(:,2)==link2(2)),3); % cid of link2
    % cid1 and cid2 are same, they are already merged
    if cid1 == cid2
        continue;
    end

    % partition density of cid1
    idx1 = (E(:,3)==cid1);
    m1 = sum(idx1); % number of links at cid1
    n1 = length(unique(E(idx1,1:2))); % number of nodes at cid1
    D1 = partdens(m1,n1);
    % partition density of cid2
    idx2 = (E(:,3)==cid2);
    m2 = sum(idx2);
    n2 = length(unique(E(idx2,1:2)));
    D2 = partdens(m2,n2);

    % record the merging point
    linkage(cnt,1:3) = [cid1,cid2,S_link(kpair)];
    % update cid
    tmp_cid=tmp_cid + 1; new_cid = tmp_cid;
```

```matlab
        % assign a new cid to merged community
        E((E(:,3)==cid1),3) = new_cid; E((E(:,3)==cid2),3) = new_cid;

        % partition density of merged community
        idx3 = (E(:,3)==new_cid);
        m3 = sum(idx3);
        n3 = length(unique(E(idx3,1:2)));
        D3 = partdens(m3,n3);

        % update partition density
        tmp_D = tmp_D + (D3 - D1 - D2)*(2/nlink);
        D(cnt) = tmp_D;

        % reassign community id
        D_cid(cnt,1) = tmp_D;
        tmp_comid = unique(E(:,3));
        for icom = 1:length(tmp_comid)
            D_cid(cnt,find(E(:,3)==tmp_comid(icom))+1) = icom;
        end
        % update
        cnt=cnt+1;
end



%%% list threshold and corresponding partition density
S_prev = -1; % predefine S
S_all = linkage(:,3); % define threshold at each step

Thr_D = zeros(unique(length(S_all)),2); % cutoff and partition density
for irow = 1:length(S_all);
    % if next similarity is not identical previous similarity,
    if S_all(irow) ~= S_prev
        best_S = S_all(irow);
        % at first merging point, D is 0
        if irow == 1
            best_D = 0;
        else
            best_D = D(irow-1);
        end
        Thr_D(irow,1:2) = [best_S,best_D];
        S_prev = S_all(irow);
    end
end

% remove unfilled rows
Thr_D((Thr_D(:,1)==0),:) = [];
% when the cutoff value is 0
Thr_D = [Thr_D;[0,D(end)]];

%%% bring the community id at maximum partition density
max_D =  max(Thr_D(:,2)); % maximum D
cutoff = Thr_D((Thr_D(:,2)==max_D),1); % cutoff corresponding to max D
cutoff = cutoff(end);
```

```matlab
% rearrange according to assigned community ID
cid = D_cid(D_cid(:,1)==max_D,2:end);
edge_cid = [link(:,1:2),cid(1,:)'];
[~,I] = sort(edge_cid(:,3));
edge_cid = edge_cid(I,:);

%%% draw a dendrogram & plot partition density plot
disp('display a dendrogram and partition density plot...');
% set figure size as screen size
h = figure(1);
ss = get(0,'ScreenSize'); % screen size
set(h,'Position',[0,0,ss(3),ss(4)]);

% dendrogram
subplot(1,2,1);
dendrogram([linkage(:,1:2),1-linkage(:,3)],0,'colorthreshold',1-cutoff);
hold on;
axis([0,size(link,1),0,1]);
% overlay cutoff level on the dendrogram
plot([0,size(link,1)],[1-cutoff,1-cutoff],'k','Linewidth',2,'Linestyle','--');
axis off;

% plot partition density vs cutoff
subplot(1,2,2); hold on;
plot(Thr_D(:,2),Thr_D(:,1));
xlabel('Partition Density'); ylabel('Threshold');
title('Partition Density vs Threshold');
% axis setting
axis([0,1,0,1])
set(gca,'YDir','reverse')
plot([0,1],[cutoff,cutoff],'k','Linewidth',2,'Linestyle','--');
end

%%% partition density of link community
% m: number of links
% n: number of nodes
function D = partdens(m,n)
    if n ~= 2
        D = (m*(m-n+1))/((n-2)*(n-1));
    elseif n == 2
        D = 0;
    end
end

%%% check if the matrix A is weighted
function b = isweight(A)
    weight = sum(unique(A));
    if weight ~= 1
        b = 1;
    else
        b = 0;
    end
end
```

```matlab
%% check if the matrix A is square
function b = issquare(A)
    if size(A,1) ~= size(A,2)
        b = 0;
    else
        b = 1;
    end
end
```