

Fall 2009

Improvement of EEG based brain computer interface by application of tripolar electrodes and independent component analysis

Hongbao Cao
Louisiana Tech University

Follow this and additional works at: <https://digitalcommons.latech.edu/dissertations>



Part of the [Biomedical Engineering and Bioengineering Commons](#)

Recommended Citation

Cao, Hongbao, "" (2009). *Dissertation*. 441.
<https://digitalcommons.latech.edu/dissertations/441>

This Dissertation is brought to you for free and open access by the Graduate School at Louisiana Tech Digital Commons. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of Louisiana Tech Digital Commons. For more information, please contact digitalcommons@latech.edu.

**IMPROVEMENT OF EEG BASED BRAIN COMPUTER
INTERFACE BY APPLICATION OF TRIPOLAR
ELECTRODES AND INDEPENDENT
COMPONENT ANALYSIS**

by

Hongbao Cao, M.S.

A Dissertation Presented in the Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

COLLEGE OF ENGINEERING AND SCIENCE
LOUISIANA TECH UNIVERSITY

November, 2009

UMI Number: 3399267

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3399267

Copyright 2010 by ProQuest LLC.

All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

LOUISIANA TECH UNIVERSITY

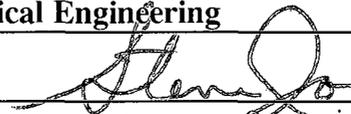
THE GRADUATE SCHOOL

October 16, 2009

Date

We hereby recommend that the dissertation prepared under our supervision
by Hongbao Cao
entitled Improvement of EEG Based Brain Computer Interface by Application of
Tripolar Electrodes and Independent Component Analysis

be accepted in partial fulfillment of the requirements for the Degree of
Doctor of Philosophy in Biomedical Engineering



Supervisor of Dissertation Research



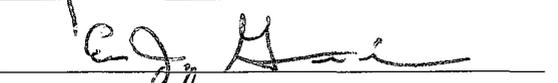
Head of Department

Department

Recommendation concurred in:



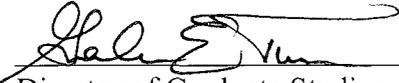




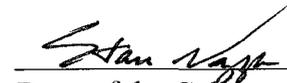


Advisory Committee

Approved:



Director of Graduate Studies



Dean of the College

Approved:



Dean of the Graduate School

APPROVAL FOR SCHOLARLY DISSEMINATION

The author grants to the Prescott Memorial Library of Louisiana Tech University the right to reproduce, by appropriate methods, upon request, any or all portions of this Dissertation. It is understood that "proper request" consists of the agreement, on the part of the requesting party, that said reproduction is for his personal use and that subsequent reproduction will not occur without written approval of the author of this Dissertation. Further, any portions of the Dissertation used in books, papers, and other works must be appropriately referenced to this Dissertation.

Finally, the author of this Dissertation reserves the right to publish freely, in the literature, at any time, any or all portions of this Dissertation.

Author Cao Hongbin
Date 10/19/2009

ABSTRACT

For persons with severe disabilities, a brain computer interface (BCI) may be a viable means of communication, with scalp-recorded electroencephalogram (EEG) being the most common signal employed in the operation of a BCI. Various electrode configurations can be used for EEG recording, one of which was a set of concentric rings that was referred to as a Laplacian electrode. It has been shown that Laplacian EEG could improve classification in EEG recognition, but the complete advantages of this configuration have not been established.

This project included two parts. First, a modeling study was performed using Independent Component Analysis (ICA) to prove that tripolar electrodes could provide better EEG signal for BCI. Next, human experiments were performed to study the application of tripolar electrodes in a BCI model to show that the application of tripolar electrodes and data-segment related parameter selection can improve EEG classification ratio for BCI.

In the first part of work, an improved four-layer anisotropic concentric spherical head computer model was programmed, then four configurations of time-varying dipole signals were used to generate the scalp surface signals that would be obtained with tripolar and disc electrodes. Four important EEG artifacts were tested: eye blinking, cheek movements, jaw movements and talking. Finally, a fast fixed-point algorithm was used for signal-independent component analysis (ICA). The results showed that signals

from tripolar electrodes generated better ICA separation than signals from disc electrodes for EEG signals, suggesting that tripolar electrodes could provide better EEG signal for BCI.

The human experiments were divided into three parts: improvement of the data acquisition system by application of tripolar concentric electrodes and related circuit; development of pre-feature selection algorithm to improve BCI EEG signal classification; and an autoregressive (AR) model and Mahalanobis distance-based linear classifier for BCI classification. In the work, tripolar electrodes and corresponding data acquisition system were developed. Two sets of left/right hand motor imagery EEG signals were acquired. Then the effectiveness of signals from tripolar concentric electrodes and disc electrodes were compared for use as a BCI. The pre-feature selection methods were developed and applied to four data segment-related parameters: the length of the data segment in each trial (LDS), its starting position (SPD), the number of trials (NT) and the AR model order (AR Order). The study showed that, compared to the classification ratio (CR) without parameter selection, the CR was significantly different with an increase by 20% to 30% with proper selection of these data-segment-related parameter values and that the optimum parameter values were subject-dependent, which suggests that the data-segment-related parameters should be individualized when building models for BCI. The experiments also showed that that tripolar concentric electrodes generated significantly higher classification accuracy than disc electrodes.

Keywords: Brain-computer interface (BCI), electroencephalogram (EEG) classification, Laplacian estimation, parameter selection, tripolar electrode.

TABLE OF CONTENTS

ABSTRACT.....	iii
LIST OF TABLES.....	viii
LIST OF FIGURES	ix
ACKNOWLEDGEMENTS.....	xii
CHAPTER 1 INTRODUCTION	1
1.1 Functional Structure and Classification of BCI.....	1
1.1.1 Functional Structure of a BCI.....	2
1.1.2 Classification of BCIs.....	2
1.2 Research Objectives and Contents of the Dissertation	5
1.2.1 Objectives of the Work.....	5
1.2.2 Works Included.....	6
1.3 Organization of the Dissertation.....	7
CHAPTER 2 BACKGROUND	9
2.1 EEG Classification in BCI.....	9
2.1.1 Feature Extraction for BCI.....	9
2.1.2 Classification Algorithms	10
2.2 Tripolar Concentric Ring Electrodes and Laplacian Estimation	10
2.2.1 Laplacian EEG.....	10
2.2.2 Tripolar Concentric Electrodes and Laplacian Estimation.....	11
2.2.3 Tripolar Electrodes vs. Disc Electrodes.....	12
2.3 Independent Component Analysis for EEG in BCI.....	13
CHAPTER 3 A MODEL STUDY.....	15
3.1 Four-Layer Anisotropic Concentric Head Model.....	15
3.1.1 Structure of the Head Model and Its Application.....	15
3.1.2 Numerical Calculation Algorithm for the Head Model.....	16
3.2 Improvement of the Head Model.....	18
3.2.1 Theory of Coordinate Change.....	18
3.2.2 Method for Coordinate Change	19
3.2.3 Database Method for Calculation Speed Improvement.....	20

3.2.4 Disc-Electrode and Ring-Electrode Potentials Generation.....	21
3.3 Independent Component Analysis Model.....	21
3.3.1 Improved Fast Algorithm for ICA	22
3.3.2 Data Sources Simulation.....	23
3.4 ICA Result.....	26
3.4.1 Higher Spatial Sensitivity of Tripolar Electrodes.....	26
3.4.2 Influence of Electrode Number and Source Number on ICA.....	27
3.4.3 Relation Between Independence of Sources and ICA Results	31
3.4.4 Tripolar Electrodes vs. Disc Electrodes for ICA	34
3.4.5 Signal Strength Affects ICA	37
3.5 Conclusions and Discussion	40
3.5.1 Conclusion	40
3.5.2 Further Discussion on Three Interesting Facts of ICA	42
3.5.3 Further Discussion on Four-Layer Head Model and Fast ICA Algorithm	47
 CHAPTER 4 HUMAN EXPERIMENT	 48
4.1 Structure of the BCI.....	48
4.2 EEG Data Acquisition and Signal Pre-Processing	49
4.2.1 Data Acquisition and Hardware Description	50
4.2.2 Data Acquisition Protocol and Software description.....	51
4.2.3 Data Pre-processing	53
4.3 Individualization of Data-Segment-Related Parameters.....	53
4.3.1 Introduction.....	53
4.3.2 Auto Search Algorithm for Parameter Selection	55
4.3.3 Results and Discussion for Data-Segment-Related Parameters Selection.....	57
4.4 Feature Selection for BCI	58
4.4.1 Introduction on Feature Selection for BCI	58
4.4.2 Feature Selection Using an AR Model	59
4.4.3 Feature Selection Using Spectrum Characters.....	61
4.5 EEG Signal Classifier for BCI.....	65
4.5.1 Introduction on EEG Signal Classification for BCI	65
4.5.2 Mahalanobis Distance Based Classifier	68
4.6 BCI Classification Results	69
4.6.1 Influence of Pre-Feature Selection Algorithm.....	69
4.6.2 Influence of the Application of Tripolar Electrodes	72
4.7 Conclusion and Discussion	73
4.7.1 Improved BCI Classification by Pre-Feature Selection.....	73
4.7.2 Improved BCI Classification by Application of Tripolar Electrodes	75
 CHAPTER 5 CONCLUSION AND FUTURE WORK	 79
5.1 Conclusions of the Work	79
5.1.1 Application of Tripolar Electrodes	79

5.1.2 Conclusions for ICA	80
5.1.3 Improved Four-Layer Head Model.....	82
5.1.4 Pre-feature Selection Method for BCI	82
5.2 Future Work	83
5.2.1 Improvement of Tripolar Electrodes.....	83
5.2.2 More on BCI	84
5.2.3 Improvement of Pre-Feature Selection	85
APPENDIX A Matlab Code For ICA and Computer Head Model.....	86
REFERENCES	101

LIST OF TABLES

Table 3.1	ICA result of four dipole sources with different electrodes number	30
Table 3.2	Normalized covariance of the ICA results and the source signals	37
Table 3.3	ICA result of four dipole sources with different numbers of electrodes	38
Table 4.1	Range of each parameter tested	57
Table 4.2	Comparison of EEG bands	64
Table 4.3	CR of the two data sets with/without parameter individualization.....	70
Table 4.4	The subject-dependency of the optimum parameter value	72

LIST OF FIGURES

Figure 1.1	A typical functional structure of a human BCI	2
Figure 2.1	Typical structure of a tripolar electrode	11
Figure 3.1	Structure of layered head model.....	16
Figure 3.2	The positive rotation of each axis.....	18
Figure 3.3.	The spherical coordinate system	20
Figure 3.4	Artifact waveforms of (a) eye blinking, (b) cheek movements, (c) jaw movements and (d) talking	25
Figure 3.5	International Electrode Placement System.....	26
Figure 3.6	Calculated signals from (a) tripolar and (b) disc electrodes with no added noise.	27
Figure 3.7	Four signal dipoles with twenty electrodes. The first row was the original source signals. The second row was the ICA results with the tripolar electrode signals and the third row was the ICA results with the disc electrode signals. Figure 3.7 to Figure 3.10 have the same layout	28
Figure 3.8	Four signal dipoles with ten electrodes	28
Figure 3.9	Four signal dipoles with six electrodes	29
Figure 3.10	Four signal dipoles with four electrodes	29
Figure 3.11	Four signal dipoles with three electrodes	30
Figure 3.12	Relation between ICA results and electrode numbers, the y-axis was the averaged cross-covariance between the ICA separated signals and the original signals	31
Figure 3.13	Five dipole sources with the fourth source = (source 1 + source 2)/2.....	32

Figure 3.14	Five dipole sources with the fourth source = $(\text{source } 1 + \text{source } 2)/2$ and the fifth source = $(\text{source } 2 + 3 \times \text{source } 3)/4$	32
Figure 3.15	Five dipole sources with the fourth source $(2:N) = (\text{source } 1(1:N-1) + \text{source } 2(1:N-1))/2$ and fourth source $(1) = (\text{source } 1(N) + \text{source } 2(N))/2$	33
Figure 3.16	Five dipole sources with two sine waves of same frequency and amplitude, but different time series.....	34
Figure 3.17	ICA Separation results using 10 electrodes with talking artifacts (a) dipole source waves;(e1) Tripolar electrode ICA results; (e2) Disc electrode ICA results.....	35
Figure 3.18	ICA results from electrodes potentials with four artifacts (a) dipole source waves;(x1) ICA results from the tripolar electrode signals; (x2) ICA results from the disc electrode signals. x was from b to e, with respect to the four artifacts: (b) eye blinking, (c) cheek movements, (d) jaw movements and (e) talking. (Vertical axis – arbitrary units, horizontal axis – time in ms.).....	36
Figure 3.19	ICA results with unity amplitude source signals.....	38
Figure 3.20	ICA results with rising cosine wave strengthened 10 times in amplitude	39
Figure 3.21	ICA results with rectangular wave strengthened 10 times in amplitude	39
Figure 4.1	A typical functional structure of a human BCI	48
Figure 4.2	Structure of the work of BCI research.....	49
Figure 4.3	Configuration and dimensions of a tripolar concentric electrode (a), electrode positions (b) and 10/20 International Electrode Placement System(c)	50
Figure 4.4.	Timing diagram of the events during the experimental protocol	51
Figure 4.5	Software structure for the data acquisition.....	52
Figure 4.6	PSD (a) and BP (b) of left/right hand imaged-movement-related EEG....	63
Figure 4.7	Structure of a Linear Discriminant Analysis	66
Figure 4.8	Structure of a Support Vector Machine.....	67

Figure 4.9	CR for Data set 1 and 2 with/without parameter individualization.....	70
Figure 4.10.	The influence of LDS, SPD, Ar Order and NT on CR. The solid traces were from Data Set 1 (tripolar) and the dashed traces were from Data Set 2 (virtual disc).....	71

ACKNOWLEDGEMENTS

This dissertation was dedicated to all of the many people who have shown me support and encouragement. I want to express my sincere gratitude to my advisors, Dr. Jones and Dr. Besio, for their patient guidance, generous help and abundant encouragement. It was my great luck and honor to be their student. Without their directions and suggestions, this dissertation could not be completed.

I owe a lot to my parents for their unconditional love and selfless sacrifice. They trust, respect and support me and try their best to offer me comfortable conditions for living and studying. Also, I want to express my appreciation to all my friends.

CHAPTER 1

INTRODUCTION

1.1 Functional Structure and Classification of BCI

A brain-computer interface (BCI), sometimes called a direct neural interface or a brain-machine interface, was a communication system that was aimed at assisting, augmenting or repairing human cognitive or sensory-motor functions without requiring any peripheral muscular activity (Wolpaw, et al., 2002). For persons with severe disabilities (e.g., spinal cord injury, amyotrophic lateral sclerosis, brainstem stroke, etc.), a brain-computer interface (BCI) may be the only feasible method for communicating with others and for environmental control (Wolpaw, et al., 2000, 2002).

Research on BCIs began in the 1970s at the University of California Los Angeles (UCLA) under a grant from the National Science Foundation followed by a contract from Defense Advanced Research Projects Agency DARPA (Vidal, 1973, 1977). Many experiments on BCI have been conducted since, mostly toward neuroprosthetics applications that aim to restore damaged hearing, sight and movement. Thanks to the remarkable cortical plasticity of the brain, signals from implanted prostheses can, after adaptation, be handled by the brain, like natural sensor or effector channels (Levine, et al., 2000).

1.1.1 Functional Structure of a BCI

A typical functional structure of a human BCI was shown in Figure 1.1. The figure shows that BCI function was divided into three parts: the first was signal acquisition, the second was signal processing and classification/translation and the third was BCI application. The works of this paper was mainly focused on the first two parts of the BCI.

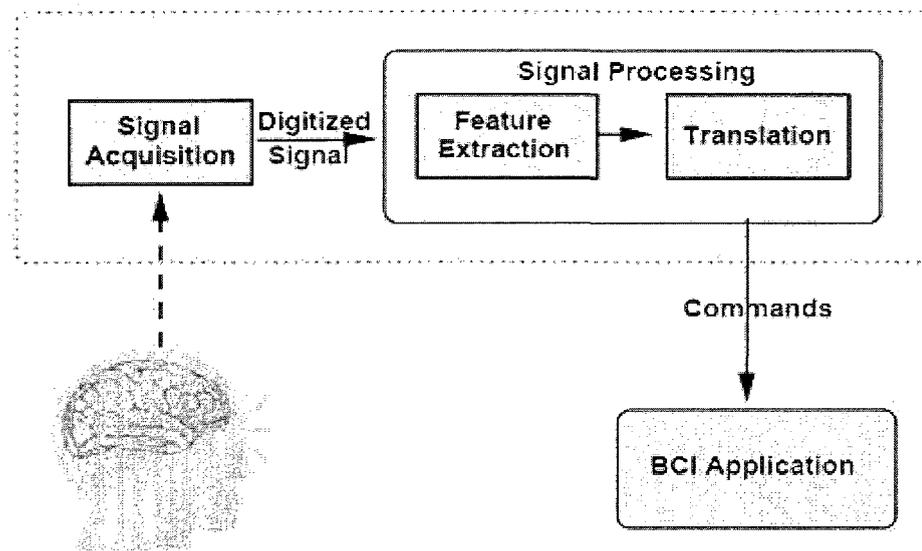


Figure 1.1 A typical functional structure of a human BCI

1.1.2 Classification of BCIs

Human BCIs can be categorized into three classes, according to the positions at which they were implanted: Invasive BCIs, Partial invasive BCIs and Non-invasive BCIs.

The first kind of BCIs was Invasive BCIs. The purpose of this kind of BCI research was to repair the damaged site and provide new functionality to persons with paralysis. During neurosurgery, invasive BCIs were implanted directly into the grey matter of the brain, which was a major component of the central nervous system, consisting of neuronal cell bodies, neuropil (dendrites and both unmyelinated axons and myelinated

axons), glial cells (astroglia and oligodendrocytes) and capillaries. As invasive BCIs rest in the grey matter that was the source of the signal or was near the source of signal, they produce the highest quality signals of BCI devices but the signals become weak or were even lost as a result of scar-tissue buildup as the body reacts to the implant as a foreign object in the brain (Lal, et al., 2005).

