Maximum Discrimination Spatial Filter for EEG-Based Brain-Computer Interface Systems

A thesis presented

by

Sheng-Han Lin

to

Institute of Computer Science and Engineering

College of Computer Science

in partial fulfillment of the requirements

for the degree of

Master

in the subject of

Computer Science

National Chiao Tung University

Hsinchu, Taiwan

2006

Maximum Discrimination Spatial Filter for EEG-Based Brain-Computer Interface Systems

Abstract

Over the past decades, people have begun to explore a new communication channel directly between human brain and computers. The brain-computer interface (BCI) has the potential to enable severely disable people to drive computers directly by brain activity rather than depend on nerves and muscles. Research into BCI systems mainly involved recording of electroencephalographic (EEG) signals using surface electrodes. By recognizing the pattern of the brain activities, the system translates the messages encoded in the brain activities into computer instructions.

In this thesis, we proposed a new spatial-filter-based feature extraction method using the concept of maximum contrast beamformer. The beamformer based on overlapping sphere head model. It can reconstruct the activation magnitude of the target source and maximize the difference between specific brain states. In asynchronous BCI systems, it is hard to decide whether the target brain source is in active state. Our method successfully maximizes the variance of signals when subject is performing motor imagery tasks and suppresses the interferences in resting states. We also verified the stability of the spatial filter. After measuring the subject's sensor position, we applied the model to beamforming method with six data sets recorded in different time. We evaluated the stability of the spatial filter by calculating the correlation coefficient between them. Finally, we used sample-by-sample analysis to simulate an asynchronous BCI system and some issues about left/right hand motor tasks are discussed in the end of this thesis. After the active state was recognized, we proposed two procedures for classifying different tasks. We proposed a spatial filter design idea by generalizing Fisher linear discriminant analysis and applied it in the procedures.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Brain-Computer Interface (BCI) System Overview

Over the past 20 years, people have more knowledge and powerful tools to explore the undiscovered area of human brain. In addition, by the growing concern of the needs and potentials of people with disabilities, more and more people feel that monitoring brain activities or other measures of brain function might provide a new non-muscular channel for sending messages and commands to the external world - a brain-computer interface [29]. A general BCI flowchart is in Fig 1.1 and a common BCI system environment is illustrated in Fig 1.2.

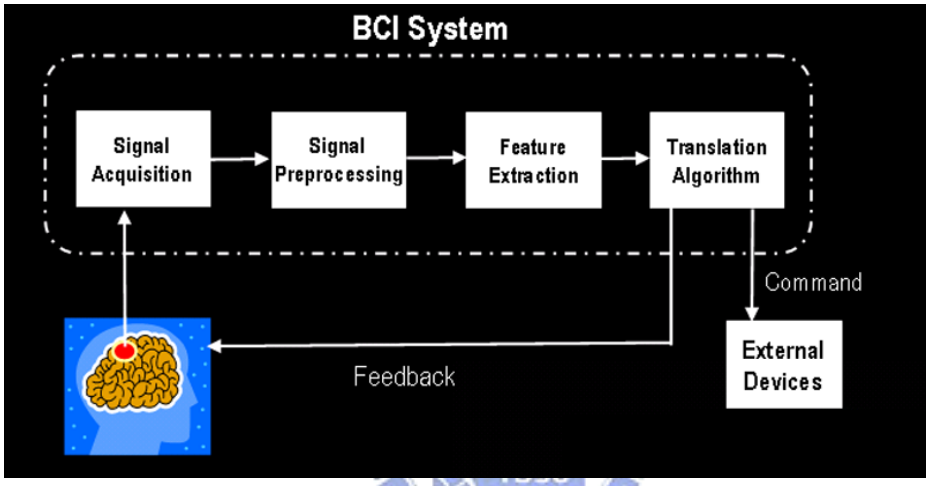


Figure 1.1: **General BCI Flowchart.** ”Signal acquisition”, ”Signal preprocessing”, ”Feature extraction”, ”Translation algorithm” are basic components of a BCI system. After translated the brain activities into device commands, the subject can control machines. And the feedback let the subject to adjust his situation dynamically.

In sections 1.1, we will briefly introduce how the system used in medical works. And in section 1.2, basic categories of current BCI systems will be illustrated.

1.1.1 Options for Restoring Function to Those with Motor Disabilities

Many different disorders will disrupt the neuromuscular channels. For instance, amyotrophic lateral sclerosis (ALS), brainstem stroke, brain or spinal cord injury, cerebral palsy, muscular dystrophies, multiple sclerosis, and numerous diseases impair the neural

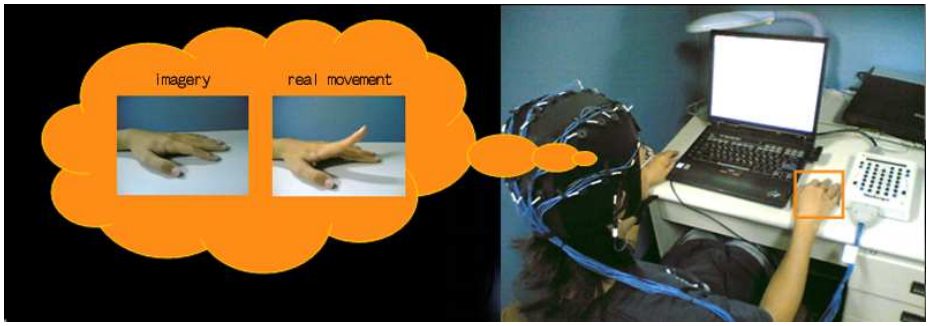
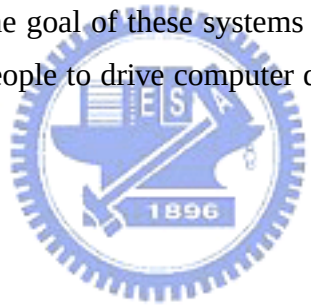


Figure 1.2: **Brain-Computer Interface Devices.**

pathways that control muscles or impair the muscle themselves [29]. Current assistant methods such as eye tracking, voice control are also finding a solution to improve the life quality of the disabled people. In recent years, BCI systems are developed in several ways such as P300 potentials, mu and beta rhythm, slow cortical potentials or ECoG recorded by implanted electrodes. The goal of these systems is to provide a communication interface for severely disabled people to drive computer directly by brain activities rather than physical means.



1.1.2 Current BCI Systems

Methods for Monitoring Brain Activity

In present days, ways to observe the brain activities have been developed. Non-invasive techniques such as functional Magnetic Resonance Imaging (fMRI), magnetoencephalography (MEG) and Positron Emission Tomography (PET) are used to monitor brain activities. On the other hand, invasive methods such as ECoG is also a method to monitor brain activities. However, fMRI, MEG, PET and ECoG are still technically demanding and expensive [29]. Generally speaking, these methods are not portable and not convenient enough for a patient to use in their daily lives and living environments.

Electroencephalography

In our research, we use electroencephalography (EEG) as the main analysis tool. Comparing to other measurements, EEG is not only an non-invasive method but also has higher time resolution and is more cheaper and portable.

The spatial resolution of EEG depends on the number of electrodes. The distribution of the electrodes is based on international 10-20 system. The ”10-20” means the relation between the location of electrodes and the underlying area of cerebral cortex.

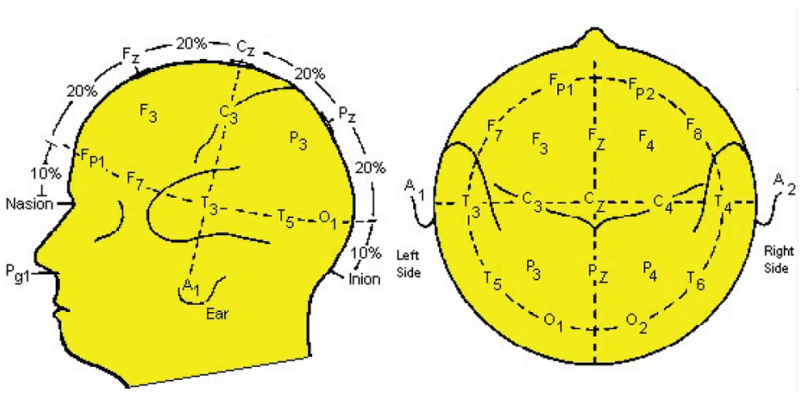


Figure 1.3: **International 10-20 System.** The ”10” and ”20” represent the distance between the electrodes. The code ”F” for frontal, ”C” for central, ”P” for parietal, ”O” for occipital and ”T” for temporal which denotes the area of human brain. (source: http://faculty.washigton.edu/chudler/1020.html)

Though using electroencephalography is a basic technique to monitor the brain activities, the details to construct a BCI system is still an open issue. According to the different electrophysiological signals that used by researchers, BCI systems can be divided into 5 groups. The first group, those using VEPs, are dependent BCIs (i.e. they depend on using muscular control of gaze direction). The other 4 groups, those using slow cortical potentials, P300 evoked potentials, mu and beta rhythms, and cortical neuronal action potentials, are called independent BCIs. For these analyzing techniques, EEG-based brain-computer systems are always facing the following problems:

1. The amplitude of the recorded EEG signals is about 10^{-6} volts, which is very small and sensitive to external interference.

2. The sampling rate is high in BCI systems, signal processing algorithms are sometimes complicated or time-consuming.

3. User's self-generated brain activities may blur the event-related activities thus mislead the recognition procedures.

4. In biomedical researches, the creature activities may varies largely. The brain pattern is always inconsistent even under the same environment and experiment.

Some details for related works will be shown in the next chapter.

1.2 Thesis Scope

In this thesis, we proposed methods in designing spatial filters which has the ability to maximize the signals of different brain states at certain brain area. In order to achieve these requirements, we proposed ideas in designing the spatial filter using beamforming techniques. It is an algorithm to estimate the source activities by MEG/EEG recordings. Based on linear constrained and minimum variance [28], we decided the dipole orientation by maximizing the signal variance between active and resting state. Furthermore, we apply the filter to EEG experiments and evaluate the performance of the filter. In the following, we briefly describe the main parts in our thesis.

1. **Designing spatial filters**: We proposed a spatial filter which is helpful to recognize the active states in asynchronous BCI systems. The spatial filter was designed by maximum contrast beamforming method which can effectively increase the variance of EEG signals corresponding to specific motor task and decrease the variance of EEG signals during resting states. It suppresses the interference from other brain areas. In designing procedure, we use morlet wavelet transform to analyze the time-frequency map of the EEG signals. By this analysis, we decide the dominant frequency band and the time period of the active state (i.e. the brain source is in activation) and control state (i.e. resting) of the signals. Then a forward model of the subject will be estimated and the power map of the whole brain will be calculated.

By this calculation, we decide the optimal source location and the spatial filter will be constructed.

2. **Performance evaluation**: Since the spatial filter is designed to maximize the difference between active and control state signal variance. The averaged variance of the signal will be illustrated and we will discuss the ability of the filter to suppress the noise from other brain areas. Furthermore, in order to verify the practicability, we compare the spatial filter designed by different signal source under the same forward model. The correlation coefficient of the spatial filters will be calculated for evaluating the stability using maximum contrast beamforming method.

3. **Simulation of an asynchronous BCI**: In the simulation of asynchronous BCI system, we use sample-by-sample analysis [26] for the evaluation of the performance.

In our work, we implemented the program for analyzing the brain signals by C/C++ and MATLAB.

1.3 Thesis Organization

In Chapter 2, we introduce the background knowledge about human brain and brain-computer interfaces. By reviewing some current BCI systems, we further discuss the role of spatial filter in signal processing procedure and discuss the designing ideas of the spatial filters. Since the goal of this thesis is to construct a spatial filter for discriminating different brain states, details of the filter will be illustrated in Chapter 3. The experiment results using our filter are shown in Chapter 4, including offline analysis and the simulation of an asynchronous BCI system. Finally, conclusions and discussions are in Chapter 5 for summarizing these methods introduced or proposed in our thesis.

Chapter 2

Background

2.1 Background Neuropsychology and Neurophysiology

In this section, we will mention about some basic issues about neuropsychology and neurophysiology. The relevant issue could be resulted from the relation between our mental state and our brain activities. However, after recording our brain signals in EEG, the signal must be translated into device commands. During this process, we also concern about the voluntarily control of their own brain activities. We will introduce the brain structure first and then discuss some mental tasks that used in BCI systems later.

2.1.1 Function of the Brain

Human brain structure can be divided into several areas. Some of the areas are in Table 2.1. In these structures, we focus on the 2-4 mm thick gray matter on the surface of the cerebrum. In the past decades, doctors and researchers have been dedicated themselves to study the functions of human brain. Up to now, the functional mapping of our brain can be illustrated in Figure 2.1. Motor cortex (the green area) concerned with controlling voluntary movements, by receiving signals from thalamic nuclei and sending output to motor control neurons in the brain stem and spinal cord. Sensorimotor cortex (the yellow area) is located prior to the central sulcus. This area responses when receiving stimulus from the external environment. Another cortex often used in BCI system is located on occipital lobe and relevant to visual stimulus.

Most of the time, subjects in BCI experiments are asked to produce and control changes in particular EEG signals by thinking about specific things or by concentrating on willing a cursor to move. Different tasks consequently activate different cortex areas. In recent studies, motor cortex acts an important role in experiments. One reason is that it controls movement of our body and directly matches the motivation that people design the non-muscular interface for disabled patients. Besides, it is a common activity in our life and the corresponding activation is concentrated on specified area. Furthermore, the right side activities of our body are basically corresponding to our left hemisphere and vice versa. These characteristics are good properties in analyzing the brain activities that encoded in brain signals.

