COMMON BRAIN LANDMARKS REFLECTING CONSISTENT STRUCTURAL AND FUNCTIONAL ARCHEITECTURE

by

DAJIANG ZHU

(Under the Direction of Professor Tianming Liu)

ABSTRACT

An unrelenting human quest regarding the brain science is: what is the intrinsic relationship between the brain’s structural and functional architectures, which partly defines what we are and who we are. Recent studies suggest that each brain’s cytoarchitectonic region has a unique set of extrinsic inputs and outputs, named as “connectional fingerprint”, which largely determines the functions that each brain area performs. However, their explicit connections are largely unknown. For example, in what extent they are inclined to be coherent with each other and otherwise they will intend to show more heterogeneity? In this dissertation, based on my previously proposed brain structural atlas which represents the most consistent structural connectome across different populations, I developed a novel group-wise optimization framework to computationally model the functional homogeneity behind them. The optimization procedure is conducted under the joint structural and functional regulations and therefore the achieved common brain landmarks reflect the consistency of brain structure and function simultaneously.

INDEX WORDS: optimization, sparse coding, fMRI, brain structure and function
COMMON BRAIN LANDMARKS REFLECTING CONSISTENT STRUCTURAL
AND FUNCTIONAL ARCHETECTURE

by

DAJIANG ZHU
B.A., Shanghai Jiaotong University, China, 2001

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA
2014
DEDICATION

This work is dedicated to all the researchers in the fields of brain imaging, computational neuroscience, psychology, psychiatry and brain mapping who need to systematically study the brain connectivity.
ACKNOWLEDGEMENTS

First, I would like to express my deepest appreciation to Dr. Tianming Liu, who is my major advisor and good friend. I will never forget how much Dr. Liu has devoted during these six years of mentoring. He has changed me to a mature and independent investigator in the brain mapping field from a fresh Ph.D. student without any sense of research. The most important thing I learned from Dr. Liu is not the knowledge or techniques, but the attitude in the research and life: be serious, be patient and never give up. Also, my super smart colleagues help me a lot and I cannot have such fruitful products without them. They are Dr. Kaiming Li, Mr. Fan Deng, Mr. Tuo Zhang, Mr. Degang Zhang, Mr. Xi Jiang, Mr. Hanbo Chen, Mr. Jinglei Lv and many others.

Furthermore, I would like to thank Dr. L. Stephen Miller, Dr. Suchendra M. Bhandarkar, Dr. Kang Li, and Dr. Qun Zhao for being the member of my committee. They gave me so many insightful suggestions and advices. One thing I want to mention is that, because of my incaution, the schedule of the final stages including comprehensive exam, prospectus and defense is very tight. I really appreciate their kindly supports which make my whole graduation procedure smooth and successful.

Finally, I want to say thank you to my wife, Jingjing and I will never make this without her full support. And to my parents and sister’s family, they are always on my side no matter what kind of challenges I am facing. Kathy, my lovely daughter, it is you give me endless power to strive and I also make this as one of your two-year birthday presents.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................................................ v
LIST OF TABLES .................................................................................................................................................. viii
LIST OF FIGURES ................................................................................................................................................ ix

CHAPTER

1 INTRODUCTION ............................................................................................................................................... 1
  1.1 Thesis Statements.................................................................................................................................... 1
  1.2 Contributions ....................................................................................................................................... 3
  1.3 Thesis Outline ....................................................................................................................................... 5

2 DICCCOL LANDMARKS ...................................................................................................................................... 7
  2.1 Motivation ............................................................................................................................................. 7
  2.2 Trace-Map Model ................................................................................................................................. 10
  2.3 Whole Brain Optimization .................................................................................................................... 15
  2.4 DICCCOL Prediction ............................................................................................................................ 19
  2.5 The Advantages of DICCCOL .............................................................................................................. 21

3 APPLICATIONS OF DICCCOL .................................................................................................................. 26
  3.1 Functional Labelling of DICCCOL ......................................................................................................... 26
  3.2 Meta-Analysis of DICCCOL .................................................................................................................. 30
  3.3 Applying DICCCOL on Brain Disease ................................................................................................... 31
4 HOLISTIC ATLASES OF FUNCTIONAL NETWORKS AND INTERACTIONS (HAFNI)........................................................................................................40

4.1 Rationale and Overview........................................................................................................40

4.2 Sparse Representations of fMRI Signals ..........................................................................41

4.3 Functional Network Component Analysis........................................................................43

4.4 Group-wise Consistent Functional Templates................................................................45

5 CONSTRUCTION OF HOLISTIC ATLASES BASED ON DICCCOL (DICCCOL-H)........................................................................................................50

5.1 Rationale and Overview.......................................................................................................50

5.2 DICCCOL Prediction and Functional Labeling ...............................................................51

5.3 Optimization on Structure and Function Simultaneously............................................51

BIBLIOGRAPHY.......................................................................................................................58
LIST OF TABLES

Table 1: Datasets for DICCCOL functional labelling .............................................................27
Table 2: Composition of “connectome signatures” ...............................................................38
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Overview</td>
<td>6</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Non-linearity of structural connectivity</td>
<td>8</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Trace-map model</td>
<td>11</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Examples of trace-map</td>
<td>12</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Comparison of trace-maps</td>
<td>13</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Validation of trace-map</td>
<td>14</td>
</tr>
<tr>
<td>Figure 7</td>
<td>DICCCOL optimization framework</td>
<td>16</td>
</tr>
<tr>
<td>Figure 8</td>
<td>One example of DICCCOL landmark</td>
<td>19</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Reproducibility and predictability of DICCCOL</td>
<td>22</td>
</tr>
<tr>
<td>Figure 10</td>
<td>The architecture of DICCCOL system</td>
<td>24</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Statistics of DICCCOL functional labelling</td>
<td>29</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Meta-analysis of DICCCOL</td>
<td>30</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Applications based on DICCCOL</td>
<td>31</td>
</tr>
<tr>
<td>Figure 14</td>
<td>Discrepant DICCCOLs in MCI</td>
<td>34</td>
</tr>
<tr>
<td>Figure 15</td>
<td>Classification using DICCCOL</td>
<td>37</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Illustration of connectome signatures</td>
<td>39</td>
</tr>
<tr>
<td>Figure 17</td>
<td>Computational pipeline of sparse representation of fMRI signals</td>
<td>42</td>
</tr>
<tr>
<td>Figure 18</td>
<td>HAFNI templates – part I</td>
<td>47</td>
</tr>
<tr>
<td>Figure 19</td>
<td>HAFNI templates – part II</td>
<td>48</td>
</tr>
</tbody>
</table>
Figure 20: HAFNI templates – part III .................................................................49
Figure 21: Illustration of joint optimization..........................................................52
Figure 22: Results of joint optimization ...............................................................55
Figure 23: Spatial distribution of optimized DICCOL-H landmarks.......................56
CHAPTER 1

INTRODUCTION

1.1 Thesis Statements

An unrelenting human quest regarding the brain science is: what is the intrinsic relationship between the brain’s structural and functional architectures, which partly defines what we are and who we are. Recent studies suggest that each brain’s cytoarchitectonic area has a unique set of extrinsic inputs and outputs, called the “connectional fingerprint” [1], which largely determines the functions that each brain area performs. This close relationship between structural connectivity pattern and brain function has also been confirmed and replicated in recent studies in the literature [1, 3] and our own works [4-9]. However, their explicit connections are largely unknown. For example, in what extent they are inclined to be coherent with each other and otherwise they will intend to show more heterogeneity [10]? We need a comprehensive understanding of the principles that regulate the information processing (function) in a particular structural pattern, and between the interacting structural units in the brain as a whole.

In current stage, Functional Magnetic Resonance Imaging (fMRI) [11-16] is the most popular method that can examine the functional activities of the whole brain when people performing a specific task or having rest, due to its non-invasive and in vivo nature. After decades of active research, there has been mounting evidence [17-21] that the total human brain function emerges from and is realized by the interaction of multiple
concurrent neural processes or networks, each of which is spatially distributed across specific structural substrate of neuroanatomical areas [22-23]. However, due to the lack of effective computational brain mapping approaches, it is very challenging to robustly and faithfully reconstruct concurrent functional networks from fMRI (either task fMRI or resting state fMRI) data and quantitatively measure their network-level interactions. In other words, many important hidden functional characteristics are ignored by current model based fMRI analysis and which might lead to potential inaccurate or biased comprehensions of cognitive response either in task or resting state.

The second challenging issue that has received intensive attention recently is how to establish a stable and neuroscience grounded foundation or substrates for synthetically and quantitatively measuring connectivity and dynamic interactions, either within individual brains or comparing them across populations. Due to the scarcity of ground-truth data, researchers have to validate and replicate data from multiple subjects so that sufficient statistical power can be achieved. However, this population-level data pooling step requires the determination of accurate correspondences between regions of interest (ROIs) across different brains, which is a major barrier in human brain mapping [24] and neuroimaging for several decades [25-29]. Despite the enormous efforts on exploring the most consistent structural connectivity patterns among different populations [5], nevertheless, it is still considerable challenging to integrate the structural information with the corresponding consistent functional profiles. This challenge is not only come from the remarkable individual variability of cortical anatomy, connection and function, but also come from the critical lack of effective computational model for robustly and comprehensively estimating the brain structural and functional relationship.
To tackle these longstanding challenges, by using diffusion tensor imaging (DTI), functional magnetic resonance imaging (fMRI) techniques and the state-of-the-art sparse learning method, we construct DICCCOL based holistic atlases (DICCCOL-H) that reflect both group-wise consistent structural connectivity patterns and functional homogeneity of the human brain.

