Brain connectivity analysis with ICA

Yang, Kanyan

2005

Yang, K. (2005). Brain connectivity analysis with ICA. Master's thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/2546>

<https://doi.org/10.32657/10356/2546>

Nanyang Technological University

Downloaded on 26 Apr 2023 21:00:41 SGT

8357843

Brain Connectivity Analysis with ICA

Yang Kanyan

School of Computer Engineering

A thesis submitted to the Nanyang Technological University
in fulfilment of the requirement for the degree of
Master of Engineering

2005

RC
386.6
-M34
422
2005

Acknowledgements

I would like to express my sincere gratitude to Associate Professor Jagath C. Rajapakse, my supervisor, for his many suggestions and invaluable guidance during my research. I have learned lots of things from him ranging from academic knowledge to research experiences.

I should really thank Max-Planck-Institute of Cognitive Neuroscience, Leipzig, Germany and Department of Diagnostic Radiology, University of Freiburg, Germany for providing the fMRI datasets in my experiments.

I am also grateful to Professor Ed Bullmore, Department of Psychiatry, University of Cambridge, UK for sharing his knowledge and providing the source code of implementations in his paper.

All the members of Bioinformatics Research Center should be thanked for their friendship and creating a very pleasant research environment, especially my good partner Bobby.

Besides, I would enthusiastically thank my family and my beloved Peggy, for their constant support for my work and warm care for my life.

Finally, great thanks to School of Computer Engineering, Nanyang Technological University for offering me the award of Research Scholarship and first-rate research facilities.

Table of Contents

Acknowledgements	
Table of Contents	ii
Summary	iv
Author's Publications	vi
List of Figures	vii
Glossary	x
1 Introduction	1
1.1 Background Knowledge	3
1.1.1 Basics About the Brain	3
1.1.2 Functional MRI	5
1.1.3 Correlation Analysis	6
1.1.4 Statistical Parametric Mapping	6
1.2 Motivation	10
1.3 Major Achievements	11
1.4 Organization of the Thesis	12
2 Restoration of fMRI Using ICA	14
2.1 Introduction	14
2.2 ICA Algorithms	17
2.2.1 Fast ICA Algorithm	19
2.2.2 Maximum Likelihood Algorithm	23
2.3 ICA-Based Model of fMRI Restoration	25
2.3.1 The Conceptual Restoration Model	27

2.3.2	Basic Measures of Identifying Components	29
2.3.3	Identification and Elimination of Specific Components in fMRI	32
2.4	Reduction of Noise and Its Implementation	37
2.4.1	Visual Task	37
2.4.2	Memory Retrieval Task	40
2.4.3	Synthetic Data	41
2.5	Reduction of Physiological Fluctuations and Its Implementation	43
2.5.1	Method of Identification	43
2.5.2	Experiment and Results	46
2.6	Conclusion	49
3	Higher-Order Functional Brain Connectivity	52
3.1	Introduction	52
3.2	Connectivity in the Brain	56
3.3	Second-Order Functional Connectivity and PCA	59
3.4	ICA and Higher-Order Functional Connectivity	63
3.5	Experiments and Results	68
3.5.1	Eigenimage Analysis	68
3.5.2	Connectivity Analysis with ICA	70
3.6	Conclusion and Discussion	73
4	Exploratory Approach to Modeling Neuronal Interactions with SEM	76
4.1	Introduction	76
4.2	Structural Equation Modeling	78
4.3	Confirmatory and Exploratory Modeling Approaches	83
4.4	Methods	85
4.4.1	Activation Detection with ICA	85
4.4.2	Automated Search for the Best Path Model	86
4.5	Experiments and Results	90
4.6	Conclusion and Discussion	93
5	Conclusion and Discussion	96
5.1	Summary of Contributions	96
5.2	Directions of Future Work	100
5.2.1	Enrichment of the Restoration Model	100
5.2.2	Improvements in Structural Equation Modeling	101
5.2.3	Detection of Different Neural Systems	102
	Bibliography	103

Summary

Functional Magnetic Resonance Imaging (fMRI) is increasingly utilized to explore brain networks and neuronal interactions underlying brain functions. Although the concept of functional connectivity has been introduced to analyze brain connections for many years, this is a measure relying on the pattern of temporal correlations that exist between distinct neuronal units. In this research, we are going to extend the definition of brain connectivity into a higher-order statistical sense. Apart from this, two more contributions are made including a novel restoration model and a fully exploratory approach to investigating effective connectivity.

Independent Component Analysis (ICA) has been demonstrated as an effective method to decompose fMR images into independent processes and extract useful neural signals. With this basic technique, a restoration model for fMRI is proposed and constructed and much investigation is placed on identification of a large number of artifacts and noises in fMRI signals. Experiments of reducing noises and physiological fluctuations in fMRI demonstrate that our model is possible for restoration of fMR images, confirmed by the improvement in efficacy of activation-detecting techniques. Cleaning our fMRI datasets is an important part to enhance the analysis of brain functions.

Besides, ICA makes it possible to look into functional connectivity in a higher-order statistical sense and leads us to develop the definition of higher-order functional connectivity. Utilizing ICA, we are able to detect numerous highly connected brain regions, without being affected by any interfering factors in the fMR images. Since brain connectivity is usually the result of neuronal synchronization along anatomical

pathways, higher-order connectivity can provide a closer insight into brain interactions.

Last, we propose an exploratory approach to effective connectivity analysis, in which the regions involved are selected from the component maps of SICA decomposition and an automated search for best fitting model is employed to seek for the best path model. From the experiment on visual data, we conclude it as an useful exploratory approach to analyzing effective connectivity for a specific brain operation.

A main characteristic in this thesis is the novel application of ICA technique in various analysis of neuroimaging study, from restoration of images to modeling effective connectivity, in comparison with classical and traditional methods. Especially, the most significant contribution should be highlighted on the new definition of higher-order connectivity, extending from previous second-order connectivity, where ICA plays a tremendous role in finding the connectivity patterns.

Author's Publications

Journal Paper

1. Kanyan Yang and Jagath C. Rajapakse, "ICA Gives Higher-Order Functional Connectivity of Brain", *Neural Information Processing - Letters and Reviews*, Vol.2, No.2, pp.27-32, 2004

Conference Papers

1. Kanyan Yang and Jagath C. Rajapakse, "Denoising of functional MRI using ICA", *IEEE 3rd International Symposium on Image and Signal Processing and Analysis (ISPA 2003)*, Italy, September, 2003
2. Kanyan Yang and Jagath C. Rajapakse, "An Exploratory Approach to Modeling Effective Connectivity", *2004 IEEE International Symposium on Biomedical Imaging (ISBI2004)*, 2004

List of Figures

1.1	The SPM graphical user interface	9
2.1	Conceptual model of restoration of fMRI data	27
2.2	Kurtosis of independent components on visual task data	38
2.3	Maps of activation on visual task data, threshold of correlation for both maps:0.5 (a)before denoising (b)after denoising; colored blobs show the activation	39
2.4	Maps of activation on visual task data: (a)before denoising, threshold of correlation:0.4, (b)after denoising, threshold of correlation:0.5 . . .	39
2.5	Activations obtained from a representative subject in the memory retrieval task experiment using SPM approach (a)before denoising (b)after denoising; colored blobs show the activation	41
2.6	Kurtosis of independent components on synthetic fMRI data	41
2.7	Maps of activation on synthetic data, threshold of correlation for both maps:0.5 (a)before denoising (b)after denoising	42
2.8	The ROC curves for the synthetic data	43
2.9	Histogram of the dominant frequency components: (a) within the respiratory frequency range (b) within the cardiac frequency range . . .	46
2.10	Estimated waveforms: (a) respiratory signal (b) cardiac signal	47
2.11	The power spectrums of estimated waveforms: (a) respiratory signal (b) cardiac signal	47
2.12	Fourier fitting of (a) respiratory motion (b) cardiac motion	48

2.13 The mapping of activation and their comparison: (a) before (b) after physiological artifact reduction (c) the ROC curves 48

3.1 (a) The first eigenimage with eigenvalue=573.5; (b) the second eigenimage with eigenvalue=19.0. The eigenimages were transformed into a z-map and thresholded at z-value=2.8 with a minimum blob size of 2. 69

3.2 Voxel intensity distribution of a representative component map 71

3.3 Component maps of neural sources obtained from ICA: (a) map with $\kappa = 37.8$ (b) map with $\kappa = 24.5$ (c) map with $\kappa = 23.7$. The maps were transformed into a z-map and thresholded at p-value ≤ 0.02 with a minimum blob size of 2. 71

4.1 Decomposition of effects 79

4.2 Schematic representation of methods involved in structural equation modeling of a neural system: (a) Path diagram of a simple network with four regions; (b) The information about the correlations of activities; (c) Path equations showing how the correlations between regions can be decomposed to solve for the path coefficients; (d) Structural equations showing the variance in activity of each region as a function of the weighted variance of other brain regions and a residual influence. . . 82

4.3 Conceptual model of exploratory modeling of effective connectivity from fMRI 85

4.4 Activation maps whose time-series show a significant correlation with the stimulation time blocks. Colored voxels demonstrate great activations. 90

4.5 Measures of model goodness for a series of automatically specified path models: (a) chi square (b) p-value (c) Alaike's information criterion against the increasing number of nonzero path coefficient. 91

4.6 Path model for the fMRI data from visual experiment: (a) best fitted model having $q = 2$ nonzero path coefficients (b) a confirmatory model with fully connected edges 92

Glossary

ABS Absolute Value

ANOVA Analysis of Variance

BSS Blind Source Separation

CBF Cerebral Blood Flow

CSF cerebrospinal fluid

EEG electroencephalography

EPI Echo-Planar Imaging

fMRI functional Magnetic Resonance Imaging

GRF Gaussian Random Field

HOS Higher Order Statistics

HRF Hemodynamic Response Function

IC Independent Component

ICA Independent Component Analysis

i.i.d. independent identical distribution

MI Mutual Information

MRF Markov Random Field

MSE Mean Square Error

PC Principal Component

PCA Principal Component Analysis

pdf probability density function

PET Positron Emission Tomography

SEM Structural Equation Modeling

SICA Spatial ICA

SNR Signal-to-Noise Ratio

SPM Statistical Parametric Mapping

TICA Temporal ICA

Chapter 1

Introduction

Functional Magnetic Resonance Imaging (fMRI) has enabled scientists to look for the first time into the human brain, in vivo, to literally "watch it while it works". This has revealed exciting insights into the spatial and temporal changes underlying a broad range of brain functions, such as how we see, feel, move, understand each other and lay down memories [1].

Throughout the early history of neurology and neuroscience, most theoretical accounts of brain function have emphasized aspects of localization or distributed properties [2]. Instead, modern connectivity views focus extensively on the structure and dynamics of large-scale neuronal networks, leading to the anatomical, functional and effective connectivity of the human brain. All the previous work of finding functional connectivity used the measure of second-order correlation coefficient, including PCA analysis with PET data [3], later pairwise temporal correlation [4] and eigenimage analysis on fMRI data[5]. As they were conducted on mixed neuroimaging datasets, the pattern of functional connectivity may not be accurately obtained, which was affected by various kinds of other interferences in fMR images.

Another novel trend in the analysis of interactions among brain regions has been

focused on the covariance-based method: structural equation modeling (SEM) [6], in which connections between brain areas are based on known neuroanatomy and the interregional covariances of activity are used to calculate path coefficients representing the magnitude of the influence of each directional path. However, this approach to investigating the effective connectivity is only able to affirm or refute the neural networks in a previously known anatomical model, where the neural elements involved are already given. As a confirmatory approach, it does not allow analysis of the behavior of the network due to an abnormal event in a brain structure.

In the past decade, many techniques have been exploited in the detection of significant activation regions. Correlation techniques [7] for detecting task-related activation in fMRI are based on the assumption that voxels indexing brain regions participating in the cognitive processing of the given experimental and control task should show different fMRI signals levels during the performance of these tasks. Analysis of Variance/Covariance (ANOVA/ANCOVA) approaches including Statistical Parametric Mapping (SPM) [8] test the signal at each voxel using univariate measures (e.g. t-tests or F-tests) under the null hypothesis that the values are distributed under a known probability distribution (typically Gaussian). Recently, Independent Component Analysis (ICA) has been increasingly explored as a data-driven approach for the analysis of fMRI data. Some experiments and results have indicated that ICA can be used to reliably separate fMRI datasets into meaningful constituent components, including consistently and transiently task-related physiological changes, nontask-related physiological phenomena, and machine or movement artifacts [9], which suggests a new powerful approach to the restoration of fMRI datasets.

In this chapter of introduction, a background knowledge review will be given next,

from the biochemical anatomy to a statistical perspective. Then the motivation of my research will be described, following which is an outline of my achieved work.

1.1 Background Knowledge

This section gives some basics related to the fMRI techniques, which serves as the threshold leading me into the area of what I started to research.

1.1.1 Basics About the Brain

The brain forms the essence of what defines us as human beings. Our brain is composed of three types of tissue: gray matter, white matter, and cerebrospinal fluid (CSF) [10].

Gray matter is called "gray" because it looks relatively dark in anatomical brain specimens (postmortem tissue). Its dark color is produced by densely packed cell bodies of nerve cells (neurons), which perform the basic "command functions" within our brains. Nerve cell bodies are highly concentrated on the surface of the brain, giving it the appearance of being covered with bark, and so this outer surface is called the cerebral cortex. Small concentrated islands of nerve cells also occur deep inside the brain, and these are called subcortical regions.

Neurons act as the integrating units of the nervous systems, and although they share many characteristics of other cells in the body, they also have special characteristics that help them perform their functions. Once they formed, they do not regenerate, and unless they suffer lethal damage, most live as long as the person in which they are found. Neurons vary enormously in size and shape, the differences bearing evidence of each neuron's particular adaption to its specialized function. Neurons are

aggregated into communities, or nuclei, each of which makes a special contribution to behavior. Neurons change their behavior with experience: they learn, remember and forget. During development, far more neurons are created than will ultimately survive, and so cell death is a pronounced developmental stage in sculpting the adult brain.

White matter is called "white" because it looks much lighter (because of blood) than grey matter in both postmortem tissue and most MR images. Neurons are connected to one another by long "wires" that project out of them, called axons, which permit individual neurons to send messages back and forth to one another across relatively long distances. The axon originates in the cell body at a transition point called the axon hillock. Each cell has only one axon, which can vary in length from a few microns to more than a meter. Most axons have branches called collaterals. At the end of the axon and its collaterals are fine terminations called teleodendria. The teleodendria end in little knobs, called terminals, which make junctions with other cells. The axons can grow new teleodendria and new terminals and so form the other half of the equation for mediating improved skills and learning. Because each neuron has only one axon, the cell can send only one message, but the collaterals and teleodendria allow that message to be sent to a number of destinations.

The brain is bathed inside and out by CSF, a fluid that contains nutrients and byproducts of brain activity. The regions inside the brain that contain CSF are called ventricles.

The surface of the human brain looks very wrinkled, reflecting the phylogenetic maturity or advanced age of the human mind. The ridges and furrows covering the brain are called gyri and sulci.

1.1.2 Functional MRI

Magnetic Resonance Imaging (MRI) technology is an unique approach to mapping the human brain in 3D, non-invasively. MRI of human brain depends on a magnetic property of hydrogen nucleus, known as the spin angular momentum. Because the water contents (and also the hydrogen contents) differ in different brain matter, MRI produces different image intensities for gray and white matter, and CSF. This results in a structural MR brain image that provides clear views of various anatomical structures at a high spatial resolution.

Functional MRI (fMRI) delivers images sensitive to what we think and perform. From our birth, different regions of the brain develop and organize to perform specific functions. When we perform a particular task, networks of neurons in the corresponding regions of the brain act to process signals received from sensory organs and other parts of the brain. The collective action of activated neurons of the brain decides our subsequent physical and mental acts.

Metabolic activities concomitant with neuronal activation occur by consuming oxygen and glucose in order to supply energy for electrochemical activities of the neurons. Hemoglobin molecules carry most of the oxygen in blood. Therefore at the outset of neuronal activation, there is an increase of de-oxyhemoglobin, which is followed by a rise in the concentration of oxyhemoglobin in order to supply the demand for oxygen. Because de-oxyhemoglobin is paramagnetic and oxyhemoglobin is diamagnetic, the change of magnetic properties of the blood around an activated area alters MR image intensities in the vicinity of neuronal activation.

1.1.3 Correlation Analysis

Correlation analysis technique exploits a priori knowledge of the expected time course of task-related changes in the signal to determine their intensity and spatial extent [7]. Many current fMRI experiments use a block design in which the subject is instructed to perform experimental (E) and control (C) tasks in an alternating sequence of 20-40s blocks (e.g., CECECEC...). A reference function is constructed by convolving a square wave matching the time course of the experimental/control task blocks with a fixed model of the hemodynamic response function [11] (an estimate of the fMRI signal changes evoked by a brief burst of neural activity). This reference function is then correlated with the time series recorded from each voxel as follows:

$$cor(k) = \frac{\mathbf{m}_k^T \mathbf{y}}{\sqrt{\mathbf{m}_k^T \mathbf{m}_k} \sqrt{\mathbf{y}^T \mathbf{y}}} \quad (1.1.1)$$

where $cor(k)$ is the correlation coefficient for the k^{th} voxel, $\mathbf{m}_k = (m_{k1} \ m_{k2} \ \dots \ m_{kn})^T$, where m_{ki} is the intensity value of the k^{th} voxel in the i^{th} scan subtracted by $\bar{m}_k$, $\mathbf{y} = (y_1 \ y_2 \ \dots \ y_n)$ where y_i is the value of reference function at the time point corresponding to the i^{th} scan subtracted by $\bar{y}$. The symbols $\bar{m}_k$ and $\bar{y}$ denote averages that run over all the n scans of the k^{th} voxel and all the time points of the block design, respectively. Those voxels, whose signals are positively correlated with the reference function above a preselected threshold, are designated "areas of activation".

1.1.4 Statistical Parametric Mapping

Statistical Parametric mapping (SPM) is a spatially extended statistical process that is used to test hypotheses about regionally specific effects in neuroimaging data [12]. The most established sorts of statistical parametric maps are based on linear models

and t-tests.

The general linear model for a response variable x_{ij} (such as rCBF) at voxel j is:

$$x_{ij} = \beta_{1j}g_{i1} + \beta_{2j}g_{i2} + \dots + \beta_{Kj}g_{iK} + e_{ij} \quad (1.1.2)$$

where $i = 1, \dots, I$ indexes the observation (e.g. scan). The general linear model assumes the errors (e_{ij}) are independent and identically distributed normally $[N(0, \sigma_j^2)]$. Here the β_{kj} are K unknown parameters for each voxel j . The coefficient g_{ik} are explanatory variables relating to the conditions under which the observation i is made.