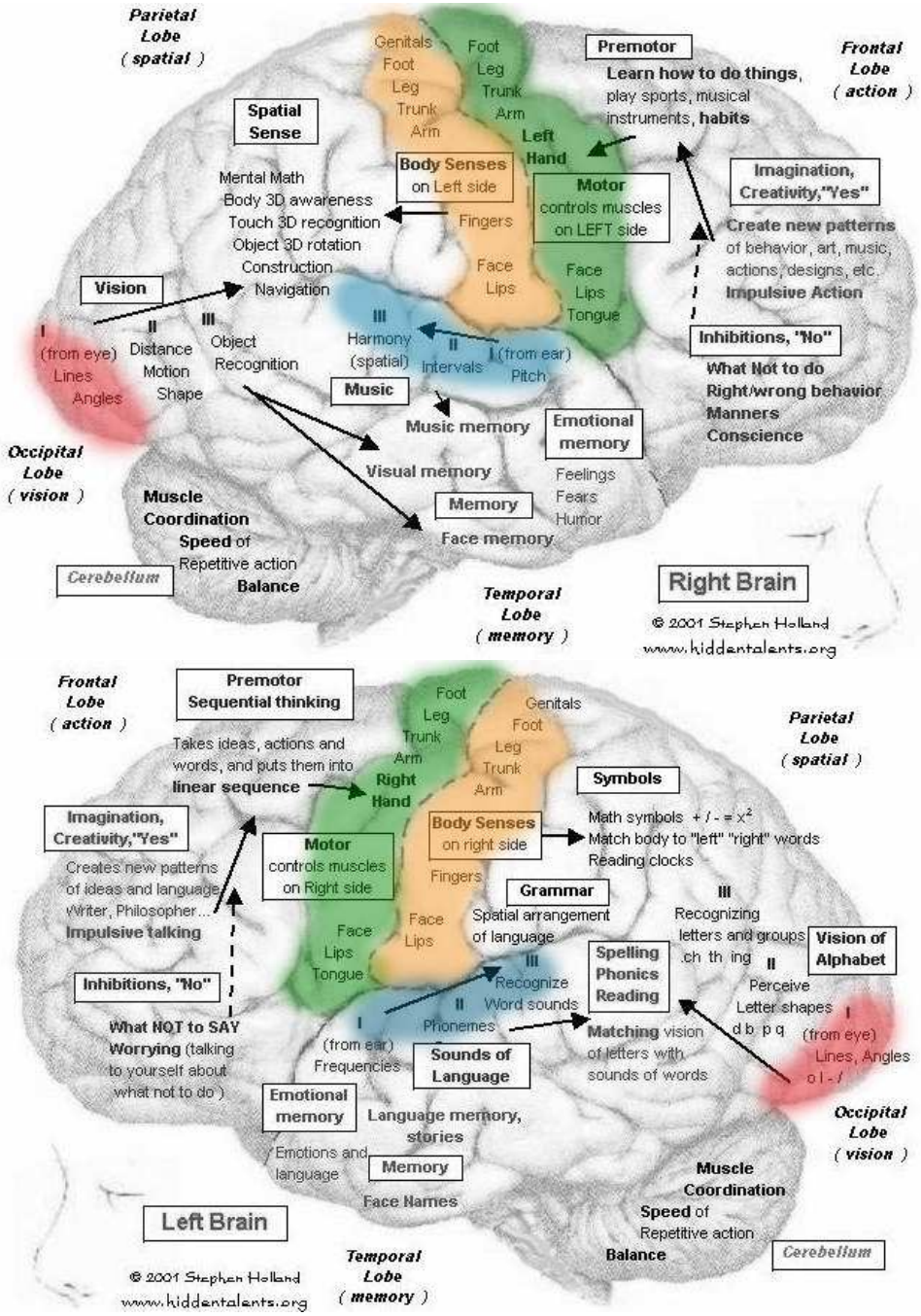


Figure 2.1: **Brain function map.** (source: http://members.shaw.ca/hidden-talents/brain/113-maps.html)

Brain structure	Function
Cerebral cortex	The outermost layer of the cerebral hemisphere is composed of gray matter. Cortices are asymmetrical. Both hemispheres are able to analyze sensory data, perform memory functions, learn new information, form thoughts and make decisions.
Left hemisphere	**Sequential Analysis:** systematic, logical interpretation of information. **Interpretation and production of symbolic information:** language, mathematics, abstraction and reasoning. Memory stored in a language format.
Right hemisphere	**Holistic Functioning:** processing multi-sensory input simultaneously to provide ”holistic” picture of one’s environment. Visual spatial skills. Holistic functions such as dancing and gymnastics are coordinated by the right hemisphere. Memory is stored in auditory, visual and spatial modalities.
Corpus callosum	Connects right and left hemisphere to allow for communication between the hemispheres. Forms roof of the lateral and third ventricles.

Table 2.1: **Basic brain structures.** (source: http://www.waiting.com/brainfunction.html)

2.1.2 Event-related Brain Signals

The history of event-related brain signals might be traced back to 1930 since Berger found that some event could block the ongoing alpha activities [7]. It is about 50 years later than the beginning of the history of cognitive science since Wundt established the first laboratory of psychology in 1879. Later in 1947, Dawson used the skill that overlaid the evoked potential (EP) and create the new generation of neural electrophysiology. In the follow paragraphs, event-related potentials (ERP) and event-related desynchronization and synchronization (ERD/ERS) techniques will be illustrated [3].

Event-related Potentials

ERP is usually defined by ”The potential variation in brain evoked by specified stimulus from external world.” The ”specified stimulus” may be the start of the stimulus or the drawback of the stimulus. In 1951, Dawson first averaged the recordings of evoked potentials to analyze the signals. This method is also called ”*Average Evoked Potentials (AEP)*” [3].

The evoked potential of EEG is about 2-10 μV [3] and hard to be observed in the ongoing signals. However, the response time and waveform of the event-related potential is fixed. With the ongoing background EEG treated as a zero-mean noise, we can increase the signal-to-noise ratio $n^{1/2}$ times after averaging EEG trials. Some common ERPs are used in BCI systems, such as P300 and visual-evoked potential (VEP). Some details are as below.

1. *P300* : P300 was found by Sutton et al. in 1965 [3]. It was initially found a positive peak appears 300ms after the onset time. This component could be detected under an oddball paradigm and the magnitude of the peak is in proportion to the effort of mental activity. The properties of P300 have described in many papers through different points [8]. Some BCI systems implemented by P300 are also discussed in [6]. Fig 2.2 is an typical P300-based BCI system, subject was told to gaze at the target word, the matrix flashes once a row and column. The P300 ERP was found at Cz and Pz and the largest magnitude evoked while the target word is hit.

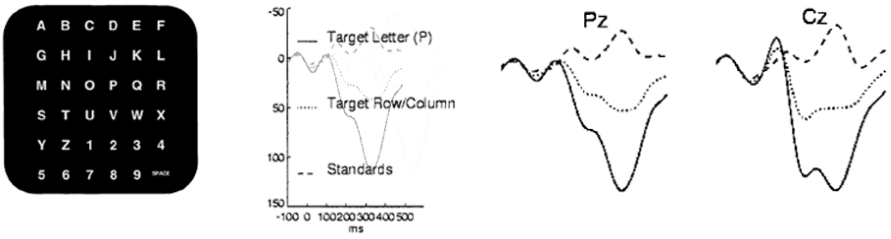


Figure 2.2: **Typical P300-based BCI system.** ”ERP’s at the midline electrodes sites recorded from the wheel-chaired subject and the able-bodies subjects. The data associated with the three types of trials are superimposed.”

2. *VEP and SSVEP* : Visual evoked potential (VEP) is a typical ERP and is easy to be detected when the eyes receive a flashlight stimulus. The response time is about 100-200ms after the visual task. Moreover, if the flashlight is not just once but appeared continuously and regularly, then steady-state visual evoked potential (SSVEP) will be generated. A typical experiment is a phase-reversing checkerboard pattern [11]. If the checkerboard reversed in certain frequency, we can see a component on the frequency in power spectrum density of the recorded signals. In Fig 2.3, the subject gaze at the center of monitor which the two checkboards will switch between each other at a random time and cause a ”flash” event.

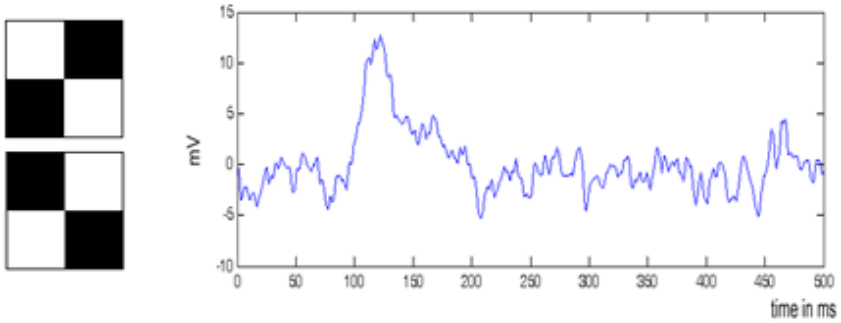


Figure 2.3: **Visual evoked potential experiment.** The signal is averaged by 100 trials, VEP was found at about 100ms where 0ms is the time the picture performs a flash.

Event-related Desynchronization and Synchronization

EEG activities could be divided into two categories - *phase-locked* and *non-phase-locked*. Phase-locked means that the responded waveform of the event-related tasks is fixed, for instance, P300 has a positive peak at 300ms after the event. All types of event-related potentials is the former case [9]. A typical non-phase-locked event is motor-related tasks, the power of mu-rhythm will decrease during the planning stage. Opposite to phase-locked potentials, this kind of phenomenon will keep a few seconds but not only ”several peak at several time”. These time-locked but non-phase-locked signals can not be extracted by linear operations such as addition, subtraction or averaging, thus frequency analysis is required [7].

Event-related desynchronization and synchronization (ERD/ERS) researches were proposed by Pfurtscheller at 1977. By the definition in [7], ERD/ERS is described below.

1. *when referring to ERD/ERS of the EEG/MEG it is necessary to specify the frequency band;*

2. *the term ERD is only meaningful if the baseline measured some seconds before the event represents a rhythmicity seen as a clear peak in the power spectrum. Similarly, the term ERS only has a meaning if the event results in the appearance of a rhythmic component and therewith in a spectral peak that was initially not detectable (Lopes da Silva and Pfurtscheller, 1999).*

Basically, there are two methods in calculating ERD - *bandpower method* and *intertrial variance method*. The former is for both phase-locked and non-phase-locked EEG activities, the later is for non-phase-locked signals only. Details of these methods could be found in [9] and thus not illustrated here.

2.2 BCI Research and Associated Motor Related Tasks

Since voluntary control of brain activity is always an essential condition for BCI systems, choosing a useful task is an important issue. Discussion about the events were mainly focused on motor cortex and motor related task in several present-day BCI systems, such as BCI2000 by Wadsworth center in New York and Graz BCI system implemented by the Austrian team.

2.2.1 Wadsworth BCI

Studies at the Wadsworth Center over the past 20 years have shown that people with or without motor disabilities can learn to control the amplitude of μ and β rhythms in EEG activity recorded from the scalp over sensorimotor cortex and can use that control to move a cursor on a computer screen in one or two dimensions [30]. 8-12 Hz μ and 13-28 Hz β rhythms are produced in sensorimotor cortex and associated areas. Their research

concentrated on defining the topographical, spectral, and temporal features of mu- and beta-rhythm control and optimizing the mutually adaptive interactions between the user and the BCI system. The analysis relies largely on the measure r^2 [30], the proportion of the total variance in mu- or beta-rhythm amplitude that is accounted for by target position and thereby reflects the user's level of EEG control. In recent years, they developed a general-purpose BCI system called BCI-2000 and have made it available to other research groups, which will be briefly described in the following paragraphs.

Essential Features

In practical BCI system design, there are many problems must be concerned. Wadsworth Center have brought up some important concepts and some experience during BCI2000 system design [24]. The essential features in designing the system are:

1. *Common Model* : A model with four modules including source, signal processing, user application and operator interface. Modules can communicate through network based on TCP/IP.

2. *Interchangeability and Independence* : The goal of this feature is to design a system which maximize the independence among the components. In other words, when changing a component, it will not effect other components.

3. *Scalability* : Scalability means that the modules do not put constraints on the parameters such as sampling rate or channel numbers.

4. *Real-Time Capability* : A real-time BCI system have to provide the ability to process the signal in a very short time. In BCI2000, they minimize the effect on response time of the operating system or other devices, which may cause delay.

5. *Support and Facilitation of Offline Analyses* : Most of the time, offline analysis needs a lot of information such as event time, event types or feedbacks during the experiment. BCI2000 records sufficient information for comprehensive analysis.

6. *Practicality* : BCI2000 not only integrates many different BCI methods but also provides documentation for researchers from engineers to end users.

Modules

In the common model, each of the components will be descried in the following paragraph. Fig 2.4 shows the flow chart illustrating general prototype in many BCI systems [24]. It can be divided into the following components:

1. *Source Module* : This module digitizes and stores the brain signals and relevant information of the acquired data. At least five modules have been created in the system to support different manufactures.

2. *Signal Processing Module* : In detailed, signal processing module could be divided into four steps. They are *feature extraction*, *feature selection*, *classification* and *translation* in order. During the first two signal processing progresses, spatial filters including Laplacian derivation, common average, independent component analysis and common spatial patterns [14] are applied. Another important technique is temporal filter, to date, they have implement five variations: slow wave filter, autoregressive spectral estimation, a finite-impulse response filter and a filter that averages evoked responses. The other two steps trains a classifier to classify and translate the inputs signals into device commands. The interdependence of the module is a critical problem in system design and BCI2000 also did a lot of effort in this module.

3. *User Application Module* : User application is the interface that user learn and control the system. BCI2000 has already implemented seven applications using different tasks.

4. *Operator Module* : This module allows the investigator to control the experiment and adjust the system parameters such as some signal processing variables.

2.2.2 Synchronous and Asynchronous BCI Systems

Generally speaking, one way to classify the BCI system is to depend on the source of stimulus - internal-paced or external-paced. Internal-paced means that the subject could

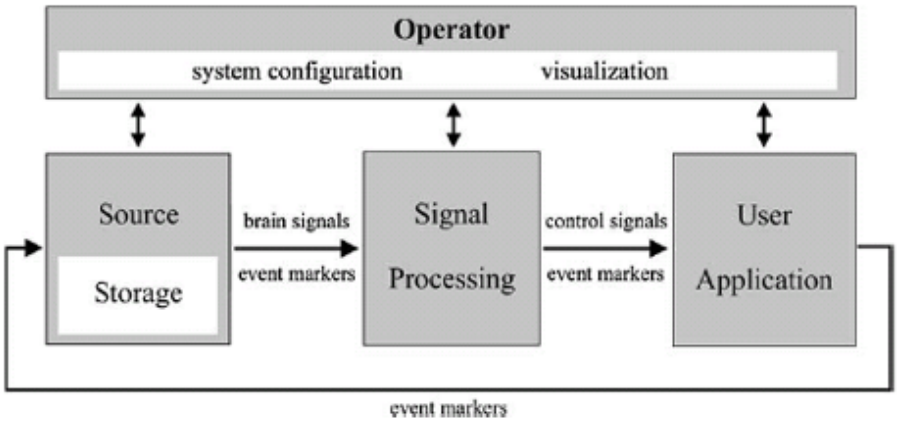


Figure 2.4: **BCI2000 design.** The four modules are the components of BCI2000. Notice that the operator component maintains the system configuration and online presentation for the investigator [24].

decide when to start the 'mental activity' by himself and in external-paced system the subject is given a cue to indicate when the 'mental activity' should start. [5].

As described in previous sections, researches about EEG depends largely on event-related tasks in the beginning stage, that is, we observe the changes of waveform after a specified event in a specific time. In a synchronous BCI system, we can get the start time of a trial by asking the subject to response after a visual or audio cue. But in practical usage of a BCI system, a cue-based system may not be suitable to the disabled.

In the past five years, many groups have tried to develop systems that can dynamically recognize the activities of subjects. Base on the understanding of brain signals until today, some basic asynchronous system have been applied to practical usage [17] [18]. The BCI user can perform a mental activity whenever he wished to submit a command to the system. The difficulties in asynchronous system may be arisen from that **we not only have to recognize where is the onset time of an active state (user performs a mental activity) but also have to classify the period of signal into certain predefined task from noise and other brain states (user is resting, idling or thinking something that may bring about some EEG activities).**

However, integrating all BCI components into a real-time system is also a challenge. Offline analysis could be complicated and time-consuming in processing signals in order to

achieve high accuracy in task recognizing or observe the detail of the signals. But in practical system, all processes must be finished in a very short time. These requirements are called ”online” in custom. Generally speaking, the goal of an online BCI system is to use as less channel as possible but keeps satisfying accuracy and efficiency. All signal processing procedures are willing to increase the signal-to-noise ratio and find useful features to support the design of BCI systems.

2.3 Spatial Filter in BCI Systems

Spatial filter is widely used in analyzing EEG signals [14] [13] [27], it is an inevitable process in many occasions. In this section, some spatial filters will be introduced.

2.3.1 Introduction to Spatial Filter

In spatial domain, suppose that the number of EEG sensor is N, then measured data from a sensor can be viewed as a vector. Therefore the N sensors is the basis and spans a vector space in the spatial domain which derives the name *”spatial filter”*. The concept of spatial filter has been discussed earlier then 1995, both spatial and temporal filter method are proposed for increasing signal-to-noise ratio. A filter can effectively suppress the noise and signals from undesired location is always of necessity in analyzing brain signals.