In vision science, direct brain implants have been used to treat non-congenital (acquired) blindness. One of the first scientists to construct a working brain interface to restore sight was private researcher William Dobbelle. BCIs focusing on motor neuroprosthetics aim to either restore movement in individuals with paralysis or provide devices to assist them, such as interfaces with computers or robot arms. Researchers at Emory University in Atlanta led by Philip Kennedy and Bakay were the first to install a brain implant in a human that produced signals of high enough quality to simulate movement. Their patient, Johnny Ray, suffered from 'locked-in syndrome' after suffering a brain-stem stroke. Ray's implant was installed in 1998 and he lived long enough to start working with the implant, eventually learning to control a computer cursor (Kennedy and Bakay, 1998).

Tetraplegic Matt Nagle became the first person to control an artificial hand using a BCI in 2005 as part of the first nine-month human trial of Cyberkinetics Neurotechnology's BrainGate chip-implant. Implanted in Nagle's right precentral gyrus (area of the motor cortex for arm movement), the 96-electrode BrainGate implant allowed Nagle to control a robotic arm by thinking about moving his hand as well as a computer cursor, lights and TV (Leigh, et al., 2006).

The second kind of BCIs was partially invasive BCIs. This kind of BCI devices (Serruya and Donoghue, 2003; Hill, et al., 2006) could provide better resolution signals than non-invasive BCIs since they were implanted inside the skull and not affected by the bone tissue which deflects and deforms signals. Also, compared to fully-invasive BCIs, partially invasive BCIs have a lower risk of forming scar-tissue in the brain since they stay out of the brain.

Electrocorticography (ECoG) measures the electrical activity of the brain taken from beneath the skull in a similar way to non-invasive electroencephalography, but the electrodes were embedded in a thin plastic pad that was placed above the cortex, beneath the dura mater (Hill, et al., 2006). Eric Leuthardt and Daniel Moran from Washington University in St Louis first tried ECoG technologies in humans in 2004.

ECoG was a promising intermediate BCI modality because it has higher spatial resolution, better signal-to-noise ratio, wider frequency range and less training requirements than scalp-recorded EEG. It also has lower technical difficulty, lower clinical risk and probably superior long-term stability than intracortical single-neuron recording (Serruya and Donoghue, 2003; Hill, et al., 2006).

The third kind of BCIs was Non-invasive BCIs. This kind of BCIs uses neuroimaging technology-based interfaces that have been developed during the past years. The most common signal employed for this kind of BCIs has been the scalp-recorded electroencephalogram (EEG) (Wolpaw, et al., 2000; Pfurtscheller, et al., 2000). Unfortunately, the EEG lacks high spatial resolution, primarily due to the blurring effects of the volume conductor with disc electrodes. It has also been shown that conventional EEG signals recorded with disc electrodes have reference electrode problems as idealized

references were not available with EEG (Nune, et al., 1994). A common average reference and concentric electrodes have been proposed to resolve the reference electrode problems as discussed by Nunez, since concentric electrodes act like closely spaced bipolar recordings (Nunez, et al., 1994). However, in the common average reference recordings, components present in most of the electrodes, but absent or minimal in the electrode of interest, may appear as “ghost potentials” (Desmedt, Chalklin and Tomberg, 1990).

1.2 Research Objectives and Contents of the Dissertation

1.2.1 Objectives of the Work

In this work, the objective was divided into two parts that focus on the first two parts of a functional BCI (see Figure 1.1). The first part of the objective was to prove that tripolar electrodes could provide a better EEG signal for BCI; the second part of the objective was to prove that the application of tripolar EEG and data-segment related parameter selection could improve the EEG classification ratio for BCI.

To achieve the objective, two parts of the work were done accordingly: Firstly, a numerical modeling study was conducted, applying Independent Component Analysis (ICA) on tripolar EEG signals and disc EEG signals, which were simulated from a concentric four-layer head model to prove that tripolar electrodes could provide better EEG signals for BCI. Secondly, human experiments using a BCI model to show that the application of tripolar EEG and data-segment-related parameter selection could improve the EEG classification ratio for BCI.

1.2.2 Works Included

The work contained in the dissertation included:

1. Development of an improved four-layer anisotropic concentric spherical computer head model.
2. Development of a PCA and ICA model for EEG signal separation.
3. Comparison of the ICA of signals from tripolar electrodes signals and disc electrodes with four important EEG artifacts: eye blinking, cheek movements, jaw movements and talking.
4. Improvement of ICA separation of dipole sources by Laplacian estimation using tripolar concentric electrodes in signal processing.
5. Development of tripolar concentric electrodes and related circuit that provided high spatial resolution and better SNR EEG signal.
6. Design of an EEG signal collection system for brain computer interface (BCI)
7. Development of an autoregressive (AR) model and Mahalanobis distance-based classifier.
8. Development of a pre-feature selection algorithm for BCI.
9. An increase in the classification ratio of left/right hand motor imagery EEG signals up to 84% (we did improved the CR, the average was 78.73 and highest was 84%).

The creative points of the work included:

1. An improved four-layer anisotropic concentric spherical head model that generate tripolar Laplacian estimation and traditional disc electrodes surface EEG signals simultaneously; Generate potentials such that the dipole sources could be placed at any position within the head with the dipole sources' moments oriented in any

direction; Use a newly-established database for the head model that increased the numerical calculation speed by thousands of times.

2. An improved fast calculation algorithm for the ICA model by embedding PCA into it which makes it faster and more stable.
3. Application of tripolar electrodes in BCI, which was proved to generate better signal classification results for BCI.
4. Development of a pre-feature selection algorithm in EEG classification for BCI, which proved to be beneficial for BCI signal classification.

1.3 Organization of the Dissertation

In Chapter 1, a general review of the brain computer interface was given, including the classification and functional structure of BCI. The objective of the dissertation and the creative points of the works were given.

Chapter 2 provides some background for this research, including concepts and recent development of tripolar concentric ring electrodes and Laplacian estimation, independent component analysis for EEG in BCI and methods for EEG classification in BCI.

In Chapter 3, a numerical model, including four-layer head model and ICA model, was given for the test of the application of tripolar electrodes estimated Laplacian EEG in to the ICA. Comparison between tripolar electrodes and traditional disc electrodes were conducted and results were discussed.

Chapter 4 gives the human experiment for BCI. It was the main part of the dissertation. Content in this chapter includes EEG data acquisition and signal pre-processing, pre-feature selection algorithm development, feature selection for BCI (AR

model) and EEG signal classification for BCI. Tripolar Electrodes and traditional disc electrodes were also compared and the results were discussed. Conclusions about the application of tripolar EEG and pre-feature selection on BCI were drawn.

Chapter 5 provides the conclusions of our work and suggestions for future research.

CHAPTER 2

BACKGROUND

2.1 EEG Classification in BCI

To compare the effect of the tripolar electrodes and disc electrodes in the application of BCI, we designed a signal classification system for the BCI, which included signal acquisition, signal pre-feature selection, feature selection and signal classification. A brain-computer interfaces (BCI) could be seen as a pattern recognition system and its performance depends on both the feature extraction algorithm and the classification algorithm employed (Lotte, et al., 2007).

2.1.1 Feature Extraction for BCI

The first step of a pattern recognition was the feature selection, including what features were used, what their properties were and how they were used. Many features of EEG signals have been used in the design of BCIs, such as amplitude values (Kaper, et al., 2004), band powers (BP) (Pfurtscheller, et al., 1997), power spectral density (PSD) values (Millan and Mourino, 2003), autoregressive (AR) and adaptive autoregressive (AAR) parameters (Penny, et al., 2000; Pfurtscheller, et al., 1998), time-frequency features (Wang, Deng and He, 2004) and inverse model-based features (Qin, Ding and He, 2004; Kamousi, Liu and He, 2005; Congedo, Lotte and Lecuyer, 2006). In this work, we used AR model features and BP model features for feature selection in the human BCI experiments.

2.1.2 Classification Algorithms

Many feature extraction methods and classification algorithms have been applied to BCIs. Classification algorithms were divided into five categories: linear classifiers, neural networks, nonlinear Bayesian classifiers, nearest neighbor classifiers and combinations of classifiers (Lotte, et al., 2007). Among those classification algorithms, linear classifiers were probably the most popular for BCI applications. Linear classifiers were discriminant algorithms that use linear functions to distinguish classes. The two main kinds of linear classifiers that have been used in BCI design were linear discriminant analysis (LDA) and support vector machine (SVM). The current work uses the LDA method for EEG classification in the human BCI experiments.

2.2 Tripolar Concentric Ring Electrodes and Laplacian Estimation

2.2.1 Laplacian EEG

Recently, the application of surface Laplacian electrodes to EEG was introduced to help alleviate the blurring effects. Surface Laplacian mapping has been shown to enhance the high spatial frequency components and spatial selectivity of the electrical activity located close to the observation point (He, 1999). The Laplacian was the second spatial derivative of the potentials on the body surface which reduces the blurring effect. The application of the Laplacian method to EEG began with Hjorth (Hjorth, 1975) using a five-point method (FPM). He (He, 1999) performed the surface Laplacian with Hjorth's technique derived from an array of disc electrodes measuring surface potentials. Several other approaches have been shown to perform well, including a) the spline Laplacian algorithm (Perri, Bertrand and Pernier, 1987), b) the ellipsoidal spline Laplacian

algorithm (Law, Nunez and Wijesinghe, 1993), c) realistic Laplacian estimation techniques (Babiloni et al., 1995, 1996) and d) realistic geometry Laplacian algorithms (He, Lian and Li, 2001).

2.2.2 Tripolar Concentric Electrodes and Laplacian Estimation

However, the gains from the above-mentioned application of the Laplacian depend on conventional disc electrodes, which were based on the same technology Hans Burger used in 1924. There has been little effort to improve the electrodes. To our knowledge, Fattorusso and Tilmant (Fattoruss and Tilmant, 1949) were the first to report the use of concentric electrodes. Figure 2.1 shows the typical structure of a tripolar electrode.

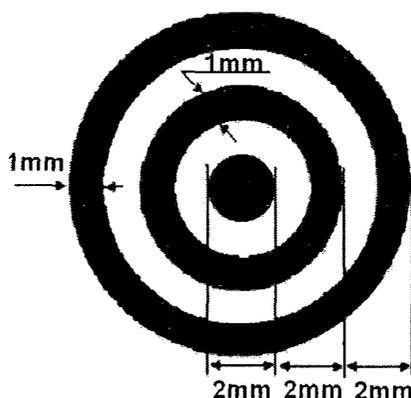


Figure 2.1 Typical structure of a tripolar electrode

Concentric electrodes were symmetrical, alleviating electrode orientation problems (Farino and Cescon, 2001). They act as high-pass spatial filters reducing the low spatial frequencies, accentuating localized activity increasing the spatial selectivity (He, 1999). Concentric electrodes outperform disc electrodes with higher signal-to-noise ratio (SNR), higher spatial selectivity and lower mutual information (MI) which should be beneficial for the field of EEG (He, 1999; Farino and Cescon, 2001; Koka and Besio, 2007). Further,

McFarland, et al. concluded that the common average and the Laplacian derivative yield good performance on EEG classification (McFarland, et al., 2007). Babiloni, et al. demonstrated that surface Laplacian transformation of EEG signals could improve the recognition scores of imagined motor activity (Babiloni, et al., 2000). Besio, et al. developed Laplacian estimation using tripolar electrodes (Besio et al., 2006b) and showed that the tripolar electrode generated significantly higher classification accuracy than disc electrodes (Besio, Cao and Zhou, 2008). Equation (2.1) was the Laplacian estimation that was developed by Besio, et al., where S was the estimated Laplacian signal, P_1 , P_2 , P_3 were the potentials from the outer ring, medium ring and center of the tripolar electrode.

$$S = 16(P_2 - P_3) - (P_1 - P_3) \quad (2.1)$$

2.2.3 Tripolar Electrodes vs. Disc Electrodes

Since the tripolar concentric electrode has significant advantages over disc electrodes, in this paper a comparison of the classification of left/right hand imagery was performed between signals from disc electrodes and tripolar concentric electrodes. Two bipolar signals were acquired from each tripolar concentric electrode and then combined to estimate the Laplacian (Besio et al., 2006 b). An autoregressive (AR) model (Penny, et al., 2000) for feature extraction was built. A Mahalanobis distance-based linear classifier (Mahalanobis, 1936) was used for classification, which was previously established for BCI classification (Cincotti, et al., 2002).

To compare the two electrode configurations fairly, the maximum classification ratio was searched for each data set. An exhaustive search algorithm was utilized to find the

best factors for each subject that generated the highest classification ratio. The results showed that signals from tripolar concentric electrodes generated significantly higher classification ratios than did signals from disc electrodes (Besio, Cao and Zhou, 2008).

2.3 Independent Component Analysis for EEG in BCI

Independent component analysis (ICA) was a computational method for separating a multivariate signal into additive subcomponents, assuming there was mutual statistical independence of the non-Gaussian source signals (James, 2005). To the best of our knowledge, ICA was first applied to encephalography (EEG) by (Makeig, et al., 2002) and was now widely accepted in the EEG research community, most often to detect and remove stereotyped eye, muscle and line noise artifacts (Jung, et al., 1999, 2000). Ventoura, et al. used ICA for reconstructing averaged event-related potentials (ERPs) in the time window of the P600 component, selecting a subset of independent components' projections to the original electrode recording positions (Ventouras, et al., 2004). Basically, Ventoura, et al. used ICA as a filter.

However, ICA also has been used to separate biologically plausible brain sources whose activity patterns were distinctly linked to behavioral phenomena (Delorme, et al., 2006). Many of the biologically plausible sources ICA identifies in EEG data have scalp maps nearly fitting the projection of a single equivalent current dipole (Jung, et al., 2001; Makeig, et al., 2002) and were, therefore, compatible with the projection to the scalp electrodes of synchronous local field activity within a connected patch of cortex.

Fast ICA was an efficient and popular algorithm invented by Aapo Hyvärinen at Helsinki University of Technology (Aapo and Erkki, 1997). The algorithm has cubic convergence speed and does not require parameter adjustment.

CHAPTER 3

A MODEL STUDY

3.1 Four-Layer Anisotropic Concentric Head Model

3.1.1 Structure of the Head Model and Its Application

The computation of the electric potential generated by current density sources in the brain was the so-called EEG forward problem (Vatta, Bruno and Inchingolo, 2005). In order to obtain an accurate solution of the brain tissues, it was necessary to correctly model the shape of the head and the electrical conductivity. A mathematical dipole was commonly used to describe the source. This function provides an adequate description because, if the recorded potentials were caused by an extended source, the error so induced was small (Zhou and Van, 1992). The head was generally described as a volume conductor with piecewise constant conductivity to mimic the different conductivities in different parts of the head. Several versions have been reported: the homogeneous sphere, the three-sphere and four-sphere models, the homogeneous spheroid and the so-called realistic models (De Munck, 1988). A typical layered head model structure was shown as Figure 3.1. For this chapter, a Four-Layer Anisotropic Concentric Head Model was developed for the study (De Munck, 1988).

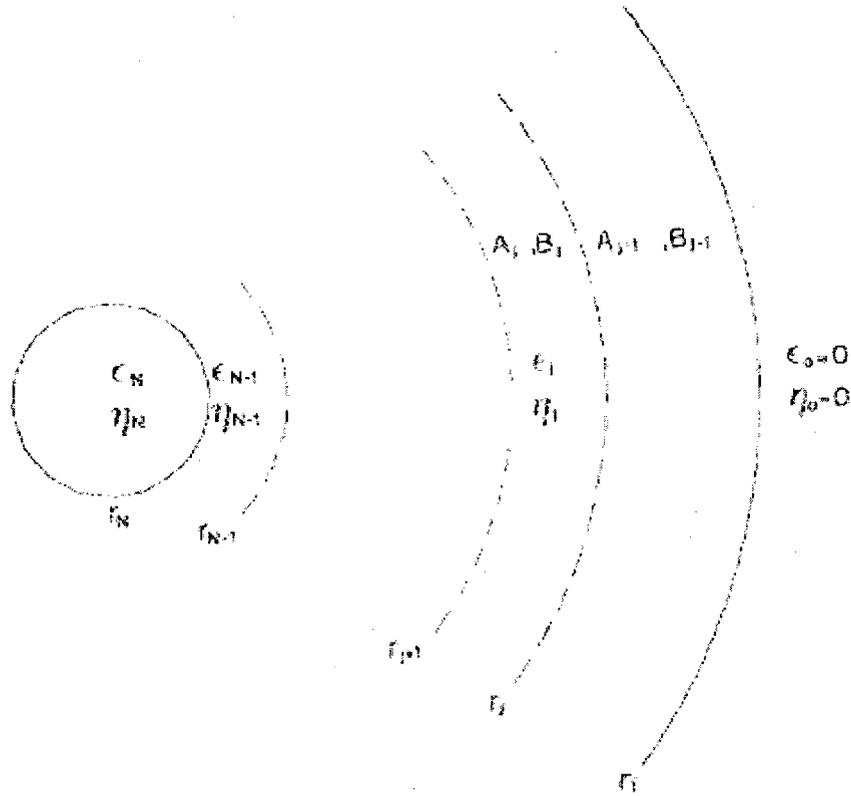


Figure 3.1 Structure of layered head model

3.1.2 Numerical Calculation Algorithm for the Head Model

A modified expression for computing the potential distribution on the exterior surface of a four-layer anisotropic spherical volume conductor, with the dipole lying within the center of innermost sphere was shown in Equations (3.1) (Zhou and Van, 1992):

$$\psi_{dip} = \frac{1}{4} \sum_{n=1}^{\infty} \frac{(2n+1)v_4}{(2v_4+1)\epsilon_4} \frac{2v_1+1}{v_1} \frac{e_{1n}e_{2n}e_{3n}e_{4n}}{D_n r_1} \{M_r r_0^{-1} P_n^0(\cos\theta) + M_\theta v_4^{-1} P_n^1(\cos\theta) \cos\phi\} \quad (3.1)$$

and the parameters were calculated using Equation (3.2) to Equation (3.7).