We can rephrase our model in terms of matrices:

$$\mathbf{x}_j = \mathbf{G}\boldsymbol{\beta}_j + \mathbf{e}_j \quad (1.1.3)$$

Here $\mathbf{x}_j = (x_{1j}, x_{2j}, \dots, x_{Ij})^T$ is the rCBF data vector for voxel j , $\mathbf{G}$ is the design matrix, comprised of the coefficient g_{ik} , with each row for every scan and each column for every effect (condition) in the model. $\boldsymbol{\beta}_j = (\beta_{1j}, \beta_{2j}, \dots, \beta_{Kj})^T$ is the parameter vector for voxel j and $\mathbf{e}_j$ is a vector of normally distributed error terms.

The least square estimate of $\boldsymbol{\beta}_j$, say $\hat{\boldsymbol{\beta}}_j$, is:

$$\hat{\boldsymbol{\beta}}_j = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{X} \quad (1.1.4)$$

The remaining errors can be worked out by:

$$\mathbf{e}_j = \mathbf{x}_j - \mathbf{G}\hat{\boldsymbol{\beta}}_j \quad (1.1.5)$$

We can assess whether there will be a linear relation linking effect k with our fMRI intensity values. We test this against the null hypothesis: there is no relationship between effect k and the voxel data. On the null hypothesis, β_{kj} will not be

significantly different from zero. We can test this by making a t statistic, where the t statistics:

$$t = \frac{\hat{\beta}_{kj}}{\varepsilon_{kj}} \quad (1.1.6)$$

The standard error ε_{kj} can be worked out from the original analysis of variance, using the remaining error e_j above. SPM will do the calculation above for every voxel in the brain, and thus, for every voxel, there will be a separate $\hat{\beta}_{kj}$. This t-statistic will be large and positive if the linearity is significantly greater than 0, and large and negative if the linearity is significantly less than 0.

After calculating the t-statistic, SPM technique converts the t-statistics to Z-scores. Z-scores are a way that SPM uses to display and analyze the p-values from the t-statistics, which are the numbers from the unit normal distribution that would give the same p value as the t statistic. Then, SPM will show a picture of the Z-statistics, thresholded at a given p-value.

SPM2 is the latest software of using construction and assessment of spatially extended statistical process to test hypotheses about neuroimaging data from PET (Positron Emission Tomography) and fMRI. Fig.(1.1) gives the SPM graphical user interface which is divided into four panels: the upper for spatial pre-processing of data, the two middle for model specification/parameter estimation and statistical inference, and the lower for general utilities.

The spatial pre-processing functions are:

- Realign: realignment of functional time-series.
- Coregister: Coregistration of same modality and multi-modality image volumes.
- Slice timing: Adjustment for timing differences in multi-slice image acquisition

Figure 1.1: The SPM graphical user interface

(fMRI).

- Normalize: Spatial normalization of image volumes to a template.
- Smooth: Convolution of image volume with a Gaussian kernel.
- Segment: Segmentation of MRI volumes into CSF, grey and white matter.

Model specification and parameter estimation:

- Basic models: basic statistical models for independent data.
- fMRI: set-up of modality-specific models.
- Review design: review a previously specified model.

- Estimate: estimation of a previously specified model and configuration.

Statistical inference:

- Results: analysis and display of regional effects.

The lower panel includes miscellaneous and general utility functions.

1.2 Motivation

Conventionally, functional brain connectivity is measured by temporal correlations between neuronal elements, which is regarded as second-order statistics. Motivated by the complicated network and neuronal mechanism underpinning brain functions, we aim to discover more intimate and strong connection beyond second-order sense. In this research, we are going to extend the definition of brain connectivity into a higher-order statistical sense. The research primarily circles around Independent Component Analysis (ICA). The choice of ICA technique is due to the reason that ICA is able to extract spatially independent neural systems, related to a given task, in a data-driven manner and without any interferences from artifacts and noises in the neuroimages. Approaches will be proposed and experiments performed to demonstrate the utility of ICA in analyzing this brain connectivity.

By this means, a more complete restoration for neuroimages is also possible of implementation. It has been demonstrated that ICA is capable of extracting different source signals underlying fMRI data: nontask-related signal components, movements and other artifacts as well as consistently or transiently task-related fMRI activations. Thus, based on ICA separation, a restoration model for fMRI data could be derived.

Besides, Structural Equation Modeling (SEM) has proved to be a powerful way to combine functional neuroimaging data with anatomical circuitry to determine effective connectivity underlying a particular task. However, previous work usually aimed to affirm or refute a known model. Recently, an automated search for the best fitting model that can be found to account for the data was proposed by Bullmore [13], but the logical basis of this approach requires that the system under investigation is already well understood by the analyst. Therefore, a more exploratory approach to path analysis of fMRI data may circumvent some of these problems. Since ICA is an essentially exploratory technique where little constraints or a priori information is placed on the system, the method can be used to find brain regions that are functionally involved in a neural network in a data-driven manner.

1.3 Major Achievements

The major achievements of this thesis are listed as entries below:

- An ICA-based scheme of restoring fMRI has been constructed and implemented. First, a wide range of main artifacts that affect fMR images have been illustrated. In the experiments, noises were detected and removed on the basis of kurtosis values. Physiological fluctuations in fMRI were also reduced by detecting the independent component with the associated time-series that has the highest correlation with a phase model. The noisy dataset and the restored one were both processed by correlation technique and SPM to compare their capabilities of activation detection. From this study it can be concluded that

ICA technique is possible for restoration of fMR images thus improving the efficacy of detection techniques of activation. Because the tiny neural fluctuations in fMRI signals are easily affected by uninteresting confounds, this restoration model will help the accuracy of analysis results avoid being corrupted.

- An improved concept of functional connectivity has been developed in the higher-order statistical sense and ICA technique used for detecting connected brain regions. The fMRI datasets from resting brain were processed with eigen-image analysis and ICA, respectively, which showed an accordance with previous observations. The applicability and superiority of ICA was validated by finding higher-order functional connectivity with more number of components and less noisy imperfection.
- Structural Equation Modeling (SEM) has been applied into a functional network, where the regions involved in this model are selected from the component maps of SICA decomposition. An automated search for best fitting model, proposed by Bullmore, is employed to seek for the path model which can best describe the interactions among brain regions. From an experiment on visual-task data, it can be concluded to be an useful exploratory approach to analyzing effective connectivity for a specific brain operation.

1.4 Organization of the Thesis

This thesis is organized so as to describe my work from basic backgrounds to specific topic details. Totally it consists of 5 chapters.

Chapter 1 begins with a background review of my knowledge in the related area,

then gives the motivation of my research, with an outline of my achieved work.

In chapter 2, the mathematical framework of ICA technique is covered. Then, the descriptions of how ICA is applied to analyzing and restoring fMRI datasets are expressed.

Chapter 3 illustrates the basic concepts and the previous works in brain connectivity analysis, following which a higher-order functional connectivity definition is developed, with the application of ICA.

SEM approach to modeling neuronal interactions is described and an exploratory attempt is made in chapter 4.

Finally, the last chapter concludes this thesis and gives future directions, including enrichment of restoration model, some improvements in SEM, and etc.

Chapter 2

Restoration of fMRI Using ICA

This chapter provides a novel approach to full restoration of fMRI data. In the first part, the framework and mathematical principles of ICA is outlined. Later, a general ICA-based restoration model is described and various kinds of artifacts are illustrated for the first time. Then, the following two sections present the results of the experiments on reduction of noise and physiological fluctuations. The importance of this process lies on its effect of preventing the connectivity analysis from being affected by uninteresting confounds.

2.1 Introduction

In MR imaging system, it is desirable that most of the intensity changes occur from the neural sources. Nevertheless, different artifacts and noises continue to be a major hindrance to the accurate detection of neuronal activity. The confounding artifacts in fMRI mean signals that are not generated by neuronal activity, but by some external disturbances including head motion, blinking, etc., while those remaining signals of random environmental disturbances are referred to as noises. The interference of these factors, especially those completely uncorrelated with task-related changes, may tend

to mask the effects of activations related to task-performance, reducing the sensitivity and specificity of correlation analysis.

Sources of motion artifacts include gross involuntary subject motion and physiology-related motion such as respiration and cardiac pulsation. While gross motion can be minimized by physically constraining the subject, compensation of physiological motion is more difficult. A finite impulse response band-reject digital filters were designed to remove the physiological fluctuations, where the cardiac and respiratory information was obtained from a pulse oximeter [14]. Later, a retrospective approach was developed by directly subtracting estimated effects of Fourier series to remove the physiological artifacts [15]. However, both methods require external physiological monitoring of respiration and heartbeat. An image-based estimation and correction method that was free from this constraints was developed and described in [16].

An approach that modeled and removed movement-related artifacts from fMRI time-series was presented in [17], where movement-related effects were divided into those that were a function of position of the object in the frame of reference of the scanner and those that were due to movement in previous scans. In order to assess the effect of stimulus correlated motion, conventional visual and motor protocols were each performed and an image co-registration technique was used to retrospectively monitor subject motion [18]. Other attention has been placed on how the signal to noise ratio (SNR) influences the functional maps [19][20] of brain. A low-pass spatial filter has been attempted and validated to remove high-frequency noise from the images [21] and a frequency-window filtering method to observe oscillations linked to vasomotor [22].

Recently, independent component analysis (ICA) has become an rapidly evolving technique in analysis of biomedical signals (e.g. EEG, fMRI, optical imaging) [23][24][25]. The importance of the underlying independence assumption of ICA and its applicability has been evaluated in [26] [27] [28]. Some experiments and results have indicated that ICA can be used to reliably separate fMRI data sets into meaningful spatially-independent constituent components, including consistently and transiently task-related physiological changes, nontask-related physiological phenomena, and machine or movement artifacts [9]. A novel ICA paradigm, say ICA with reference (ICA-R), was proposed to produce only component maps corresponding to the input stimuli with much less computational and memory requirements [29].

The capability of ICA to decompose neuroimaging data into different components that are due to the various sources gives rise to a promising approach to restoring the datasets with minimum interferences of artifacts and noises. In the past, much investigation has been placed on eliminating interferences in EEG, EMG and ECG data. Using ICA, eye movement artifacts, muscle artifacts and cardiac contaminations have been removed from EEG and ECG signals [30][31][32][33].

However, these methods of interference cancellation for EEG, together with the previous approaches for fMRI [9][34][35], all require manual detection and classification of specific artifacts or components. In Cichocki's research [36], some relatively simple techniques to automatically detect and eliminate noise and artifacts were proposed for EEG recordings, using criteria for classification, ordering and detection of noisy signals. The work focused on several techniques: ICA and Higher Order Statistics (HOS) measure of Gaussianity (to detect and eliminate Gaussian noise), linear predictor (to detect i.i.d. sources and classify temporal structured sources) and Hurst

exponent (to detect randomness in independent components). Thus, the purpose of our work is to extend these automatic techniques with ICA to identify and eliminate broad kinds of artifacts and noises from fMRI dataset and then investigate the effect of the restoration method on the activation detection capability of correlation analysis and SPM.

2.2 ICA Algorithms

Independent component analysis (ICA) is essentially a method for extracting individual signals from mixtures of signals. It is worth stressing here that ICA does not incorporate any knowledge specific to the signals; in order to work, it requires simply that the individual signals or underlying components are statistically independent from each other, which is a stronger property than "uncorrelatedness".

Mathematically, statistical independence is defined in terms of probability densities. The random variables x and y are said to be *independent* if and only if

$$p_{x,y}(x, y) = p_x(x)p_y(y) \quad (2.2.1)$$

In words, the joint density $p_{x,y}(x, y)$ of x and y must factorize into the product of their marginal densities $p_x(x)$ and $p_y(y)$.

Independent random variables satisfy the basic property:

$$E\{g(x)h(y)\} = E\{g(x)\}E\{h(y)\} \quad (2.2.2)$$

where $E\{ \}$ denotes expectation and $g(x)$ and $h(y)$ are any absolutely integrable functions of x and y , respectively, while uncorrelatedness implies:

$$E\{xy\} = E\{x\}E\{y\} \quad (2.2.3)$$

This equation can be obtained from independence property Eq.(2.2.2) as a special case where both $g(x)$ and $h(y)$ are linear functions, and takes into account second-order statistics (correlation or covariance) only. This demonstrates why independence is a stronger sense of uncorrelatedness.

Essentially, if two variables are independent, then the value of one variable provides absolutely no information about the value of the other variable. By contrast, even though two variables are uncorrelated, the value of one variable can still provide information about the value of the other variable.

Another very important principle of ICA estimation is that there is at most one Gaussian component in the mixture. This is because if the random variables have gaussian distributions, independence and uncorrelatedness become the same thing.

The general strategy underlying ICA can be summarized as follows:

1. It is assumed that different physical processes (e.g. two speakers) give rise to independent source signals. Specifically, sources are assumed to be statistically independent.
2. A measured signal (e.g. a microphone output) usually contains contributions from many different physical sources, and therefore consists of a mixture of unrelated source signals.
3. Unrelated signals are usually statistically independent, with a maximum entropy. Therefore, if a set of signals with maximum entropy can be recovered from a set of mixtures, then such signals are independent.

For computational and conceptual simplicity, the problem can be sought as a linear transformation of the original data. Assume that there is a k -dimensional zero-mean

vector $\mathbf{c}(t) = (c_1(t), c_2(t), \dots, c_k(t))^T$, such that the components $c_i(t)$ are mutually independent. The vector $\mathbf{c}(t)$ corresponds to k independent scalar-valued source signals $c_i(t)$. A zero-mean m -dimensional data vector $\mathbf{x}(t) = (\mathbf{x}_1(t), \dots, \mathbf{x}_m(t))^T$ is observed at each time point t , such that

$$\mathbf{x}(t) = \mathbf{M}\mathbf{c}(t) \quad (2.2.4)$$

where $\mathbf{M}$ is a full-rank $m \times k$ scalar matrix. Then the problem is to determine a constant (weight) matrix $\mathbf{W}$ so that the elements of linear transformation of the observation variables

$$\mathbf{c}(t) = \mathbf{W}\mathbf{x}(t) \quad (2.2.5)$$

are maximally independent of each other.

In what follows, we are going to describe two main algorithms for estimating the ICA model, which are used in my work for fMRI analysis. Readers who are interested in details of other ICA algorithms can refer to the book [37].

2.2.1 Fast ICA Algorithm

The fast ICA algorithm is an approach to computing the demixing matrix $\mathbf{W}$ in Eq.(2.2.5) by minimizing mutual information [38]. First, the differential entropy H of the vector of components $\mathbf{c} = (c_1, \dots, c_n)^T$ with density $p(\cdot)$ is defined as follows:

$$H(\mathbf{c}) = - \int p(\mathbf{c}) \log p(\mathbf{c}) d\mathbf{c} \quad (2.2.6)$$

Differential entropy can be normalized to give rise to the definition of negentropy, which can be interpreted as a measure of nongaussianity [39]:

$$J(\mathbf{c}) = H(\mathbf{c}_{gauss}) - H(\mathbf{c}) \quad (2.2.7)$$

where $\mathbf{c}_{gauss}$ is a Gaussian random vector of the same covariance matrix as $\mathbf{c}$. Using the concept of differential entropy, one can define the mutual information I between the n random variables $c_i, i = 1 \dots n$ [39]:

$$I(c_1, c_2, \dots, c_n) = J(\mathbf{c}) - \sum_i J(c_i) \quad (2.2.8)$$

Mutual information is a natural measure of the dependence between random variables. It is particularly interesting to express mutual information using negentropy, constraining the variables to be uncorrelated.

Since mutual information is a natural measure of the dependence between random variables, the ICA of the random vector $\mathbf{x}$ as an invertible transformation $\mathbf{c} = \mathbf{W}\mathbf{x}$ is to determine the demixing matrix $\mathbf{W}$ so that the mutual information of the transformed components c_i is minimized. Note that mutual information is not affected by multiplication of the components by scalar constants. Therefore, the definition only defines the independent components up to some multiplicative constants.

In the simplest case, the negentropy could be approximated in the form [40]

$$J(c_i) \approx [c\{E\{G(c_i)\} - E\{G(\nu)\}\}]^2 \quad (2.2.9)'$$

where G is practically any non-quadratic function, c is an irrelevant constant, and ν is a Gaussian variable of zero mean and unit variance.

The approximation of negentropy given above gives readily an objective function for estimating the ICA transform in our framework. First, to find one independent component, we maximize the function J given by

$$J(w) = [E\{G(\mathbf{w}^T \mathbf{x})\} - E\{G(\nu)\}]^2 \quad (2.2.10)$$

with $\mathbf{w}$ is an m -dimensional weight vector constrained so that $E\{(\mathbf{w}^T \mathbf{x})^2\} = 1$ (we can fix the scale arbitrarily). The other independent components can be estimated

one by one using a deflation scheme. Here we choose the contrast function $G(u) = \frac{1}{4}u^4$ and its derivative $g(u) = u^3$. The choice of the function G is deferred to the end of this section.

Second, using the approach of minimizing mutual information, the above one-unit contrast function can be extended to compute the whole matrix $\mathbf{W}$ in Eq.(2.2.5). Recalling from Eq.(2.2.8) that mutual information is minimized when the sum of the negentropies of the components is maximized. Maximizing the sum of n one-unit contrast functions, and taking into account the constraint of decorrelation, one obtains the following optimization problem:

$$\begin{aligned} & \text{maximize } \sum_{i=1}^n J(\mathbf{w}_i) \\ & \text{under constraint } E\{(\mathbf{w}_k^T \mathbf{x})(\mathbf{w}_j^T \mathbf{x})\} = \delta_{jk} \end{aligned} \quad (2.2.11)$$

Thus we have defined our ICA estimator by an optimization problem.

A fixed-point iteration algorithm is chosen for maximizing the contrast function in Eq.(2.2.10). According to the Kuhn-Tucker conditions, the optima of $E\{G(\mathbf{w}^T \mathbf{x})\}$ under the constraint $E\{(\mathbf{w}^T \mathbf{x})^2\} = \|\mathbf{w}\|^2 = 1$ are obtained at points where

$$E\{\mathbf{x}g(\mathbf{w}^T \mathbf{x})\} - \beta \mathbf{w} = 0 \quad (2.2.12)$$

where β is a constant that can be easily evaluated to give $\beta = E\{\mathbf{w}_0^T \mathbf{x}g(\mathbf{w}_0^T \mathbf{x})\}$, where $\mathbf{w}_0$ is the value of $\mathbf{w}$ at the optimum. Solving this equation by Newton's method, the fixed-point algorithm is given:

$$\begin{aligned} \mathbf{w}^+ &= E\{\mathbf{x}g(\mathbf{w}^T \mathbf{x})\} - E\{g'(\mathbf{w}^T \mathbf{x})\}\mathbf{w} \\ \mathbf{w}^* &= \mathbf{w}^+ / \|\mathbf{w}^+\| \end{aligned} \quad (2.2.13)$$

The above one-unit algorithm can be used to construct a system of n neurons to estimate the whole ICA transformation using the multi-unit contrast function. To

prevent different neurons from converging to the same maxima we must decorrelate the outputs $\mathbf{w}_1^T \mathbf{x}, \dots, \mathbf{w}_n^T \mathbf{x}$ after every iteration.