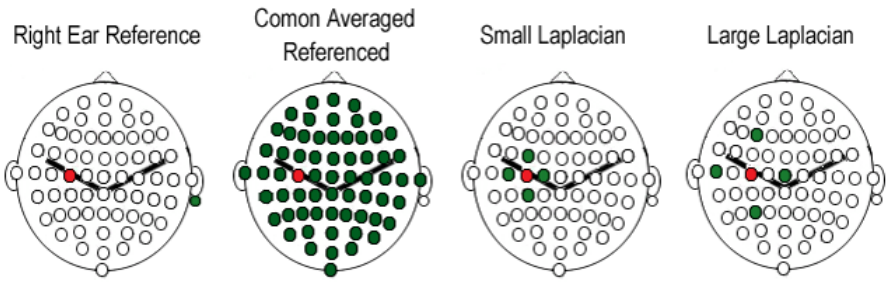


Figure 2.5: **Examples of spatial filter.** Suppose the target channel is C3 (red),for CAR filter, the recordings of C3 is subtracted by the mean of all the green channels. For Laplacian, only the mean of four channels will be subtracted. [13]

To date, some widely used spatial filters are *common average reference (CAR)*, *small Laplacian*, *large Laplacian* and *common spatial patterns*. Specifically, the design idea of each filter is as below and an example that take C3 channel as the desired channel is in Fig 2.5.

1. *Common average reference* : When applying CAR, mean of measurements from all channels will be subtracted. The filter based on the concept that uniform and entirely channel covered components are usually not representative enough in specific brain activity. In this way, global components will be removed and thus highly emphasizes focal distributions. Because focal distributions usually have higher frequency, CAR can be viewed as a function of high pass filter.

2. *Laplacian method* : Laplacian method filters the signal at a certain location by the concept of Laplation derivation. In [13], McFarland et al. use a finite difference method, which approximates the second derivative by subtracting the averaged signals surrounding the target channel. The formula is

$$\mathbf{V}_i^{LAP} = \mathbf{V}_i^{ER} - \sum_{j \in Si} g_{ij} \mathbf{V}_j^{ER} \tag{2.1}$$

   where

$$g_{ij} = 1/d_{ij} / \sum_{j \in Si} g_{ij} 1/d_{ij} \tag{2.2}$$

   Si is the set of electrodes surrounding the ith electrode and d_{ij} is the distance between electrode i and j. When d_{ij} is 3cm, the method is called small Laplacian and 6cm for large Laplacian. Most of the time, the distance is decided by the electrode on the EEG equipment such as 64-channel or 40-channel system.

3. *Common spatial patterns* : BCI group in Graz have developed an optimal spatial filter before 1999 [14] [20] [27]. Until now, this filter is used in many EEG analysis for classifying different tasks and has high classification rate. CSP is also called *common spatial subspace decomposition (CSSD)*. CSP is a feature extraction method and *"can be realized by projections of the high-dimensional, spatial-temporal raw signals onto very few specifically designed spatial filters"* [14]. It uses variance of

the two classes of data as the discriminative information. An example is in Fig 2.6, the spatial filter in the figure is for discriminating left and right motor imageries. All the details of CSP could be found in [14] and thus doesn't describe here.

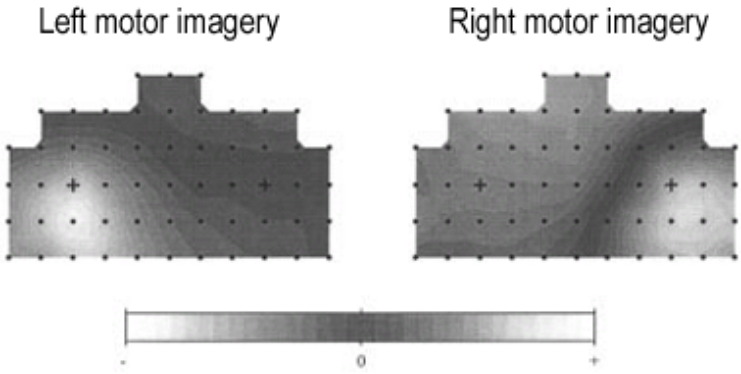


Figure 2.6: **Topography calculated by Common Spatial Pattern method.** CSP constructs spatial filters for discrimination of two populations. The topography represents the coefficient of channels. [19]

2.3.2 Model-Based Spatial Filter

Generally speaking, methods in designing spatial filter could be divided into **data-driven** and **model-based** categories. The concept of them are as follows.

1. *Data-driven analysis* : Data-driven analysis performs computations in an order dictated by data dependencies. When applying these methods to design spatial filter, the algorithm do not have to know the channel information. Take Laplacian for example, the method is defined by "subtracts the averaged signal around the target channel". And in CSSD, the algorithm calculate the projection matrix by simutaneous diagonalization of two covariance matrices [20].

2. *Model-based analysis* : A model-based analysis based on fundamental knowledge of the design and function of an object. For example, Beamforming method depends largely on the sensor positions, that is, it uses the information of the channels. The method can estimate the source signals precisely with a robust forward model.

In our thesis, we used the later method in designing our filter and details will be illustrated in next chapter.

Chapter 3

Model-Based Spatial Filter For Asynchronous BCI Systems

3.1 EEG Forward Prediction

3.1.1 Development in Forward Model

Before we start to discuss the forward prediction problems, we have to construct appropriate head model to describe our head. The problem is not simple because human brain consist of different tissues, such as cortex (gray matter and white matter), cerebrospinal fluid (CSF), skull and scalp. The history of the head model construction start from the researches about *single-shell model* which assume that human head is a simple shell. In the beginning, least-square estimation is used to find out the best-fit sphere among all EEG sensors and then map the sensor locations onto the best-fit sphere surface. In this way, we can estimate the surface potential. Furthermore, *multi-shell model* is proposed when taking different tissues into consideration. Based on this model, P. Berg, et al. [2] and Mingui Sun [25] have proposed two different ways to enhance the efficiency of multi-shell forward model. Both methods improve the efficiency of the calculation but lose some accuracy.

Nevertheless, there is still a problem - human head is not just a sphere. Prediction error will raise while the distance between sensors and sphere surface increases. A method to solve this problem is boundary element model (BEM). It uses the Magnetic Resonance Imaging (MRI) information to reconstruct the surface of tissues. Though it's very accurate but it is time-consuming. In order to improve the drawbacks, Mosher et al. presented the idea of *overlapping sphere* rather than using traditional *sensor-fitted sphere* on MEG problems [12] [23]. We then extended the method and applied on EEG works.

3.1.2 Forward Model Using Overlapping Sphere

The idea of the method is using multi-shell geometry rather than BEM model to estimate the overlapping sphere. By assuming that human head had m layers and estimate the surface potential by the second kind Fredholm integral. We use digitizer to measure the surface of realistic head and then calculate the overlapping sphere for each EEG sensor by minimizing the difference between the multi-shell sphere and realistic head. The details for constructing the spatial filter using the forward model will be introduced in next chapter.

3.2 Introduction to Beamforming Techniques

Beamforming [4] [10] is a method to localize the signal source during array signal processing. It was developed in middle of 20th century and widely used in different field such as sonar, radar and astronomical telescope array systems. The aim of this method is to calculate a set of weighting of the physical channels, called beamforming coefficients. By linearly combine the recording signals with corresponding coefficients, we can create a virtual sensor at a specified position with a specified dipole orientation. In [28], Van Veen proposed a linearly constrained minimum variance (LCMV) method for implementation of beamforming on EEG/MEG. In section 3.2.1, we will briefly introduce the data model used in beamforming technique, a simple concept of beamforming is in Fig 3.1. In section 3.2.2, the detail for calculating the dipole orientation will be explained.

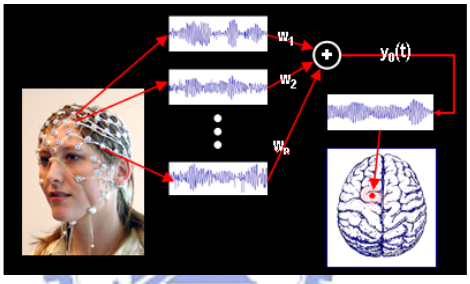


Figure 3.1: **Beamforming techniques used in EEG analysis.** The spatial filter is a set of weighting linearly combine the signals recorded by sensors.

3.2.1 Data Model

Under the system of N channel EEG sensors, the measured surface potential $\mathbf{m}$ at an instant time can be regarded as an $N \times 1$ vector expressed by

$$\mathbf{m} = \mathbf{G}(\mathbf{r})\mathbf{q} = \mathbf{G}(\mathbf{r})\frac{\mathbf{q}}{\|\mathbf{q}\|}\|\mathbf{q}\| = \mathbf{l}(\mathbf{r};\mathbf{q})\|\mathbf{q}\| \tag{3.1}$$

where $\mathbf{G}(\mathbf{r})$ is the gain matrix calculated by forward model and $\mathbf{l}(\mathbf{r};\mathbf{q})$ is the leadfield. More precisely, leadfield means the measurement with the dipole source located at $\mathbf{r}$ with

dipole moment $\mathbf{q}$ which composed by dipole orientation $\frac{\mathbf{q}}{\|\mathbf{q}\|}$ and dipole strength $\|\mathbf{q}\|$. Furthermore, when there are $\mathbf{k}$ dipole sources at an instant time, we model the noise as an $N \times 1$ vector $\mathbf{n}$. The measured data can be rewritten as

$$\mathbf{m} = \sum_{i=1}^{k} \mathbf{l}(\mathbf{r_i}; \mathbf{q_i})\|\mathbf{q_i}\| + \mathbf{n} \tag{3.2}$$

where $\mathbf{q_i}$ ($i = 1,2,\ldots,k$) is the ith dipole moment.

Notice that the equation above represents the measurement at an instant time. In time domain, bio-medical signal is often modeled as a random signal and thus we take temporal information into consideration and we use first and second order statistics to describe the dipole as

$$\bar{\mathbf{q}}_\mathbf{i} = E\{\mathbf{q_i}\} \tag{3.3}$$

$$c_{q_i} = E\{\,[q_i - \bar{q}_i][q_i - \bar{q}_i]^T\}\,, \tag{3.4}$$

respectively, where E stands for expectation. Furthermore, the mean and covariance matrix of the measurement are

$$\bar{\mathbf{m}} = E\{\mathbf{m}\} = \sum_{i=1}^{k} \mathbf{l}(\mathbf{r_i}; \mathbf{q_i})\bar{q}_i \tag{3.5}$$

$$\mathbf{C} = E\{\|\mathbf{m}(\mathbf{q_i}) - \bar{\mathbf{m}}\|\|\mathbf{m}(\mathbf{q_i}) - \bar{\mathbf{m}}\|^T\} = \sum_{i=1}^{L} \mathbf{l}(\mathbf{r_i}; \mathbf{q_i})c_{\mathbf{q_i}}\mathbf{l}^T(\mathbf{r_i}; \mathbf{q_i}) + \mathbf{C}_n \tag{3.6}$$

respectively, where $\mathbf{C}_n$ is the covariance of the noise under an assumption of zero mean. Practically, $\mathbf{C}$ is calculated by using recorded EEG signals.

$$\mathbf{C} = \frac{1}{N-1}\mathbf{M}_{MDF}\mathbf{M}_{MDF}^T \tag{3.7}$$

where T is the sampling number and $\mathbf{M}$ is an $N \times T$ matrix which represents the recorded EEG signals. The subscript "MDF" denotes Mean-Deviation Form which each element is substituted by the mean of row of original matrix (i.e. averaged potential of each sensor).

3.2.2 Spatial Filter Design

As mentioned in previous section, beamforming is designed to reconstruct the source activation by linearly combine the recordings from each EEG sensor. The idea can be written as

$$y = \mathbf{w}^T(\mathbf{r_0}; \mathbf{q_0})\mathbf{m} \tag{3.8}$$

where y is the reconstructed moment with dipole location r_0 and dipole orientation $\frac{\mathbf{q}}{\|\mathbf{q}\|}$, and $\mathbf{w}^T(\mathbf{r_0}; \mathbf{q_0})$ is an $N \times 1$ vector which denotes the spatial filter. By LCMV, there are two constraints in finding w. The first one is **linearly constrained**:

$$\mathbf{w}^T(\mathbf{r_0}; \mathbf{q_0})\mathbf{l}(\mathbf{r_0}; \mathbf{q_0}) = 1 \tag{3.9}$$

which extracts the target source ($\mathbf{r} = \mathbf{r_0}$ and $\mathbf{q} = \mathbf{q_0}$) and suppresses other sources ($\mathbf{r} \neq \mathbf{r_0}$ and $\mathbf{q} \neq \mathbf{q_0}$). This constraint is also called *unit gain constraint* because after filtering the predicted potential, we would get the original source. The second idea of LCMV is **minimum variance**:

$$\min_{\mathbf{w}(\mathbf{r_0};\mathbf{q_0})} c_y s.t. \mathbf{w}^T(\mathbf{r_0}; \mathbf{q_0})\mathbf{l}(\mathbf{r_0}; \mathbf{q_0}) = 1 \tag{3.10}$$

where c_y is the variance of the estimated signal. The reason to minimize the variance of the filtered signal is that if forward model is exactly correct and without noise, then

$$y_0 = \mathbf{w}^T(\mathbf{r_0}; \mathbf{q_0})\mathbf{m} = \mathbf{w}^T(\mathbf{r_0}; \mathbf{q_0})\mathbf{l}(\mathbf{r_0}; \mathbf{q_0})\mathbf{q_0} = 1 \times \mathbf{q_0} = \mathbf{q_0} \tag{3.11}$$

where $\mathbf{q_0}$ is the true source moment at the target position. The details in solving the filter $\mathbf{w}$ are in [28] and the equation is:

$$\mathbf{w} = (\mathbf{C} + \alpha\mathbf{I})^{-1}\mathbf{l}(\mathbf{l}^T(\mathbf{C} + \alpha\mathbf{I})^{-1}\mathbf{l})^{-1} = \frac{(\mathbf{C} + \alpha\mathbf{I})^{-1}\mathbf{l}}{\mathbf{l}^T(\mathbf{C} + \alpha\mathbf{I})^{-1}\mathbf{l}} \tag{3.12}$$

where α is a regularization parameter, $\mathbf{C}$ is the covariance matrix explained in previous section and $\mathbf{I}$ is the identity matrix. Here we omit $(\mathbf{r_0}; \mathbf{q_0})$ for simplicity.

However, there is still one question - "How do we know the dipole orientation?" . In accordance with this question, LCMV decomposes the orientation solution space with 3

orthogonal basis in 3D space [28]. Robinson and Vrba proposes *synthetic aperture magnetometery (SAM)* method to search the orientation such that the resultant value of z-deviate is maximum [22]. However, we proposed a method to calculate the optimal dipole orientation analytically. The following section will illustrate the details in this method and explain how we use it in designing a filter for an asynchronous BCI system.

3.3 Spatial Filters for Asynchronous BCI Systems

All the following topics in designing our asynchronous BCI system could be divided into two parts.

1. *Recognition of active state* : As mentioned in subsection 2.2.2, we want to design a filter that can recognize the active state during the execution of BCI system. In subsection 3.3.1, we use *maximum contrast beamformer* (MCB) method for this situation. Details of MCB will be illustrated later in subsecion 3.3.1.

2. *Classification of different tasks* : After we recognize the active state, we can further classify the signals. We proposed a filter construction method in order to increase the distance between each group. In subsection 3.3.2, we will introduce how to use Fisher's criterion to design this filter.