$$e_{1n} = \left(\frac{r_2}{r_1}\right)^{v_1}, \quad e_{2n} = \left(\frac{r_3}{r_2}\right)^{v_2}, \quad e_{3n} = \left(\frac{r_4}{r_3}\right)^{v_3}, \quad e_{4n} = \left(\frac{r_4}{r_3}\right)^{v_4} \quad (3.2)$$

$$\begin{aligned}
Dn = & \left\{ d_{11}(e_{1n}e_{2n}e_{3n})^2 \frac{r_4}{r_1} + d_{12}(e_{1n}e_{3n})^2 \frac{r_4 r_2}{r_3 r_1} \right. \\
& \left. + d_{13}(e_{1n}e_{2n})^2 \frac{r_3}{r_1} + d_{14}e_{3n}^2 \frac{r_2}{r_1} \right\} \frac{v_1 + 1}{v_1} \\
& + d_{21}(e_{1n}e_{3n})^2 \frac{r_4}{r_2} + d_{22}e_{3n}^2 \frac{r_4}{r_3} \\
& + d_{23}e_{2n}^2 \frac{r_3}{r_2} + d_{24}
\end{aligned} \tag{3.3}$$

and,

$$\begin{aligned}
d_{11} &= c_4(2,1)c_3(1,1)c_2(1,1) \\
d_{12} &= c_4(2,1)c_3(1,2)c_2(2,1) \\
d_{13} &= c_4(2,2)c_3(2,1)c_2(1,1) \\
d_{14} &= c_4(2,2)c_3(2,2)c_2(2,1) \\
d_{21} &= c_4(2,1)c_3(1,1)c_2(1,2) \\
d_{22} &= c_4(2,1)c_3(1,2)c_2(2,2) \\
d_{23} &= c_4(2,2)c_3(2,1)c_2(1,2) \\
d_{24} &= c_4(2,2)c_3(2,2)c_2(2,2)
\end{aligned} \tag{3.4}$$

$$\begin{aligned}
c_j(1,1) &= \frac{\varepsilon_{j-1}v_{j-1} + \varepsilon_j(v_j + 1)}{(2v_j + 1)\varepsilon_j} \\
c_j(1,2) &= \frac{-\varepsilon_{j-1}(v_{j-1} + 1) + \varepsilon_j(v_j + 1)}{(2v_j + 1)\varepsilon_j} \\
c_j(2,1) &= \frac{-\varepsilon_{j-1}v_{j-1} + \varepsilon_j v_j}{(2v_j + 1)\varepsilon_j} \\
c_j(2,2) &= \frac{\varepsilon_{j-1}(v_{j-1} + 1) + \varepsilon_j v_j}{(2v_j + 1)\varepsilon_j}
\end{aligned} \tag{3.5}$$

for $j=2,3,4$,

$$P_n^1(\cos \theta) = i \frac{n+1}{\pi} \int_0^\pi (\cos \theta + i \sin \theta)^n \cos t dt \tag{3.6}$$

$$P_n^0(\cos \theta) = \frac{1}{\pi} \int_0^\pi (\cos \theta + i \sin \theta \cos t)^n dt . \tag{3.7}$$

See Appendix A, function FixPointAlgor() for Matlab programs that implement the algorithm.

3.2 Improvement of the Head Model

The algorithm proposed above has two major problems. First, it calculates the potential generated by a unit dipole only if the moment of the dipole was on a positive z-axis direction. However, we need to calculate multi dipoles placed at any position within the innermost layer, with arbitrary dipole moment directions. Second, the calculation speed was really slowing since a large number of integrations were needed, as was shown in Equations (3.6) and (3.7). In order to solve these problems, the Coordinate Rotating Method and Database Method were introduced into the algorithm.

3.2.1 Theory of Coordinate Change

For 3D rotation, there was a 3×3 matrix for rotation about each principal axis. The direction of positive rotation was determined by the right-hand rule, as shown in Figure 3.2.

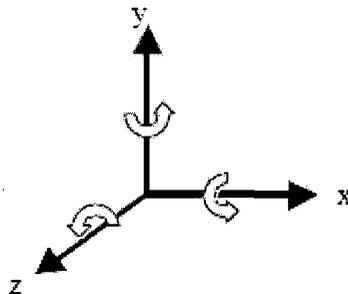


Figure 3.2 The positive rotation of each axis

Let $[p_x, p_y, p_z]$ and $[p'_x, p'_y, p'_z]$ denote the coordinates of a point in the original coordinate system and in the coordinates system after rotation, respectively. The relation between these two coordinates can be expressed by Equation 3.8.

$$\begin{aligned}
 \text{Z-axis: } \begin{bmatrix} p'_x \\ p'_y \\ p'_z \end{bmatrix} &= \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \\
 \text{X-axis: } \begin{bmatrix} p'_x \\ p'_y \\ p'_z \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \\
 \text{Y-axis: } \begin{bmatrix} p'_x \\ p'_y \\ p'_z \end{bmatrix} &= \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}.
 \end{aligned} \tag{3.8}$$

where θ was the counterclockwise angle of rotation about the given axis. These Equations allow the dipole to be rotated to any arbitrary angle.

3.2.2 Method for Coordinate Change

As shown in Figure 3.3, the moment of the dipole was pointing to a location point on the brain surface with the spherical coordinates (R, θ, ϕ) . It was a straightforward exercise to easily obtain the corresponding Cartesian coordinates (x, y, z) . In order to rotate the z-axis to this electrode and realize the coordinate calculation, the coordinate system was rotated in two steps. The first was a rotation of $+\theta$ about the z-axis. The second was a rotation of $+\phi$ about the new y-axis. The center of the electrode was placed on the z-axis. These rotations were repeated for each electrode and the potential of each electrode was then calculated under the new coordinate system. See Appendix A, function ChangeCoordinates for Matlab programs in realizing the algorithm. Note: In Matlab, the rotation function was $(X,Y,Z) = \text{SPH2CART}(TH,PHI,R)$, where TH was the

counterclockwise angle in the x-y plane measured from the positive x axis. PHI was the elevation angle from the x-y plane, which differs from ϕ in Figure 3.3.

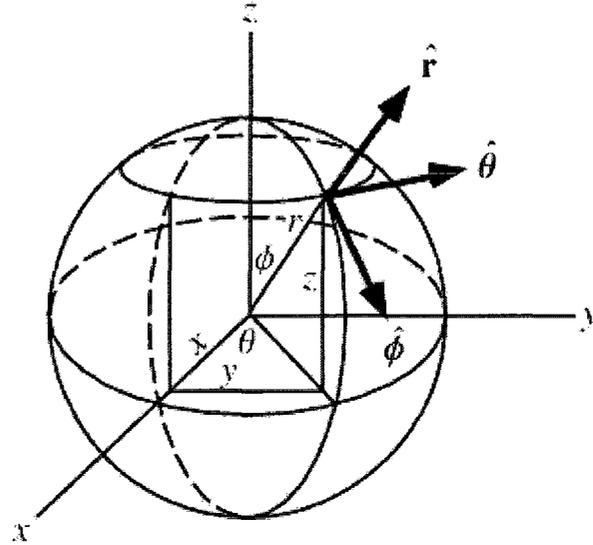


Figure 3.3. The spherical coordinate system

3.2.3 Database Method for Calculation Speed Improvement

As was shown in Equations (3.1; 3.6) and (3.7), to calculate a dipole's potential on each point of the electrodes (the potential of the electrode rings were calculated by averaging N point potential on the ring), $2N$ symbol integrations need to be calculated. Consider N_1 points on each electrode to be computed, with N_2 electrodes and N_3 dipoles, the total number of integrations will be $2NN_1N_2N_3 \approx 7200$. This number of integrations will be computationally expensive. For example, if we calculate the potentials for a single dipole on six electrodes, several hours were required to reach the results with satisfactory precision.

However, it was noticed that in Equations (3.1), for a given calculation precision, N was limited and, for each n in Equations (3.6) and (3.7), the symbol integration results

will be same. Thus, a database that contains high enough order ($n \geq N$) integration results could be built and used repeatedly in the algorithm instead of calculating the integration each time. Compared to the original method, the computation, after using the database method, decreased the computation time by more than three orders of magnitude, making the computation time effectively negligible.

3.2.4 Disc-Electrode and Ring-Electrode Potentials Generation

For tripolar ring electrodes, N (usually set at 50) points of potential on each ring were calculated and the average of those potential values was taken as the potential on the ring. The Laplacian estimation was performed using Equation (3.9) to get the final tripolar signal:

$$S_t = 16(P_2 - P_3) - (P_1 - P_3), \quad (3.9)$$

where S_t was the tripolar signal and P_1 , P_2 and P_3 were the signals from the outer ring, middle ring and center disc, respectively.

For disc electrodes, the potentials of the three rings were averaged using Equations (3.10) to get the virtual disc electrodes signal S_d :

$$S_d = (P_1 + P_2 + P_3)/3 \quad (3.10)$$

3.3 Independent Component Analysis Model

In the work of Chapter 3, a fast fixed-point ICA algorithm was developed and typical source signals were simulated as dipole moments. Then, potentials from those dipole sources were collected using tripolar electrodes and disc electrodes, respectively. After that, ICA was applied using signals from tripolar and disc electrodes.

3.3.1 Improved Fast Algorithm for ICA

The improved fast fixed point algorithm (Aapo and Erkki, 1997) used in this study has cubic convergence speed, has no learning rate or other adjustable parameters and calculates the components of both the negative kurtosis and the positive kurtosis. In addition, it solved the problem present in the original algorithm that a blind number of sources must be set at the beginning of the ICA. The following gives its algorithm:

1. Original data transform using equation (3.11):

$$x = Mv \quad (3.11)$$

where $x = (x_1, x_2, \dots, x_n)^T$ was the transformed signal which will be used as the input of the ICA. Its elements, x_i , were mutually uncorrelated and all have unit variance, which means its correlation matrix equals unity: $E\{xx^T\} = I$. The vector $v = (v_1, v_2, \dots, v_n)^T$ was the observed signals from each electrode. v_i was the signal from the i^{th} electrode. M was the transform matrix, defined as Equation (3.12):

$$M = DP^T, \quad (3.12)$$

where P was the vector of the principal components of v , which was an orthonormal transformation matrix and D was calculated using Equation (3.13):

$$D = \text{diag}(D_1, D_2, \dots, D_n) = 1/\sqrt{\text{cov}(y)}, \quad (3.13)$$

$$y = P^T v, \quad (3.14)$$

where y in Equation (3.14) was the projected vectors, the covariance matrix of which was a diagonal matrix and D was a diagonal matrix.

2. Independent component separation:

The independent components can be expressed by Equation (3.15):

$$\begin{cases} S = B^T x \\ s_i = (b_i)^T x \end{cases} \quad (3.15)$$

where $S = \{s_1, \dots, s_i, \dots, s_n\}$ was the independent component and $B = \{b_1, \dots, b_i, \dots, b_n\}$ was its corresponding transform matrix, which can be obtained by the following steps:

- i) Take a random initial vector, $B^0 = \{b_1^0, b_2^0, \dots, b_n^0\}$, with a norm of 1. Let $k=1$;
- ii) Let $b_i^k = E\{x(b_i^{k-1T} x)^3\} - 3b_i^{k-1}$. where $E\{*\}$ was the expectation, which could be estimated from a large sample;
- iii) Orthogonalize the projection $b_i^k = b_i^k - \overline{B} \overline{B}^T b_i^k$, where $\overline{B} = (b_1, b_2, \dots, b_{i-1})$ were the transform matrix previously found;
- iv) Divide b_i^k by its norm;
- v) If $|b_i^{kT} b_i^{k-1}|$ was not close enough to 1, let $k = k + 1$ and go back to step 2. Otherwise, output the vector $b_i = b_i^k \cdot b_i$, which was the transform vector that was found for the i^{th} independent component.

The transform matrix B found by this algorithm was an orthogonal unit matrix, which means $BB^T = I$. The algorithm will separate, as much as possible, independent sources as long as the number of collected signal channels (electrodes) was not less than the number of sources.

3.3.2 Data Sources Simulation

The first step was dipole signal sources simulation. Sine waves with different frequency, rectangular signals, rising cosine signals and Gaussian white noise signals, were used for the dipole sources, where sine waves were sub-Gaussian signals and rectangular signals and rising cosine signals were super-Gaussian signals. The signals were generated with different frequencies (from 1 to 100 Hz) that covered the frequency range of EEG signals with typical noises. The sampling frequency for each signal was set

at 200 Hz, which was the sample rate for EEG signals that were obtained under clinically real conditions (Kees, et al., 1994).

The second step was typical EEG artifact sources simulation. Four artifacts were recorded from human experiments (Kakkeri, 2005): eye blinking, cheek movements, jaw movements and talking. Their waveforms were used for the simulation of EEG artifact with the amplitude modified according to the SNR in the human experiments condition:

$$snr_t = (13, 18, 23, 28); \quad (3.16)$$

$$snr_d = (-28, -17, -13, -7); \quad (3.17)$$

$$SNR(dB) = 20 \log_{10} \left(\frac{A_{signal}}{A_{noise}} \right). \quad (3.18)$$

where snr_t and snr_d were the SNR of the tripolar and disc electrodes respectively, which have different SNR in the real condition. A_{signal} and A_{noise} were the amplitudes of the signal and noise, respectively. Figure 3.4 shows the waveforms of those signals.

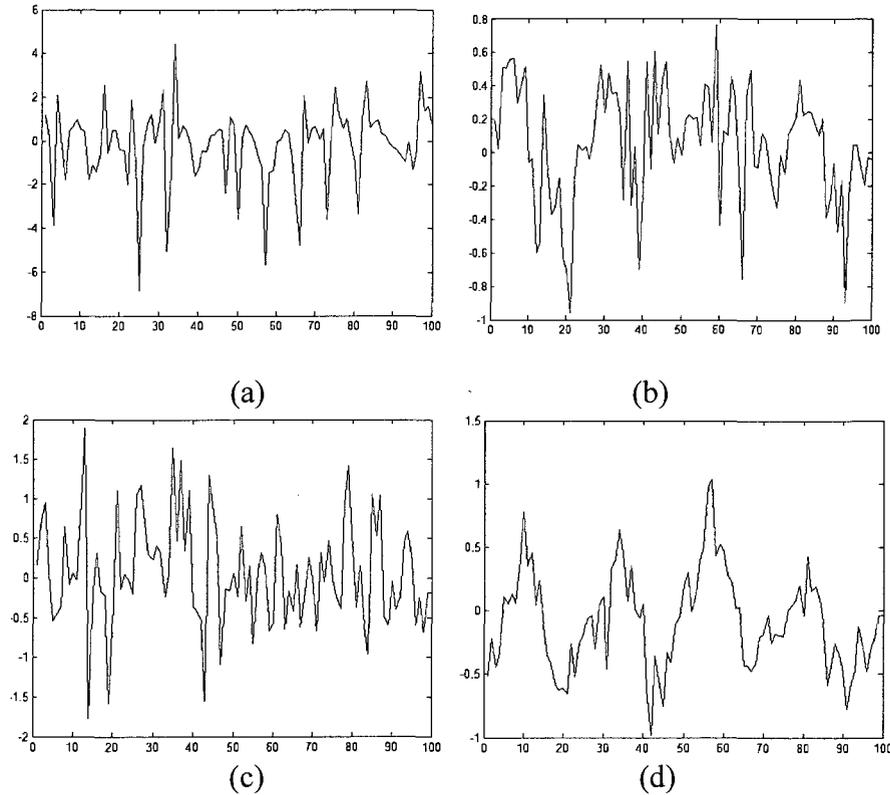


Figure 3.4 Artifact waveforms of (a) eye blinking, (b) cheek movements, (c) jaw movements and (d) talking

The amplitude modification was processed as follows: in the four layer head model, suppose the amplitude of the EEG potential generated by the dipole sources at an electrode was A_{signal} and the SNR for given kind of noise was snr . From the definition of SNR (Equation (3.18)), the amplitude of the noise A_{noise} was calculated use Equation (3.19).

$$A_{noise} = \frac{A_{signal}}{10^{snr/20}}. \quad (3.19)$$

The third step was electrode distribution and number of electrodes design. For the model, the electrodes can be at any position on the surface of the scalp and the number of electrodes should not be less than the number of sources. In real conditions, the number

of electrodes should be more than the desired independent components. The five electrodes in the model were placed at C3, C4, Cz, T3, though T4 (see Figure 3.5) and other positions where a number of electrodes from 4 to 10 were tested. Electrodes for the artifacts recorded by (Kakkeri, 2005) were placed at C3, C4 and Cz.