1.2 Contributions

**Representation of common brain architectures by DICCCOL:** Each DICCCOL landmark possesses intrinsically-established structural/functional correspondences across individuals, and the DICCCOL map collectively offers a universal representation of common structural brain architectures across individuals and populations (chapter 2) [5]. Also, an effective and efficient DICCCOL prediction framework (chapter 2) [5,114] will automatically predict all DICCCOLs in a new, single brain, thus offering an individualized reference system for human brains and enabling numerous applications. In the absence of task-based fMRI data, for example, in cases where it is impractical to acquire large-scale fMRI data in real-world situations such as the MCI populations [4] mentioned in this thesis (chapter 3), the DICCCOL models can be used instead in predicting functional brain regions based on the widely available DTI data. In comparison with image registration, including group-wise and multi-atlases image registration algorithms, and cortical parcellation methods, the DICCCOL map and its prediction framework offer a novel solution to the immensely challenging problem of accurately localizing functional regions in individuals and automatically establishing their correspondences [5]. In comparison with Talairach or other atlases that encodes
functional localizations by stereotaxic coordinates, the DICCCOL brain reference system encodes functional localizations of common brain structures by consistent fiber connection patterns, which are much more reproducible and predictable across brains [5].

**Sparse representation of whole brain fMRI signals:** In most previous fMRI studies, researchers have mainly relied on a common practice of averaging or smoothing single fMRI signals within a neighborhood [115,116]. In this thesis, on the contrary, we decomposed fMRI signals into linear combinations of multiple components based on the sparse representation of whole-brain fMRI signals. This novel data-driven strategy naturally accounts for the fact that a brain region might be involved in multiple functional processes [17, 20, 21, 34] and thus its fMRI signal is composed of various components. Our results (chapter 4) have demonstrated that this novel strategy can effectively and robustly reconstruct multiple simultaneous functional networks, including both task-evoked networks and RSNs, which can be well reproduced across individual brains.

**Reconstruction of multiple concurrent interacting functional networks in the brain:** GLM-based activation detection and ICA-based clustering have been arguably the dominant methods in task fMRI and resting state fMRI data analyses. Unfortunately, both methods are limited in reconstructing multiple concurrent, interacting functional networks. As a consequence, it is still largely uncertain, during either task performance or resting state, whether/how functional networks spatially overlap/interact with each other and whether the functional brain architecture is composed of highly-specialized components, or it is general-purpose machinery. In this thesis (chapter 4), novel sparse representation and dictionary learning methodology effectively infer the spatial overlap patterns among those brain networks, which are represented by the time series of the
over-complete basis dictionaries. The results have also revealed the common and widespread spatial overlaps within and among both task-evoked and resting state networks.

**DICCCOL-H based optimization:** Different from the classic DICCCOL system, which optimizes and predicts each DICCCOL landmark only based on the group-wise structural connectivity consistency, the proposed DICCCOL-H (chapter 5) [117] will optimize the landmark’s structure and function simultaneously: the functional constraint is come from the group-wise agreement of those consistent functional network components identified by an innovative fMRI signals sparse representation. During the optimization, the landmarks will move towards the locations which possess more group-wise functional homogeneity. At the same time, the structural constraint ensures that the established structural similarity will not be destroyed. To our knowledge, this framework is the first one to formally consider the group-wise structural consistency and functional homogeneity at the same time.

1.3 Thesis Outline

As illustrated in Fig.1, this thesis contains four sections which are corresponding to chapter 2 to chapter 5. First, I will introduce DICCCOL system (chapter 2) and its applications (chapter 3), since it is the foundation of the proposed joint structural/functional optimization process; Then, an innovative sparse representation of the whole brain fMRI signals will be presented, in which the most consistent functional networks are recovered and identified as functional templates (chapter 4); Finally, we use these templates as functional regulations to optimize our previous DICCCOL system to
achieve the Holistic Atlases of Brain Structure and Function, also named as DICCCOL-H (chapter 5).

Fig.1. Overview. The four components are in colored boxes, which will be detailed in chapter 2-5.
CHAPTER 2
DICCCOL LANDMARKS

2.1 Motivation

When measuring structural and/or functional brain connectivity, network nodes, or regions of interests (ROIs), provide the structural substrates for measuring connectivity within individual brain and for comparing them across different populations [24]. Thus, identification of reliable, reproducible and accurate ROIs that are consistent across different brains is critically important for the success of connectivity mapping [6, 8, 9, 118]. However, from our perspective, determination of corresponding brain ROIs in different brains is perhaps one of the foremost challenges in human brain mapping, due to four critical reasons [24]. 1) The functional and/or cytoarchitectural boundaries between cortical regions are unclear [24, 28, 119]; 2) The individual variability of cortical structure and function is remarkable [24, 29]; 3) The properties of ROIs are highly nonlinear [6, 8, 9, 24]. For instance, a slight change of the size or location of a ROI might dramatically alter its structural and/or functional connectivity profiles (e.g., shown in Fig.2) [6]. 4) It is even more challenging to identify accurate ROIs in some brain disease patients in that the brain architecture might have been altered during neurodevelopment [120-125].
Current approaches for identifying ROIs in brain imaging can be broadly classified into four categories [24,126]. The first is manual labeling by experts based on their domain knowledge [127]. While widely used, this method is vulnerable to inter-subject and intra-subject variation and its reproducibility may be low. The second method is to cluster ROIs from the brain image itself and is data-driven [128-129]. However, these data-driven approaches are typically sensitive to the clustering parameters used, and their neuroscience interpretation is not clear. The third one is to predefine ROIs in a template brain, and warp them to the individual space using image registration algorithms [130, 131]. The accuracy of these atlas-based warping methods is limited due to the variability of neuroanatomy across different brains. The fourth method uses task-based fMRI paradigms to identify activated brain regions as network ROIs [132]. This methodology is regarded as the benchmark approach for ROI identification. However, task-based fMRI

*Fig.2. Non-linearity of structural connectivity. (a) The size changes to the yellow bubble, or the location moves from the red one. (b)-(d): Fiber connections (in white) before the movement (b), after the enlargement (c), and after the movement (d).*
is demanding and time-consuming [120, 112] and it is impractical to acquire extensive fMRI data for large-scale brain networks.

In response to the challenges of mapping a common brain architecture and inspired by the connectional fingerprint concept [1] and fiber clustering literature [2,133,135], we hypothesize that there is a common human brain architecture that can be effectively represented by group-wise consistent structural fiber connection patterns. To test this hypothesis, I extensively extended my previous work (section 2.2) [6] which used DTI datasets to discover the dense and common cortical landmarks likely present across all human brains. We have dubbed this strategy: Dense Individualized and Common Connectivity-based Cortical Landmarks (DICCCOL). The basic idea is that we optimize the localizations of each DICCCOL landmark in individual brains by maximizing the group-wise consistency of their white matter fiber connectivity patterns. This approach effectively and simultaneously addresses the above-mentioned three challenges in the following ways. 1) The DICCCOLs provide intrinsically-established correspondences across subjects, which avoids the pitfall of seeking unclear cortical boundaries. 2) Individual structural variability is effectively addressed by directly determining the locations and sizes of DICCCOL landmarks in each individual’s space. 3) The nonlinearity of cortical connection properties is adequately addressed by a global optimization and search procedure, in which group-wise consistency is used as an effective constraint.
2.2 Trace-Map Model

**Bundle description based on the trace-map model:** Many algorithms, such as spectral clustering [133], normalized cut clustering [134] and atlas-based clustering [2], have been developed to cluster white matter fibers into different bundles. However, an open problem remains: how can a fiber bundle be described quantitatively? In this thesis, we need a quantitative fiber bundle descriptor or model to represent fibers and compare their similarities within and across different subjects. Hence, we have proposed a novel method by which to describe the fiber bundle; we call this method the trace-map model.

The core idea of trace-map contains three steps: first, each fiber curve was divided into segments and each segment was composed of a collection of points. Then, the Principal Component Analysis (PCA) was used to find the principal direction of each segment, represented as a vector as showed in Fig. 3(a). Finally, the vectors were translated to the origin of a global spherical coordinate system and shoot from the origin to the surface of a unit sphere centered at the origin. In this way, we can have a trace point on the sphere, and then the same procedure was performed on the segments of all other fibers in each bundle (Fig. 3(b) and (c)). Fig. 3(d) shows two examples. The top image is a U-shape fiber bundle and its corresponding trace-map. The bottom image is a line-shape case.

There are two issues to be noted here. One is that all subjects’ brains must be aligned. In our implementation, the principal direction of each brain was calculated using PCA. This principal direction was then used to align different brains into a randomly selected template subject. Thus, fiber bundles with similar shapes but different orientations can be differentiated by the different trace-point distributions on the standard sphere surface. The second issue is that one of the two ends of the fiber bundles needs to
be assigned as the start point. Since each fiber bundle was extracted from a small region on the cortical surface, we selected the end that was closest to the center of the region. This is very important to ensure that the trace-maps of one fiber at different optimization procedures are consistent.

The proposed trace-map model has the following advantages. 1) It is an effective way to represent and compare fiber bundles. Essentially, the trace-map model transforms a fiber bundle to a set of points distributed on the surface of a unit sphere. It projects the complex, geometric features of the fiber onto point distribution patterns in a standard space, in which different fiber bundles from different subjects can be compared quantitatively. The patterns reflect the accumulation of the strength of the fiber bundle in different directions. To a certain extent, it is similar to the idea of inflating the convoluted cortical surface onto a standard sphere: after projecting the cortical surface to a standard

---

Fig.3. Trace-map model. (a) Calculation of the principal direction for one segment of each fiber. (b) Each segment could be represented by a series of vectors. (c) After translation to the origin of a global coordinate system, each vector shoots to a unit sphere whose center is the origin. (d) Two examples of fiber bundles and their trace-maps. The top row is a U-shape fiber bundle example and the bottom row is a line-shape one. For both cases, the left are fiber bundles and the right are their trace-map representations.
sphere surface, the folding patterns across different subjects can be compared and analyzed. 2) The trace-map model is not sensitive to the small variations of the fiber bundles. This is a very important property when performing comparisons across different subjects, because we are more interested in comparing the overall shapes of the fiber bundles.