3.3.1 Maximum Contrast Beamformer for Brain State Analysis

Maximum Contrast Beamformer

The decision of dipole orientation is an important issue in beamforming techniques. A correct dipole orientation can successfully suppress the undesired noise. The idea of MCB is finding the optimal dipole orientation by maximizing the ratio of active state and control state. In the beginning, recall that the definition in section 3.2.2 the leadfield $\mathbf{l} = \mathbf{G}(\mathbf{r})\frac{\mathbf{q}}{\|\mathbf{q}\|}$ can be rewritten as $\mathbf{l} = \mathbf{G}\mathbf{j}$ and substitute it into Eq 3.12 we have

$$\mathbf{w} = \frac{(\mathbf{C}+\alpha\mathbf{I})^{-1}\mathbf{l}}{\mathbf{l}^T(\mathbf{C}+\alpha\mathbf{I})^{-1}\mathbf{l}} = \frac{(\mathbf{C}+\alpha\mathbf{I})^{-1}\mathbf{G}\mathbf{j}}{\mathbf{j}^T\mathbf{G}^T(\mathbf{C}+\alpha\mathbf{I})^{-1}\mathbf{G}\mathbf{j}} \doteq \frac{\mathbf{A}\mathbf{j}}{\mathbf{j}^T\mathbf{B}\mathbf{j}} \tag{3.13}$$

where $\mathbf{A} = (\mathbf{C} + \alpha\mathbf{I})^{-1}\mathbf{G}$ and $\mathbf{B} = \mathbf{G}^T(\mathbf{C} + \alpha\mathbf{I})^{-1}\mathbf{G}$. Notice that the dipole orientation j could be extracted. In the idea of MCB, we maximize the ratio between active and control state by using F statistic for the criterion in deciding the ratio. The formula is

$$\mathbf{F} = \frac{\mathbf{w}^T\mathbf{C_a}\mathbf{w}}{\mathbf{w}^T\mathbf{C_c}\mathbf{w}} \tag{3.14}$$

After substituting Eq 3.13 into Eq 3.14, the formula can be translated as:

$$\tilde{\mathbf{j}} = arg\max_{\mathbf{j}} \frac{(\frac{\mathbf{Aj}}{\mathbf{j}^T\mathbf{Bj}})^T\mathbf{C_a}(\frac{\mathbf{Aj}}{\mathbf{j}^T\mathbf{Bj}})}{(\frac{\mathbf{Aj}}{\mathbf{j}^T\mathbf{Bj}})^T\mathbf{C_c}(\frac{\mathbf{Aj}}{\mathbf{j}^T\mathbf{Bj}})} = \max_{\mathbf{j}} \frac{\mathbf{j}^T\mathbf{A}^T\mathbf{C_a}\mathbf{Aj}}{\mathbf{j}^T\mathbf{A}^T\mathbf{C_c}\mathbf{Aj}} \doteq \max_{\mathbf{j}} \frac{\mathbf{j}^T\mathbf{Pj}}{\mathbf{j}^T\mathbf{Qj}} \tag{3.15}$$

where $\mathbf{P} = \mathbf{A}^T\mathbf{C_a}\mathbf{A}$ and $\mathbf{Q} = \mathbf{A}^T\mathbf{C_c}\mathbf{A}$. Now we can know that it is a traditional optimization problem in solving j and the solution is the eigenvector with respect to the maximum eigenvalue of matrix $\mathbf{Q}^{-1}\mathbf{P}$. Therefore, we determined the source orientation with deterministic computational steps.

MCB in asynchronous BCI system

However, in biomedical engineering domain, the intersubject variability is always a critical problem. In section 2.2, we can realize that not only the peak frequency but also peak latency [15] of subjects are always different. Therefore, when take the theory into practical use, we must take the following topics into consideration.

1. *Frequency band* : The peak frequency band of the subject must be decided, we use continuous wavelet transform (CWT) for peak frequency selection. Morlet wavelet transform is applied and proper band width will be decided according to the time-frequency map.

2. *Signal range for active/control state* : While selectiing the time period of active state , we tested different signal ranges and 0.5 second active range is found to be the best width. On the other hand, resting state is more stable when the selected signal range changes.

3. *Dipole location* : As mentioned in section 3.2.1, dipole location is used to calculate forward model. But in reality, we don't know the exact position of the brain source.

In our methods, we searched the brain and find the location with maximum contrast beamformer.

4. *Optimal regularization parameter* : The regularization parameter refers to the α in Eq 3.12. A value of 10^{-7} is applied in the general cases.

5. *Feature extraction* : After the signal was filtered by the spatial filter, we extract the feature for classification by calculating the variance of the signal. The variance of a signal at an instant time is calculated by the nearest 500ms periods.

6. *Classification of brain states* : We use a threshold to define the boundary of active state and control state. We determine that the subject is performing motor related tasks when the variance is larger than the threshold or the subject should be resting.

All the results of experiments and the details for deciding the parameters above will be shown in next chapter.

3.3.2 Different Tasks Classification Using Maximum Discrimination Beamfroming Techniques

Fisher Linear Discriminant

This linear discriminant method is proposed by Ronald A. Fisher in 1936 [21] and widely used in many classification problems such as image recognition. We may begin to introduce this method by considering a problem that projects data from d dimensions onto a line. The idea of the method is in [21]. It is basically designed for two-class classification. The goal is to find the most discriminant projection vector w. Fig 3.2 is an example, the two populations can be regarded as right group and left group after the projection [21]. The projected data would have two properties:

1. The distance between the two projected group means will be maximized.

2. The variance of each group will be minimized.

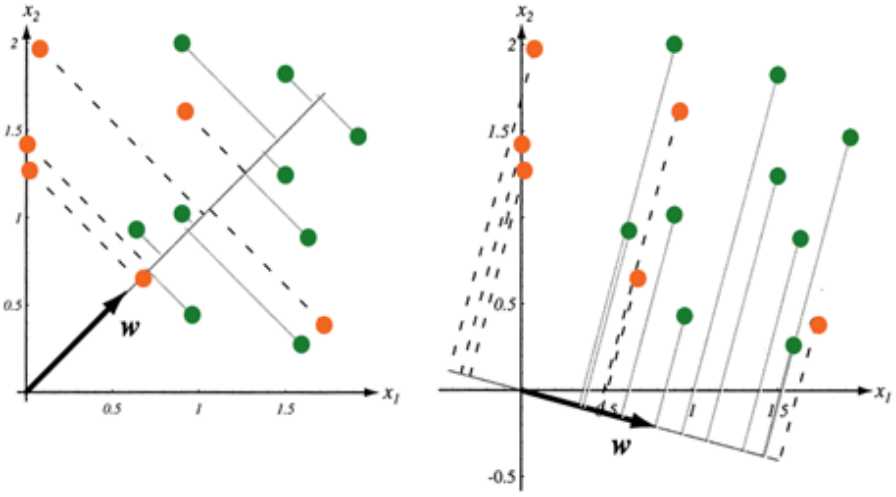


Figure 3.2: **Projections in different directions.** A same set of data was projected onto two different lines in direction $\mathbf{w}$. The projected points were more separable on the right figure [21].

Suppose we have two sets of d-dimensional samples $\mathbf{x_1}, ..., \mathbf{x_n}$, n_1 in set D_1 and n_2 in set D_2. Then the projection can be written as

$$\mathbf{y_i} = \mathbf{w^T x_i}, \forall i = 1, ..., n \tag{3.16}$$

where $\mathbf{w}$ and $\mathbf{x_i}$ are $d \times 1$ vectors, and the projected data $\mathbf{y_i}$ is in G_1 and G_2 corresponding to D_1 and D_2 respectively. The mean $\mathbf{m_1}$ of set D_1 is given by

$$\mathbf{m_1} = \frac{1}{n_1} \sum_{x_i \in D_1} \mathbf{x_i} \tag{3.17}$$

and the mean $\mathbf{m_1'}$ which projected from data in set D_1 is given by

$$\mathbf{m_1'} = \frac{1}{n_1} \sum_{y_i \in G_1} \mathbf{y_i} = \frac{1}{n_1} \sum_{x_i \in D_1} \mathbf{w^T x_i} = \mathbf{w^T m_1} \tag{3.18}$$

which is the projection of m_1 and we can have the projected mean of set D_2 by the same way. Now we can define the distance between the two groups by

$$|\mathbf{m_1'} - \mathbf{m_2'}| = |\mathbf{w^T}(\mathbf{m_1} - \mathbf{m_2})| \tag{3.19}$$

Remember the second property of Fisher linear discriminant is to minimize the variance of projected data. Take set G_1 for example, rather than using simple variance $\mathbf{s_1}$, we define the *scatter* instead.

$$\mathbf{s_1}^2 = \sum_{y_i \in G_1} (\mathbf{y_i} - \mathbf{m'_1})^2 \tag{3.20}$$

Similarly, the variance $\mathbf{s_2}$ of set G_2 could be calculated in the same way. Thus, the criterion function of Fisher linear discriminant can be defined by

$$\mathbf{J(w)} = \frac{|\mathbf{m'_1} - \mathbf{m'_2}|^2}{\mathbf{s_1}^2 + \mathbf{s_2}^2} \tag{3.21}$$

Now the idea of Fisher's criterion is illustrated, and then we adapt the criterion function in order to find the optimal $\mathbf{w}$. For this purpose, we modify the numerator and denominator of the equation by

$$\begin{aligned} |\mathbf{m'_1} - \mathbf{m'_2}|^2 &= (\mathbf{w^T m_1} - \mathbf{w^T m_2})^2 \\ &= \mathbf{w^T}(\mathbf{m_1} - \mathbf{m_2})(\mathbf{m_1} - \mathbf{m_2})^T\mathbf{w} \\ &= \mathbf{w^T S_B w} \end{aligned} \tag{3.22}$$

where $\mathbf{S_B} = (\mathbf{m_1} - \mathbf{m_2})(\mathbf{m_1} - \mathbf{m_2})^T$ and

$$\begin{aligned} \mathbf{s_1}^2 + \mathbf{s_2}^2 &= \sum_{y_i \in G_1} (\mathbf{y_i} - \mathbf{m'_1})^2 + \sum_{y_i \in G_2} (\mathbf{y_i} - \mathbf{m'_2})^2 \\ &= \sum_{x_i \in D_1} (\mathbf{w^T x_i} - \mathbf{w^T m_1})^2 + \sum_{x_i \in D_2} (\mathbf{w^T x_i} - \mathbf{w^T m_2})^2 \\ &= \sum_{x_i \in D_1} \mathbf{w^T}(\mathbf{m_1} - \mathbf{m_2})(\mathbf{m_1} - \mathbf{m_2})^T\mathbf{w} + \sum_{x_i \in D_2} \mathbf{w^T}(\mathbf{m_1} - \mathbf{m_2})(\mathbf{m_1} - \mathbf{m_2})^T\mathbf{w} \\ &= \mathbf{w^T S_1 w} + \mathbf{w^T S_2 w} \\ &= \mathbf{w^T S_W w} \end{aligned} \tag{3.23}$$

where $\mathbf{S_W} = \mathbf{S_1} + \mathbf{S_2}$. We call S_B the *between-class scatter matrix* and S_W the *within-class scatter matrix*. Consequently, Fisher's criterion function can be rewritten as

$$\mathbf{J(w)} = \frac{\mathbf{w^T S_B w}}{\mathbf{w^T S_W w}} \tag{3.24}$$

The goal of Fisher linear discriminant is to find an optimal $\mathbf{w}$ to maximize the object function $\mathbf{J}(\mathbf{w})$.

Maximum Discrimination using Fisher linear discriminant

Base on Fisher's method, we further apply the concept to classify different EEG tasks. Recall that our goal is to maximize the difference between different tasks, now we reform the procedure as the following.

Suppose the channel number is N and the number of sample points of a trial is T, we generalize the d-dimensional data to our $N \times T$ EEG data. Similar to the previous paragraph, we define $\mathbf{x_i}$ as $N \times T$ EEG trials. For the two classes of tasks, number of trials n_1 is in set D_1 and n_2 is in set D_2. Therefore, the mean of each group is the same as Eq 3.17, and notice that $\mathbf{x_i}$ is a matrix. After the data was projected by $\mathbf{w}$ (i.e. the $N \times 1$ spatial filter), we would have the projected data $\mathbf{y_i}$ (an $1 \times T$ vector) which represents the filtered signal.

Under this situation, the calculation of projected mean is the same as Eq 3.18, but notice that $\mathbf{m_1'}$ and $\mathbf{m_2'}$ are both $1 \times T$ vectors which mean the averaged projected signal. Then the remaining problem is how to define the distance between group means and how to calculate the group variance. We generalize the concept "distance" from Eq 3.22 by

$$\begin{aligned} \|\mathbf{m_1'} - \mathbf{m_2'}\|^2 = & \|\mathbf{w^T m_1} - \mathbf{w^T m_2}\|^2 \\ & = \mathbf{w^T}(\mathbf{m_1} - \mathbf{m_2})(\mathbf{m_1} - \mathbf{m_2})^T\mathbf{w} \\ & = \mathbf{w^T S_B w} \end{aligned} \tag{3.25}$$

where $\mathbf{S_B}$ is the new *between-class scatter matrix*.

It is easy to be realized that if we consider the original Fisher's linear discriminant as data at "an instant time" in EEG. The concept of distance is simply generalized to "a series of time". By the same way, we generalize the *scatter* by

$$\mathbf{s_1}^2 = \sum_{y_i \in G_1} \|\mathbf{y_i} - \mathbf{m_1'}\|^2 \tag{3.26}$$

and the same with $\mathbf{s_2}^2$. Thus the *within-class scatter matrix* can be defined by

$$
\begin{aligned}
\mathbf{s_1}^2 + \mathbf{s_2}^2 &= \sum_{y_i \in G_1} \|\mathbf{y_i} - \mathbf{m_1'}\|^2 + \sum_{y_i \in G_2} \|\mathbf{y_i} - \mathbf{m_2'}\|^2 \\
&= \sum_{x_i \in D_1} \|\mathbf{w^T x_i} - \mathbf{w^T m_1}\|^2 + \sum_{x_i \in D_2} \|\mathbf{w^T x_i} - \mathbf{w^T m_2}\|^2 \\
&= \sum_{x_i \in D_1} \mathbf{w^T}(\mathbf{m_1} - \mathbf{m_2})(\mathbf{m_1} - \mathbf{m_2})^T\mathbf{w} + \sum_{x_i \in D_2} \mathbf{w^T}(\mathbf{m_1} - \mathbf{m_2})(\mathbf{m_1} - \mathbf{m_2})^T\mathbf{w} \\
&= \mathbf{w^T S_1 w} + \mathbf{w^T S_2 w} \\
&= \mathbf{w^T S_W w}
\end{aligned}
\tag{3.27}
$$

and we may notice that the form of criterion function $\mathbf{J(w)}$ is the same as Eq 3.14 and thus we calculated the projection vector in the same way with Eq 3.14.

By the end of this chapter, we summarize the idea for designing the maximum discriminant spatial filter. In our methods, we use Fisher's criterion to find an optimal dipole orientation which can maximize the difference between the two classes of tasks. We generalize the data format of Fisher discriminant analysis from d-dimesional sample to $N \times T$ dimensional sample (i.e. from "an instant time" to "a series of time"). In next chapter, we will show the results by applying the filter.

Chapter 4

Experiment Results

4.1 Experiments

4.1.1 Experiment Paradigm

The experiment paradigm is in Fig 4.1. At the beginning of a trial, a fixation cross is shown in the middle of screen and disappeared at 2s. After the warning tone, a visual cue (an arrow points to the left or right) will appear and last for 1.25 seconds. The subject is told to do the finger movement/imagery task after the cue disappeared. Left arrow is for left finger lifting task and vice versa. In this experiment, a session is about 20 minutes and the period of a trial is 8 seconds. Left or right arrow appears randomly during the experiment.

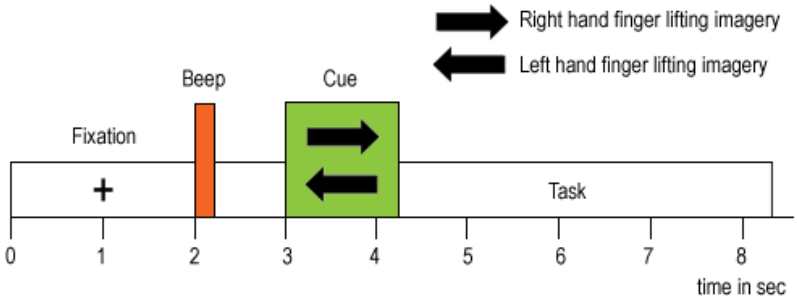


Figure 4.1: **Experiment paradigm.** The experiment is for right/left finger lifting tasks

4.1.2 Data Sets

In our experiment, we have data sets from 24 years old male. Six data sets was acquired about once per month, three of them are experiments of real finger lifting task (labeled R01-03) and the others are of imagery task (labeled I01-03).

