Multimodal Neuroimaging Integrating Functional Magnetic Resonance Imaging and Electroencephalography

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF THE UNIVERSITY OF MINNESOTA
BY

ZHONGMING LIU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

BIN HE, Advisor
Professor of Biomedical Engineering

Month and Year of Degree Clearance
September, 2008
Acknowledgements

I would like to express my most sincere appreciation to my advisor and mentor, Dr. Bin He, who directed me to this research with enormous support and resources throughout the program.

I am also eager to thank our collaborators, Dr. Wei Chen and Dr. Nanyin Zhang, from Center of Magnetic Resonance Research at the University of Minnesota, for scientific discussions and professional supports in MRI/fMRI technologies. I also wish to thank Dr. Chang-Hwan Im and Dr. Xiaoxiao Bai for their valuable contributions to the research included in this dissertation, to Ms. Lin Yang, Mr. Fedja Kecman, Ms. Cristina Rios and Ms. Han Yuan for their assistance in EEG/fMRI experimentation and data analysis, and to Mr. Christopher Wilke for technical discussions and English proofreading. Many thanks also go to Ms. Baharan Kamousi, Dr. Nobuyuki Yamawaki and Mr. Kun Wang for fruitful collaborations in neuroimaging and neuroscience researches. I am also grateful to Mr. Xu Li, Ms. Lei Qin, Dr. Yuan Lai, Dr. Lei Ding and Dr. Yuan Xu for sharing knowledge and experience in biosignal and image processing.

Furthermore, I would like to take this opportunity to acknowledge Dr. Steven Pogwizd for his collaboration and scientific discussions in cardiac mapping experiments on animals. Special thanks go to Dr. Xin Zhang for helping me start up my research in the lab, to Mr. Chenguang Liu and Mr. Chengzong Han for their great assistance in developing cardiac imaging methodologies. I also appreciate the help from many others with whom I have worked during the past 5 years.

Last but not least, I would like to express appreciation to my wife, Mrs. Fang Wu, for her continuous love, support and patience, to my daughter, Miss Angie Liu, for being a gorgeous gift of happiness to my life, to my parents, Mr. Dingran Liu and Mrs. Shulan Liu for giving me intelligence and courage, and to my sister and her family, Mrs. Jiayu Liu, Mr. Ming Zhao and Miss. Zhengwei Zhao for their understanding and encouragement.
Abstract

Functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) are two widely used neuroimaging modalities with complementary merits and limitations. FMRI has low temporal resolution but high spatial resolution, while EEG has low spatial resolution but high temporal resolution. The researches included in my dissertation aim at developing an fMRI-EEG integrated multimodal functional neuroimaging approach with significantly enhanced spatiotemporal resolution relative to fMRI or EEG alone. Towards this goal, i) we have established the capability to simultaneously record EEG and fMRI signals; ii) we have mathematically modeled the interactions between stimuli (or tasks), synaptic currents and BOLD responses, and have developed a model-driven linear regression method for quantifying BOLD fMRI signals to characterize event-related electrophysiological responses; iii) we have developed two novel algorithms based on an adaptive Wiener filter and the Twomey regularization, respectively, for the fMRI-constrained cortical current density imaging; iv) we have employed the developed fMRI-EEG technique to investigate the cortical visual pathway and the cortical substrates responsible for the bilateral visual integration. The outcome of these researches is encouraging. Significant progresses have been achieved in both technical and fundamental aspects of the fMRI-EEG integrated neuroimaging, which holds potential to open a unique window to non-invasively image and investigate electrophysiological brain activity, as well as its relationship to cerebral hemodynamics.
Table of Contents

Acknowledgements .............................................................................................................. i
Abstract ................................................................................................................................ ii
Table of Contents ................................................................................................................ iii
List of Figures ........................................................................................................................ vi

Chapter 1 Introduction ........................................................................................................ 1
  1.1 Overview .................................................................................................................... 1
  1.2 Motivation .................................................................................................................. 4
  1.3 The Organization of this Thesis ............................................................................. 6

Chapter 2 Basics of EEG and fMRI .................................................................................... 8
  2.1 Introduction .............................................................................................................. 8
  2.2 Electroencephalograph .......................................................................................... 8
    2.2.1 Physiological origin of EEG ......................................................................... 8
    2.2.2 EEG Forward Problem .............................................................................. 12
    2.2.3 EEG Inverse Problem ............................................................................. 16
  2.3 Functional Magnetic Resonance Imaging .............................................................. 19
    2.3.1 MRI Physics ............................................................................................... 19
    2.3.2 BOLD fMRI .............................................................................................. 20
    2.3.3 fMRI Statistic Parametric Mapping ......................................................... 22

Chapter 3 Simultaneous EEG-fMRI ................................................................................. 25
  3.1 Introduction ............................................................................................................. 25
  3.2 Simultaneous fMRI-EEG Recordings .................................................................... 28
    3.2.1 Hardware ..................................................................................................... 28
    3.2.2 Signal Processing ....................................................................................... 29
  3.3 Evaluation of fMRI-EEG Simultaneous Recordings .............................................. 36
    3.3.1 Visual Evoked Potential ............................................................................. 36
    3.3.2 Steady-State Visual Evoked Potential ....................................................... 38
    3.3.3 Resting State Alpha-Wave Modulation ..................................................... 39
  3.4 Electrophysiological Correlates to Negative BOLD ................................................ 40
    3.4.1 Background ............................................................................................... 41
    3.4.2 NBR to Small Circular Stimuli ..................................................................... 42
    3.4.3 EEG Correlates to the NBR ....................................................................... 45
  3.5 Conclusion & Discussion ......................................................................................... 50
Chapter 4 Neurovascular Coupling

4.1 Introduction ........................................................................................................ 53
4.2 Theoretical Modeling ......................................................................................... 56
  4.2.1 Systematic Perspective ............................................................................... 56
  4.2.2 BOLD Quantification ................................................................................ 57
4.3 Experimental Investigation .............................................................................. 61
  4.3.1 Full-screen Checkerboard Experiment ....................................................... 61
  4.3.2 Responses to Variable Visual Contrast ....................................................... 63
4.4 Conclusion & Discussion ................................................................................... 67

Chapter 5 fMRI-EEG Integrated Neuroimaging: Methodology

5.1 Introduction ........................................................................................................ 70
  5.1.1 fMRI-Constrained EEG Source Imaging ......................................................... 70
  5.1.2 EEG-Informed fMRI Analysis ......................................................................... 73
5.2 Methods for fMRI-Constrained Electrical Source Imaging ............................... 74
  5.2.1 Wiener Filter for EEG Source Reconstruction ........................................... 75
  5.2.2 fMRI-Weighted Current Density Estimation .............................................. 77
  5.2.3 fMRI-Constrained Adaptive Wiener Filter ................................................. 79
  5.2.4 Twomey Regularization .............................................................................. 87
  5.2.5 fMRI-Seeded Dipole Fitting ....................................................................... 89
5.3 Evaluation of fMRI-Constrained Electrical Source Imaging ............................. 91
  5.3.1 Simulation for fMRI-Constrained Adaptive Wiener Filter ......................... 91
  5.3.2 Experiments for fMRI-Constrained Adaptive Wiener Filter .................... 107
  5.3.3 Simulation for Twomey Algorithm .......................................................... 110
  5.3.4 Simulation for fMRI-Seeded Dipole Fitting ............................................. 114
5.4 Discussion ........................................................................................................ 115
  fMRI-Constrained Adaptive Wiener Filter ............................................................. 115
  Twomey Algorithm ................................................................................................. 119

Chapter 6 fMRI-EEG Integrated Neuroimaging: Application

6.1 Introduction ...................................................................................................... 122
6.2 Materials and Methods ..................................................................................... 126
  6.2.1 Subjects ..................................................................................................... 126
  6.2.2 Stimuli ....................................................................................................... 126
  6.2.3 Data Acquisition ........................................................................................ 126
6.3 Data Analysis ................................................................................................... 127
List of Figures

Figure 1.1 Spatial and temporal resolution of various noninvasive and invasive functional imaging/recording techniques ................................................................. 3
Figure 1.2 Illustration of the fMRI-EEG integrated multimodal neuroimaging ....... 5
Figure 2.1 Electrophysiological behavior of a single neuron: action potential vs. post-synaptic potential ................................................................. 9
Figure 2.2 Intracellular and extracellular currents associated with the activation potential propagation ................................................................. 11
Figure 2.3 Spatial scale and frequency range of six electrophysiological recording techniques ........................................................................... 12
Figure 2.4 Post-synaptic electrophysiological process ................................... 13
Figure 2.5 EEG forward and inverse problems ............................................ 17
Figure 2.6 Principle of BOLD-fMRI for neuroimaging .................................. 22
Figure 3.1 Interference between EEG and fMRI recordings ......................... 26
Figure 3.2 Experimental setup for simultaneous fMRI-EEG recordings ........... 28
Figure 3.3 Gradient Artifacts .................................................................. 30
Figure 3.4 Time-frequency features of GA and the GA removal ................. 31
Figure 3.5 Cardiac Ballistic Artifacts ....................................................... 33
Figure 3.6 Independent component analysis for identifying and removing the artifact and noise components ........................................................ 35
Figure 3.7 MRI/fMRI image quality for simultaneous fMRI-EEG ............... 37
Figure 3.8 Visual Evoked Potentials ......................................................... 38
Figure 3.9 Steady-State Visual Evoked Potentials ..................................... 39
Figure 3.10 Alpha Modulation with an eye-open-and-close task ................... 40
Figure 3.11 Negative BOLD response to small visual stimuli ..................... 43
Figure 3.12 NBR independent to stimulus locations .................................. 44
Figure 3.13 Spatial and time-frequency analyses of EEG based on the independent component analysis ............................................................. 47
Figure 3.14 Identified independent components in correlation with the NBR .... 48
Figure 3.15 Spatial and temporal consistency between the NBR and the alpha- and delta-band EEG source activity........................................................................................................49
Figure 4.1 Linear systems modeling the interactions between stimuli, neuronal responses and BOLD responses........................................................................................................56
Figure 4.2 Simplified models relating stimuli with the BOLD signal change...........58
Figure 4.3 Procedures for quantifying BOLD fMRI signals to characterize event-related synaptic source activity.................................................................................................61
Figure 4.4 Comparison between the spatial distributions of the time integral of current density power estimated from the VEP signals, and the BOLD effect sizes derived from the fMRI data................................................................................................................62
Figure 4.5 Pattern-reversal visual stimuli with variable contrasts.........................63
Figure 4.6 BOLD-to-VEP relationship in a single subject........................................66
Figure 4.7 BOLD-to-VEP relationship in a group of subjects.................................67
Figure 5.1 fMRI-weighted EEG source imaging.......................................................78
Figure 5.2 Flow-chart of the fMRI-EEG co-registration..........................................81
Figure 5.3 Source signals over repeated EEG epochs..............................................83
Figure 5.4 Head and cortical models for computer simulation studies...................92
Figure 5.5 Simulating the BOLD response from the EEG source activity.............94
Figure 5.6 SVD analysis on the segmented EEG epochs.........................................96
Figure 5.7 BOLD effect sizes vs. source power estimated from EEG....................100
Figure 5.8 Source distribution reconstruction and source waveform estimation with fMRI false positives.................................................................................................................102
Figure 5.9 Source distribution reconstruction and source waveform estimation with fMRI false negatives................................................................................................................104
Figure 5.10 Estimated source waveforms in comparison with the real source waveforms with different temporal features.................................................................106
Figure 5.11 EEG-fMRI integrated imaging of the cortical responses to visual stimulation.................................................................109
Figure 5.12 Cortical source imaging by the Twomey algorithm.............................111
Figure 5.13 Point spread function (PSF) for fMRI visible and invisible sources......112
Figure 5.14 Localization error for the Twomey algorithm.............................113
Figure 5.15 Computer simulation on the fMRI-seeded dipole fitting..................115
Figure 6.1 Scenarios for investigating the cortical activity underlying the bilateral visual integration.................................................................124
Figure 6.2 VEP responses to unilateral stimuli for a single subject...............131
Figure 6.3 Topographic similarity and dissimilarity between the VEP responses to the LVF and RVF stimuli.................................................................133
Figure 6.4 fMRI activation and RC maps (LVF vs. RVF)...............................135
Figure 6.5 fMRI activation maps in response to LVF, RVF and bilateral stimuli for 4 subjects.................................................................136
Figure 6.6 RC maps (LVF vs. RVF) individually obtained from 6 subjects.......137
Figure 6.7 Group-averaged RC maps visualized on the inflated cortical surface...138
Figure 6.8 VEP responses to bilateral stimuli vs. the sum of responses to separate unilateral stimuli.................................................................139
Figure 6.9 Under-summation of VEP responses to bilateral stimuli...............141
Figure 6.10 Spatially nonlinear fMRI responses to bilateral stimuli for a single subject.................................................................142
Figure 6.11 Under-summed fMRI responses for 5 subjects.........................143
Figure 6.12 Spatiotemporal cortical activity underlying the unilateral visual processing and the bilateral visual integration........................................144
Chapter 1 Introduction

1.1 Overview

The past decades have witnessed an explosive development of functional neuroimaging methodologies, which aim at noninvasively imaging the human brain activity and connectivity associated with sensory and cognitive processes in both normal and pathological conditions. The modern neuroimaging techniques rely on various “source” signals that change across different spatial and temporal scales in accompany with underlying neural activity. For instance, neural activity elevates electrophysiological signals, such as action potentials (AP) and post-synaptic potentials (PSP), which serve as the primary messengers for communications among neurons. It is generally agreed that neural activity is most well characterized by neuronal electrophysiology, which follows tightly with neuronal dynamics. Additionally, neural activity is also coupled with metabolic and hemodynamic processes. The human brain function requires sustained blood flow to supply oxygen to compensate for cerebral metabolic energy consumption. As a result, changes in neural activity often induce cascaded changes in cerebral metabolic rate for oxygen (CMRO$_2$), cerebral blood flow (CBF), oxygen extraction fraction (OEF) and cerebral blood volume (CBV) etc. In contrast to electrophysiological signals, such metabolic and hemodynamic responses are much slower and reflect the indirect and secondary effects of neural activity.

Electroencephalography (EEG) (Berger, 1929) and magnetoencephalography (MEG) (Cohen, 1968) are two popular functional neuroimaging techniques based on electrophysiological principles. For EEG and MEG, electric and magnetic sensors, respectively, are placed on/above the scalp surface to record electrical potentials and magnetic fluxes. Such electromagnetic signals arise collectively from all neuronal currents within the brain, and propagate (virtually) instantaneously from the brain volume to the external head surface through a volume conduction process (Nunez, 1981; Hämäläinen et al. 1993). The instantaneous nature of EEG/MEG indicates remarkable temporal resolution and precision for studying the brain function in the neuronal time scale. However, the collective nature suggests insufficient spatial resolution and
specificity to precisely distinguish different neural assemblies activated concurrently (Nunez and Srinivasan, 2005). This is regardless of recent advancements in electromagnetic source imaging (ESI) (Baillet et al. 2001; He and Lian 2002; Michel et al. 2004), which has led to great strides in improving the spatial resolution of EEG/MEG to a scale of centimeter or even smaller, by modeling the head volume conductor and solving an ill-posed inverse problem.

Functional magnetic resonance imaging (fMRI), based on hemodynamic and metabolic principles, provides another powerful imaging tool for visualizing brain areas engaged in certain sensory or cognitive functions. Neural activity induces the concomitant alternation of the local oxyhemoglobin vs. deoxyhemoglobin content, giving rise to a so-called blood oxygen level dependent (BOLD) MR signal change (Ogawa et al. 1990). Since its advent in the early 1990’s (Ogawa et al. 1992; Kwong et al. 1992; Bandettini et al. 1992), the BOLD-contrast fMRI has rapidly gained a dominant position in neuroimaging and neuroscience researches (Raichle and Mintun 2006). This technique has superiorities over competing electrophysiological neuroimaging techniques, due to its whole brain coverage, relatively uniform sensitivity, high spatial resolution and specificity. These advantages are primarily attributable to well-established MR imaging strategies (e.g. echo-planar imaging) that allow for the frequency and phase encoding of the spatial distribution of BOLD-contrast MR signals. However, fMRI is limited by the poor temporal resolution as well as its indirect nature with respect to neuronal activity. These limitations often lead to concerns to any simple interpretation of the BOLD signal as a surrogate index of neural activity, as the neurovascular coupling mechanism is complex and remains uncertain in various aspects.

In addition to fMRI and EEG/MEG, other noninvasive functional brain mapping techniques include positron emission tomography (PET) (Fox and Raichle 1986; Fox et al. 1988), single-photon emission computed tomography (SPECT), near-infrared spectroscopy (NIRS). These neuroimaging techniques as well as fMRI share closely related metabolic or hemodynamic principles. As such, all of them suffer from limited temporal resolution regardless of their different sampling rates, while fMRI has better (or
at least equivalent) spatial resolution than NIRS, PET or SPECT, without the need of applying radioactive tracers as required by PET and SPECT.

In summary, virtually all modern neuroimaging modalities are based on either electrophysiological or hemodynamic/metabolic principles. Their strengths and limitations depend largely upon the spatiotemporal characteristics of the measured “source” signals in relation to underlying neuronal activity, as well as many diverse kinds of sensing and imaging methods applied to individual modality. For an illustrative overview, we graph the existing noninvasive neuroimaging methodologies, in comparison with some representative invasive experimental techniques, with respect to their spatial and temporal resolution, as shown in Fig. 1.1.

**Figure 1.1** Schematic illustration of the ranges of spatial and temporal resolution of various noninvasive (in blue) and invasive (in red) imaging/recording techniques. The invasive techniques include intracranial electrophysiological recordings of single-unit activity (SUA), multi-unit activity (MUA) and local field potential (LFP).
1.2 Motivation

As stated in the above overview, no existing noninvasive neuroimaging modality can achieve, by itself alone, high resolution in both spatial and temporal domains. One technique may have merits in the temporal aspect but limitations in the spatial aspect (such as EEG/MEG), or vice versa (such as fMRI/PET/SPECT). The complementary features of existing techniques motivate recent efforts that attempt to integrate multiple neuroimaging modalities, particularly EEG/MEG and fMRI (Liu et al. 2006a; Dale and Halgren 2001; Gotman et al. 2006; Ritter and Villringer 2006; Shibasaki 2008), namely multimodal functional neuroimaging.

The merits of multimodal functional neuroimaging are at least two folded. First, electrophysiological and hemodynamic/metabolic signals reflect distinct but closely related aspects of the underlying neural activity. Combining data acquired through multiple modalities allows us to study brain functions from different perspectives. Convergent evidences, on one hand, definitely lead to much more confident inference and conclusions with regard to neural activity. Contradictory observations, on the other hand, also make obvious the need to pose new hypotheses to guide further investigations of the human neural system. Secondly, establishing a multimodal neuroimaging technique promises to fill in an otherwise vacant range of spatial and temporal resolution that fails to be achieved by any existing technique alone. For instance, the integration of fMRI and EEG/MEG holds the potential to reach millimeter-scale spatial resolution and millisecond-scale temporal resolution, as illustrated in Fig. 1.1. Such enhanced spatiotemporal resolution would provide us a unique, and importantly noninvasive, window to noninvasively investigate neural activities and their interactions.

In the context of multimodal neuroimaging, the present thesis focuses upon the integration of fMRI and EEG for the following reasons. i) Among existing neuroimaging techniques based on hemodynamic/metabolic principles (i.e. fMRI, PET, SPECT and NIRS), fMRI is the most widely used approach with the highest spatial resolution and reasonable temporal resolution for resolving the neural activity-induced hemodynamic changes. ii) Between the existing electrophysiological imaging techniques (i.e. EEG and MEG), EEG has comparable spatial resolution as MEG while it is feasible to record EEG
(but not MEG) simultaneously with fMRI. As illustrated in Fig. 1.2, the multimodal neuroimaging that combines fMRI and EEG is based on the close coupling between electrophysiological and hemodynamic responses.

![Illustration of the fMRI-EEG integrated multimodal neuroimaging.](image)

**Figure 1.2** Illustration of the fMRI-EEG integrated multimodal neuroimaging.

However, the integration of fMRI and EEG also faces great challenges that remain to be solved. These challenges cause uncertainties and concerns that largely compromise the intuitive promise and benefits such multimodal techniques would possibly bring to neuroimaging and neuroscience fields. For a brief summary of the most critical challenges and problems in this regard, we formulate the following questions. What is the physiological and physical relationship between hemodynamic and electrophysiological responses? Or more specifically, how should we quantitatively interpret the signal in one modality in terms of the signal in the other modality? How can EEG be reliably recorded during fMRI scans without losing the signal quality for both modalities? How should we fuse the recorded fMRI-EEG data in a principled way to achieve enhanced neuroimaging capability? How should we design an fMRI-EEG experiment to address a specific neuroscience problem by the cross-modal integration? How should we interpret multimodal imaging results with what kinds of cautions be taken?
All of these questions have attracted tremendous attention and interdisciplinary efforts from scientists and engineers in neuroscience, biomedical engineering, biophysics, electrical engineering and mathematics etc. In the past 5 years, we have made considerable progresses towards answering these questions, by i) modeling the neurovascular coupling relationship and quantifying BOLD fMRI signals to characterize electrophysiological responses, ii) exploring and evaluating simultaneous EEG-fMRI recordings, iii) developing fMRI-EEG integrated cortical current density imaging algorithms, iv) applying the developed fMRI-EEG integration techniques to studying the human visual system.

1.3 The Organization of this Thesis

This thesis covers some of the research outcomes we have achieved with regard to the fMRI-EEG integrated multimodal functional neuroimaging.

In chapter 2, we review the principles of electroencephalography, electrical source imaging, and functional magnetic resonance imaging.

In chapter 3, we introduce the recording and signal processing techniques for the simultaneous acquisition of both fMRI and EEG data. The techniques are evaluated with three well-controlled experimental paradigms, and employed to investigate the EEG correlates to the negative BOLD responses.

In chapter 4, we present a theoretical model describing the interactions between stimuli, neuronal responses and hemodynamic signals. Stemming from this model, we rigorously derive a model-driven method for quantifying BOLD signals to characterize event-related electrophysiological responses. The theoretical results are tested against real fMRI and EEG data in response to visual stimuli with variable contrasts. Furthermore, we discuss the range of conditions in which our modeling assumption is valid (or invalid).

In chapter 5, we describe two novel fMRI-EEG integration algorithms in comparison with the conventional fMRI-weighted current density imaging and the fMRI-seeded dipole fitting. The first algorithm is specifically based on the theoretical model introduced in chapter 4, while the second algorithm is based on the Twomey regularization. For these algorithms, the results from a series of computer simulations and preliminary experimental data are reported and discussed.
In chapter 6, the developed fMRI-EEG multimodal approaches are used to study the human visual pathway and the bilateral visual integration.
Chapter 2 Basics of EEG and fMRI

2.1 Introduction

Most of existing multimodal techniques are based on strategies that incorporate the information from one modality into the analysis of the other modality. Thus, it is prerequisite and logic to first introduce the physiological origins of EEG and fMRI, and the imaging principles of both modalities.

2.2 Electroencephalograph

Electroencephalography (EEG) is a collective term for both electroencephalogram and electroencephalograph. Electroencephalogram is the waveform of electrical potentials recorded from a single electrode in direct contact with the scalp surface; electroencephalograph is the spatial map of electrical potentials recorded from a number of spatially distributed scalp-surface electrodes. Therefore, EEG represents the spatiotemporal distribution of scalp electrical potentials.

2.2.1 Physiological origin of EEG

What generates the EEG signal? A general answer is the electrical activity inside the brain. However, answering what specific aspects of the electrical activity serve as the primary EEG sources is not straightforward.

The EEG signals reflect the extracranial electrical field arising from the neuronal currents throughout the entire brain. For a given electrode, the recorded electrical potential is the weighted sum of neuronal current sources at every possible location inside the brain. Each weighting factor depends largely upon the proximity of the corresponding current source to the electrode location as well as the current direction relative to the head geometry.

EEG is a global index of brain electrical activity. As such, we imply that EEG reflects the synchronized electrical behavior of an assembly of neurons within a certain region or even across regions. Asynchronized electrical activities have little contribution to the observed EEG signal, because their random consequences are virtually cancelled.
out when summed over locations. Therefore, the feature of large-scale synchrony restricts the aspects of brain electrophysiology that primarily contribute to EEG.

**Figure 2.1** Electrophysiological behavior of a single neuron primarily exhibits in two forms: the neuronal spiking activity and the synaptic activity. The spiking activity is a train of action potentials that propagate along the axon. The synaptic activity modulates the post-synaptic potential by releasing neurotransmitters to act upon the ion channels at the post-synaptic neuron. Note that the synaptic process often behaves as a low-pass filter, resulting in a lower frequency of the post-synaptic potential than that of the spiking activity. (This figure is modified from a figure in www.answers.com)

The brain electrophysiological process takes place in a variety of spatial and temporal scales. Here, let us start our discussion from the level of a single neuron. The electrophysiological activity of a single neuron is mainly represented in two forms: the neuronal spiking activity and the post-synaptic potential (PSP), as illustrated in Fig. 2.1.A and 2.1.B respectively. The neuronal spiking activity refers to the action potentials that fire along the axon. The spiking activity reaches and accumulates at the synapse through which one neuron connects to another. When the accumulated spiking activity is beyond
a certain threshold, it controls the amount of chemical transmitters released from the synapse. The released neurotransmitters further control the gating of the ion channels at the post-synaptic neuron and hence modulates the PSP. For the spiking activity above the threshold, the synapse effectively behaves as a low-pass filter, which is attributable to its accumulative effect and the delay due to the electrochemical conversion. Under such circumstances, the pre-synaptic spiking activity may be temporally correlated with the PSP. However, for the sub-threshold spiking activity, such correlation does not necessarily exist.

EEG is insensitive to action potential sources due to the bidirectional currents with short signal duration. When the action potential propagates along the axon, the associated electrical currents flow in opposite directions in both intracellular and extracellular spaces (see Fig. 2.2). The neuronal spiking activity is only observable when the recording electrode is placed at a close vicinity to the neuron. When the electrode is relatively distant from the neuron (such as on the scalp surface), the electrical potentials generated by opposing currents are virtually cancelled out. Moreover, the brief duration of the action potential requires a very strong degree of synchrony with a great temporal precision in order to be measurable from the scalp surface. However, high-frequency components (>120Hz) corresponding to the brief duration of neuronal spikes are highly underrepresented in EEG.

In fact, the primary EEG sources are the synaptic currents flowing through the apical dendrites of the pyramidal neurons located within the cortical gray matter (Eccles 1951; Nunez 1985; Dale and Sereno 1993; Kandel et al. 2000; Baillet et al. 2001). The columnar organization of the apical dendrites of large pyramidal neurons makes them feasible to form synchronized regional synaptic currents (Nunez and Srinivasan 2005). For more details, see the following sub-section. Moreover, some observations have been reported in as early as 1960’s that the frequency spectrum of the PSP is close to that of the scalp potentials (Klee et al. 1965; Creutzfeldt et al. 1966).
Figure 2.2 Current flows in both intracellular and extracellular spaces associated with the action potential propagation along a myelinated axon. From the active area, the intracellular currents flow to the neighboring resting areas along the propagation direction as well as the neighboring repolarized areas along the anti-propagation direction. The currents continue in the extracellular space and flow from the resting and repolarized areas to the active area. As a result, both the intracellular and extracellular currents are bidirectional. This figure is adapted from (Malmivuo and Plonsey, 1995)

However, the extracranial potentials measured with EEG have much smaller amplitudes and lower frequencies than the PSP recorded intracranially, such as local field potentials (LFP) and electrocorticograms (ECoG). This discrepancy is due to the different spatial scale to which each recording technique is applied. In general, the low-frequency components tend to be synchronized in a larger spatial scale, whereas the synchronized high-frequency activity is often confined to a smaller scale (Nunez and Srinivasan 2005). In Fig. 2.3, we graph the electrophysiological recording techniques with respect to their spatial scale and frequency range.
Figure 2.3 Spatial scale and frequency range of six electrophysiological recording techniques. Noninvasive EEG and MEG are outlined in blue. The other four invasive techniques (i.e. ECoG, LFP, MUA and SUA) are outlined in red.

### 2.2.2 EEG Forward Problem

In this subsection, we review the methods for solving a so-called EEG forward problem. The EEG forward problem deals with 1) how to model synaptic currents within the brain volume, and 2) how to model the head volume conduction process in order to quantitatively link current sources with scalp potentials.

**Current Source Model**

As discussed in the previous subsection, the primary EEG sources are the post-synaptic currents flowing through the apical dendritic trees of cortical pyramidal cells. When the neurotransmitters act upon the post-synaptic neuron, they control the ions to flow into (or out of) the intracellular space, effectively forming an extracellular current sink (or source). The ionic flows continue in the intracellular space until exiting (or re-entering) at the other side of the dendrite, effectively forming an extracellular current source (or sink). As shown in Fig. 2.4, such an electrophysiological process, when
viewed from a location on the scalp surface that is relatively remote to where the process takes place, can be physically modeled as an electrical current dipole composed of a pair of current source and sink with infinitely small inter-distance.

**Figure 2.4** Illustration of the post-synaptic electrophysiological process. For a pyramidal cell assembly, such a process can be well represented by a current dipole generating the extracellular electrical field. This figure is adapted from (Baillet et al. 2001).

Typically, distributed current source models are employed to represent the whole-brain bioelectric activity. The brain activity with any distribution of bioelectric currents can be approximately represented by a source model consisting of thousands of current dipoles that are evenly placed within the entire brain volume. At each location, three orthogonal dipoles are used, and the weighted combination of them is capable of representing the regional current flow in an arbitrary direction.