A simple way of achieving decorrelation is a deflation scheme based on a Gram-Schmidt-like decorrelation:

$$\begin{aligned} 1. \text{ Let } \mathbf{w}_{p+1} &= \mathbf{w}_{p+1} - \sum_{j=1}^p \mathbf{w}_{p+1}^T \mathbf{w}_j \mathbf{w}_j \\ 2. \text{ Let } \mathbf{w}_{p+1} &= \mathbf{w}_{p+1} / \sqrt{\mathbf{w}_{p+1}^T \mathbf{w}_{p+1}} \end{aligned} \quad (2.2.14)$$

This means that we estimate the independent components one by one. When p independent components have been estimated, the one-unit algorithm for $\mathbf{w}_{p+1}$ is run, and after every iteration step the 'projections' $\mathbf{w}_{p+1}^T \mathbf{w}_j \mathbf{w}_j, j = 1, \dots, p$ of the previously estimated p vectors are subtracted from $\mathbf{w}_{p+1}$, and then $\mathbf{w}_{p+1}$ is renormalized.

Hence, after deriving the demixing matrix $\mathbf{W}$, the matrix of components $\mathbf{C}$ can then be obtained by computing Eq.(2.2.5).

Regarding the question of choosing the contrast function G , two important criteria are adopted in particular. First, the contrast function should be fast to compute. It is noted that polynomial functions tend to be faster to compute than, say, the hyperbolic tangent. Second, in the case where kurtosis is used as a contrast function, i.e., if $G(u) = u^4$, the convergence of the algorithm in Eq.(2.2.13) is global and it does converge to the right extrema. For details of the proof, please see [38].

Actually, in contrast to many other ICA methods, the framework provides estimators that work for any distributions of the independent components and for any choice of the contrast function. The choice of the contrast function is only important if one wants to optimize the performance of the method.

2.2.2 Maximum Likelihood Algorithm

An extension of the infomax algorithm of Bell and Sejnowski [34] has been proposed that is able to blindly separate mixed signals with sub- and supergaussian source distributions [41].

At this point, we consider the case where the number of sources is equal to the number of sensors $m = k$ and at most one source is normally distributed. The goal of ICA is to find a linear mapping $\mathbf{W}$ such that the unmixed signals $\mathbf{u}$,

$$\mathbf{u}(t) = \mathbf{W}\mathbf{x}(t) = \mathbf{W}\mathbf{M}\mathbf{c}(t) \quad (2.2.15)$$

The learning algorithm can be derived using the maximum likelihood formulation. The probability density function (p.d.f.) of the observations $\mathbf{x}$ can be expressed as:

$$p(\mathbf{x}) = |\det(\mathbf{W})|p(\mathbf{u}) \quad (2.2.16)$$

where $p(\mathbf{u}) = \prod_{i=1}^k p_i(u_i)$ is the hypothesized distribution of $p(c)$. The log-likelihood of Eq.(2.2.16) is:

$$L(\mathbf{u}, \mathbf{W}) = \log|\det(\mathbf{W})| + \sum_{i=1}^k \log p_i(u_i) \quad (2.2.17)$$

An efficient way to maximize the log-likelihood is to follow the natural gradient:

$$\Delta \mathbf{W} \propto \frac{\partial L(\mathbf{u}, \mathbf{W})}{\partial \mathbf{W}} \mathbf{W}^T \mathbf{W} = [\mathbf{I} - \varphi(\mathbf{u})\mathbf{u}^T] \mathbf{W} \quad (2.2.18)$$

where

$$\varphi(\mathbf{u}) = -\frac{\frac{\partial p(\mathbf{u})}{\partial \mathbf{u}}}{p(\mathbf{u})} = \left[-\frac{\frac{\partial p(u_1)}{\partial u_1}}{p(u_1)}, \dots, -\frac{\frac{\partial p(u_k)}{\partial u_k}}{p(u_k)} \right]^T \quad (2.2.19)$$

A symmetric strictly subgaussian density can be modelled using a symmetrical form of the Pearson mixture model:

$$p(u) = \frac{1}{2} \left(N(\mu, \sigma^2) + N(-\mu, \sigma^2) \right) \quad (2.2.20)$$

where $N(\mu, a^2)$ is the normal density with mean μ and variance a^2 . Defining $a = \frac{\mu}{\sigma^2}$ and applying into Eq.(2.2.20), we may write for $\varphi(u)$:

$$\varphi(u) = -\frac{\frac{\partial p(u)}{\partial u}}{p(u)} = \frac{\mu}{\sigma^2} - a \left(\frac{\exp(au) - \exp(-au)}{\exp(au) + \exp(-au)} \right) \quad (2.2.21)$$

Using the definition of the hyperbolic tangent, we can write

$$\varphi(u) = \frac{u}{\sigma^2} - \frac{\mu}{\sigma^2} \tanh\left(\frac{\mu}{\sigma^2} u\right) \quad (2.2.22)$$

Setting $\mu = 1$ and $\sigma^2 = 1$, the learning rule for strictly subgaussian source is now (Eq.(2.2.18) and Eq.(2.2.22)):

$$\Delta \mathbf{W} \propto \left[\mathbf{I} + \tanh(\mathbf{u})\mathbf{u}^T - \mathbf{u}\mathbf{u}^T \right] \mathbf{W} \quad (2.2.23)$$

In the case of unimodal supergaussian sources, we adopt the following density model:

$$p(u) \propto p_G(u) \text{sech}^2(u) \quad (2.2.24)$$

where $p_G(u) = N(0,1)$ is a zero-mean gaussian density with unit variance. The nonlinearity $\varphi(u)$ is now:

$$\varphi(u) = -\frac{\frac{\partial p(u)}{\partial u}}{p(u)} = u + \tanh(u) \quad (2.2.25)$$

The learning rule for supergaussian sources is (Eq.(2.2.18) and (Eq.(2.2.25)):

$$\Delta \mathbf{W} \propto \left[\mathbf{I} - \tanh(\mathbf{u})\mathbf{u}^T - \mathbf{u}\mathbf{u}^T \right] \mathbf{W} \quad (2.2.26)$$

The difference between the supergaussian learning rule in Eq.(2.2.26) and the subgaussian learning rule in Eq.(2.2.23) is the sign before the tanh function:

$$\Delta \mathbf{W} \propto \left[\mathbf{I} - \mathbf{K} \tanh(\mathbf{u})\mathbf{u}^T - \mathbf{u}\mathbf{u}^T \right] \mathbf{W} \begin{cases} k_i = 1 & : \text{supergaussian} \\ k_i = -1 & : \text{subgaussian} \end{cases} \quad (2.2.27)$$

where k_i are elements of the k -dimensional diagonal matrix $\mathbf{K}$. The switching parameters k_i can be derived from the generic stability analysis of separating solutions:

$$k_i = \text{sign}\left(E\{\text{sech}^2(u_i)\}E\{u_i^2\} - E\{\tanh(u_i)u_i\}\right) \quad (2.2.28)$$

2.3 ICA-Based Model of fMRI Restoration

The fMRI signal associated with a given voxel is affected by a subject's general arousal levels, the experimental task being executed, drifting sensor outputs, artifacts and other noises. Thus the signal at each voxel presumably consists of a mixture of underlying source signals, which leads to the basic model of ICA. If a signal mixture can be decomposed into a set of statistically independent, non-Gaussian signals, then these signals are likely to be the source signals of the mixed fMRI signals.

In fMRI analysis, ICA can be used in two complementary ways to decompose an image sequence into a set of images and a corresponding set of time-varying image amplitudes, which are defined *Spatial ICA* (SICA) and *Temporal ICA* (TICA), respectively.

Denote by $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_m)^T$ the matrix of noisy fMRI data, where $\mathbf{x}_i = (x_{i1}, \dots, x_{in})$ is an intensity brain map scanned in each time and x_{ij} is the measured j^{th} voxel intensity value in the i^{th} scan of the noisy signals. SICA embodies the assumption that $\mathbf{X}$ can be expressed rigorously by writing a linear equation relating the component maps and their time courses to the measured fMRI signals [9]:

$$\mathbf{X} = \mathbf{M}\mathbf{C} \quad (2.3.1)$$

where $\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_k)^T$ is the matrix of k spatially independent component maps and $\mathbf{M} = (\mathbf{m}_1, \dots, \mathbf{m}_k)$ the $m \times k$ mixing matrix. Thus the voxel values for each

of the component maps are placed in separate rows of matrix $\mathbf{C}$, and each column of $\mathbf{M}$ represents a time-series of a corresponding component. Hence, the matrix of component maps is computed by multiplying the observed data matrix $\mathbf{X}$ by an unmixing matrix $\mathbf{W}$:

$$\mathbf{C} = \mathbf{W}\mathbf{X} \quad (2.3.2)$$

If $\mathbf{W}$ is a square matrix of full rank, its inverse $\mathbf{M} = \mathbf{W}^{-1}$ is well-defined. ICA technique then aims to determine the unknown unmixing matrix $\mathbf{W}$ to decompose the observed fMRI signals into statistically independent component maps.

On the other hand, TICA embodies the assumption that the transform of fMRI scans $\mathbf{X}_T = \mathbf{X}^T$ can be decomposed as:

$$\mathbf{X}_T = \mathbf{M}_T \mathbf{C}_T \quad (2.3.3)$$

In the above equation, $\mathbf{M}_T$ is an $n \times k$ mixing matrix in which each column is an image, and $\mathbf{C}_T$ is an $k \times m$ matrix of k statistically independent temporal sequences. Thus, m_{ij} represents the magnitude of contribution of the j th source time-series to the j th voxel.

Functional organization of the brain is based on two complementary principles, localization and connectionism [42]. Localization implies that each psychomotor function is performed principally in a small set of brain areas. The complementary principle of connectionism posits that the brain regions involved in a given psychomotor function may be widely distributed.

Consistent with these principles, it was suggested that the multi-focal brain areas activated by performance of a psychomotor task should be unrelated to the brain areas whose signals are affected by artifacts, such as physiological pulsations, subtle head movements, and machine noise which may dominate fMRI experiments [9]. Each

Figure 2.1: Conceptual model of restoration of fMRI data

of these separate processes may be represented by one or more spatially-independent components, and associated with a single time course of enhancement or suppression and a component map. This was just the implication of SICA.

2.3.1 The Conceptual Restoration Model

It has been demonstrated that ICA is capable of extracting different sources signals underlying fMRI data: nontask-related signal components, movements and other artifacts as well as consistently or transiently task-related fMRI activations, based only on weak assumptions about their spatial distributions and without a priori assumptions about their time courses [9].

Up to now, researchers mainly used SICA for fMRI analysis. Because the number of voxels in the scanned images is far more than that of time points, the computation of TICA requires huge workload of time and memory.

By this means, a model for eliminating noise and other undesirable artifacts from

noisy fMRI data and restoring it by using ICA approach can be conceived, which is depicted in Fig(2.1). In our work, "restoration" means processing the fMRI dataset to convert it into its original or "clean" condition, which is, by and large, composed of only neural signals or subject to minimum interferences of artifacts and noises.

First, a demixing process of ICA is performed using a robust neural network algorithm (here we employs Fast ICA [38]) by linear transformation of the noisy data as given in Eq.(2.3.2). It is based on the assumption that the brain activity, the artifacts and noises are spatially separate processes, and their independence is reflected in the statistical relation between the fMRI intensity signals generated by those processes. In order to perform ICA, it is necessary to have at least as many mixtures as there are independent sources ($m \geq k$).

In the next stage, these independent source signals are extracted and classified in terms of their different properties under certain criteria and those undesirable components are thereby removed by switching corresponding switches "off", i.e., setting the values of the corresponding rows of $\mathbf{C}$ to zeros. This step can be formulated as :

$$\hat{\mathbf{C}} = \mathbf{S}\mathbf{C} \quad (2.3.4)$$

where $\mathbf{S}$ is a k -by- k switching matrix in which $s_{ii} = 1$ if the i th switch is "on" meanwhile the remainder elements in the matrix are all zeros. The identification methods of a number of artifacts and noises which are deemed as undesirable signals will be further discussed in the later part of this section.

Finally, a reconstruction process of projecting the useful independent components $\hat{\mathbf{c}}_i$ back onto the mixed signals is done by applying the linear inverse transformation:

$$\hat{\mathbf{X}} = \mathbf{M}\hat{\mathbf{C}} \quad (2.3.5)$$

where $\mathbf{M}$ is pseudo inverse of demixing matrix $\mathbf{W}$ and $\hat{\mathbf{C}}$ is the filtered matrix of independent components. Thus it is reasonable to assume that if adequate information is available on the characteristics of underlying signals, a full restoration of fMRI data is possible by this approach which depends on selection and rejection of the components. Therefore, the major challenge of the construction of our model resides in the identification of individual artifacts and noises.

2.3.2 Basic Measures of Identifying Components

Kurtosis

Kurtosis is a fourth order statistics [37] defined in the zero-mean case by the equation:

$$\kappa(x) = E\{x^4\} - 3[E\{x^2\}]^2 \quad (2.3.6)$$

The expectation could be calculated from the average over all the elements of each row of matrix $\mathbf{C}$ which corresponded to each independent component.

A very important feature of kurtosis is that it is the simplest statistical quantity for indicating the nongaussianity of a random variable. The kurtosis value is zero if the variable has a Gaussian distribution. Since random noises are always supposed to bear a nearly gaussian distribution, the component maps of noises in fMRI can then be identified and removed in terms of their kurtosis values by using a cut-off value as the noise removal threshold.

Frequency Spectrum

Different independent components linked to neural activations and physiological fluctuations are temporally characterized by distinct magnitudes in the frequency domain. With a digital filter, some experiments have demonstrated that physiological signals belong to ranges of relatively high frequencies [14] while the neural signals are usually low-frequency fluctuations caused by the sensitivity of blood flow and blood oxygenation.

Hurst Exponent

The Hurst exponent H (and associated fractal dimension $D = 2 - H$) is one possible parameter for characterizing a time series [43]. Hurst et al. developed the rescaled range (R/S) analysis for a time series $y(t)$. The range R was defined as a difference between maximum and minimum "accumulated" values:

$$R = \max_{1 \leq T \leq N} \{Y(T)\} - \min_{1 \leq T \leq N} \{Y(T)\} \quad (2.3.7)$$

where

$$Y(T) = \sum_{t=1}^T \left(y(t) - \bar{y}(t) \right) \quad (2.3.8)$$

and second, the standard deviation S was estimated from the observed value $y(t)$:

$$S = \left(\frac{1}{N} \sum_{t=1}^N [y(t) - \bar{y}(t)]^2 \right)^{\frac{1}{2}} \quad (2.3.9)$$

In the above equations, N is the number of samples and $\bar{y}(t)$ denotes the average over the first t samples. Hurst found that the ratio R/S is very well described for a large number of phenomena by the following nonlinear empirical relation:

$$\frac{R}{S} = (cN)^H \quad (2.3.10)$$

where c is some constant (typically $c = \frac{1}{2}$) and H is the Hurst exponent in the range from 0 to 1.

For calculation of H , we used a recurrent method from [44] [30]. It follows Eq.(2.3.10) that:

$$H = \frac{\log_{10}(\frac{R}{S})}{\log_{10}(\frac{N}{2})} \quad (2.3.11)$$

The $Y(T)$ is recursively computed as

$$Y(T) = Y(T-1) + \frac{T-1}{T} \left(y(t) - \bar{y}(t-1) \right) \quad (2.3.12)$$

and $\bar{y}(t)$ is averaging of the process $y(t)$:

$$\bar{y}(t) = \frac{t-1}{t} \bar{y}(t-1) + \frac{1}{t} y(t) \quad (2.3.13)$$

The standard deviation S is obtained by recursively computing its square according to:

$$S = \left(D(N) \right)^{\frac{1}{2}} \quad (2.3.14)$$

where $D(t)$ is computed as:

$$D(t) = \frac{t-1}{t} D(t-1) + \frac{(t+1)^2}{t^3} \left(y(t) - \bar{y}(t-1) \right)^2 \quad (2.3.15)$$

The initial condition is $\bar{y}(1) = y(1)$.

The values of the Hurst exponent range between 0 and 1. A value of 0.5 indicates a true random walk (a Brownian time series). In a random walk there is no correlation between any element and a future element. A Hurst exponent value H , $0.5 < H < 1$ indicates "persistent behavior" (e.g., a positive autocorrelation). If there is an increase from time step $t-1$ to t there will probably be an increase from t to $t+1$. The same is true of decreases, where a decrease will tend to follow a decrease. A Hurst exponent

value $0 < H < 0.5$ will exist for a time series with "anti-persistent behavior" (or negative autocorrelation). Here an increase will tend to be followed by a decrease. Or a decrease will be followed by an increase. At the limit of $H = 0$ the time series must change direction every sample and a straight line with nonzero slope will have the Hurst exponent of 1.

The Hurst exponent H could be employed to detect and remove undesirable ICs. The time series of most interesting or desirable components have been studied to have a Hurst exponent in the range $H = 0.70 - 0.76$ [30]. Furthermore, it is found by extensive computer experiments that some artifacts, such as those from eye blinking or heart beats, have a specific value of H [30]. Thus, they could be automatically identified and removed on the basis of their Hurst exponent values.

2.3.3 Identification and Elimination of Specific Components in fMRI

In fMRI images, besides the activation components, there exist other confounding artifacts and noises. The tiny signal change coming from neuronal activity is vulnerable to various kinds of artifacts and noises. ICA gives a method for separating these components which may be independent from each other, mixed in the noisy dataset. Here we give a brief description of various kinds of possible components affecting the fMRI signals and the exploratory approaches to classifying and detecting them. The "0/1" states of the soft switches in Fig(2.1) can then be automatically determined by certain values of the properties of the separated independent components.

Activation Induced by Scanner Acoustic Noise

A concern in functional MRI studies is that brain activation produced by MRI gradient switching acoustic noise may corrupt results primarily in studies of auditory cortex activation and of cognitive processes that may be modulated by ambient noise [45]. In that experiment, two types of time series was compared. The first, considered "task", involved applying only EPI gradients for 20s without the application of RF pulses, then without pause, starting image collection. The second, called "control", involved typical sequential image acquisition without the prior gradient pulses. Signal enhancement was demonstrated bilaterally in the auditory cortex for the first 5s, during which the BOLD signal in the "control" series was in the process of increasing as a result of the EPI gradient noise while the BOLD signal in the "task" series was already saturated in the "on" state.

Therefore, the component of acoustic-noise-induced activation may be detected by a signal enhancement in the first several seconds of the time series which has a evident effect on the left and right auditory cortex.

Cardiac Pulsation

Because of the finite cranial volume, the influx of pumped blood results in a dynamic interaction between the competing space requirements of the blood volume, cerebrospinal fluid (CSF) pools and brain tissue [46][47][48]. These physical changes can have important implications for cognitive experiments done with fMRI. Bulk brain motion or changes in blood volume at the capillary level could cause widespread global fluctuations of measured signal intensity with the cardiac cycle. Large-vessel pulsatility may affect the measured signal intensity in the areas adjacent to the vessels, both

by causing tissue movement and by producing an influx of unsaturated blood into the slice of interest. These causes of increased signal variation can reduce the ability to detect hemodynamic changes related to neural activity using fMRI.