4.2 Implementation of Spatial Filter

Recent studies show that cognitive tasks and motor functions are relevant to specific frequency band of EEG signals [15] [16]. In our methods, it is also important to decide the dominant frequency band. Details will be illustrated and discussed in this section.

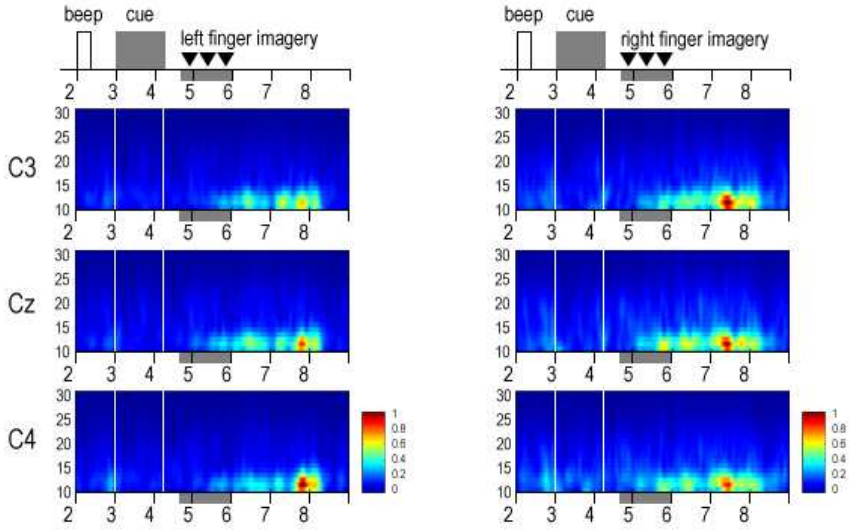


Figure 4.2: **Time-frequency map of I01.** The map has been normalized to [0 1] respectively. The dominant frequency is around 10-15Hz and the post-movement power appears about 1 second later.

4.2.1 Time-frequency analysis

Frequency Band Selection

In order to estimate the dominant frequency band, we applied continuous wavelet transform (CWT) to EEG signals. CWT is a time-frequency analysis method in signal processing procedures and each trials was calculated using morlet wavelet. In Fig 4.2, , the frequency with range from 10 to 30 Hz was illustrated and upper alpha band (10-13 Hz) could be determined as a dominant frequency band.

In our method, we use 10-15 Hz as the dominant frequency band for this subject. A band pass filter was applied to the raw EEG signals before the proceeding calculations.

Active State Selection

In maximum contrast beamforming method, trials are used to calculate the covariance matrix of active and control states. Thus an appropriate period of signals which can represents the states is very important. Different ranges of active state are tested and we found that 0.5 second width (covered by the red area in the time-frequency map) is the best range

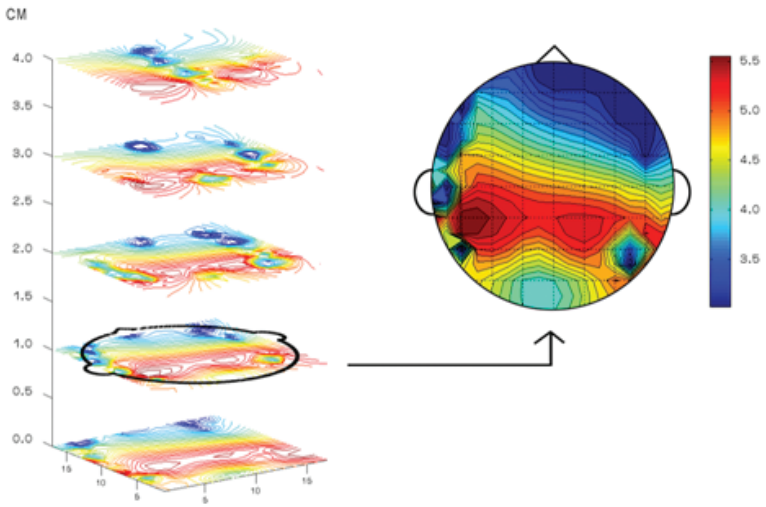


Figure 4.3: **Tomography of data R01.** The brain was divided into grids and each F value was calculated using maximum contrast beamforming methods.

for the maximum contrast. On the other hand, we used 1-second range for resting state signals. Since the signal of the resting state is more stable (the subject do nothing with eyes opened), the range of resting state didn't effect the result very much.

4.2.2 Brain Activation Tomography

Tomography Using Maximum Contrast Beamformer

As mentioned in section 3.3.1, the spatial filter was calculated under a given dipole location (i.e. activated brain source). Since we will not know the exact dipole location of the subject, we searched the whole brain using MCB.

After divided the brain into grids, we calculated the optimal dipole orientation which maximizes the active and control state for each grid. In Fig 4.3, a set of tomography of right hand real movement task (data set R01) was calculated. The peak was calculated at motor cortex which located at left hemisphere. By this tomography, the activated brain source could be determined and we can further use the dipole location in designing our spatial filter.

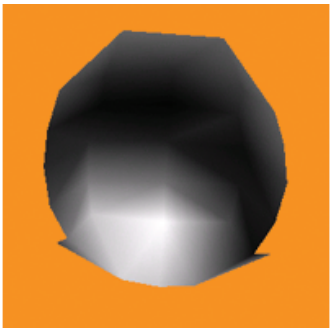


Figure 4.4: **Topography of data R01.** The topography of data R01 (right hand movement trials) was used to calculate the spatial filter. Notice that the activation brain source is at the left hemisphere with respect to right hand movement.

Topography of Spatial Filter

After the pre-processing procedures in previous sections, the spatial filter $\mathbf{w}$ introduced in section 3.3.1 could be calculated. Recall that $\mathbf{w}$ is the weighting corresponding to the EEG sensors and the filtered signal can be viewed as a linear combination of all the EEG signals. The topography of the spatial filter calculated using data R01 (right hand movement trials) is in Fig 4.4.

Stability of Spatial Filter

Furthermore, we considered a question that a EEG subject may encounter - ”How to estimate the sensor positions which constructed the forward model?”. The measurement of the sensor position is not easily to be acquired at any time. In accordance with this concern, we applied the sensor positions estimated during the experiment R01 to each data set. In Fig 4.5, right hand imagery/movement trials were calculated under the same gainmatrix (i.e. Using the same sensor position and dipole location).

In Fig 4.6, left hand trials were used to calculate the spatial filters. We can see that the correlation coefficients are more similar to those calculated using right hand trials. It is reasonable that non-dominant hand is always more distinguishable in active and control states [1].

In this section, we illustrated the details in constructing spatial filters and we verified

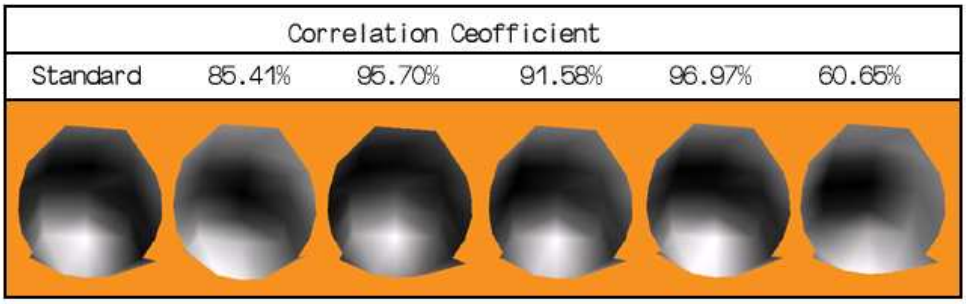


Figure 4.5: **Topography using the same sensor position. (Right finger imagery/movement trials)** We took the first topography to be the standard spatial filter and other data sets were compared with it by calculating correlation coefficient . Data labeled from left to right were R01, R02, I01, I02, I03, R03.

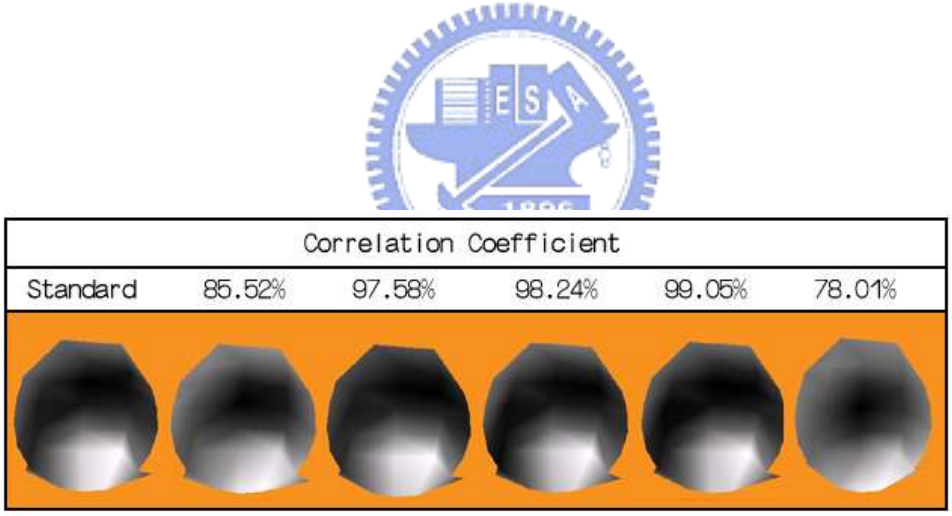


Figure 4.6: **Topography using the same sensor position. (Left finger imagery/movement trials)** We took the first topography to be the standard spatial filter and other data sets were compared with it by calculating correlation coefficient . Data labeled from left to right were R01, R02, I01, I02, I03, R03.

the stability of maximum contrast beamforming with different data sets. In next section, we will evaluating the performance of the filter by simulation of asynchronous BCI systems.

4.3 Performance Evaluation

4.3.1 Unbalanced features

When evaluating the performance in asynchronous BCI systems, a common problem is that the time period of active state is always far less than resting states. In [26], Pfurtscheller et al., proposed an evaluating method using operating characteristics curves (ROC) to define the accuracy of an asynchronous BCI system and we use sample-by-sample analysis as the measurement. The two axes of the ROC curves are true positive rate (TPR) and false positive rate (FPR). TPR is used for measurements of sensitivity and FPR is used for selectivity [26]. The definition is:

$$\mathbf{TPR} = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FN}}, \mathbf{FPR} = \frac{\mathbf{FP}}{\mathbf{TN} + \mathbf{FP}} \tag{4.1}$$

where TP, FN, TN, and FP are the number of true positive, false negative, true negative, and false positive respectively. Note that the type is defined sample-by-sample and are illustrated in Fig 4.14.

4.3.2 Recognition of Active State

The flowcharts for recognition is in the following. In Fig 4.7, we applied our spatial filter to EEG signals (8-15Hz). We extracted the feature by calculating the signal variance and then used a threshold to recognize the active brain states.

Applying Spatial Filter

In Fig 4.8 and Fig 4.9, we take a look at continuous 20 trials form data set I02 (left hand finger imagery). Data set I02 is the most clean data set which has least noise interference. The left column are the first continuous trials recorded at C4 channel and the right column is the data filtered by our maximum contrast filter. In trial 07, the raw data has some noise

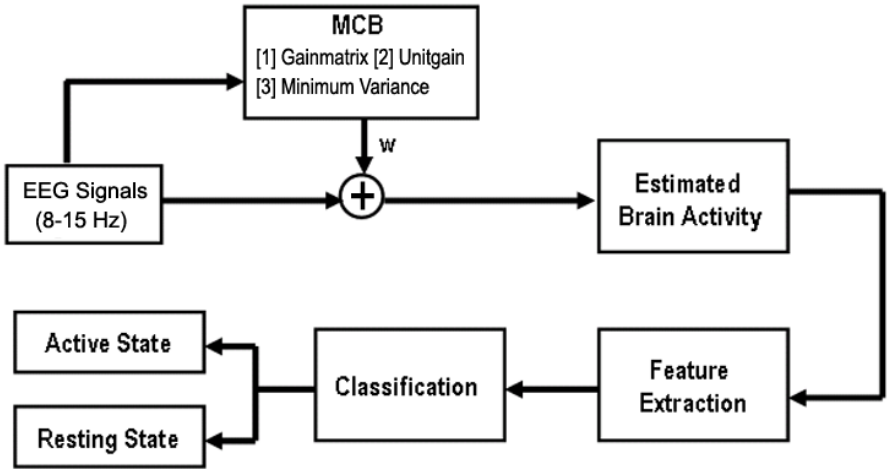


Figure 4.7: **Flowchart - Recognition of Active State.**

at 0-2000 sample points (i.e. 2 seconds before the cue appear). After applying the spatial filter, the noises were eliminated and the variance is stable around 5000-7000 sample points. Similar cases can be found in trial 01, 11 and 19. For trial 03, 07, 10, 16, we can see that the variance of the signals is largely increased.

In contrast to data set I02, we also show the continuous 20 trials from data set R02 which is largely interfered by the environment and the subject. In Fig 4.10 and Fig 4.11, we can see the noise is larger than I02 by observing the signals around 0-3000ms. After applying our filter, we successfully suppressed the noise.

Averaged Variance Ratio

In our method, we extracted our feature by calculated the signal variance. In order to evaluate the changes of the signal, we calculated the variance of the filtered signal by a 500ms-width sliding window (Fig 4.12). Firstly, We calculated the signal variance curve and then divided it by the variance of resting states. We called it the *ratio curve* and we used it as our feature. In Fig 4.13, we illustrates the averaged ratio for each data sets (right hand tasks). We can see that the ratio at active states is at least 3 times larger the resting

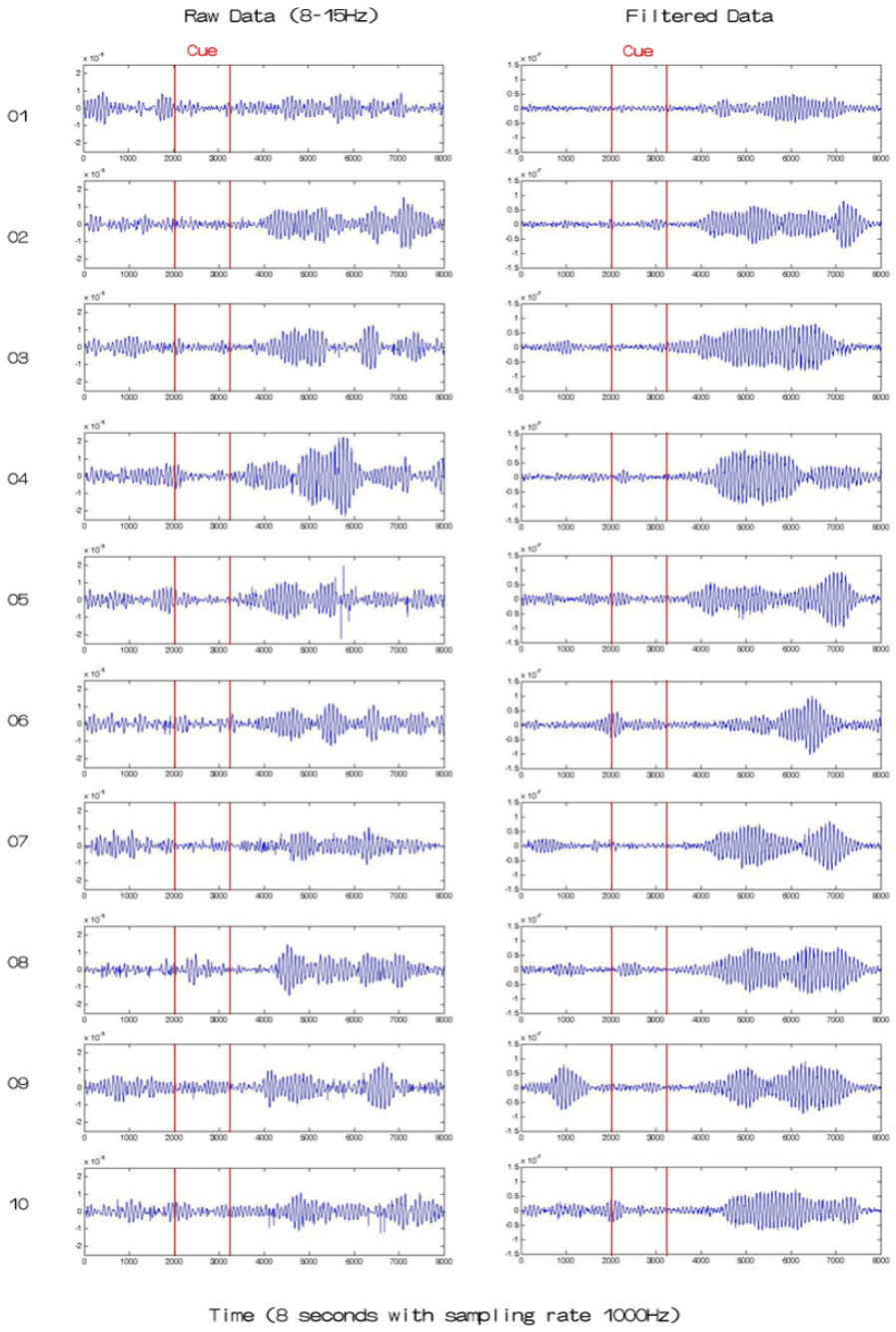


Figure 4.8: **Applying filter to raw data.** The signals are from data set I02 (trail 1-10), left finger imagery trials. The left column is the raw data recorded at channel C4, and the right column is the filtered signal.