In addition, we can also utilize the brain anatomical information to constrain the current source space to the cortical gray matter due to its dominant presence of large pyramidal cells. Such anatomical constraints can be obtained from existing structural neuroimaging modalities. Particularly, T1-weighted MRI provides high spatial resolution and great contrast, which allow us to differentiate the cortical gray matter from the white-matter and the cerebrospinal fluid (CSF). We may further choose to constrain the current source orientations to be perpendicular to the cortical surface, because of i) that the
columnar organization of neurons within the cortical gray matter constrains the regional current flow in either outward or inward normal direction with respect to the local cortical patch (Dale and Sereno 1993), and ii) that the gray-matter thickness (about 2-4 mm) (Fischl and Dale 2000) is much smaller than the “source-to-sensor” distance (Nunez and Srinivasan 2005).

All of the above source models are often referred to as distributed current density models. Physically, any bioelectric source activity can be represented by a continuous distribution of primary current source density. Mathematically, both current density and current dipole share the identical form of equations for computing the extracellular potential. A discrete current dipole distribution serves to approximate the continuous current density distribution. Distributed current density models can be further categorized as the volumetric current density (VCD) model and the cortical current density (CCD) model, depending on how the MRI-derived anatomical information is applied to constraining the current source locations and orientations. The VCD and CCD models constrain the distributed dipoles to the brain volume and the brain surface, respectively.

When the brain electrical activity is confined to a few focal regions, one may prefer to use a source model only with a few discrete dipoles, rather than a more general distributed source model. Multiple dipole sources can be localized via nonlinear optimization algorithms (Kavanagh 1978; Scherg and von Cramon 1985; He et al. 1987; Cuffin 1995; Uutela et al. 1998; Musha and Okamoto 1999) or sub-space scanning procedures (Mosher et al. 1992; Mosher and Leahy 1999; Sekihara et al. 1997; Xu et al. 2004; Ding and He 2006). However, the assumption of focal sources is not always valid, or known a priori. Even if valid, the appropriate number of dipoles is often difficult to be determined from the EEG data alone (Bai and He 2005, 2006).

**Volume Conductor Model**

The electrical field arising from post-synaptic neuronal currents is governed by the Maxwell equation in a quasi-static condition (Plonsey 1969; Malmivuo and Plonsey 1995), as expressed by Eq. (2.1)

\[ \nabla \cdot \mathbf{j}(\mathbf{r}) = \nabla \cdot \sigma(\mathbf{r}) \nabla \phi(\mathbf{r}) \]  \hspace{1cm} (2.1)
where \( \mathbf{j} \) stands for impressed current density, \( \phi \) stands for electrical potential, \( \sigma \) stands for electrical conductivity and \( \mathbf{r} \) denotes location. Eq. (2.1) is subject to the boundary conditions that the electrical potential \( \phi \) and the normal component of the conductive current density \( \sigma(\mathbf{r}) \nabla \phi(\mathbf{r}) \) remain continuous across the boundary between any tissue/media compartments of different conductivities.

Based upon this governing equation, the EEG forward problem is to compute the distribution of \( \phi \) on the scalp surface, given any known distribution of \( \mathbf{j} \) inside the brain as well as the conductivity values throughout the head volume.

The forward solution to Eq. (2.1) cannot be derived analytically except for in a few simplified (and perhaps over-simplified) cases, which assume piece-wise homogeneous conductivity distribution within concentric three-shell (Rush and Driscoll 1969; Perrin et al. 1987; Wang and He 1998) and four-shell (Cuffin and Cohen 1979; Sun 1997) spherical models. More accurate numeric forward solutions can be obtained by means of the boundary element method (BEM) (He et al. 1987; Hämäläinen and Sarvas 1989; Fuchs et al. 1998; Kybic et al. 2005), finite element method (FEM) (Yan et al. 1991; Weinstein et al. 2000; Zhang et al. 2004) or finite difference method (FDM) (Neilson et al. 2005), all of which allow us to incorporate the realistic geometries and conductivities of the head. As such, these numerical techniques have to utilize the anatomical information provided by other structural imaging modalities, particularly T1 and proton density weighted MRI, in order to segment different brain tissues (i.e. the gray matter and the white matter) and head structures (i.e. CSF, skull and scalp).

In addition, we may further improve the accuracy of the forward solution by incorporating the anisotropic conductivity of the cerebral white matter. The white-matter anisotropy is caused by the bundled axon fibers that restrict the direction of ionic movements (Wolters et al. 2006). Recently, we and other labs have developed methods to extract the anisotropic conductivity distribution from diffusion tensor magnetic resonance imaging (DT-MRI) (Wang et al. in press; Tuch et al. 2001; Wolters et al. 2006). Our approach is based on the fact that the transportation processes for both water molecules and electrical charges are restricted by the same microscopic structure (Basser et al. 2000).
1994), which is composed of multiple compartments of axons and glial cells aligning along different orientations (Wang et al. in press).

Based upon the above modeling and computation methods, the scalp potential measurements can be linked with the current sources through a linear transform matrix.

\[ \mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{b}(t) \]  \hspace{1cm} (2.2)

where \( \mathbf{A} \) is an \( N_x \)-by-\( N_s \) matrix (\( N_x \) is the number of EEG sensors, \( N_s \) is the number of current sources), \( \mathbf{s}(t) \) is an \( N_s \)-by-1 vector, \( \mathbf{x}(t) \) and \( \mathbf{b}(t) \) are \( N_x \)-by-1 vectors. Note that the \( i \)-th column of \( \mathbf{A} \) (for any \( 1 \leq i \leq N_x \)) represents the scalp potential distribution generated by the \( i \)-th unitary dipole in the source model.

### 2.2.3 EEG Inverse Problem

As discussed in the previous subsection, solving the EEG forward problem is to quantitatively link the source model with scalp potential measurements. Such a linkage is linear, leading to a unique solution of scalp potentials that can be computed from any known distribution of bioelectrical current density inside the brain. Conversely, the electrical source imaging is the inverting process of the forward computation. That is, computing the distribution of unknown bioelectrical current density from measured scalp potentials. Therefore, the electrical source imaging is also referred to as the EEG inverse problem. Fig. 2.5 illustrates both forward and inverse processes in relation to the current source model, the volume conductor model and the EEG signals.

As opposed to the uniqueness of the forward solution, the EEG inverse solution is non-unique because the number of unknowns is usually much larger than the number of recording channels. The inverse problem is highly ill-posed, since there are infinite sets of source configurations that are equally compatible with the EEG data. In addition, the ill-posedness also implies a high condition number of the transfer matrix, whose pseudo-inverse matrix largely amplifies the effect of the recording noises.
Figure 2.5 Illustration of the forward and inverse problems in relation to the current source model, the volume conductor model and the EEG recording and mapping.

In order to obtain a unique inverse solution, one has to apply additional constraints to the desirable solution. For instance, one may specify the optimal solution as the most energy efficient one among those fitting equally well with the data. In a noise-free condition, this constraint leads to the linear least-squares source estimate. With recording and background noises, this constraint should be incorporated as a minimum norm side constraint, giving rise to the minimum norm estimate (MNE) (Hämäläinen and Ilmoniemi 1984). Other variations of the MNE include the lead-field normalized weighted minimum norm (WMN) (Wang et al. 1992), low-resolution brain electromagnetic tomography (LORETA) (Pascual-Marqui et al. 1994), and etc. The common feature shared by these algorithms is the linearity of the inverse solution, meaning that the inverse solution can be obtained through transforming the measurements through a linear system (or inverse matrix).

The linearity of the inverse computation also allows us to characterize the statistics of source estimates. In practice, it is often feasible to estimate the noise statistics in the sensor space (e.g. by using data during a signal-free period). As the inverse matrix
also transforms the noise to the source space, it is a simple linear computation to further obtain the noise statistics in the source space. By normalizing the source estimates with respect to the corresponding noise sensitivity, one can assess the statistical significance of the inverse solution and obtain a map of source estimate statistics. Although this line, Dale et al. and Pascual-Marqui et al. have developed two statistical functional mapping techniques, known as dynamic statistical parametric mapping (dSPM) (Dale et al. 2000) and standardized LORETA (sLORETA) (Pascual-Marqui 2002), based on the MNE and LORETA algorithms respectively.

The linear inverse solutions often end up with low-resolution images of current density or its statistics. One way to improve the spatial resolution is to generate images with focal source distribution by iteratively repeating the linear inverse computation (Gorodnitsky et al. 1995, 1997; Yao and He 2001; Liu et al. 2004, 2005). For each step during the iteration, the inverse solution linearly computed from the previous step is used as the weighting factors to constrain the source estimates in the current step. As such a recursive process continues till convergence, the estimated source distribution tends to be more focalized. Moreover, these iterative methods have been suggested to be equivalent to a nonlinear inverse solver that minimizes the cost function formulated with L-p norms for both the data fitting term and the side constraint (Rao and Kreutz-Delgado 1999). This is in line with other nonlinear inverse algorithms that specifically utilize the solution constraint other than the L-2 norm, such as the minimum current estimation (MCE) (Uutela et al. 1999), L1 norm (Matsuura and Okabe 1995; Huang et al. 2006) and L-p norm (Beucker and Schlitt 1996).

The imaging performance of all linear and nonlinear EEG inverse algorithms varies from case to case, depending on the nature of brain electrical activity associated with the sensory and cognitive functions of interest. The methodological choice is up to the users’ own judgment as to whether the activity is sparse or extended, stationary or non-stationary, etc. Although the evaluation of some typical algorithms have been done in a few experimental and simulation cases (Yao and Dewald 2005; Bai et al. 2007; Grova et al. 2006), a general conclusion or agreement regarding a single optimal source
imaging approach, if possible at all, has not been reached (Fuchs et al. 1999; Michel et al. 2004).

2.3 Functional Magnetic Resonance Imaging

In the past decade, our ability in functional brain mapping has been most significantly advanced by the invention of fMRI (Ogawa et al. 1992; Kwong et al. 1992; Bandettini et al. 1992). Although fMRI is an exciting and powerful imaging technique with numerous neuroscience applications, many fundamental aspects of this technique remain unclear and under active research.

2.3.1 MRI Physics

An atomic nucleus possesses an intrinsic angular momentum called spin. A spinning nucleus behaves as a tiny magnet. All nuclei that contain odd numbers of protons or neutrons (e.g. $^1$H, $^{17}$O, $^{19}$F and $^{31}$P) have a net spin or magnetic moment (Rabi et al. 1938). $^1$H is the primary focus of most MRI methodologies and applications due to its natural abundance in biological systems (such as the brain).

A bulk of spins exhibits a phenomenon when it is subject to certain external static and oscillating magnetic fields (Rabi et al. 1938; Bloch et al. 1946; Purcell et al. 1946). This physical phenomenon has been known as nuclear magnetic resonance (NMR). In a static magnetic field ($B_0$), all of the protons within a volume precess around the $B_0$ field axis. The precession frequency, typically in the radio frequency (RF) range, is proportional to $B_0$. Therefore, the net bulk magnetization ($M_0$) is non-zero along $B_0$ (i.e. the longitude direction) but effectively zero along directions perpendicular to $B_0$ (i.e. the transverse direction) due to the random and asynchronized precession phases. If being applied a second transverse magnetic field oscillating in the precession frequency, all of the spins tend to precess synchronously, producing a rotating transverse magnetization ($M_{xy}$) and a variable longitude magnetization ($M_z$). Note that the direction of the net magnetization relative to the $B_0$ axis, as measured by the flip angle in the rotating frame, depends on the duration and strength of the excitation RF pulse. When the excitation pulse is turned off, the spins gradually return to the equilibrium state. Both of the time courses with which $M_z$ recovers and $M_{xy}$ decays are approximately exponential, with the
corresponding time constants known as $T_1$ and $T_2$ respectively. The decaying and oscillating transverse magnetization is measurable through a receive coil that senses the electrical current induced by the magnetic flux change.

One of the most important developments in MRI is the ability to localize the MR signal in the three-dimensional brain volume through the spatial-frequency encoding (Lauterbur 1973). This is accomplished by using the magnetic gradients in which the field strength changes gradually along an axis. There are typically three orthogonal gradients for the slice-selection, phase-encoding and frequency-encoding respectively. The first gradient coordinates with the RF excitation frequency to selectively enable a single slice to be on resonance. The other two gradients encode the image of the selected slice into a so-called k-space. The k-space and image space are linked by a pair of inverse and forward Fourier transformation.

2.3.2 BOLD fMRI

In terms of functional brain mapping, MRI techniques have been used to map regional hemodynamic changes induced by elevated neural activity. Most notably is the gradient-echo MRI, which creates $T_2^*$ weighted images reflecting local susceptibility changes produced by changes in the deoxyhemoglobin content in blood vessels. This effect is called the blood oxygenation level dependent (BOLD) contrast (Ogawa et al. 1990).

The BOLD contrast is attributable to some basic biophysical properties of hemoglobin and the susceptibility effects in NMR. Deoxyhemoglobin is paramagnetic and its concentration inversely depends on the blood oxygenation level (Pauling and Coryell 1936). The paramagnetic property of the blood causes the bulk susceptibility difference between a blood vessel and the surrounding neural tissue. The intravoxel susceptibility difference produces the local magnetic field inhomogeneity, which in turn results in resonance frequency shifts and additional phase dispersion for extra-vessel molecules. This BOLD-dependent effect is pronounced in gradient echo MR images where higher oxygenation level leads to larger MR signals (Ogawa et al. 1990). In addition, echo-planar imaging (EPI) sequence further provides a remarkable temporal resolution to record dynamic BOLD signals in real-time (Turner et al. 1991). In the first
three pioneer studies, the BOLD fMRI has been independently demonstrated to be capable of noninvasively mapping neurophysiologic changes in response to the visual stimulation (Ogawa et al. 1992; Kwong et al. 1992) or the motor task (Bandettini et al. 1992).

In most fMRI studies, brain regions involved in carrying out a task or responding to a stimulus show the BOLD signal increase relative to the resting state. Such BOLD signal increase is known as the positive BOLD response (PBR). The PBR is believed to result from the over-compensation for the oxygen usage by the cerebrovascular system, which causes an influx of oxygenated blood in excess of oxygen consumption by regional brain activation (Fox and Raichle 1986; Fox et al. 1988). In fact, the BOLD signal reflects the combined effect of cerebral blood flow (CBF), cerebral blood volume (CBV) and cerebral metabolic rate of oxygen (CMRO₂). Several models have been proposed to account for the contribution to the BOLD signal from neural activity-induced changes in CBF, CBV, CMRO₂ (Ogawa et al. 1993; Buxton et al. 1998, 2004; Mandeville et al. 1999; Friston et al. 2000; Aubert and Costalat 2002). In short, the BOLD signal is contributed positively by CBF and negatively by CBV and CMRO₂ changes.

Understanding the neurophysiologic origin of BOLD is crucial to the interpretation of BOLD signals in terms of neuronal events. There have been debates regarding whether the BOLD signal arises from neuronal inputs or outputs (i.e. synaptic or spiking activities). Early studies have found a linear relationship between human fMRI response and primate neuronal spike activity in the visual areas (V5 and V1) (Rees et al. 2000; Heeger et al. 2000). In contrast, important findings obtained from simultaneously recorded BOLD and intracranial electrical signals on primates suggest that the BOLD response is more correlated with the power of local field potential (LFP), which represents the synchronized synaptic inputs of a given neural population (Logothetis et al, 2001). Since then, evidence has increasingly suggested that the BOLD fMRI signal primarily reflects synaptic activity rather than neuronal spike activity (Arthurs and Boniface 2002; Lauritzen and Gold 2003; Martindale et al. 2003). This is also in light of the fact that the hemodynamic response is driven by the metabolic energy demand, nearly all of which is imposed by synaptic activity instead of action potential firing (Mathiesen
et al. 1998; Arthurs and Boniface 2002; Nair 2005). The observed correlation between the BOLD signal and the neuronal spiking rate (Rees et al. 2000; Heeger et al. 2000) might be explained by the post-synaptic current flow in correlation with the spike activity of the pre-synaptic neurons (Heeger and Ress 2002; Arthurs and Boniface 2002). In addition, the energy contained in the synaptic current flow (proportional to the square of current density or LFP) can be further thought of as the physical energy correlate of metabolic energetics.

2.3.3 fMRI Statistic Parametric Mapping

The vast majority of fMRI applications use fMRI for functional localization of the brain regions engaged in specific sensory processing or cognitive functions (Logothetis 2008). FMRI as a brain mapping tool depends on experimental designs, data analysis methods and an explicitly or implicitly assumed relationship between BOLD signals and neural responses. Fig. 2.6 summarizes the principle of fMRI.

**Figure 2.6** Principle of BOLD-contrast fMRI. The regional neuronal activity alters the local CBF, CBV and CMRO2, which collectively leads to changes in the blood oxygen level. The increase of oxygen level,
meaning the decrease of local field inhomogeniety, produces a longer T2 or T2* relaxation and therefore larger MR signals. The frequency-and-phase encoding technique (e.g. echo-planar imaging) allows for the fast acquisition of a so-called k-space data, which can be transformed to the original image space through Fourier transformation. The signal is most frequently analyzed voxel by voxel, yielding statistic maps indicating regions with significant hemodynamic effects related to external stimuli/tasks or internal events. These regions, arguably define the activated neuronal populations.

FMRI is a comparable technique in nature. An fMRI experiment has to include at least two conditions in contrast. A task (or stimulus) condition places specific demands to the brain, while a user-defined control condition may involve a “baseline” task or a resting state. The signal difference between the task and control conditions is evaluated, typically voxel by voxel, to identify the regions engaged in the task execution. Popular fMRI experimental designs define the alternation of conditions in either a block-design or event-related manner. In line with earlier PET paradigms, the block-design involves prolonged task and control periods, so as to observe sustained BOLD signal changes with a high contrast-to-noise ratio (CNR). The event-related design focuses on the averaged single-trial BOLD response in ways analogous to event-related potentials/fields (ERP/ERF) in EEG/MEG (Buckner et al. 1996; Dale and Buckner 1997).

For both types of study designs, the BOLD signal at an “activated” voxel is expected to change in a way related to the transitions from one condition to another. This rationale allows us to make inference about regionally specific effects in response to the task/stimulus. For experiments with only two conditions, voxel-wise statistical inference may be simply based on a Student’s t-test or period cross correlation (Bandettini et al. 1992). A more general and flexible approach is based on the general linear model (GLM) (Friston et al. 1994; 1995a). Stimulus functions encoding the occurrence of a particular event or experimental state (e.g. boxcar-functions) are convolved with a hemodynamic response function (HRF) to form regressors in the GLM. Fitting the GLM to the data
allows for the estimation of model parameters and the statistic inference against a null hypothesis (i.e. the voxel is not activated). Such inference is classic in statistics, as opposed to more recent methods based on the Bayesian inference which provides the posterior probability that the voxel is activated given the data (Friston et al. 2002). All these methods discussed so far are model-driven in a sense that they require specific assumptions about the time courses of the processes contributing to the measured signals and/or a priori statistical distributions of the signal and the noise. To remove such model dependence, a data-driven method has been implemented using independent component analysis (ICA) (McKeown et al. 1998; Calhoun et al. 2001).
Chapter 3 Simultaneous EEG-fMRI

3.1 Introduction

It is often desirable to acquire EEG and fMRI signals simultaneously in studies that integrate both modalities or investigate the cross-modal relationship. This is as opposed to the individual fMRI/EEG acquisition through separate sessions. Separate acquisition may cause concerns as to whether or not separately acquired multimodal datasets correspond to the identical brain status or process. Discrepancies are expected (or perhaps unavoidable), when studying high-level cognitive functions or neuropathological mechanisms, since the related brain activities are hardly reproducible across different recording sessions.

For instance, the simultaneous fMRI-EEG acquisition is indispensable for mapping the fMRI correlates to interictal epileptic activities, e.g. interictal EEG spikes (Ives et al. 1993; Gotman et al. 2004, 2005, 2006). Under these circumstances, the fMRI response of interest is not manipulated by user-defined external stimuli or tasks; but rather, it is driven by spontaneous biological events. The surface EEG provides the electrophysiological markers for such internal events, providing the timing information for the event-related fMRI analysis. Some other examples are also seen in studies of mapping the fMRI correlates to continuous rhythmic EEG modulations during the resting state (Goldman et al. 2002; Laufs et al. 2003; Moosmann et al. 2003) or sleep (Kaufmann et al. 2006; Horovitz et al. 2008). These studies require high-quality EEG signals within a wide frequency range of interest.

However, the simultaneous fMRI-EEG recording faces great technical challenges. As illustrated in Fig. 3.1, EEG recordings inside the MR scanner are subject to strong interferences from the fMRI environment. The electromotive force is formed in wired loops between electrodes due to the time-varying magnetic fields and/or the motion of EEG leads that change the looped cross-sectional area perpendicular to the static magnetic field (Goldman et al. 2000; Allen et al. 2000). During the fMRI scanning, the time-varying magnetic fields are the dominant sources of artifacts to concurrent EEG recordings. The fMRI scans require radio-frequency (RF) excitation pulses and rapidly
switching magnetic gradients, both of which result in large $dB/dt$ (i.e. the temporal derivative of the magnetic field) and induce voltage artifacts of large magnitude (about 1000 times of the normal EEG magnitude). In addition to the fMRI pulse-induced artifacts, the pulsatile lead movements related to heart beats can cause ballistocardiographic artifacts, which are often several times larger than the neural activity-induced EEG signal (Allen et al. 1998). As the EEG electrodes are in direct contact with the scalp, such artificial current induction may also raise concerns as of the subject safety, since the head tissues surrounding the electrodes may be heated or damaged by the large non-biological currents (Lemieux et al. 1997; Angelone et al. 2004, 2006). From the MRI/fMRI perspective, the fMRI-EEG simultaneous recording may also affect the MRI/fMRI image quality due to the field inhomogeneity introduced by the EEG devices placed inside a bore magnet and the MR susceptibility of scalp-surface electrodes etc (Krakow et al. 2000; Bonmassar et al. 2001a). All of these practical concerns considerably compromise the otherwise favorable simultaneous recording strategy in fMRI-EEG combined researches.

Figure 3.1 EEG artifacts due to the interference from the static and time-varying magnetic fields in the fMRI scanning environment. According to Faraday’s law, the electromotive force that drives the artificial currents depends on the areas of wired loops ($S$), the rate of their changes over time ($dS/dt$), the magnetic field strength ($B$) and its changing rate over time ($dB/dt$).
(dB/dt). Any source leading to changes related to these four factors causes artifacts to EEG recordings.

Active researches have been conducted to resolve the above practical issues to enable the simultaneous fMRI-EEG recording (for reviews, see Ritter and Villringer 2006; Gotman et al. 2006; Laufs et al. 2008). Special efforts have been put forth upon developing the MR-compatible EEG caps and amplifiers (Goldman et al. 2000; Allen et al. 2000), synchronization of EEG sampling with the MR pulse sequence (Mandelkow et al. 2006), EEG lead design and placement (Lemieux et al. 1997; Baumann and Noll 1999; Vasios et al. 2006), cable wiring (Goldman et al. 2000), signal transmitting (Allen et al. 2000; van Audekerke et al. 2000), subject comfort and safety (Lemieux et al. 1997; Benar et al. 2003) and etc. On the software side, signal processing algorithms have been developed to remove the MR gradient artifacts (Sijbers et al. 1999; Allen et al. 2000; Hoffmann et al. 2000; Garreffa et al. 2003; Negishi et al. 2004; Niazy et al. 2005; Wan et al. 2006a) and the cardiac ballistic artifacts (Allen et al. 1998; Goldman et al. 2000; Benar et al. 2003; Ellingson et al. 2004; Kim et al. 2004; Niazy et al. 2005; Srivastava et al. 2005; Nakamura et al. 2006) from the EEG simultaneously recorded with fMRI. For certain sensory stimulation paradigms, interleaved fMRI/EEG acquisition strategies have also been employed, occasionally, to acquire fMRI data during the delayed time windows that are of no interest for the ERP (Bonmassar et al. 1999, 2001b; Anami et al. 2003).

In line with existing developments and investigations, we have performed a series of simultaneous fMRI-EEG experiments with well-controlled paradigms to evaluate the data quality of both fMRI and EEG. In 3.2, we describe the experimental settings and signal processing algorithms employed in our studies. In 3.3, we present the results obtained from three experiments pertaining to the visual evoked potential, steady-state visual evoked potential and alpha-wave power modulation, respectively. In 3.4, we discuss the electrophysiological correlates to the negative BOLD response observed through simultaneous fMRI-EEG recordings in a visual stimulation experiment.
3.2 **Simultaneous fMRI-EEG Recordings**

3.2.1 **Hardware**

In our fMRI-EEG simultaneous experiments, we used a 64-channel MR-compatible EEG system (BrainProduct, Munich, Germany) inside both 3-T and 4-T magnets. Fig. 3.2 shows the key components in the system setup.

![Figure 3.2 Illustration of the experimental setup for simultaneous fMRI-EEG recordings. See text for details.](image)

Carbon ring flat electrodes were fixated on the EEG cap in order to minimize the subjects’ discomfort. An inline 5-kΩ current-limiting resistor was embedded at the junction between each electrode and the connecting wire to drastically reduce the magnitude of induced currents towards the amplifier and avoid burning the surrounding...
head tissues (Lemieux et al. 1997). Non-magnetic copper wires connecting electrodes on the cap were twisted and bundled. The twisted wiring enabled the induced currents to flow in opposite directions and thus largely cancelled each other; the bundled wires helped avoid the loop formation. A unipolar electrode was placed on the subject’s back to record the ECG signal. The wire connected to the ECG lead was merged with other scalp EEG leads into a wire bundle. All of the EEG and ECG channels were referenced to FCz. An occipital electrode was used as the ground.

From the EEG cap, two branches of wire bundles were connected to two MR-compatible adapters, where they were converted to two short and flat ribbon cables. The ribbon cables were in turn connected with two 32-channel shielded and non-magnetic preamplifiers. These amplifiers were 16-bit with the resolution of 0.5μV/bit, enabling a dynamic range of ±32.767 mV. Through the pre-amplifiers, the signals were locally amplified, low-pass filtered (<250 Hz), and digitized at 5K Hz. The sampling clock on the amplifiers was synchronized with the clock driving the MR scanner’s gradient switching system to ensure the phase synchrony of gradient artifacts.

The digitized EEG signals were transmitted to an integration unit in the control room via long optical fibers. The EPI volume onsets were also recorded by connecting the TTL output from the MRI scanner to the integration unit. The integration unit also received inputs pertaining to the onsets of stimuli or tasks from the stimulation device. With these input signals being merged together, a USB cable was used to transmit the signals to a PC, wherein the signals and triggers were monitored in real-time.

3.2.2 Signal Processing

The raw EEG data acquired with fMRI were largely contaminated by gradient artifacts and ballistocardiographic artifacts, both of which needed to be removed prior to any other conventional EEG or ERP analyses. In what follows, we discuss the key features of these artifacts, based on which artifact reduction algorithms were chosen to obtain clear EEG signals, hopefully of equivalent quality as the EEG acquired alone.
Gradient Artifacts

The time-varying magnetic fields involved in the fMRI acquisition induce strong artifacts that dominate the real EEG signals. The fMRI scanner-induced artifacts have been commonly known as gradient artifacts (GA).

Since the identical pulse sequence was repeatedly used to acquire a time series of EPI volumes, the GA were periodic and stable over volumes. The regularity and periodicity were self-evident from the raw EEG recordings during continuous fMRI scans. In fact, the GA pattern was closely correlated with the EPI pulse sequence. For the example shown in Fig. 3.3, ten EPI slices were acquired within each volume of 1-sec TR, resulting in ten evenly spaced artifacts. The artifact pattern corresponding to a single EPI slice contained a large low-frequency oscillatory signal followed by high-frequency oscillations of smaller amplitude. The former corresponded to the slice-selection gradient while the latter corresponded to phase-encoding and readout gradients.

Figure 3.3 Gradient artifacts phased locked to the EPI volume and slice acquisition. The 1st row shows the TTL triggers, indicating the onsets of EPI volume acquisitions, with a volume TR of 1 sec. The 2nd row shows the raw EEG recordings during continuous fMRI scans. The 3rd row shows
the GA within a time window during which a single EPI volume was acquired. The 4th row shows the GA during the acquisition of a single EPI slice, which is temporally correlated with the MR gradients.

The regularity and periodicity of GA allowed us to isolate the GA from the raw EEG recordings, so as to restore the GA-free EEG using an average template subtraction algorithm (Allen et al. 2000). In addition to the virtually constant GA that were phased locked to the fMRI image acquisition, the raw EEG recordings also contained the non-GA components including the biological signals of interest, the recording noise and other kinds of MR-related artifacts. Since the non-GA signals were temporally independent upon the scanner triggers, averaging the raw EEG data segmented with respect to the scanner triggers resulted in a template signal exclusively of GA, while the averages of all non-GA components should equal zero. It followed that the GA-free signals could be obtained by subtracting the averaged GA template from the original raw recordings. Fig. 3.4 shows an example of the averaged GA template represented in time and frequency, as well as the EEG before and after subtracting the GA template. Clearly, the average template subtraction algorithm effectively removed the gradient artifacts.
Figure 3.4  A) GA template; B) frequency spectrum of the GA template; C) EEG data before removing GA; D) EEG after removing GA.

Cardiac Ballistic Artifacts

However in Fig. 3.4.D, we could still observe a large periodic artifact signal that existed even after removing the GA. This was because the EEG recordings inside the MR scanner were also subject to the artifacts originating from the interference between heart beats and the static magnetic field. Such artifacts have been known as the ballistocardiographic artifact or cardiac ballistic artifacts (CBA).

The source of CBA was more complex than that of GA. One possible source was the electrode movements adjacent to blood vessels as a result of the pulsatile blood flows pumped from the beating heart. Such pulsatile electrode movements likely changed the areas of wired loops that interfered with the main magnetic field (Allen et al. 1998). It was also possible that the ionic molecules carried by the blood gave rise to a so-called Hall-voltage underlying the CBA (Ellingson et al. 2004).