**Fiber bundle comparison based on trace-map model:** Our rationale for comparing fiber bundles through trace-maps is that similar fiber bundles have similar overall trace-map patterns. Fig. 4 shows four examples. Figs. 4(a) and 4(b) are a pair of fiber bundles that are similar by visual inspection. We can see that their trace-maps, Fig. 4(e) and 4(f), are also similar. Fig. 4(c) and 4(d) show another pair of similar fiber bundles and their corresponding trace-maps are shown in Fig. 4(g) and 4(h). Again, we can clearly see the similar patterns of the point distributions in the trace-maps.
After arriving at a trace-map representation of the fiber bundles, the bundles can be compared by defining the distance between their corresponding trace-maps, as shown in Fig. 5. For each point, $P_1$, in one trace-map, its corresponding location, $P_{1'}$, in the other trace-map can easily be found in terms of the same location. The point density, denoted by $\text{den}(P_1)$, is then calculated as follows:

$$\text{den}(P_1) = \frac{n_1}{N}$$

$n_1$ is the number of points in the trace-map whose center is $P_1$ with radius $d$, which is in the range of 0-1.0 since the standard sphere onto which we project the fiber bundles is a unit sphere surface; in this work we empirically choose $d=0.3$. $N$ is the total number of points in the trace-map. As shown in Fig. 5, we calculate the point density in the red circle. The total distance of two trace-maps is defined as:

$$D(T_1, T_2) = \frac{\sum_{i=1}^{n} |\text{den}(P_i) - \text{den}(P_{i'})|}{n} + \frac{\sum_{j=1}^{m} |\text{den}(P_j) - \text{den}(P_{j'})|}{m}$$

$T_1$ and $T_2$ are two trace-maps. $P_i$ is a point in $T_1$ and $P_{i'}$ is its corresponding point in $T_2$ which share the similar location. $P_j$ is a point in $T_2$ and $P_{j'}$ is its corresponding point in $T_1$. $n$ and $m$ are numbers of points in $T_1$ and $T_2$ respectively. Intuitively, Eq. (2) means...
that we iterate over all data points in one trace-map, and measure the density within a circle centered at the data point in consideration and also a circle placed in the corresponding location in the other trace-map. This iterative process is repeated over the data points in the other trace-map, and the same procedure is iterated over all possible locations in each trace-map. Notably, we simplified the computation by only considering locations where a data point is present in one or the other trace-map.

To evaluate the effectiveness and distinctiveness of the trace-map model, we randomly chose a subject and extracted the fiber bundles from all possible ROIs whose centers are the vertices of the cortical surface with a certain scale of neighborhood (4-ring mesh vertex neighborhood in this work). The fiber bundles were then represented by trace-maps and the distances between the trace-maps of the selected ROIs and the rest were calculated. The distance between the trace-map of the selected ROI and the trace-maps of all other ROIs on the cortical surface are shown in Fig. 6.

From the result, we can see that: 1) most of the fiber bundles emanating from other ROIs have significant differences in comparison with the selected ROI. That is, most of the regions in the cortex are blue. 2) Considering the small neighborhood of the ROI we
chose, the trace-map distances between the selected ROI and others roughly follow a Gaussian distribution. This result suggests that the trace-map of an ROI is quite distinctive, which is critical to unambiguously characterize the current ROI.

2.3 Whole Brain Optimization

**Initialization and overview of the DICCCOL discovery framework:** Similar to the above mentioned work, we randomly selected one subject as the template and generated a dense, regular map of 3D grid points within the boundary box of the reconstructed cortical surface. The intersection locations between the grid map and the cortical surface were used as the initial landmarks. As a result, we generated 2056 landmarks on the template (Figs. 7a-7b). Then, we registered this grid of landmarks to other subjects by warping their T1-weighted MRI images to the same template MRI image using the linear registration algorithm FSL FLIRT. This linear warping is expected to initialize the dense grid map of landmarks and establish their rough correspondences across different subjects (Figs. 7a-7b). The aim of this initialization was to create a dense map of DICCCOL landmarks distributed over major functional brain regions.

Then, we extracted white matter fiber bundles emanating from small regions around the neighborhood of each initial DICCCOL landmark (Figs. 7c-7g). The centers of these small regions were determined by the vertices of the cortical surface mesh, and each small region served as the candidate for landmark location optimization. Fig. 7d shows examples of the candidate fiber bundles we extracted. Afterwards, we projected the fiber bundles to a standard sphere space, called trace-map [9, 10], as shown in Fig. 7e, and calculated the distance between any pair of trace-maps in different subjects within the
group. Finally, we performed a whole space search to find one group of fiber bundles (Fig. 7f) which gave the least group-wise variance. Fig. 7g shows examples of the optimized locations (red bubble) and the DICCCOL landmark movements (yellow arrow).

Fig. 7. DICCCOL optimization framework. (a)-(b): Illustration of landmark initialization among a group of subjects. (a) We generated a dense regular grid map on a randomly selected template. (b) We registered this grid map to other subjects using linear registration algorithm. The green bubbles are the landmarks. (c)-(g): The workflow of our DICCCOL landmark discovery framework. (c) The corresponding initialized landmarks (green bubbles) in a group of subjects. (d) A group of fiber bundles extracted from the neighborhood of the landmark. (e) Trace-maps corresponding to each fiber bundle. (f)
The optimized fiber bundle of each subject. (g) The movements of the landmarks from initial locations (green) to the optimized locations (red). Step (1): Extracting fiber bundles from different locations close to the initial landmark. Step (2): Transforming the fiber bundles to trace-maps. Step (3): Finding the group of fiber bundles which make the group variance the least. Step (4): Finding the optimized location of initial landmark (red bubble). (h)-(j): Illustration of trace-map distance. (h) A sphere coordinate system for finding the sample points. We totally have 144 sample points by adjusting angle $\Phi$ and $\theta$. (i) A sphere with 144 sample points. (j) Two trace-maps. The two red circles belong to the same sample point and will be compared based on the point density information within red circles.

**Optimization of landmark locations:** We formulate the problem of optimization of landmark locations and sizes as an energy minimization problem, which aims to maximize the consistency of structural connectivity patterns across a group of subjects. By searching the whole space of landmark candidate locations and sizes, we can find an optimal combination of new landmarks that ensure the fiber bundles from different subjects have the least group variance. Mathematically, the energy function we want to minimize is defined as:

$$E(S_1, S_2, \ldots, S_m) = \sum_{k \neq l} E(S_k, S_l), \quad k \neq l, \quad l = 1, 2, \ldots, m$$  \hspace{1cm} (3)

$S_1 \ldots S_m$ are $m$ subjects. We let $E(S_k, S_l) = D(T_k, T_l)$, and rewrite the Eq. (3) as below:

$$E(S_1, S_2, \ldots, S_m) = \sum_{k \neq l} \sum_{i=1}^{n} (T_{ki} - T_{li})^2, \quad k \neq l, \quad l = 1, 2, \ldots, m$$  \hspace{1cm} (4)

For any two subjects $S_k$ and $S_l$, we transformed them to the corresponding vector format, $T_k$ and $T_l$, of trace-maps. $T_{ki}$ and $T_{li}$ are the $i$th element of $T_k$ and $T_l$ respectively. Intuitively, we aim to minimize the group distance among fiber shapes defined by trace-maps here.

In our implementation, for each landmark of the subject, we examined around 30 locations (surface vertices of 5-ring neighbors of the initial landmark) and extracted their corresponding emanating fiber bundles as the candidates for optimization. Then, we
transformed the fiber bundles to trace-maps. After representing them as vectors, we calculated the distance between any pair of them from different subjects. Thus, we can conduct a search in the whole space of landmark location combinations to find the optimal one which has the least variance of fiber bundles shapes within the group. The optimization procedure (Eq.(4)) is performed for each of those 2056 initial landmarks separately.

**Determination of consistent DICCCCOLs:** Ten subjects were randomly selected and were equally divided into two groups. The above mentioned steps were performed separately in these two groups. Due to that the computational cost of landmark optimization procedure via global search grows exponentially with the number of subjects used [6], we can more easily deal with 5 subjects in each group at current stage. As a result, we obtained two independent groups of converged landmarks. For each initialized landmark in different subjects in two groups, we used both quantitative (via trace-map) and qualitative (via visual evaluation) methods to evaluate the consistency of converged landmarks. First, for each converged landmark in one group, we sought the most consistent counterparts in another group by measuring their distances of trace-maps and ranked the top 5 candidates in the decreasing order as possible corresponding landmarks in two groups. Then, we used an in-house batch visualization tool (illustrated in Fig. 8) to visually examine all of the top 5 landmark pairs in two separate groups. If the fiber shape patterns were determined to be the most consistent across two independent groups, the landmark pair was determined as a DICCCOL landmark. In addition, the trace-map distances between any pair of DICCCOL landmarks across subjects were also checked to verify that the landmark was similar across groups of subjects. Finally, we
determined 358 DICCCOL landmarks by two experts independently by both visual evaluation and trace-map distance measurements, and a third expert independently verified these results. If any of the subjects in two separate groups exhibited substantially different fiber shape pattern, that landmark was discarded. Therefore, all of the discovered 358 DICCCOLs were independently confirmed in two different groups of subjects, and their fiber connection patterns turned out to be very consistent. The visualizations of all 358 DICCCOLs are released online at: http://dicccol.cs.uga.edu.

2.4 DICCCOL Prediction

It has been shown in the literature that prediction of functional brain regions via DTI data has superior advantages since a DTI scan takes less than ten minutes and is widely available [114]. Here, we are motivated to predict the 358 DICCCOL landmarks in a single subject’s brain. The prediction of DICCCOLs is akin to the optimization procedure...
in Section 2.3. We will transform a new subject (on MRI image via FSL FLIRT) to be predicted to the template brain which was used for discovering the DICCCOLs and perform the optimization procedure following the Eq. (4). It is noted that there is a slight difference from Section 2.3 since we already have the locations of DICCCOLs in the model brains. Therefore, we will keep those DICCCOLs in these models unchanged, and optimize the new subject only to minimize the trace-map difference among the new group including the models and the subject to be predicted. Specifically, \( S_{m1}, S_{m2}, \ldots, S_{m10} \) and \( S_p \) represent the model dataset and the new subject to be predict, respectively.

Formally, we summarize the algorithm as bellow:

1) We randomly select one case from the model dataset as a template \( (S_{mi}) \), and each of the 358 DICCCOL landmarks in the template is roughly initialized in \( S_p \) by transforming them to the subject via a linear registration algorithm FSL FLIRT.