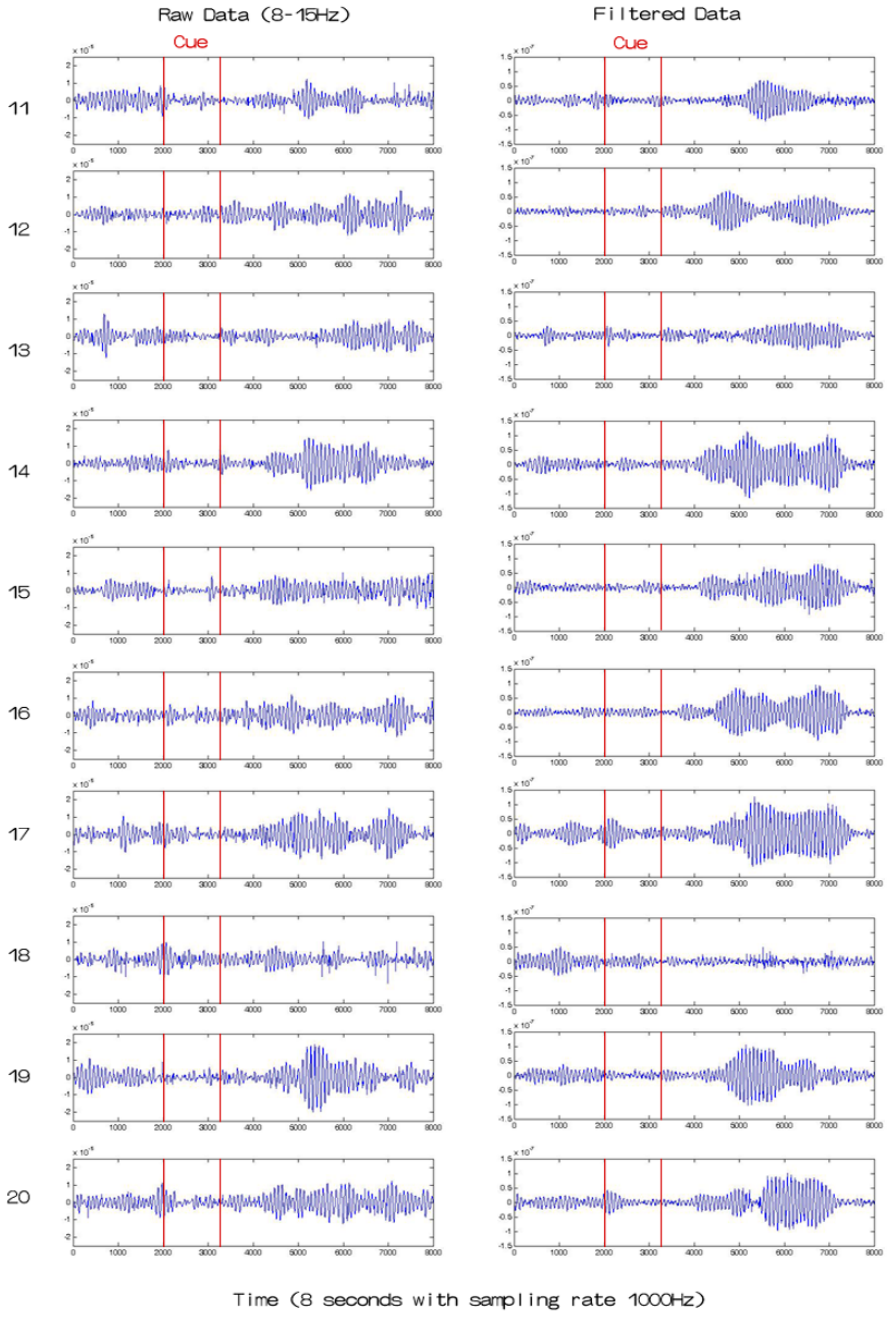


Figure 4.9: **Applying filter to raw data.** The signals are from data set I02 (trial 11-20), left finger imagery trials. The left column is the raw data recorded at channel C4, and the right column is the filtered signal.

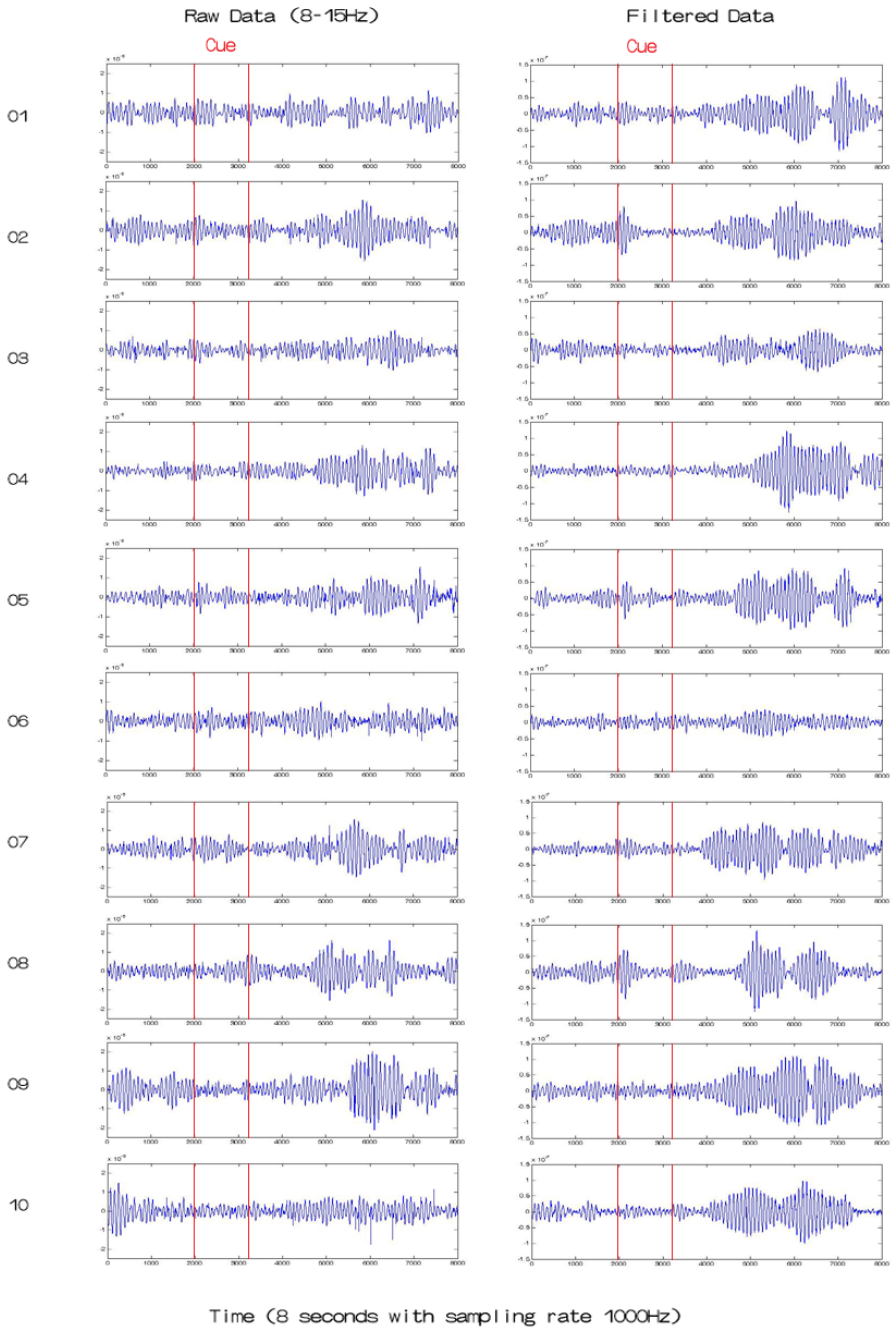


Figure 4.10: **Applying filter to raw data.** The signals are from data set R02 (trail 1-10), left finger imagery trials. The left column is the raw data recorded at channel C4, and the right column is the filtered signal.

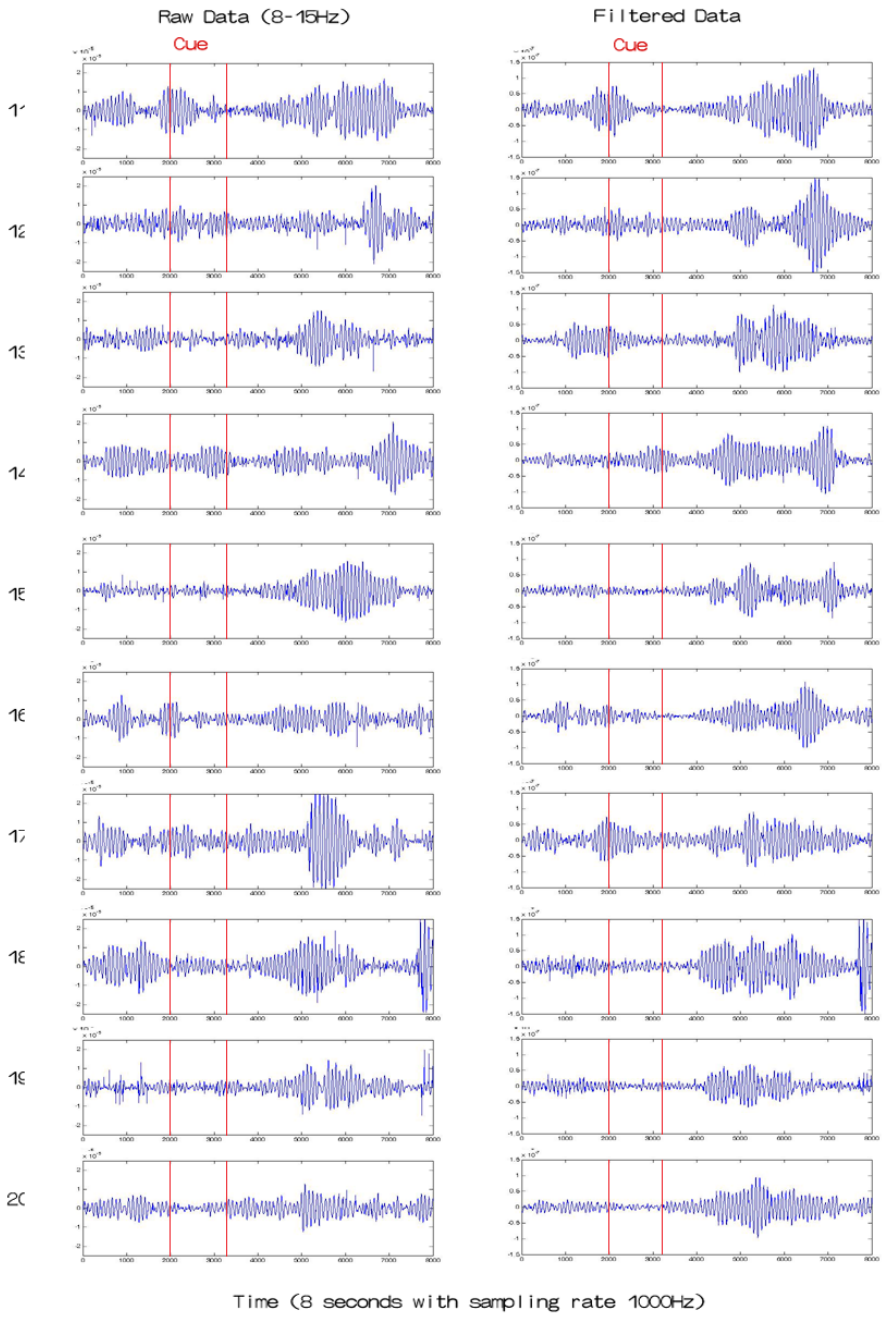


Figure 4.11: **Applying filter to raw data.** The signals are from data set R02 (trial 11-20), left finger imagery trials. The left column is the raw data recorded at channel C4, and the right column is the filtered signal.

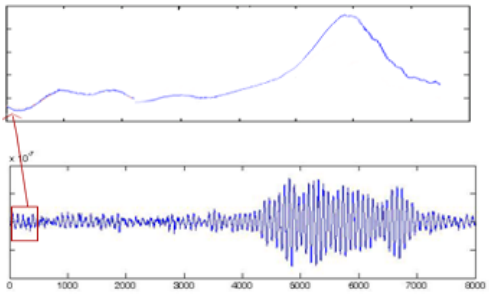
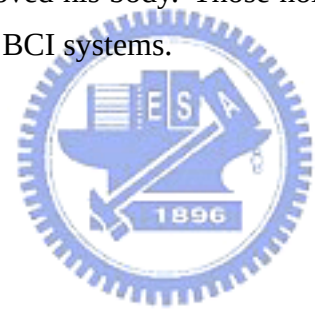


Figure 4.12: **Calculating variance using sliding window.** A sliding window with 500ms width is used to calculate the curve of signal variance. This curve will divide the variance of resting state signal for the ratio curve for classification.

states (blue curve) which is larger than the ratio of raw data (red curve). Besides, there are some noises before the cue in data set R02 and R03. That may be result from the eye movement or the subject moved his body. Those noises are largely suppressed and this is also a practical property for BCI systems.



Evaluating ROC Curves

In Fig 4.14, we used the ratio curve as our feature. While evaluating the performance, a threshold was adjusted and the Eq 4.1 will be calculated. The classification result was in Fig 4.15 and Fig 4.16. The red curve is the result of raw EEG data (8-15Hz) and the blue curve is the result of data applied by our spatial filter. The point marked on the curve is the point which nearest to point (0,1). Note that the point (0,1) means that the classification is perfect because it is exactly correct in ”event period” and no error in ”non-event period”. In our experiment, (TPR, FPR) is (0.7363, 0.2001) for averaged right hand tasks and (0.7658, 0.1983) for averaged left hand tasks. From the results, we conclude that left hand is more effective in discriminating active and control state signals because both TPR ad FRP of left hand tasks are better than right hand tasks. This result can be explained by the researches in [1] which says that ERS is more differentiable in non-dominant hand.