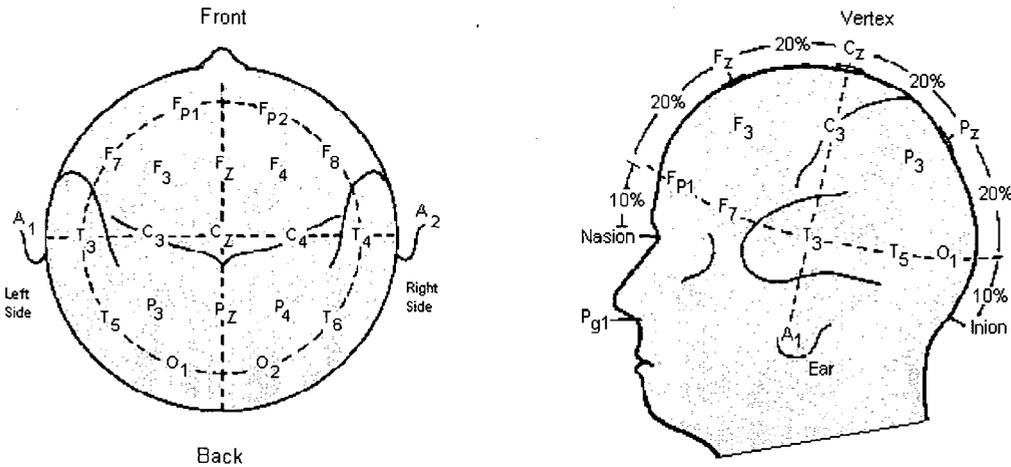


Figure 3.5 International Electrode Placement System

3.4 ICA Result

3.4.1 Higher Spatial Sensitivity of Tripolar Electrodes

Figure 3.6 shows the potentials produced by a vertically oriented unit dipole located at (58,0,0) mm. Those potentials were recorded by a disc and a tripolar electrode located at the surface of the sphere at different angular positions from $\phi = 0$ ("north pole" of the sphere, (75, 0, 0)) to $\phi = \pi$ ("south pole", (75, 0, π)). Because the signal from the tripolar electrode was a difference of potentials, the magnitude was approximately 2 to 3 orders of magnitude below that of the disc electrode. And, due to the high spatial sensitivity of the tripolar electrodes, it gave better ICA results, which was shown later.

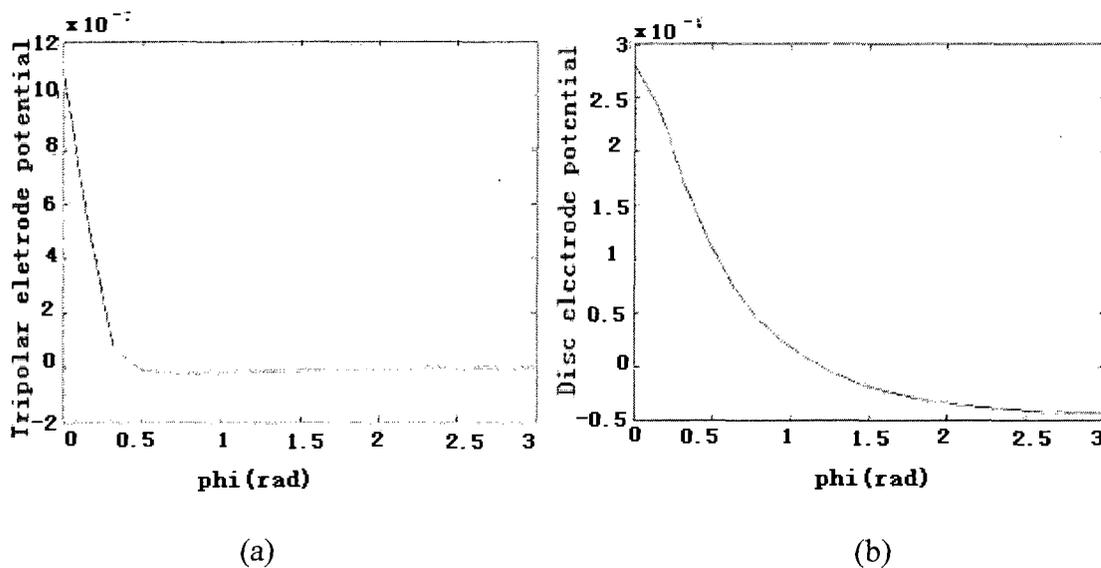


Figure 3.6 Calculated signals from (a) tripolar and (b) disc electrodes with no added noise.

3.4.2 Influence of Electrode Number and Source Number on ICA

Experiments show that the number of electrodes should be no less than the number of source dipoles to fully recover the independent sources. On the other hand, many more electrodes than sources will not improve the ICA results much, which could be seen from Figure 3.12. Figure 3.7 through Figure 3.11 shows the ICA results when there were four dipole sources (ring cosine, rectangular, sine and Gaussian white noise) with 20, 10, 6, 5, 4 and 3 electrodes that were distributed evenly along Cz-C3-T3-A1 line, as shown in Figure 3.5. The ICA were taken with artifact noise added to the electrodes, with SNR of 28 dB and -7 dB, respectively. Table 3.1 shows ICA results of four dipole sources with different numbers of electrodes, in which the 'ICA Cov' states the normalized covariance of the ICA results and the original signal. The closer this value was to 1, the better was the ICA result.

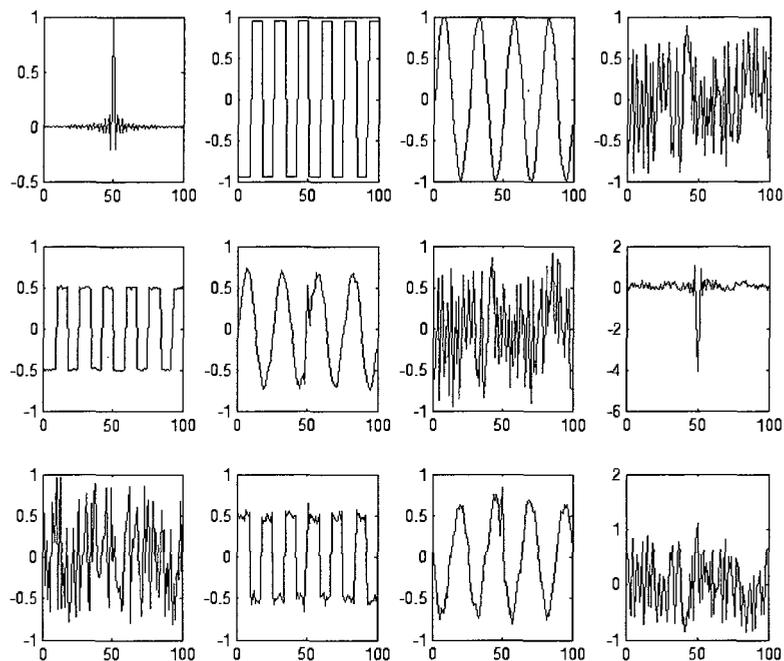


Figure 3.7 Four signal dipoles with twenty electrodes. The first row was the original source signals. The second row was the ICA results with the tripolar electrode signals and the third row was the ICA results with the disc electrode signals. Figure 3.7 to Figure 3.10 have the same layout

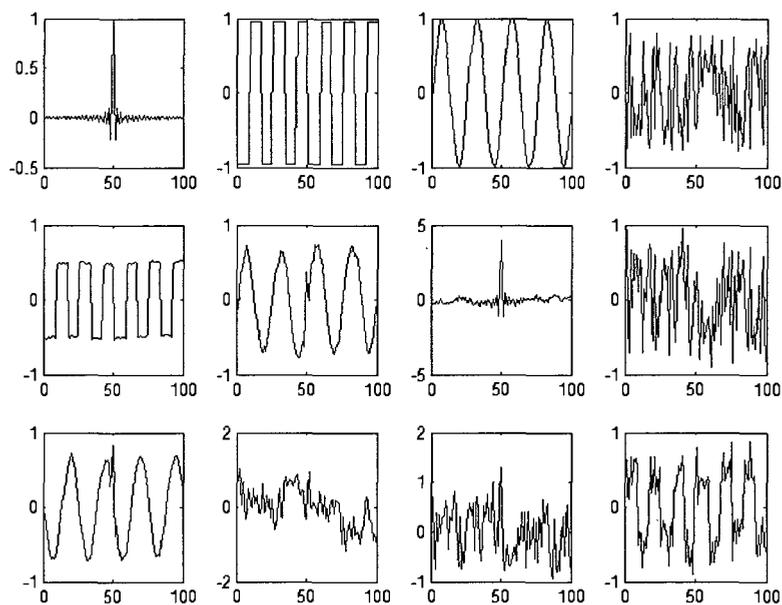


Figure 3.8 Four signal dipoles with ten electrodes

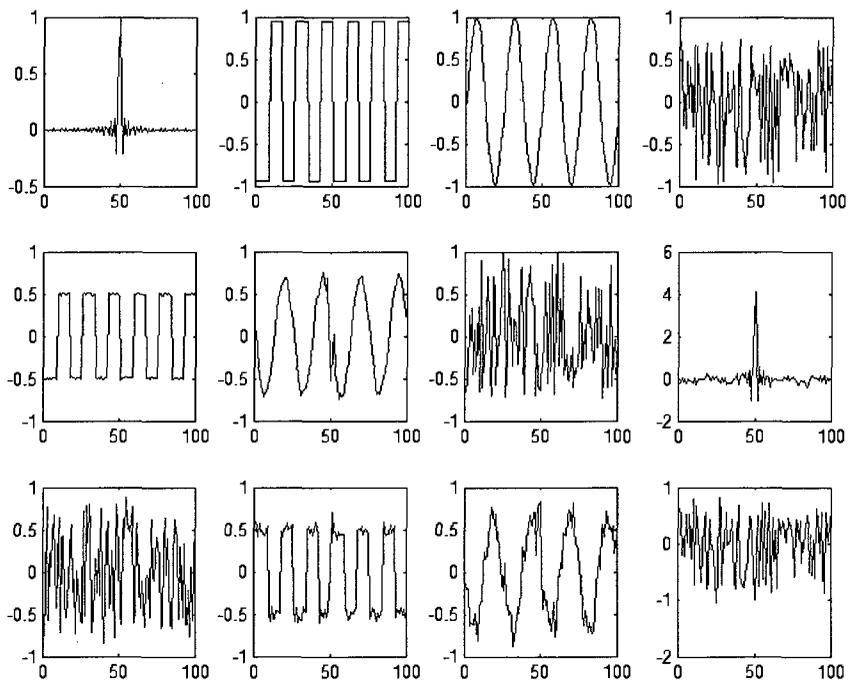


Figure 3.9 Four signal dipoles with six electrodes

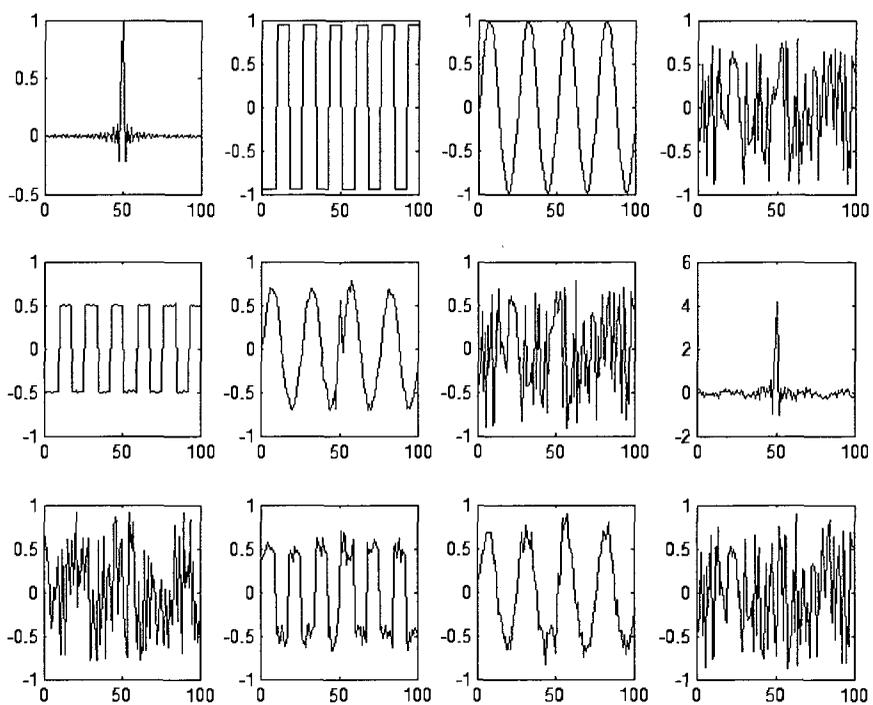


Figure 3.10 Four signal dipoles with four electrodes

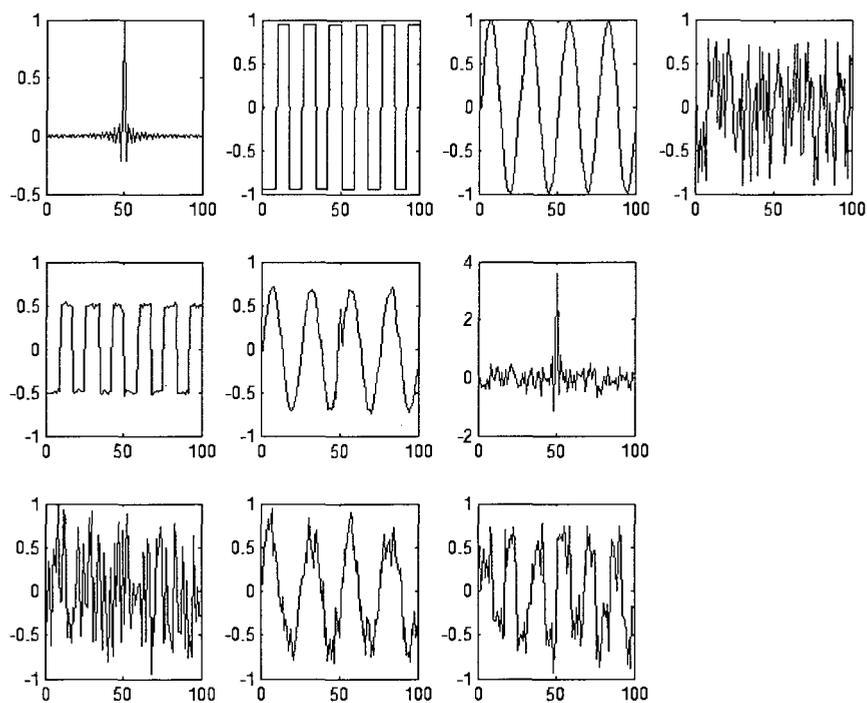


Figure 3.11 Four signal dipoles with three electrodes

Table 3.1 ICA result of four dipole sources with different electrodes number

ICA Cov		Sources			
		Rising Cosine	Rectangular	Sine	White noise
Electrode No.					
Tripolar electrodes	20	0.9712	0.9998	0.9847	0.9806
	10	0.9618	0.9998	0.9846	0.9963
	6	0.9738	0.9998	0.9864	0.9898
	5	0.9680	0.9998	0.9847	0.9896
	4	0.9660	0.9998	0.9822	0.9873
	3	0.8449	0.9992	0.9875	
Disc electrodes	20	0.3042	0.9888	0.9803	0.9717
	10	0.3010	0.9933	0.9863	0.9603
	6	0.2469	0.9921	0.9843	0.9693
	5	0.2402	0.9889	0.9472	0.9497
	4	0.1710	0.9809	0.9697	0.9472
	3	0.1313	0.6647	0.6990	

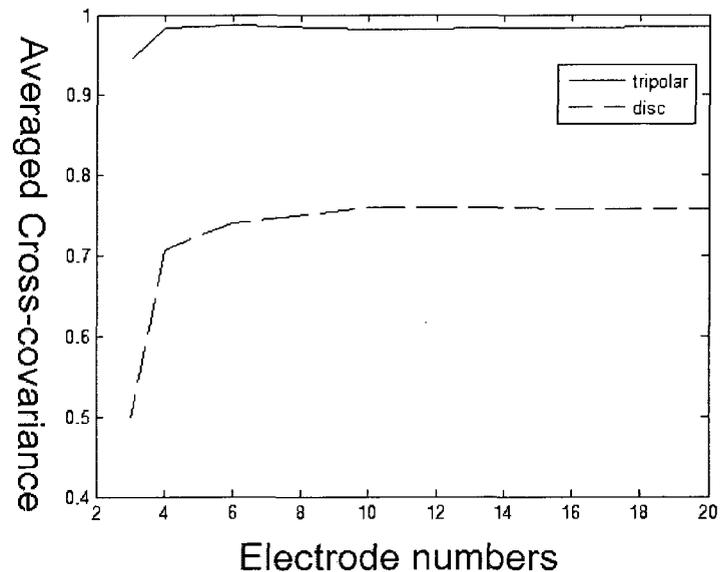


Figure 3.12 Relation between ICA results and electrode numbers, the y-axis was the averaged cross-covariance between the ICA separated signals and the original signals

3.4.3 Relation Between Independence of Sources and ICA Results

Figure 3.12 demonstrates that too many electrodes does not substantially improve ICA separation. However, the experiment also showed that as long as there were enough electrodes (not less than the independent sources) the independent sources would be fully recovered. Dependent sources will not appear in the ICA results. In other words, if there were signals that were controlled by other sources, even if they were located at different places, they will not be found in the ICA results. A dependent signal was one that was produced by the same sources and function without being controlled by other signal sources. They could be collected from different locations (e.g., ECG could be collected from the chest as well as from forehead), but those signals will still be taken as one separate signal. Figure 3.13 and Figure 3.14 show the ICA results with one and two dependent sources.