Regardless of the exact sources, the CBA appeared in the scalp EEG after a relatively fixed time delay with respect to each heart beat, and remained relatively stable over time. Therefore, we employed an average template subtraction algorithm (Allen et al. 1998), similar to the GA removal algorithm. The algorithm started with setting temporal markers for every heart beat using the recorded ECG signal. A single-beat ECG recorded inside the MR scanner contained the P-wave, QRS complex and augmented T-wave. Since the QRS complex evolved much faster than the P- and T-waves, the R-peak possessed the best temporal precision to serve as a reliable marker for every heart beat. The R-peaks were detected automatically by searching for the maximum correlation between the recorded ECG in a sliding window and a single-beat reference ECG.

Fig. 3.5 shows the examples of the ECG with the identified R-peaks and the scalp EEG from a single channel with the CBA. Superimposed to the biological EEG signals, the CBA were repeated with a relatively stable pattern following every single heart beat. We also noticed that the time points with the largest CBA amplitudes in the scalp EEG were consistently delayed relative to the R-peaks. Such a delay remained almost constant,
and represented the time needed for the blood pumped out from the heart to flow through the scalp so as to exert the largest effect upon the EEG. We detected the averaged delay, referred to as the CBC delay, as the time difference between the R-peak and the maximum global field power for those channels containing the CBA.

**Figure 3.5** CBA reduction. The 3rd row shows the ECG with identified R-peaks. The 2nd row shows the scalp EEG with the CBA. The largest CBA is delayed from the R-peak. The CBA template is obtained from averaging adjacent 10~20 epochs segmented with respect to the delayed R-peaks. The CBA template is further subtracted from the EEG data to reduce the artifact. The 4th row shows the EEG before and after the CBA reduction.

As illustrated in Fig. 3.5, we further obtained averaged CBA templates for each EEG channel. This was achieved by first segmenting the EEG with respect to the R-peaks plus the detected CBC delay. Averaging these segments was expected to yield a template exclusively representing the CBA, assuming the neural activity-induced EEG was independent upon the heart beats. However, the CBA were less stable than the GA due to
the variable heart rate etc. To account for the variability of the CBA, we computed a series of time-dependent CBA templates, each of which was obtained from a moving average of the adjacent 19 segments. The biological EEG of interest was remained after subtracting the CBA templates from the recorded EEG. Fig. 3.5 also shows the comparison between the EEG data before and after reducing the CBA. Obviously, the CBA were significantly reduced or removed.

**Independent Component Analysis**

Although the EEG data quality after the above GA and CBA reduction were often comparable to the EEG recorded outside the fMRI environment, some residual artifacts may still be present. The residual artifacts were caused by the non-phase-locked or drifted GA as well as the non-stationary CBA, all of which could not be completely removed by the template subtraction algorithms described above. These residual artifacts may affect, in various degrees, the ensuring ERP, frequency or time-frequency EEG analyses.

Therefore, we applied the independent component analysis (ICA) to further remove the residual artifacts in the processed EEG data. When ICA was applied to the continuous multi-channel EEG, it decomposed the data into independent components (ICs), which were composed of spatial components and corresponding time courses that were independent to each other. ICA allowed us to isolate different EEG sources including the neural signals of interest, the recording noise, or the residual artifacts. By examining the spatial, time and frequency features of each IC, judgment was made as to whether or not each IC corresponded to noise or artifacts and thus should be removed.

Specifically, three criteria were employed to select artifact or noisy ICs. 1) The spatial component fails to satisfy the discrete Picard condition (Hansen 1990). A failure of Picard condition suggests the spatial component is not smooth enough to be produced by any current distribution inside the brain, as the spatial smoothness is imposed by the head volume conduction (see 2.2.1); 2) The time course of an IC contains obvious outliers that represents a sudden and large signal change likely due to the head movement etc; 3) The frequency spectrum resembles the typical spectrum of the GA template. An artifact or noise IC may satisfy more than one of these three criteria.
Fig. 3.6 shows some examples of the artificial and noise ICs, identified according to the above criteria. Both of the components in 3.6.A) and 3.6.B) had noisy spatial components that failed to satisfy the discrete Picard condition, and had frequency spectra similar to that of the averaged GA template (see Fig. 3.4.B). The component in Fig. 3.6.C) was regarded as the noise artifact because its time course contained several outliers.

![Figure 3.6](image)

**Figure 3.6** Three noise or artifact ICs, identified according to three criteria applied to the spatial distribution, time course and frequency spectrum of the decomposed ICs.

After the noise and artifact ICs were identified, we removed them and transformed the remaining ICs back to multi-channel EEG signals, using a linear transformation that was precisely the inverting process of the IC decomposition.
3.3 Evaluation of fMRI-EEG Simultaneous Recordings

The techniques described in 3.2 hold promise to enable simultaneous fMRI-EEG recordings. However, whether the data quality would be sufficient and reliable enough for specific EEG-fMRI studies is still open to question. To address this question, we have performed a series of simultaneous fMRI-EEG experiments, in which visual evoked potential (VEP) (Im et al. 2005; Bai et al. in revision), steady-state visual evoked potential (SSVEP) and spontaneous alpha-wave modulation were assessed. These experiments served as control experiments exemplifying typical fMRI-ERP or fMRI-EEG applications.

3.3.1 Visual Evoked Potential

Five healthy human subjects (aged 18~29, males and right-handed) participated in the fMRI-VEP experiments with written consent. A full rectangular checkerboard pattern (6×6 black-and-white contrast; average luminance: 20 cd/m²; flickering at 2 Hz) was delivered to the subjects through a LCD monitor outside the MRI scanning room, or back mirrored through a DLP projector inside the scanner. The horizontal and vertical visual angles of the checkerboard pattern were 40° and 30°, respectively. The subjects were instructed to fixate at a cross mark at the center of the screen during the experiment. Three sets of EEG data were acquired (outside the MRI scanner, inside the scanner without fMRI scanning, and inside the scanner during fMRI scanning), using a 64-channel MR compatible EEG system (BrainAmp MR 32 Plus, BrainProducts, Germany). Both structure MRI (sMRI) and fMRI data were collected using a 3T MRI system (Siemens Trio, Siemens, Germany). The whole-head T1-weighted MR images (FOV 256 mm × 256 mm, in-plane resolution 1 mm × 1 mm, 1-mm slice thickness) were acquired using the Turboflash sequence (TR/TE=20 ms/5 ms). The T2*-weighted fMRI data were acquired from ten axial slices (matrix size 64 × 64, in-plane resolution 4 mm × 4 mm, 5-mm thickness) covering the visual cortex using the gradient-echo echo planar imaging (EPI) sequence (TR/TE=1000 ms/ 35 ms). Two 60-sec stimulation periods were alternated with three 40-sec resting periods with only the fixation on a dark gray
background. The period cross-correlation method (Bandettini et al. 1992) was applied to obtain the fMRI activation map with the correlation coefficient $\geq 0.5$. 

Fig. 3.7 shows the fMRI activation maps overlaid on the anatomical MR images for 2 representative subjects. The visual areas responding to the visual stimuli were correctly highlighted. No distortion to either the MRI or fMRI data was found, suggesting that the EEG electrodes and other devices do not affect the MR image quality.

![Figure 3.7](image)

**Figure 3.7** fMRI activation maps (for 2 subjects) resulting from the MRI/fMRI data during the concurrent EEG acquisition.

The EEG data were first preprocessed using the methods described in 3.2 before the conventional VEP analysis. Around 400 trials were segmented with respect to the stimulus onsets. The EEG segments were subject to the linear de-trend, baseline correction and then averaged to yield the VEP signals. We compared the VEP waveforms recorded outside the MRI scanner, inside the scanner without or with concurrent fMRI scans. The VEP waveforms at two occipital electrodes (O1, O2) for a single subject are shown in Fig. 3.8 (a) and (b), respectively. The VEP waveforms were consistent with the previously reported VEP morphology (Di Russo et al. 2005; Bonmassar et al. 2001). More importantly, the VEP recorded in three different conditions closely resembled each other, with slight differences in the P100 amplitude and latency. These results suggest the EEG data simultaneously recorded with fMRI has reliable, reproducible and high quality for the conventional ERP analysis.
3.3.2 Steady-State Visual Evoked Potential

As rhythmic EEG signals are often of interest in many EEG studies, we also tested whether certain frequency components could be restored from the EEG data simultaneously acquired with fMRI. For this purpose, we employed the similar pattern-reversal checkerboard stimuli as what we used in 3.3.1, except that the reversing frequency was 9 Hz instead of 2 Hz. The 9-Hz stimuli lasted for 3 minutes and evoked sustained rhythmic EEG signals at 9 Hz, known as the steady-state visual evoked potential (SSVEP).

Fig. 3.9 shows the power spectrum at an occipital channel Oz and the spatial distribution of the 9-Hz EEG power. As expected for the SSVEP, the Oz power spectrum had a peak precisely at the stimulus frequency as well as its 2\textsuperscript{nd} harmonic frequency. Spatially, the 9-Hz power was primarily distributed at the middle occipital lobe, which corresponded to the location of the primary visual cortex.
Figure 3.9 A) Power spectrum for the SSVEP recorded at Oz, B) Scalp map of the 9-Hz SSVEP power.

3.3.3 Resting State Alpha-Wave Modulation

We further tested whether the alpha-wave modulation induced by an eye-open-eye-close task could be observed from the recorded EEG during concurrent fMRI scans. It has been widely known that the alpha wave increases (or decreases) when a subject closes (or opens) his/her eyes. This experiment was of great interest, because it resembled a class of rhythmic modulation phenomena that may occur spontaneously at rest (Goldman et al. 2002; Laufs et al. 2003; Moosmann et al. 2003), in response to certain sensory stimulation or task (Pfurtscheller and Lopes da Silva 1999), or during and after the epileptic seizure etc.

In the eye-open-and-close task, we instructed the subjects to close/open his/her eyes in a self paced manner. The experiments were performed with a 3-T MR scanner inside a dimly lighted environment. The fMRI and EEG were simultaneously and continuously recorded during the entire experiment.

Fig. 3.10 shows the typical results from a single subject. From the signals at the occipital electrodes, we computed the EEG power as a function of time and frequency, namely the time-frequency representation (TRF). As shown in Fig. 3.10.A, the occipital power within the alpha-band (8~12 Hz) increased when eyes were closed and decreased when eyes were opened. Similarly, the map of BOLD responses in correlation with the
eye-open-and-close alternations (Fig.3.10.B), identified occipital and frontal regions that exhibited significantly larger BOLD responses during the eye-open periods than during the eye-closed periods (p<0.01, corrected).

**Figure 3.10** A) Time-frequency representation of the EEG data from the occipital electrodes. The blue-to-red colors encode the EEG power as a function of time and frequency. B) BOLD activation map contrasting the eye-open vs. eye-close conditions (p<0.01 after Bonferroni correction)

### 3.4 Electrophysiological Correlates to Negative BOLD

Furthermore, we applied the simultaneous fMRI-EEG recording and processing techniques to investigate the electrophysiological correlates of the negative BOLD response. After a brief introduction to the scientific background, we present some preliminary results obtained from the simultaneous fMRI-EEG experiments with small circular pattern-onset visual stimuli.
3.4.1 Background

Most fMRI studies rely on the task-induced positive BOLD response (PBR) as the hemodynamic index of increased neuronal activity (Bandettini et al. 1992; Kwong et al. 1992; Ogawa et al. 1992). The negative BOLD response (NBR), commonly observed as the sustained BOLD signal decrease relative to the resting-state baseline, is occasionally observed in both humans and animals undergoing sensory (Shulman et al. 1997; Tootell et al. 1998; Allison et al. 2000; Smith et al. 2000, 2004; Shmeul et al. 2002, 2006; Stefanovic et al. 2004; Whittingstall et al. 2007) and cognitive (Amedi et al. 2005) tasks.

Regardless of the increasing observations of the negative BOLD, its origin and relationship to neuronal activity remains poorly understood and controversial. The NBR may originate from reduced neuronal activity (Shulman et al. 1997; Gusnard & Raichle, 2001; Gold and Lauritzen 2002; Stefanovic et al. 2004; Smith et al. 2004), or vascular blood stealing (Harel et al. 2002; Kannurpatti & Biswal 2004), or perhaps both (Shmeul et al. 2002, 2006). In addition, the NBR can be either dependent or independent upon the task or stimulus content (Gusnard & Raichle 2001; Raichle & Mintun 2006). For example, in response to a particular stimulus, the NBR may be found at non-stimulated regions within the relevant sensory system (Tootell et al. 1998; Allison et al. 2000; Smith et al. 2000, 2004; Shmeul et al. 2002, 2006; Stefanovic et al. 2004). Such a task-dependent NBR is often hypothesized to reflect neuronal suppression at regions complementary to those with increased neural activity (Shmeul et al. 2006). On the other hand, decreases in BOLD (or CBF) signals are also observable at locations that change little across a wide variety of tasks (Shulman et al. 1997; Raichle et al. 2001; Mazoyer et al. 2001). These task-independent NBR regions are believed to be functionally active (Raichle et al. 2001; Gusnard & Raichle, 2001) and connected (Fox et al. 2005) in the resting state, collectively responsible for a default-mode brain function (Binder et al. 1999; Raichle et al. 2001). At these regions, the NBR may be induced as a consequence of decreased default-mode activity when the resting-state brain function is interrupted by the execution of attention demanding tasks (Raichle & Mintun 2006).

By simultaneously recording fMRI and EEG signals, we investigated the EEG-correlates to the sustained NBR induced by small circular pattern-onset visual stimuli.
(Smith et al. 2000, 2004). Specifically, we attempted to answer: 1) could we observe the visual-evoked NBR? 2) If yes, how was the NBR located in relation to the stimulus location? 3) If yes, could we observe a component(s) of continuous EEG signals that demonstrated a modulation in correlation with that of the NBR? 4) If yes, was the NBR-correlated EEG modulation frequency-dependent?

3.4.2 NBR to Small Circular Stimuli

Four subjects participated in the experiments with written consent. The visual stimuli consisted of 1 or 2 circular black-and-white checkerboard(s) presented on a gray background. The circular stimuli had small sizes with the diameter of 10°. The stimuli were located at ±8° or ±12° along the horizontal and vertical directions. The stimuli were pattern-onset, meaning that checkerboards appeared for a short duration of 50 msec and then disappeared. Subjects were instructed to gaze at a central fixation cross. The simultaneous fMRI-EEG experiments were conducted with a block design. In each stimulus block, either unilateral or bilateral stimuli were presented for 30 sec with an inter-stimulus-interval (ISI) of 0.5 sec. Several 30-sec control blocks with only the fixation cross were inserted between the stimulus blocks.

Fig 3.11 shows the BOLD activation maps contrasting three stimulus conditions vs. the control condition for a single subject. As shown in Fig. 3.11.A and 3.11.B, a unilateral stimulus evoked focal PBR at the contralateral V1 and V2, as well as the bilateral extrastriate areas (e.g. V5). The location and extent of the positive V1 activation agreed with the stimulus location and size according to the retinotopic relationship. The NBR was found at the ipsilateral and non-stimulated contralateral V1 areas, particularly at the deep anterior region close to the parietal-occipital sulcus. When two symmetric and bilateral stimuli were presented simultaneously, the PBR were focally located at the V1 and V2 areas within both hemispheres, while the regions with the NBR still covered a large portion of the anterior-median occipital cortex.
In line with the findings obtained in previous studies (Smith et al. 2000, 2004; Shmeul et al. 2002, 2006), our results demonstrate that in response to small visual stimuli, the NBR may be observed at the areas that roughly correspond to the stimulus-free part of the visual field, while the PBR are highly specific to the stimulus property (e.g. location and size). The fact that the NBR was found on the different hemisphere from the PBR tends to favor the neuronal rather than vascular origin of the NBR, because the two cerebral hemispheres are supplied by largely separate vasculatures.

We further varied the number and locations of the visual stimuli in different stimulus blocks. We found that most of the V1 demonstrated consistent NBR to different
stimuli. This was perhaps due to the small stimulus size, which gave rise to largely overlapped non-stimulated visual field for different stimulus blocks. Fig. 3.12 shows the BOLD activation map contrasting the sum of all the stimulus conditions vs. the control condition. The bilateral extrastriate areas exhibited the PBR across all of the six stimulus conditions, whereas the V1 areas mostly exhibited sustained NBR.

![BOLD Activation Map](image)

**Figure 3.12** Top: BOLD activation map contrasting the sum of all stimulus conditions vs. the control (resting) condition; Bottom: positive (in red) and negative (in blue) BOLD time courses averaged over voxels within the activated (in red-to-yellow) and deactivated (in blue-to-green) regions, respectively.
3.4.3 EEG Correlates to the NBR

In order to confirm the neuronal origin of the NBR, we searched the EEG data simultaneously recorded with the fMRI for some components that demonstrated the same modulation as that of the NBR. As shown in Fig. 3.12, the NBR at a large portion of the V1 were sustained across different stimulus blocks, whereas the PBR was found at variable locations according to the stimulus location in the visual field (as shown in Fig. 3.11). Therefore, if the NBR were coupled with decreased electrophysiological activities, the activities at the negative BOLD areas should generate some scalp EEG components that modulated in a similar way as the NBR did.

To test this hypothesis, we attempted to search for the NBR-correlated EEG signals (or components) using a data-driven approach based upon the independent component analysis (ICA). Here, ICA was used to decompose the spatiotemporal EEG data into a series of scalp potential maps associated with their corresponding time courses. In what follows, we first describe this ICA-based approach.

Let us start with denoting the continuous multi-channel EEG signals as an $N_x$-by-$N_T$ matrix $X$, where $N_x$ and $N_T$ are the numbers of EEG channels and temporal sampling points, respectively. The $t$-th column of $X$ represents the scalp potential map at the time $t$. The independent component (IC) decomposition of $X$ can be expressed as

$$X = Y W T$$

(3.1)

where $Y$ is an $N_x$-by-$N_x$ matrix, $W$ is an $N_x$-by-$N_x$ diagonal matrix, $T$ is an $N_x$-by-$N_T$ matrix. The matrix $W$ can be estimated from the data matrix $X$ using a logistic Infomax ICA algorithm (Bell and Sejnowski, 1995).

Eq. (3.1) can be expanded as Eq. (3.2)

$$X = \sum_{i=1}^{N_x} Y_i w_i T_i$$

(3.2)

where $Y_i$ is the $i$-th column of $Y$, $T_i$ is the $i$-th row of $T$, and $w_i$ is the $i$-th diagonal element of $W$. Mathematically, the IC decomposition ensures $T_i \perp T_j$ for $i \neq j$.

Eq. (3.2) suggests that the spatiotemporal EEG data can be equivalently represented as a weighted sum of a group of scalp potential maps $\{Y_i\}$ multiplied by a series of time courses $\{T_i\}$ that are independent to each other.
Note that the EEG data are linearly linked with the spatiotemporal source activity, as specified by the following EEG forward model.

$$\mathbf{X} = \mathbf{A} \mathbf{S} + \mathbf{B}$$  \hspace{1cm} (3.3)

where $\mathbf{A}$ is the source-to-sensor transfer matrix obtained through modeling the current source space and the head volume conductor (see 2.2.2 for details), $\mathbf{S}$ is the spatiotemporal source activity and $\mathbf{B}$ represents the noise.

Let $\mathbf{G}$ stand for a pseudo-inverse or regularized inverse matrix of $\mathbf{A}$ (see 2.2.3 and 5.2.1 for details). Using $\mathbf{G}$, we can reconstruct the spatiotemporal source distribution from the EEG data.

$$\hat{\mathbf{S}} = \mathbf{G} \mathbf{X}$$  \hspace{1cm} (3.4)

Replace $\mathbf{X}$ by Eq. (3.2), then Eq. (3.4) can be re-written as Eq. (3.5)

$$\hat{\mathbf{S}} = \sum_{i=1}^{N_x} \mathbf{G} \mathbf{Y}_i w_i \mathbf{T}_i$$  \hspace{1cm} (3.5)

Eq. (3.10) indicates that the reconstruction of the spatiotemporal source activity from the EEG data can be equivalently represented as the sum of a group of source spatial patterns $\{\mathbf{G} \mathbf{Y}_i\}$ multiplied by a series of independent time courses $\{\mathbf{T}_i\}$ weighted by a set of weighting coefficients $\{w_i\}$.

Importantly, the decomposed independent source components have separated spatial and time aspects, namely the space-time separation. It allows us to assess the modulations of independent source activities in time and/or frequency domains, separately from imaging the corresponding spatial source distributions.

Based on the aforementioned ICA-based approach, Fig. 3.13 schematically illustrates the separated spatial and temporal analyses for investigating the EEG correlates to the NBR. As discussed in 3.2, we were able to remove the MR-related artifacts from the EEG data recorded during concurrent fMRI scans. The resultant EEG data had equivalent quality as if recorded alone with the fMRI. The ICA was performed to decompose the multi-channel EEG data into a number of independent components, each of which contained a spatial map and a corresponding time course of the same length as the EEG recording period. The spatial map implied a certain source pattern, which could be imaged by solving the EEG inverse problem. The time course indicated how the amplitude of the source pattern changed over time. We further computed the
time-frequency representation (TFR) of each component’s time course using the short-time Fourier transformation. The TFR indicated how the frequency-specific power modulated over time for each decomposed source pattern. In line with the use of ICA for removing the residual artifacts (see 3.2.2), here ICA also allowed us to identify noisy or artifact components (such as eye blinks or muscle artifacts).

Figure 3.13 Schematic illustration of the separated spatial and time-frequency analyses based on the independent component analysis.

We identified a NBR-correlated IC if 1) its source distribution overlapped with (or was close to) the NBR regions and 2) it contained a frequency band(s) over which the
power changed similarly as the NBR did. We found three ICs satisfying both criteria, as shown in Fig. 3.14. Commonly among these three ICs, their underlying source distributions were mainly at the V1, and there were one or multiple frequency component(s) that exhibited decreased power during all of the stimulus blocks and increased power during the control block.

**Figure 3.14** Three ICs (shown as three rows) correlated with the NBR. The source image, scalp potential pattern and the time-frequency representation corresponding to each IC are shown from left to right.
Figure 3.15 A) and B) are the source image and TFR of the sum of the NBR-correlated ICs; C) Power time courses for alpha and delta bands; D) Predicted BOLD responses derived from delta and alpha power modulations are compared with the NBR time course.

According to Eq. (3.5), we summed the three NBR-correlated ICs, resulting in a single source image and the corresponding TFR, as shown in Fig. 3.15.A) and 3.15.B),
respectively. We could observe the alternative power modulation in the low frequency range, particularly at the delta (2~4 Hz) and alpha (8~12 Hz) bands. We further averaged the time courses of the power within each of these two frequency bands (as shown in Fig. 3.15.C) and convoluted the resultant delta-band and alpha-band power waveforms with a canonical hemodynamic impulse response function (HRF). Such convolution gave rise to two predicted BOLD responses coupled with the power modulation in the delta and alpha bands. We found both predicted responses fitted the NBR time course reasonably well, with the cross-correlation higher than 0.4. These results suggest that the NBR at V1 are coupled with the decrease of alpha and delta electrophysiological activities at V1.

3.5 Conclusion & Discussion

In summary, the simultaneous EEG-fMRI recording is challenging but necessary in many fMRI-EEG integrated neuroimaging and neuroscience researches. Advances mainly in signal processing techniques have allowed us to remove various MR-related artifacts from the EEG data recorded during concurrent fMRI scans. We have explored and developed some of these signal processing algorithms, which are demonstrated to be reliable and effective for retrieving the event-related potentials (e.g. VEP), frequency-specific EEG component (e.g. SSVEP) and time-frequency-specific EEG component (e.g. alpha-wave modulation). The usefulness of simultaneous fMRI-EEG recordings is also demonstrated in a study investigating the EEG correlates to the negative BOLD responses.

Although the simultaneous fMRI-EEG recording is desirable for a variety of purposes, it is not always necessary when studying some passive sensory evoked responses. In this regard, one has to base his/her own choice on the reproducibility of the task (or stimulus) of interest vs. the possible risk of dealing with largely contaminated EEG data if recorded simultaneously with fMRI. Regardless of the theoretical efficacy of the artifact reduction algorithms developed by us and others, the outcome of these algorithms is after all partly artificial and, after all, “worse” than artifact-free signals. For some simple paradigms only involving passive stimulus-evoked responses, it might not be always worthwhile to choose to have likely “worse” signals rather than ensuring clear signals while bearing with separate but highly reproducible experiments.
Through simultaneous fMRI-EEG experiments, the NBR related to small visual stimuli were found to be correlated with delta- and alpha-band EEG modulations. This finding further confirms the neuronal origin of the NBR, since the NBR-correlated EEG modulations are independent upon cerebral vasculature but solely depends on the electrophysiological neural activity.

Both the NBR and its correlated EEG components arose from the median occipital cortex, instead of the lateral occipital lobe (i.e. extrastriate cortical areas). The median occipital cortex contains the primary visual areas (V1). A region in V1 is specifically responsive to a part of the visual field, which is known as its receptive field. When the visual input is small in the visual field, it accordingly activates a small region whose receptive field contains the visual input. For neurons with non-stimulated receptive fields, their activities are likely suppressed to allow for the optimal reallocation of attentional resources to facilitate the visual processing. The coordinated activation and deactivation may be enabled by the horizontal connections between the neurons within the same hemisphere or by the cross-hemisphere interaction through the corpus callosum, as the NBR can occur at the hemisphere contralateral to the PBR. Alternatively, both the PBR and NBR may be controlled by the sub-cortical areas.

Perhaps, one would argue that the identified EEG components which we interpreted to be correlated with the NBR, might be actually coupled with the PBR while assuming a negative correlation between the alpha and beta power and the BOLD signal. However, we believe this argument cannot hold up, since the alpha and beta power of three identified ICs decreased during all of the stimulus blocks. The visual stimuli within different stimulus blocks were from different locations, resulting in both the PBR and its correlated electrophysiological responses occurring at different V1 regions. As a result, any PBR-correlated EEG component (if exists at all) should modulate differently across blocks.

Our findings further suggest a non-zero baseline activity. The NBR is referred to the BOLD signal decrease relative to the resting state. The existence of the NBR implies that some of the brain areas demonstrate non-zero brain signals (e.g. EEG and fMRI) in the resting state, which allows for the possibility of further signal decrease from the
resting state to the states with stimuli or tasks. In addition, the EEG components in correlation with the NBR contained posterior alpha activity, which is well known to be an electrophysiological signature of the resting-state brain function.
Chapter 4 Neurovascular Coupling

4.1 Introduction

Functional MRI, as a neuroimaging tool, explicitly or implicitly assumes a linear relationship between neural activity and fMRI signals. Only with this assumption can quantitative inference be made, regarding the underlying neural activity, based on noninvasively measured fMRI signals and specific experimental designs. Regardless of numerous fMRI applications, this underlying assumption has not been fully validated. Increasing attention has been directed upon the coupling between BOLD signals and neuronal electrophysiological activity, namely the neurovascular coupling. Quantitative neurovascular coupling models can not only guide the interpretation of BOLD signals, but also provide the theoretical basis for developing multimodal neuroimaging techniques that combine fMRI with electrophysiological measurements such as EEG and MEG.

Firstly, it is logic to ask what aspects of neuronal behaviors account for BOLD fMRI responses. The answer is still under debate and controversy, as both the spiking activity (i.e. neuronal output) and the synaptic activity (i.e. neuronal input) have been suggested to serve as the neurophysiological origin of BOLD signals.

Early experimental studies (Rees et al., 2000; Heeger et al., 2000) suggest the dependence of BOLD responses on the neuronal spiking activity, as they found that the peak magnitudes of BOLD signals measured at human V5 (Rees et al. 2000) and V1 (Heeger et al. 2000) were linearly correlated with the averaged spiking rates invasively recorded at analogous areas in monkeys. More recently, Mukamel et al. also found the close correlation between the BOLD signals measured from normal subjects and the neuronal spiking activity recorded from epilepsy patients in response to the identical stimulation (Mukamel et al. 2005). However, the conclusions drawn from these studies might be confounded by the fact that the BOLD and neuronal spiking signals were acquired from different species or subjects.

Alternatively, many other studies have demonstrated that the BOLD fMRI reflects the synaptic activity rather than the spiking activity (Logothetis et al. 2001; Lauritzen 2001; Thomsen et al. 2004; Viswanathan and Freeman 2007). For instance, Logothetis et
al. recorded the spiking activity (i.e. single-unit activity (SUA) and multi-unit activity (MUA)) and the synaptic activity (i.e. local field potential (LFP)) simultaneously with BOLD fMRI signals from monkeys (Logothetis et al. 2001). They found that through linear transformation systems, the LFP yielded better estimates of BOLD responses than the MUA did. From these observations, they concluded that BOLD signals are reflection of synaptic activities (Logothetis et al. 2001; Logothetis 2003; Logothetis and Wandell 2004). Following this study, similar findings were also reported in several independent studies, such as (Lauritzen 2001; Thomsen et al. 2004; Viswanathan and Freeman 2007).

Between these two competing views, agreement has been reached in general, but not in details, with regard to the synaptic origin of BOLD signals (Raichle and Mintun 2006). In fact, hemodynamic responses are driven by metabolic energy demands, nearly all of which is imposed by synaptic activity instead of action potential firing (Mathiesen et al., 1998; Arthurs and Boniface, 2002). The observed correlation between the BOLD signal and the neuronal spiking rate (Rees et al., 2000; Heeger et al., 2000) might be explained by the fact that the pre-synaptic neuronal spiking activity is indeed correlated with the synaptic current under certain circumstances, as discussed in 2.2.1 and (Heeger and Ress, 2002; Arthurs and Boniface, 2002). However, when the LFP and MUA are disassociated with each other, the BOLD response is primarily correlated with the LFP (Thomsen et al. 2004; Viswanathan and Freeman 2007).