Due to their characteristic of frequency, these signals could be detected in terms of their frequency spectrum when transforming the time course into the frequency domain. Previous studies have measured the heart rate to be around from 0.75 to 1.5 Hz corresponding to heart rates varying from 45 to 90 beats/min [14].

An alternative way of detection is on the basis of the Hurst exponent H , which is one possible parameter for characterizing a time series of living organism as complex nonlinear dynamic systems [43]. The heart-beat artifacts are usually characterized by $H = 0.64 - 0.69$ [30]. The accuracy of the detection may be enhanced if both of these methods are performed to the dataset.

Respiratory Motion

Respiratory artifact is caused by the variation of magnetic field distribution in the brain. It is found to produce both global changes in phase image and localized variations in magnitude image, especially near ventricles. In addition, regions closer to the chest are more severely affected. Moreover, as we move to high-field systems for more BOLD contrast, these artifacts increase with field strength. In the frequency domain, the respiratory rate was measured 0.1-0.5Hz [14], corresponding to respiration rates varying from 6 to 30 cycles/min.

Eye Blinking

In an experiment, some physiological activities of the subject are hard to be inhibited or controlled. Those signals should therefore be taken into account when analyzing the components in fMRI, one of which is the eye blinking (or wink). From extensive computer experiments [30], it was found that the Hurst exponent of eye blinking artifacts $H = 0.58 - 0.64$. The time course of the components corresponding to eye blinking is hypothesized to have sharp variations from time to time.

Saccade

In analogy to the eye blinking, the saccade artifacts are also due to the untractable eye movements in fMRI experiments. Sharp variations over some particular time points would occur in the time series of this category of artifacts. More further studies need to be done for the detection.

Biting

Apart from the eye movements, subjects in fMRI experiments would inevitably perform acting of biting during the trials. Closely grouped oscillations or variations are supposed to arise in the time course of the components grounded on a small number of times of biting during a cycle of trial. From the perspective of their spatial distribution, this kind of artifact is probably influential in the motor area.

Head Motion

Head motions are very common and somehow significant confounding phenomena of a subject because to date it is still difficult to position the subject to avoid any movement of the head. The ICA decomposition of some trials [9] shows components which have abrupt changes in their time course and/or ring-like spatial distributions could be tentatively interpreted as arising from small abrupt or gradual head movements during the trials.

Vasomotor Oscillations

In addition to physiological oscillations, some low frequency fluctuations are found arising out of fluctuations in hemodynamics. These fluctuations are at least in part due to vasomotor oscillations with frequencies around at 0.1 Hz [1].

Drifts of CSF

The brain is bathed inside and out by cerebrospinal fluid (CSF), a fluid containing nutrients and byproducts of brain activity. Because of its fluidity, the substances drifting over the CSF would unavoidably interfere the signals in the brain areas, yet with relatively slow frequencies.

Other Random Noise

Different from the various kinds of artifacts discussed above, these signals are not underlyingly related to the activities or changes caused by human body. They may be thus categorized into random noises for the sake of ease which are usually generated by the field strength, MR manufacturer, installation environment, coil used for the

fMRI experiments, or other random fluctuations of the external environment.

The identification of noises are based on the fourth order statistics called the *kurtosis*, which is the simplest measure of nongaussianity. Since random noises are always supposed to bear a nearly gaussian distribution, the component maps of noises in fMRI will statistically have a kurtosis value close to zero.

2.4 Reduction of Noise and Its Implementation

As aforementioned, noise is an ubiquitous source of interference in the neuroimaging data. This confounding factor, due to its randomness and instability, is completely uncorrelated with task-related changes, and may tend to mask the effects of activations related to task-performance, reducing the sensitivity and specificity of some statistical analysis on the fMRI signals, e.g., correlation analysis. In this section, the results of denoising fMRI datasets collected in visual task, memory retrieval task experiments and synthetic dataset are presented. Although a full restoration of fMRI signals is possible, only noise removal based on kurtosis value is demonstrated here.

2.4.1 Visual Task

An 8-Hz alternating checkerboard pattern with a central fixation point was projected on a LCD system, and subjects were asked to fixate on the point during simulations. Images were acquired at three axial levels of the brain at the visual cortex of five subjects [11]. The fast ICA algorithm was applied to the image datasets being implemented in MATLAB 6.5 to obtain the transformation matrix $\mathbf{W}$ and the component maps $\mathbf{C}$ in Eq.(2.3.2).

Figure 2.2: Kurtosis of independent components on visual task data

Fig. (2.2) shows the kurtosis of the independent component maps of a representative visual task dataset. A cut-off value of $kurt = 3.0$ was used as the noise removal threshold after which nearly two-thirds of the component maps were removed.

After performing the denoising scheme on the dataset, the correlation analysis technique was employed to investigate the effect of noise elimination. As has been illustrated in Section 1.1.3, those voxels, whose signals are positively correlated with the reference function above a preselected threshold, are designated "areas of activation".

Although this method is both computationally simple and reasonably effective, it has several major drawbacks, of which the most significant is the bulk of the measured signals produced by other time-varying phenomena. It is hypothesized that a denoising preprocessing may enhance the accuracy of correlation analysis.

In Fig.(2.3), a threshold of 0.5 corresponding to a p-value of 5% was selected to detect the activation area to the visual task datasets, which indicates that the activation map after denoising appeared to have a higher extent of localization of

Figure 2.3: Maps of activation on visual task data, threshold of correlation for both maps:0.5 (a)before denoising (b)after denoising; colored blobs show the activation

Figure 2.4: Maps of activation on visual task data: (a)before denoising, threshold of correlation:0.4, (b)after denoising, threshold of correlation:0.5

activation voxels and more voxels were able to detected. If the threshold value was reduced to 0.4, the activation map from the noisy images produced as many activated voxels as that after denoising at a threshold of 0.5, however a trade-off of worse localization occurred shown in Fig.2.4(a). These findings suggest that the denoising to the mixed images can increase the correlation coefficient value of the activated voxels hence improving the accuracy of detection.

2.4.2 Memory Retrieval Task

In this memory task experiment, four subject learned three different sets (sizes of 4,6, and 8) of letters prior to the actual experiment with a corresponding cue for each set. During each trial of the experiment, a cue for the set and a letter were presented and the subject decided whether the letter corresponds to the indicated set; 48 trials were presented each for every 15.3s and a delay of 2.0s was maintained between the presentation of the cue and probe. The processes involved in the brain during the experiment include encoding of the cue and probe, retrieval of information from secondary memory, scanning of primary memory, response selection, and response execution. Fourteen slices were acquired with 19.2cm FOV, 5mm thickness, 2mm spacing, and 64x64 matrix using single shot gradient echo-planer imaging (EPI) sequences (TE=46ms and TR= 1700ms).

The brain maps obtained at an axial level using SPM which show significant activated voxels, one without being processed by our noise removal technique and the other having been denoised, are shown in Fig.(2.5). Both maps show activation in the anticipated cortical areas of the brain. The activations detected by SPM after denoising are more focal, local to the memory cortical areas, and more acceptable

Figure 2.5: Activations obtained from a representative subject in the memory retrieval task experiment using SPM approach (a)before denoising (b)after denoising; colored blobs show the activation

and neurophysiologically correct. Thus, the activation patterns after a restoration of fMRI data, are less noisy.

2.4.3 Synthetic Data

Figure 2.6: Kurtosis of independent components on synthetic fMRI data

Because with real data, quantitative measures are not possible, we need synthetic dataset to further investigate our approach. A synthetic fMRI dataset was created by

Figure 2.7: Maps of activation on synthetic data, threshold of correlation for both maps:0.5 (a)before denoising (b)after denoising

extracting brain scans acquired in rest conditions from a dataset obtained in the above visual experiment. A ground truth of activation was generated by processing the original visual dataset using correlation analysis. Box-car time series were designed for the activated pixels and the responses were generated by convolving the box-car time series with a gamma HRF, which were then added to the activated voxels with a contrast-to-noise ratio between 1-10%.

Fig.(2.6) shows the kurtosis of the components in synthetic data where a cut-off value of $kurt = 2.0$ was adopted. The activation maps of synthetic data are shown in Fig.(2.4). Some tiny blobs which represented activations over the brain image were excluded after denoising, which reduced the value of false positive rate in ROC analysis. Also, the ROC curves of the relationship between the true-positive rate (proportion of correctly detected activations to all added activations) and the false-positive rate (proportion of pixels that were incorrectly recognized as active in all pixels without added activations) for both maps are compared in Fig.(2.8) by adjusting different thresholds of correlation. The experiment on synthetic data

Figure 2.8: The ROC curves for the synthetic data

gives quantitative results where the ROC curve after denoising obviously indicates superiority in accuracy of detection.

2.5 Reduction of Physiological Fluctuations and Its Implementation

In this section, the approach to identification and elimination of physiological sources, primarily linked to heart and respiratory fluctuations is proposed and its experiment result presented, being implemented under our restoration model.

2.5.1 Method of Identification

Although the time-series of the cardiac and respiratory changes can be featured by distinct powers in their frequency spectrum, it still requires the manual detection of the corresponding components. We are going to propose an approach to identifying

and removing the physiological components automatically from the fMRI data.

The approach consists of (a) finding the two dominant frequencies that are appeared in the largest number of pixels (b) looking for the pixels that contain strong physiological fluctuations as basic cardiac and respiratory silhouettes (c) ordering the data into unit respiratory and cardiac cycles based on detected timing profiles (d) estimating corresponding physiological effects by nonlinear least-square fitting (e) detecting the independent component with the associated time series that has the highest correlation with the phase model (f) removing the detected components by subtraction.

An efficient approach to estimation of physiological signals without synchronization means with external monitoring were proposed previously [16]. First, each pixel time-course in the brain is normalized by its mean value across all the time points. Then the dominant frequency components in the respiratory and cardiac frequency ranges are identified from the spectrum of each pixel. The respiratory frequency range is from 0.1 to 0.5 Hz, corresponding to respiration rates varying from 6 to 30 cycles/min; the cardiac frequency range is from 0.75 to 1.5 Hz, corresponding to heart rates varying from 45 to 90 beats/min. Both kinds of motions are the most prominent periodic signal sources in the brain, the periodicity should be observed in the signal time-courses of most brain regions. Hence the respiratory and cardiac frequencies can be estimated by the frequency components that most pixel time-courses contain. Two histograms of dominant frequency components in the respiratory and cardiac frequency ranges were generated. The corresponding physiological frequency could be determined by the frequency with the largest number of pixels in the histogram.

To estimate respiratory and cardiac signals, we first look for the pixels that contain

strong physiological fluctuations. The pixels with dominant frequency components around the identified physiological frequencies are selected and their relative powers are computed to determine the amount of physiological fluctuations. The relative power is calculated by dividing the power within 0.1 Hz, which is the width of the usual frequency peak, centered at the dominant frequency, by the power within the respiratory and cardiac frequency range mentioned in the previous paragraph. Then the estimated physiological signals are obtained by band-pass filtering the pixel time-courses with the highest respiratory or cardiac power.

Then, we used a method similar to that proposed by Hu [15] but applied it to subtracting independent components from mixed signals. In this method, images are reordered into normalized respiratory and cardiac cycles by their relative phases (denoted as $\theta(i)$, where i represents the i th image) within the corresponding physiological period:

$$\theta_i = \frac{\mathbf{T}(i) - \mathbf{T}_s(j)}{\mathbf{T}_s(j+1) - \mathbf{T}_s(j)} \quad (2.5.1)$$

where $\mathbf{T}(i)$ is the time that the i th image is sampled, and $\mathbf{T}_s(j)$ and $\mathbf{T}_s(j+1)$ are the starting and ending points of the j th physiological cycle, respectively, satisfying $\mathbf{T}_s(j) < \mathbf{T}(i) < \mathbf{T}_s(j+1)$. The starting and ending points were defined by the peaks of the identified physiological silhouette signals. From the ordered data, the effects of physiological motion are estimated by fitting to a Fourier series:

$$F[\theta_i] = A_0 + \sum_{n=1}^2 A_n \cos(2n\pi\theta_i) + \sum_{n=1}^2 B_n \sin(2n\pi\theta_i) \quad (2.5.2)$$

where A_0, A_n and B_n are coefficients to be fitted.

After these steps, the independent components that represent the physiological fluctuations could be detected by correlating their time-series with the time-course derived from Eq.(2.5.2).

2.5.2 Experiment and Results

To evaluate the performance of this artifact reduction technique in the presence of functional activation, a synthetic fMRI dataset was simulated by adding a convolved box-car time series of truth activation to a series of resting-state scans. The scanner parameters were TE=40ms, TR=250ms and matrix=64x64, and 1024 time steps were extracted. The functional responses were block designs with stimulating/resting times of 1s. The simulated responses were superimposed on voxels chosen from the visual cortex. After processing by our artifact elimination approach, the correlations with the original HRF were calculated and compared to evaluate the effectiveness.

Figure 2.9: Histogram of the dominant frequency components: (a) within the respiratory frequency range (b) within the cardiac frequency range

To find the two dominant frequencies that were supposed due to respiratory and cardiac motions, two histograms of the dominant frequency components were plotted within the two frequency ranges, respectively. As shown in Fig.(2.9a), within the respiratory frequency range from 0.1 to 0.5 Hz, it is apparent that most pixels show a strong frequency component near 0.30 Hz; within the cardiac frequency range from

Figure 2.10: Estimated waveforms: (a) respiratory signal (b) cardiac signal

Figure 2.11: The power spectrums of estimated waveforms: (a) respiratory signal (b) cardiac signal

Figure 2.12: Fourier fitting of (a) respiratory motion (b) cardiac motion

Figure 2.13: The mapping of activation and their comparison: (a) before (b) after physiological artifact reduction (c) the ROC curves

0.75 to 1.5 Hz, most pixels contain a frequency component of 1.11 Hz (Fig.(2.9a)).

Fig(2.10) shows the estimated respiratory and cardiac waveforms obtained from pixels that contain the physiological fluctuation, whose time-series were then filtered by a pass-band filter within a width of 0.1 Hz, centered at the dominant physiological frequency. The power spectrums of the estimated respiratory/cardiac signals are shown in Fig.(2.11), which confirms that the estimated physiological signals are very close to the actual physiological cycles.

Fig(2.12) shows the result of Fourier fitting after re-ordering the estimated physiological time courses into the normalized cardiac/respiratory phases, respectively.

After obtaining the fitted Fourier series, SICA was then applied to decompose the fMRI scans into a number of spatially independent components. The component time-series corresponding to physiological artifacts were detected by a highest correlation with the Fourier phase model, fitted in the previous step.

Fig.(2.13a) and Fig.(2.13b) give the activation maps of synthetic data. Some tiny blobs which represented activations over the brain image were excluded after removing the physiological artifacts, thus reducing the value of false positive rate in ROC analysis. The ROC curves of both datasets are plotted in Fig(2.13c) by adjusting different thresholds of correlation, which further illustrates improvement in accuracy of detection after our artifact reduction technique.

2.6 Conclusion

In the unified framework depicted in Fig.(2.1), a demixing process of ICA is performed to decompose the original data into independent source signals. After this

stage, using the detection methods illustrated in the preceding two sections, undesirable source signals corresponding to random noises and physiological fluctuations can be extracted. By setting their switches to "off", as formulated in Eq.(2.3.4), and reconstructing other independent components back into the mixed signals, the removal of both noises and physiological fluctuations could be achieved under this model.

Our results demonstrate that ICA can extract task-related, nontask-related, and artifact components with little a priori knowledge of their temporal and spatial structure. This property of the ICA algorithm warrants its description as providing "blind separation" of the data into spatially independent components.

The outcomes of our experiments lend support to the assumption that the result of some activation detection techniques, such as correlation analysis and SPM, could be improved with a high correlation coefficient and good localization, after a restoration of the fMRI dataset using ICA. This study confirmed the feasibility of ICA for eliminating non-task related sources and noises. It is shown that this approach with ICA is a valid preprocessing technique in fMRI analysis especially detecting brain regions that are activated by input stimuli.

Although ICA is capable of "blind separation" into independent components, the subsequent interpretation of the separated components requires additional knowledge on the part of the experimenter. In our current trials, which utilized some statistical criteria of the components, the target component can be easily found by comparing its time series to a reference function, or calculating a statistic of its spatial distribution.

However, the proposed model in our work is still a heuristic approach to the objective restoration of functional MR images. More endeavor should be placed on collecting further information to acquire a clearer and more comprehensive view of

the various artifacts in fMRI to enrich and extend the model. This would indeed provide a new effective method to view into the underpinning of fMRI signals.

Chapter 3

Higher-Order Functional Brain Connectivity

This chapter first introduces the main divergences of connectivity in brain studies. Then we explain the established definition of second-order functional connectivity and its analysis method, following which a new concept of higher-order functional connectivity is proposed and defined. Then we show how ICA can be used to analyze higher-order connectivity of the human brain. After that, the results of these concepts and techniques on the fMR images obtained in resting brain experiments will be presented and discussed.

3.1 Introduction

Connectivity approaches to understanding the integration mechanism of brain function have gained an important growth over the past decade. There is mounting evidence that dynamical patterns generated by brain networks underlie all of cognition and perception [49][50][51]. Disruptions of the wiring of these networks may result in severe and specific alternations of mental and perceptual function.

In the quite beginning, the concept of functional connectivity was elaborated in the analysis of multiunit recordings of separable neuronal spike trains, recorded simultaneously from different brain areas [52][53]. Temporal coherence among the activity of different neurons was commonly measured by cross-correlating their spike trains. The resulting correlograms were then interpreted as the signature of functional connectivity.

The distributed brain systems associated with performance of a verbal fluency task were first identified in a non-directed correlational analysis of neurophysiological data obtained with PET [3]. PET is in a unique position to acquire data for this sort of analysis because it samples the entire brain state in a uniform fashion. This allows all possible functional connections to be assessed using serial measurements of the same subject in different brain states. But the shortcoming of PET include its relatively poor spatiotemporal resolution and the exact nature of the dependency of measured regional CBF (rCBF) on neural discharge rates.

There was a close relationship between effective connectivity and efficacy: "It is useful to describe the effective connectivity with a connectivity matrix of effective synaptic weights. Matrix elements would represent the effective influence by neuron i on neuron j [53]." It was also been proposed that "the notion of effective connectivity should be understood as the experiment and time-dependent, simplest possible circuit diagram that would replicate the observed timing relationships between the recorded neurons [54]." These definitions were essential and useful abstractions but lacked operational significance. In Friston's work [3], he reserved the term functional connectivity to mean the observed temporal correlation between two electro/neurophysiological measurements from different parts of the brain. Effective connectivity referred to the

underlying efficacy (W_{ij}). This analysis used a recursive Principal Component Analysis (PCA), which suggested that the variance in neurophysiological measurements could be accounted for by two principal components.