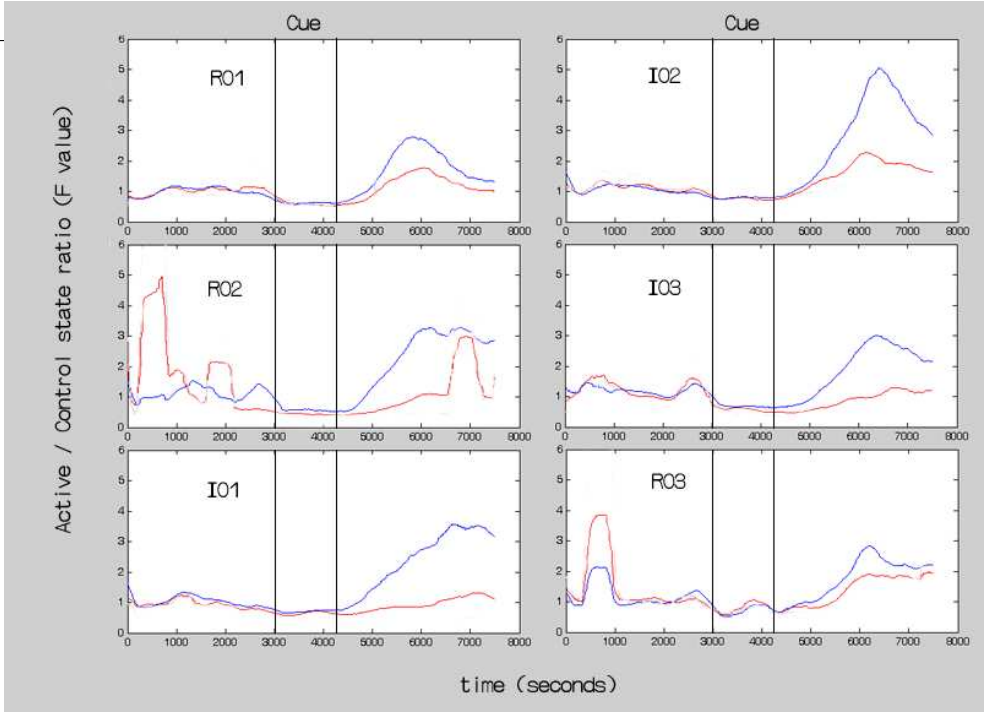


Figure 4.13: **Averaged variance of trials.** In the figure, the averaged ratio for each data set is illustrated. The result is from right finger movement/imagery trials. The x-axis is time and y-axis is the value that the current signal variance divides the signal variance at resting states. The red line is for raw data recorded at channel C3, the blue line is for filtered signals.

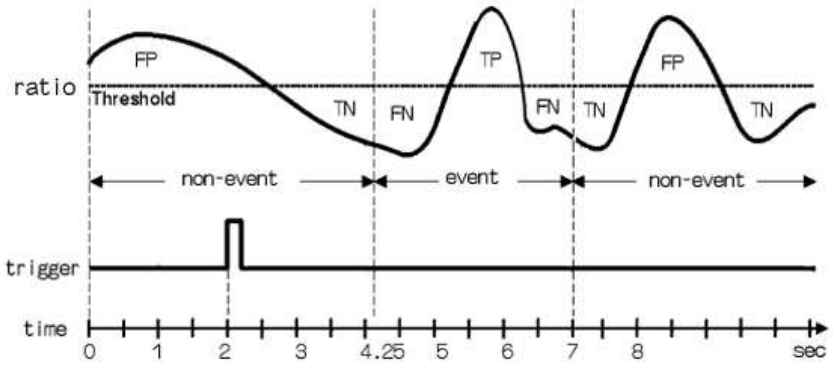


Figure 4.14: **Paradigm for sample-by-sample analysis.** Sample-by-sample evaluates the true positive rate (TPR, denotes the accuracy in ”event period” which is the higher the better) and the false positive rate (FPR, denotes the error in ”non-event period” which is the lower the better). The event period are defined by 2.25-5.0s after the arrow appeared. [26]

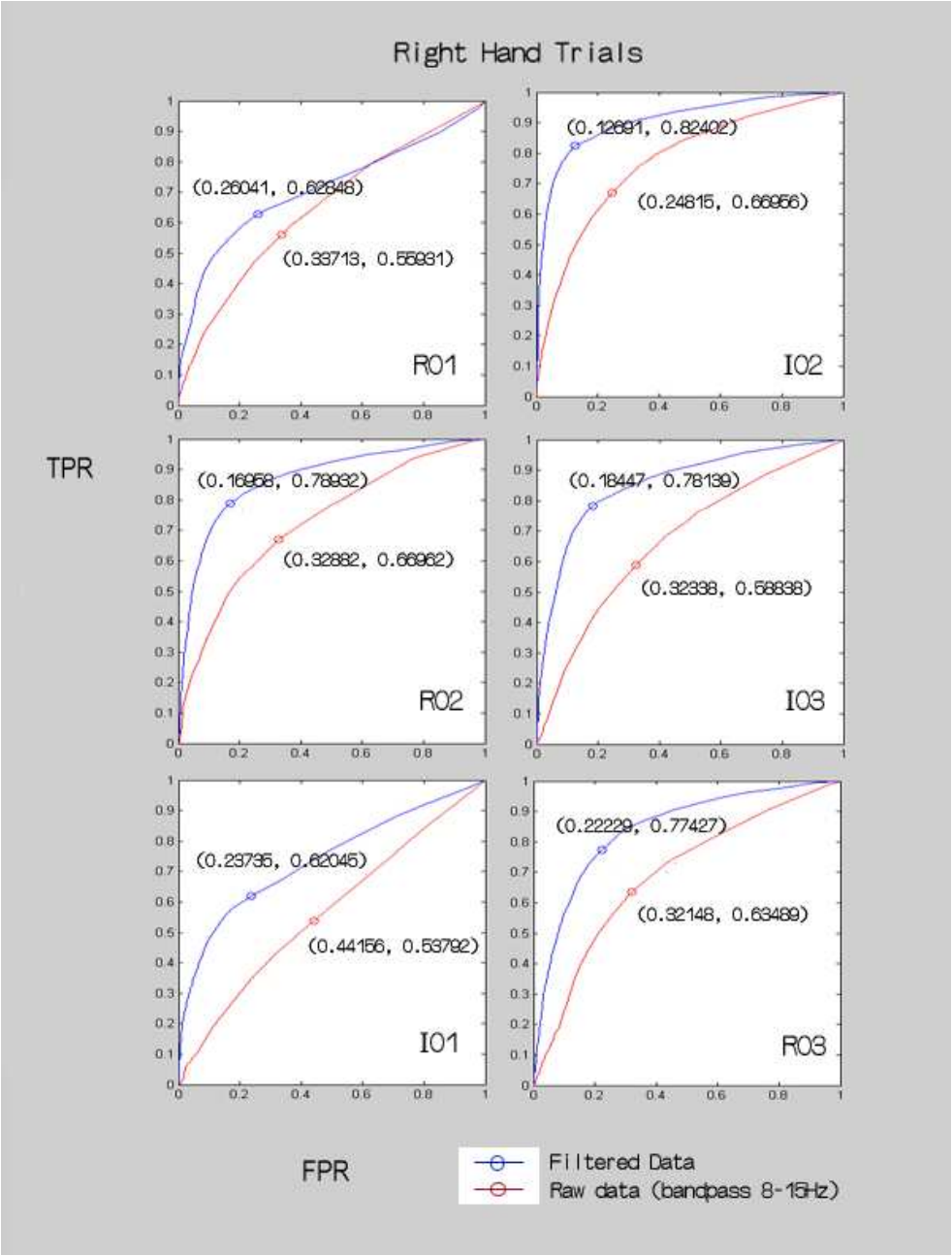


Figure 4.15: **ROC curves for Data sets R01, R02, I01.** Sample-by-sample evaluates the true positive rate (TPR, denotes the accuracy in ”event period” which is the higher the better) and the false positive rate (FPR, denotes the error in ”non-event period” which is the lower the better). The event period are defined by 2.25-5.0s after the arrow appeared.

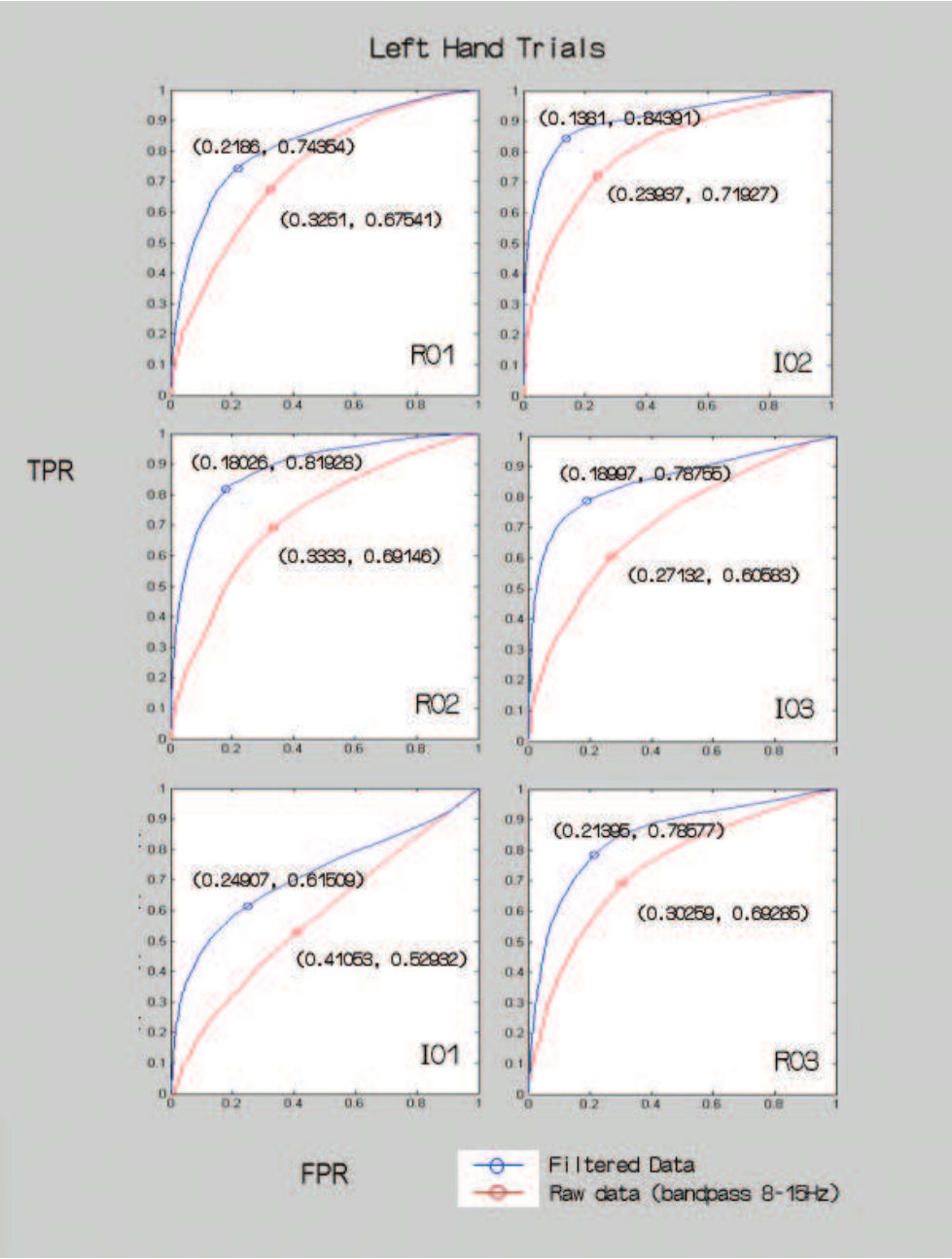


Figure 4.16: **ROC curves for Data sets I02, I03, R03.** Sample-by-sample evaluates the true positive rate (TPR, denotes the accuracy in ”event period” which is the higher the better) and the false positive rate (FPR, denotes the error in ”non-event period” which is the lower the better). The event period are defined by 2.25-5.0s after the arrow appeared.

4.4 Discussions

4.4.1 Imagery and Real Movement Tasks

In our thesis, we use six data sets for analysis. Three of them is imaginary of finger lifting and three of them are real finger lifting. Motor imagery is very similar to real movement and was discussed since a few years ago.

"Imaginary of right and left hand movements result in desynchronization of mu and beta rhythms over the contralateral hand area, very similar to planning and execution of real movements (Neuper and Pfurtscheller, 1999)" [15]

In our experiments, both of them are successful in calculating the spatial filter. The correlation coefficient can prove that it is also true in our data.

4.4.2 Brain Activities on Left/Right Hemisphere

Research in [1] shows that the post-movement power (calculated by ERD/ERS) is more differentiable in non-dominant hand . In our left hand imagery experiments, the averaged power of C4 is obviously larger than C3. In right hand imagery, the averaged power of C3 is larger then C4, but the difference between C3 and C4 is not a large as left hand imagery. The value of the power in Fig 4.2 have been normalized to [0 1] respectively. In Fig 4.17, the maximum power illustrate the interaction between tasks and hemispheres.

These properties and important in analyzing movement imagery tasks was discussed in [1] [15]. Classifying different motor tasks is always an problem concerned with finding a good feature. Knowing more information about the subject is helpful in designing the BCI system for the subject.

4.4.3 Spatial Filter Using Maximum Discriminant Beamformer

Remind about sec 3.3.2, we design the filter pair for classification of left/right hand movement. Fig 4.18 illustrates the topography of the two spatial filter. The left topography is the filter with dipole position located at right motor area and the other is at left motor area.

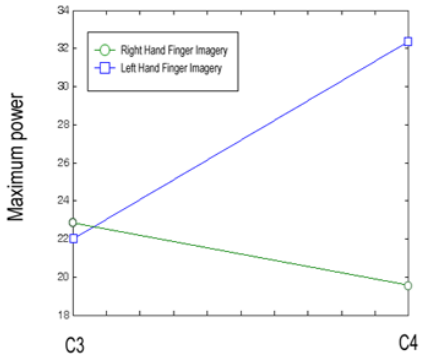


Figure 4.17: **Interaction between tasks and hemispheres.** Wtih left hand finger imagery, the activity is obviously focus at the contralateral hemisphere (C4). When performing dominant hand tasks (the subject is right-handed), cortex activity at both hemisphere is not significantly different.

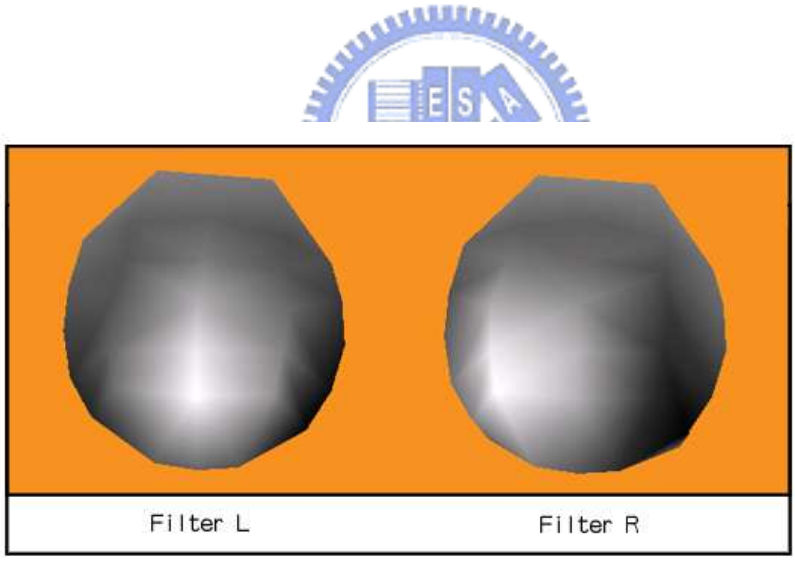


Figure 4.18: **Filter pair designed by maximum discriminant beamformer.** The topography pair is designed by maximum discriminant beamforming method. The idea is to maximize the difference of power at left and right motor activities.