In line with the above experimental studies, the neurovascular coupling is often approximated by a linear transformation model or hemodynamic impulse response function (HRF). The model input is the neuronal synaptic activity. The model output is the hemodynamic response. More specifically for the BOLD-contrast fMRI, the model input should be the power of synaptic currents (Liu and He, 2008) or potentials (Logothetis et al. 2001), which can be regarded as the physical correlates of metabolic energetics (Nangini et al. 2007; Wan et al. 2006b; Liu and He 2008).

A linear HRF at best serves as an approximation of the complex interactions between neuronal activity, metabolic demand and blood flow and oxygenation (Heeger and Ress 2002). With approximations made at various degrees, the HRF can be simply a Gauss function, gamma function (Boynton et al. 1996) or double-gamma canonical
function (Friston et al. 1995b, 1998). In particular, the canonical HRF can well describe most of the key features of the neurovascular coupling, including the onset time, initial dip, overshoot and undershoot etc.

The use of a linear HRF largely simplifies the analysis and interpretation of BOLD fMRI data, with or without in combination with other electrophysiological modalities. However, when only the fMRI data are available, the HRF has to take the stimuli (or tasks) as the model input while assuming a linear relationship between stimuli and neuronal responses, which may or may not be valid depending on particular circumstances. Some studies have reported a nonlinear stimulus-to-BOLD relationship (Vazquez and Noll 1998), which may be caused by the nonlinear neuronal response to stimuli (such as neuronal refractory effect), and/or the nonlinear vascular response to neuronal activity (such as vascular refractory effect). The BOLD signal can be quantified by its peak height, steady-state height or time integral. However, it remains unclear how to interpret these different measures in which quantitative aspects of the underlying neuronal response. Due to this uncertainty, one must be cautious to avoid using mismatched quantifications of BOLD and neuronal signals when assessing the linearity or nonlinearity of the relationship between them.

It is often of interest to quantify BOLD signals to characterize event-related electrophysiological responses (Liu and He 2008b). Invasive or noninvasive assessments of stimulus evoked (or task related) electrophysiological responses are often conducted in an event-related manner. This is because neuronal responses often take place within a very short time window, as opposed to the sluggishness and long duration of hemodynamics. For the purpose of integrating ERP and fMRI or investigating the relationship between them, it is particularly important to develop a theoretically driven approach to interpret fMRI signals as physically and physiologically meaningful indices for ERP.

Towards this end, we have developed mathematical models describing the interactions between stimuli (or tasks), neuronal synaptic currents and BOLD responses. This model allows us to derive the relationship between event-related synaptic responses
and BOLD fMRI signals. It also formulates the theoretical guidance and hypothesis necessary for assessing the BOLD linearity and nonlinearity.

4.2 Theoretical Modeling

4.2.1 Systematic Perspective

From the systematic perspective, the cascade interactions from stimuli to BOLD signals are illustrated in Fig. 4.1. Neural responses evoked by stimuli can be expressed by the convolution of the stimuli with a neural impulse response function (NRF), or more specifically the event-related synaptic activity. The total synaptic activity is the sum of the evoked response and the spontaneous activity presumably irrelevant to the stimuli. Through another linear system (i.e. HRF), the total synaptic activity is linked to the BOLD fMRI signal, which consists of the BOLD response coupled with the evoked neuronal response, and the baseline (or resting-state) BOLD signal related with the spontaneous neuronal activity.

Figure 4.1 Linear systems modeling the interactions between stimuli, neuronal responses and BOLD responses.

This system is assumed to be linear; however, the linearity of the NRF and HRF may be affected by neuronal and vascular refractory effects, respectively. Neuronal
refractory effect is referred to the fact that the total neuronal response to multiple stimuli is not the superimposition of individual response to each stimulus. This often occurs due to a too short inter-stimulus-interval (ISI), such that the neuronal response to a stimulus fails to completely recover so as to repeatedly respond to ensuing stimuli. In most cases, the neuronal refractory effect can be avoided by using a long ISI. Due to the short duration of the NRF itself, it is straight-forward to set the ISI longer than the NRF duration. Such a setting also allows us to simply average the responses to repeated stimuli, yielding the event-related electrophysiological signal, as in most ERP studies. As opposed to neuronal refractory effect, vascular refractory effect is more complex and still under controversy. Here, let us first ignore the vascular refractory effect, so that the linearity of the HRF holds up. We will re-visit this issue later.

As discussed in 2.3.3, fMRI is comparative in nature because the fMRI analysis is always based upon comparing one condition vs. the other. Most typically, the comparison is between a stimulus (or task) condition and the resting-state control condition. Such a comparison is useful in finding brain regions engaged in the stimulus processing. In the resting control state, the neuronal signals are non-zero, collectively reflecting the spontaneous behavior of the brain. Here, we assume the spontaneous brain activity does not change from the resting-state condition to the stimulus condition. Accordingly, the fMRI signal change during the stimulus condition relative to the resting-state control condition exclusively results from the stimulus evoked neuronal response.

In short, three primary assumptions are made to ensure the system linearity: 1) linear NRF, with a sufficiently long ISI, 2) linear HRF, ignoring potential vascular refractory effect, and 3) no change to the baseline activity.

4.2.2 BOLD Quantification

With these assumptions, we simplify the system in Fig. 4.1 into the one in Fig. 4.2. Comparing Fig. 4.1 and Fig. 4.2, one can find that the simplified system is linear without involving the spontaneous activity and the BOLD baseline. In what follows, we derive the relationship between the NRF and the BOLD signal.
Since many fMRI experiments are often conducted in a block-design manner (as discussed in 2.3.3), we start our derivation by representing the stimuli as a train of delta functions. Later, we will generalize the derived relationship to the event-related design.

Assume a single stimulus at time 0 evokes synaptic current sources \( s \), where \( r \) indicates location in the brain and \( t \) indicates time. The power of such single-stimulus evoked currents serves as the NRF. The duration of the NRF, denoted as \( T_s \), typically ranges from several tens to several hundred milliseconds, depending on particular circumstances. Then the source activity evoked by a block of \( N \) stimuli, denoted as \( g \), can be written as Eq. (4.1).

\[
\bar{g}(r, t) = \sum_{n=1}^{N} \delta(t - nT_{ISI}) \otimes s(r, t)
\]  

(4.1)

where \( \delta(t) \) is a delta function, \( T_{ISI} \) is the inter-stimulus-interval (ISI) and \( \otimes \) denotes convolution. As previously discussed, Eq. (4.1) is valid when the ISI is longer than the neuronal refractory period. It is particularly so in most ERP studies, in which the ISI is designed to be even longer than \( T_s \) in order to ensure that the electrical activity responding to preceding stimuli recovers to the resting state (or the baseline) before the response to next stimulus is elicited.

Assume the induced BOLD-fMRI response \( f(r, t) \) relates to the power of local synaptic current by a linear system that is characterized by the hemodynamic impulse response function (HRF) \( h(t) \), plus noise. Such a linear system is illustrated in Fig. 1 and mathematically expressed as Eq. (4.2) to Eq. (4.4):

\[
f_j(r, t) = g^2(r, t) \otimes h(t)
\]  

(4.2)

\[
f_j(r, t) = \sum_{n=1}^{N} \delta(t - nT_{ISI}) \otimes s^2(r, t) \otimes h(t)
\]  

(4.3)
\[ f(r,t) = f_s(r,t) + f_n(r,t) \]  \hspace{1cm} (4.4)

where \( g(r,t) \) and \( s(r,t) \) are the magnitudes of \( \tilde{g}(r,t) \) and \( \tilde{s}(r,t) \) respectively, \( f_s(r,t) \) is the “signal” part of the fMRI response and \( f_n(r,t) \) is the “noise” part.

Eq. (4.3) can be re-organized as Eq. (4.5):

\[ f_s(r,t) = s^2(r,t) \sum_{n=1}^{N} \delta(t - nT_{isi}) \cdot h(t) \]  \hspace{1cm} (4.5)

We define a predictor signal \( p(t) \) as Eq. (4.6).

\[ p(t) = \sum_{n=1}^{N} \delta(t - nT_{isi}) \cdot h(t) \]  \hspace{1cm} (4.6)

Then, Eq. (4.5) can be further re-written as Eq. (4.7) and Eq. (4.8).

\[ f_s(r,t) = s^2(r,t) \cdot p(t) \]  \hspace{1cm} (4.7)

\[ f_s(r,t) = \int_{T_s}^{s^2(r,t)p(t-\tau)d\tau} \]  \hspace{1cm} (4.8)

Since the hemodynamic impulse response \( h(t) \) evolves much slower than the electrophysiological impulse response, it keeps approximately constant over the duration of the electrical source signal evoked by a single stimulus. Mathematically, we assume \( p(t) \approx p(t-\tau) \) for \( 0 \leq \tau \leq T_s \). Then Eq. (4.8) can be simplified as Eq. (4.9)

\[ f_s(r,t) = p(t) \int_{T_s}^{s^2(r,t)p(t-\tau)d\tau} \]  \hspace{1cm} (4.9)

We define \( \beta(r) \) as Eq. (4.10) and re-write Eq. (4.9) as Eq. (4.11)

\[ \beta(r) = \int_{T_s}^{s^2(r,t)d\tau} \]  \hspace{1cm} (4.10)

\[ f_s(r,t) = \beta(r)p(t) \]  \hspace{1cm} (4.11)

For a time series of fMRI response with \( N_f \) discrete samples over time, it immediately follows from Eq. (4.4) and Eq. (4.11) that an over-determined linear regression model can be described using a vector notation as Eq. (4.12).

\[ F(r) = \beta(r) \cdot P + F_n(r) \]  \hspace{1cm} (4.12)
where $F(r) = [f(r, 1) f(r, 2) \ldots f(r, N_r)]^\top$ and $F_n(r) = [f_n(r, 1) f_n(r, 2) \ldots f_n(r, N_f)]^\top$ are a vector of the fMRI time series and its “noise” part respectively, and $P = [p(1) \ p(2) \ldots \ p(N_f)]^\top$ is a vector of the predictor signal.

Based on Eq. (4.12), the minimum least-squares estimate of $\beta(r)$, denoted as $\hat{\beta}(r)$, can be computed from Eq. (4.13)

$$\hat{\beta}(r) = \left(P^\top P\right)^{-1} P^\top F(r)$$  \hspace{1cm} (4.13)

As expressed in Eq. (4.13), $\hat{\beta}(r)$ is the BOLD effect size that exclusively depends upon the BOLD response, the experimental protocol and the hemodynamic impulse response function. More importantly, $\hat{\beta}(r)$ can be interpreted, by the definition of $\beta(r)$ in Eq. (4.10), as an estimate of the time integral of current source power over the period of $T_r$. Fig. 4.3 graphically illustrates the above derivations that lead to a method for quantifying BOLD signals to characterize event-related electrophysiological responses.

This interpretation is particularly meaningful and valuable in the context of electrical source imaging. And it is explicitly used later when incorporating the fMRI into solving the EEG inverse problem (see 5.2.3 for details).

The above method can also be applied to the quantification of the event-related fMRI response, simply by changing the predictor function in Eq. (4.6) to Eq. (4.14).

$$p(t) = \delta(t) * h(t) = h(t)$$ \hspace{1cm} (4.14)

The above method can be further generalized to more complicated experimental designs with multiple or mixed tasks. Correspondingly, multiple predictor functions need to be defined by convolving the mathematically expressed experimental protocol with the HRF. If the quantification of BOLD responses has a multivariate nature, the regression parameters should be defined for each stimulus or task independently. And we can set up a general linear model (GLM) (Friston et al., 1995) similarly as Eq. (4.12) and compute the regression parameters using a multivariate linear regression algorithm.
4.3 Experimental Investigation

Through theoretical modeling the interactions between stimuli, neuronal and BOLD responses, we have derived that the time integral of the event-related synaptic power is proportional to a scaling factor (i.e. BOLD effect size) by which the modeled (or predicted) BOLD response is scaled to best fit the real BOLD signal. In this subsection, we describe two experimental paradigms through which we have tested the theoretically derived linear relationship.

4.3.1 Full-screen Checkerboard Experiment

The first experimental paradigm simply employed a full-screen checkerboard visual stimulation. In response to the visual stimulation, the fMRI and EEG data were
acquired simultaneously. The detailed experimental procedures and setup are described in 3.3.1.

From the VEP signals, we imaged and averaged the power of cortical current distribution during the entire 500-ms post-stimulus period. From the fMRI data, we obtained the BOLD activation map, which approximately represents the spatial distribution of BOLD effect sizes. The comparison between these two maps “spatially” evaluated the theoretical relationship between the integral (or equivalently the average) of synaptic power and the BOLD effect size.

![Figure 4.4](image)

**Figure 4.4** Comparison between the spatial distributions of the time integral of current density power estimated from the VEP signals, and the BOLD effect sizes derived from the fMRI data.

Fig. 4.4 shows the results obtained from two subjects. General consistency can be seen between the cortical map of the integral of estimated current density power and that of the BOLD effect size. This result supports our theoretical model, whereas the mild difference between two sets of maps may be partly due to the spatial inaccuracy of the EEG source imaging.
4.3.2 Responses to Variable Visual Contrast

As opposed to the first paradigm, the second paradigm employed variable visual contrast which provided a wide dynamic range to assess the theoretically derived linearity.

Subjects & Stimuli

Seven subjects (aged 19–31, 3 females) participated in both fMRI and EEG experiments with written consent according to a protocol approved by the institutional review board (IRB) at the University of Minnesota.

Figure 4.5 Illustration of the pattern-reversal visual stimuli with variable contrasts.

As summarized in Fig. 4.5, the visual stimulus was a quarter of circular grating bar at the lower-right visual field with the contrast ranging from 5% (or 10% for 3 subjects) to 100%. The 100% contrast represented a black-and-white contrast. The stimuli were pattern-reversal with an ISI of 500 ms, meaning that the contrast pattern reversed every 500 msec. When the stimuli were presented, a red solid circle was shown
on the center of the screen. The circle was still shown even without any stimulus. All of the subjects were trained and instructed to maintain the fixation at the red circle. Both fMRI and EEG experiments employed a block design with interleaved stimulus and control blocks each lasting for 30 s. During the control blocks, only the fixation circle was displayed. From the first to the last stimulus blocks, the visual contrast was either increasing or decreasing.

**Data Acquisition**

For 4/7 subjects, the EEG and fMRI data were acquired during separate sessions, while for the other 3 subjects simultaneous fMRI-EEG experiments were performed. The EEG-alone studies were conducted in an electrically shielded room. The fMRI or fMRI-EEG experiments were conducted in a 3-T/90 cm bore magnet (Siemens Trio, Siemens, Germany) equipped with an eight-channel phase array head volume coil. Each experiment included six repeated runs. In different runs, increasing or decreasing visual contrast was used alternatively.

In both EEG and fMRI-EEG experiments, the scalp potentials from 64 electrodes (referenced to FCz and placed according to the extended international 10/20 system) were recorded at 1000 Hz and filtered (0.3~70 Hz) through a pair of amplifiers (BrainAmp MR 64 Plus, BrainProducts, Germany). Eye blinks and movements were monitored with horizontal and vertical electrooculographic (EOG) electrodes. The electrode locations and three anatomical landmarks (left/right preauricular points and nasion) were digitized through a three-dimensional (3-D) RF localizer (Polhemus Fastrak, Colchester, VT). The detailed experimental setup for simultaneous fMRI-EEG experiments are described in 3.2.1.

In both fMRI and fMRI-EEG experiments, the whole-head anatomy was first acquired with 256 sagittal T1-weighted MR images (matrix size: 256×256; in-plane resolution: 1×1 mm²; slice thickness: 1 mm; no gap between slices) using a TurboFLASH sequence (TR/TE = 20/5 ms). The BOLD fMRI data was acquired with 16 axial T₂*-weighted images (matrix size: 64×64; in-plane resolution: 4×4 mm²; slice thickness: 5 mm; no gap between slices) covering both the occipital and parietal lobes using a gradient-echo echo-planar imaging (EPI) sequence (TR/TE = 1000/35 ms).
Data Processing

Following the MR-artifact removal (see 3.2.2 for details), the EEG was preprocessed in BrainVision Analyzer (BrainProducts, Gilching, Germany). The recorded EEG signals were sequentially subject to ocular artifact rejection (by visual inspection), band-pass filtering (0.3 – 40 Hz), segmentation from -100 to 500 ms around the stimulus onsets, pre-stimulus baseline correction, linear trend removal and response averaging to obtain the contrast-specific VEP data. To quantify the time integral of the VEP power arising from the V1 activity, we summed the VEP power within a period around the P100 component (from about 60 ms to 140 msec, slightly variable across subjects). This time range has been known to reflect the visual response at V1.

The fMRI data was analyzed using the general linear model (GLM) approach described in 4.2.2. The BOLD activation map corresponding to the 100% visual contrast was first obtained with p<0.01 after Bonferroni correction. A region of interest (ROI) covering the V1 activation was defined. Both the BOLD time course and the quantified BOLD effect size were averaged within the defined V1 ROI.

For the V1 area, the time integral of the VEP power and the averaged BOLD effect size were plotted for each individual subject and all of the subjects as a group, in order to assess the linearity between them as suggested by the theoretical model.

Results

Fig. 4.6 shows the results obtained from a representative subject. With a high significance level (p<0.01 corrected), the BOLD activation map (Fig. 4.6.A) revealed mainly the activation at V1 areas. The V1 BOLD response elevated with increasing visual contrast, as shown in Fig. 4.6.B. However the increase of BOLD responses was not a linear function of the visual contrast, as larger increases were seen at lower contrast while smaller increases at higher contrast. As shown in Fig. 4.6.C, the VEP signal at Oz (at the middle occipital lobe) demonstrated a similar contrast modulation as the V1 BOLD response. That is, higher visual contrast gave rise to larger VEP magnitude particularly around 100 msec. The scatter-plot of the averaged BOLD effect size vs. the
time integral of the VEP power, as shown in Fig. 4.6.D, suggests a linear relationship between them with the linear correlation of about 0.94.

**Figure 4.6** A) BOLD activation map corresponding to 100% visual contrast, B) BOLD time course averaged within the identified ROI at V1, C) VEP at Oz for different visual contrast, D) the scatter plot of the BOLD effect size vs. the time integral of the VEP power.

Fig. 4.7 shows the BOLD and VEP comparison for every individual subject. Good linear fit to the data were found in all 7 subjects, with the correlation coefficient ranging from 0.72 to 0.94. These results indicate the reproducibility of the observed linear relationship.

In summary, the comparison between the BOLD and VEP responses to visual stimuli with variable contrast confirmed the linear relationship between the BOLD effect
size and the time integral of synaptic current power, as suggest by the theoretical models described in 4.2.

**Figure 4.7** Scatter plots of the BOLD effect size vs. the time integral of the VEP power for all 7 subjects.

### 4.4 Conclusion & Discussion

Through modeling the interactions between stimuli, neuronal responses and BOLD signals, we derive a theoretical relationship between the BOLD effect size and the time integral of the event-related synaptic power. Our model and related computational method provides a theory-driven approach to quantify BOLD signals to characterize event-related electrophysiological responses. Since the model assumes a linear neurovascular coupling, the theoretical results obtained in our study would guide the assessment of the linear vs. nonlinear neurovascular coupling relationship.

We also evaluated the theoretically derived relationship against real data in two visual stimulation paradigms. The experimental results suggest that the spatial
distribution of the time integral of current density power estimated from the VEP signals generally agrees with that of the BOLD effect size quantified from the fMRI data, and that the BOLD-VEP linearity holds up within a wide dynamic range resulting from the variable visual contrast.

**Assessment of Neurovascular Coupling**

The theoretical model presented in this Chapter is based on a linear neurovascular coupling. When theoretical results are rigorously derived from modeling assumptions, evaluation of the theoretical results against experimental data provides a hypothesis driven way to test the modeling assumption.

This is as opposed to many existing studies that investigate the neurovascular coupling through mismatched quantifications of BOLD and neuronal responses. For instance, it is problematic to assess the linearity or nonlinearity through comparing 1) the peak (or steady-state) height of BOLD response vs. the time integral of synaptic current during a long time period during which both BOLD and neuronal signals are collected, or 2) the time-integral or the peak (or steady-state) height of BOLD response vs. the magnitude of event-related neuronal response (e.g. ERP or single-trial LFP) at a single peak latency. Depending on particular circumstances, these mismatched indices for BOLD and neuronal response quantification may exhibit nonlinear features even though the linear neurovascular coupling is valid, or vice versa.

**Baseline Activity**

The theoretical model also assumes that the baseline activity does not change from the resting-state control condition to the stimulus condition. However, it is possible that the baseline activity changes in response to certain stimulus. It all depends on the location of interest as well as the stimulus type and property. For instance, there are a number of default-mode regions that are functionally active during the resting state but likely suppressed given certain stimulus or task that requests attentional resources. It is also possible that the baseline activity at some brain regions spontaneously modulates over time, with or without any relationship to the external stimulus. As a result, if the stimulus is given when the baseline activity modulation reaches a lower (or higher) level
than during the resting-state control state, the corresponding BOLD signal change reflect not only the stimulus evoked neuronal response but also the modulation of the spontaneous baseline activity.

**Vascular Refractory Effect**

In our experiments with variable visual contrast, we observed that the linear function fitting the BOLD effect size vs. the time integral of VEP power failed to pass through the origin, different from what would be expected from the theoretical results. Instead, a positive intercept was found for all 7 subjects. As discussed before, such deviation against the theoretical results necessarily reflects one or more invalid modeling assumptions.

Among three modeling assumptions, the linear NRF should be valid in this study since a 500-ms ISI was employed to allow for a recovery period long enough to avoid the neuronal refractory effect. We may also exclude the baseline activity modulation from concerns. A suppression of baseline activity would more likely give rise to a negative, instead of positive, intercept, and the self-modulation of baseline activity should possibly causes both negative and positive intercepts, while only the positive intercept was observed in our data. In contrast, the observed “nonlinearity” most likely attributed to the vascular refractory effect. By taking this effect into account, one should expect the best-fit linear function be much close to the origin.
Chapter 5 fMRI-EEG Integrated Neuroimaging: Methodology

5.1 Introduction

As discussed in the previous chapters, fMRI and EEG have complementary advantages in imaging either the spatial or temporal aspect of the brain function, and vice versa for their limitations. This greatly motivates recent methodological developments that attempt to integrate both modalities to establish multimodal functional neuroimaging methods with high spatiotemporal resolution. Some of these methods are grounded by the close coupling between BOLD and EEG source signals, which suggests that brain regions showing increased BOLD response are also on average more electrically active over time (Dale and Sereno 1993; Liu et al. 2006a; Liu and He 2008a).

Methods for the fMRI-EEG integrated neuroimaging can be categorized into two types: the fMRI-constrained EEG source imaging and the EEG-informed fMRI analysis. Briefly, the fMRI-constrained EEG source imaging rests on the EEG source imaging, while incorporating hopefully helpful spatial information from fMRI. The EEG-informed fMRI analysis utilizes the time and/or frequency specific information to derive the regressors in the fMRI analysis.

5.1.1 fMRI-Constrained EEG Source Imaging

The methods we are mainly taking belong to the former category. The existing fMRI-constrained EEG source imaging methods typically utilize fMRI activation maps, resulting from statistical analyses of fMRI time series, as \textit{a priori} information of where EEG sources are likely located. Depending on different EEG source models, an fMRI map can be used to constrain the locations of multiple current dipoles, namely the fMRI-constrained dipole fitting (Ahlfors et al. 1999; Korvenoja et al. 1999; Fujimaki et al. 2002; Vanni et al. 2004a), or to constrain the distributed source distribution on the folded cortical surface or in the 3-D brain volume, namely the fMRI-constrained current density imaging (Liu et al. 1998; Dale et al. 2000; Wagner et al. 2000; Babiloni et al. 2003; Ahlfors and Simpson 2004; Sato et al. 2004; Phillips et al. 2005; Liu et al. 2006b;
Methods for the fMRI-constrained dipole fitting and current density imaging are described in details in \textbf{5.2.5} and \textbf{5.2.1} through \textbf{5.2.4}, respectively.

While applications of dipole fitting techniques are often questionable when brain activities are spatially distributed instead of being confined to focal regions, the distributed source imaging is generally applicable. The existing methods for the fMRI-constrained current density imaging have been implemented under different theoretical frameworks such as Wiener estimation (Dale and Sereno 1993; Liu et al. 1998; Dale et al. 2000), weighted minimum norm (Wagner et al. 2000; Babiloni et al. 2005; Ahlfors and Simpson 2004), Bayesian estimation (Sato et al. 2004; Phillips et al. 2005; Mattout et al. 2006) and Twomey regularization (Liu et al. 2006b).

All of these methods have limitations in two important aspects. First, there is no generalized method to quantify the fMRI signal with an explicit physical interpretation in the context of the EEG source imaging (Liu and He 2008a). It is perhaps such limitation that makes it difficult to develop a principled way to incorporate the fMRI data in solving the EEG inverse problem. Moreover, when the prior spatial constraint is derived from the fMRI activation map after applying a statistical threshold as in most existing methods, the fMRI-constrained source reconstruction is also subject to the choice of the threshold as well as a variety of methods for the fMRI analysis (Ahlfors and Simpson 2004). Secondly but more importantly, an fMRI-derived “time-invariant” spatial constraint (Lin et al., 2006) is applied when imaging the temporally variable current source distribution. However, such a time-invariant spatial constraint may entail both fMRI false positives and false negatives, as a result of unavoidable mismatches between locations of fMRI activations and instantaneous source activities, namely the fMRI-EEG mismatches.

Most of the fMRI-EEG mismatches are fundamentally caused by highly different temporal scales in which fMRI and EEG data are generated and collected. Neural activities evolve so fast that the brain function is always carried out readily. In response to a single “event”, the evoked neural activity is most substantial within a very short period of time ranging from tens to hundreds milliseconds. Since neuronal events are accompanied with instantaneous electrical responses, the scalp EEG signals collected
with a sufficiently fast sampling rate carry the information about the “current” status of underlying neural activities (Nunez and Srinivasan 2005). On the other hand, the hemodynamic response measured by fMRI is much slower with a substantial delay. It has been shown that the BOLD response only happens about 1~2 seconds after the event onset (Boynton et al. 1996). Such a delay is significant relative to the short duration of neural activity. One can only infer from the fMRI data the “past” status of neural activity. The sluggish hemodynamics may be better appreciated by considering the neurovascular coupling system as a low-pass filter (or a temporal point spread function) (Logothetis et al. 2001; Friston et al. 1994), which effectively smooths out the rapid neural dynamics. As a result, the fMRI response reflects the energetic effect of neural activity averaged (or accumulated) over time (Boynton et al. 1996; Mathiesen et al. 1998; Martindale et al. 2003; Wan et al. 2006). Lastly, the fMRI acquisition is also limited by the scanning speed (typically up to 50 ms per slice), which is often too slow to probe the temporal aspect of neural activity. In short, the spatial locations of fMRI activations cannot be simply equated with those of electrical activities at every millisecond, considering that fMRI has much lower temporal resolution and specificity than EEG.

The fMRI-EEG mismatches can be further categorized into three types, namely fMRI extra sources, fMRI invisible sources and the fMRI displacement (Liu et al. 1998; Wagner et al. 2000; Liu et al. 2006b). The fMRI extra sources represent the source regions that are deemed as active in fMRI but do not contain the sources for the EEG at a certain time instant. During a short period of interest following the event onset, the fMRI activations have to be thought of as “static” (or time-invariant) while the EEG signals are variable and the source imaging is carried out instant by instant. The electrical source activity, in a time window with a scale of milliseconds, may only involve a subset of the activated fMRI areas, whereas other areas may appear as false positives if including them all in the prior spatial constraint (Liu et al. 2006b; Liu and He 2008a). The fMRI invisible sources are the real EEG sources but not deemed as active by fMRI. A transient current source may generate observable EEG signals whereas it may last too briefly to induce a sustained BOLD response. In this condition, the fMRI-derived time-invariant spatial constraint includes false negatives, which often result in the under-estimation of fMRI
invisible sources as reported in several independent studies (Liu et al. 1998; Liu et al. 2006b; Liu and He 2008). The fMRI displacement is referred to the spatial difference between vascular and electrophysiological sources (Wagner et al. 2000; Liu et al. 2006b). Such displacement is fundamental and hardly attributable to the different temporal scale or resolution of both modalities. In summary, dealing with such unavoidable spatiotemporal mismatches is essential to establishing a reliable method for the fMRI-constrained EEG source analysis.

5.1.2 EEG-Informed fMRI Analysis

The EEG-informed fMRI analysis rests on the fMRI analysis frameworks (e.g. the GLM analysis) while incorporating temporal- or frequency-specific information available in EEG. Such a multimodal strategy differs from the conventional fMRI analysis in its unique ability to selectively localize the fMRI correlates to specific neuronal events or rhythms. This ability directly benefits from the EEG/MEG measurements which reflect the mass neural responses, whereas the analysis based on fMRI alone has to rely on the timing of stimuli or tasks. Note that, the approaches in this category usually prefer EEG to MEG due to the highly desired simultaneous recordings of fMRI and electrophysiological signals.

Techniques employing this strategy are particularly useful in stimulus- or task-free experimental conditions. Perhaps, the most typical example is the interictal spike related fMRI mapping, in which the interictal epileptiform events manifest themselves as spike-like discharges in EEG (Ives et al. 1993; Gotman et al. 2004, 2005, 2006). Some other examples are witnessed in studies exploring the neural substrates underlying the rhythmic modulations in the resting or pathological brain (Goldman et al. 2002; Laufs et al. 2003; Moosmann et al. 2003; Kaufmann et al. 2006; Horovitz et al. 2008). In these studies, predictors for the fMRI regression analysis are derived from frequency-specific EEG modulations. By assessing the correlation between the fMRI signals and the EEG-defined predictors, one may localize the neural regions responsible for the generation of the rhythmic modulations of interest.