Other than using PET, Biswal first employed fMRI data in resting human brain for functional connectivity analysis in the motor cortex [4]. During the resting state acquisitions, the subjects were instructed to refrain from any cognitive, language, or motor tasks as much as possible. The experiment was designed to discover whether time courses of low frequency in resting brain had a temporal correlation within regions associated with motor function. For mapping of motor function, a square wave reference with 20s periods of "task activation" assigned a value of 1 alternating with 20s periods of "rest" assigned a value of 0 was used. All pixels that passed a threshold of statistical significance value were identified as belonging to the finger motor cortex. Each of the resting brain time courses from the regions that had been identified by fMRI was used a reference to produce, from each resting brain study, approximately 60 correlation-coefficient images. Pixels in each of these images that passed the correlation-coefficient of 0.35, which corresponded to a significance value of $P < 10^{-3}$, were counted. They have discovered that low-frequency fluctuations in resting brain from regions of the primary sensory motor cortex that were associated with hand movement were strongly correlated both within and across hemispheres.

Similar to PCA approach, a method of computing eigenimages of fMR images obtained from resting data can also be attempted to get functional connectivity patterns of the brain [5]. The approach circumvents the intensive computation of most contemporary techniques of analyzing fMRI data, which compared time-series extracted from the brain voxels conjointly with one another until pixels having high mutual

correlations were identified [4][55]. As a multivariate approach, this technique does not require identification of the specific regions to determine the connectivity before the experiments. In summary, the above approaches all extracted information about neural interactions through decomposition of interregional covariances of activity.

Other than PCA and eigenimage analysis, there are also some novel but heuristic approaches in analyzing function connectivity. Replicator dynamics was proposed to detect close functional networks, different from standard clustering approach to analyzing fMRI data [56]. The novelty of this method is that these networks had the property that every network member was closely connected with every other member. The basic assumption is that during the course of an fMRI experiment, several brain regions are active and interact with each other and, thus, form a functionally coherent network. These networks can be detected by analyzing correlations between fMRI time courses. The method has the advantage of only using pairwise similarity measurements rather than an explicit measurement vector in each pixel, as the entities that are being processed are very high-dimensional vectors of time courses.

Another approach is the application of random field theory [57], which looks at the entire 6D matrix of correlations between all voxels and search for 6D local maxima, as opposed to correlation analysis between voxel measurements from either PET CBF or BOLD fMRI images in 3D. In stead of traditional linear PCA method, a nonlinear PCA has been presented that identifies underlying sources causing the expression of spatial modes or patterns of activity in neuroimaging time series where these sources can interact to produce second-order modes [58].

In comparison with studies of functional connectivity that utilize task manipulations, the analysis of correlations in steady state (or resting) data is less susceptible to

confounds arising when functionally unrelated brain regions respond in similar way to changes in task. Thus, many experiments have investigated functional connectivity via correlation analysis between regional signals recorded in fMRI data obtained in a steady state [59][60][61].

However, in all of the above work, the functional connectivity was measured by the second-order statistic of correlation coefficient. In this work, we propose a ICA-based technique that measures the functional connectivity in the complete statistical sense. Previously, ICA technique has been demonstrated applicable and efficient in detecting activated brain regions from task/control experiments [9] [29]. Here, the applicability of ICA in analyzing functional connectivity is attempted. We do not aim to quantify interactions among brain regions, but to automatically detect patterns of functionally connected brain regions without any a priori information on the anatomical configuration. Besides, this technique is capable of eliminating other interferences in the fMRI which would affect the magnitude of correlation in previous second-order functional connectivity. In our study, the technique was tried on fMRI data from resting brain only.

3.2 Connectivity in the Brain

Formally, the neural system is unique in that it is composed of numerous interconnected elements ranging from single neurons to entire ensembles. The neural connections range from local intra-regional connections among neurons, to interregional connections among ensembles of neurons across brain areas. Within small and localized region of the brain, neurons form characteristic sets of connections, so-called local circuits. These communication between nerve cells is carried out along physical

connections, often linking cells that are separated by large distances. Signals within these connections consist of series of action potentials (spikes) of unit magnitude and duration. The arrival of an action potential at a synaptic junction triggers numerous biochemical and biophysical processes, ultimately resulting in transmission of electrical signals to the postsynaptic (receiving) cell. Neurons in the cerebral cortex maintain thousands of input and output connections with other neurons, forming a dense network of connectivity spanning the entire thalamocortical system.

Brain networks are not random, but form highly specific patterns. A predominant feature of brain networks is that neurons tend to connect predominantly with other neurons in local groups. At this high level of scale, connection patterns formed by these local, intra-areal networks are thought to be responsible for the specific processing requirements of each area. Considering the entire brain, the large-scale organization of the cortex is characterized by patterns of interconnections linking brain areas within and between specific sensory and motor systems.

It is worth noting that communications between and along neural elements underlie the brain function. It was suggested by McIntosh [6] that the brain function is the result of changes in the covariance among neural elements. Looking from higher-order criteria, we attempt to demonstrate the contribution of a collection of synchronized brain units to understanding brain activities beyond second-order sense.

As we are going to develop a definition in brain connectivity, it is useful at this point to introduce and clarify some conventional and latest terminology and investigate into their relations. Hitherto, basically three kinds of connectivities have drawn the attentions of scientists and researchers in brain studies: anatomical, effective and functional connectivity. In summary, the definitions are given as [3][62]:

Definition 3.2.1. *Anatomical connectivity* is determined by the set of physical or structural connections linking neuronal units at a given time, which gives the simple neuroanatomical organization of clusters of cortical areas in the brain.

Definition 3.2.2. *Functional connectivity* is defined as the temporal correlation between pairs of neurophysiological (functional) measurements that exists between distinct neuronal units.

Definition 3.2.3. *Effective connectivity* refers to the influence of one brain region on another region through direct and indirect anatomical model.

Although there is no simple one-to-one mapping among these three kinds of connectivity, they are yet interdependent in some ways. Anatomical connectivity is a necessary underpinning for structural models of extrinsic cortical connections and also has been used to infer functional connectivity [63] and effective connectivity [6]. Analysis of intra-areal patterns of connections would involve "connection bundles" or "synaptic patches" linking local neuronal populations. Analysis of large-scale connection patterns would focus on connection pathways linking segregated areas of the brain. In both analysis, a choice has to be made on the level of the spatial scale at which the analysis is to be performed.

Assigning numerical weights to the connections to anatomical model leads to the functional connectivity. In functional connectivity, the temporal correlations are often the result of neuronal interactions along anatomical or structural connections; in some cases, observed correlations may be due to common input from an external neuronal or stimulus source. It is easy to see that anatomical connectivity is a major constraint on the kinds of patterns of functional connectivity that can be generated. In the other

direction, functional connectivity can contribute to the shaping of the underlying anatomical substrate.

Effective connectivity gives an accurate quantification to describe the connectivity between brain regions, which could be further divided into direct and indirect effects through the anatomical model [6]. In some aspects, the functional model is close to the notion of effective connectivity since they both measure the dependence between brain regions. But effective connectivity deals with more a cause-to-effect influence. To this extent, effective connectivity is much closer to achieving the goal of investigating and modelling the interactions of neural networks. Thus, the interdependence and reciprocity among these three connectivity concepts deserves emphasis as it captures some of the unique aspects of brain networks.

3.3 Second-Order Functional Connectivity and PCA

As illustrated by Friston [3], the pattern of functional connectivity are represented by uncorrelated brain systems within which there are significant temporal correlations between different brain regions (or voxels). The brain units denoted as regions or voxels within the same system are highly correlated by their temporal variations and different brain systems are largely uncorrelated. Based on the pairwise correlation coefficients, this can be referred to as a second-order connectivity.

Some previous work has reported PCA or eigenimage analysis offers a good technique to analyze this sort of functional connectivity [5], which decomposed the fMR images into a series of orthogonal patterns that embodied in a decreasing manner the amounts of the functional connectivity. PCA extracts the important features of

the correlation matrix in terms of principal components or eigenvectors. These vectors are the linear combinations that account for orthogonal amounts of variance in the observed data. Only a few principal components are usually required to explain the majority of observed variance. In terms of functional connectivity, a principal component represents a truly distributed brain system within with there are high inter-correlations. Furthermore, because any one component is orthogonal to the remaining, these systems are functionally unconnected from each other.

The functional MR image can be represented in matrix format as $\mathbf{X} = \{x_{ij}\}_{m \times n}$ where x_{ij} is the intensity of the j th voxel of the i th scan of the fMR image, m denotes the total number of scans and n is number of voxels in an image scan. Then the time-series obtained at the i th voxel is denoted by vector $\mathbf{x}_i = (x_{1j}, x_{2j}, \dots, x_{mj})^T$. Let the average intensity of j th voxel be $\bar{x}_j = \frac{1}{m} \sum_{i=1}^m x_{ij}$ and the mean corrected scan denoted by $\tilde{\mathbf{x}}_j$. Then the mean corrected functional image is given by $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_n]^T$.

Consider the covariance matrix $\mathbf{A}$ of functional data:

$$\mathbf{A} = \tilde{\mathbf{X}}^T \tilde{\mathbf{X}}$$

where $\mathbf{A} = \{a_{ij}\}_{n \times n}$ and $a_{ij} = \mathbf{x}_i^T \mathbf{x}_j$ which indicates the covariance between the i th and j th voxel. For simplicity let's normalize the covariance matrix and obtain the correlation matrix $\mathbf{R} = \{r_{ij}\}_{n \times n}$ where

$$r_{ij} = \frac{a_{ij}}{\sqrt{a_{ii}a_{jj}}} \quad (3.3.2)$$

From the well-known Karhunen-Loeve transform (KLT), $\mathbf{R}$ can be expanded as [64]:

$$\mathbf{R} = \sum_{i=1}^n \lambda_i \mathbf{e}_i \mathbf{e}_i^T \quad (3.3.3)$$

where λ_i and $\mathbf{e}_i = (e_{i1}, e_{i2}, \dots, e_{in})^T$ are the i th eigenvalue and eigenvector of $\mathbf{R}$ respectively.

Theorem 3.3.1. In the sense of second-order statistics, the functional connectivity is represented by a series eigenimages $\mathbf{e}_i$ corresponding to relatively large eigenvalues λ_i .

Proof. Let us arrange $\lambda_1 > \lambda_2 > \dots > \lambda_n$. If λ_1 is significantly large compared to $\lambda_2, \dots, \lambda_n$, then matrix $\mathbf{R}$ can be written as

$$\mathbf{R} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \delta \quad (3.3.4)$$

where δ is the error matrix with elements having negligible values. Therefore $\lambda_1 \mathbf{e}_1 \mathbf{e}_1^T$ matrix represents almost the matrix $\mathbf{R}$ and

$$r_{ij} \approx \lambda_1 e_{1i} e_{1j} \quad (3.3.5)$$

If e_{1i} and e_{1j} are large, then the r_{ij} is also large, i.e., the correlation coefficient of i th and j th voxel is high. Thus according to the definition of functional connectivity, it can be concluded that the i th voxel and the j th voxel have a high functional connectivity. So the voxels corresponding to significant intensities in the first eigenimage represent a map of functional connectivity.

Let consider a situation where the q largest eigenvalues are very significant in contrast to the remainder. The correlation matrix can be expanded as follows:

$$\mathbf{R} = \sum_{i=1}^q \lambda_i \mathbf{e}_i \mathbf{e}_i^T + \delta \quad (3.3.6)$$

where δ is a matrix with elements having negligibly small values. In this case,

$$r_{ij} \approx \sum_{k=1}^q \lambda_k e_{ki} e_{kj} \quad (3.3.7)$$

Some previous works have reported utilizing PCA and eigenimage technique to analyze functional connectivity from PET [3] and MRI dataset [5]. Both employed the presumption that the largest q eigenvectors corresponding to relatively high eigenvalues would represent the functional connectivity. Nevertheless, no strict mathematical derivations have been made for it.

Actually, this presumption really holds for fMRI dataset in that from previous studies [3] [5], those uncorrelated brain systems almost do not have overlapping voxels in common. In this case, if the i th and j th voxel of the p th eigenimage have high values, then according to Eq.(3.3.7) and the above notion of non-overlapping of eigenimages, we have $e_{ki} \approx 0$, $e_{kj} \approx 0$ for all k that $1 < k < q$, $k \neq p$, and thus $r_{ij} \approx \lambda_p e_{pi} e_{pj}$ also has high value. Therefore, these two voxels can be said to show functional connectivity. ■

The eigenimage analysis or PCA only provides an approximate pattern of functional connectivity measured by a relatively high value of second-order correlation. In fMRI data, besides the neural sources, there exist other confounding artifacts and noises. The tiny signal change coming from neural activity may be vulnerable to various kinds of artifacts and noises. Since the correlation is measured on the fMRI dataset which is mixed by different source fluctuation, the high correlation value may arise from a connectivity of interfering signals other than the neural sources. A typical interference is the physiological fluctuation due to respiratory and cardiac motion.

Another motivation for further development is to extend the correlation from second order to higher order. For any two voxels i, j , let's denote their time variation by variables $\mathbf{v}_i$ and $\mathbf{v}_j$. The second-order correlation coefficient between v_i and v_j is

defined as:

$$c_1(i, j) = \frac{\tilde{\mathbf{v}}_i^T \tilde{\mathbf{v}}_j}{\sqrt{\tilde{\mathbf{v}}_i^T \tilde{\mathbf{v}}_i} \sqrt{\tilde{\mathbf{v}}_j^T \tilde{\mathbf{v}}_j}} \quad (3.3.8)$$

where $\sim$ implies mean-corrected. Here we aim to extend it to a higher-order correlation $c_q(i, j)$.

Definition 3.3.1. *The higher-order correlation between voxel i and j is defined as:*

$$c_q(i, j) = \frac{\tilde{\mathbf{v}}_i^{qT} \tilde{\mathbf{v}}_j^q}{\sqrt{\tilde{\mathbf{v}}_i^{qT} \tilde{\mathbf{v}}_i^q} \sqrt{\tilde{\mathbf{v}}_j^{qT} \tilde{\mathbf{v}}_j^q}}$$

Here, $c_q(i, j)$ is also called $2q$ -order correlation.

In Definition (3.2.2), we say voxel i and j are functionally connected if $c_1(i, j)$ is significantly large between the interval $[0, 1]$. But it does not suffice a stricter requirement of significant correlation value for any higher order q .

3.4 ICA and Higher-Order Functional Connectivity

Independent Component Analysis (ICA) is a method for finding underlying factors or components from multivariate (multidimensional) statistical data. The mathematics of this technique is elaborated in the preceding chapter.

It has been suggested that the separate processes underlying fMRI datasets may be represented by one or more spatially-independent components, and associated with a single time course of enhancement or suppression. The problem can be expressed rigorously by writing a linear equation relating the components maps and their time

courses to the measured fMRI signals [9]:

$$\mathbf{X} = \mathbf{M}\mathbf{C} \quad (3.4.1)$$

where $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_k]^T$ is the matrix of component maps and $\mathbf{M} = [\mathbf{m}_1, \dots, \mathbf{m}_k]$ the mixing matrix. Thus the voxel values for each of the component maps are placed in separate rows of matrix $\mathbf{C}$, and each column of $\mathbf{M}$ represents a time series of a corresponding component. As a convention, this separation process of ICA is usually termed spatial ICA (SICA).

Previous experiments have demonstrated the feasibility of ICA in extracting task-related component and discarding other artifacts and noises [9] [29], which provides us with a picture of brain voxels and regions that are connected by the same neural source without any interference from artifacts. From this perspective, we could further develop the concept of functional connectivity as:

Definition 3.4.1. *Higher-order functional connectivity* is defined as brain systems within which the different voxels or regions are activated by the same dependency of temporal variation in the complete statistical sense.

Unlike the definition of second-order, which measures the interdependence between brain units(e.g., voxels or functional regions), the connectivity in Definition(3.4.1) is more a dependence between a brain unit and a certain brain system. Here the brain system is characterized by a definite time course.

Theorem 3.4.1. For a certain time course $\mathbf{m}_p$ related to neural activation, the functional connectivity is represented by a set of voxels $\{v_i | |c_{pi}| > \alpha, i = 1, 2, \dots, m\}$ where $\mathbf{c}_p = (c_{p1}, c_{p2}, \dots, c_{pn})^T$ is a vector of component map from SICA and c_{pi} is the intensity value of voxel v_i in the component map $\mathbf{c}_p$. α is a preselected threshold

value, which determines the significance of connectivity from a given P -value, and $|\cdot|$ stands for absolute value.

Proof. Choose the p th component $\mathbf{c}_p = (c_{p1}, c_{p2}, \dots, c_{pn})^T$, and its corresponding time series $\mathbf{m}_p = (m_{1p}, m_{2p}, \dots, m_{mp})^T$. The component map is described by a distribution of values c_{pi} , one for each voxel. These values represent the relative amount a given voxel is modulated by the activation of the source signal. Under the null hypothesis of no activation, the connectivity maps have a Gaussian distribution of intensity values [5]. To find and display voxels contributing significantly to a particular component map, the map values should be scaled to z-scores (the number of standard deviations from the map mean). Voxels whose absolute z-score are greater than some threshold of significance P can be considered to be "connected" voxels for that component. The threshold of intensity value, α , can thereby be determined from the z-score threshold. Negative z-scores indicate voxels whose fMRI signals are modulated opposite to the time course of activation for that component. In accord with the Definition(3.4.1), all these connected voxels in the component map constitute a brain system of *higher-order functional connectivity* at a significance level of P . $\square$

Theorem.(3.4.1) gives a rigorous expression of *higher-order functional connectivity* in a mathematical form and shows how SICA could be applied to finding the higher-order functional connectivity.

As aforementioned in the preceding section, there are two advantages that feature the newly defined functional connectivity over Definition(3.2.2). The first is its accuracy in the measurement that are unaffected by any confounding factors which has been a prominent concern in the previous work of fMRI data analysis. This feature highlights the concept "functional" because the connectivity is termed in the light of

neural sources and brain "function" is right believed to be a result of integration of neuronal activations in brain units.

Second, this connectivity is investigated in a complete statistical sense, i.e., the brain unites are connected in any higher order. Let's reveal its stronger property of dependence than previous definition quantitatively by the statistic defined in Definition.(3.3.1).

Theorem 3.4.2. Any two voxels i, j having significantly large intensity values in the same component map satisfies: $c_q(i, j) = 1$ for any $q > 0$.