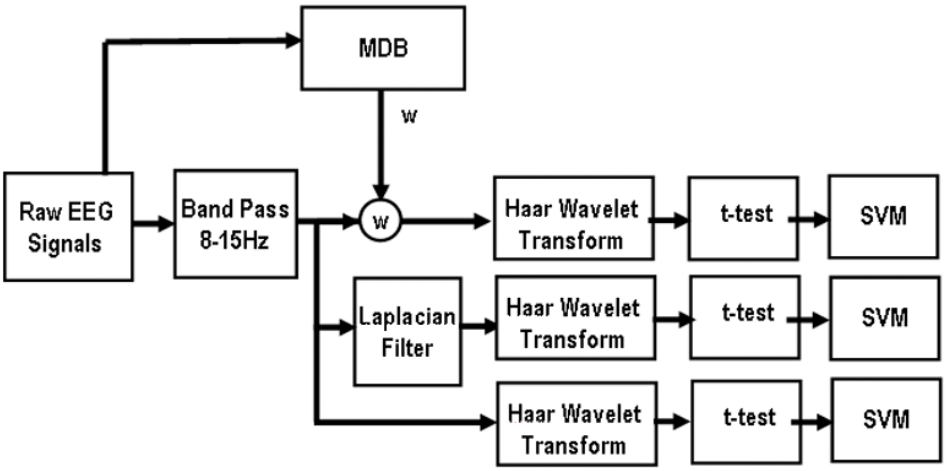


Figure 4.19: **Flowchart(1) for classifying left/right finger lifting tasks.** In this procedure, we compared the effect among raw data and data filtered by Laplacian or MDB.

In our experiments, we used two procedures for different tasks classification. The first one is in Fig 4.19, we used the 8-15Hz raw data (sampling rate 1000) and then use haar wavelet transform as our feature extraction method. T-test is used to find the significant components and then we used SVM (Support Vector Machine) as the classifier. In our experiences, the optimal number of components is around 70 to 100. We compared the effect by applying Laplacian spatial filter and MDB spatial filter. The result is in Fig 4.20, we can see the result of MDB is not very stable. It may because that the filter is sensitive to the parameters selected during the construction of the spatial filter. How to find good parameters and make it stable is an important problem in our future works.

The second procedure is in Fig 4.21, we used five pass-band for feature extraction - delta (0-3Hz), theta (4-8Hz), alpha (9-12Hz), beta (13-30Hz) and gamma bands(30-40Hz). For each pass band, we filtered the signals by MDB filter pairs. Then we calculated the ERD/ERS as our features. In the last stage, we combined the features and we used linear discriminant analysis (LDA) as our classifier. The goal of LDA is to project data into two groups by calculating the optimal projection vector. (The concept is introduced in sec 3.3.2.)

Data Set	Raw Data	Laplacian	MDB
I01	70%	80%	76%
I02	66%	79%	88%
I03	71%	82%	81%

Figure 4.20: **Results of SVM cross validation.**

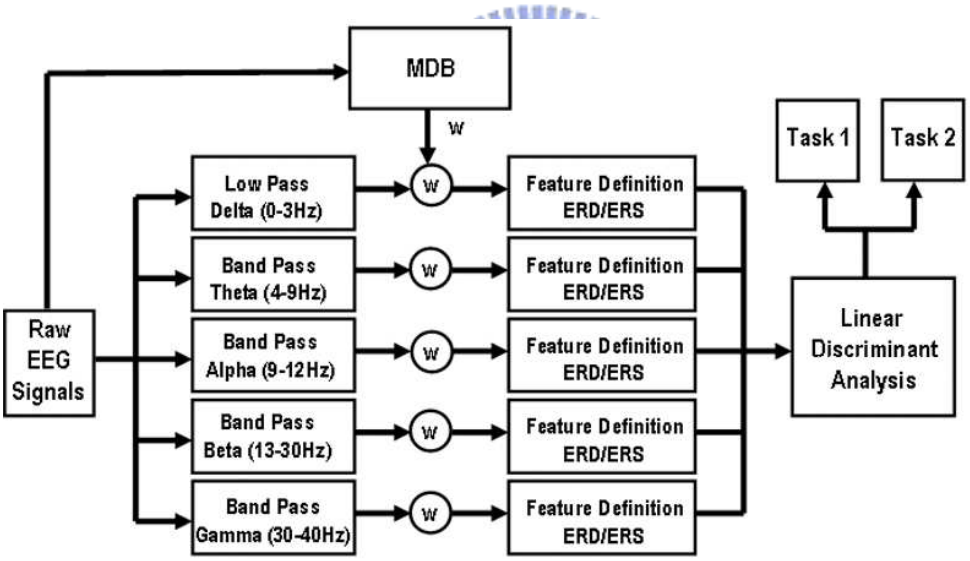


Figure 4.21: **Flowchart(2) for classifying left/right finger lifting tasks.** Five band-pass filter for 0-3Hz, 4-9Hz, 10-13Hz, 14-30Hz, 30-40Hz was applied. After filtered by MDB filter, we used ERD/ERS for feature definition and linear discriminant analysis to classify the two tasks.

In our experiments, data set I02 was chose for tesing. In this data set, there are 80 trials of left finger lifting and 84 trials of right finger lifting. We took 50 trials each (total 100 trials) for training the projection vector and then applied it to the other trials. In Fig 4.22, red points are trials of left finger lifting and blue points are trials of right finger lifting. The value of y-axis is the value after applying the projection vector. Our results showed that the trials for training could be distinguished but the others could not. We thought that it may because the vector is over-fitted to the training sets. Though we can have high classification rate for self-testing, it is still difficult to classify the two tasks. This is also an important problem in our future works.

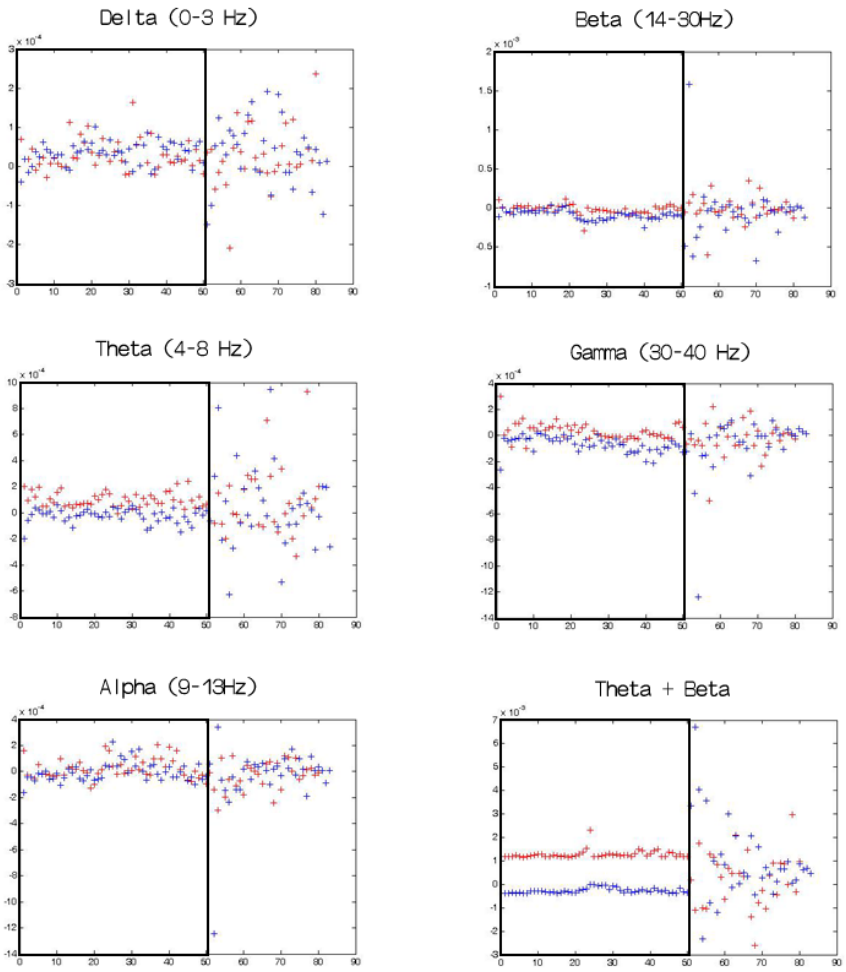


Figure 4.22: **Results of LDA.** Five bands were tested. Red points are left finger lifting trials and blue points are right finger lifting trials. The previous 100 trials are training data (in the black frame) and the rests are for testing.

Chapter 5

Conclusions

In asynchronous BCI systems, the most difficult problem is the determination of active periods. Motor related tasks such as hand, arm or finger movement/imagery are important tasks in BCI systems. In our thesis, we proposed a spatial filter which can efficiently recognizes the active periods of brain sources. In chapter 2, we introduced the background of neuropsychology and neurophysiology. Besides, motor related knowledge is also of significance in current BCI systems. In chapter 3, we used overlapping-sphere model to construct the forward model. Furthermore, we used maximum contrast beamforming method to calculate the optimal dipole orientation which can maximize the variance of active and control state signals.

In chapter 4, we used continuous wavelet transform to preprocess the raw EEG signals. After the dominant frequency band and the ranges of signals were decided, we calculated the spatial filter and further applied it to EEG signals. The filter successfully maximized the variance of the signals at active period and minimized the variance in resting states.

In our method, feature extraction and classification is simple but effective. We use signal variance as features and we only use a threshold to classify the brain states.

The major advantages of our filter is as follows.

1. **Noise suppression**: Our spatial filter can successfully suppress the noises from other brain areas. And the optimal brain source can be acquired by searching the whole brain with beamforming method.

2. **Accommodation**: For each BCI system users, they may have their own spatial filters which are adjusted by their own parameters. In section 4.2.2, we verified the stability of the spatial filter. Users can only measure the sensor position once and apply it to proceeding procedures.

3. **High recognition rate**: In section 4.3, we successfully increased the signal variance which is important in recognition of brain states. In our results, the ROC curves shown the ability of the spatial filter.

However, how to construct asynchronous BCI systems is still a difficult problem, and we may base on the spatial filter for more research in future works.

Bibliography

[1] C. Andrew and G. Pfurtscheller. Lack of bilateral coherence of post-movement central beta oscillations in the human electroencephalogram. *Neuroscience Letters*, 273(2):89–92, 1999.

[2] P. Berg and M. Scherg. A fast method for forward computation of multi-shell spherical head models. *Electroencephalography and Clinical Neurophysiology*, 20, 1994.

[3] E. K. P. Chong and S. H. Zak. *Event-Related Brain Potentials, THE COGNITIVE ERP TEXTBOOK*. Wiley, 2nd edition, 2001.

[4] H. Cox, R. M. Zeskind, and M. M. Owen. Robust adaptive beamforming. *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-35(10):1365–1376, October 1987.

[5] E. A. Curran and M. J. Stokes. Learning to control brain activity: A review of the production and control of eeg components for driving brain-computer interface (bci) systems. *Brain and Cognition*, 51(3):326–336, 2003.

[6] Kevin M. Spencer Emanuel Donchin and Ranjith Wijesinghe. The mental prothesis: Assessing the speed of a p300-based brain-computer interface. *IEEE Transactions on Rehabilitation Engineering*, 8(2):174–179, 2000.

[7] F.H. Lopes da Silva G. Pfurtscheller. Event-related eeg/meg synchronization and de-synchronization: basic principles. *Clinical Neurophysiology*, 110:1842–1857, 1999.

[8] B. H. Jansen, A. Allam, P. Kota, K. Lachance, A. Osho, and K. Sundaresan. An exploratory study of factors affecting single trial p300 detection. *IEEE Transactions on Biomedical Engineering*, 51(6):975–978, 2004.

[9] G. Pfurtscheller J.Kalcher. Discrimination between phase-locked and non-phase-locked event-related eeg activity. *Electroencephalography and clinical Neurophysiology*, 94:381–384, 1995.

[10] H. Krim and M. Viberg. Two decades of array signal processing research. *IEEE Signal Proc. Mag.*, pages 67–94, July 1996.

[11] E. C. Lalor, S. P. Kelly, C. Finucane, R. Burke, R. Smith, R. B. Reilly, and G. McDarby. Steady-state vep-based brain-computer interface control in an immersive 3d gaming environment. *Eurasip Journal on Applied Signal Processing*, 2005(19):3156–3164, 2005.

[12] J.C. Mosher M. X. Huang and R.M. Leahy. A sensor-weighted overlapping-sphere head model and exhaustive head model comprison for meg. *Physics in Medicine and Biology*, 44:423–440, 1999.

[13] D. J. McFarland, L. M. McCane, S. V. David, and J. R. Wolpaw. Spatial filter selection for eeg-based communication. *Electroencephalography and Clinical Neurophysiology*, 103(3):386–394, 1997.

[14] J. Muller-Gerking, G. Pfurtscheller, and H. Flyvbjerg. Designing optimal spatial filters for single-trial eeg classification in a movement task. *Clinical Neurophysiology*, 110(5):787–798, 1999.

[15] C. Neuper and G. Pfurtscheller. Event-related dynamics of cortical rhythms: frequency-specific features and functional correlates. *International Journal of Psychophysiology*, 43(1):41–58, 2001.

[16] C. Neuper and G. Pfurtscheller. Evidence for distinct beta resonance frequencies in human eeg related to specific sensorimotor cortical areas. *Clinical Neurophysiology*, 112(11):2084–2097, 2001.

[17] G. Pfurtscheller, R. Leeb, C. Keinrath, D. Friedman, C. Neuper, C. Guger, and M. Slaterc. Walking from thought. *Brain Research*, 1071(1):145–152, 2006.

[18] G. Pfurtscheller, G. R. Muller, J. Pfurtscheller, H. J. Gerner, and R. Rupp. 'thought' - control of functional electrical stimulation to restore hand grasp in a patient with tetraplegia. *Neuroscience Letters*, 351(1):33–36, 2003.

[19] G. Pfurtscheller and C. Neuper. Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89(7):1123–1134, 2001. Sp. Iss. SI.

[20] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller. Optimal spatial filtering of single trial eeg during imagined hand movement. *IEEE Transactions on Rehabilitation Engineering*, 8(4):441–446, 2000.

[21] David G. Stork Richard O. Duda, Peter E. Hart. *Pattern Classification*. Wiley, 2nd edition, 2001.

[22] S. E. Robinson and J. Vrba. Functional neuroimaging by synthetic aperture magnetometry (SAM). Technical report, CTF Systems Inc., Port Coquitlam, BC, Canada, 1998.

[23] J. J. Ermer S. Baillet, J. C. Mosher and R. M. leahy. Rapid recomputable eeg forward models for realistic head shapes. *Physics in Medicine and Biology*, 46:1256–1281, 2001.

[24] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw. Bci2000: A general-purpose, brain-computer interface (bci) system. *IEEE Transactions on Biomedical Engineering*, 51(6):1034–1043, 2004.

[25] Mingui Sun. An efficient algorithm for computing multishell spherical volume conductor model in eeg dipole source localization. *IEEE Transactions on Biomedical Engineering*, 16:1243–1252, 1997.

[26] G. Townsend, B. Graimann, and G. Pfurtscheller. Continuous eeg classification during motor imagery - simulation of an asynchronous bci. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 12(2):258–265, 2004.

[27] G. Townsend, B. Graimann, and G. Pfurtscheller. A comparison of common spatial patterns with complex band power features in a four-class bci experiment. *IEEE Transactions on Biomedical Engineering*, 53(4):642–651, 2006.

[28] B. D. Van Veen, W. V. Drongelen, M. Yuchtman, and Akifumi Suzuki. Localization of brain electrical actvity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering*, 44:867–880, 1997.

[29] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6):767–791, 2002.

[30] J. R. Wolpaw, D. J. McFarland, T. M. Vaughan, and G. Schalk. The wadsworth center brain-computer interface (bci) research and revelopment program. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):204–207, 2003.