A more sophisticated approach is based on parametric task manipulations and the single-trial EEG-fMRI co-variation (Eichele et al. 2005; Debener et al. 2006). A range of
parametrically graded experimental conditions are employed to identify cortical regions for which the BOLD response shows the same modulation across conditions as a specific single-trial ERP component. More specifically, external stimuli are designed to induce the variation of single-trial EEG responses. This single-trial variability is defined with respect to each time point within a trial. The signal varying over trials can serve as the predictor in the fMRI regression analysis, yielding a map of fMRI correlates. Repeating the analysis at multiple or potentially all latencies, one can obtain a series of fMRI maps, each of which is associated with a specific instant within the time scale of ERP. This approach, although interesting, is demanding in signal processing. The single-trial variability can result from spontaneous activity and noise that may or may not be differentiable from that driven by external stimuli. Moreover, it requires a careful design of stimulus properties to allow for an effective control of single-trial amplitudes.

Compared to the fMRI-constrained EEG/MEG source imaging, the EEG-informed fMRI analysis appears to be wholly different. In fact, both types of methods share an important commonality in their fundamental assumption—that is, the neuronal electrophysiological response is linearly correlated with the BOLD fMRI signal. To date, it is difficult to judge which method is superior to the other, although the EEG-informed fMRI analysis is formulated in a better posed manner than the fMRI-constrained source imaging.

### 5.2 Methods for fMRI-Constrained Electrical Source Imaging

In this subsection, we describe several methods for the fMRI-constrained electrical source imaging. Since all of these methods are some kinds of algorithms for solving the EEG inverse problem with helpful inputs derived from fMRI, we first introduce a generic linear EEG inverse solver known as the Wiener filter (5.2.1). Deriving from this generic inverse solver, we propose two novel algorithms (5.2.3 and 5.2.4) that fuse the cross-modal information in a more principled way, as compared to the conventional fMRI-weighted current density reconstruction method (5.2.2). Finally we describe the fMRI-seeded dipole fitting technique in 5.2.5.
5.2.1 Wiener Filter for EEG Source Reconstruction

As discussed in 2.2.2, a transfer matrix $A$ can be numerically computed to link the underlying current source distribution, $s(t)$, to the recorded scalp potentials, $x(t)$, with the existence of recording noise, $b(t)$, for every time instant $t$.

$$x(t) = As(t) + b(t)$$ (5.1)

where $A$ is an $N_x$-by-$N_s$ matrix ($N_x$ is the number of EEG sensors, $N_s$ is the number of current sources), $s(t)$ is an $N_s$-by-1 vector, $x(t)$ and $b(t)$ are $N_x$-by-1 vectors.

At a certain time instant, the spatial vectors $s(t)$, $x(t)$ and $b(t)$ can be viewed as stochastic processes with their index sets over source or sensor locations. Assuming both $s(t)$ and $b(t)$ have zero means, we define the source auto-covariance matrix and the noise auto-covariance matrix as Eq. (5.2) and Eq. (5.3), respectively.

$$C_s(t) = E\{s(t)s^T(t)\}$$ (5.2)

$$C_b(t) = E\{b(t)b^T(t)\}$$ (5.3)

If a priori information is given to both $C_s(t)$ and $C_b(t)$, a linear inverse operator $G(t)$ can be used to estimate $s(t)$ from $x(t)$, denoted as $\hat{s}(t)$.

$$G(t) = C_s(t)A^T(AC_s(t)A^T + C_b(t))^{-1}$$ (5.4)

$$\hat{s}(t) = G(t)x(t)$$ (5.5)

In practice, the noise covariance matrix, $C_b(t)$, can be estimated directly from the EEG data (Fuchs et al. 1998), whereas a priori knowledge of $C_s(t)$ is usually unavailable, and hence it is typically assumed to be proportional to an identity matrix (Hämäläinen and Ilmoniemi 1984) or a spatial Laplacian operator (Pascual-Marqui et al. 1994). It is worthwhile to emphasize that in the Wiener filter formulation, the stochastic process is referred to the spatial domain instead of the time domain.

The Wiener filter as Eq. (5.4) is equivalent to a linear operator solving the following weighted minimum norm cost function.

$$\Omega(s(t)) = \|W_s(As(t) - x(t))\|^2_2 + \lambda(t)\|W_s s(t)\|^2_2$$ (5.6)
where $W_s$ and $W_w$ are the square weighting matrices in the sensor and source spaces respectively, $\lambda(t)$ is the regularization parameter that balances the contributions to the cost function from the first data fitting term and the second weighted minimum norm regularization term.

Since Eq. (5.6) is quadratic in nature, the optimal solution of $s(t)$ that minimizes $\Omega(s(t))$ is a linear function of $x(t)$ and can be computed as follows.

$$\frac{d\Omega(s(t))}{ds} = 0$$

$$s(t) = H(t)x(t)$$

$$H(t) = (W^T_s W_s)^{-1} A^T W^T_s \left[ W_x A (W^T_s W_s)^{-1} A^T W^T_s + \lambda(t)I \right]^{-1} W_x$$

Assume $W_s$ and $W_w$ are invertible, Eq. (5.9) can be re-written as Eq. (5.10)

$$H(t) = (W^T_s W_s)^{-1} A^T \left( (W^T_s)^{-1} \right) \left[ W_x A (W^T_s W_s)^{-1} A^T W^T_s + \lambda(t)I \right]^{-1} (W^T_s)^{-1}$$

$$= (W^T_s W_s)^{-1} A^T \left[ A (W^T_s W_s)^{-1} A^T + \lambda(t) (W^T_s W_s)^{-1} \right]^{-1}$$

$$= (W^T_s W_s)^{-1} A^T \left[ A (W^T_s W_s)^{-1} A^T + \lambda(t) (W^T_s W_s)^{-1} \right]^{-1}$$

Eq. (5.10) is equivalent to Eq. (5.4), if satisfying

$$\left( W^T_s W_s \right)^{-1} = C_s(t)$$

$$\lambda(t) \left( W^T_s W_s \right)^{-1} = C_b(t)$$

If we further assume $W_s$ and $W_w$ are diagonal matrices, meaning that the weight factor is separately applied to each individual sensor or source location without any cross-location term, the weighting matrices can be determined as

$$W_s = (C_s(t))^{-1/2}$$

$$W_w = \sqrt{\lambda(t)} \cdot (C_b(t))^{-1/2}$$

Eq. (5.13) suggests that at each source location, the weighting factor is inversely proportional to the source standard deviation. Since a higher source standard deviation means a higher likelihood of having a non-zero source activity, such a setting for $W_s$ favors source locations with larger prior variances over those with smaller prior variances.
Eq. (5.14) suggests that at each sensor location, the weighting factor is inversely proportional to the noise standard deviation. Since a higher noise standard deviation means a less reliability of the recorded data, the source estimates based on such a setting for $W_x$ relies more on the data recorded from channels with smaller noise sensitivity. The above interpretation justifies the settings for the weighting matrices, in consistency with the Wiener filter, in the weighted minimum norm formulation.

### 5.2.2 fMRI-Weighted Current Density Estimation

In this subsection, we describe the methods for the fMRI-weighted current density estimation, which originate from the earliest development of the fMRI-EEG integrated neuroimaging (Dale and Sereno 1993; Liu et al. 1998). To date, these methods are still widely used in neuroimaging and neuroscience fields, perhaps because of two popular and commercialized electromagnetic source imaging software: CURRY (Neuroscan, Charlotte, NC, USA) and BESA (MEGIS Software GmbH, Gräfelfing, Germany) that include the implementations of these algorithms.

The fMRI-weighted current density estimation has been proposed and formulated in the context of either the Wiener filter (Dale and Sereno 1993; Liu et al. 1998; Dale et al. 2000) or the weighted minimum norm (Wagner et al. 2000; Babiloni et al. 2003; Ahlfors and Simpson 2004). Based on the discussion in 5.2.2, these two different ways of implementation are fundamentally equivalent and interchangeable.

Regardless of the different theoretical framework, the key idea in such methods is to derive spatial prior constraints from the fMRI activation map. For instance, in the context of the Wiener filter, the source covariance matrix is assumed to be time-invariant, and the diagonal elements of the source covariance matrix are determined by whether the corresponding locations are within fMRI activations or not. For an activated fMRI voxel, a weighting factor $f$ is assigned; otherwise, a control weighting factor 1 is used.
Figure 5.1 Illustration of the construction of fMRI weighting matrix for the fMRI-weighted current density estimation. \( f \) is the fMRI-weighting factor, which specifies the degree of preference to the source locations inside the fMRI activations, relative to those outside.

Mathematically, we assume the source covariance matrix is proportional to a static feature matrix, denoted as \( R \), by a factor of \( \gamma(t) \).

\[
\mathbf{C}_i(t) = \gamma(t) \mathbf{R}
\]  

(5.15)

The feature matrix is diagonal, and the diagonal elements are derived from the fMRI activation map, as illustrated in Fig. 5.1.

\[
\mathbf{R} = \text{diag}(r_1, r_2, \ldots, r_N)
\]  

(5.16)

\[
r_i = \begin{cases} 
  f & \text{the } i\text{th voxel is inside the fMRI activation} \\
  1 & \text{the } i\text{th voxel is outside the fMRI activation}
\end{cases}
\]  

(5.17)

Here, the proportional factor \( \gamma(t) \) serves as the regularization parameter, similarly as \( \lambda(t) \) in Eq. (5.6) and Eq. (5.9). Both parameters are related as expressed in Eq. (5.18).

\[
\lambda(t) = 1/\gamma(t)
\]  

(5.18)
Both $\gamma(t)$ and $\lambda(t)$ can be determined from the EEG data by using a variety of methods, such as L-curve (Hansen 1992), CRESO, discrepancy principle.

Once the regularization parameter is chosen, the fMRI weighting factor $f$ is the only unknown parameter that remains to be determined. The value of $f$ controls the degree of preference to the source locations deemed as active in fMRI relative to those outside the fMRI activated regions. In practice, the choice of $f$ is often empirical instead of theoretically driven. Based upon results from several independent computer simulation studies, the optimal value of $f$ has been proposed to be 3 (Babiloni et al. 2003), 6.7 (Wagner and Fuchs 2001) or 10 (Liu et al. 1998; Babiloni et al. 2003; Ahlfors and Simpson 2004). Particularly when $f = 10$, the algorithm is known as the 90% fMRI-weighted current density imaging, applications of which have been well received in neuroimaging and neuroscience fields (Dale et al. 2000; Babiloni et al., 2005).

The limitations of the fMRI-weighted current density imaging are two folded.

a) The fMRI weighting factor is chosen empirically instead of in a data-driven or principled manner. It leads to complications as of the interpretation of the reconstructed source images.

b) The assumption that the source auto-covariance matrix is proportional to a time-invariant matrix, as expressed in Eq. (5.15), is not technically valid. A time-invariant source covariance matrix implies that the source signal at any specific location is a stationary stochastic process over time and its sampled values at all time points are drawn independently from an identical statistical distribution. However, such an assumption is practically valid because neural networks in the brain are always carried out by source signals with rapidly and coordinated temporal evolvement that are certainly not exchangeable over time.

5.2.3 fMRI-Constrained Adaptive Wiener Filter

To overcome the above limitations, we propose a new algorithm for the fMRI-EEG data fusion. This algorithm utilizes the quantitative relationship between BOLD and EEG source signals, derived and discussed in Chapter 3, and uses the Wiener filter in an
adaptive and instant-by-instant manner to solve the EEG inverse problem (Liu and He 2008).

**fMRI-EEG Co-registration**

The quantitative relationship derived in Chapter 3 allows us to co-register the time integral of the estimated EEG source power with the BOLD effect size. As implied by the linearity of neurovascular coupling, the BOLD effect size quantified from the fMRI time series is proportional to the time integral of the EEG source power during a short event-related period. Mathematically, it is expressed as

$$\int_0^{T_s} s^2(r, t) \, dt \propto \beta(r)$$

(5.19)

where $s^2(r, t)$ is the source power at the location $r$ and the time $t$, $\beta(r)$ is the quantified BOLD effect size at the location $r$. $T_s$ is the duration of the electrophysiological response induced by a transient stimulus or task, which is referred to as the “event-related period”.

Unfortunately, $s^2(r, t)$ is not accessible from either EEG or fMRI. However, the shape of its time course can be reasonably well retrieved by solving the EEG inverse problem, although the absolute magnitudes of source estimates are often distorted and under-estimated due to the ill-posedness of the inverse problem. As shown in both Eq. (5.5) and Eq. (5.8), the EEG inverse solution results from the spatial de-convolution of measured scalp potential maps. Since the inverse problem solver is essentially a spatial operator, it causes significantly less distortion or bias to the reconstructed temporal “waveform” than to the spatial “image”.

80
Separate analyses of fMRI and EEG provide the complementary but closely related information about the spatiotemporal source power distribution. Specifically for the source power at each voxel, its magnitude-independent temporal evolvement is assumed to be identical as the normalized time course of the source power estimate obtained from EEG alone, while the time integral of the source power during the event-related period is constrained by the BOLD effect size quantified from fMRI alone. Such a process that simply merges the results from separate fMRI and EEG analyses is referred to as the fMRI-EEG co-registration.

Mathematically, we further describe the above fMRI-EEG co-registration process in Eq. (5.20) and Eq. (5.21). Let $G$ stand for the linear EEG inverse operator derived from the EEG forward model. When applying $G$ to the EEG data, we transform the data in the sensor space to the source space, yielding a set of current source estimates, denoted as $\hat{s}_{EEG}(t)$.
\[ \tilde{s}_{\text{EEG}}(t) = Gx(t) \] (5.20)

We normalize the power of the source estimates using Eq. (5.21), yielding a set of fMRI-EEG co-registered source power, denoted as \( \tilde{s}^2(t) \). Note that such a normalization process ensures that at every source location \( r \), \( \tilde{s}^2(r,t) \) satisfies the neurovascular coupling relationship specified in Eq. (5.19).

\[
\tilde{s}^2(r,t) = \tilde{s}_{\text{EEG}}^2(r,t) \beta(t) \int r_{\text{EEG}}^2(r,t) dt 
\] (5.21)

The fMRI-EEG co-registration is also illustrated in Fig. 5.2. Encouragingly, the co-registration process brings high spatial resolution comparable to fMRI and high temporal resolution comparable to the EEG source imaging.

**Adaptive Wiener Filter**

Although the aforementioned fMRI-EEG co-registration method has considerable merits as compared to fMRI or EEG alone, it fails to consider the possible difference between locations of the vascular sources identified by fMRI and the neuronal current sources underlying EEG, or to handle the possible inaccuracy of the neurovascular coupling model, or to reduce the temporal cross-talk in source estimates obtained from the EEG alone. This is essentially because the spatial and temporal aspects of neural activity are separately obtained from fMRI and EEG, respectively.

Since the electrophysiological activity is of ultimate interest in the spatiotemporal functional neuroimaging, we propose to perform an additional step to re-fit the source activity to the EEG data while taking the fMRI-EEG co-registered source power estimates as time-variant prior spatial constraints. This additional step also ensures that the spatiotemporal source reconstruction is primarily dependent upon EEG measurements, and thus provides the robustness against inaccurate or mismatched fMRI constraints.

Specifically, we estimate the current density distribution instant by instant by using the Wiener filter, in which the source covariance is variable over time and derived from the fMRI-EEG co-registered source power. The time-variant source covariance is based upon the statistical interpretation of the fMRI-EEG co-registered source power as the time-dependent variance of source activity over the “epochs”, which is defined by the
onsets of repeated stimuli or tasks. In contrast to the conventional Wiener filter method (or the fMRI-weighted methods) with a time-invariant source covariance matrix (Dale and Sereno 1993; Liu et al. 1998; Dale et al. 2000; Lin et al. 2006), we shall refer to the proposed approach as the *Adaptive Wiener Filter* (AWF), since the time-variant source covariance matrix is derived in a data-driven manner (Liu and He 2008). In the ensuing mathematical derivations, we emphasize the statistical interpretation of the fMRI-EEG co-registered spatial prior constraints in the context of Wiener filter (see 5.2.1).

As a common practice in ERP experiments, the electrophysiological response is repeatedly induced by a train of stimuli (or tasks). The EEG response in each repetition is called an EEG “epoch”, which is usually time-locked to the stimulus onset.

![Figure 5.3](image)

**Figure 5.3** The source signal at any *i*-th location *s∗(_i_, t) is not a stationary stochastic process over time, but its values at any given time *t_j* sampled in different epochs are i.i.d. Thus, *s*(t, k), for 1 ≤ k ≤ N_e, represents an independent observation of a stochastic process *s*(t) for each individual time *t* (note that *s*(t) is a stochastic process with its index sets over source locations).
Let \( x(t,k) \), \( s(t,k) \) and \( b(t,k) \) stand for the random column vectors of EEG recordings (\( N_x \)-by-1), source signals (\( N_s \)-by-1) and measurement noise (\( N_x \)-by-1), respectively, at the time \( t \) and in the \( k \)-th epoch. A valid assumption is that source signals in different epochs represent independent observations of the identical stochastic process, \( s(t) \), for any specific time \( t \). Fig. 5.3 illustrates this concept.

By collecting signals at the time \( t \) from all \( N_e \) epochs, we define

\[
X(t) = [x(t,1), x(t,2), \ldots, x(t,N_e)]
\]

(5.22)

\[
S(t) = [s(t,1), s(t,2), \ldots, s(t,N_e)]
\]

(5.23)

\[
B(t) = [b(t,1), b(t,2), \ldots, b(t, N_e)]
\]

(5.24)

And, we re-write the forward equation in a matrix notation.

\[
X(t) = AS(t) + B(t)
\]

(5.25)

The singular value decomposition (SVD) of \( X(t) \) is written as Eq. (5.26)

\[
X(t) = U(t)A(t)V^T(t)
\]

\[
= \sum_{q=1}^{N_e} \sigma_q(t)U_q(t)V_q^T(t)
\]

(5.26)

where \( N_q = \min(N_x, N_e) \), the singular vector \( U_q(t) \) is the \( q \)-th column vector in the matrix \( U(t) \) that represents a spatial component of scalp potentials at the time \( t \), the singular vector \( V_q(t) \) is the \( q \)-th column vector in the matrix \( V(t) \) that represents the variation (over epochs) of the corresponding spatial component \( U_q(t) \).

We truncate the spatial components that do not satisfy the discrete Picard condition (Hansen 1990). The truncated components are not spatially smooth enough to be associated with any source activity, and hence are dominated by noise perturbation.

After the truncation, we compute the maximum likelihood estimates (MLE) of distributed sources that account for each of the \( Q \) remaining spatial components.

\[
\tilde{s}_q(t) = A^T (AA^T + r_q I)^{-1} U_q(t) \quad (1 \leq q \leq Q)
\]

(5.27)

where \( \tilde{s}_q(t) \) is an \( N_x \)-by-1 vector that represents the estimated source distribution underlying the \( q \)-th spatial component \( U_q(t) \) at the time \( t \), and the regularization
parameter $r_q$ can be obtained by using the methods like the “L-curve” method (Hansen 1992) and etc.

Combining the MLE for all Q components, we can obtain a set of source estimates for the time $t$ in all $N_e$ epochs, collectively denoted as $\tilde{S}(t)$.

$$\tilde{S}(t) = \sum_{q=1}^{Q} \lambda_q(t) \tilde{s}_q(t) \mathbf{V}_q^T(t)$$

(5.28)

where $\tilde{S}(t)$ is an $N_s$-by-$N_e$ matrix with each row representing the variation of estimated source signals over epochs.

According to Eq. (5.2), the MLE of source variance at the $i$-th location for the time $t$, denoted as $\tilde{\sigma}_i(t)$, can be written as Eq. (5.29).

$$\tilde{\sigma}_i(t) = \frac{1}{N_e} \tilde{s}_i(t) \tilde{S}_i^T(t)$$

(5.29)

where $\tilde{s}_i(t)$ is the $i$-th row vector in the matrix $\tilde{S}(t)$.

Substitute $\tilde{S}(t)$ by Eq. (5.28) and re-write Eq. (5.29) as Eq. (5.30).

$$\tilde{\sigma}_i(t) = \frac{1}{N_e} \left( \sum_{q=1}^{Q} \lambda_q(t) \tilde{s}_q(t) \mathbf{V}_q^T(t) \right)^T \left( \sum_{q=1}^{Q} \lambda_q(t) \tilde{s}_q(t) \mathbf{V}_q^T(t) \right)$$

$$= \frac{1}{N_e} \left( \sum_{q=1}^{Q} \lambda_q(t) \tilde{s}_q(t) \mathbf{V}_q^T(t) \right)$$

(5.30)

Since $\mathbf{V}_q^T(t)\mathbf{V}_q(t) = 1$ and $\mathbf{V}_p^T(t)\mathbf{V}_q(t) = 0$ for any $p \neq q$, Eq. (5.30) can be simplified as Eq. (5.31).

$$\tilde{\sigma}_i(t) = \frac{1}{N_e} \sum_{q=1}^{Q} \left( \lambda_q(t) \tilde{s}_q(t) \right)^2$$

(5.31)

Repeat the above procedures for all the time points during the period $T_s$, a time course of source variance estimates can be computed for every source location $i$, collectively denoted as a vector $\tilde{\sigma}_i$.

$$\tilde{\sigma}_i = [\tilde{\sigma}_i(1), \tilde{\sigma}_i(2), \ldots, \tilde{\sigma}_i(T_s)]$$

(5.32)
Similar to Eq. (5.21), we can normalize $\tilde{\sigma}_i$ so that its time integral equals the quantified BOLD effect size, as expressed by Eq. (5.33).

$$\tilde{\sigma}_i(t) = \tilde{\sigma}_i(t) \beta / \sum_{\text{rel}_i} \tilde{\sigma}_i(t)$$ (5.33)

Eq. (5.33) combines the fMRI constraint to the time integral of source variance and the MLE of source variance. Therefore, it provides an estimate of the time-variant source covariance matrix, denote as $\tilde{\mathbf{C}}_s(t)$, using the multimodal datasets. Assuming no a priori cross-correlation between different source locations, we define

$$\tilde{\mathbf{C}}_s(t) = \text{diag}(\tilde{\sigma}_1(t), \tilde{\sigma}_2(t), ..., \tilde{\sigma}_N(t))$$ (5.34)

Note that $\tilde{\mathbf{C}}_s(t)$ may differ from the true source covariance matrix by a proportional factor $\gamma_s(t)$, as expressed in Eq. (5.35). This is because the BOLD effect size $\beta_i$ is proportional (instead of equal) to the time integral of the source variance, whereas Eq. (5.33) enforces the equality between them.

$$\mathbf{C}_s(t) = \gamma_s(t) \cdot \tilde{\mathbf{C}}_s(t)$$ (5.35)

In addition to the conventional regularization techniques, we can estimate $\gamma_s(t)$ from the EEG data covariance matrix and the noise covariance matrix, using the method described as below.

$$\mathbf{C}_s(t) = E\{\mathbf{x}(t)[\mathbf{x}(t)]^T\}$$

$$= E\{[\mathbf{A}\mathbf{s}(t) + \mathbf{b}(t)][\mathbf{A}\mathbf{s}(t) + \mathbf{b}(t)]^T\}$$

$$= E\{[\mathbf{A}\mathbf{s}(t)][\mathbf{s}(t)]^T \mathbf{A}^T + \mathbf{A}\mathbf{s}(t)[\mathbf{b}(t)]^T + \mathbf{b}(t)[\mathbf{s}(t)]^T \mathbf{A}^T + \mathbf{b}(t)[\mathbf{b}(t)]^T\}$$ (5.36)

Assuming that the source signal and the recording noise are independent to each other, we have $\mathbf{s}(t)[\mathbf{b}(t)]^T = 0$ and $\mathbf{b}(t)[\mathbf{s}(t)]^T = 0$. This assumption also implies that $\mathbf{C}_b(t)$ should be time independent since the time $t$ is specifically referred to the onset of the stimulus or task execution.

We can re-write Eq. (5.36) as Eq.(5.37), and compute $\gamma_s(t)$ from Eq. (5.38).

$$\mathbf{C}_s(t) = \mathbf{A}\mathbf{C}_s(t)\mathbf{A}^T + \mathbf{C}_b$$

$$= \gamma_s(t)\mathbf{A}\tilde{\mathbf{C}}_s(t)\mathbf{A}^T + \mathbf{C}_b$$ (5.37)
\[ \gamma_s(t) = \frac{tr(C_s(t)) - tr(C_b)}{tr(\bar{A}C_s(t)A^T)} \]  

(5.38)

where \( tr(\cdot) \) computes the sum of the diagonal elements in an input matrix.

Accordingly, the spatiotemporal source activity can be reconstructed by using the adaptive Wiener filter, written as Eq. (5.39).

\[ G(t) = \gamma_s(t)\bar{C}_s(t)A^T(\gamma_sA\bar{C}_s(t)A^T + C_b)^{-1} \]  

(5.39)

As described in this subsection and the previous subsection, one could estimate the source variance simply from the averaged EEG signals (i.e. ERP) or the segmented signals before response averaging, in combination with the BOLD effect size quantified from fMRI data. It is worth emphasizing the difference and commonality between them. In fact, the source variance estimation based on the ERP (see Eq. (5.20)) is a special case of the generalized method based on the segmented epochs (see Eq. (5.26) through (Eq. (5.31)). Since ERP is obtained from the average of the segmented EEG data, it is composed of the DC components of the inter-trial variable EEG signals viewed at every time instant. Such DC components usually appear in the first (or at least first few) component(s) after the SVD of the spatial-trial data matrix (see Eq. (5.26)). However, the generalized method described in this subsection, is capable of retrieving the non-DC single-trial variability with slow fluctuation, which may reflect the non-phased locked response or other system-level modulations.

### 5.2.4 Twomey Regularization

As opposed to the time-variant spatial constraint used in the fMRI-constrained AWF algorithm (see 5.2.3), we have developed another algorithm based on the static fMRI-derived spatial constraint (Liu et al. 2006b). This algorithm is similar to the fMRI-weighted current density estimation (see 5.2.2), in the sense that the spatial constraint is exclusively derived from the fMRI activation map. However, instead of applying different empirical weighting factors for locations inside vs. outside fMRI activations, our algorithm consists of two steps. In the first step, the EEG source space is confined to the regions deemed as active in fMRI, yielding to a set of “hard” constrained source images. In the second, the source images obtained from the first step is re-entered as the
initial guess of the desired solution into an EEG least-squares fitting procedure with Twomey regularization (Twomey 1963). As such, this algorithm is called Twomey algorithm (Liu et al. 2006b).

In the first step, the inverse solution is obtained through the fMRI-weighted current density estimation (see 5.2.2) with the fMRI weighting factor as large as 100, as expressed in Eq. (5.40) and Eq. (5.43).

\[
\hat{s}_0(t) = G_0 x(t) \tag{5.40}
\]

\[
G_0 = \gamma(t) R_f A^T (\gamma(t) A R_f A^T + C_b)^{-1} \tag{5.41}
\]

\[
R_f = [r_1, r_2, \ldots, r_{N_3}] \tag{5.42}
\]

\[
\tau_i = \begin{cases} 
1, & \text{inactive fMRI voxel} \\
100, & \text{active fMRI voxel}
\end{cases} \tag{5.43}
\]

As would be expected, such an over-constrained high-resolution cortical source reconstruction is highly sensitive to invalid fMRI priors, arising from the fMRI-EEG mismatches discussed in 5.1.1. To overcome the resulting distortion in the second step, \(\hat{s}_0(t)\) is re-entered as an initial estimate into the EEG least square fitting procedure with Twomey regularization (Twomey 1963), which minimizes the difference between the desired solution and the initial guess. The cost function for the Twomey-regularized EEG least-squares fitting can be written as (5.44)

\[
\Omega(t) = \| C_b^{-1/2} [A s(t) - x(t)] \|^2_2 + \lambda(t) \| s(t) - \hat{s}_0(t) \|^2_2 \tag{5.44}
\]

We can see that the influence from fMRI spatial priors has been softened in the second step of Twomey algorithm, which yields the robustness of the estimated solution in face of the possible fMRI-EEG misspecifications. If \(\lambda\), the regularization parameter that balances the two terms in the cost function, is chosen to be a large value, the solution is forced to be very close to \(\hat{s}_0(t)\), and the solution is dominated by the fMRI spatial prior and may be still sensitive to the presence of fMRI-EEG mismatches sources. Conversely, if \(\lambda\) is small, the solution tends to shift away from \(\hat{s}_0(t)\) in a way that further reduces the residual norm. As a result, the source estimation has good chance to be corrected against the influence from fMRI-EEG mismatches through better fitting to EEG recordings. And if \(\lambda\) is virtually close to zero, the solution turns to a purely least square inverse solution, suffering from unstableness and low spatial resolution.
Clearly, the trade-off is controlled by $\lambda$. We choose $\lambda$ by using the “L-curve” approach, which plots the first term of residual norm versus the norm of the discrepancy between the desired solution and the initial solution in log-log scale. The value of regularization parameter at the corner of the L-shaped curve represents the desired setting that compromises the minimization of the two terms in (Hansen 1992). Once $\lambda$ is determined, the solution that minimizes the cost function in Eq. (5.44) can be obtained from Eq. (5.45) or Eq. (5.46).

$$\hat{s}(t) = (A^T C^{-1}_b A + \lambda I)^{-1} \left( A^T C^{-1}_b x(t) + \lambda \hat{s}_0(t) \right)$$  \hspace{1cm} (5.45)

$$\hat{s}(t) = \hat{s}_0(t) + A^T (AA^T + \lambda(t) C_b)^{-1} (x(t) - A \hat{s}_0(t))$$  \hspace{1cm} (5.46)

Although both equations above are equivalent, Eq. (5.46) is more computationally efficient than Eq. (5.45), since it only requires computing the inverse of an $N_x$-by-$N_x$ matrix instead of a much larger $N_s$-by-$N_s$ matrix.