2) For \( S_p \), we extract white matter fiber bundles emanating from small regions around the neighbourhood of each initialized DICCCOL landmark. The centers of these small regions will be determined by the vertices of the cortical surface mesh, and each small region will serve as the candidate for landmark location optimization.

3) For \( S_{mi} \), each of the 358 model DICCCOLs will be fixed for the optimization.

4) We project the fiber bundles of the candidate landmarks in \( S_p \) to a standard sphere space, called trace-map, as shown in Fig. 3. For each landmark to be optimized in \( S_p \), we calculate the trace-map distances between the candidate landmark and those DICCCOL landmarks in the model subjects within the group.
5) For each landmark, we performed a whole space search to find one group of fiber bundles which gives the least group-wise variance. The candidate landmark in $S_p$ with the least group-wise variance is selected as the predicted DICCCOL landmark.

As we can see, even though the prediction is an exhaustive search algorithm in which the performance is dependent on how many candidates we choose from $S_p$, it can be finished within linear time because we will not move the DICCCOLs in the model brains. Therefore, the DICCCOL prediction in a new brain with DTI data is very fast, typically around ten minutes on a desktop computer.

2.5 The Advantages of DICCCOL

**Reproducibility and predictability:** The 358 DICCCOLs were identified via a data-driven whole brain search procedure (Sections 2.2-2.4) in ten randomly selected subjects (equally and randomly divided into two independent groups), as shown in Fig. 9a. As an example, we randomly selected five DICCCOLs (five enlarged color spheres in Fig. 9a) and plotted their emanating fibers in these ten brains (Figs. 9b-9f). It can be clearly seen that the fiber connection patterns of the same landmark in ten brains are very consistent, suggesting that DICCCOLs represent common structural brain architecture. Importantly, by visual inspection, all of these 358 DICCCOLs have consistent fiber connection patterns in these ten brains. For more details, the visualization of all of these 358 landmarks is available online at: http://dicccol.cs.uga.edu. In addition to visual evaluation, we quantitatively measured the differences of fiber shape patterns represented by the trace-maps (Section 2.2) for each DICCCOL within and across two groups (Figs.
Fig.9. Reproducibility and predictability of DICCCOL. (a): The 358 DICCCOLs. (b)-(f): DTI-derived fibers emanating from 5 landmarks (enlarged color bubbles in (a)) in 2 groups of 5 subjects (in 2 rows) respectively. (g)-(k): The predicted 5 landmarks in 2 group of 5 subjects (in 2 rows) and their corresponding connection fibers. (l): Average trace-map distance for each landmark in the first group (rows in (b)-(f)); the color bar is on top of (o)-(p). (m): Average trace-map distance for each landmark in the second group (rows in (b)-(f)); (n): Average trace-map distance for each landmark across 2 groups in (b)-(f); (o)-(p): Average trace-map distance for each landmark in the 2 predicted groups in (g)-(k), respectively. (q): The decrease fraction of trace-map distance before and after optimization (the color bar on the top of (q)). The initialization was performed via a linear image warping algorithm.

The average trace-map distance is 2.19, 2.05 and 2.15 using Eq. (4). It is evident that the quantitative trace-map representations of fiber bundles for each DICCCOL has similar patterns within and across two separate groups, demonstrating the consistency of DICCCOL’s fiber connection patterns.

In addition to the remarkable reproducibility of each DICCCOL in Figs. 9b-9f, the 358 DICCCOLs can be effectively and accurately predicted in a single, separate brain
with DTI data, as exemplified in Figs. 9g-9k. Here, each landmark was predicted in ten separate test brains (Figs. 9g-9k) based on the template fiber bundles of corresponding landmarks (Figs. 9b-9f). We can clearly see that the predicted landmarks have quite consistent fiber connection patterns in these test brains (Figs. 9g-9k) as those in the template brains (Figs. 9b-9f), indicating that the DICCCOLs are predictable across different brains. Quantitatively, the predicted landmarks have similar quantitative trace-map patterns as those in the template brains, as shown in Figs. 9o-9p. The average trace-map distance is 2.27 and 2.17. As a comparison, the predicted landmarks have much more consistent fiber trace-map patterns than the linearly registered ones via FSL FLIRT (Fig. 9q). The average decrease fraction of trace-map distance is 15.5%. These results support the DICCCOL as an effective, quantitative representation of common structural brain architecture that is reproducible and predicable across subjects and populations.

**Unified ROI Solution:** As summarized in Fig. 10, our data-driven discovery approach has identified 358 DICCCOLs that are consistent and reproducible across over 143 brains based on DTI data. Extensive studies have shown that these 358 landmarks can be accurately predicted across different subjects and populations. Our work has demonstrated that there is deep-rooted regularity in the structural architecture of the human brain, which has been jointly and spontaneously encoded by the DICCCOL map. The DICCCOL map has been evaluated by four independent multimodal fMRI and DTI datasets which contained over 143 subjects covering different age groups, i.e., adolescent, adult, and elderly. In total, 121 consistent and stable functional ROIs derived from eight task-based fMRI network (auditory, attention, emotion, empathy, fear, semantic decision making, visual and working memory networks) and one R-fMRI network (default mode...
Fig. 10. The architecture of DICCCOL system. Spheres in orange (total 6), red (total 8), brown (total 9), pink (total 8), blue (total 27), yellow (total 14), cyan (total 14), purple (total 16), and black-red (total 19) colors stand for landmarks in empathy, default mode, visual, auditory, attention, working memory, fear, emotion, and semantic decision making networks that are identified from fMRI datasets. The green spheres (totally 263) stand for landmarks that are not functionally-labeled yet. The DICCCOLs serve as structural substrates to represent the common human brain architecture. For instance, nine different functionally-specialized brain networks ((b)-(j)) identified from different fMRI datasets are integrated into the same universal brain reference system (a) via DICCCOL. Then, the functionally-labeled DICCCOLs in the universal space can be predicted in each individual brain with DTI data such that the DICCCOLs and their functional identities can be readily transferred to a local coordinate system (k).

network), shown in Fig. 10b-10j, were used to functionally label the predicted DICCCOLs for individuals (section 3.1). Our extensive experimental results demonstrated that the DICCCOL representation of functional ROIs is accurate, robust, consistent and reproducible in multiple multimodal fMRI and DTI datasets. With the
universal DICCCOL brain reference system, different measurements of the structural and functional properties of the brain, e.g., morphological measurements derived from structural MRI data and functional measurements derived from fMRI data, can be reported, integrated, and compared within the DICCCOL reference system. For instance, we can report fMRI-derived activated regions by their corresponding closest DICCCOL IDs, instead of their stereotaxic coordinates in relation to the Talairach or MNI coordinate system.

In a broader sense, the DICCCOL map provides a general platform to aggregate and integrate functional networks from different multimodal DTI and fMRI datasets to the universal DICCCOL map, the sum of which can then be transferred to a new, separate individual or population via DTI data. For instance, the functional labeling of a portion of the DICCCOLs in an individual dataset, e.g., in Fig. 10b-10j, can be readily transferred to the universal template space (Fig. 10a), and then be propagated to other individual brains, as shown in Fig. 10k. In this way, specific functional localizations on the DICCCOL map achieved in one multimodal fMRI and DTI dataset (e.g., Fig. 10b-10j) can contribute to the same functional localization problem in other brains, once DTI data, on which the DICCCOL map prediction can be accurately performed, is available (e.g., Fig. 10k). This common DICCCOL platform offers an alternative approach and can be complementary to current methods, such that contributions from different labs can be effectively integrated and compared.
CHAPTER 3
APPLICATIONS OF DICCCOL

3.1 Functional Labelling of DICCCOL

**Data:** In total, we used four different multimodal DTI/fMRI datasets for functional labelling of the DICCCOL map, as summarized in Table 1. In brief, dataset 1 included the DTI, R-fMRI (resting-state fMRI), and five task-based fMRI scans of eleven healthy young adults recruited at The University of Georgia (UGA) Bioimaging Research Center (BIRC) under IRB approval. The scans were performed on a GE 3T Signa MRI system using an 8-channel head coil at the UGA BIRC. The five task-based fMRI scans were based on in-house verified paradigms including emotion, empathy, fear, semantic decision making, and working memory tasks at UGA BIRC. The dataset 2 included twenty three healthy adult students recruited under UGA IRB approval. Working memory task-based fMRI and DTI scans were acquired for these participants at the UGA BIRC. The dataset 3 included twenty elderly healthy subjects recruited and scanned at the UGA BIRC under IRB approval. Multimodal DTI and Stroop task-based fMRI datasets were acquired using the same imaging parameters as those in datasets 1 and 2. The dataset 4 included multimodal DTI, R-fMRI and task-based fMRI scans for 89 subjects including three age groups of adolescents (28), adults (53) and elderly participants (23). These participants were recruited and scanned on a 3T MRI scanner in West China Hospital, Huaxi MR Research Center, Chengdu, China under IRB approvals.
Datasets | Types | Networks
---|---|---
Dataset 1 | DTI, R-fMRI, five task-based fMRI scans | Emotion, Empathy, Fear, Semantic decision making, Working memory
Dataset 2 | DTI, one task-based fMRI scan | Working memory
Dataset 3 | DTI, one task-based fMRI scan | Attention
Dataset 4 | DTI, R-fMRI, two task-based fMRI scans | Default mode, Visual, Auditory

**Table 1.** Datasets for DICCCOL functional labelling.

**Functional localizations of DICCCOLs:** In total, we were able to identify 121 functional ROIs that were consistently activated from nine brain networks (working memory, default mode, auditory, semantic decision making, emotion, empathy, fear, attention, and visual networks) based on the fMRI datasets. To examine the functional co-localizations of 358 DICCCOLs, we mapped the 121 functionally-labelled brain ROIs onto the DICCCOL map. Surprisingly, 95 out of the 358 DICCCOLs were consistently co-localized in one or more functional brain networks determined by fMRI datasets across different subjects and/or populations (see Fig. 11). Specifically, 76 of them are located adjacently to one functional network, 16 of them are located within two functional networks, and 3 of them are located inside three functional networks.