Proof. For any two voxels i, j having large intensities in the k th component, the time series of activation k on these two voxels can be denoted by $\mathbf{v}_i = c_{ki}\mathbf{m}_k = (c_{ki}m_{1k}, c_{ki}m_{2k}, \dots, c_{ki}m_{mk})^T$ and $\mathbf{v}_j = c_{kj}\mathbf{m}_k = (c_{kj}m_{1k}, c_{kj}m_{2k}, \dots, c_{kj}m_{mk})^T$. For any q , denote

$$\mathbf{v}_i^q = ((c_{ki}m_{1k})^q, (c_{ki}m_{2k})^q, \dots, (c_{ki}m_{mk})^q)^T$$

$$\mathbf{v}_j^q = ((c_{kj}m_{1k})^q, (c_{kj}m_{2k})^q, \dots, (c_{kj}m_{mk})^q)^T$$

The mean values of the variables are:

$$E\{v_i^q\} = c_{ki}^q \frac{1}{m} \sum_{t=1}^m m_{tk}^q \quad (3.4.2)$$

$$E\{v_j^q\} = c_{kj}^q \frac{1}{m} \sum_{t=1}^m m_{tk}^q \quad (3.4.3)$$

Let us denote

then

$$E\{v_i^q\} = c_{ki}^q \mu \quad (3.4.4)$$

$$E\{v_j^q\} = c_{kj}^q \mu \quad (3.4.5)$$

So the average-corrected vector:

$$\begin{aligned}\tilde{\mathbf{v}}_i^q &= (c_{ki}^q(m_{1k}^q - \mu), c_{ki}^q(m_{2k}^q - \mu), \dots, c_{ki}^q(m_{mk}^q - \mu))^T \\ &= c_{ki}^q \tilde{\mathbf{m}}_k \\ \tilde{\mathbf{v}}_j^q &= (c_{kj}^q(m_{1k}^q - \mu), c_{kj}^q(m_{2k}^q - \mu), \dots, c_{kj}^q(m_{mk}^q - \mu))^T \\ &= c_{kj}^q \tilde{\mathbf{m}}_k\end{aligned}$$

where $\tilde{\mathbf{m}}_k = (m_{1k}^q - \mu, m_{2k}^q - \mu, \dots, m_{mk}^q - \mu)^T$. Then the standardized correlation of v_i^q, v_j^q is:

$$\begin{aligned}c_q(i, j) &= \frac{\tilde{\mathbf{v}}_i^{qT} \tilde{\mathbf{v}}_j^q}{\sqrt{\tilde{\mathbf{v}}_i^{qT} \tilde{\mathbf{v}}_i^q} \sqrt{\tilde{\mathbf{v}}_j^{qT} \tilde{\mathbf{v}}_j^q}} \\ &= \frac{c_{ki}^q \tilde{\mathbf{m}}_k^T \tilde{\mathbf{m}}_k c_{kj}^q}{c_{ki}^q \sqrt{\tilde{\mathbf{m}}_k^T \tilde{\mathbf{m}}_k} c_{kj}^q \sqrt{\tilde{\mathbf{m}}_k^T \tilde{\mathbf{m}}_k}} \\ &= 1\end{aligned}$$

In second-order correlation, voxel i and j are connected if $c_1(i, j)$ is large between the interval $[0, 1]$. However, Theorem.(3.4.2) demonstrates that the connectivity derived from the components of neural sources by ICA suffices for any order q , $c_q(i, j) = 1$, which attains the maximum value of correlation. Therefore, this connectivity satisfies much stronger requirements for the interdependence between two brain regions.

Furthermore, the approach of ICA to finding higher-order brain connectivity has the additional advantage that it does not require identification of the specific regions to determine the connectivity areas before the experiment and could provide a closely

correct or exact pattern of connectivity automatically, dependent on the degree of independence of the component maps.

3.5 Experiments and Results

Experiments were conducted for each subject in a Bruker Medspec S200 Advance 2T scanner with a slew rate 200T/ms/s at the Department of Diagnostic Radiology of the University of Freiburg, Germany, in July 1998 (data is not publicly available on the Internet). Volunteers were positioned in the scanner to best scan the position of the individual motor cortex of the brain. The images were acquired while the subject was resting in the scanner, where the subject were instructed to refrain from any cognitive, language, or imaging visual or motor task as much as possible. The axial slices of the brain approximately at the motor cortex were scanned using fast EPI sequence with repetition rates of $TR=250\text{ms}$. Two brain slices, one with $FOV=25.6\text{mm}$ at $+47.5\text{mm}$ from the AC-PC plane and other with $FOV=19.2\text{mm}$ at $+26\text{mm}$ from the AC-PC plane were acquired for the TR values. The scanner parameters were $TE=40\text{ms}$ and $\text{matrix}=64 \times 64$, and 1024 time steps were extracted. Ten datasets of the resting brain experiments acquired from healthy volunteers were analyzed. In what follows we present the results of our analysis on a representative dataset.

3.5.1 Eigenimage Analysis

As demonstrated by Biswal et al. [4], only the low frequency components of the brain responses are most likely to indicate functional connectivity. Therefore, the images were filtered by low-pass filtering, which removed the interferences mainly due to

cardiac and respiratory fluctuations [14] prior to computing the eigenimages.

Figure 3.1: (a) The first eigenimage with eigenvalue=573.5; (b) the second eigenimage with eigenvalue=19.0. The eigenimages were transformed into a z-map and thresholded at z-value=2.8 with a minimum blob size of 2.

It has been demonstrated that the voxel intensity distribution of the eigenimage, which is obtained under the null hypothesis (H_0) of no connectivity, is close to a normal distribution [5]. This provided us with a thresholding mechanism for eigenimages at a given z-value for the inferencing. The confidence of these areas was further enhanced by considering the fact that functional connectivity must occur as a group of voxels instead of isolated ones [65]. These significant areas were identified and classified depending on where these blobs were located in the brain.

Two principal eigenimages with the largest eigenvalues were obtained and superimposed on an anatomical image as shown in Fig.(3.1). The eigenimages were thresholded at a z-value of 2.8 with a minimum blob size set to 2. The eigenvalues corresponding to the eigenimages were 573.5 and 19.0, respectively. Only the first two eigenimages were capable of producing significant connectivity while the rest of the eigenimages did not account for any significant connectivity in the brain.

As we see in Fig.(3.1a), the voxels in the ipsi and contra lateral PMA (primary

motor area) in each hemisphere are detected in the resting state. SMA (supplementary motor area) are also clearly identified in the center gyrus, that can be seen on the midline of the image. Voxels posterior to SMA are identified as belonging to the paracentral lobule. This may be an extension of the primary sensorimotor cortex into the inner hemispheric fissure. Basically, these identified areas are matched to the findings by Biswal [4] and Jeong [59]. Some voxels that are not coincident with the motor cortex areas as defined by fMRI seem nevertheless to be a manifestation of functional connectivity.

Fig.(3.1b) shows the colorized voxels in the region of visual cortex in both hemispheres, which indicates the functional connectivity between the left and right visual cortices in the resting state. The same areas were also detected in the results of [59] and [5].

3.5.2 Connectivity Analysis with ICA

The images were then analyzed using SICA technique. After the ICA process, a large number of components were extracted due to the high amount of scans. In order to pick out the component maps which are most likely neurally-related, we based our identification of ICs on a fourth order statistic called *kurtosis* (k) [37]. The kurtosis value is zero if the variable has a Gaussian distribution. Since random noises are always supposed to bear a nearly Gaussian distribution and due to the specificity of spatial distributions of useful brain signals, the component maps of neural sources in fMRI usually have a kurtosis value far more than 0.

Besides, the ICs corresponding to physiological motions and other artifacts were

Figure 3.2: Voxel intensity distribution of a representative component map

Figure 3.3: Component maps of neural sources obtained from ICA: (a) map with $\kappa = 37.8$ (b) map with $\kappa = 24.5$ (c) map with $\kappa = 23.7$. The maps were transformed into a z-map and thresholded at $p\text{-value} \leq 0.02$ with a minimum blob size of 2.

further explored and removed using other identification techniques. Here the component maps corresponding to the ten largest kurtosis values were scaled to z-values and overlaid on the anatomical slice for visualization and investigation of connected regions.

The voxel intensities of a representative component map of no connectivity are plotted in Fig.(3.2). As seen in the figure, the voxel intensity distribution of the component map, which is obtained under the null hypothesis of no connectivity, is close to a normal distribution. This provides us with a thresholding mechanism at a given P -value and we use $P = 0.05$ significance for inferencing the connectivity.

The confidence of these areas is enhanced by considering the fact that functional connectivity must occur as a group of voxels instead of isolated ones. In order to find the significantly connected voxels, a blob size threshold of 2 were used for the independent component maps.

As some artifacts may not be removed by thresholding the kurtosis value, we relied on the findings of previous experiments of detecting functional connectivity in resting brain data [4][5][59] to select the components which were likely enough to be neural sources. Fig(3.3) presents two component maps that were chosen to show strong connectivities in motor system and visual system, respectively. In Fig.(3.3a) with $k = 37.8$, clear large blobs of brain voxels are detected in the left and right hemisphere which are assigned to motor cortex. Similar colored blobs are also found in SMA and even in premotor areas which are in the anterior central gyrus close to the outside. The apparent discrepancy with Fig.(3.1a) is that more voxels of Fig.(3.1a) lay outside the motor cortex boundary, although contiguous to it.

The component map in Fig.(3.3b) with a corresponding kurtosis value of 24.5 indicates functional connectivity in visual cortices. In the occipital lobe, these areas are visualized with a better localization and spatial specificity in contrast to Fig.(3.1b). Apparently, the imperfection of noisy voxels colorized in the eigenimages is further ameliorated in the component maps from ICA, which suggests that ICA is useful to find higher-order connectivity without being affected by interfering signals.

Fig(3.1c) presents a map of functional connectivity mainly between primary somatosensory areas that was not detected by eigenimages. The inability of eigenimage to find this connectivity may be caused by the small change of the time series in these areas relative to the interfering signals on these regions, because second-order

connectivity is evaluated by the correlation coefficient of the mixed intensities of two voxels. This finding demonstrates the strong capability of ICA to detect functional connectivities across cortical areas even though the intensity values are interfered with other uninteresting sources.

3.6 Conclusion and Discussion

In this chapter, we proposed a new definition of functional connectivity with the approach of ICA to finding the connectivities underlying fMR images obtained in resting brain experiments. The signal changes of resting brain fMRI are subtle and interfered with other artifacts and random noises. Although eigenimages with low eigenvalues are assumed as noises, PCA could not thoroughly analyze the connectivity without being affected by additive confounds, as the correlation between voxels is computed from the mixed value of intensities. However, ICA technique is capable of decomposing the fMR images into a number of independent source signals, which could purely be artifacts, noises or neural sources and hence the neural source components are noise-free. As the connectivity is totally determined by one source signal in a complete statistical sense, this analysis method is robust to any interferences.

Functional connectivity is often the result of neuronal interactions along anatomical or structural connections; in some cases it may be due to common input from an external neuronal or stimulus source. Thus, after captured in fMR images, this connectivity is more likely to be demonstrated with a same neural source signal or component. In this sense, we are confident to say that higher-order connectivity is more appropriate to describe the underlying connections among neuronal elements.

In cases of task-involved experiments, some connectivity may be due to the stimulus source designed for a particular task. Usually, if the component map has a time course which is significantly correlated with the time blocks of stimuli, this connectivity pattern is known as activation map [9]. However, other components obtained in ICA decomposition of fMR images may also be related to connected components. These connectivities are probably caused by some common modulation of psychological or hemodynamical mechanism, which is not evidently correlated with the stimulus source.

On the other hand, higher-order connectivity could not completely take the place of second-order connectivity. It is believed that second-order connectivity still takes the dominant role in analyzing brain interactions within a neural system. A neural system is usually composed of a number of brain regions, that are responsible for a certain brain function. In a certain neural system, most connected regions could be interpreted by a temporal correlation, as they are involved in the same brain function. However, it is not required and not realistic that the connected regions should have the same dependency of temporal variation. In this sense, higher-order connectivity is just a much stronger property than second-order connectivity.

The different eigenimages extracted from PCA represent uncorrelated neural systems. However, ICA aims to separate the fMR images based on a maximum independence criteria among different components, which is a stronger property than uncorrelatedness. Due to the spatial independence of different sources underlying fMRI datasets, ICA could perform a better separation of these signals and give a pattern of functional connectivity closer to the sources than PCA.

Conventionally, ICA has been mainly exploited in studies of stimulated brain analysis, both on fMRI data [9] and EEG signals [30], and different algorithms of ICA been attempted. In our work, ICA was first used to analyze resting state brain to obtain patterns of higher-order connectivity. Up to date, spatial ICA has dominated fMRI analysis. The neural sources in our experiment were independent of each other in their spatial distributions, while eigenimages were temporally analyzed. Some experiments have provided that temporal ICA and spatial ICA would converge to same results if the source signals are both spatially and temporally independent [26]. Much further effort may be needed to look into the applicability of TICA to brain connectivity analysis and investigate the connectivity in task-involved brain, to discover more higher-order connectivities around the brain.

Chapter 4

Exploratory Approach to Modeling Neuronal Interactions with SEM

In this chapter, a data-driven structural equation modeling approach is proposed. The detailed description of structural equation modeling is provided after introduction. The exploratory detection method of SICA, combined with an automated best model search is explained after that. This is followed by an experiment on the visual-task fMRI data. We conclude this chapter by suggesting that this exploratory method will give a more dynamic model for the interactions among functional brain network, especially in the event of brain abnormalities.

4.1 Introduction

It is worth noting that communications between neural elements underlie brain function. It follows from this that a change in the observed activity of any element results from a change in the communication with one or more connected elements. In other words, an activity change at any central nervous system site must come about through a change in the influences of one or more afferent pathways.

The analysis of these interactions among brain regions extract information through decomposition of interregional covariances of activity. One of these methods is Structural Equation Modeling (SEM) or path analysis. It has proved to be a powerful way to combine functional neuroimaging data with anatomical circuitry to determine the functional neuroanatomy underlying a particular task. SEM has been applied to 2-DG and FDG data obtained from rats in different behavioral paradigms [66] and human brain imaging data obtained from PET measures of regional cerebral blood flow (rCBF) [67][68].

Later, the application of SEM to functional brain imaging data was explained [6]. This application made use of anatomically-based model to define a network and express the interactions among brain regions. The effects of an incomplete anatomical model, and the omission of regions that had an impact on other regions included in the model was examined using simulated dataset. It was suggested for the first time that covariance-based methods of SEM would provide a more realistic picture of brain operations in the context of functional interactions.

More recently, an automated search for the best fitting model that can be found to account for the data has been proposed [13]. This work is concerned with evaluating goodness-of-fit of a path analytic model to an interregional correlation matrix derived from fMRI data. The algorithm starts from the null model, in which all path coefficients are zero, and iteratively unconstrains the coefficient which has the largest Lagrangian multiplier at each step until a model is identified which has maximum goodness by a parsimonious fit index.

It seems much more natural, given the current state of biophysical knowledge data of fMRI, to adopt an exploratory rather than confirmatory approach, because to

affirm or refute a previously known model is not enough to allow for the conditions of brain disorders. Although a data-driven approach of searching a best fitting model is available [13], the logical basis of this approach is that the system under investigation is already well understood by the analyst, who can strongly predict a certain set of brain regions involved in this system. A predicament also lingers that he or she will have only incomplete knowledge about the system and its element. The predefined model may be invalid if a lesion occurs around the cortices of brain.

In this work, we develop a more exploratory approach to path analysis of fMRI data, which may circumvent some of these problems. It uses SICA to find the brain regions that are functionally involved in a neural network for a particular brain operation. The automated search, proposed by Bullmore [13], which aims to a parsimonious fit index of both the discrepancy between observed and modelled data and its cost, is exploited to find the best preferable path model, together with its fitting path coefficients. This method is then experimented on fMRI data acquired from subjects performing a periodically designed visual-stimulation task.

4.2 Structural Equation Modeling

The principles underlying decomposition of effects are both the major strength of path analytic approaches and essential for understanding the class of approaches called structural equation modeling (SEM) [69]. For any model, the relationships between variables can be decomposed into causal effects and noncausal relationships by using the logic introduced by path analysis. Furthermore, within causal and non-causal, the effects can be broken down even more. For causal effects, there are effects

Figure 4.1: Decomposition of effects

that go directly from one variable to a second variable (**direct effects**) and effects between two variables that are mediated by one or more intervening variables (**indirect effects**). For noncausal relationships, there are relationships between two variables that occur (a) because both are caused by a third variable (these are referred to as noncausal reflecting common causes or **noncausal due to shared antecedents**) and (b) because in models with more than one independent variable there can be relationships among them in which cause and effect are not theoretically articulated (often called **unanalyzed prior associations**). If all independent variables in a model are unrelated to one another, then there is no variability of this type. Fig(4.2) illustrates the different types of associations.

The traditional form of data analysis in structural equation modeling has been to look at the interregional covariances [6]. Expressed in terms of neural systems, a measure of covariance represents the degree to which the activities of two regions are related to each other, or how they vary together. A high covariance between areas A and B means that if area A increases its activity, so probably will B (in the case of a positive covariance).

The relationship between brain areas can be described using a simple linear mathematical expression of the variance of a region as influenced by the variance of another. The equation for this is:

$$Y = \alpha + \beta_{y \cdot x} X + \psi \quad (4.2.1)$$

where Y and X are measures of the activity in two interconnected brain regions across a sample. The equation also contains the slope of the line $\beta_{y \cdot x}$, indicating the size of the influence of X on Y. The final term ψ is the residual representing the variance in Y that is not determined by X. The value of $\beta_{y \cdot x}$ can be computed directly from the covariance of Y and X, and when the measures of Y and X are standardized (transformed to z-scores), the intercept α becomes zero.

When more dependent regions are included and the effects from multiple influences on regional activity are derived, the general linear model can be further expanded to define the interconnections of an entire network, expressed in matrices:

$$\begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ w & 0 & 0 & 0 \\ v & y & 0 & 0 \\ x & z & 0 & 0 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} + \begin{bmatrix} \psi_A \\ \psi_B \\ \psi_C \\ \psi_D \end{bmatrix} \quad (4.2.2)$$

In this equation, the variances of all regions A,B,C,D, are represented as a vector that is determined by the weighted influence of the other regions plus the residual

influences. The zero values in the weight matrix represent connections that do not exist in the model.

There are many commercially available computer packages that are specifically designed for structural equation modeling including LISREL [70] and EQS [71]. Fig.(4.2) illustrates processes and features of structural equation models defined by Eq.(4.2.2). The system, made up of four variables, has a causal structure indicated by the arrows. The regions and connections define the anatomical model. By using this anatomical model, the correlation matrix can be decomposed to assign the path coefficients to each of the arrows, given by letters v-z. The path equations in Fig.(4.2c) and structural equations in Fig.(4.2d) are mathematically equivalent.

In such model, the variables with arrowheads pointing to are often called dependent or **endogenous variables**. B,C,D in Fig.(4.2A) are endogenous variables. Variables with no causal arrows pointing toward them are called independent or **exogenous variables**. Variable A in Fig.(4.2A) is exogenous variable.