Eq. (5.46) also indicates that Twomey algorithm ends up with a linear inverse operator, written as Eq. (5.47).

$$G = G_0 + A^T (AA^T + \lambda(t) C_b)^{-1} (I - AG_0)$$  \hspace{1cm} (5.47)

The linear inverse operator facilitates the evaluation of the resultant inverse solution through a well-defined index, known as the “point-spread function” (Liu et al. 1998; Dale et al. 2000). See 5.3.3 for more details.

### 5.2.5 fMRI-Seeded Dipole Fitting

All of the fMRI-EEG integration methods, described in 5.2.1 through 5.2.4, aim at imaging the spatiotemporal current density distribution. The end products of these methods can be a movie of images showing where the brain activity is located and how it evolves over time. In this subsection, we describe the fMRI-seeded dipole fitting technique. As opposed to the current density imaging techniques, the primary goal of this technique is not to image brain activity. Rather, it aims at retrieving the time course of brain activity at well-defined fMRI activation foci.

If the neural activity underlying a certain sensory or cognitive brain function is focal (or relatively focal), we can identify a number of fMRI foci (or hotspots) from the fMRI activation map. The focal nature allows us to model the electrical activity at each
fMRI hotspot as an equivalent regional current dipole. The dipole locations are fixed but unknown whereas the dipole orientations and moments are variable over time. All the dipole parameters including location, orientation and moment can be estimated from the EEG data.

Firstly, a nonlinear optimization problem needs to be solved to determine the dipole locations such that the scalp potentials generated by the modeled dipoles can best explain the measured EEG signals. Assume \( m \) dipoles at \( m \) fMRI foci. Each dipole is represented by a pair of dipole location and moment, denoted as \( (\vec{r}_i, \vec{p}_i(t)) \), with 6 parameters (i.e. 3 for the location and 3 for the moment). In total, the multi-dipole model contains \( 6m \) unknown parameters to be determined from EEG. The number of unknowns can be further reduced to \( 3m \), because once the locations are known, the dipole moments can be uniquely computed from EEG through the linear least-squares estimation. In other words, the dipole moments are functions of the dipole locations and the EEG data.

Mathematically, the lead-field matrix \( F \) for all \( m \) dipoles can be obtained from the head volume conductor model (see 2.2.2), and \( F \) is a function of dipole locations.

\[
F(\vec{r}_1, \vec{r}_2, ..., \vec{r}_m) = \\
[\begin{array}{c}
g(\vec{r}_1, \vec{a}_x), g(\vec{r}_1, \vec{a}_y), g(\vec{r}_1, \vec{a}_z), ..., g(\vec{r}_m, \vec{a}_x), g(\vec{r}_m, \vec{a}_y), g(\vec{r}_m, \vec{a}_z) 
\end{array}]
\]  

(5.48)

where \( g(\cdot) \) represents the function for the computation of the forward solution, \( \vec{a}_x, \vec{a}_y, \vec{a}_z \) are the unitary vector in x-, y- and z-direction respectively.

It follows that the estimation of the dipole moments is an over-determined linear inverse problem, if \( 3m \leq N_x \).

\[
[\vec{p}_1(t), \vec{p}_2(t), ..., \vec{p}_m(t)] = (F^TF)^{-1}F^T\mathbf{x}(t) = G(\vec{r}_1, ..., \vec{r}_m)\mathbf{x}(t) 
\]  

(5.49)

The dipole locations are estimated by minimizing the following nonlinear cost function.

\[
\Omega(\vec{r}_1, ..., \vec{r}_m) = \sum_{t=1}^{T} ||G(\vec{r}_1, ..., \vec{r}_m)\mathbf{x}(t) - \mathbf{x}(t)||_2^2 
\]  

(5.50)

The nonlinearity of the above cost function suggests that the fitted dipole locations are highly dependent on the initial locations with which the optimization procedure begins. It is such a demand that well justifies the usefulness of fMRI for “seeding” the initial dipole locations, since the fMRI foci are presumably close to where
the neural activity takes place. Therefore, we take the fMRI seeding locations as the centers of gravity of every fMRI hotspot.

5.3 Evaluation of fMRI-Constrained Electrical Source Imaging

In order to compare and evaluate the above fMRI-constrained electrical source imaging algorithms, we performed a series of computer simulation and well-controlled experimental studies.

5.3.1 Simulation for fMRI-Constrained Adaptive Wiener Filter

Computer Simulation Setting

To test the imaging performance of the fMRI-constrained AWF algorithm (described in 5.2.3), we simulated focalized brain activity at three locations on the posterior cortical surface embedded in a realistically shaped head volume conductor model. The geometries of the head and cortex were extracted from high resolution T1-weighted MR images of a human subject (256 slices, matrix size: 256×256, voxel size: 1×1×1 mm³). The electrical conductivities of the scalp, skull and brain were set to be 0.33, 0.0165, and 0.33 S/m, respectively (Lai et al. 2005; Oostendorp et al. 2000; Zhang et al. 2006). The cortical current density source model consisted of around 7,000 current dipoles evenly placed on the cortical surface. 128 electrodes with a standard montage were co-registered to the boundary element head model.

Based on the above settings, we simulated current source signals, BOLD fMRI signals and EEG recordings. We computed the inverse solutions from the simulated data using three algorithms. The detailed procedures are described step by step as follows.

Current source activities

3 current dipoles were placed on 3 selected locations on the folded cortical surface. To simulate the source activity, representative waveforms were assigned to these dipole sources such that the source signals were simulated to be temporally uncorrelated or correlated, or to be transient or sustained, or to have the same frequency but different phases or the same phase but different frequencies, or to have time courses obtained in a real experiment.
Figure 5.4 Realistic cortical surface and boundary element models with 128 scalp surface electrodes, as used in the computer simulation studies.

**BOLD-fMRI responses**

The BOLD-fMRI signals were generated from the simulated current sources. The source waveforms were repeated every 500 ms for a total of 30 s. For each source location, the BOLD signal was simulated by convolving the time course of source power with a gamma-function HRF suggested by (Boynton et al. 1996).

\[
h(t - \tau) = \frac{(t/\tau)^{n-1} e^{-(t/\tau)}}{\tau(n-1)!}
\]  

(5.51)

where \( n = 3 \), \( \tau = 1.25 \) sec and \( \varepsilon = 2.5 \) sec as in the default settings of a widely used fMRI analysis software - BrainVoyager QX (BrainInnovation, Netherlands).

The spatial distribution of the simulated BOLD signals was then convoluted with a Gaussian spatial kernel with a given full-width-half-maximum (FWHM). Here the FWHM was chosen to be 4 mm, which represented a reasonable value consistent with the known spatial resolution of fMRI according to previous experimental studies (Engel et al. 1997). Gaussian white noise (GWN) was added into the simulated BOLD time courses so that the BOLD response had a signal-to-noise ratio (SNR) equal to 10.

**EEG Signals**

From the simulated spatiotemporal current source distributions, we further simulated scalp EEG signals at 128 electrodes using the BEM-based forward computation (Hämäläinen and Sarvas 1989). 250 trials of EEG data were simulated as the
summation of the forward solution and GWN. The SNR in each trial was around 0.6 and the ERP signal averaged from all the trials had a SNR equal to 10.

Source imaging

Finally, we reconstructed the spatiotemporal source distribution based on the simulated EEG alone using the minimum-norm algorithm (Hämäläinen and Ilmoniemi 1984), or based on both the simulated fMRI and EEG using the 90% fMRI-weighted current density estimation (described in 5.2.2) and the fMRI-constrained adaptive Wiener filter (described in 5.2.3).

Results

BOLD-fMRI simulation and quantification

Fig. 5.5 shows the simulated BOLD fMRI signals and the quantified fMRI maps. The simulated fMRI responses were induced by a 30-sec block of stimuli with an ISI of 500 ms (Fig. 5.5.A). The electrical current sources were located at the lateral occipital sulcus (blue), the medial occipital dorsal area (red) and the intraparietal sulcus (green) on the right hemisphere (Fig. 5.5.B, right). Representative source locations were selected to be within the sulcal fundus, the sulcal wall and the gyral crown, respectively. The source waveforms shown in corresponding colors were simulated as 3 temporally uncorrelated gauss functions (Fig. 5.5.B left). After applying a 2-D spatial gauss smoother with a given FWHM (Fig. 5.5.C), the simulated BOLD responses emerged within 3 extended regions. The noise-contaminated fMRI signals surrounding the “red” source are plotted in Fig. 5.5.D). These time courses had a similar shape as the predictor signal (Fig. 5.5.E) such that they could be fitted with the predictor signal simply after scaling. The quantified fMRI response at each cortical location was precisely the scaling factor that allowed for the best fit between the scaled predictor signal and the measured BOLD response. This scaling factor was also found approximately proportional to the averaged BOLD signal change at the steady state with stimuli relative to the resting state without stimuli.
The map of quantified fMRI responses is shown in Fig. 5.5.F). Three activated regions were revealed. The extent of these regions depended upon the fMRI spatial resolution (or specificity), which was inversely proportional to FWHM. With an increasingly larger FWHM, the fMRI response extended from a point source to a larger region, in which the quantified values of fMRI responses became gradually smaller than the time integral of corresponding source powers (Fig. 5.5.G).


**Time-variant spatial constraints**

Fig. 5.6 demonstrates the method (described in 5.2.3) for probing the temporal change of source variance based on the segmented EEG epochs. To simplify the illustration, let us look closely at a period of time while a single source is activated.

The source location and waveform is displayed in Fig. 5.6.A). Since the ERP signal after response averaging had a SNR of 10, the EEG data in every epoch had an average SNR of 0.6. Fig. 5.6.B) shows the SVD components obtained from the data at the 40-ms latency in all the epochs. Only 3 components with largest singular values are shown. The singular value for the first decomposed component was found to be much larger than those for the second and third components. Clearly, the first spatial component\(^1\) more likely represented the scalp potential field generated by internal brain sources, whereas the second and third spatial components attributed to external recording noises. The first temporal component\(^2\), which reflected the variation of the global field strength over epochs, had a DC offset with small variation. However, dramatic oscillations around 0 were observed in the second and third temporal components. Based on these observations, we inferred that the first SVD component contained the scalp potential field arising from the source activity at 40 ms, as well as its variation over epochs. Applying the discrete Picard condition also ended up with the same conclusion that only the first component at this time instant satisfied the Picard condition while all other components should be truncated.

---

\(^1\) The spatial components are the left singular vectors in Eq. (26).

\(^2\) The temporal components are the right singular vectors in Eq. (26). Although called “temporal”, they are actually not the signals over time but rather the signals at the same latency over epochs.
Figure 5.6 An example of the SVD analysis on the segmented epochs. A) The EEG data was simulated from a single source activity, with the source location on the cortex shown on the left and the source waveform shown on the right; B) the largest 3 SVD components of the data at the peak latency (40 ms) over all the epochs. The singular values are shown in the left column, the decomposed spatial components shown in the middle and the corresponding variations over epochs shown in the right; C) the plot of
the largest singular value as a function of time; the spatial components associated with the largest singular value at 7 representative time points are displayed along the plotted time course.

For a better understanding of this conclusion and its implication, it is worthwhile to mention an important feature of SVD. With the singular value gradually decaying from its maximum value toward zero, both the left and right singular vectors necessarily tend to have more and more oscillations. Therefore, if the signal is less variable over epochs than the noise, which are the case in practical ERP studies, then the SVD component(s) with the largest singular value(s) mostly reflect the variance of source activities instead of that of the noise. On the other hand, these components also tend to be spatially smooth and more likely to satisfy the Picard condition. Accordingly, the DC line shown in Fig. 5.6.B) indicates that the source signal at 40 ms is repeated over epochs with little variation.

After repeating the SVD analysis to every time instant, we plotted the largest singular value as a function of time shown in Fig. 5.6.C). For several representative time points (from 10 to 70 ms stepped by 10 ms), the corresponding first spatial component was also displayed along the plotted time course. The source amplitudes at 10 and 70 ms were very small as shown in Fig. 5.6.A). The decomposed spatial components at these 2 latencies also appeared to be too noisy to satisfy the Picard condition. At other selected time points, the spatial patterns appeared to be similar and the resultant inverse solutions of these spatial patterns should also be spatially consistent to each other. These results demonstrate that the spatial component(s) with largest singular value(s), if satisfying the Picard condition, reflect the potential fields projected from the source space to the sensor space and contain the information on where the source is most likely located at the latency of interest.

By comparing Fig. 5.6.C) with Fig. 5.6.A), we also noticed that the temporal change of largest singular values was in general agreement with the time course of the real source amplitude. This finding further supported our previous argument that the principal SVD components nicely isolated the signal from the noise.
While the spatial component(s) with largest singular value(s) indicate the source locations at different latencies, the time course of largest singular values represents the temporal variation of source amplitude. In addition, the decomposed spatial components had the field strength equal to 1 because the singular vectors were always unitary vectors. Note that the absolute amplitudes of source estimates for the principal SVD components should be proportional to the largest singular values. Therefore, the similarity between Fig. 5.6.A) and 5.6.C) further justify the plausibility of estimating the time-variant source variance from the segmented epochs.

To assess the accuracies of the estimated time-variant source variances in both time and space, Fig. 5.7 summarizes the results of source variance estimation based on the EEG or the combination of fMRI and EEG. The source configuration and stimulation protocol used here were identical as in Fig. 5.5. 3 temporally uncorrelated (Fig. 5.7.B) sources were selected from the cortex (Fig. 5.7.A) to generate the BOLD fMRI and EEG signals. The quantified fMRI map with a 4-mm FWHM is shown in Fig. 5.7.D).

The MLE of source variances were computed from the segmented epochs. Fig. 5.7.E) plots the time courses of the source variance MLE at each of 3 source locations respectively. The shape of these time courses agreed with the corresponding source waveforms with mild distinction. The maximal source variance MLE arrived at exactly the same latencies (40, 100 and 160 ms) as the real source activities. 2 “bumps” were observed in the estimated time course of the “green” source, which was caused by the cross-talks in the inverse solution from the “blue” and “red” sources respectively. However, the absolute values of the estimated source variances tended to be much smaller than those of the real source signals, except for the “red” source which was located at the gyral crown.

Fig. 5.7.F) shows the spatial distribution of the source variance MLE at 3 peak latencies (40, 100 and 160 ms) identified from Fig. 5.7.E). From the segmented EEG, the estimated map of source variances clearly spread out and the spatial maximum was not correctly localized to the original location of the source that was activated at the corresponding time instant. Fig. 5.7.C) shows the spatial distribution of the mean MLE of source variance averaged over the entire time period. This spatial distribution is
qualitatively consistent with the quantified fMRI map (Fig. 5.7.D), although Fig. 5.7.C) exhibits spatial ambiguities and biases.

All these results based on the EEG alone demonstrate that the temporal change of source variances could be well retrieved from the segmented EEG epochs, but the absolute values of source variance were usually underestimated or significantly biased due to the ill-posed nature of the EEG inverse problem.

We further derived a set of time-variant spatial constraints by combining the “static” constraint from the fMRI and the “dynamic” information from the EEG. Fig. 5.7.G) shows the spatial distribution of the fMRI-EEG derived source variances at 3 peak latencies. In contrast to Fig. 5.7.D), Fig. 5.7.G) provided more accurate spatial prior constraints for each of 3 latencies respectively than the time-invariant spatial constraints based on the fMRI alone. The fMRI false-positives that occurred in the time-invariant fMRI constraints were successfully excluded from affecting the subsequent inverse solution. As opposed to Fig. 5.7.F), Fig. 5.7.G) also provided more localized and accurate prior constraints than those based on the EEG alone, by taking advantage of the high-spatial resolution of fMRI.
Figure 5.7 A) Source locations, B) source waveforms, C) the spatial distribution on the cortex of the average MLE of source variance averaged over the entire time period, D) the map of quantified BOLD fMRI responses, E) the time courses of the MLE of source variances at three source locations respectively, F) the instantaneous cortical distributions of the MLE of source variances at 3 peak latencies (40, 100 and 160 ms)
identified from E), G) the instantaneous cortical distributions of time-variant source variances derived from combination of fMRI and EEG at the 3 peak latencies.

**fMRI-EEG integrated source imaging with fMRI false positives**

Following the above example, we further imaged the cortical current density distribution at 3 peak latencies (40, 100 and 160 ms) using 3 algorithms: the proposed adaptive Wiener filter, the Wiener estimation using a time-invariant fMRI constraint and the minimum-norm estimation based on the EEG alone. Fig. 5.8.A) shows the reconstructed current density visualized on the cortical surface. Obviously, the use of time-variant spatial constraints in the adaptive Wiener filter algorithm resulted in the most accurate imaging results among all 3 algorithms. As shown in the top row in Fig. 5.8.A, 3 focal sources were revealed at each of 3 latencies respectively without showing any interference among them, in agreement with the fact that all 3 sources were temporally uncorrelated. The conventional Wiener filter also identified 3 sources with high spatial resolutions, benefiting from the use of fMRI spatial constraints. However, spurious sources were clearly observed at 40 ms and 100 ms. On the other hand, the minimum-norm solution without using any fMRI constraint had a low spatial resolution as often experienced in the EEG (or MEG) source imaging.
Figure 5.8 Source distribution reconstruction and source waveform estimation under fMRI false positives. The source locations and simulated source waveforms are the same as used in Fig. 5.7. A) Reconstructed cortical current density distribution at 3 peak latencies (40, 100 and 160 ms) shown in 3 columns, by using 3 different algorithms: the adaptive Wiener filter (1st row), the conventional Wiener filter using time-invariant spatial constraints (2nd row) and the weighted minimum norm based on the EEG alone (3rd row); B) Estimated source waveforms at each of 3 source locations color-coded as in Fig. 5.7. The estimated waveforms are shown in solid lines and the simulated original waveforms are shown in dashed lines. The 3 rows also correspond to 3 algorithms.

The time courses of source estimates at these 3 source locations using different algorithms are plotted (in solid lines) in Fig. 5.8.B) in comparison with the real source waveforms (in dashed lines). Note that the real and estimated waveforms are plotted in different scales as indicated at the left and right axes respectively. Among all 3 algorithms, the source waveforms reconstructed by using the adaptive Wiener filter were
the closest in both shape and absolute amplitude to the real source waveforms. The
estimated time courses using the conventional Wiener filter were clearly more accurate
than the minimum-norm solution; however compared to the adaptive Wiener filter,
obvious positive and negative “bumps” were observed obviously at 40 ms and slightly at
160 ms, which were essentially because of the fMRI false positive priors that amplified
the cross-talks among sources in the inverse solution.

_fMRI-EEG integrated source imaging with fMRI false negatives_

We further investigated the performance of the proposed adaptive Wiener filter
approach in dealing with the fMRI missing sources (i.e. fMRI false negatives). Again, 3
sources were selected from the cortical surface (see Fig. 5.9.A). The source waveforms
consisted of 2 gauss functions (for the “red” and “green” sources) that were temporally
uncorrelated and a short-duration pulse (for the “blue” source) with smaller peak
amplitude (see Fig. 5.9.C).

Since the “blue” source had an almost transient and smaller activity that was
insufficient to induce a large and sustained BOLD response, we could not identify this
source from the map of quantified BOLD responses (Fig. 5.9.B) and hence the “blue”
source behaved as an fMRI invisible source. This fMRI invisible source became
problematic in the conventional Wiener estimation of current density. At 40 ms when the
“blue” source was activated and reached its peak amplitude, the source activity at (or
around) the “blue” source location could not be imaged, as shown in the middle row of
Fig. 5.9.D). Instead, the spurious activity at an instantaneous fMRI false positive region
(around the “green” source) was observed. At the other 2 latencies, spurious source
activities with small amplitudes still emerged in the estimated current density distribution
when the Wiener filter algorithm with a time-invariant spatial constraint was used.
Figure 5.9 Source distribution reconstruction and source waveform estimation under fMRI false negatives. A) 3 source locations, B) quantified BOLD fMRI map, C) the simulated source waveforms, D) Reconstructed cortical current density distribution at 3 peak latencies (40, 100 and 160 ms) shown in 3 columns, by using 3 different algorithms: the adaptive Wiener filter (1st row), the conventional Wiener filter using time-invariant spatial constraints (2nd row) and the weighted minimum norm based on the EEG alone (3rd row); E) Estimated source waveforms at each of 3 source locations. The estimated waveforms are shown in solid lines and the simulated original waveforms are shown in dashed lines. The 3 rows also correspond to 3 algorithms as in Fig. 5.9.D).

Although the “blue” source was transient, the EEG-based source imaging could still detect this source activity due to the high temporal resolution of EEG. But again, the imaged source distribution had a low spatial specificity. By taking full advantage of the
temporal information from the EEG, the adaptive Wiener filter could clearly image and localize the “blue” source activity at 40 ms, as well as the other 2 source activities at 100 ms and 160 ms respectively.

The superiority of the adaptive Wiener filter was further confirmed by its ability to estimate the waveforms of both fMRI visible and invisible sources. As shown in Fig. 5.9.E), only the adaptive Wiener filter approach could recover the transient source activity in the estimated source time courses, without sacrificing the estimation accuracies to other sustained source activities.

Reconstructing source activities with complicated temporal dynamics

To assess the applicability of the proposed adaptive Wiener filter approach to reconstructing the source activities with complicated temporal dynamics, we simulated 3 source activities with a variety of temporal features. These 3 sources were located at the same locations as in the previous stimulations. Then we compared the source waveforms estimated by using the adaptive Wiener filter approach (in solid lines) with the originally simulated source waveforms (in dashed lines). The results are summarized in Fig. 5.10.

In Fig. 5.10.A), all 3 sources had gauss-function time courses that were overlapped over time. The shapes of the reconstructed source waveforms were close to gauss functions and the onset and peak latencies were also consistent with the simulated real source activities. However, slightly more fluctuations were observed in the reconstructed waveforms than in the real source waveforms.

In Fig. 5.10.B), a transient source activity (“red”) was included in addition to two overlapping gauss-function waveforms (“blue” and “green”). The “red” source was also an fMRI invisible source which occurred at the time point when both the other two sources were also active. In this situation, the time courses of both the fMRI visible and invisible sources were also well reconstructed.
Figure 5.10 Estimated source waveforms compared to the real source waveforms with different temporal features. 3 sources locations are the same as in Fig. 5.9. The originally simulated waveforms are shown in dashed lines and the estimated source waveforms are shown in solid lines.

In Fig. 5.10.C), the “green” source was a transient source temporally overlapping with a sustained “red” source. The “red” source also had a short-duration peak around the same time when the sustained “blue” source reached its peak amplitude. Clearly in this situation, both the fMRI false positives and false negatives existed with and without overlaps over time. Nevertheless, we could still estimate the source waveforms that were consistently accurate for both the sustained and transient source activities.

In addition, we simulated 3 rhythmic activities that had the same frequency but different phases (Fig. 5.10.D) or the same initial phase but different frequencies (Fig. 5.10.E). In both situations, the frequency and phase information were well reflected in the estimated source waveforms without significant distortion, whereas we also observed distinctions in absolute amplitudes between the estimated and real source waveforms.

In Fig. 5.10.F), we simulated the source activities by using the “realistic” dipole source waveforms estimated from a real ERP data set acquired in a visual experiment. In this case, the source activities had fairly complicated temporal features. The estimated...
source waveforms were highly correlated with the simulated real source waveforms, without losing even detailed temporal dynamics.

The results in Fig. 5.10 all demonstrate good performances of the proposed adaptive Wiener filter algorithm to accurately estimate the source activities with a high temporal resolution, based on simulated fMRI and EEG data.

**5.3.2 Experiments for fMRI-Constrained Adaptive Wiener Filter**

**Experimental Setting**

To assess the applicability of the proposed adaptive Wiener filter algorithm, we also employed real fMRI and EEG data collected during a preliminary experiment with a unilateral checkerboard visual stimulus.

The experiment was conducted with a healthy subject (male, age 22) under the approval of the institutional review board (IRB) at the University of Minnesota. Informed consent was obtained from the subject before the experiment. The experiment included 2 separate sessions with the identical visual stimuli for the EEG and fMRI data collection respectively. The visual stimulation was a rectangular checkerboard within the lower left quadrant of the visual field; the checkerboard pattern was reversed at 2 Hz. In the EEG experiment, 6-second breaks without stimulation were randomly inserted into the otherwise 4-minute continuous visual presentation, such that the subject had a break about every 20 seconds on average to avoid the neural adaptation. In the fMRI experiment, the visual stimuli were presented in six 30-second blocks separated by seven 30-second resting blocks without stimulation. The stimuli delivered through a DLP projector were back-mirrored to the subject inside the MRI scanner. For both EEG and fMRI experiments, the subject was instructed to always gaze at a central fixation point.

The EEG signals collected from a 64-channel system (BrainAmp MR 64 plus, BrainProducts, Germany) with a 1000-Hz sampling rate were sequentially subject to the ocular artifact rejection, band-pass filtering (0.3 – 40 Hz), segmentation with respect to the stimulus onsets, pre-stimulus baseline correction, linear trend removal. After these preprocessing steps, the data in 380 segmented epochs was averaged to yield the visual evoked potentials (VEP). The anatomical MRI and fMRI data were collected in a 3-T
MRI system (Siemens Trio, Siemens, Germany). The whole-head T1-weighted MR images (matrix size 256×256, 1mm slice thickness) were acquired using the Turboflash sequence (TR/TE = 20 ms/5 ms). The T2*-weighted fMRI data was acquired from 16 axial slices (matrix size 64×64, 5mm thickness) covering the visual cortex using the echo planar imaging (EPI) sequence (TR/TE = 1000 ms/35 ms). The MRI and fMRI data were analyzed using BrainVoyager QX (Brain Innovation, Netherlands). The EPI volumes underwent several preprocessing steps including three-dimensional motion correction, slice scan time correction and linear trend removal. Then, the fMRI data was aligned with the anatomical MR images. The fMRI activation map was obtained by statistical analysis using a general linear model.

Three cortical current density imaging algorithms were applied to image the cortical responses on the hemisphere (right) contralateral to the stimulation (left). The 90% fMRI-weighted Wiener estimation (Liu et al. 1998; Dale et al. 2000) and the proposed algorithm used both EEG and fMRI data, whereas the minimum-norm algorithm used the VEP data lone.

Results

In response to the unilateral visual stimulation (Fig. 5.11.A), the activated cortical areas at the contralateral hemisphere were revealed in the fMRI activation map (Fig. 5.11.B). The fMRI activation map indicated a dorsal visual pathway covering V1, V2, dorsomedial areas (such as V3 and V7), intraparietal sulcus (IPS) as well as medial temporal (MT) area (also known as V5). The top row of Fig. 5.11.C shows the time course of global field power of VEP, which indicates three VEP peak latencies (76, 112 and 212 ms). The 2nd through 4th rows of Fig. 5.11.C show the reconstructed contralateral CCD distribution using three imaging algorithms, respectively. From the CCD images reconstructed by only using the VEP data, the dorsal pathway was seen gradually extending from lower-tier visual areas to high-tier visual areas. By using our proposed adaptive Wiener filter to integrate the fMRI and EEG data, a consistent sequence of activities was observed with a much enhanced spatial resolution, showing the pathway starting from V1/V2 and then V3/V3a and finally V5/V7 and IPS. The observed cortical
visual pathway was generally in agreement with the well-known hierarchical organization of the visual system (Felleman and Van Essen 1991). In contrast, the imaging results obtained by using the conventional 90% fMRI-weighted approach also had improved spatial resolution compared to the EEG-alone source imaging. However, it imposed a false positive source region in and around V1/V2 at the latency of 212 ms, whereas a more likely high-tier EEG source around V5 observable from the EEG data was missed.

**Figure 5.11** A) The pattern-reversal checkerboard visual stimulation, B) fMRI activation map with a corrected threshold p<0.01, and C) the global field power of VEP and the dynamic cortical source distribution at three VEP latencies (76, 112, 212 ms after the visual onset) imaged from EEG.
alone (1st row), or fMRI-EEG integration using our proposed adaptive
wiener filter (2nd row) and the conventional 90% fMRI weighted
algorithm (3rd row). Both the source images and the fMRI activation map
are visualized on an inflated representation of cortical surface.

5.3.3 Simulation for Twomey Algorithm

We have developed a Twomey regularization algorithm to correct the fMRI-EEG
mismatches (Liu et al. 2006b). Fig. 5.12 shows the effect of fMRI-EEG mismatches (see
B.4) on instantaneous cortical source imaging using three different inverse algorithms:
weighted minimum norm (WMN), 90% fMRI-constrained Wiener estimation and the
Twomey algorithm. A 128-channel scalp EEG map was simulated at SNR=5. The upper
row shows an example with five unitary dipole sources (Fig. 5.12.A). The fMRI
activations covered three sources, leaving the other two sources (marked by blue dotted
circles) to behave as fMRI missing sources (Fig. 5.12.C). The WMN had extended
cortical source reconstruction (Fig. 5.12.B) and thereby suffered from ambiguity of
localizing the “true” sources. With valid fMRI priors, the 90% fMRI-constrained Wiener
estimation substantially improved the spatial specificity of imaging fMRI visible sources;
however the fMRI invisible sources were considerably underestimated (Fig. 5.12.D). The
Twomey solution had a similar imaging performance for the fMRI visible sources
compared to the Wiener estimation, and it also had a much improved estimation for the
fMRI invisible source activities (Fig. 5.12.E). This example demonstrated the efficacy of
the Twomey algorithm in revealing fMRI invisible sources.