To quantitatively evaluate the functional localization accuracy by the 95 DICCCOLs, we measured the Euclidean distance between the centers of each DICCCOL and each fMRI-derived landmark, and reported the results in Fig. 11 [5]. There are 9 sub-figures corresponding to the 9 functional networks identified using fMRI datasets, that is, working memory (Fig. 11a), default mode (Fig. 11b), auditory (Fig. 11c), semantic decision making (Fig. 11d), emotion (Fig. 11e), empathy (Fig. 11f), fear (Fig. 11g), attention (Fig. 11h), and visual networks (Fig. 11i) respectively. In each sub-figure, the
fMRI-derived landmarks are highlighted by white spheres, while the corresponding DICCCOLs are highlighted in other colors. The distances (measured in mm) between the centers of fMRI landmarks and DICCCOLs are shown in the bottom panel, in which the horizontal axis indexes activations and the vertical axis is the distance in the unit of mm. Each bar represents the median (interface between the red and yellow bars), minimum and maximum value (two ends of the white line), 25% (bottom of the red bar) and 75% (top of the yellow bar) of the distances for each fMRI activation peak. The average distances for the nine functional networks are 6.07 mm, 5.43 mm, 6.48 mm, 6.25 mm, 6.12 mm, 6.41 mm, 5.93 mm, 5.94 mm, and 7.59 mm respectively. On average, the distance is 6.25 mm. The results in Fig. 11 demonstrate that the DICCCOLs are consistently co-localized with functional brain regions, and the DICCCOL map itself offers an effective and quantitative representation of common functional brain architecture that is reproducible across subjects and populations.
Fig. 11. Statistics of DICCCOL functional labelling. Specifically, 76 of them are located adjacently to one functional network, 16 of them are located within two functional networks, and 3 of them are located inside three functional networks. (a): Working memory network (dataset 2). White spheres represent fMRI-derived benchmarks, and yellow spheres represent corresponding DICCCOLs. The distances between centers of fMRI benchmarks and DICCCOLs are shown in the bottom panel, in which the horizontal axis indexes activations and the vertical axis is the distance in the unit of mm. Each bar represents the median (interface between the red and yellow bars), minimum and maximum value (two ends of the white line), 25% (bottom of the red bar) and 75% (top of the yellow bar) of the distances for each fMRI activation peak. The average distance is 6.07 mm. (b)-(i): results for default mode (dataset 1), auditory (dataset 4), semantic decision making (dataset 1), emotion (dataset 1), empathy (dataset 1), fear (dataset 1), attention (dataset 3), visual networks (dataset 4), respectively. In (b)-(i), white spheres stand for fMRI benchmarks and other colors represent corresponding DICCCOLs. The average distances between centers of fMRI benchmarks and DICCCOLs in these networks are 5.50 mm, 6.48 mm, 6.25 mm, 6.12 mm, 6.41 mm, 5.93 mm, 5.94 mm, and 7.59 mm, respectively.
3.2 Meta-Analysis of DICCCOL

Based on our previous work [5, 137], we successfully labeled DICCCOLs with corresponding functional roles (involved functional networks) through meta-analysis. In brief, we registered the average coordinates of each DICCCOL to a standard atlas space and searched in a small range to check if any functional task activation reports for this location existed [137]. If one or more activation reports were found in the considered range they were assigned to this DICCCOL as the corresponding functional roles. All the functional tasks (networks) used to label DICCCOLs were divided into five categories: action, perception, cognition, interoception and emotion [138]. For example, action
includes eight sub-functional networks such as execution, imagination and inhibition. In total we labeled 339 DICCCOLs with 55 sub-functional networks.

3.3 Applying DICCCOL on Brain Disease

We already successfully applied DICCCOL on multiple brain disease or disorders as showed in Fig.13, such as Mild Cognitive Impairment (MCI) [4], Post-traumatic Stress Disorder (PTSD) [60], Schizophrenia (SZ) [51], Prenatal Cocaine Exposure (PCE) [7] and many others [52]. Here, I will use MCI as an example to illustrate how to perform connectivity analysis based on DICCCOL landmarks.

**Discrepant DICCCOLs – landmarks with white matter alterations in MCI:**

Many previous studies have shown that some structure alterations including gray matter loss and/or white matter disruptions can be repetitively observed in MCI across different datasets and labs. Since DICCCOLs are defined based on the group-wise consistency of white matter profiles, we hypothesized that the DICCCOLs related to those altered white matter bundles would show different patterns between MCI patients and aged controls. In our experiment, we used two independent datasets including the MCI patients and aged

Fig.13. Applications based on DICCCOL.
controls. After applying the DICCCOL prediction procedure [5] on both of them, some predicted DICCCOLs showed abnormal characteristics compared to the others. That is, these DICCCOLs displayed higher group trace-map distances with MCI patients compared to the aged controls, indicating their fiber bundle patterns have higher variability at these locations. By using simple t-tests to evaluate and explore those abnormal DICCCOLs that have significantly (p=0.05) higher distributions of trace-map distance in MCIs, we obtained 56 and 95 discrepant DICCCOLs for the two datasets. The results are illustrated in Fig.14.

These discrepant DICCCOLs are plotted on the cortical surface using green and red bubbles (Fig 14(a) and (b)) for two independent datasets. Though distributed over the whole cortex, they still show some clear assembling patterns and, as we expected, most of them are located in areas which are consistent with previous findings: orange and purple arrows show some DICCCOLs located at the cingulate region and entorhinal cortex [139-142], respectively. The magenta arrows highlight the prefrontal areas [143] and dorsal part of the cortex which might be involved in the alteration of the corpus callosum [144-146]. In general, the discrepant DICCCOLs are located near the regions that have previously been proved to be associated with atrophy/alteration of either GM or WM. One discrepant DICCCOL was randomly selected within each dataset and its corresponding fiber bundles were shown on the top and bottom of Fig.14 (a) (dataset 1) and (b) (dataset 2) as examples. The locations of the selected ones are marked with black circles. Examples of severely altered white matter bundles of MCI patients are highlighted with red boxes. To quantitatively measure the difference of those discrepant DICCCOLs between MCI patients and aged controls, the average value and the standard
deviation of the trace-map distance within aged controls and MCI patients of the two datasets were calculated and displayed in Fig. 14(c). From this visualization, we can see that the average trace-map distances of MCI patients are significantly higher than those of aged controls (p<0.05).

Note that those discrepant DICCCOLs are no longer capable of providing consistent structural connectivity patterns, their intrinsic correspondences across different individuals are much less accurate. Hence in the following classification and functional network analysis, we discarded these discrepant DICCCOLs and constructed the functional connectomes only based on those “normal” DICCCOLs. However, those discrepant DICCCOLs may warrant further investigations.
Fig. 14. Discrepant DICCCOLs in MCI. In total, we obtained 56 for Dataset 1 and 95 for Dataset 2. The discrepant DICCCOLs are displayed as colored bubbles on the cortex. We randomly chose one as an example and showed the white matter bundles extracted from the selected DICCCOL on the top and bottom of (a) (dataset 1) and (b) (dataset 2). Some cases with significant differences between MCI subjects and normal controls are highlighted using red boxes. Sub-figure (c) shows the comparison of trace-map distance of the discrepant DICCCOLs between MCI and normal controls.
Classification based on functional connectomes: for each preserved DICCCOL, we can effectively acquire its fMRI time series by averaging in a small neighborhood (3 rings of surface mesh and the radius is approximate 3mm). We evaluated the functional connectivity (FC) between each pair of DICCCOLs with Pearson correlation coefficients and constructed an M×M symmetric matrix for later analysis. Here, M equals the number of preserved DICCCOLs. For dataset 1 M=302, and M=263 for dataset 2.

Since we only have two classes (MCI subjects and normal controls), we adopted a simple t-test (p<0.05) in the first stage to remove the connectivity without significant differences between two disease/control classes. However, the t-test evaluates the features separately, which means it does not consider the relevance among the features and thus it cannot capture the redundancy of these preserved features. To tackle this problem, we employed the Correlation-based Feature Selection (CFS) [147] algorithm as the second-stage feature selection. The core idea of CFS is that through a heuristic process it evaluates the merit of a subset of features by considering the goodness of individual features for predicting the class along with the degree of inter-correlation among them. Unlike the first stage t-test, CFS will compute feature-class and feature-feature correlations simultaneously. Given a feature subset S with k features, the Merit_s is defined as follows:

$$Merit_s = \frac{k\overline{Corr}(c,f)}{\sqrt{k+(k-1)\overline{Corr}(f,f)}}$$ (5)

where $\overline{Corr}(c,f)$ and $\overline{Corr}(f,f)$ are the mean feature-class correlation and the average feature-feature inter-correlation, respectively.

Once we obtained the most common and discriminative connectivity following this two-stage feature selection procedure, a support vector machine classifier [148] with
linear kernel was employed for solving the classification problem. Due to the limited numbers of subjects in the two datasets, we adopted the commonly used “leave-one-out” cross-validation strategy to evaluate the sensitivity (proportion of patients correctly predicted) and specificity (proportion of healthy controls correctly predicted) of our selected features.

We summarized the number of preserved features at each feature selection stage and its corresponding classification results in Fig. 15. In general, we had 45451 and 34453 features (pair-wise functional connectivity) initially. After the first stage feature selection (t-test) those features with no significant differentiation power were discarded and 2106/3691 connections passed through the significance test for Datasets 1 and 2, respectively. Interestingly, there were 134 common features across both datasets and they were treated as the input for the second stage feature selection (CFS). The feature training in the second stage was conducted within each dataset and all the subjects in the dataset were used for the training process. After that, we achieved 33 and 45 connectivity patterns for the two datasets, which served as “connectomics signatures” for the subsequent disease/control classification and neuroscience interpretation. One important issue that should be noted is that some useful features may also be discarded in the first stage given the fact that a subset of features could have strong differentiation power together even if they could not pass the significance test alone. However, due to the computation power we did not utilize feature-feature relations given the large search space. On the other hand, if these features made it through the first stage feature selection, they would be captured by the CFS (second stage) algorithm.
To better demonstrate the advantages of our method, we reported not only the classification accuracy with the final feature set, but also the number of survived connectivities and the intermediate classification results at each stage of the whole feature selection procedure. With the most relevant and discriminative connectivity selected, the classification accuracy substantially improved and we achieved 100% and 95.8% accuracies for the two datasets. Specifically, using the common (134, in total) functional connectivity of the t-test results from the two datasets, the classification accuracy was not decreased (the accuracy did not change for Dataset 1 and improved for Dataset 2). This supports the feasibility of using common connectivity of different datasets to constrain the feature space in the following step. The functional networks involved in the finally-preserved functional features (functional connectomes) were analyzed in the next section.