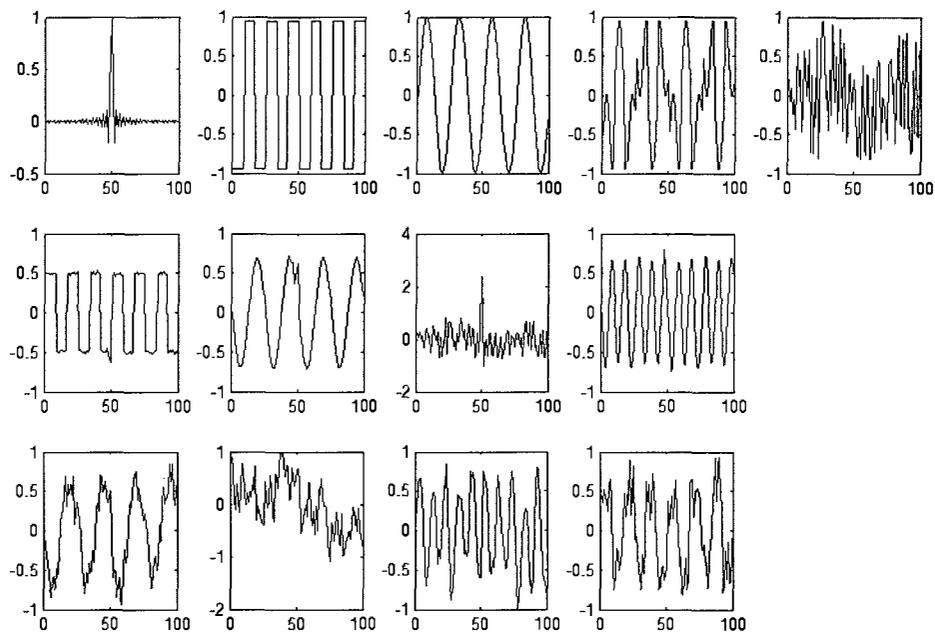


Figure 3.13 Five dipole sources with the fourth source = (source 1 + source 2)/2

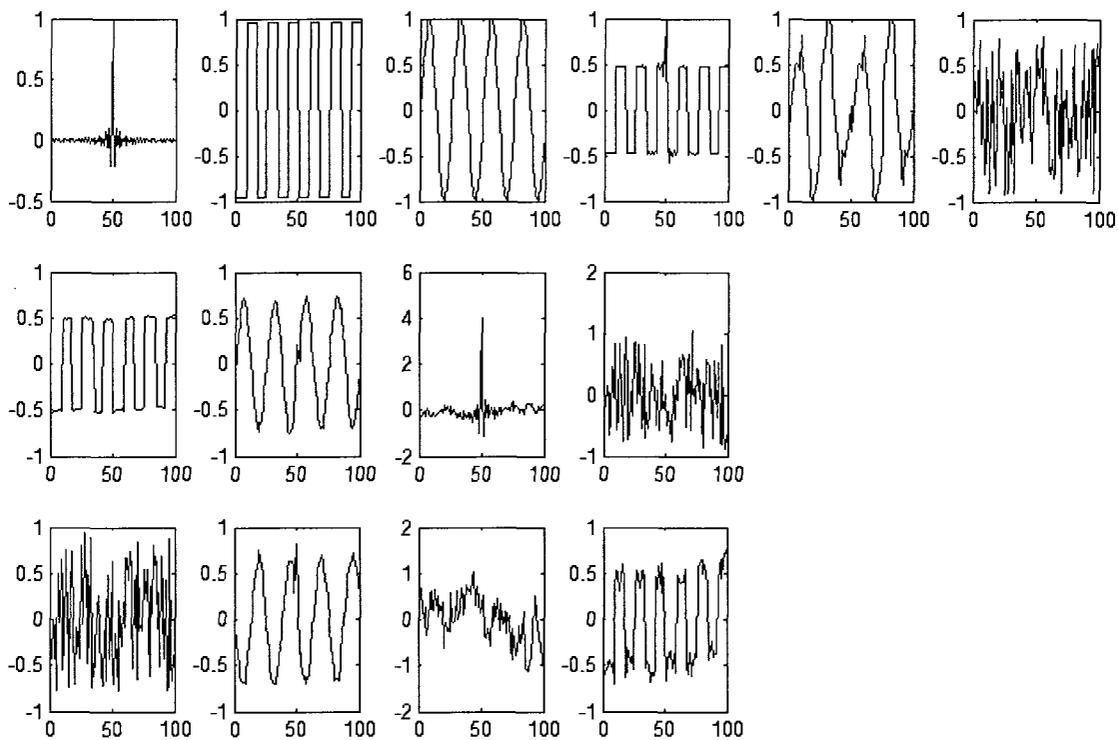


Figure 3.14 Five dipole sources with the fourth source = (source 1 + source 2)/2 and the fifth source = (source 2 + 3×source 3)/4

However, if the sources were from different sources, even if they were in the same kind of waveform, they will still be considered as different sources and be separated into the ICA results. The conditions used to generate Figure 3.14 were nearly identical to those used to generate Figure 3.13 except that the first point of the fourth signal, which depends on source 1 and source 2 in the conditions used for Figure 3.14, was moved to the end of the signal, which made it independent from source 1 and source 2. Though the waveform was nearly unchanged, it was considered to be one independent component and was separated into the ICA results. This effect was clearer in Figure 3.16, where there were two sine waves with the same frequency and same amplitude, but different time series. These signals were considered to be two independent signals and were separated in ICA results.

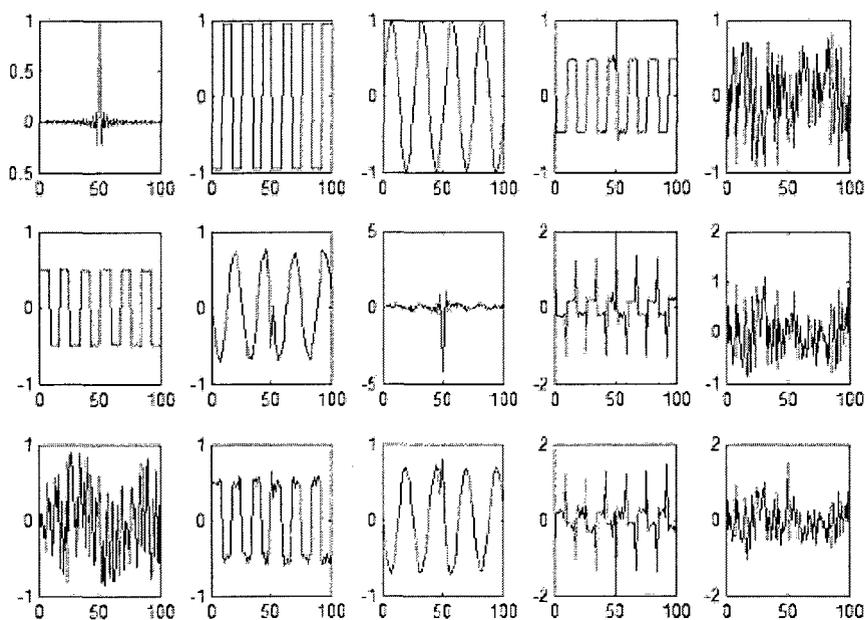


Figure 3.15 Five dipole sources with the fourth source(2:N) = (source 1(1:N-1) + source 2(1:N-1))/2 and fourth source(1) = (source 1(N) + source 2(N))/2

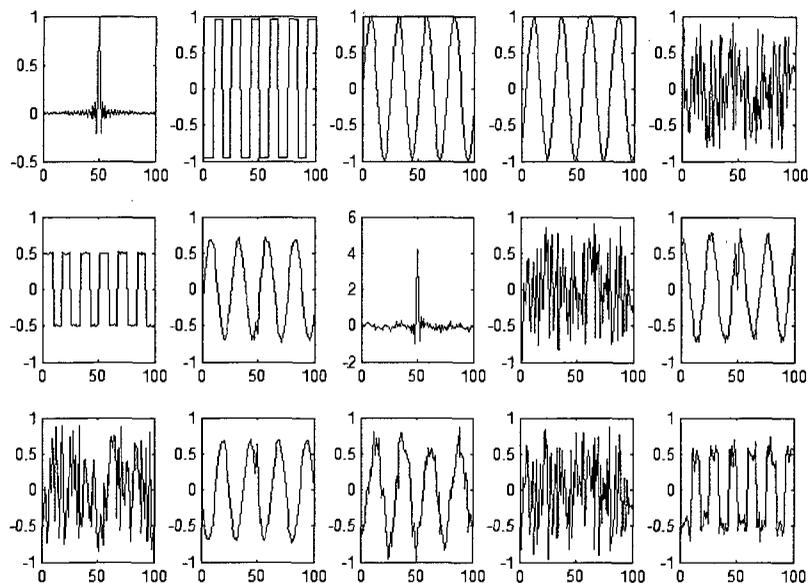


Figure 3.16 Five dipole sources with two sine waves of same frequency and amplitude, but different time series

3.4.4 Tripolar Electrodes vs. Disc Electrodes for ICA

Sections 3.3.2 and 3.3.3 demonstrated that if independent sources (function independently at different locations) generated signals on the electrodes, given enough electrodes, ICA could separate out all the independent components. However, signals with common mode noises will give poor ICA results.

Tripolar Electrodes have been shown to have less mutual information and higher spatial resolution compared to disc electrodes (Koka and Besio, 2007). These advantages gave tripolar signals higher separation ability for ICA and greater common mode noise rejection, which could be seen in Equation (3.9). Further discussion about why tripolar electrodes give better ICA results was made in Section 3.5.

Figure 3.17 was the ICA results for 10 electrodes, including artifacts, under the same conditions as in Figure 3.18 e1 and e2. Figure 3.15 shows the wave forms of the dipoles

(row a), the ICA results of signals from tripolar electrodes (row x1) and the ICA results of the signal from disc electrodes (row x2), where x was from b to e, with respect to the electrode potentials with four artifacts: (b) eye blinking, (c) cheek movements, (d) jaw movements and (e) talking. Table 3.2 gives the normalized covariance for the ICA separation results and the source signals.

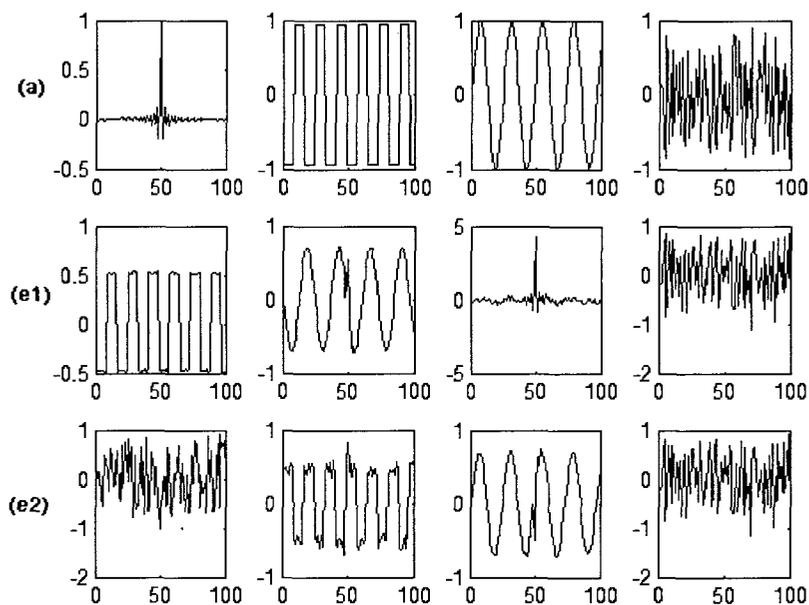


Figure 3.17 ICA Separation results using 10 electrodes with talking artifacts (a) dipole source waves;(e1) Tripolar electrode ICA results; (e2) Disc electrode ICA results

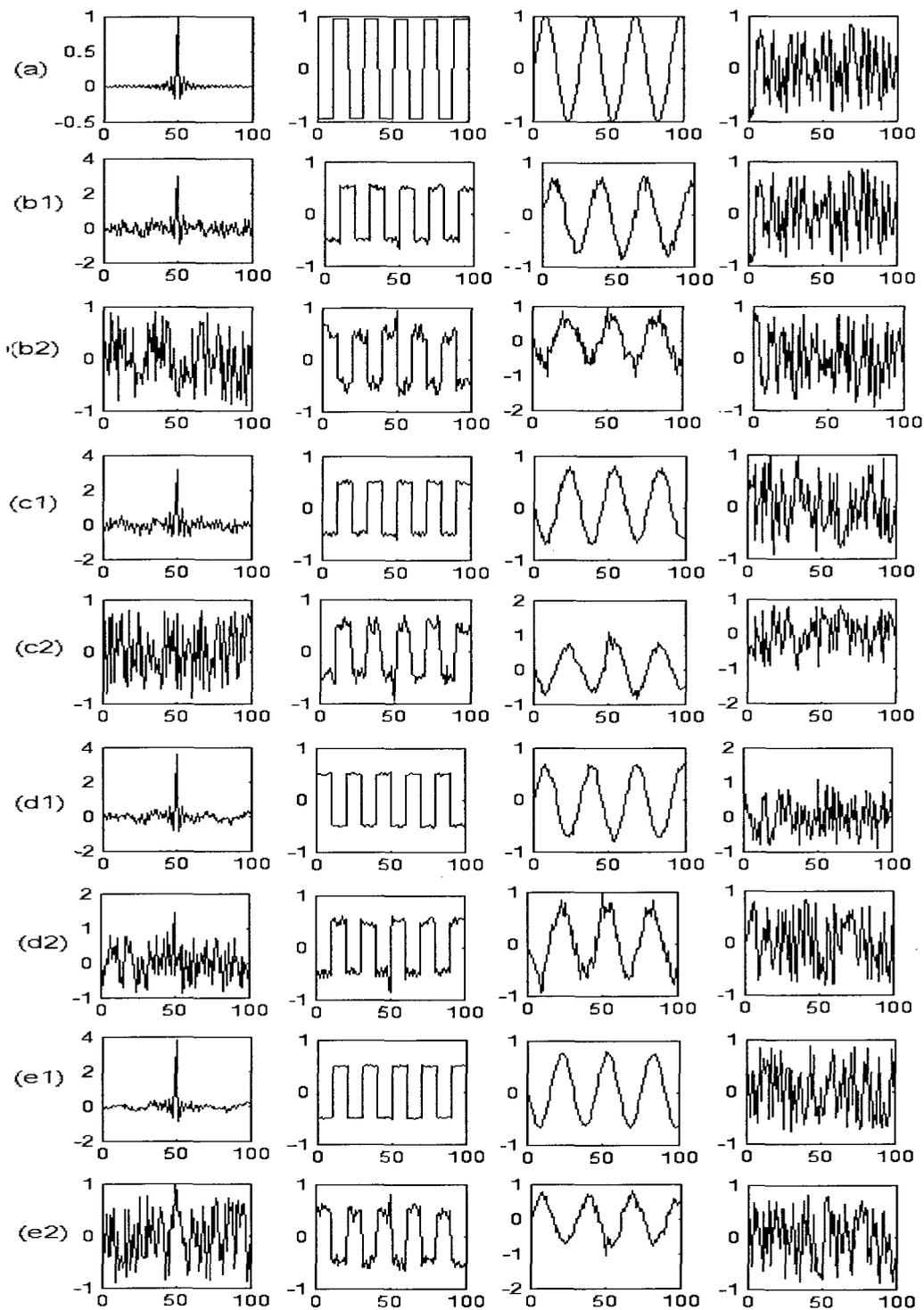


Figure 3.18 ICA results from electrode potentials with four artifacts (a) dipole source waves;(x1) ICA results from the tripolar electrode signals;(x2) ICA results from the disc electrode signals. x was from b to e, with respect to the four artifacts: (b) eye blinking, (c) cheek movements, (d) jaw movements and (e) talking. (Vertical axis – arbitrary units, horizontal axis – time in ms.)

Table 3.2 Normalized covariance of the ICA results and the source signals

IC Co		Dipol sourc		Rising cosine signal	Rectangular signal	Sine wave signal	White noise
		Tripolar	Disc				
Eye blink	Tripolar	0.636	0.996	0.985	0.976		
	Disc	0.231	0.993	0.966	0.962		
Cheek move	Tripolar	0.846	0.998	0.987	0.973		
	Disc	0.231	0.985	0.980	0.972		
Jaw move	Tripolar	0.942	0.999	0.997	0.974		
	Disc	0.237	0.989	0.977	0.960		
talk	Tripolar	0.966	0.999	0.992	0.942		
	Disc	0.393	0.993	0.985	0.907		

3.4.5 Signal Strength Affects ICA

Signal strength strongly influences ICA results. Figure 3.19 was the ICA results with all source signals set to 1, while the in Figure 3.20 and Figure 3.21, the rising cosine wave and the rectangular wave were increased in amplitude. In the ICA results for these conditions, the corresponding independent component improved greatly.