The lower row shows an example with three unitary dipole sources (Fig. 5.12.F).
The fMRI activations included two additional regions without EEG source activity (Fig.
5.12.H). Similar as the first example, the WMN solution had low spatial resolution (Fig.
5.12.G). In both the Wiener and Twomey solutions, the fMRI visible sources could be
clearly observed, whereas slightly spurious sources also appeared in the extra fMRI
regions in the Wiener estimates (Fig. 5.12.I). By use of the Twomey algorithm, the
spurious sources induced by the extra fMRI prior had even smaller intensity in the
reconstructed cortical source distribution (Fig. 5.12.J). This example indicated that the
fMRI extra sources did not significantly distort the fMRI-EEG integrated cortical imaging using the Twomey algorithm.

![Brain Images](image)

**Figure 5.12** Typical examples of effects of fMRI invisible and fMRI extra sources on three cortical source imaging methods: the EEG-alone weighted minimum norm, the 90% fMRI-constrained Wiener filter and the Twomey algorithm. The upper row shows the results of simulation on fMRI invisible sources; the lower row for fMRI extra sources.

Fig. 5.13 shows the PSF values for all three algorithms (WMN, Wiener and Twomey) in 100 random trials with 5 unitary point sources and SNR=5. The simulated fMRI activations covered 4 sources, leaving one source to behave as an fMRI missing source. Fig. 5.13.A) shows the mean PSF averaged over the four fMRI visible sources in each trial. Clearly, both Wiener and Twomey resulted in much smaller PSF of fMRI visible sources than WMN. However, for the fMRI invisible source (as shown in Fig. 5.13.B), considerably larger PSF was found for Wiener than for WMN. The values of PSF of the fMRI invisible source for Twomey was much smaller than those for Wiener, and in many trials they were even smaller (or at least comparable) than WMN solution.
Fig. 5.13.C) illustrates the PSF of the fMRI invisible source for both Wiener and Twomey algorithms as a function of the respective PSF value in WMN solution in the same trial.

Clearly, as for the fMRI invisible sources that were more “visible” (with smaller PSF values) in WMN solution, it was more obvious to observe the improvement by use of Twomey algorithm relative to Wiener. If the fMRI invisible sources could hardly be imaged even in non-fMRI-biased WMN solution, Twomey algorithm was also less helpful in revealing such fMRI-missed sources.

**Figure 5.13** Point Spread Function (PSF) for either fMRI visible sources or fMRI invisible sources in 100 random trials.
Fig. 5.14 summarizes the effects of displacement of fMRI active region relative to the associated electric current dipole source on the localization error of fMRI-EEG integrated imaging approaches (Wiener and Twomey), in comparison with EEG-alone WMN solution. The bars in Fig. 5.14 represent the mean LE averaged over 100 random, with or without different levels of fMRI displacement (5, 10, 15 or 20 mm). The WMN solution had a mean localization error around 1-cm, and the error tended to decrease with higher SNR. With accurate fMRI priors, the fMRI-constrained solutions in both Wiener and Twomey had much smaller mean LE (about 4 to 5-mm) than WMN, and Wiener gave slightly smaller LE than Twomey. If the fMRI activation was displaced away from the electrical source, the LEs for both Wiener and Twomey increased when the displacement was enlarged. With a smaller fMRI displacement (e.g. 5-mm), the LEs for Wiener and Twomey were very close, whereas with a relatively larger fMRI displacement (e.g. >1.0cm), Twomey gave smaller LE than Wiener. Moreover, even with the fMRI displacement as large as 2-cm, Wiener and Twomey still had smaller LE than WMN solution, except at SNR=10 Wiener had larger LE than WMN. This finding suggests that the incorporation of fMRI spatial prior even with less than 2-cm displacement was still helpful to localize the electrical sources, compared to the EEG-alone WMN solution.

![Figure 5.14 The average localization error as a function of METHODS (Twomey, Wiener and WMN) and fMRI displacement (distance of fMRI](image)
activation away from the electrical source)

5.3.4 Simulation for fMRI-Seeded Dipole Fitting

We have conducted preliminary computer simulations to evaluate the fMRI-seeded dipole fitting technique in resolving the time courses of multiple current sources. Fig. 5.15 shows some representative results. Three cortical patches with spatial extents of 20 mm were simulated (Fig. 5.15.A). The current density distribution within each patch was simulated to have a Gaussian profile, in which the center of patch had the largest current density amplitude (Fig. 5.15.B). All current directions were perpendicular to the local cortical surface. The integral of the current density in each patch had an assigned time course as shown with different colors in the top row of Fig. 5.15.C)-F). Three source waveforms were simulated to have no temporal overlap (Fig. 5.15.C), or be temporally overlapped (Fig. 5.15.D), or have the identical frequency with different phases (Fig. 5.15.E) or different frequencies with the same initial phase (Fig. 5.15.F). The distributed patch activities were modeled by 3 equivalent dipoles. Assuming the fMRI-seeded initial dipole locations had a distance of 20 mm to the centers of patches (namely the fMRI localization error), the dipole locations were first searched by a non-linear procedure and then the equivalent dipole moments were linearly estimated. As shown in Fig. 5.15.C)-F), the estimated time courses were highly correlated to the “true” waveforms of distributed patch activities, with only a mild distortion observed in Fig. 5.15.F). Note that the estimated dipoles were not necessarily inside the corresponding patch, whereas their spatial relationship depended upon the activated patch location and extent as illustrated in Fig. 5.15.B). The last row in Fig. 5.15.C)-F) shows the estimation accuracy, quantified as the correlation coefficient (CC) between the estimated and “true” source waveforms, as a function of source extent (0~30 mm) and fMRI localization error (0~20 mm). These simulation results suggest that by using the fMRI-seeded dipole fitting one can estimate the waveforms of multiple extended source regions, and the estimated waveforms may be further taken as inputs to the connectivity estimation for investigating the causal interactions among regions of interest.
Figure 5.15 A) Three distributed sources each with a spatial extent of 20 mm, B) the spatial relationship between the distributed sources with certain current density profile and the localized regional dipoles, C)-F) results of time course estimation when the integral of activities inside each patch were assigned with four different types of temporal waveforms.

5.4 Discussion

fMRI-Constrained Adaptive Wiener Filter

The proposed adaptive Wiener filter method on the fMRI-EEG integration significantly differs from the existing methods in two important aspects: 1) interpreting the BOLD effect size as proportional to the time integral of EEG source power (see details in Chapter 4), 2) estimating the time-variant source covariance matrices from both fMRI and EEG data. Our pilot simulation results have demonstrated that the proposed method is capable of handling the mismatches between locations of fMRI activations and EEG source activities at any time instant (including fMRI false positives and false negatives), as well as probing the millisecond temporal dynamics of source activities. Our preliminary experimental application also suggests a robust performance of improved
spatial resolution (relative to the EEG-alone inverse estimates) and temporal resolution (relative to the fMRI).

In previously reported fMRI-EEG multimodal neuroimaging studies, the fMRI data analysis yields an fMRI activation map, which is then used to constrain the EEG inverse solutions in a subsequent step; however, the pixel values in the fMRI activation map fail to provide any quantitative interpretation that is also physically relevant to the following step of the fMRI-constrained EEG source imaging or localization. In contrast, the fMRI data in the present study are quantified as well-defined physical constraints to the EEG inverse solutions. The time integral of source power (over the period of ERP) is constrained to be proportional to the quantified BOLD effect size at each source location. In this sense, the proposed fMRI quantification method also represents a more specific inference regarding underlying neural activities. Existing fMRI analysis methods are often used to quantify the level by which the BOLD signal is statistically correlated to the designed protocol of tasks of interests. Based upon such a statistical analysis, inferences are often made voxel by voxel regarding how likely a region in the brain is involved in the neural computation underlying the tasks. Clearly, such an inference is conceptual rather than physical. In the present study, we quantitatively relate the fMRI signal over a time scale of seconds to the averaged electrical signal over a much shorter time scale of tens of milliseconds. In short, we suggest using the proposed fMRI quantification method in all fMRI-EEG multimodal neuroimaging studies, or even in fMRI studies.

Note that a linear neurovascular coupling is assumed in this approach. Although such an assumption is commonly made in the fMRI analysis and interpretation, the nonlinear aspect of fMRI response has also been observed, such as under repetitive stimulation (Janz et al, 2001). The nonlinearity may arise from the neural and/or hemodynamic adaptation or habituation (Janz et al, 2001), or the vascular refractory effect (Cannestra et al, 1998), or the interplay of multiple factors. However, our recent investigation of neurovascular coupling based on the VEP and CBF data under graded brain suppression shows that the linear model is still a good approximation of neurovascular coupling relationship, although it contains a subtle nonlinear component (Zhang et al, 2007). Furthermore, the fMRI response nonlinearity is probably less of
concern in most experiments combining ERP and fMRI, as the adjacent stimuli (in ERP paradigms) are often required to be sufficiently separated in time to avoid the interference among stimuli.

In most existing fMRI-EEG (or -MEG) integrated source imaging algorithms, the fMRI-derived spatial constraints have to be assumed to be time-invariant within a short time period during which the source imaging is carried out instant by instant. This is essentially because the source temporal information cannot be resolved by fMRI. However, when the “static” fMRI spatial information is taken as a time-invariant weighting matrix (in the fMRI-weighted minimum norm formulation) or source covariance matrix (in the Wiener filter formulation) for every time point, it is almost unavoidable that the imaging results would be affected by the fMRI-EEG mismatches, which may happen at any time and also vary over time. Previous efforts have been mainly focused upon choosing an empirical weighting factor for locations inside the fMRI activation relative to those outside (Liu et al, 1998; Babiloni et al, 2003; Ahlfors and Simpson, 2004). In such a way, fMRI constraints are weakened such that distortions due to fMRI false positives and false negatives may be alleviated but perhaps never removed. However, this type of methods, no matter what the fMRI weighting factor is chosen to be, has a tradeoff between the robustness against invalid fMRI prior constraints and the spatial resolution of the instantaneous inverse solution. If the fMRI weighting factor is chosen to be 1, the inverse solution is effectively the minimum norm solution; if the fMRI weighting factor is chosen to be infinitely large, the inverse solution can be highly biased by the fMRI-EEG mismatches. On the other hand, the proposed adaptive Wiener filter approach is aimed at correcting fMRI-EEG mismatches in a much more fundamental way, in consideration of the fact that the mismatches are essentially caused by the different time scales at which the fMRI and EEG signals are generated and collected respectively. The proposed adaptive Wiener filter method holds the promise to remove the mismatches instead of attempting to alleviate the resulting distortions with a cost of lower spatial resolution.

In addition, the assumption of a time-invariant source covariance matrix is also not technically valid. A time-invariant source covariance matrix implies that the source
signal at any specific location is a stationary stochastic process over time and its sampled values at all time points are drawn independently from an identical distribution. Such an assumption is not valid in practice because neural networks in the brain are always carried out by source signals with rapidly and coordinated temporal evolution that are certainly not exchangeable over time. In contrast, the estimation of time-variant source covariance matrices in the proposed adaptive Wiener filter method relies on the assumption that at any time instant, the scalp potential maps recorded in different epochs are independent observations of a stochastic process with the index set over channels, which is a reasonable assumption when the electrical responses are repeatedly induced over epochs by the same stimulus or task.

More factors other than those addressed in this chapter may also give rise to the spatial mismatches between the fMRI activations and instantaneous source activities. The fMRI spatial specificity and the cross-modal co-registration error may lead to a difference between locations of neural activities and those of the detected fMRI activations. However, these factors are probably more technical rather than fundamental, in a sense that they may be resolved (or bypassed) before applying the integrated imaging algorithm. For example, one may use a spin-echo sequence and increase the field strength to eliminate the contribution from large draining veins, such that the collected fMRI signals originate exclusively from the microvasculature (Yacoub et al, 2003), resulting in an improved fMRI specificity. One may also use advanced geometric manipulations with optimization procedures that combine the landmark and surface point fitting procedures to ensure a good enough co-registration of difference coordinate systems respectively for the fMRI and EEG data.

In the present study, we assume a linear transform model to describe the relationship in time between BOLD fMRI signals and the power of electrical responses. However, recent studies also investigated the correlation between BOLD signals and the EEG (and LFP) power spectrum (Mukamel et al, 2005; Niessing et al, 2005; Kilner et al, 2005). It was suggested that the BOLD responses were positively correlated with the high-frequency components of EEG (or LFP) signals and negatively correlated with the low-frequency components. Such an obvious difference between the models described in
time and in frequency does not necessarily mean they are contradictory to each other. Instead, we believe that their difference only reflects the different aspects of the hemodynamic-to-electrophysiological coupling addressed in these two types of models. The model in time as used here describes how the BOLD signals relate to the event-related electrical response, whereas a model in frequency describes how the BOLD signals reflect the change of continuous electrical signals. It remains unclear how the fMRI-EEG/MEG multimodal neuroimaging should benefit from these recently reported findings (Mukamel et al., 2005; Niessing et al., 2005) on the correlation of BOLD signals with frequency components of EEG power spectrum, although it is indeed desirable.

Although the singular value decomposition of the recorded EEG over epochs also allows computing the noise covariance matrix in a time-variant manner, it may not be necessary under most circumstances since the noise is not time-locked to the “events” and can be assumed to be time-invariant. Therefore, one can estimate the noise covariance matrix simply from the pre-onset part of EEG epochs. For every pre-onset time point, a covariance matrix can be computed from the multi-channel EEG data over epochs. Then, the noise covariance matrix can be obtained by averaging the computed covariance matrices at all pre-onset time points.

It may also be worthwhile to mention that the information from fMRI and EEG have also been proposed to be fused in a symmetric fashion (Trujillo-Barreto et al., 2001; Daunizeau et al., 2007), in which a common hierarchical Bayesian model are formulated for both fMRI and EEG. These symmetric Bayesian approaches may represent another feasible way that potentially leads to a reliable fMRI-EEG integrated neuroimaging.

**Twomey Algorithm**

The proposed Twomey algorithm is a two-step approach to combine the fMRI and EEG data. The solution with hard fMRI constraint resulting from the first step is re-entered to the second step as the initial estimate in a modified Tikhonov formulation utilizing Twomey regularization. The reason for having the second step is to enhance the robustness of the CCD imaging against inconsistent fMRI information is because that in the second step fMRI constraint is not directly involved any more. Noticeably, the eventual inverse operator is still a linear one. This property essentially differentiates this
algorithm from other iterative algorithms such as FOCUSS (Gorodnitsky and Rao, 1997; Rao and Kreutz-Delgado, 1999), in which the previous solution is re-entered to the current estimation procedure as a weighting matrix and eventually ends up with a nonlinear inverse operator equivalent to L-p norm (Rao and Kreutz-Delgado, 1999).

From the perspective of Bayesian estimation, the two-step Twomey algorithm provides a different way of rendering the prior Gaussian probability density of current density at each source location, by controlling its source variance and mean separately in two steps. In the first step, which is identical to the conventional Wiener filter method, the probability densities are assumed to be zero-mean with the variance depending on whether the source location is inside the fMRI activation or not. Specifically, the source variance is larger for fMRI active locations than that for fMRI non-active locations. As a consequence, the prior probability distribution for fMRI invisible locations are much shaper and more concentrated to its mean value 0 while the prior probability distribution for fMRI visible locations have much wider spread. Therefore, the large amplitude of non-zero current density estimate is penalized in the fMRI invisible locations but rewarded in the fMRI visible locations. That is also the reason why the fMRI invisible sources tend to be underestimated in the Wiener filter method. However, in the second step, the prior probability distribution for different source locations is assumed to have the identical variance but different means as given by the initial source estimate in the first step. The use of Twomey regularization will yield more robust solution in that the location-wise hyperparameters of the prior probability are the means in the second step, which are carried by the initial estimate from the first step and convey the integrated information from both fMRI and EEG, instead of fMRI alone. Since the locations with the prior means far away from zero are more likely having a larger absolute amplitude of current density estimate, the source locations that are favored by the fMRI spatial constraint and have larger estimate in the first step are still indirectly favored in the second step, resulting in the low PSF for fMRI visible sources for the Twomey algorithm. However, the relaxation from the constraints on the variance makes the prior probability distribution equally spread out, providing better chance for the non-fMRI locations to
have a large non-zero source estimate so that the fMRI-invisible sources may be recovered.

The proposed Twomey algorithm can reduce the distortions due to the fMRI-EEG mismatches. Our simulation results indicate that the most problematic fMRI-EEG mismatch for the Twomey algorithm is fMRI invisible EEG sources. Although applying Twomey regularization can alleviate the distortion due to fMRI invisible sources, the fMRI invisible sources are still underestimated as the PSF of fMRI invisible sources is still higher than that of fMRI visible sources and even higher than the PSF in the weighted minimum norm solution at some cases. In this sense, the fMRI-constrained adaptive Wiener filter algorithm is superior to the Twomey algorithm.
Chapter 6 fMRI-EEG Integrated Neuroimaging: Application

6.1. Introduction

Humans readily perceive a unified visual world in spite of its fragmented and discontinuous internal representation at the primary visual cortex (V1). In particular, retinal inputs from the left and right hemifields are separately projected to the contralateral V1 area, and consequently the V1 retinotopy is split at the midline between the two cerebral hemispheres (Lavidor and Walsh 2004). This fact points to an important question as to how the bilaterally separated visual representations are integrated into an intact percept, namely the bilateral visual integration (BVI).

The traditional visual hierarchy theory suggests a serial processing through a bottom-up cascade of discrete visual areas (Felleman and van Essen 1991). In this view, the BVI process has to be initiated at the higher-tier functional areas in which the receptive fields (RF) of neurons are large enough to cover both the left and right visual fields (LVF and RVF respectively). It also implies that these areas are activated later than lower-tier areas with unilateral RF, such as V1/V2. However, among the areas with bilateral RF, it remains unclear which is the first cortical substrate to merge the bilateral spatial information while the other areas successively operate upon the whole visual scene to perform more sophisticated processing (e.g. visual object recognition). To tackle this question, it is clearly necessary to look into not only the receptive field properties of all cortical visual areas but also the temporal sequence of their activations in response to bilateral visual inputs.

Nevertheless, investigating the cortical BVI process exclusively based on the traditional visual hierarchy theory is challenged by the new theories proposing a parallel top-down and bottom-up information flow (Lamme and Roelfsema 2000). Evidence has increasingly suggested the existence of functional as well as anatomical feedback connections from higher-tier areas targeting lower-tier areas (Salin and Bullier 1995; Hupé et al. 1998; Lamme et al. 1998). As a result, neural activities at lower visual areas with traditionally contralateral RF may indirectly respond to inputs from the ipsilateral visual field through a top-down stream stemming from higher areas with bilateral RF.
(Tootell et al. 1998; Ban et al. 2006). The neural response at a higher visual area may also precede the activations or re-activations of some cortical areas at lower levels of the hierarchy (Buchner et al. 1997; Hupé et al. 2001; Barnikol et al. 2006). These complications beyond the traditional view amount to further uncertainties with regard to the timing and locations of the cortical BVI process (Ban et al. 2006; Vanni et al. 2004).

Since distributed cortical visual areas in different hierarchical levels may likely contribute to the BVI, it is necessary to simultaneously monitor the dynamics of neural responses from the entire visual system, which covers numerous regions within the occipital, parietal and temporal lobes. As such, high spatial and temporal resolution is desirable for mapping the spatiotemporal cortical activity related to BVI. However, all of the existing functional neuroimaging modalities have limitations in either spatial or temporal aspect. For example, functional magnetic resonance imaging (fMRI) (Kwong et al. 1992; Ogawa et al. 1992; Bandettini et al. 1992), by measuring hemodynamic and/or metabolic responses, has the advantage of revealing spatial details of neural activations but with limitations concerning its low temporal resolution in the order of seconds. Conversely, electroencephalography (EEG) (or magnetoencephalography, MEG) can detect rapid electrophysiological responses whereas the EEG source imaging often suffers from poor spatial resolution due to its inherent mathematical difficulty (Baillet et al. 2001; He and Lian 2002). In light of the complementary advantages of the high-spatial-resolution fMRI and the high-temporal-resolution EEG, many methodological developments have been focused upon the cross-modal integration (Dale and Halgren 2001; Liu et al. 2006a, 2006b; Liu and He 2008; Gotman et al. 2006), which has the potential to significantly advance our knowledge in understanding the sensory and cognitive neurosciences (Dale et al. 2000; Eichele et al. 2005).

Such a multimodal approach was applied in the present study to address when and where the bilateral visual inputs are integrated into a cohesive percept. For the analyses with respect to the locations and latencies of visual evoked cortical responses, we acquired the blood oxygenation level dependent (BOLD) fMRI and scalp EEG (visual evoked potential, VEP) signals induced by the same visual stimulation. A pair of identical pattern-reversal (i.e. counter-phase flicking) checkerboards were presented
separately, simultaneously or in an interleaved manner at vertically symmetric positions within the lower visual field. Such stimuli have been shown to activate the elementary components of the visual system without a substantial involvement of high-level psychological processes (Tobimatsu and Celesia 2006).

**Figure 6.1** Scenarios for investigating the cortical activity underlying the bilateral visual integration. A) Two visual pathways respectively responsive to the left and right visual fields (i.e. LVF and RVF) converge at their common regions where the bilateral visual information is
integrated. B) Brain signals (including both fMRI and EEG) associated with the cortical BVI process can be isolated from those associated with other independent and unilaterally responsive processes by subtracting the responses to the bilateral stimulus from the sum of responses to both unilateral stimuli.

The experimental design and data analysis were carried out along two scenarios illustrated in Fig. 6.1. The first scenario (Fig. 6.1.A) suggests that the BVI is enabled by the anatomical convergence of two cortical visual pathways responsible for the LVF and RVF respectively. Accordingly, these convergent areas should become co-activated by both LVF and RVF unilateral stimuli, namely “equal responses to unilateral stimuli”. Under the second scenario (Fig. 6.1.B), the non-integrative early visual areas should exhibit independent responses to the LVF and RVF unilateral stimuli respectively. At such areas, the sum of unilaterally induced responses should strictly equal the response to bilateral stimuli presented simultaneously. However, at the cortical regions involved in the BVI, a deviation from the linear summation is expected due to their lack of dependence upon visual inputs obtained exclusively from the contralateral hemifield. Previous studies have suggested that such spatial nonlinearity is often demonstrated by a smaller response to the bilateral stimulus than the sum of individual response to each unilateral component, namely “the under-summed responses to the bilateral stimulus” (Vanni et al. 2004; Miniussi et al. 1998; Murray et al. 2001; Steger et al. 2001; Supek et al. 1999).

With these scenarios, we analyzed the BOLD fMRI and VEP signals evoked by unilateral (i.e., two checkerboards presented separately from the LVF and RVF respectively) and bilateral stimuli (i.e., simultaneous or interleaved presentation of both checkerboards). The visual areas involved in the BVI were obtained by searching for the cortical regions that showed approximately equal BOLD responses to both unilateral stimuli and demonstrated nonlinearly under-summed BOLD responses to the bilateral stimulus. The timing of the BVI was revealed by seeking the latencies with high cross correlation between the VEP topographies elicited by the LVF stimulus and those by the
RVF stimulus, as well as with considerable components after subtracting the VEP response to the bilateral stimulus from the sum of responses to both unilateral components. The spatiotemporal cortical processes responsible for the unilateral visual processing and the BVI were further reconstructed from the corresponding VEP data (or components) with or without using the information from the fMRI data.

6.2. Materials and Methods

6.2.1 Subjects

Ten healthy right-handed subjects (age 24±6 years, 8 male and 2 female) participated in the EEG study. Six of the subjects also participated in the fMRI study. All of the subjects had normal or corrected-to-normal vision and gave written, informed consent in accordance with a protocol approved by the institutional review board at the University of Minnesota.

6.2.2 Stimuli

The visual stimuli consisted of one or two rectangular black-and-white pattern-reversal checkerboards (size: 12º horizontal and 10º vertical; reversing frequency: 2 Hz; spatial frequency: 1.0 cycle/degree; mean luminance: 20 cd/m²) displayed on a dark gray background (luminance: 5 cd/m²) with a yellow central fixation point. Each rectangle was presented 2º below the horizontal meridian and 4º left or right of the fixation point (measured from the near edge). In the two unilateral conditions, a single pattern-reversal checkerboard was presented in either the lower-left or lower-right quadrant of the visual field. In the two bilateral conditions, a pair of checkerboards presented within both lower quadrants were reversed either simultaneously or in an interleaved manner.

6.2.3 Data Acquisition

The EEG study was first conducted individually in an electrically shielded room for all 10 subjects. Following the EEG study, the fMRI data was acquired for 6/10 subjects in a 3-T/90 cm bore magnet (Siemens Trio, Siemens, Germany) equipped with an eight-channel phase array head volume coil. All of the subjects were trained and
instructed to maintain sustained visual gaze upon the central fixation point during the experiment.

The EEG experiment included six repeated 3.5-min runs. During each run, the unilateral stimuli and simultaneous bilateral stimuli were presented in a mixed sequence interspersed with six 4-sec periods without stimuli. The scalp potentials from 64 electrodes (referenced to FCz and placed according to the extended international 10/20 system) were recorded at 1000 Hz and filtered (0.3–70 Hz) through a pair of amplifiers (BrainAmp MR 64 Plus, BrainProducts, Germany). Eye blinks and movements were monitored with horizontal and vertical electrooculographic (EOG) electrodes. The electrode locations and three anatomical landmarks (left/right preauricular points and nasion) were digitized through a three-dimensional (3-D) RF localizer (Polhemus Fastrak, Colchester, VT).

In the following fMRI experiment, the whole-head anatomy was first acquired with 256 sagittal T1-weighted MR images (matrix size: 256×256; in-plane resolution: 1×1 mm²; slice thickness: 1 mm; no gap between slices) using a TurboFLASH sequence (TR/TE = 20/5 ms). The functional study included six repeated 4.5-min runs. Each run consisted of four 30-sec task blocks (with left/right unilateral stimuli and simultaneous/interleaved bilateral stimuli respectively) separated by five 30-sec control periods without stimuli. The BOLD fMRI data was acquired with 16 axial T₂*-weighted images (matrix size: 64×64; in-plane resolution: 4×4 mm²; slice thickness: 5 mm; no gap between slices) covering both the occipital and parietal lobes using a gradient-echo echo-planar imaging (EPI) sequence (TR/TE = 1000/35 ms).

6.3. Data Analysis

6.3.1 Visual Evoked Potentials

The EEG preprocessing was conducted in BrainVision Analyzer (BrainProducts, Gilching, Germany). The recorded EEG signals were sequentially subject to ocular artifact rejection (by visual inspection), band-pass filtering (0.3 – 40 Hz), segmentation from -100 to 500 ms around the stimulus onsets, pre-stimulus baseline correction, linear trend removal and response averaging to obtain the stimulus-specific VEP data. The
global field power (GFP) at every time instant was computed as the standard deviation of instantaneous scalp potentials at all 64 channels. The VEP topographies evoked by the left and right stimuli (denoted as $\Phi_L$ and $\Phi_R$ respectively) were quantitatively compared instant by instant through their correlation coefficient (denoted as $r$) and relative difference defined as $\frac{\|\Phi_L - \Phi_R\|}{\|\Phi_L\|}$. In a similar fashion, we also compared the topographies responding to the bilateral stimulus with the sum of those responding to two unilateral stimuli. The difference between them obtained by a simple subtraction represents the nonlinearly under-summed VEP component to the bilateral stimulus.

### 6.3.2 BOLD-fMRI Analysis

The fMRI data was processed using BrainVoyager (Brain Innovation, Maastricht, Netherlands). The EPI volumes underwent preprocessing steps including head motion correction, slice scan time correction, linear trend removal and high-pass filtering (3 cycles per scan) and were then averaged over six functional runs. One subject with considerable head motion artifact was rejected from subsequent fMRI data analysis. The functional volumes were aligned to the anatomical images and re-sampled to a voxel size of $3\times3\times3$ mm$^3$ with a trilinear interpolation.

For each individual subject, the fMRI data was analyzed using a general linear model (GLM) (Friston et al. 1995; Worsley and Friston 1995; Liu and He 2008). Note that in our GLM analysis, the design matrix was defined by the convolution of stimulus-specific trains of delta functions (representing repeated transient visual stimuli) with the canonical hemodynamic impulse response function. Using the regressors defined as above, we computed the BOLD effect size (i.e. regression coefficient) associated with each type of task through multivariate linear regression. The quantified BOLD effect sizes also characterized the time integral of the power of the corresponding visual-evoked electrophysiological responses, according to our previous theoretical modeling study (Liu and He 2008). Based upon the outputs from the GLM analysis, we further computed and visualized the statistical parametric maps contrasting a task vs. the control (i.e. no stimulus) or multiple tasks (using $t$ statistic, thresholded at $p<0.01$ after Bonferroni correction).
The BOLD responses to the left and right unilateral stimuli were compared through an index of relative contribution (RC), which indicates how strong one unilateral condition relative to the other contributed to the fMRI time series at every activated voxel. Specifically, the RC value was defined as $\frac{\beta_L - \beta_R}{\beta_L + \beta_R}$ where $\beta_L$ and $\beta_R$ were the BOLD effect sizes associated with the left and right unilateral stimuli respectively. The nonlinear under-summation of BOLD response to the bilateral stimulus was evaluated in two different ways: 1) contrasting the sum of responses to both LVF and RVF unilateral stimuli vs. the response to the combined bilateral stimuli presented simultaneously, denoted as “LVF+RVF–Both”, and 2) contrasting the responses to the interleaved vs. simultaneous bilateral stimuli, denoted as “L2R–Both”. Group averaging was done through a cortical matching approach by aligning the curvature information of every cortical surface (Fischl et al. 1999).