**Fig.15.** Classification using DICCCOL.

**Functional connectomes with high differentiation power:** Based on the meta-analysis results (section 3.2) of DICCCOL, we quantitatively analyzed the composition of the acquired “connectome signatures” and the details are summarized in Table 2 and Fig. 16. Not surprisingly, cognition-related DICCCOLs played the most critical role
within the signatures in both datasets. Those DICCCOLs involved in perception tasks also stood for a relatively high proportion (Table 2).

<table>
<thead>
<tr>
<th>Involved Connectivity</th>
<th>Involved DICCCOLs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tota</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Dataset 1</td>
<td>33</td>
</tr>
<tr>
<td>Dataset 2</td>
<td>45</td>
</tr>
</tbody>
</table>

Table 2. Composition of “connectome signatures”.

Fig. 16 is a visual presentation of the “connectome signatures”. The green ticks in the middle ring indicate 358 DICCCOLs and they are roughly arranged according to the axial projection of the cortex surface: from top to bottom the ticks represent the DICCCOLs located at frontal, parietal, temporal and occipital lobes. The red and green curves represent increased and decreased connectivities, respectively. From the figure, we can see many increased ones (red curves) in both datasets. In fact in dataset 2, only one decreased connectivity exists. This result is consistent with previous studies [149-151] that increased connectivity is a common symptom in MCI and early stage AD, which is interpreted as a compensatory mechanism for reallocation or recruitment of cognitive resources to maintain routine performance in MCI/AD patients.
Fig.16. Illustration of connectome signatures. Representation of the derived “connectome signatures” (left) and their corresponding Functional Relation Matrix (FRM) (right). The green ticks in the middle ring indicate 358 DICCCOLs. The red and green curves represent increased and decreased connectivity, respectively. Connectivity histogram shows the degree of connectivity at a specific DICCCOL. Four colored rings in the outer layer represent four categories of functional networks: perception, action, cognition and emotion. The heat map between DICCCOLs and the functional network shows the total frequency of involvement in all of the functional networks.
CHAPTER 4

HOLISTIC ATLASES OF FUNCTIONAL NETWORKS AND INTERACTIONS

(HAFNI)

4.1 Rationale and Overview

Traditionally, the subtraction approach (contrast between task and baseline epochs) has been the dominant methodology in both task-based fMRI paradigm design and fMRI data analysis [30-31]. Despite its remarkable successes and significant neuroscientific insights, nevertheless, it has considerable difficulty in reconstructing concurrent, interacting functional networks, as it was already widely recognized and pointed out in the literature that spatially overlapping networks subserving different functions are likely to be unnoticed by the blocked subtraction paradigms [32-33]. Meanwhile, from a human neuroscience perspective, it has been widely reported that a variety of brain regions and networks exhibit strong functional diversity and heterogeneity [17, 20, 21, 34]. That is, the same brain region could participate in multiple functional processes/domains simultaneously and a single functional network might recruit various neuroanatomic areas at different stages as well. Besides task-based fMRI, resting state fMRI has been another major neuroimaging technique to examine the intrinsic functional activities when the subject did not perform any task [35-37]. A variety of computational methods including independent component analysis (ICA) [38-39], normalized cut [40] and other clustering algorithms [41], have been employed to estimate resting state networks (RSNs) [35, 37, 42]. It should be pointed out, however, that virtually all current RSNs
identification methods did employ the strategy of spatially clustering fMRI signals [38-40] and assumed that RSNs are not spatially overlapping with each other. In addition, there has been increasing interest in examining the relationship/interaction between task-evoked and intrinsic resting state functional networks [43-44].

Recently, we have applied an effective sparse coding algorithm on a dataset which includes 19 subjects performing a working memory task [45-46]. The results are very interesting that we can observe some stimulus correlated or anti-correlated components after the signals decomposition. Inspired by this finding and in order to effectively address the abovementioned fundamental questions and bridge the current significant neuroscience knowledge gaps, we developed an innovative computational framework of sparse representation of whole-brain fMRI signals and apply it on the recently publicly released Human Connectome Project (HCP) high-quality fMRI data [47].

4.2 Sparse Representations of fMRI Signals

Given a collection of data vectors $X = [x_1, \ldots, x_n] \in \mathbb{R}^{m \times n}$, if there exists a linear combination of a small size of $d_i$ that can effectively represent $X$, we call $X$ admits a sparse approximation over dictionary $D$, where $D = [d_1, \ldots, d_k] \in \mathbb{R}^{m \times k}$. For each fMRI dataset having $n$ voxels with $t$ time length, we are aiming to learn a neuroscience meaningful and over-complete dictionary $D \in \mathbb{R}^{t \times m}$ ($m > t$ and $m < n$) for the sparse representation of whole brain signals $S$, where $S = [s_1, s_2, \ldots s_n] \in \mathbb{R}^{t \times n}$. The loss function is defined as:

$$
\min_{\alpha \in \mathbb{R}^m} \frac{1}{2} ||s_i - D\alpha_i||^2_2 + \lambda ||\alpha_i||_1
$$
Similar to LASSO [48], $\lambda$ is a regularization parameter which is used as a trade-off item between the sparsity level of coefficient ($\alpha_i$) and the regression residual.

As illustrated in Fig.17, sparse representation includes three steps: first, for each single subject’s brain, we extract pre-processed fMRI signals of all gray matter (GM) voxels by using individual GM mask. Then after normalization by using zero mean and standard deviation of 1, these signals are resembled into a big matrix $S \in \mathbb{R}^{t \times n}$ (Fig.17.a), where $t$ is the number of time points (fMRI volume numbers) and $n$ columns represent fMRI signals extracted from $n$ GM voxels. Finally, by applying a publicly available efficient online dictionary learning methods [49], each fMRI signal (each column vector) in $S$ can be modeled as a linear combination of a small size of learned dictionary atoms (Fig.17.b) with corresponding coefficients (Fig.17.c). For example, a specific signal

![Fig.17. Computational pipeline of sparse representation of fMRI signals.](image)
vector, $s_i$, can be represented as the product of basis dictionary $D$ and $\alpha_i$, where $\alpha_i$ is the corresponding coefficient vector in the coefficient weight matrix.

A particularly important characteristic of this framework is that the reference weight matrix [49] naturally reveals the spatial overlap patterns (Fig.17-c) among those reconstructed brain networks, which are represented by the time series of the over-complete basis dictionaries (Fig.17-b). It turned out that this novel methodology can effectively and robustly uncover multiple functional networks, including both task-evoked and Resting State Networks (RSNs), that can be well-characterized and interpreted in spatial, temporal and frequency domains.

4.3 Functional Network Component Analysis

As each functional network has its own spatial pattern and time series that serve as the basis for sparsely representing the whole-brain fMRI signals, a natural question arises: what are the neuroscience meanings of those hundreds of network components? To address this question, we propose a synthetic analysis method to identify those neuroscience meaningful components and establish their correspondences across different brains.

**Temporal-Frequency analysis:** For task fMRI (block design), the frequency of a cycle between the task and the baseline in the stimulus $F_{\text{stimulus}}$ is calculated by:

$$F_{\text{stimulus}} = \frac{1}{\text{average length of task} + \text{average length of rest}} \times \frac{1}{TR}$$  \hspace{1cm} (7)

where TR is repetition time. For the time series of the j-th network component $TS_j$, we can obtain its frequency spectrum $FS_j$ by using fast Fourier transform on its signal, and
calculate the energy concentration $E_{Fre,j}$ of the stimulus curve frequency over all frequency ranges:

$$E_{Fre,j} = \frac{FS_{Fre_{stimulus,j}}}{\sum_i FS_{i,j}}$$  \hspace{1cm} (8)

Where $FS_{Fre_{stimulus,j}}$ is the energy of the stimulus frequency in the spectrum, and $FS_{i,j}$ denotes the energy of the i-th position in the spectrum of the j-th network component. Intuitively, a large $E_{Fre,j}$ suggests that this network component is more likely to be responsive to the task stimulus and should be considered as the task related network.

Meanwhile, we can obtain Pearson correlation between the signal of each network component with the stimulus curve, which is defined as:

$$E_{corr,j} = \text{corr}(TS_j, Curve_{stimulus})$$  \hspace{1cm} (9)

Essentially, $E_{corr,j}$ measures the temporal similarity between the component’s time series and the stimulus curve, where a large value indicates better correspondence between the component and the stimulus.

**Spatial analysis:** For both task and resting state fMRI, we do have reliable pre-knowledge, such as the contrast templates derived from general linear model (GLM) and typical RSN templates [50]. The spatial similarity is defined by the overlapping rate $R$ between the spatial patterns (A) of the network component and the above mentioned templates (T):

$$R(A, T) = \frac{|A \cap T|}{|T|}$$  \hspace{1cm} (10)

Temporal-frequency and spatial analysis can effectively compensate each other since they reflect the intrinsic properties of the network components with different domains. Especially for resting state fMRI with which the temporal and frequency characteristics
have not been fully understood or quantitatively described, spatial analysis becomes a driving force for successful identification of those RSNs under both task and resting state.

4.4 Group-wise Consistent Functional Templates

We applied the sparse representation framework on HCP-Q1 dataset [47] (68 subjects). The primary goals of HCP task fMRI datasets were to identify as many core functional regions in the brain as possible that can be correlated to structural and functional connectomes. HCP dataset can be considered as one of the most systematic and comprehensive mapping of connectome-scale functional networks for a large population in the literature so far. Totally it contains seven tasks and within each task multiple contrasts (sub tasks) are covered.