Table 3.3 ICA result of four dipole sources with different numbers of electrodes

ICA Cov Electrode No.		Sources			
		Rising Cosine Signal	Rectangular Signal	Sinewave Signal	White Noise
Tripolar Electrodes	All unity	0.9405	0.9998	0.9837	0.9605
	Rising cosine (10)	0.9978	0.9934	0.9987	0.9820
	Rectangular (10)	0.9683	1.0000	0.9847	0.9900
Disc Electrodes	All unity	0.2307	0.9919	0.9872	0.9265
	Rising cosine (10)	0.9636	0.9399	0.7719	0.6418
	Rectangular (10)	0.2297	0.9999	0.9455	0.9503

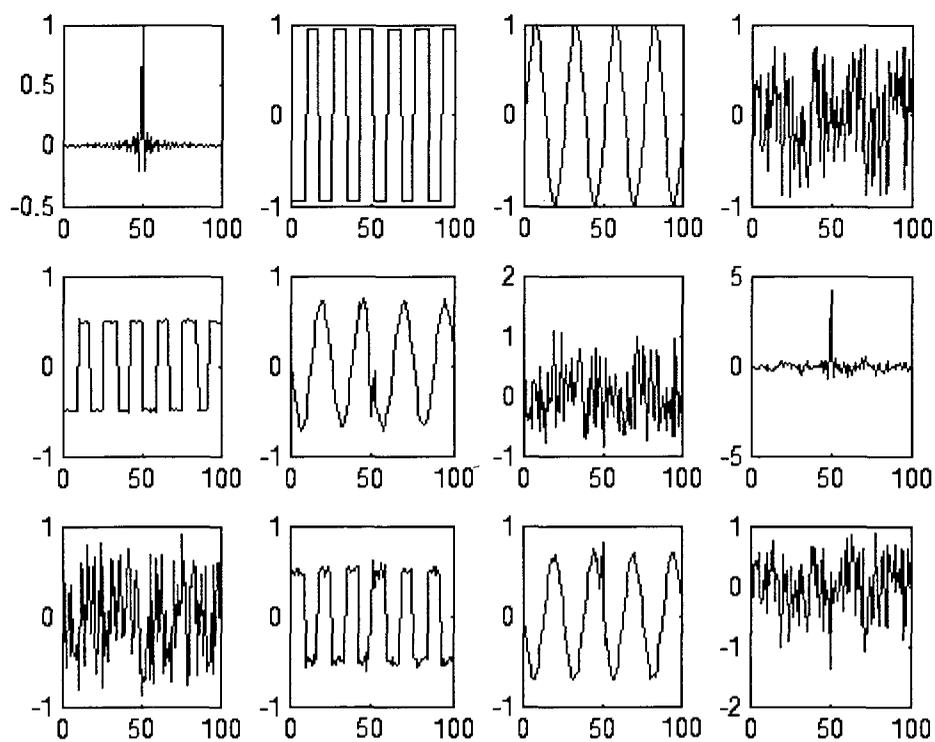


Figure 3.19 ICA results with unity amplitude source signals

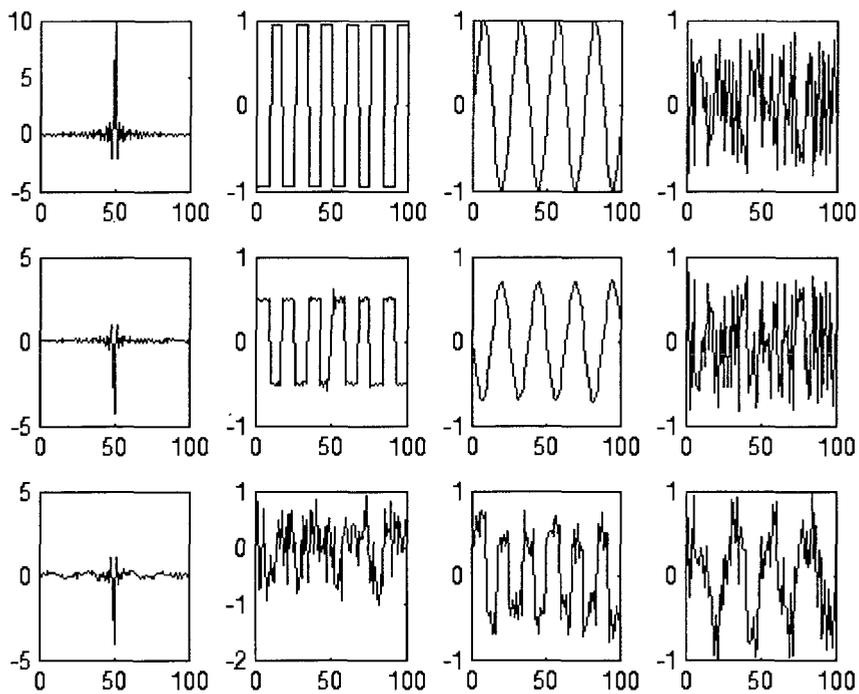


Figure 3.20 ICA results with rising cosine wave strengthened 10 times in amplitude

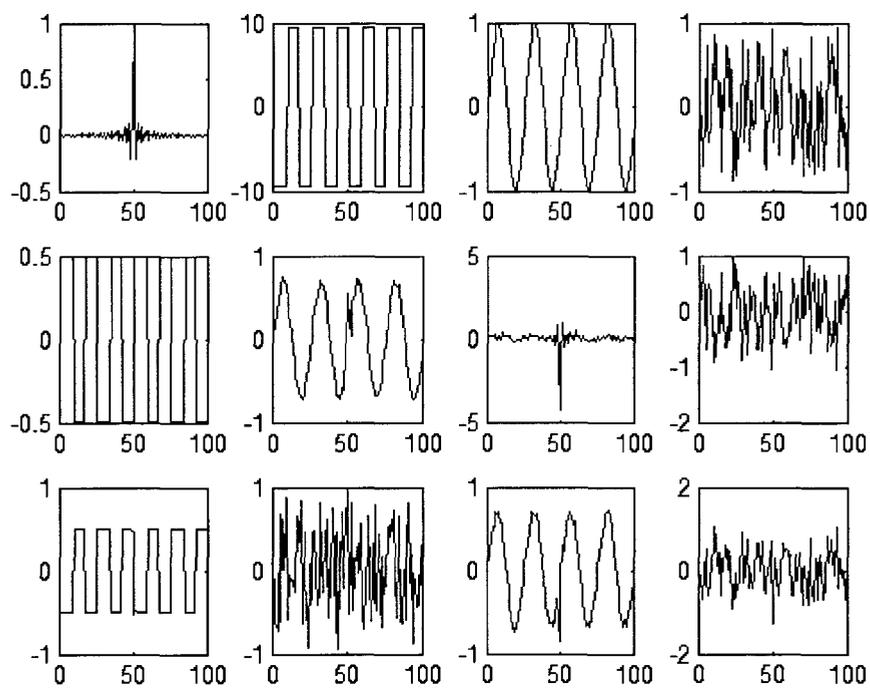


Figure 3.21 ICA results with rectangular wave strengthened 10 times in amplitude

3.5 Conclusions and Discussion

3.5.1 Conclusion

In this chapter, an improved four-layer anisotropic concentric head model was developed, which could calculate tripolar and disc head surface potentials generated by dipoles placed at arbitrary positions within the inner most layer and with arbitrary dipole moments. Then, an ICA analyses model was developed with an improved fast calculation algorithm. Four signals (sine wave signals, rectangular signals, rising cosine signals and Gaussian white noise signals) were simulated as the dipole sources, among which there was one kind of sub-Gaussian signals (sine waves) and there were two kinds of super-Gaussian signals (rectangular signals, rising cosine signals). The ICA results from the tripolar and disc electrodes were compared. Two more interesting points were studied: the influence of the number of electrodes and the number of dipole sources, the relation between the independence/dependence of the sources and the ICA results.

We draw conclusions as follows and then give further discussion:

1. Tripolar Electrodes generate better separation results. This improvement may be due to the tripolar electrodes having higher spatial resolution, thus they were more sensitive to the source spatial distribution and provide more uncorrelated signals for ICA. The improvement may also be due to the higher common mode noise rejection of the tripolar electrodes compared to disc electrodes.
2. The number of independent sources that could be found for the ICA algorithm developed in this chapter was no more than the number of electrodes used in the ICA.
3. More electrodes will help slightly in the separation results when the number of electrodes was greater than or equal to the number of sources.

4. The necessary number of electrodes should not be less than the number of sources.
5. The number of electrodes should keep increasing if the number of independent components was equal to the number of electrodes. In this way, the independent sources could also be found, which was the maximum number of independent components found.
6. The number of electrodes should be slightly higher than the number of independent sources.
7. Source signals that function dependently (differ in time series) will be considered as independent source signals and will be separated into the ICA results, even if they were the same kind of signals with the same frequency and amplitude.
8. Signals with exactly the same time series wave pattern (only differ in amplitude) will be considered to be the same signal patterns, no matter how many they were and where they originate from.
9. Signals that depend on other sources (linearly composed of other sources) will not be found.
10. Common mode noises could not be separated, since they do not have spatial difference. Thus, they will be the main noises in the signal for ICA separation.
11. Noises from a single source could be seen as independent sources. Thus, an ICA process does not consider the pathway through which the signal was produced on the electrodes.
12. Sources with relatively strong signal amplitude will allow ICA to more readily separate the corresponding independent components.

3.5.2 Further Discussion on Three Interesting Facts of ICA

The work of this chapter provided answers to three interesting questions related to the ICA. First, what kind of sources could be separated? Second, what advantages do using tripolar electrodes provide? Third, how many electrodes should be used? Now we give further discussion about those facts.

For the first question, what were the expected separation sources? The work of this chapter showed that sources with the following characters would be or would not be separated in the ICA results:

1. Sources function independently, having no relation with other function sources.
2. Sources giving the same potentials/signal on all electrodes could not be separated. These sources include common mode noises, which could not be seen as coming from one single source and could not be found.
3. Sources with same time series in waveform, despite their different amplitude and locations, will be considered to be one independent source.
4. Sources with different time series in waveform, despite their same amplitude, same locations and same waveform, will be considered to be different independent sources, as shown in Figure 3.13.
5. Sources with strong signals/amplitudes, will generate relatively better corresponding independent components.

From above we can see that for a source to be separated by ICA, it must have three properties:

1. They have spatial difference, that was, they give different signals to spatially different distributed electrodes. The higher the spatial difference, the better the ICA results.
2. It must have a time sequence that differs from the other signals.
3. They must have sufficiently high amplitude to be separated.

From above discussion we conclude that:

1. Common mode noises were not separable, since they do not have spatial differences. Thus, they will be main noises in the signal for ICA separation.
2. Noises from a single source could be seen as independent sources. Thus, the ICA process does not consider the pathway through which the signal was generated on the electrodes.
3. Sources with relatively strong signals/amplitudes will help in ICA for the separation of their corresponding independent components.

However, why tripolar electrodes give better ICA? From Table 3.1 and Table 3.2, the ICA results from tripolar signals were always better than the disc electrode signals, collected under the same conditions. This result could also be seen from Figure 3.6 to Figure 3.19, where all the independent components of the results of the ICA process were taken under the same conditions. In those figures, ICA results from the tripolar signals could find all the independent components with high similarity, while ICA from the disc electrodes give relatively poor independent component separation. This result may be caused by the following advantages of the tripolar electrodes:

1. Higher signal to noise ratio (SNR);
2. Especially higher common mode noise rejection;

3. Higher spatial sensitivity;
4. Less mutual information.

These characteristics, together with the conclusions under the heading ‘Expected separation sources,’ can affect the ICA process, as discussed below.

ICA was mainly used for separating linearly combined signals, which serves as a filter. A combined signal P that was acquired from an electrode can be expressed by

$$P_i = a_1s_1 + a_2s_2 + \dots + a_ns_n, \quad (3.20)$$

where s_i were the uncorrelated, zero-mean, unit variance signals. Thus the weighted coefficients a_i can be seen as the amplitude of the source signal. The greater the difference between a_i , the better the results of ICA filtering (separating). For tripolar electrodes, since they have a high signal to noise ratio (SNR), the signal amplitudes will be relatively higher, which will help in ICA (see Figure 3.16-3.18). Further explanations of the influence of this high SNR of the tripolar electrodes on ICA were provided in Conclusion 3 of the ‘Expected separation sources’ Section.

Since tripolar electrodes have higher spatial resolution (as was shown in Fig 3.4) and lower mutual information, they will give greater differences of coefficients a_i for each electrode, thus lead to better ICA filtering results, which can be seen from Figure 3.15. In Figure 3.15, under the same conditions, the ICA results of the tripolar electrode signals extracted all the independent components with high similarity (see Table 1), while the ICA results of the disc electrodes caused inaccuracy in each independent component extracted. In particular, a rising cosine component was not recognizable in the results from the disc electrodes (Fig 3.15 b2-e2 1st column). This improved separation with the

tripolar electrodes could be further explained by property 1 of the separable sources in ‘Expected separation sources’ Section.

The higher fidelity of the tripolar electrodes may also be caused by their high rejection of common mode noise, which can be seen from the tripolar potential calculation Equations (3.9) (Besio et al., 2006b):

$$PL=16(P_{Middle}-P_c)-(P_{Outer}-P_c), \quad (3.9)$$

where P_{Middle} , P_{Outer} and P_c were potentials from the middle ring, the outer ring and the center disc, respectively. Since those three electrode elements were close to each other, they have nearly the same common mode noise, which becomes sharply attenuated when Equations (3.9) was used in the estimation of the Laplacian tripolar potential PL . When we consider sources, such as the AC wall mains, which were generally distant from the electrodes compared to the signal source in the brain, the common mode noise rejection of the tripolar concentric electrode was beneficial. Further explainaion about this influence of common mode noise rejection to ICA could be seen in conclusion 1 of ‘Expected separation sources’ Section.

In the study, we noticed that the number electrodes to be used in ICA should be selected correctly. Large increases in the number of electrodes beyond the number of sources does not improve the ICA separation of the signal sources and causes unnecessary additional calculations. In contrast, too few electrodes, especially if the number was less than the number of sources, will cause inaccurate ICA separation, as could be seen from Section 3.3.2, ‘Influence of Electrode Number and Source Number on ICA’. From Figure 3.6 through Figure 3.10 and Table 3.1, with the increase of electrode numbers from 5 to 20 for the separation of four sources, the ICA separation accuracy did

not greatly increase. But, if the electrode number fell below the number of sources, the ICA could hardly give a correct separation result. This gave the first rule of electrode number selection: the number of electrodes must not be less than the independent source numbers.

However, we do not generally know a priori how many independent sources will be present. But, from part 3.3.3, 'Relation Between Independence of Sources and ICA Results', Figure 3.11 and Figure 3.12, we could see that if the source signals were not independent of others, they will not be found. Also from part 3.3.2, 'Influence of Electrode Number and Source Number on ICA', Figure 3.6 through Figure 3.9, even if there were more electrodes than independent sources, the ICA results contain no more independent components than the number of independent sources. Thus, we get the second rule of electrode number selection: The number of electrodes should be increased until the number of independent components in the ICA results does not increase.

In addition, from Table 3.1 and comparing Figure 6 through Figure 9 (20 electrodes to 4 electrodes), using more electrodes than the number of sources did help somewhat, which suggested that more electrodes than the minimum number (the number of sources) should be used for ICA.

In conclusion, three rules should be followed in selecting the number of electrodes for ICA:

1. The number of electrodes should not be less than the number of sources.
2. The number of electrodes should be increased if the number of independent components was equal to the number of electrodes. In this way, the number of

independent sources could also be found, which was the maximum number of the found independent components.

3. The number of electrodes should be slightly greater than the number of independent sources.

3.5.3 Further Discussion on Four-Layer Head Model and Fast ICA Algorithm

First, there were some improvement in four-layer anisotropic concentric head model that worth mention. To get an accurate solution in an EEG forward problem, it was necessary to correctly and effectively model the shape of the head and tissues electrical conductivity and dipole distribution (Vatta, Bruno and Inchingolo, 2005). The improved four-layer anisotropic concentric head model developed in this work allows the user not only to set the radial conductivity and the tangential conductivity of each layer, but also to set the number, position, moment direction of the dipole sources. The potential generated on the surface of the head could be calculated for both tripolar and disc electrodes with arbitrary positions. For tripolar electrodes, Laplacian estimation was also given according to the tripolar Laplacian EEG Equations (3.9). In addition, the numerical calculation speed of the model was increased by thousands of times by the introduction of the integration database method.

Secondly, the fast ICA algorithm developed in this study not only has the high convergence speed (cubic), but also was stable and simple, requiring no parameter adjustment. In addition, this study shows that the number of blind independent sources could be determined, which solved the problem in the original algorithm that the blind number sources must be set at the beginning of the ICA.

CHAPTER 4

HUMAN EXPERIMENT

4.1 Structure of the BCI

As we described in Chapter 1, a typical functional structure of a human BCI could be shown as in Figure 4.1. From the figure it can be seen that BCI function was divided into three parts: the first was signal acquisition, the second was signal processing and classification/translation and the third was BCI Application. This paper focuses on the first two parts of the Hardware and Software description BCI.

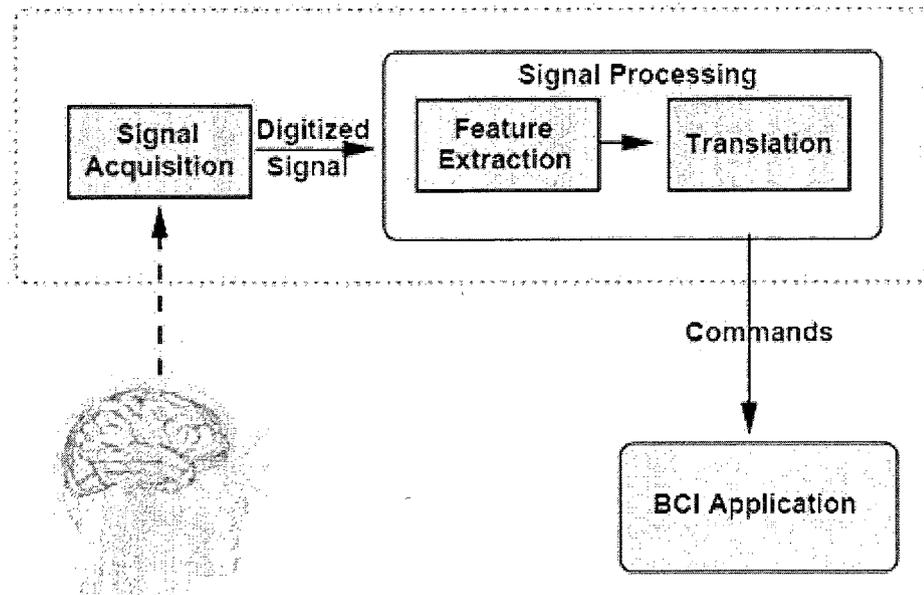


Figure 4.1 A typical functional structure of a human BCI

Figure 4.2 gives the structure of the work of this chapter. First, we design a data acquisition system, which includes electrode design, hardware design and data acquisition protocol and software design. Then, we pre-processed the EEG data, including data filtering, removing eye-blink noise, etc. After that, we studied the data segment-related parameter selection and developed a fast search algorithm to find the best object-specific data segment related parameters for the BCI. In the following, we developed an autoregressive (AR) model for feature extraction and a Mahalanobis distance-based linear classifier for classification. Finally, we compared the results of using tripolar and disc electrodes with the BCI. Conclusions and discussion for the work were provided at the end of this Chapter.