### 6.3.3 Cortical Source Imaging

A boundary element model consisting of triangulated surfaces of the scalp, skull and brain was built for each individual subject after segmenting the subject’s anatomical MRI. Distributed cortical current density (CCD) was modeled by around 7,000 dipoles placed on the cortical surface (i.e. the boundary between the white matter and the gray matter) with dipole orientations along the outer-normal direction of local cortical patches. The boundary element method (Hämäläinen and Sarvas 1989) was employed to compute the scalp potential distribution arising from each individual unitary dipole, which collectively sets up a spatial linear system describing the source-to-signal transformation. The minimum norm algorithm (Hämäläinen and Ilmoniemi 1984) was used to reconstruct the spatiotemporal CCD distribution from the measured scalp potential maps.

The spatiotemporal CCD distribution was also reconstructed by using an advanced multimodal neuroimaging algorithm (Liu and He 2008) integrating the complementary information from both fMRI and VEP data. With this technique, the quantified BOLD effect size in each voxel was used to constrain the time integral of the power of regional synaptic current source estimates during the post-stimulus VEP duration (i.e. 0–400 ms).
The imaged CCD distribution obtained from either the EEG alone or the fMRI-EEG integration was visualized on both the folded cortical surfaces.

6.4. Results

6.4.1 Equal Responses to Unilateral Stimuli

Traditional visual hierarchy theory states that unilateral retinal inputs are first processed independently at early functional areas along their responsive visual pathways before converging onto some common cortical regions where the spatial information from both hemifields is pooled together (see Fig 6.1.A for a simplified illustration). Assuming an equal speed of information flow via the LVF and RVF pathways, both processes are expected to reach the bilaterally integrative areas at the same latencies following the visual stimulus. If so, we can accordingly identify the timing of the BVI as the latencies when the instantaneous VEP topographies evoked unilaterally exhibit identical (or closely similar) spatial patterns, given the fact that the scalp potential distribution is essentially a low-pass filtered indicator of the locations of the underlying cortical generators.
Figure 6.2 (A) Waveforms of the global field power (GFP) in response to the LVF (blue) and RVF (red) unilateral stimuli for a single subject (S1). Light blue bar at time 0 indicates the stimulus onset. Three GFP peaks were found at 76, 112 and 212 ms for both conditions. (B) Electrical potential distributions at the peak latencies identified in (A) are displayed on the subject’s realistically shaped scalp surface.

The temporal evolvement of multi-channel VEPs was observed through the time course of the global field power (GFP). Fig. 6.2.A shows the GFP waveforms for one subject (S1) responding to the LVF and RVF stimuli respectively. Three GFP peaks were observed in both time courses at almost identical latencies (76, 112, 212 ms), suggesting
an equal processing speed for both pathways. The VEP topographies at these peak latencies were displayed on the subject’s realistically-shaped head surface (Fig. 6.2.B). At the early latencies (76 and 112 ms), the VEP responses to the LVF and RVF stimuli had very different yet symmetric spatial patterns (cross correlation $r = -0.08$ and 0.53). These early VEP responses demonstrated the retinotopic relationship of underlying cortical generators that were symmetrically arranged within the right and left hemispheres respectively. At the late latency (e.g. 212 ms), the LVF and RVF stimuli produced highly correlated VEP topographies ($r = 0.97$) with bilaterally extended distribution of negative potentials.
Figure 6.3 (A) GFP time courses averaged over all the subjects (n=10). (B-C) Cross-subject average of the VEP topographic cross correlation (B)
and relative difference (C) between LVF and RVF conditions. Error bars indicate the inter-subject standard deviation.

The above findings in one subject were confirmed with the VEP group analysis (n=10). Fig. 6.3.A shows the cross-subject average of the GFP waveforms under two unilateral conditions. We quantitatively measured the similarity and dissimilarity between the VEP topographies evoked by the LVF and RVF stimuli by means of the spatial cross correlation (Fig. 6.3.B) and relative difference (Fig. 6.3.C) computed on an instant by instant basis. Consistently among all of the subjects, both the maximum correlation (0.87±0.09) and minimum relative difference (0.29±0.13) occurred around 200 ms (within a range from 170 to 250 ms). These results tend to suggest that the unilateral visual information from either the LVF or RVF passes on to anatomically convergent areas with bilateral RF during late visual processing.

We also analyzed the fMRI responses in a similar manner to localize the cortical regions that were equally responsive to both unilateral stimuli. Fig. 6.4 shows the fMRI results from one subject (S5). As shown on an axial slice along the AC-PC plane, the unilateral stimulus induced significant BOLD activations in the medial occipital and lateral occipito-temporal (LOT) regions, which corresponded to the V1 and MT+ (or V5) areas respectively (Barnikol et al. 2006; Watson et al. 1993; Dukelow et al. 2001). Clearly, the V1 activation was dominant in the contralateral (to the stimulus) hemisphere, whereas the LOT activations occurred bilaterally without an observable difference between the LVF and RVF stimuli. The BOLD activations produced by the simultaneous bilateral stimuli revealed all of the areas that were activated by either the LVF or RVF stimulus. Although the V1 activations by the bilateral stimulus were much enhanced relative to the ipsilateral V1 activation by either unilateral stimulus, the LOT activations remained almost unchanged regardless of whether the stimulus was unilateral or bilateral, suggesting a spatially invariant response at the LOT areas. Moreover, we compared the BOLD responses to the LVF and RVF unilateral stimuli through a quantified relative contribution (RC) value indicating how strong one unilateral condition relative to the other contributed to the fMRI BOLD signal at every activated voxel. The RC map (LVF
vs. RVF) showed that V1 areas with large positive (red) and negative (blue) RC values were dominantly responsive to the contralateral visual field. The LOT areas had RC values around 0 (yellow), particularly at the anterior sub-region which has previously been suggested to be the medial superior temporal (MST) area (Dukelow et al. 2001), and responded almost equally to both ipsilateral and contralateral stimuli.

**Figure 6.4** For a single subject (S5), the fMRI activation maps under the LVF (top-left) and RVF (top-right) unilateral as well as simultaneous bilateral stimuli (bottom-left) and the LVF vs. RVF relative contribution (RC) map (bottom-right) are displayed on an axial MRI slice along the AC-PC plane. All activation maps were obtained through a general linear model (GLM) analysis. The color map shows the corresponding t values with a corrected threshold set at $p \leq 0.01$. A voxel with a positive RC value (yellow to red) means stronger BOLD response to LVF than to RVF, and
vice versa for a negative RC value (blue to yellow). Thus, regions shown in yellow (RC≈0) in the RC map are deemed as equally responsive to both LVF and RVF unilateral stimuli.

Figure 6.5 fMRI activation maps in response to LVF, RVF and bilateral stimuli, individually obtained from 4 subjects.

Fig. 6.5 shows the fMRI activation maps for four individual subjects in response to LVF, RVF and bilateral stimuli respectively. Fig. 6.6 shows the individual subject’s RC map visualized on the cortical surfaces. We can tell from these images that the findings obtained in Fig. 6.4 are reproducible across different subjects.
The RC map was then averaged across subjects (n=5), as shown in Fig. 6.7. We found the contralateral dominance was most prominent in the striate cortex but became less perceivable or even disappeared in multiple extra-striate areas (such as the intraparietal cortex, V7 and MT+), which are often known with high level visual functions. We also found that the areas demonstrating approximately equal fMRI responses to both unilateral stimuli (i.e. RC close to 0) were symmetrically located in both hemispheres. This finding coincided with our observation that equal VEP responses to both unilateral stimuli were characterized by bilaterally distributed scalp potentials occurring during the late latencies (Fig. 6.2.B).
The convergent evidence from both fMRI and EEG suggests that at the late stage of visual processing, the unilateral information is processed more bilaterally (and maybe symmetrically) by extra-striate higher-tier visual areas without a demonstrable retinotopy rather than primarily on the hemisphere contralateral to the visual input (Tootell et al. 1998; Nelles et al. 2002).

6.4.2 Under-Summed Response to the Bilateral Stimulus

As previously described, two unilateral stimuli activate independent cortical processes until reaching some integrative regions which group the visual information across the vertical meridian. Since the bilaterally integrative regions tend to have spatially invariant responses, their regional responses evoked by the bilateral stimulus should be smaller than the sum of responses to both unilateral components. In contrast, independent cortical activities exclusively for processing the unilateral visual information are simply summed up when two unilateral stimuli at both hemifields are presented together. As illustrated in Fig 6.1.B, we utilized this scenario to isolate the BVI-related brain signals from those associated with other independent and unilaterally responsive processes, by subtracting the fMRI and EEG responses to the simultaneous bilateral
stimuli (denoted as “both”) from the sum of the corresponding responses to the two unilateral stimuli (denoted as “sum”).

Figure 6.8 (A) Waveforms of the global field power (GFP) of VEP responses to bilateral stimuli (denoted as “both”) and the sum of responses to both unilateral stimuli (denoted as “sum”) for one subject. (B) Scalp potential maps at three GFP peak latencies (76, 112, 212 ms) in (A) are shown in the subject’s head surface. (C) Time courses of the cross
correlation and relative difference between the VEP topographies of “both” and “sum”.

The spatial nonlinearity of the VEP response to the bilateral stimulus was first demonstrated by the data from a representative subject, as shown in Fig. 6.8.A through 6.8.C. By comparing two GFP waveforms in Fig. 6.8.A, we found the global VEP amplitude for the “both” was smaller than that of the “sum” mostly at the late latencies (after 150 ms). At three peak latencies (76, 112, 212 ms), the scalp potential maps of the “both” and “sum” are displayed for comparison in Fig. 6.8.B. The VEP topographies at the early latencies (76 and 112 ms) were closely correlated ($r = 0.96$ and 0.97) with similar global amplitudes. At 212 ms, the global amplitude of the “sum” was about twice as large as that of the “both” in spite of a high cross correlation ($r = 0.97$). Fig. 6.8.C plots the topographic correlation and the relative difference between “both” and “sum” as functions of time. Although high correlation was always found from 70 to 365 ms, large relative differences were mainly observable at 200 ms and later. These results strongly suggest that the bilaterally integrative cortical processes with spatially invariant responses are the dominant generators of the VEP signals at late post-stimulus latencies.

Nevertheless, subtly smaller responses to the bilateral stimulus were still observed in Fig. 6.8.A for two early VEP components at 76 and 112 ms respectively. Across subjects, similar under-summation phenomena were observed for the first component at 79±3 ms in 7/10 subjects and for the second component at 123±14 ms in 5/10 subjects. Fig. 6.9.A shows the group averaged GFPs for the “sum” and “both” respectively. We further extracted the under-summed VEP responses by subtracting the “both” from the “sum”. Fig. 6.9.B shows the group average of the under-summed VEP, depicted as both a GFP time course and a series of 2-D topographies at several representative latencies. We observed two early components with peak latencies at 60 and 96 ms respectively, in addition to a sustained late component (from 160 to 290 ms with the peak at 235 ms). The scalp topographies indicate focal positive potentials at the lateral occipital and posterior temporal regions (PO7 and PO8) around 60 ms; focal negative potentials at the medial occipital and posterior parietal regions (Oz and POz) around 96 ms; and bilaterally
extended negative potentials around 235 ms. These results suggest that early BVI processes emerge around 60 ms even before V1 activations. This is in light of the fact that cortical responses at V1 to the pattern-reversal visual stimulation have been believed to generate a VEP component with a 75-ms peak latency, commonly known as N75 (Barnikol et al. 2006; Tobimatsu and Celesia 2006; Di Russo et al. 2005). Correspondingly, this N75 component was also found in the present study at an average peak latency of 79 ms. The late under-summed VEP component (after 200 ms) had a qualitatively similar map as the corresponding late components under both unilateral conditions (right, Fig. 6.2.B), which further suggests that the late visual processing is bilaterally extended and symmetric and acts upon integrated visual inputs regardless of their spatial locations and extents.

**Figure 6.9** (A) Time courses of the group averaged GFPs for “both” and “sum”, respectively. (B) Time course of the group averaged GFP of the VEP component after subtracting “both” from “sum”. The corresponding 2-D topographic maps at several representative peak latencies are shown together with the time course.
The under-summed BOLD responses to the bilateral stimulus were examined through two contrasts: “LVF+RVF–Both” and “L2R–Both” (L2R denotes the interleaved bilateral stimuli). For a representative subject (S1), the t-statistic maps of both contrasts (Fig. 6.10.A) exhibited consistent spatial distributions. After overlaying these two maps on a flattened cortical surface (Fig. 6.10.B), we found a considerable overlap between them, especially at the medial, posterior and lateral occipital and inferior temporal areas. These results suggest that both contrasts reflect the same quantification for the time integral of the power of the under-summed cortical electrophysiological responses to the bilateral stimulus (Liu and He 2008). Consistent results were also obtained from other subjects as shown in Fig. 6.11.

Figure 6.10 (A) For a representative subject (S1), the t-statistic maps of two different contrasts [“LVF+RVF–Both” (left) and “L2R–Both” (right)] are shown in red-to-yellow and blue-to-green respectively. (B) Two maps
in (A) are overlaid on flattened cortical surfaces showing consistency between them.

![Figure 6.11](image.png)

**Figure 6.11** Cortical maps of under-summed BOLD response to the bilateral stimulus, for 5 individual subjects.

### 6.4.3 Spatiotemporal Cortical Source Imaging

To discern the spatiotemporal cortical activities underlying the unilateral visual processing and the bilateral visual integration, we estimated the cortical current density (CCD) distribution from the VEP signals evoked by the unilateral stimulus and the nonlinearly under-summed VEP response to the bilateral stimulus, respectively. The CCD reconstruction was performed with and without incorporating the corresponding fMRI data.

For a single subject, the reconstructed CCD images at several representative latencies are displayed in Fig. 6.12. As shown in Fig. 6.12.A, the cortical activity in response to the left stimulus was dominated at the contralateral right hemisphere at early latencies, whereas the ipsilateral activations were mainly found at late latencies. For the
late unilateral visual processing, the cortical activity was located more bilaterally rather than solely on the hemisphere contralateral to the stimulus. Such a trend was consistently observed in both the EEG-alone and the fMRI-EEG-combined imaging results, whereas the fMRI-EEG integrated analysis resulted in much higher spatial specificity.

Figure 6.12 Spatiotemporal cortical activity underlying the unilateral visual processing (A) and the bilateral visual integration (B), respectively. For both panels, images in the 1st row visualize the cortical current density (CCD) estimates obtained by using the minimum norm algorithm based on the VEP data, while those in the 2nd row visualize the fMRI-EEG integrated CCD imaging results.

Fig. 6.12.B shows the CCD estimates that accounts for the under-summed VEP (or VEP-fMRI) responses. According to our scenario illustrated in Fig. 1.B, these images indicate the spatiotemporal cortical activity underlying the bilateral visual integration.
Similar to Fig. 6.9.A, the CCD images obtained from the EEG alone and the fMRI-EEG integration were generally consistent. However, the fMRI-EEG integrated analysis provided source images with relatively more focal activations and higher spatial resolution. We can see from these images that the LOT areas on both hemispheres were activated at early latencies around 50–60 ms, suggesting that the bilateral visual information is first integrated by neurons in the LOT areas. The V1 activations were clearly observed around 70–80 ms, shortly following the LOT activations. At later latencies after around 150 ms, the cortical activity spread out onto more visual areas (such as V1, V3/V3A, V5, V7 and the intraparietal sulcus) within the medial, dorsal and lateral occipital cortex as well as the posterior parietal lobe. It suggests the recruitment of an extended network of visual areas for late visual processing. Moreover, we also found that at late latencies the cortical BVI activity appeared to be similar to the corresponding activity in response to the unilateral stimulus, which supports our previous interpretation that the late visual processing acts upon the whole visual field.

6.5. Discussion

The present results are generally consistent with the traditional view of the hierarchical visual processing (Felleman and van Essen 1991). In response to unilateral visual stimuli from the LVF and RVF, the fMRI data demonstrate that the cortical areas from the striate to the extrastriate cortex have gradually larger receptive fields crossing the vertical meridian, and that the bilaterally responsive areas are symmetrically located on both cortical hemispheres. The VEP data further suggest that the cortical visual processing tends to be independent of the spatial locations of visual inputs at the late latencies from 170 to 250 ms, and that these spatially invariant cortical responses generate bilaterally symmetrical scalp potential distributions. This converging fMRI and VEP evidence, in line with previous fMRI studies (Tootell et al. 1998; Nelles et al. 2002), suggests that spatially integrative visual processes primarily take place at the bilateral extrastriate cortex.

Nevertheless, the integrative cortical processing for the bilateral visual information is also considered beyond a solely bottom-up sequence. As observed and imaged from the spatially nonlinear VEP components, the cortical BVI process starts
even before 60 ms at the LOT regions within both hemispheres, giving rise to focal and positive potentials at the posterior and bilateral scalp surfaces. In comparison with the 79-ms peak latency of the V1 activation, visual responses at the higher-tier LOT regions may even precede the activations or re-activations in the lower-tier striate cortex. Additionally, these LOT regions have bilateral RF and demonstrate smaller BOLD-fMRI responses to bilateral stimuli relative to the sum of the BOLD responses to unilateral stimuli. Gathering these results from both VEP and fMRI, we conclude that the LOT regions are the first visual areas that pool the information from both left and right visual hemifields.

We further identify the LOT region as the MT+ complex which has been assumed to encompass both middle temporal (MT) and medial superior temporal (MST) areas (Hupé et al. 1998; Buchner et al. 1997; Hupé et al. 2001; Barnikol et al. 2006; Vanni et al. 2004; Watson et al. 1993; Dukelow et al. 2001; Nelles et al. 2002; Di Russo et al. 2005; Ffytche et al. 1995; Nowak and Bullier 1997). Anatomically, the MT+ complex is located at the juncture of the ascending limb of the inferior temporal sulcus and the lateral occipital sulcus (Watson et al. 1993; Dukelow et al. 2001) which coincides with the locations of the LOT activities observed in the present study. Neurons in the MT+ also have bilateral RF especially for the anterior sub-region MST (Tootell et al. 1998; Nelles et al. 2002; Dukelow et al. 2001). This is in agreement with the fMRI RC (LVF vs. RVF) maps shown in Fig. 2. Previous studies have also reported the MT+ activations under the elementary pattern-reversal or grating visual stimulation similar as used in the present study (Tootell et al. 1998; Buchner et al. 1997; Barnikol et al. 2006; Nelles et al. 2002; Di Russo et al. 2005), and the MT+ activities have also been found to occur very early as observed through the intracranial recordings (Lamme and Roelfsema 2000; Nowak and Bullier 1997). In addition, the fact that retinal outputs can project to bilateral MT+ areas through the fast magnocellular cortical pathway (Vanni et al. 2004) or by bypassing the V1 area via direct connections from subcortical structures (Lamme and Roelfsema 2000; Barnikol et al. 2006; Nowak and Bullier, 1997) may enable the MT+ areas to be activated earlier than most other high-tier visual areas or even the V1 area as reported in this study. Moreover, since the feedback connections from the MT+ areas have been demonstrated to be capable of modulating the neural responses at V1, V2 and V3 (Hupé et al. 1998;
Hupé et al. 2001), the MT+ areas indeed offer a reasonable source of explanation of the observed involvement of the early visual areas in the bilateral visual integration as well as in the contextual modulation (Ban et al. 2006). These results reveal an essential role of MT+ in the visual perception and brain cognition processing.

Extracting the spatial non-linear components of the visual evoked responses allows isolation of the brain signals associated with bilaterally responsive activities from other independent and unilaterally responsive activities. This provides an effective means to investigate the locations and latencies of the cortical BVI. For instance, the focal and bilateral VEP distribution arising from the LOT activities is much less observable in the VEP responses to either the unilateral or bilateral stimuli since it is temporally overlapped with the emerging V1 activation. Furthermore, the use of nonlinear VEP components in the spatiotemporal neuroimaging has the benefit of excluding both spatial and temporal interferences from non-integrative electrical activities. The ill-posedness of the EEG inverse problem results in cross-talk among the current source estimates, particularly at neighboring locations (Dale et al. 2000; Liu et al. 2006a). Since most visual areas are closely clustered at the occipital lobe, the accompanying unilaterally responsive cortical activities almost unavoidably introduce spurious source estimates and false positive temporal correlation at other cortical regions. Alternatively, linearly separating the under-summed VEP components removes the non-integrative activities in the signal space. Since the head volume conductor model is a linear quasi-static system, this procedure precludes the possibility of confounding cross-interference among source estimates without affecting the nature (i.e. locations and latencies) of the cortical BVI process. This procedure also points to an important distinction from other related electrical source imaging studies with single vs. paired visual stimuli (Vanni et al. 2004; Steger et al. 2001).

Finally, the present results also demonstrate the superiority of the fMRI-EEG integrated spatiotemporal neuroimaging (Liu and He 2008). As demonstrated in the present study, such a multimodal neuroimaging technique could provide higher resolution to both spatial and temporal domains than by using two modalities separately. Although the gained advantage of high temporal resolution relative to that of fMRI is self evident,
the enhanced spatial resolution and specificity relative to the EEG source imaging is due
to the reduced point spread function by use of additional fMRI information (Liu et al.
2006a, 2006b; Liu and He 2008; Dale et al. 2000). More specifically, the EEG source
estimates without the prior spatial constraint unavoidably spread out from the activated
neural regions onto the surrounding areas as attributable to the intrinsic ill-posedness of
the EEG inverse problem (Dale et al. 2000). In the fMRI-EEG integrated imaging, the
fMRI response is quantified to characterize the time integral of the local source power
(Liu and He 2008). Such fMRI-derived spatial constraints dampen the surrounding
spurious EEG sources and thus give rise to more focal source images. By maximizing the
mutual information of fMRI and EEG, the fMRI-EEG integrated source images are
anticipated to provide an enhanced spatial resolution and specificity to investigate the
spatiotemporal cortical activity.
Chapter 7 Conclusion and Perspectives

7.1. Conclusion

Noninvasive functional neuroimaging, as an important tool for basic neuroscience research and clinical diagnosis, has never been good enough to stop improving the spatial and temporal resolution and specificity. While existing neuroimaging modalities might approach their limits in imaging capability mostly due to fundamental as well as technical reasons, it becomes increasingly attractive to integrate multiple complementary modalities in attempt to reach significantly enhanced spatiotemporal resolution that cannot be achieved by any individual modality.

In this regard, the integration of fMRI and EEG/MEG has received most interests as well as debates. Electrophysiological and hemodynamic/metabolic signals reflect distinct but closely coupled aspects of the underlying neural activity. Combining fMRI and EEG/MEG data allows us to study brain functions from different perspectives. Convergent evidences, on one hand, definitely lead to much more confident conclusions on understanding neural mechanisms. Contradictory observations, on the other hand, also pose new hypotheses and challenges necessary to guide further investigations of the human neural system.

Our understanding in the coupling between fMRI and EEG/MEG has been substantially improved in the past decade, providing a great opportunity to combine these modalities in a more fundamental and principled way. Agreements have been reached in general regarding the neurovascular and neurometabolic coupling. Quantitative models of the cross-modal relationship, like those discussed in this review, represent important progress along this line. Although these models can at best approximate the complex interactions between hemodynamics and electrophysiology while subjected to experimental conditions and fundamental assumptions, they do serve as a sound basis for developing multimodal neuroimaging techniques, which promise to enhance the existing imaging capability, at least relative to fMRI and EEG/MEG alone.

However, this is not meant to guarantee the success of multimodal neuroimaging. Existing theories fail to explain every aspect of explosively expanding imaging datasets.
documented in thousands of research articles. To date, the primary bottleneck is still more fundamental than technical. Methodologies that rest on an assumed or modeled physiological linkage between fMRI and EEG/MEG, almost certainly fail under particular circumstances when the linkage is invalid. On one hand, theoretical modeling and experimental investigation would need to be performed across microscopic, mesoscopic and macroscopic scales, and proceed in parallel to further consolidate the physiological and physical basis for the multimodal integration. On the other hand, considerations should be taken carefully in experimental designs to self-justify the rationale of combining different modalities, regardless of whatever algorithms used to fuse the multimodal datasets. Cautions have to be taken as well to the interpretation of the imaging results.

7.2. Perspectives

In what follows, we will discuss two important remaining issues that deserve systematic investigations and may represent critical challenges as well as opportunities to future developments in the multimodal neuroimaging.

A. Frequency-dependent Neurovascular Coupling

As discussed in this review, the vast majority of methodologies integrating fMRI and EEG/MEG are based upon a linear neurovascular coupling. However, recent studies propose alternative hypothesis (Mukamel et al. 2005; Kilner et al. 2005). That is, a negative correlation between low-frequency (e.g. alpha-band) electrophysiological signals and BOLD fMRI signals (Mukamel et al. 2005; Goldman et al. 2002; Moosmann et al. 2003; Laufs et al. 2003), whereas high-frequency components (e.g. gamma-band) contribute positively to the BOLD signal (Mukamel et al. 2005; Siegel and König 2003). Although this relationship remains highly controversial and often varies considerably across subjects (Goncalves et al. 2006), it does point to some cross-modal relationship that has been ignored or unexplained by conventional neurovascular coupling models. This alternative view, although being speculative so far, may help revising the neurovascular model by extending the model from the spatiotemporal domain to the spatial-frequency or spatial-temporal-frequency domain. Along this line, an interesting paper deserves attentions (Kilner et al. 2005). In this analytical work, Friston and his
colleagues propose a heuristic model suggesting that an increase in hemodynamic signals is associated with a shift of electrophysiological power spectrum from low to high frequencies (i.e. a loss in low-frequency power relative to a gain in high-frequency power). As the authors admit, the model may be oversimplified. Its value to neuroimaging also remains to be demonstrated. Moreover, the validity of this model is obviously challenged by the fact that the stimulus-driven low-frequency power increase is associated with an increase in the BOLD response. For instance, visual stimuli presented with low temporal frequency (e.g. 4~5 Hz) induce a noticeable increase in the occipital EEG/MEG power at the stimulus frequency in accompany with a positive increase of the BOLD fMRI signals in the primary visual cortex.

In short, the frequency-dependent neurovascular coupling remains to be verified. Even if true, its impact on interpretation of fMRI, EEG/MEG and their combined imaging remains to be investigated.

**B. Imaging Brain Functional Connectivity**

One of the important applications of the multimodal fMRI-EEG/MEG imaging shall be in imaging brain functional connectivity. Static images indicating brain regions responsible for the execution of particular tasks do not convey sufficient information with respect to how these regions communicate with each other. The concept of brain connectivity now plays an important role in neuroscience, as a way to understand the organized behavior of brain regions or to reveal the functional brain circuitry (Horwitz, 2003; David et al, 2004; Ioannides 2007).

Investigators have computed cortical connectivity patterns based on hemodynamic or metabolic measurements such as fMRI (Buchel and Friston 1997; McIntosh and Gonzalez-Lima 1994, Schlosser et al, 2003, Arthurs et al, 2007), whereas the indirect nature of fMRI signals confounds the interpretation of fMRI-derived connectivity in terms of neuronal interaction (Otte and Halsband 2006). The cortical networks are formed and characterized by organized euronal oscillations that span several orders of magnitude in frequency (Buzsaki and Draguhn, 2004). As discussed in this review, the neurovascular coupling behaves as a temporal low-pass filter with a cut-off
frequency at around 0.4 Hz. High-frequency oscillations may be effectively excluded from the connectivity patterns estimated from fMRI.

EEG/MEG holds promise to reveal dynamic connectivity since it is sensitive to transient neural activities occurring on the order of milliseconds (Gevins et al, 1989; Jerbi et al, 2007). A variety of techniques have been used, most of which have amounted to evaluating the cross-correlation or phase synchronization of signals between pairs of scalp electrodes or sensors (Lachaux et al, 1999). Graph-theory-based tools from the study of complex network have also been developed to describe the connectivity of large-scale networks (Strogatz 2001). However, the relation between the observed connectivity pattern in the sensor space and that in the source space is complicated by the dispersion of electromagnetic signals from the cortex to the sensors.

Multimodal fMRI-EEG/MEG integrated neuroimaging approaches hold the potential to greatly enhance our ability to reveal the brain functional connectivity, due to the combined high spatial and temporal resolution. The more precise in our ability to image the brain functional anatomy, the better we would be able to pinpoint neural connectivity to specific brain regions. The higher temporal resolution we can achieve, the better chance we would be able to view a wide range of dynamics in functional interactions within a local or large scale. Efforts have been made in integrating fMRI with electromagnetic source imaging (Babiloni et al, 2005) using the directed transfer function (DTF) approach. Other existing methods, such as structural equation modeling (SEM) (Astolfi et al, 2005) and Partial Directed Coherence (PDC) (Baccalà and Sameshima, 2001), which have been used to extract functional connectivity from EEG source imaging data, are also applicable to fMRI-EEG/MEG imaging data (Astolfi et al, 2007).

It is also important to investigate the relationships between the functional connectivity as derived from fMRI data and from EEG/MEG data. Due to the different time scales and spatial resolutions of these two modalities there has not been clearly established consistency between the results obtained from fMRI or EEG/MEG alone. Perhaps, this is simply owing to the fact that these modalities reflect the different and complementary consequences of the same neurophysiological origin. Combining both
modalities promises to offer a more complete conclusion on the connectivity among assemblies of neurons.

Retrieving connectivity patterns from functional neuroimaging data essentially deals with a problem of data-driven system identification. That is to find a directional or non-directional functional network that can explain or predict the relationship among the data at various brain locations. While neuronal interactions are physically enabled by anatomical connections, which may be reconstructed from diffusion tensor MRI (Basser et al. 1994), the consistency between functional and anatomical connectivity remains to be investigated. Moreover, neuronal circuitry at microscopic scales entails inhibitory and excitatory connections. The dynamic yet balanced behaviors of neuronal inhibition and excitation fulfill the regional computation, as well as the large-scale synchrony and interaction among regions. The linkage between the large-scale functional network and neuronal microcircuits remains unclear. Interpretations of functional connectivity in terms of excitation and inhibition are currently missing, but potentially important.
Reference


Publication List

Peer-reviewed journal articles


**Conference Presentations / Posters**