As shown in Fig.(4.2D), structural equation modeling also allows for influences not measured or not measurable to be incorporated in the model as residuals. Residual influences can be represented in at least two forms. One representation of residuals is best thought of as including the combined influences of regions outside the model and the influence of a brain region upon itself [66][72]. Brain regions that do not have a residual influence imply that all the variance in that region is accounted for by the connections with other regions in the model. Another representation of residuals is as a variable having a direct path to one or more of the regions within the model, which conceptually represent a region that has a strong influence on areas within the model, but could not be included as an influence that is exogenous to the model.

Figure 4.2: Schematic representation of methods involved in structural equation modeling of a neural system: (a) Path diagram of a simple network with four regions; (b) The information about the correlations of activities; (c) Path equations showing how the correlations between regions can be decomposed to solve for the path coefficients; (d) Structural equations showing the variance in activity of each region as a function of the weighted variance of other brain regions and a residual influence.

The technical definition of a path coefficient is the direct proportional functional influence one region has on another through their direct anatomical connection, with all other regions in the model left unchanged. This is the expected change in the activity of one region given a unit change in the region influencing it. If covariances between neural elements underlie brain operation, then path coefficient may be interpreted as an indication of whether there are task-related differences in functional influences within the same anatomical pathways when evaluated across tasks.

4.3 Confirmatory and Exploratory Modeling Approaches

Confirmatory approaches examine whether or not existing data are consistent with a highly constrained a priori structure that meets conditions of model identification. In fact, a model never can be confirmed. It just can be disconfirmed (it does not fit the observed data), or it can fail to be disconfirmed (it fits). The most important points for confirmatory SEM in recent experiments [6], are that these approaches begin with a theoretical model that has to be identified and must attempt to see whether or not data are consistent with that theoretical model.

For example, as mentioned in McIntosh's experiment of applying SEM to brain network analysis [6], anatomical constrained are usually employed in the modeling approaches. Six regions were simulated from a unit normal distribution and specific values for the path coefficients and residuals were given by a base model. Three further simulations were conducted to determine the effect of model modifications, which included adding or eliminating a path and even omission of a region. All

the effects were assessed by examination of changes in path coefficients from the base model and by evaluation of modification indices. Although the modeling steps are not totally restricted to the base model, it is completely possible that different networks will be derived if base model differs or different anatomical constraints are imposed.

In the central nervous system, there are numerous ways one area can have a functional impact on another. Many neural systems have a parallel anatomical organization by which connections between areas can be both direct and indirect. Effects decomposition of anatomically based structural equation models allows for the evaluation of whether the influence of one region on another is through a direct effect or is mediated through one or more indirect routes. Moreover, when evaluating experimental differences, it is entirely possible for the total effects of two regions to be the same while direct and indirect effects differ, or for total effects to differ when direct effect do not.

Therefore, in this sense, an exploratory approach is superior in being not so restrictive to a theoretical model, especially to the predefined paths and their coefficients. Bullmore introduced an alternative approach based on an algorithm for automatic identification of the best fitting model that can be found to account for the data [13]. Five main regions of generic activation, consistent with the results of previous research, were studied. However, in brain studies, a more realistic account of the fMRI data analysts' predicament is that we have only incomplete knowledge about the system. Besides, does it unquestionably follow that a model derived from a normal brain could provide the correct basis for the abnormality study? The answer, in our view, is no. In our work, some extensions are given to this exploratory approach to more automatically define the brain regions involved, in order to make a small but

Figure 4.3: Conceptual model of exploratory modeling of effective connectivity from fMRI

really contributive step in modeling effective connectivity.

4.4 Methods

As depicted in Fig.(4.3), our work consists of two steps: activation region detection and path model fitting. At the first step, we mainly adopt ICA (Independent Component Analysis) to automatically find the brain regions that are involved in a certain task. After the activation maps have been obtained, an automated search for the best fitting model is applied to the activated areas, in order to find the best neural network with fitted path coefficients.

4.4.1 Activation Detection with ICA

Denote by $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_m)^T$ the matrix of fMRI data, where $\mathbf{x}_i = (x_{i1}, \dots, x_{in})$ is an intensity brain map scanned in each time and x_{ij} is the measured j^{th} voxel intensity value in the i^{th} scan. SICA embodies the assumption that $\mathbf{X}$ can be expressed rigorously by writing a linear equation relating the component maps and their time

courses to the measured fMRI signals [9]:

$$\mathbf{X} = \mathbf{M}\mathbf{C} \quad (4.4.1)$$

where $\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_k)^T$ is the matrix of k spatially independent component maps and $\mathbf{M} = (\mathbf{m}_1, \dots, \mathbf{m}_k)$ the $m \times k$ mixing matrix. Thus the voxel values for each of the component maps are placed in separate rows of matrix $\mathbf{C}$, and each column of $\mathbf{M}$ represents a time series of a corresponding component.

In order to find the components that are due to the neural sources activated by a particular brain function, we first detect and remove the component maps of uninteresting signals. The identification of noisy components could be based on a fourth order statistic, kurtosis (κ). A very important feature of kurtosis is that it is the simplest statistical quantity for indicating the nongaussianity of a random variable [37]. Due to the Gaussianity of random noises and the specificity of spatial distributions of useful brain signals, the component maps of neural sources in fMRI usually have a kurtosis value far more than zero. The details of computing the independent components and denoising process have already been illustrated in Chapter 2.

Then correlation analysis is performed to select those components having a significant correlation with a designed reference time-series. Those voxels, whose signals are significantly correlated with the reference function above a preselected threshold, are designated "areas of activation".

4.4.2 Automated Search for the Best Path Model

In the second stage, the automated search for best fitting model, proposed by Bullmore [13], is adopted to find the best neural network with fitted path coefficients.

We presume there are n main regions of activation, say $v_1, v_2, v_3, \dots, v_n$. A convenient way of doing this is to write down the model as a set of simultaneous regression equation:

$$\mathbf{v} = \mathbf{K}\mathbf{v} + \mathbf{\Psi} \quad (4.4.2)$$

where $\mathbf{v} = (v_1, v_2, \dots, v_n)^T$ denotes the vector of regional variances, $\mathbf{\Psi} = (\Psi_1, \Psi_2, \dots, \Psi_n)^T$ denotes the vector of residual variances, and $\mathbf{K} = \{\theta_{ij}\}_{n \times n}$ denotes the $n \times n$ path model matrix where θ_{ij} represents the path coefficient of influence from j th region to i th region. Eq.(4.4.2) is just a matrix form of Eq.(4.2.2), but the path model matrix here is unknown.

Our objective is to find a best fitting model with q non-zero path coefficients, which minimizes a measure of discrepancy between the observed correlation matrix $\mathbf{C}$ and the correlation matrix predicted by the model $\Sigma(\theta)$.

Eq.(4.4.2) can be rearranged to show that:

$$\mathbf{v} = (\mathbf{I} - \mathbf{K})^{-1}\mathbf{\Psi} \quad (4.4.3)$$

The correlation matrix $\Sigma(\theta)$ predicted by the path model is then given by the McArdle-McDonald equation [73]

$$\Sigma(\theta) = (\mathbf{I} - \mathbf{K})^{-1}\mathbf{\Psi}\mathbf{\Psi}^T((\mathbf{I} - \mathbf{K})^{-1})^T \quad (4.4.4)$$

Estimates of the path coefficients are found by iteratively minimizing the maximum likelihood (ML) discrepancy function

$$F = \log|\Sigma(\theta)| + \text{trace}(\mathbf{C}\Sigma^{-1}(\theta)) - \log|\mathbf{C}| - p \quad (4.4.5)$$

where p is the number of regions involved in the model.

Under the null hypothesis $H_0 : \Sigma(\theta) = C$, the value of the maximum likelihood discrepancy function F multiplied by $\nu - 1$, where ν is the number of independent observations on each variable, is approximately distributed as χ^2 on $\frac{1}{2}p(p+1) - q$ degrees of freedom.

Given by Bullmore [13], the automated search for the best fitting model starts from the worst fitting model - that is the "null model" in which all path coefficients are constrained or set to zero. The algorithm computes the Lagrangian multiplier (LM) for each constrained coefficient and allows the coefficient with the maximum LM to be nonzero. The $q = 1$ path model is fitted to the data and LM again computed for all constrained coefficients. The coefficient with maximum LM is allowed to be nonzero, this $q = 2$ model is fitted to the data, and the process is iterated until as many coefficients as possible are unconstrained.

The Lagrangian multipliers L_i are a function of the first and second order partial derivatives of the ML discrepancy function with respect to the path coefficient θ_i :

$$L = \frac{\nu - 1}{2} \left(\frac{\partial F}{\partial \theta_i} \right)^2 \left(\frac{\partial^2 F}{\partial^2 \theta_i} \right)^{-1} \quad (4.4.6)$$

The derivatives can be evaluated by a simple numerical method [74] based on the equalities:

$$\begin{aligned} \frac{\partial^2 F}{\partial^2 \theta_i} &\approx \frac{1}{2} \text{trace} \left(\Sigma^{-1}(\theta) \mathbf{C}_i \Sigma^{-1}(\theta) \mathbf{C}_i \right) \\ \frac{\partial F}{\partial \theta_i} &= \frac{1}{2} \text{trace} \left(\Sigma^{-1}(\theta) (\Sigma(\theta) - \mathbf{C}) \Sigma^{-1}(\theta) \mathbf{C}_i \right) \end{aligned} \quad (4.4.7)$$

The matrix $\mathbf{C}_i$ comprises the partial derivatives of the modelled covariance matrix with respect to the i th parameter. These matrices can be computed for each parameter by finite forward differences, i.e.,

$$\mathbf{C}_i = \frac{\Sigma(\theta + \mathbf{e}_i \eta) - \Sigma(\theta)}{\eta} \quad (4.4.8)$$

where $\mathbf{e}_i$ denotes a vector of length $p \times (p - 1)$ with only one nonzero element, which is unitary at the i th position, and η is an arbitrarily small constant, e.g., $\eta = 10^{-4}$.

The maximum Lagrangian multiplier, at any step in the search, identifies the parameter estimate for which the local slope of the discrepancy function is steepest. To avoid fitting unidentified model, i.e., to restrict the number of unconstrained parameters q less than the number of nonredundant elements in the correlation matrix, we therefore keep a recursive model. Note that recursive, in the technical sense used here, means there can be no reciprocal connections between any pair of regions.

The minimum value of the discrepancy function, and therefore chi square, will monotonically decrease as the number of paths in the model increases. However, in the eyes of many analysts, an interesting model is probably one that explains as much as possible for as little as possible cost (in number of model parameters). Chi square, combined with the number of model parameter q , can provide measures of goodness that are sensitive to its cost. One example is Akaike's information criterion A:

$$A = \chi_q^2 + 2q \tag{4.4.9}$$

As path coefficients are unconstrained, chi square for the model with q nonzero paths χ_q^2 will decrease and $2q$ will increase. If the decrement $\chi_q^2 - \chi_{q+1}^2 < 2$, then A will have a minimum at q and this can be taken as an operational definition of the best model.

4.5 Experiments and Results

FMR Images presented in this section were obtained at the MRI Center of the Max-Planck-Institute of Cognitive Neuroscience in May 1993. An 8-Hz alternating checker-board pattern with a central fixation point was projected on a LCD system, and subjects were asked to fixate on the point during stimulations. When the subject was performing the experiment, 3-6 two-dimensional T_2^* -weighted images, each with 64 scans, were acquired using a gradient-echo FLASH sequence (TR=80.5 msec: TE=40 msec; matrix=128x 128). ON and OFF stimuli were presented at a rate of 5.162 sec/sample. Each stimulation period has four successive stimulation ON sates followed by four stimulation OFF states. The stimulations were repeated for eight cycles.

The fMRI scans were decomposed into 62 independent component maps by sICA. A threshold of kurtosis value $\kappa = 5.0$ was applied to remove the noisy sources and correlation analysis technique was employed to pick out the components of neural activations.

Figure 4.4: Activation maps whose time-series show a significant correlation with the stimulation time blocks. Colored voxels demonstrate great activations.

As shown in Fig.(4.4), there are three main regions of activation, demonstrated in three component maps, respectively, which have a significant correlation value with the stimulation time blocks (> 0.1): (a) primary visual sensory (PVS) area (17), (b) right primary somatomotor (RPSA) area (4), (c) left primary somatomotor (LPSA) area.

The time-series of each region were obtained from the corresponding column of the mixing matrix $\mathbf{M}$ multiplied by the averaged weight. According to Eq.(4.4.1), denote the component map of one activated region by $\mathbf{c}_i$, and its corresponding time course is $\mathbf{m}_i$, i.e., the i th column of $\mathbf{M}$, the temporal pattern of response in this region is computed from $\frac{1}{P} \sum_p c_{ip} \mathbf{m}_i$, where P is the number of activated voxels and c_{ip} denotes all the activated voxel intensities in i th component map. The residual variance for each region was estimated by 35% of the total variance of the region.

Figure 4.5: Measures of model goodness for a series of automatically specified path models: (a) chi square (b) p-value (c) Akaike's information criterion against the increasing number of nonzero path coefficient.

The automated search of best fitting model was performed on the data of three regions. Chi square, P value and Akaike's information criterion fit index for each model in the search series are shown in Fig(4.5). The Chi square value monotonically

Figure 4.6: Path model for the fMRI data from visual experiment: (a) best fitted model having $q = 2$ nonzero path coefficients (b) a confirmatory model with fully connected edges

decreases as the number of nonzero path coefficients in the model increases from zero to three. The P value for chi square has a maximum when $q = 2$ and Akaike's information criterion has a minimum when $q = 2$.

Besides the exploratory approach, a confirmatory SEM model was also derived for comparison. The base model was simulated as fully connected, where two connections are directly drawn from PVS to LPSA and RPSA, and meanwhile, an additional path is presumed from RPSA to LPSA. Thus, the best model, in terms of maximum probability and minimum cost, has $q = 2$ nonzero path coefficients and the path model is shown in Fig.(4.6a). The result of the confirmatory SEM model was shown in Fig.(4.6b). From Fig.(4.5), we can see that the fully connected model with $q = 3$ nonzero path coefficients has a lower fit index of p-value. The best model of $q = 2$ paths clearly conforms to the process of theoretically preferred model. Since the stimuli used in this experiment were all visual patterns, the main input region is primary visual sensory (PVS) area, and a connection is drawn directly forward from

PVS to RPSA and from PVS to LPSA, respectively to model the process of visual sensory and psychomotor operation. Both effects are looked on as direct effects. This may be understood as a simple model of an inner visual circuit or neural network. PVS is putatively responsible for receiving and analyzing the input from peripheral visual stimuli and then directing the signals to the contralateral primary somatomotor areas.

4.6 Conclusion and Discussion

Although it has been concluded that path analysis methods could provide a realistic picture of the interactions of brain activity and behavior, there is still no evidence for a direct, anatomical connection between the regions in question. Use of anatomical data in this way is complicated by the fact that much of it has been obtained by study of monkeys. And it is not always certain which area of, say, the macaque brain is homologous to a given area of human brain, especially if the human brain is specialized for language or some other uniquely human function. Even in cases where translation between human and nonhuman primate anatomy is unproblematic or can be agreed by convention, the mechanisms underlying brain lesions or mental disorders yet could not be accounted for by theoretical presumptions.

In such cases, an objective- or data-driven method is superior to a purely theoretical or confirmatory approach. Independent Component Analysis (ICA) is essentially exploratory technique since no constraints or a priori information is placed upon the system. This method can be used to select regions that appear to be operating as a functional unit. A significant correlation with the stimulation time blocks can provide clues for regions that need to be included in a structural equation model. SICA

based on the spatial independence of components in fMRI data gives the underlying spatial patterns of voxel intensities and suggest regions that may be part of a functional system. This approach has been used in fMRI studies to extract consistently task-related components and transiently-related components [9]. A further extension of this is that these components can be explained as an indication of different regions of a functional system.

Although SPM (Statistical Parameter Mapping) is possible for detecting activated areas in the brain, it can only provide the significance but not exactly the information of activation intensity for an activated region. However, ICA has been demonstrated as a more powerful method [9], not only for finding activation maps, but extracting intensity values from the component maps as well, due to its ability to separate activation maps from artifacts and noises.

The amalgamation with Bullmore's automated search for best path model [13] provides a picture as to possible functional influences that could best account for the interactions among this brain operation network. This approach is indeed a breakthrough step from conventionally confirmatory approaches in structural equation modeling, where a theory-constrained model is being modulated by evaluating modification of paths and omission of regions.

However, our combined approach is not undoubtedly good enough to find a irrefutable path model for an fMRI dataset. Questions still lie on the dependence of the p-value on asymptotic distributional assumptions and other details of the analysis, such as prior estimation of residual variances and effective degrees of freedom. Theoretically anatomical network may serve as a balance for validating the exploratory

models. But it is affirmative that our dynamic modeling approaches will have its significant potentials in looking into unpredictable and abnormal events in brain functions.

Chapter 5

Conclusion and Discussion

5.1 Summary of Contributions

fMRI has developed into one of the most popular tools for analyzing and studying the cognitive processes in human brains because of its noninvasiveness, high spatial and temporal resolution. Meanwhile, ICA is becoming a popular method to represent functional MRI data because the independent signals are usually the underlying components and suitable to characterize their observations. Our research primarily focuses on employing ICA techniques to restore, study and analyze fMRI images, due to the special ability of ICA to extract spatially and temporally independent neural systems, artifacts and noises in a data-driven manner and with little a priori information.

This thesis has been addressing three main topics of work we have achieved in functional brain analysis with fMRI as follows.

First, we have studied the concepts and several efficient algorithms of Independent Component Analysis (ICA), and its utility in analyzing fMRI data. A restoration model for fMRI has been constructed and implemented, which bases on identification

and elimination of a large number of artifacts and noises from ICA separation. In this model, a demixing process of ICA is performed using a robust neural network algorithm (here we employ Fast ICA [38]) by linear transformation of the noisy data. In the next stage, these independent source signals are extracted and classified in terms of their different properties under certain criteria and those undesirable components are thereby removed by switching corresponding soft switches "off". Three basic measures, Kurtosis, frequency spectrum and Hurst exponent have been proposed for identifying components and a list of usual artifacts underlying fMRI dataset and the corresponding identification and removal methods has been contrived. Finally, a reconstruction process of projecting the useful independent components back onto the mixed signals is done by applying the linear inverse transformation of ICA decomposition. The SICA is mainly used for fMRI analysis, because the number of voxels in the scanned images is far more than that of time points, which requires huge workload of time and memory for the computation of TICA.

Experiments have been performed on reducing noises and physiological fluctuations in fMRI. The noisy dataset and the restored one were both processed by an activation detection technique to compare their capabilities of activation detection. The outcomes lend support to the assumption that the capability of some activation detection techniques, such as correlation analysis and SPM, could be improved with a high correlation coefficient and good localization, after a restoration of the fMRI dataset using ICA. This study confirmed the feasibility of ICA for eliminating non-task related sources and noises. It is demonstrated that our model is possible for restoration of fMR images and a valid preprocessing technique in fMRI analysis especially in detecting brain regions that are activated by input stimuli.