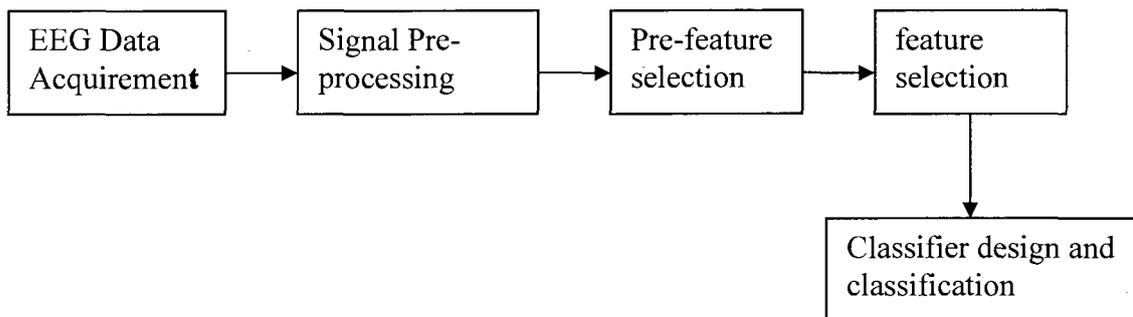


Figure 4.2 Structure of the work of BCI research

4.2 EEG Data Acquirement and Signal Pre-Processing

EEG signals from twelve healthy subjects (females=3, aged from 23 to 30) were recorded using two tripolar concentric electrodes (Figure 4.3) and two virtual disc electrodes (described later), resulting in two data sets. Then four parameters (described later in the Exhaustive Search Algorithm for Parameter Selection section) were studied.

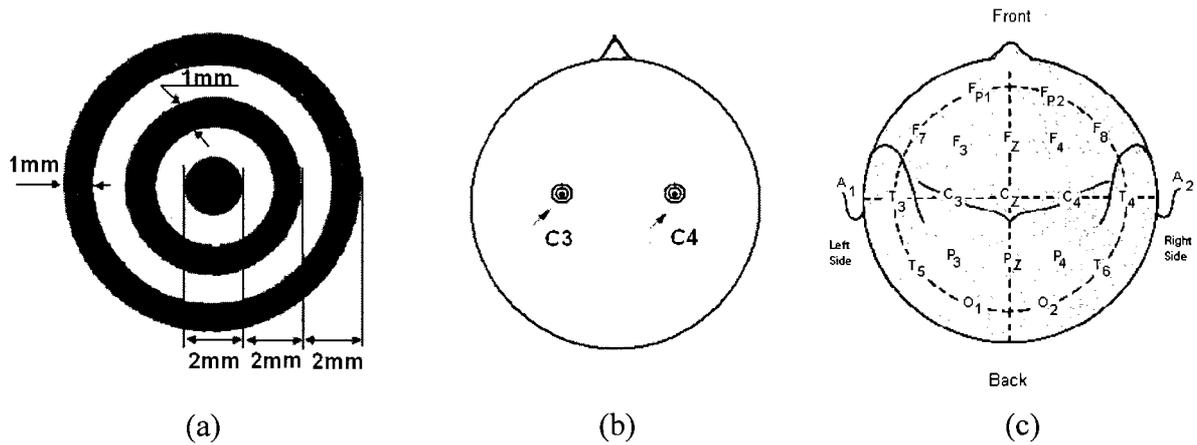


Figure 4.3 Configuration and dimensions of a tripolar concentric electrode (a), electrode positions (b) and 10/20 International Electrode Placement System(c)

4.2.1 Data Acquisition and Hardware Description

All experiments were conducted in accordance with the Institutional Review Board approved protocol (Wolpaw, et al., 2000 a). Two sets of signals from each subject were recorded with tripolar concentric electrodes (Figure 4.3. (b)). Approximately 1.0 mm of 10-20 electrode paste (D.O. Weaver & Co., Aurora, CO) was applied to each electrode prior to placing it on the scalp at C3 and C4 of the 10-20 International Electrode Placement System, as shown in Figure 4.3. (b). Those positions were used in the two channels used for feature extraction. Custom built pre-amplifiers (gain 10) along with a Grass 15LT Bipolar Portable Physiodata Amplifier System with 15A54 AC amplifiers were used for a total gain of 100 K. The filters were set from 0.5 Hz to 30 Hz with the 60 Hz notch filter on. The data were acquired (14 bit) using a DI-720 data acquisition system (DataQ Instruments, Akron, Ohio) with a sampling rate of 125 samples/second per channel.

For signal Data Set 1, two bipolar signals were recorded from each electrode (P_1-P_3 and P_2-P_3 , where P_1, P_2, P_3 were the signals from the outer ring, middle ring and center

disc, respectively). For signal Data Set 2, the outer ring, middle ring and center disc of the electrodes were shorted to make a virtual disc electrode and one signal was recorded from each virtual disc electrode with respect to the reference electrode on the forehead.

4.2.2 Data Acquisition Protocol and Software description

Figure 4.4 was a timing diagram of the protocol followed for acquiring trials of the signals. Each trial started with a visual fixation on a cross displayed on a computer monitor directly in front of the subjects. The cross was displayed for two seconds and then a short warning beep was sounded to alert the subject that a cue was about to be presented. At the third second, a cue was given that lasted for one second. After the cue the subject was required to imagine a left/right hand-lifting movement according to the cue. A random pause was selected such that the length of each trial was between 8 and 9 seconds. For each subject, 480 trials were recorded, approximately 240 each of left and right hand related signals. Half of the signals were used as training data and the other half was used for testing.

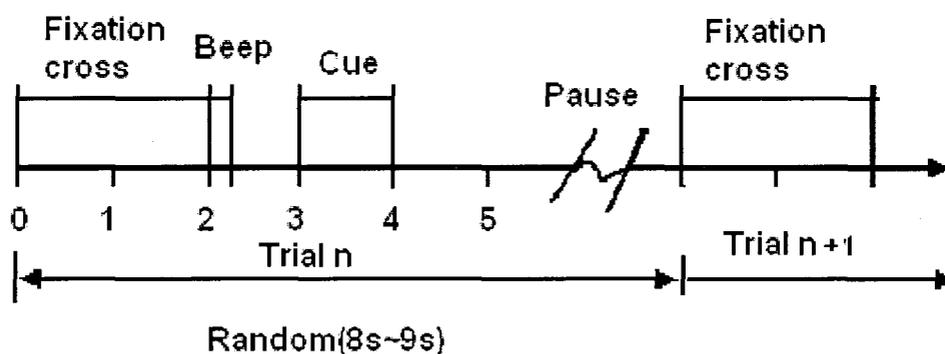


Figure 4.4. Timing diagram of the events during the experimental protocol

The software was developed under a Visual C++ platform using a graphical user interface (GUI). Figure 4.5 gives the structure of the software. It includes the Data Acquisition hardware control, screen display control, timer control, sound control, etc. The user could control the start and end of the program through the GUI.

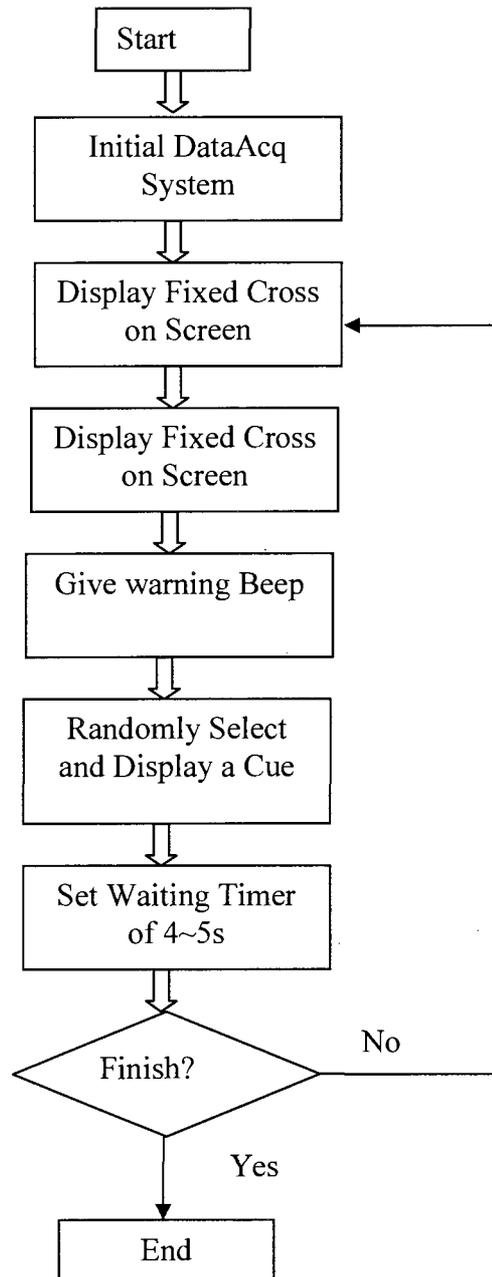


Figure 4.5 Software structure for the data acquisition

4.2.3 Data Pre-processing

The data pre-processing included two parts of work. The first one was the filter design and data filtering. Digital band-pass Butterworth filters were designed for multi-channel EEG signals and the pass band was set from 0.2 Hz to 35 Hz. See attached Matlab program BandEEG() for details.

The second part of the data pre-processing was removal of electrooculogram (EOG) noise. An important noise source in the BCI system was electrooculogram (EOG), which was caused by eye movement and blinking. This signal was much stronger than the EEG signal. It was usually removed by hand. In this work, a threshold filtering method was developed which could remove the eye-blink noises automatically. The threshold was set at 10 times the average of the signal amplitude. The trials containing EOG will be discarded.

4.3 Individualization of Data-Segment-Related Parameters

4.3.1 Introduction

The recognition procedure of EEG signals includes three steps: feature extraction, feature selection and classification. Previously the improvements of the parametric modeling techniques have mostly focused on developing more effective feature extraction and selection methods (Wolpaw, et al., 2000; Pfurtscheller, et al., 2000; Jonathan, et al., 2002; Wang, McFarland and Vaughan, 2000; Pardey, Roberts and Tarassenko, 1996; Burke, et al. 2005; Schroder, Bogdan and Rosenstiel, 2003). However, there were also works concerned with data segment/channel selection (Stastny, Sovka and Stancak, 2003; Burke, et al. 2005; Jiruska, et al., 2005; Palaniappan, 2006; Ince, Arica and Tewfik, 2006). For example, Burke, et al. stated that there might be another data

segment length employed providing better classification and that the order of ARX model needs to change according to different waveforms (Burke, et al., 2005); also Schroder, et al. showed that the ‘all channel choice’ for feature selection and classification in BCI was not the best choice (Schroder, Bogdan and Rosenstiel, 2003). Further, Jiruska, et al. showed that longer signal segments brought comparably better results and that the best results were obtained with a fixed AR model order, not the automatically chosen order by the auto order selection algorithms, such as MDL and AIC (Jiruska, et al., 2005).

Analysis of learned EEG patterns confirms that for a subject to operate satisfactorily his/her personal BCI, the personal BCI must fit the individual features of the subject (Millan, et al., 2002a, 2002b). However, no report was published for systematic analysis of the influence of data segment selection on BCI classification. In this work, the necessity of data segment selection was proven and an auto selection method was developed, which greatly increased the signal classification in BCI. In the work presented in this paper, the length of the data segment in each trial (LDS), along with its starting position (SPD) were studied. Since the ‘all channel choice’ was not the optimum choice (Schroder, Bogdan and Rosenstiel, 2003), it was possible that not using all trials in each channel may generate better results. As a matter of fact, in all the available trials, it was most likely that not all of them provide a ‘good pattern’ for the action-related EEG signal due to the changing conditions of the subjects over time. For example, when a left-cue was given, the subject was supposed to think about lifting their left hand but, if they did not focus during that time and did not think about moving the left arm/hand or thought about something else, there would not be a ‘good pattern.’ Thus, this trial to the observer may look appropriate, but it was not suitable to be used in the assembly of the signal

processing model. Trials like that should be deleted. The algorithm was simple: if adding a trial into the model building can increase the classification ratio, keep the trial; otherwise, delete it. Thus, the number of trials (NT) that were used to build the model was also analyzed.

What's more, an AR model was used for feature extraction (Ben, Bourne and James, 1981; Penny, et al., 2000; Burke, et al., 2005) and a Mahalanobis distance-based linear classifier was used for classification (Mahalanobis, 1936). Thus, selection of the model order (AR Order) was also analyzed. Since all four parameters mentioned above were within a known range, an exhaustive search algorithm was employed on each subject's EEG data to find the best value of these four parameters for EEG based BCI classification. The results showed that the four parameters had a great influence on the classification accuracy and that proper selection of those parameters' values produced a significantly better classification ratio (CR) when compared with the results without the selection process. The results also showed that the optimum value of those four parameters were subject-dependent, which suggests that parameters should be individualized for each subject.

4.3.2 Auto Search Algorithm for Parameter Selection

In this work, four data-segment-related parameters were studied and an auto search algorithm was developed. The four data-segment-related parameters were the length of the data segment in each trial (LDS), with its starting position (SPD), the number of trials (NT) and the AR model order (AR Order). Since all the parameters were within a known range, an auto search algorithm was used to select the best values of the parameters for the highest classification, which was an improved exhaustive search method. The key

was to select the proper search step length, such that the total range was covered effectively using the least amount of time. The rules for the selective-search algorithm were:

1. Set the search range equal to the total possible range of the parameters, (see Table 1 for each parameters' search range).
2. Set the step length to one-fifth the search range and find and record the best parameters' values.
3. Reduce the search length to be $2/5^{\text{th}}$ of the last search length, centered at the best parameters' values, then repeat Step 2.
4. Repeat Step 3 until the optimum parameters' values do not change or until the step length was smaller than $1/100^{\text{th}}$ of the total possible search length. Record the optimum parameters' value.

In each searching round of step 2 and 3, the classification results were recorded and the best parameters were those that generated the highest classification ratio (CR). The following was the example for the search process of LDS (other parameters were searched simultaneously with LDS). In step 1 and step 2, the initial search range of LDS was set at 0.1 to 3 s with a length of 2.9s. The searching step was then set as $2.9/5=0.58\text{s}$. Thus, LDS values of 0.1, 0.68, 1.26, 1.84, 2.42 and 3.00s were used. In step 3, suppose that the LDS of 1.26s was the best, then, the search length was reset to $2(2.9/5) = 1.16\text{s}$, which centered at 1.26s, leading the range to be 0.68 to 1.84s. Then, Step 2 was repeated until the search step was smaller than $1/100^{\text{th}}$ of the total possible search length of 2.9s.

For the auto search algorithm developed in this work, it was necessary to only set the possible search ranges. The system would then automatically decide which of the set

parameters was the optimum one. The criterion was giving the highest classification ratio (CR) that was defined by Equations (4.1).

$$CR = \frac{\text{Correctly recognized trials}}{\text{Total trials recognized}}, \quad (4.1)$$

where the total trials recognized were approximately 160 to 200 for each subject. Table 4.1 gives the initial search algorithm for each of the parameters studied in this work. Further discussion was added in section 4.7.

Table 4.1 Range of each parameter tested

Parameter	Range
LDS(s)	0.1-3
SPD(s)	4.5-7.4
NT	0.3-1.0
AR Order	3-15

ANOVA was performed to compare the CRs for the signals generated from the tripolar concentric electrodes and virtual disc electrodes. Statistics were reported as mean \pm standard deviation with P-values designated to test significance.

4.3.3 Results and Discussion for Data-Segment-Related Parameters Selection

The experiment showed that the LDS, SPD, AR Order and NT significantly influenced the CR for the EEG data recorded from the 12 subjects. For these data, the optimal values of these parameters to generate the highest CR in EEG-based BCI did exist. A selective searching algorithm was capable of finding the optimum values. Our results also show that the optimum values of these four parameters were subject-

dependent, which suggests that, when constructing a model for BCI EEG analysis, subject variances should be considered and the parameters should be individually customized.

Though the conclusions were based on the model analyzed for this work, all the parameters discussed, other than AR Order, were subject specific. Therefore, it may be possible that our conclusions were suitable for other models as well.

Detailed results and conclusions were provided in Section 4.7, where they were discussed together with the BCI classification results.

4.4 Feature Selection for BCI

4.4.1 Introduction on Feature Selection for BCI

The first step of pattern recognition was the feature selection, including what features were used, what their properties were and how they were used. There were many kinds of features that have been attempted to design BCI, such as amplitude values of EEG signals (Kaper, et al., 2004), BP (Pfurtscheller, et al., 1997), PSD values (Millan and Mourino, 2003), autoregressive (AR) and adaptive autoregressive (AAR) parameters (Penny, et al., 2000; Pfurtscheller, et al., 1998), time frequency features (Wang, Deng and He, 2004) and inverse model-based features (Qin, Ding and He, 2004; Kamousi, Liu and He, 2005; Congedo, Lotte and Lecuyer, 2006). In this work, we used AR model parameters, PSD values and BP values for feature selection in the human BCI experiments.

4.4.2 Feature Selection Using an AR Model

In statistics and signal processing, an autoregressive (AR) model was a type of random process which was often used to model and predict various types of natural phenomena. The parameters of an AR model could be used as features for the data it models.

An AR(p) model was defined as Equation (4.2); The notation AR(p) refers to the autoregressive model of order p),

$$X_t = c + \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t, \quad (4.2)$$

where $\varphi_1, \dots, \varphi_p$ were the parameters of the model, c was a constant and ε_t was white noise. The constant term was often omitted for simplicity. An autoregressive model was essentially an all-pole infinite impulse response filter with some additional interpretation placed on it. Some constraints were necessary on the values of the parameters of this model in order that the model remains stationary. For example, processes in the AR(1) model with $|\varphi_1| \geq 1$ were not stationary.

The calculation of the AR parameters was based on parameters φ_i , where $i = 1, \dots, p$. There was a direct correspondence between these parameters and the covariance function of the process and this correspondence can be inverted to determine the parameters from the autocorrelation function (which was itself obtained from the covariance). This was done using the Yule-Walker Equations (Ben, Bourne and James, 1981; Penny, et al., 2000; Burke, et al. 2005):

$$\gamma_m = \sum_{k=1}^p \varphi_k \gamma_{m-k} + \sigma_\varepsilon^2 \delta_m, \quad (4.3)$$

where $m = 0, 1, \dots, p$, yielding $p + 1$ Equations. γ_m was the autocorrelation function of X , σ_ϵ was the standard deviation of the input noise process and δ_m was the Kronecker delta function. Because the last part of the Equation was non-zero only if $m = 0$, the Equation was usually solved by representing it as a matrix for $m > 0$, thus getting Equations (4.4):

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \dots \end{bmatrix} = \begin{bmatrix} \gamma_1 & \gamma_1 & \gamma_1 & \dots \\ \gamma_2 & \gamma_1 & \gamma_1 & \dots \\ \gamma_3 & \gamma_1 & \gamma_1 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \dots \end{bmatrix}. \quad (4.4)$$

Then solve for all φ . For $m = 0$, Equation (4.3) turn into Equation (4.5),

$$\gamma_0 = \sum_{k=1}^p \varphi_k \gamma_{-k} + \sigma_\epsilon^2, \quad (4.5)$$

which allows us to solve for σ_ϵ^2 .