In total, we have identified and confirmed 5, 3, 2, 2, 2, 3 and 6 group-wise consistent task-evoked networks, or called task component templates here, for motor (M1-M5 in Fig.18), emotion (E1-E3 in Fig.18), gambling (G1-G2 in Fig.18), language (L1-L2 in Fig.18), relational (R1-R2 in Fig.19), social (S1-S3 in Fig.19), and working memory (WM) (W1-W6 in Fig.19) networks, respectively. In particular, these 23 consistent functional task templates are reproducible and consistent across all of the HCP subjects we examined.

We also went through all of the decomposed dictionary atoms and successfully identified nine reproducible and consistent RSNs in all of the seven task fMRI datasets across all of the HCP subjects. Fig.20 shows the nine RSNs (nine rows) in these seven tasks (the first seven columns) for one exemplar subject. Meanwhile, for comparison purpose, the corresponding RSNs identified by both of the dictionary learning method
and the independent component analysis (ICA) [152] method from rsfMRI data are shown in the eighth and ninth columns in Fig.20. It is evident that all of the nine RSNs derived from either task fMRI or rsfMRI data are consistent with the template [153], and thus are called RSNs templates here. Particularly, the nine RSNs templates can be robustly reconstructed across individuals. All the above mentioned 86 templates (23 for task and 63 RSNs in seven tasks) are also called HAFNI templates and they will be used in the DICCCOL-H optimization (section 5).

A fundamental difference between the HAFNI templates and GLM-based activation maps is that the HAFNI templates are simultaneously derived from the optimally decomposed fMRI signals based on the sparse representation of whole-brain data (as illustrated in Fig.17), while the GLM-based maps were obtained from individual fMRI signals based on separate model-driven subtraction procedures.
<table>
<thead>
<tr>
<th></th>
<th>HAFNI</th>
<th>GLM</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td><img src="#" alt="Images" /></td>
<td><img src="#" alt="Images" /></td>
</tr>
<tr>
<td>M2</td>
<td><img src="#" alt="Images" /></td>
<td><img src="#" alt="Images" /></td>
</tr>
<tr>
<td>M3</td>
<td><img src="#" alt="Images" /></td>
<td><img src="#" alt="Images" /></td>
</tr>
<tr>
<td>M4</td>
<td><img src="#" alt="Images" /></td>
<td><img src="#" alt="Images" /></td>
</tr>
<tr>
<td>M5</td>
<td><img src="#" alt="Images" /></td>
<td><img src="#" alt="Images" /></td>
</tr>
<tr>
<td>E1</td>
<td><img src="#" alt="Images" /></td>
<td><img src="#" alt="Images" /></td>
</tr>
<tr>
<td>E2</td>
<td><img src="#" alt="Images" /></td>
<td><img src="#" alt="Images" /></td>
</tr>
<tr>
<td>E3</td>
<td><img src="#" alt="Images" /></td>
<td><img src="#" alt="Images" /></td>
</tr>
<tr>
<td>G1</td>
<td><img src="#" alt="Images" /></td>
<td><img src="#" alt="Images" /></td>
</tr>
<tr>
<td>G2</td>
<td><img src="#" alt="Images" /></td>
<td><img src="#" alt="Images" /></td>
</tr>
<tr>
<td>L1</td>
<td><img src="#" alt="Images" /></td>
<td><img src="#" alt="Images" /></td>
</tr>
<tr>
<td>L2</td>
<td><img src="#" alt="Images" /></td>
<td><img src="#" alt="Images" /></td>
</tr>
</tbody>
</table>

**Fig. 18.** HAFNI templates – part 1.
<table>
<thead>
<tr>
<th></th>
<th>HAFNI</th>
<th>GLM</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td><img src="image" alt="HAFNI R1" /></td>
<td><img src="image" alt="GLM R1" /></td>
</tr>
<tr>
<td>R2</td>
<td><img src="image" alt="HAFNI R2" /></td>
<td><img src="image" alt="GLM R2" /></td>
</tr>
<tr>
<td>S1</td>
<td><img src="image" alt="HAFNI S1" /></td>
<td><img src="image" alt="GLM S1" /></td>
</tr>
<tr>
<td>S2</td>
<td><img src="image" alt="HAFNI S2" /></td>
<td><img src="image" alt="GLM S2" /></td>
</tr>
<tr>
<td>S3</td>
<td><img src="image" alt="HAFNI S3" /></td>
<td><img src="image" alt="GLM S3" /></td>
</tr>
<tr>
<td>W1</td>
<td><img src="image" alt="HAFNI W1" /></td>
<td><img src="image" alt="GLM W1" /></td>
</tr>
<tr>
<td>W2</td>
<td><img src="image" alt="HAFNI W2" /></td>
<td><img src="image" alt="GLM W2" /></td>
</tr>
<tr>
<td>W3</td>
<td><img src="image" alt="HAFNI W3" /></td>
<td><img src="image" alt="GLM W3" /></td>
</tr>
<tr>
<td>W4</td>
<td><img src="image" alt="HAFNI W4" /></td>
<td><img src="image" alt="GLM W4" /></td>
</tr>
<tr>
<td>W5</td>
<td><img src="image" alt="HAFNI W5" /></td>
<td><img src="image" alt="GLM W5" /></td>
</tr>
<tr>
<td>W6</td>
<td><img src="image" alt="HAFNI W6" /></td>
<td><img src="image" alt="GLM W6" /></td>
</tr>
</tbody>
</table>

Fig. 19. HAFNI templates – part II.
<table>
<thead>
<tr>
<th>RSN</th>
<th>Motor</th>
<th>Emotion</th>
<th>Gambling</th>
<th>Language/Relational</th>
<th>Social</th>
<th>Working memory</th>
<th>Resting state</th>
<th>Template</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSN1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSN2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSN3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSN4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSN5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSN6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSN7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSN8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSN9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig.20. HAFNI templates – part III.
CHAPTER 5
CONSTRUCTION OF HOLISTIC ATLASES BASED ON DICCCOL
(DICCCOL-H)

5.1 Rationale and Overview

Recently, there have been extraordinary interests and efforts in measuring large-scale, whole-brain connectivity, known as “connectomics” [4-5,7-8,51-54], and it is considered as one of the highest priority research areas in NIH’s interim report which is in response to the President Obama’s “BRAIN Initiative” project [55]. Essentially, when mapping the brain connectivity, Regions of Interests (ROIs) provide the foundation or structural substrates for measuring connectivity within individual brains and for comparing results across populations. Thus, identification of reliable, reproducible and accurate ROIs with correspondences across individual brain is critically important for the success of brain connectivity mapping.

Our previous work of DICCCOL [5] (section 2) is the first successful attempt in the field to construct group-wise ROIs by identifying the most consistent white matter connectivity patterns across different individuals. Many studies [4,7,51-52,56-57,58,43,59,60-63,64-69,70-72,73-77,78-80] already demonstrated that it is an effective and robust ROI modeling framework and has significant improvement compared to the previous registration method[5]. Despite DICCCOL system is an important advancement in human brain mapping, however, it did not consider the functional homogeneity and heterogeneity behind those brain structural consistencies. To tackle this fundamental
issue, we propose to construct a novel Holistic Atlases of Brain Structure and Function based on DICCCOL system and the brain’s functional sparse representation, and we name it as DICCCOL-H.

5.2 DICCCOL Prediction and Functional Labeling

We applied DICCCOL prediction procedure on the HCP data to acquire the initial locations of the landmarks which provided a reliable and neuroscience grounded foundation for further joint optimization under both structural and functional regulations. After we have all 358 DICCCOL landmarks in the new brain, we can functionally label them using the functional network templates derived in section 4.4. We construct a functional regulation profile for each predicted DICCCOL landmark. This functional regulation profile is a binary vector with \( L \)-dimension and \( L \) is the number of network templates (86 in this work). For example, if one DICCCOL landmark is located in the region of a specific template, the corresponding item in the regulation vector will be 1, otherwise will be 0. Through this way, the functional regulation profile can effectively encode the functional expression of every DICCCOL landmark in individual space.

5.3 Optimization on Structure and Function Simultaneously

Because the DICCCOL landmarks already possess the most consistent structural connectivity across different populations, our primary objective is to maximize the functional homogeneity and minimize the potential affection to the established structural consistency simultaneously. This process includes two steps: 1) Construction of the functional regulation model. Since for each subject and each DICCCOL landmark, we
already achieved its functional regulation profile through functional labeling. For each DICCCOL, we assemble all the functional profiles within the group (68 subjects) and do a simple but efficient voting for each template. Thus we arrive with a 358*L matrix. Here L is the number of functional network templates. Each element represents the subject number that in those subjects the current DICCCOL is consistently located in a specific functional template. The larger number means more individuals have agreement that the current DICCCOL should belong to this functional network.

In this work, we adopt a relatively strict criterion that if more than half of the group individuals commit this template, the corresponding matrix element will be considered as value 1 in the functional regulation model. 2) Joint optimization. In brief, using the predicted location as the initial searching point, we move the DICCCOL landmark within a small neighborhood and examine if there exists an appropriate location at which the

Fig.21. Illustration of joint optimization.
functional profiles become more consistent to the functional regulation model, and at the same time the structural connectivity has not significant change. The neighborhood is defined as a circle with radius of 3mm because the average registration error is considered as 6mm [5]. The basic idea is illustrated in Fig.21. The green bubble represents the initial location. According to its current functional regulation profile, the green one will move to the neighboring locations that makes it more consistent to the functional regulation model. For example, if its current functional profile is <0, 0> and the regulation model is <1, 0>, it will intend to move along the red arrow. Otherwise, it will move to the other two directions or stay at the initial location given the regulation mode as <0, 0>. It should be noted that this optimization process is performed under the structural constraints, which can effectively preserve the already established structural consistency through DICCCOL prediction.