Second, we have extended the concept of functional connectivity into a higher-order statistical sense, which is defined as brain systems within which the different voxels or regions are activated by the same dependency of temporal variation in the complete statistical sense. ICA is proposed as an efficient method for detecting connected brain regions. The experiment result of resting brain data reported the applicability and superiority of ICA in higher-order connectivity analysis where larger number of components could be found as higher-order functional connectivity and less noisy pictures of connected areas produced.

There are two advantages that feature the newly defined functional connectivity. The first is its accuracy in the measurement that are unaffected by any confounding factors which has been a prominent concern in the previous work of fMRI data analysis. The signal changes of fMRI are subtle and easily interfered with other artifacts and random noises. Traditional PCA method could not thoroughly analyze the connectivity without being affected by additive confounds, as the correlation between voxels is computed from the mixed value of intensities. However, ICA technique is capable of decomposing the fMR images into a number of independent source signals, which could purely be artifacts, noises or neural sources and hence the neural source components are noise-free. As the connectivity is totally determined by one source signal in a complete statistical sense, this analysis method is robust to any interferences.

The other merit is that this connectivity is investigated in a complete statistical sense, i.e., the brain units are connected in any higher order. Functional connectivity is often the result of neuronal interactions along anatomical or structural connections; in some cases it may be due to common input from an external neuronal or stimulus

source. Thus, after captured in fMR images, this connectivity is more likely to be demonstrated with a same neural source signal or component. In this sense, the higher-order connectivity is more appropriate to describe the underlying connections among neuronal elements.

Third, structural equation modeling has been applied into a functional network to analyze effective connectivity, where the regions involved in this model are selected from the component maps of SICA decomposition. An automated search for best fitting model, proposed by Bullmore, is employed to seek for the path model which can best describe the interactions among brain regions.

This modeling method is deemed as an exploratory approach against the previous confirmatory approach in that no or little a prior information is required on the brain regions in question. ICA is incorporated to automatically select regions that appear to be operating as functional units for a specific neural system. Only a significant correlation with the stimulation time blocks is used to provide clues for regions that need to be included in a structural equation model. Although SPM (Statistical Parameter Mapping) is possible for detecting activated areas in the brain, it can only provide the significance but not exactly the information of activation intensity for an activated region.

The amalgamation with Bullmore's automated search for best path model [13] provides a picture as to possible functional influences that could best account for the interactions among this brain operation network. From an experiment on visual-task data, our hybrid modeling approach could be concluded to be an efficient exploratory approach to analyzing effective connectivity for a specific brain operation. Since little presumption is needed on some predefined anatomical model, this method has

its tremendous potential in investigating the neuronal interactions underlying brain lesions and mental disorders.

In short, we could summarize the contributions of this thesis into three phrases: a restoration model, the higher-order connectivity conception and an exploratory modeling of effective connectivity.

5.2 Directions of Future Work

5.2.1 Enrichment of the Restoration Model

Although an ICA-based restoration model has been contrived, our restoration model has only a general framework with limited soft switches, which serve as an automatic threshold for classification and elimination. Up to now, only the approaches of reducing noises and physiological artifacts have been intensively developed and experiments been taken to examine their effectiveness. Chapter 2 illustrates various kinds of artifacts underpinning fMR images and three criteria for detecting different components, but we still encounter some unsatisfactory results when experimenting these methods on specific artifacts. For example, although Hurst exponent is one possible parameter for characterizing the randomness of a time-series, more concerns should be taken into account when applying it to specific fMRI time-series. Each kind of artifact should have its unique property, which may lead to various measures of distinctiveness. Besides, some signals underpinning fMRI are still elusive and yet to be recognized. Therefore, a quite large deal of knowledge is required to find ways of identifying other artifacts automatically. Attaining a powerful restoration model will shed great lights on the constituents of fMRI data and hereby "clean" our datasets

for further analysis.

5.2.2 Improvements in Structural Equation Modeling

Structural Equation Modeling has been used to model effective connectivity among brain units. Yet, many aspects in this analysis need further improvement. For instance, the prior estimation of residual variance for each variable in the equations is still not well determined. The residual variance is best thought of as influences that could not be accounted for by other variables within the model. Residuals are allowed to occur in the model for both theoretical and technical reasons. Brain regions that do not have a residual influence imply that all the variance in that region is accounted for by the connections with other regions in the model. This is an unlikely situation in a neural system. One typical approach for the estimation of residuals is to fix these values at 35 to 50% of the total variance, as employed in our experiment. But these residual effects can be modified if it significantly improves the fit of the model. This adjustment of residuals might lead to indeterminacy in the final solution. Thus, a best fitting estimation for the residual variance will enhance the steadiness and accuracy of the equation model.

Conventionally, all the variables involved in the equation models represent the variance of brain regions that are thought of as activations during a particular operation. However, because of the complexity of the neural mechanism, some external or environmental changes may also effect in the neuronal activity. In this sense, these effects can be looked on as a single source, in which some common inputs or influences are encompassed. In the future, a more comprehensible model to describe the effective connectivity should be constructed, where as many effects as possible can be

found and analyzed.

5.2.3 Detection of Different Neural Systems

The study of neural network and its interactions is important for relating cognitive theories with brain operations. It is highly unlikely that the psychological function is modulated by single brain constructs. On the other hand, it is also unlikely that a single brain region takes part in only one cognitive function. There may not be a single brain area that represents "attention" for instance, but there are more probably numerous brain areas whose interactions represent attention operations. The important point is that it may be possible for parts of the same anatomical network to be involved in another function when the interactions change. These notions about brain networks may suggest the reasons why some separate functions cannot occur simultaneously on human beings, since certain brain area is not allowed to effect in two different neural systems at the same time.

Therefore, it is promising work if we are able to detect distinct neural systems corresponding to separate brain functions, which consist of remote or overlapping brain areas. These neural systems can be represented by different patterns of functional connectivity, both in higher-order and second-order statistical sense. In the preceding chapters, we have illustrated that ICA is useful in finding higher-order connectivities and activated areas involved in certain brain operation. With our technique, if more experiments of various brain tasks are carried out, numerous neural systems especially those involved in higher cognitive functions of brain such as "meditation" and "attention" will be sought.

Bibliography

- [1] P. Jezzard, P. M. Matthews, and S. M. Smith. *FUNCTIONAL MRI an introduction to methods*. Oxford University Press, 2001.
- [2] S. Finger. *The Origins of Neuroscience: A History of Explorations into Brain Function*. New York: Oxford University Press, 1994.
- [3] K. J. Friston, D. D. Frith, P. F. Liddle, and R. S. J. Frackowiak. Functional connectivity: The principal-component analysis of large (PET) data sets. *Journal of Cerebral Blood Flow and Metabolism*, 13:5–14, 1993.
- [4] B. Biswal, F. Z. Yetkin, V. M. Haughton, and J. S. Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med*, 34:537–541, 1995.
- [5] J. Piyaratna, J. C. Rajapakse, C. Oesterle, and J. Hennig. Eigenimage analysis of resting brain fMR images. *Proceedings of International Conference on Information, Communications and Signal Processing*, 2001.
- [6] A. R. McIntosh and F. Gonzalez-Lima. Structural equation modeling and its application to network analysis in functional brain imaging. *Human brain mapping*, 2:2–22, 1994.

- [7] P. A. Bandettini, A. Jesmanowicz, E.C. Wong, and J. S. Hyde. Processing strategies for time-course data sets in functional MRI of the human brain. *Magn Reson Med*, 30:161–173, 1993.
- [8] K. J. Friston. Statistical parameter mapping. In R. W. Thatcher, M. Hallett, T. Zeffiro, W. R. John, and M Huerta, editors, *Functional Neuroimaging: Technical Foundations*. Academic Press, 1994.
- [9] M. J. Mckeown, A. J. Bell, T. J. Sejnowski, and et al. Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping*, 6:160–188, 1998.
- [10] N. C. Andreasen. *BRAVE NEW BRAIN, Conquering Menial Illness in the Era of the Genome*. Oxford University Press, 2001.
- [11] J. C. Rajapakse, F. Kruggel, J. M. Maisog, and D. Y. Cramon. Modeling hemodynamic response for analysis of functional MRI time-series. *Human Brain Mapping*, 6:283–300, 1998.
- [12] <http://www.fil.ion.ucl.ac.uk/spm/>.
- [13] E. Bullmore, B. Horwitz, G. Honey, M. Brammer, S. Williams, and T. Sharma. How good is good enough in path analysis of fMRI data. *NeuroImage*, 11:289–301, 2000.
- [14] B. Biswal, E. A. Deyoe, and J. S. Hyde. Reduction of physiological fluctuations in fMRI using digital filters. *Magn Reson Med*, 35:107–113, 1996.
- [15] T. H. Le and X. Hu. Retrospective estimation and correction of physiological artifacts in fMRI by direct extraction of physiological activity from MR data. *Magn Reson Med*, 35:290–298, 1996.

- [16] K. Chuang and J. Chen. IMPACT:Image-based physiological artifacts estimation and correction technique. *Magn Reson Med*, 46:344–353, 2001.
- [17] K. J. Friston, S. Williams, R. Howard, R. S. Frackowiak, and R. Turner. Movement-related effect in fMRI time-series. *Magn Reson Med*, 35:346–355, 1996.
- [18] J. V. Hajnal, R. Myers, A. Oatridge, J. E. Schwieso, I. R. Young, and G. M. Bydder. Artifacts due to stimulus correlated motion in functional imaging of the brain. *Magn Reson Med*, 31:283–291, 1994.
- [19] W. A. Edelstein, G. H. Glover, C. J. Hardy, and R. W. Redington. The intrinsic signal-to-noise ratio in NMR imaging. *Magn Reson Med*, 3:604–618, 1986.
- [20] T. B. Parrish, D. R. Gitelman, K. S. Labar, and M. M. Mesulam. Signal to noise influence on clinical fMRI. Proceedings of *NeruoImage* Human Brain Mapping 2000 Meeting, page Poster No: 532.
- [21] M. J. Lowe and J. A. Sorenson. Spatially filtering functional magnetic resonance imaging data. *Magn Reson Med*, 37:723–729, 1997.
- [22] P. P. Mitra, S. Ogawa, X. Hu, and K. Ugurbil. The nature of spatiotemporal changes in cerebral hemodynamics as manifested in functional magnetic resonance imaging. *Magn Reson Med*, 37:511–518, 1997.
- [23] A. Hyvarinen and E. Oja. Independent component analysis: algorithms and applications. *Neural Networks*, 13:411–430, 2000.
- [24] J. V. Stone. Independent component analysis: an introduction. *TRENDS in Cognitive Science*, 6:NO.2, 2002.
- [25] A. Hyvarinen. Survey on independent component analysis. *Neural Computing Surveys*, 2:94–128, 1999.

- [26] V. D. Calhoun, T. Adali, G. D. Pearlson, and J. J. Pekar. Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. *Human Brain Mapping*, 13:43–53, 2001.
- [27] M. J. McKeown and T. J. Sejnowski. Independent component analysis of fMRI data: examining the assumptions. *Human Brain Mapping*, 6:368–372, 1998.
- [28] J. V. Stone, J. Porrill, C. Bushel, and K. Friston. Spatial, temporal and spatiotemporal independent component analysis of fMRI data. *Proceedings of Spatial-temporal modelling and its applications*, 1999.
- [29] J. C. Rajapakse and W. Lu. Extracting task-related components in functional MRI. *ICA*, 2001.
- [30] S. Vorobyov and A. Cichocki. Blind noise reduction for multisensory signals using ICA and subspace filtering, with application to EEG analysis. *Biol. Cybern.*, 86:293–303, 2002.
- [31] T. Jung, S. Makeig, C. Humphries, and et al. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37:163–178, 2000.
- [32] T. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, 111:1745–1758, 2000.
- [33] R. N. Vigario. Extraction of ocular artifacts from EEG using independent component analysis. *Electroencephalography and clinical Neurophysiology*, 103:395–404, 1997.
- [34] A. J. Bell and T. J. Sejnowski. An information maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7:1129–1159, 1996.

- [35] A. Cichocki, R. Thawonmas, and S. Amari. Sequential blind signal extraction in order specified by stochastic properties. *Electronics Letters*, 33(1):64–65, 1997.
- [36] A. Cichocki and S. Amari. *Adaptive Blind Signal and Image Processing*. JOHN WILEY and SONS, 2002.
- [37] A. Hyvarinen, J. Karhunen, and E. Oja. *Independent component analysis*. John Wiley and Sons, Inc., 2001.
- [38] A. Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Tans. on Neural Networks*, 10(3):626–634, 1999.
- [39] P. Comon. Independent component analysis- a new concept? *Signal Processing*, 36:287–314, 1994.
- [40] A. Hyvarinen. New approximations of differential entropy for independent component analysis and projection pursuit. *Advances in Neural Information Processing Systems*, 10:273–279, 1998.
- [41] T. Lee, M. Girolami, and T. J. Sejnowski. Independent component analysis using an extended informax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation*, 11:417–441, 1999.
- [42] C. G. Phillips, S. Zeki, and H. B. Barlow. Localization of function in the cerebral cortex. Past, present and future. *Brain*, 107:327–361, 1984.
- [43] H. E. Hurst, R. P. Black, and Y. M. Simaika. *Long Temn Storage*. Constable, London, U.K., 1965.
- [44] S. Vorobyov and A. Shilo. The recurrent normalized range method for revelation of fractal structure and its application to analysis of EEG. *Radioelectron Infomn*, 3:162–165, 1998.

- [45] P. A. Bandettini, A. Jesmanowicz, J. V. Kylen, R. M. Birn, and J. S. Hyde. Functional MRI of brain activation induced by scanner acoustic noise. *Magn Reson Med*, 39:410–416, 1998.
- [46] D. A. Feinberg and A. S. Mark. Human brain motion and cerebrospinal fluid circulation demonstrated with MR velocity imaging. *Radiology*, 163:793–799, 1987.
- [47] D. Greitz, R. Wirestam, A. Franck, B. Nordell, C. Thomsen, and F. Stahlberg. pulsatile brain movement and associated hydrodynamics studied by magnetic resonance phase imaging. *Neuroradiology*, 34:370–380, 1992.
- [48] D. Greitz. Cerebrospinal fluid circulation and associated intracranial dynamics: A radiologic investigation using MR imaging and radionuclide cisternography. *Acta Radiol*, 34(Suppl. 386):1, 1993.
- [49] R. S. J. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan, and J. C. Mazziotta. *Human Brain Function*. Academic Press, San Diego, 1997.
- [50] S. L. Bressler and J. A. S. Kelso. Cortical coordination dynamics and cognition. *Trends in Cognitive Sciences*, 5:26–36, 2001.
- [51] A. R. McIntosh. Mapping cognition to the brain through neural interactions. *Memory*, 7:523–548, 1999.
- [52] G. L. Gerstein and D. H. Perkel. Simultaneously recorded trains of action potentials: Analysis and functional interpretation. *Science*, 164:828–830, 1969.
- [53] G. L. Gerstein, P. Bedenbaugh, and AMHJ Aersten. Neuronal assemblies. *IEEE Trans Biomed Eng*, 36:4–14, 1989.
- [54] A. Aersten and H. Preissl. Dynamics of activity and connectivity in physiological neuronal networks. In Schuster HG, editor, *Non Linear Dynamics and Neuronal Networks*, pages 281–302. New York, VCH Publishers Inc., 1993.

- [55] J. Xiong, L. M. Parsons, J. Gao, and P. T. Fox. Interregional connectivity to primary motor cortex revealed using MRI resting state images. *Human Brain Mapping*, 8:151–156, 1999.
- [56] G. Lohmann and S. Bohn. Using replicator dynamics for analyzing fMRI data of the human brain. *IEEE Transactions on Medical Imaging*, 21:No.5, 2002.
- [57] K. J. Worsley, J. Cao, T. Paus, M. Petrides, and A. C. Evans. Applications of random field theory to functional connectivity. *Human Brain Mapping*, 6:364–367, 1998.
- [58] K. Friston, J. Phillips, D. Chawla, and C. Buchel. Revealing interactions among brain systems with nonlinear PCA. *Human Brain Mapping*, 8:92–97, 1999.
- [59] W. Jeong. Functional connectivity in the resting brain using fMRI. *Proceedings of the 4th Annual Fred S. Grodins Graduate Research Symposium*, 2000.
- [60] B. Biswal, J. V. Kylen, and J. S. Hyde. Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps. *NMR IN BIOMEDICINE*, 10:165–170, 1997.
- [61] M. Hampson, B. S. Peterson, P. Skudlarski, J. C. Gatenby, and J. C. Gore. Detection of functional connectivity using temporal correlations in MR images. *Human Brain Mapping*, 15:247–262, 2002.
- [62] K. J. Friston. Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, 2:56–78, 1994.
- [63] S. Zeki. The motion pathways of the visual cortex. In C. Blakemore, editor, *Vision: Coding and Efficiency*, pages 321–345. Cambridge University Press, 1990.
- [64] J. R. Schott. *Matrix Analysis for Statistics*. Wiley, 1997.

- [65] J. Piyaratna and J. C. Rajapakse. Comparison of Gaussian and Markov random field approaches for detection of activities in fMRI. International conference on Neural Information Processing, 2000.
- [66] A. R. McIntosh and F. Gonzalez-Lima. Structural modeling of functional visual pathways mapped with 2-deoxyglucose: Effects of patterned light and footshock. *Brain Res*, 578:75–86, 1992.
- [67] A. R. McIntosh, C. L. Grady, L. G. Ungerleider, J. V. Haxby, S. I. Rapoport, and B. Horwitz. Network analysis of cortical visual pathways mapped with PET. *Journal of Neuroscience*, 14:655–666, 1994.
- [68] A. R. McIntosh, B. Horwitz, J. V. Haxby, L. G. Ungerleider, and C. L. Grady. Functional cortical networks during short-term recognition memory for faces. *Soc. Neurosci. Abstr*, 20:362, 1994.
- [69] G. M. Maruyama. *Basics of Structural Equation Modeling*. SAGE Publications, Inc., 1998.
- [70] K. G. Joreskog and D. Sorbom. *LISREL 7 User's Reference Guide*. Mooresville, IN: Scientific Software, Inc., 1989.
- [71] P. M. Bentler. *Theory and Implementation of EQS, A Structural Equations Program*. Los Angeles: BMDP Statistical Software, 1985.
- [72] A. R. McIntosh and F. Gonzalez-Lima. The application of structural modeling to metabolic mapping of functional neural systems. In F. Gonzalez-Lima, T. Finkenstadt, and H. Scheich, editors, *Advances in Metabolic Mapping Techniques for Brain Imaging of Behavioral and Learning Functions*, pages 219–258. NATO ASI Series. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1992.

- [73] J. J. Mcardle and R. P. Mcdonald. Some algebraic properties of the reticular action model for moment structures. *Br. J. Math. Stat. Psychol.*, 37:234–251, 1984.
- [74] R. Cudeck, K. J. Klebe, and S. J. Henly. A simple Gauss-Newton procedure for covariance structure analysis with high-level computer languages. *Psychometrika*, 58:211–232, 1993.