The overall optimize function is summarized as:

$$\begin{align*}
E &= \phi (\mathbf{DT}_{i,i'}) \cdot |\mathbf{FR}_{i'} - \mathbf{FR}_M| \\
\phi (\mathbf{DT}_{i,i'}) &= \begin{cases} 1, & \text{if } |\mathbf{DT}_i - \mathbf{DT}_{i'}| \leq \sigma \\ \frac{|\mathbf{DT}_i - \mathbf{DT}_{i'}|}{\sigma}, & \text{if } |\mathbf{DT}_i - \mathbf{DT}_{i'}| > \sigma \end{cases}
\end{align*}$$

(11)

Here i and i' represent the initial location and the candidate location need to be examined, respectively. FR is the functional regulation vector. DT_{i,i'} is the trace-map distance between the candidate location and the initial location. \( \phi (\mathbf{DT}_{i,i'}) \) is defined as:

$$\begin{align*}
\phi (\mathbf{DT}_{i,i'}) &= \begin{cases} 1, & \text{if } |\mathbf{DT}_i - \mathbf{DT}_{i'}| \leq \sigma \\ \frac{|\mathbf{DT}_i - \mathbf{DT}_{i'}|}{\sigma}, & \text{if } |\mathbf{DT}_i - \mathbf{DT}_{i'}| > \sigma \end{cases}
\end{align*}$$

(12)

Here \( \sigma \) is the standard deviation of the group-wise trace-map distance of the predicted DICCCOL landmarks. Intuitively, if the structural connectivity does not change much, the functional regulation item will be the driving force for the optimization process. If not, the functional regulation item will be penalized that the DICCCOL landmark will be
inclined to stay at the initial location to maintain its already established structural consistency. This process will be applied to each DICCCOL landmark separately and eventually we can achieve the optimized landmarks, which reflect both the structural and functional consistency across different individuals.

**Identification of common structural and functional landmarks:** The above mentioned optimization process has been applied to the HCP Q1 dataset. Totally 84 DICCCOL landmarks were successfully optimized under the functional regulation model which consists of 86 consistent functional templates. It is noted that this common functional regulation model, which is represented as an 84*86 binary matrix, is derived from the group-wise voting. Hence it is possible that only a part of the model contributes to the individual optimization process. For example, for a specific subject, if some of these 84 DICCCOL landmarks are already consistent with the regulation model, then these DICCCOLs will be ignored for efficiency consideration.

The optimization result is shown in Fig.22. One subject is randomly selected as an example and the optimized landmarks are displayed on an inflated cortical surface with colored bubbles. The green and red ones represent the landmarks before and after the optimization (Fig.22 (a)). Figs.22 (b-d) are three enlarged examples and the yellow arrows illustrated the direction of the landmarks’ movements during the optimization. In addition, the white matter bundles connecting to the original and optimized landmarks are also displayed. Fig.22 (e) shows the average changes of functional regulation profiles before (left) and after (right) the optimization. The rows and columns represent optimized landmarks and different functional templates. The larger value indicates more consistency. We can clearly see that the overall structural connectivity patterns do not
Fig. 22. Results of joint optimization. (a) One subject is randomly selected as an example. Green and red bubbles represent original and optimized landmarks, respectively. (b-d) show three enlarged examples and the fiber bundles before and after the optimization. Yellow arrows indicate the direction of the movement during the optimization. (e) and (f) show the average changes of functional regulation profiles and trace-map distance before and after the optimization. (g) displays the percentage regarding the changes of the trace-map distance and the improvement of functional change much (Fig. 22b-d), however, the corresponding functional consistency is significantly improved (Fig. 22e). The quantitatively analysis of structural changes and functional consistency improvement are shown in Fig. 22 (f-g). Obviously, our
optimization goal is successfully achieved that the overall functional consistency is significantly improved (30% on average of improvement) while the consistent structural connectivity is effectively preserved (2% on average of changes).

**Landmarks possessing both structural and functional consistencies:** In this section, we focus on the latent information delivered by the identified functional regulation model. As mentioned before, the functional regulation model involves 84
DICCCOL landmarks and 86 consistent functional templates. Fig. 23 (a) demonstrates the spatial locations of these 86 landmarks. Colored bubbles represent all the 358 landmarks and the green ones are those landmarks that can be optimized through our functional regulations. It is obvious that most of them are assembled in the occipital, parietal and temporal lobes and few of them are located in the frontal lobe. One explanation is that, in frontal lobe the overlapping pattern of different functional templates is much more complicated than other brain regions, which makes it difficult to satisfy the voting procedure, since frontal lobe is considered to contribute more towards the high-level brain functions. Another possibility is because of the functional templates which are used to generate the regulation model. Though the HCP Q1 data has seven sets of task data, it is possible that some brain regions were not covered by any task networks or resting state networks. This situation led to the result of insufficient functional regulations for some brain regions. Fig.23 (b) shows the effective regulation number of every DICCCOL landmarks. For each optimized landmark there are 4 regulation profiles on average. For example, the #145 landmark (highlighted with red arrow) (Fig. 23(c)) has two regulation profiles, which come from language network and motor component of resting state network [50]. One subject is selected to demonstrate the optimization process of this landmark as showed in Fig. 23(c). Red and blue areas denote two functional networks and eventually the #145 landmark in this subject was moved from the region which was covered by one functional network to the nearby overlap region according to functional regulation model.
Bibliography


Connectomics Signatures of Prenatal Cocaine Exposure Affected Adolescent Brains,
*Joint first authors, **Joint corresponding authors, Mar 28. 34(10):2494-510., Human
Brain Mapping, 2012

T. Liu, Discovering dense and consistent landmarks in the brain, IPMI, 6801:97-110,
2011

T. Liu, Optimization of functional brain ROIs via maximization of consistency of
structural connectivity profiles, International Symposium on Biomedical Imaging
(ISBI oral), 2150-2154, 2011

[10] D. Zhu, T. Zhang, X. Jiang, X. Hu, N. Yang, J. Lv, J. Han, L. Guo, T. Liu, Fusing
DTI and FMRI Data: A Survey of Methods and Applications, invited paper,

activity distinguishes Alzheimer’s disease from healthy aging: evidence from

homogeneity, functional connectivity and imaging markers of Alzheimer’s disease:

changes in the early stages of Alzheimer’s disease: A combined structural and


[24] T. Liu, A few thoughts on Brain ROIs, Brain Imaging and Behavior, in press, 2011
    Nature Reviews Neuroscience, 11(2):139-45, 2010
[35] M. D. Fox and M.E. Raichle, Spontaneous fluctuations in brain activity observed
with functional magnetic resonance imaging. Nat Rev Neurosci., 8:700-711, 2007


Functional Connectome Patterns, Medical Image Analysis, 17(8):1106-22. 2013


[61] X. Hu; D. Zhu; P. Lv; K. Li; J. Han; L. Wang; D. Shen; L. Guo, T. Liu, Fine-Granularity Functional Interaction Signatures for Characterization of Brain Conditions, in press, Neuroinformatics, 2013


[63] Y. Yuan*; X. Jiang*; D. Zhu; H. Chen; K. Li; P. Lv; X. Yu; X. Li; S. Zhang; T. Zhang; X. Hu; J. Han; L. Guo, T. Liu, Meta-analysis of Functional Roles of DICCCOLs, *Joint first authors, Neuroinformatics, 11(1):47-63, 2012


[76] B. Ge, L. Guo, T. Zhang, D. Zhu, X. Hu, J. Han, T. Liu, Construction of Multi-scale Common Brain Network Via DICCCOL, Information Processing in Medical Imaging (IPMI), 2013


[78] B. Ge, L. Guo, T. Zhang, D. Zhu, X. Hu, J. Han, T. Liu, Construction of Multi-scale Brain Networks via DICCCOL Landmarks, accepted, ISBI, 2013


[80] B. Ge, L. Guo, D. Zhu, K. Li, X. Hu, J. Han, T. Liu, Group-wise Consistent Fiber Clustering Based on Multimodal Connectional and Functional Profiles, MICCAI,


[89] L. Snook, C. Plewes and C. Beaulieu, Voxel based versus region of interest analysis
in diffusion tensor imaging of neurodevelopment, NeuroImage 34, 243–252, 2007


423, 2004

White matter damage in Alzheimer’s disease and mild cognitive impairment:
assessment with diffusion tensor MR imaging and parallel imaging techniques,
Radiology, 243: 482–492, 2007

Du, K.H. Kramer, K. Yaffe, H. Chui, W.J. Jagust, B.L. Miller and M.W. Weiner,
Diffusion tensor imaging of cingulum fibers in mild cognitive impairment and

functional impairments in aging and early Alzheimer’s disease, Alzheimer’s Dement,
4: 381–389, 2008

[99] A. Fellgiebel, M.J. M’uller, P. Wille, P.R. Dellani, A. Scheurich, L.G. Schmidt and P.
Stoeter, Color-coded diffusion tensor imaging of posterior cingulate fiber tracts in
mild cognitive impairment, Neurobiol Aging, 26: 1193–1198, 2005

[100] D.A. Medina, L. deToledo-Morrell, F. Urresta, J.D. Gabrieli, M. Moseley, D.
Fleischman, D.A. Bennett, S. Leurgans, D.A. Turner and G. T. Stebbins, White
matter changes in mild cognitive impairment and AD: a diffusion tensor imaging
study, Neurobiol Aging, 27: 663–672, 2006

[101] Y. Nakata, N. Sato, O. Abe, S. Shikakura, K. Arima, N. Furuta, M. Uno, S. Hirai,
Y. Masutani, K. Ohtomo and S. Aoki, Diffusion abnormality in posterior cingulate
fiber tracts in Alzheimer’s disease: tract-specific analysis, Radiation Medicine, 26:
466–473, 2008


http://adni.loni.ucla.edu/

T. Zhang, L. Guo, K. Li, C. Jing, Y. Yin, D. Zhu, G. Cui, L. Li, T. Liu, Predicting functional cortical ROIs based on fiber shape models, Cerebral Cortex, doi: 10.1093/cercor/bhr152, 2011


National Academy of Sciences, 106(31): 13040-13045, 2009