Equations (4.3) to (4.5; the Yule-Walker Equations) provide one route to estimating the parameters of an AR(p) model, by replacing the theoretical covariance with estimated values. One way of specifying the estimated covariance was equivalent to a calculation using least squares regression of values X_t on the p previous values of the same series.

The order of an AR model was essential when using AR model parameter as features in the classification. As one can surmise, the AR model can be any order as desired. However, it should be as accurate as possible in terms of signal representation. Intuitively, it was known that a model order, which was too small, will not represent the properties of the signal, where as, a model order which was too high will also represent noise and inaccuracies and thus, will not be a reliable representation of the true signal. Therefore, methods that will determine the appropriate model order must be used and this problem has produced many published works (Akaike, 1974; Parzen, 1974; Rissanen, 1978;

Hannan and Quinn, 1979; Aufrichtig and Pederson, 1992; Wear, Wagner and Garra, 1995; Palaniappan, 2006).

There were many criteria in the literature for determining the AR Order. Some of these were the Akaike Information Criterion (AIC) and Final Prediction Error (FPE) (Akaike, 1974) pioneered by Akaike. Other commonly used criteria were the Minimum Description Length (MDL), suggested by Rissanen (Rissanen, 1978), Criterion Autoregressive Transfer (CAT) by Parzen (Parzen, 1974) and Residual Variance (Wear, Wagner and Garra, 1995). The Hannan and Quinn (HQ) (Hannan and Quinn, 1979) criterion increases the penalty for large order models to counteract the overfitting tendency of AIC. Aufrichtig and Pederson (Aufrichtig and Pederson, 1992) have studied AIC, while Palaniappan proposed a genetic algorithm (GA) method (Palaniappan, 2006). These methods did not cover all of the AR model order selection problems for the vast amount of applications of the AR models. However, they illustrate the importance and the necessity for appropriate selection of the AR model.

In the work of this study, the auto-search algorithm was developed also for AR Order selection (see Section 4.3), which showed that for the EEG-based classification, AR Order was also subject specific and yet could be set automatically. Additional discussion will be conducted in Section 4.6 for the AR model order selection.

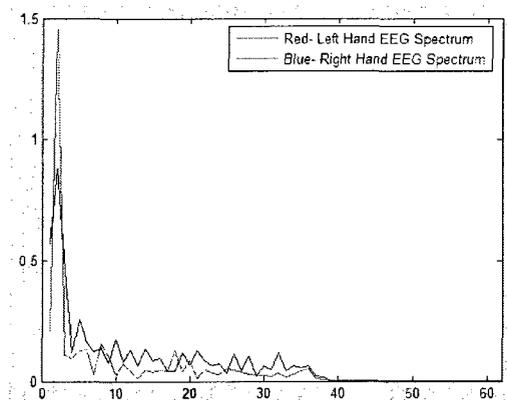
4.4.3 Feature Selection Using Spectrum Characters

The EEG was typically described in terms of (1) rhythmic activity and (2) transients. The rhythmic activity was divided into bands by frequency. To some degree, these frequency bands were a matter of nomenclature (i.e., any rhythmic activity between 8–12 Hz can be described as "alpha"), but these designations arose because rhythmic activity

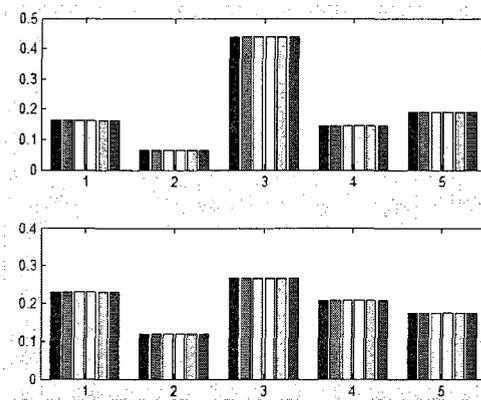
within a certain frequency range was noted to have a certain distribution over the scalp or a certain biological significance.

Most of the cerebral signal observed in the scalp EEG falls in the range of 1–20 Hz (activity below or above this range was likely to be artifactual, under standard clinical recording techniques). Table 4.2 gives the comparison of EEG bands.

Since different EEG bands were associated with certain mental activities and our BCI was based on the analysis of EEG signals associated with the imagination of left-right hand movements, BP (Pfurtscheller, et al., 1997) and PSD values (Millan and Mourino, 2003) maybe suitable for use as features for BCI classification. Figure 4.6 gives an example of PSD (a) and BP (b) of left/right hand imaged movement related EEG. The figures were generated use 60 trials from one subject and for each trial, the signals segment was from 0.5s after stimuli to 3s after stimuli. The spectra were averaged by the trial numbers that were used to generate them. In Figure 4.6 (b), The five bands were 10-12 Hz, 13-14 Hz, 15-20 Hz, 21-24 Hz and 25-28 Hz respectively and shows the plot of six channels for each band, with right hand related EEG band power at the top and left hand related EEG band power at the bottom. However, they did not yield as good results as the AR model. Further discussion presented in Section 4.7, the discussion part of this chapter.



(a)



(b)

Figure 4.6 PSD (a) and BP (b) of left/right hand imaged-movement-related EEG

Table 4.2 Comparison of EEG bands

Type	Frequency (Hz)	Location	Normally	Pathologically
Delta	up to 4	frontally in adults, posteriorly in children; high amplitude waves	<ul style="list-style-type: none"> adults slow wave sleep in babies 	<ul style="list-style-type: none"> subcortical lesions diffuse lesions metabolic encephalopathy hydrocephalus deep midline lesions.
Theta	4 – 7 Hz		<ul style="list-style-type: none"> young children drowsiness or arousal in older children and adults idling 	<ul style="list-style-type: none"> focal subcortical lesions metabolic encephalopathy deep midline disorders some instances of hydrocephalus
Alpha	8 – 12 Hz	posterior regions of head, both sides, higher in amplitude on dominant side. Central sites (c3-c4) at rest.	<ul style="list-style-type: none"> relaxed/reflecting closing the eyes 	<ul style="list-style-type: none"> coma
Beta	12 – 30 Hz	both sides, symmetrical distribution, most evident frontally; low amplitude	<ul style="list-style-type: none"> alert/working active, busy or anxious thinking, active concentration 	<ul style="list-style-type: none"> benzodiazepines
Gamma	30 – 100 +		<ul style="list-style-type: none"> certain cognitive or motor functions 	

4.5 EEG Signal Classifier for BCI

4.5.1 Introduction on EEG Signal Classification for BCI

Just as there were many feature extraction methods, there were also many classification algorithms that developed for the BCIs, which were divided into five categories: linear classifiers, neural networks, nonlinear Bayesian classifiers, nearest neighbor classifiers and combinations of classifiers (Lotte, et al., 2007). Among those classification algorithms, linear classifiers were probably the most popular algorithms for BCI applications. Linear classifiers were discriminant algorithms that use linear functions to distinguish classes. Two main kinds of linear classifiers have been used for BCI design, namely, linear discriminant analysis (LDA) and support vector machine (SVM). In this work, LDA and SVM methods and Mahalanobis distance based classifiers were employed for the EEG classification in the human BCI experiments.

The aim of LDA (also known as Fisher's LDA) was to use hyper planes to separate the data representing the different classes (Fukunaga, 1990). For a two-class problem, the class of a feature vector depends on which side of the hyperplane the vector was located (see Figure 4.7).

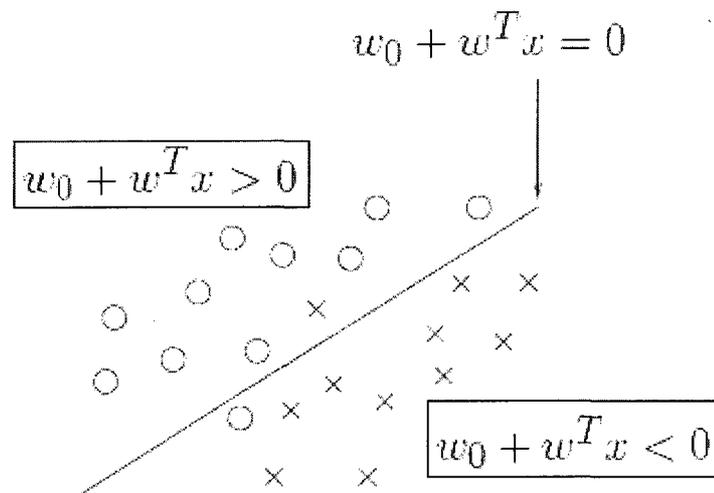


Figure 4.7 Structure of a Linear Discriminant Analysis

LDA assumes normal distribution of the data, with equal covariance matrix for both classes. The separating hyperplane was obtained by seeking the projection that maximize the distance between the means of the two classes and minimize the interclass variance (Fukunaga, 1990).

This technique has a very low computational requirement which makes it suitable for online BCI system. Moreover this classifier was simple to use and generally provides good results. Consequently, LDA has been used with success in a great number of BCI systems, such as motor imagery based BCI (Pfurtscheller and Lopes, 1999), P300 speller (Bostanov, 2004), multi-class (Garrett, et al., 2003) or asynchronous (Scherer, et al., 2004) BCI. The main drawback of LDA was its linearity requirement that can provide poor results on complex nonlinear EEG data (Garcia, Ebrahimi and Vesin, 2003).

A support vector machine (SVM) also uses a discriminante hyperplane to identify classes (Burges, 1998; Bennett and Campbell, 2001). However, concerning SVM, the selected hyper-plane was the one that maximizes the margins, i.e., the distance from the

nearest training points (see Figure 4.8). Maximizing the margins was known to increase the generalization capabilities (Burges, 1998; Bennett and Campbell, 2001). An SVM uses a regularization parameter C that enables accommodation to outliers and allows errors on the training set.

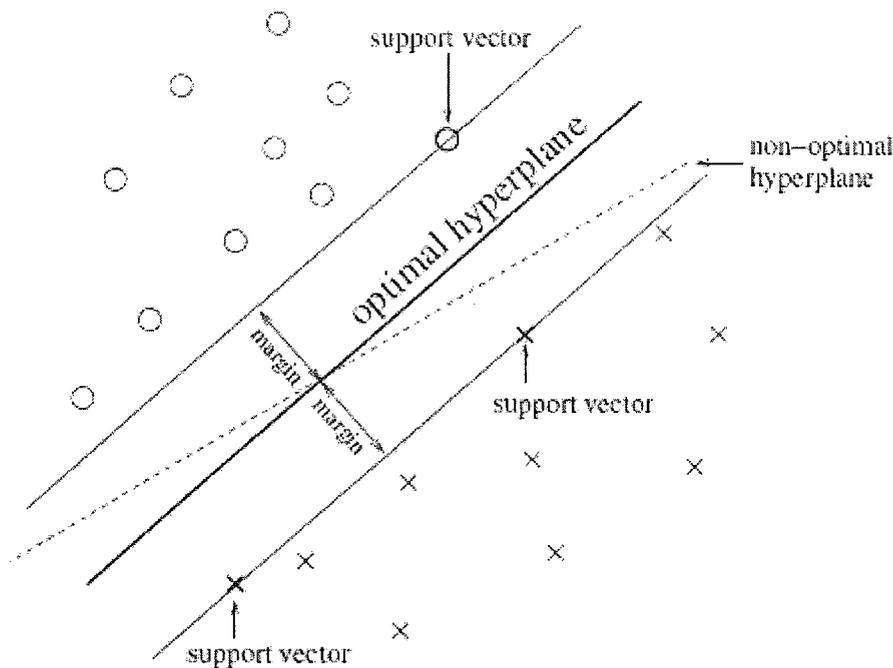


Figure 4.8 Structure of a Support Vector Machine

Such a SVM enables classification using linear decision boundaries and was known as linear SVM. This classifier has been applied, always with success, to a relatively large number of synchronous BCI problems (Blankertz, Curio and Muller, 2002; Garrett, et al., 2003). However, it was possible to create nonlinear decision boundaries, with only a small increase in the classifier's complexity, by using the “kernel trick” which consists in implicitly mapping the data to another space, generally of much higher dimensionality, using a kernel function $K(x; y)$. The kernel generally used in BCI research was the Gaussian or Radial Basis Function (RBF) kernel:

$$K(x, y) = \exp\left(\frac{-\|x - y\|^2}{2\sigma^2}\right). \quad (4.6)$$

The corresponding SVM was known as a Gaussian SVM or RBF SVM (Burges, 1998; Bennett and Campbell, 2001) RBF SVM have also given very good results for BCI applications (Kaper, et al., 2004; Garrett, et al., 2003). As LDA, SVM has been applied to multiclass BCI problems using the OVR strategy (Schlogl, et al., 2005). SVM have several advantages. Actually, thanks to the margin maximization and the regularization term, SVM were known to have good generalization properties (Garrett, et al., 2003; Jain, Duin and Mao, 2000), to be insensitive to overtraining (Jain, Duin and Mao, 2000) and to the curse-of-dimensionality (Blankertz, Curio and Muller, 2002; Garrett, et al., 2003). Finally, SVM have a few hyperparameters that need to be defined by hand, namely, the regularization parameter C and the RBF width - if using kernel 2. These advantages were gained at the expense of a low execution speed.

4.5.2 Mahalanobis Distance Based Classifier

Mahalanobis distance based classifiers assume a Gaussian distribution $N(\mu_c, M_c)$ for each prototype of the class c . If a matrix $X = (x_1, x_2, \dots, x_N)$ was defined as the feature matrix of N trials, where x_i was the feature vector of the i th trial, the Mahalanobis distance (Mahalanobis, 1936; Cincotti, et al., 2002, 2003) of a feature vector x to X was defined by equation (4.7):

$$D_M(x) = \sqrt{(x - \mu)^T M_c (x - \mu)}, \quad (4.7)$$

where x was the feature vector given in Equations (2), $\mu = \bar{X} = \frac{1}{N} \sum x_i$ was the mean of X and $M_c = E[(X - \mu)(X - \mu)^T]$ was the covariance matrix of X .

From Equations (4.7) the Mahalanobis distance of x to the left hand related group X_L and the Mahalanobis distance to the right hand related group X_R were determined. Then x was classified by the following rules:

$$\begin{cases} X_L > X_R \Rightarrow \text{right - handed} \\ X_L < X_R \Rightarrow \text{left - handed} \end{cases} \quad (4.8)$$

The Mahalanobis distance based classifier was a simple yet robust classifier and has been proven to be suitable for multiclass (Schlogl, et al., 2005) or asynchronous BCI systems (Cincotti, et al., 2003).

4.6 BCI Classification Results

This work focused on the influence of two main factors in BCI classification. The first was the application of a pre-feature selection algorithm. The second was the application of tripolar electrodes.

4.6.1 Influence of Pre-Feature Selection Algorithm

The pre-feature selection algorithm improved the CR by parameter individualization. To illustrate the reliance of the CR on the data-segment-related parameters, CRs with/without parameter individualization were shown in Figure 4.9.

The values of the parameters without individualization were LDS = 1.2 s, SPD = 1 s, AR Order = 11 s and NT = 1, as recommended by previous works (Palaniappan, 2006; Jiruska, et al., 2005; Schroder, Bogdan and Rosenstiel, 2003). The value NT = 1 represents all of the available data. The averages of CRs with/without parameter

individualization were 78.73 ± 3.30 and 57.68 ± 14.99 , respectively, as was shown in Table 4.3.

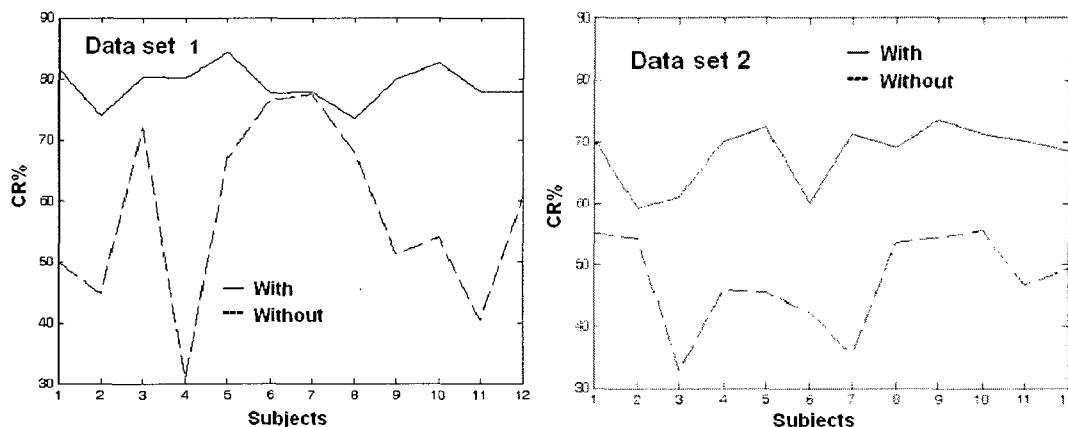


Figure 4.9 CR for Data set 1 and 2 with/without parameter individualization

Table 4.3 CR of the two data sets with/without parameter individualization

	Data Set 1		Data Set 2	
	Without	With	Without	With
CR Avg.	57.68	78.73	47.60	68.01
CR Std.	14.99	3.30	7.68	4.97
P value	8.7×10^{-5}		1.0×10^{-7}	

Another noticeable phenomenon in the pre-feature selection is the influence of LDS, SPD, AR Order and NT on CR. Figure 4.10 shows the influence of LDS (a), SPD (b), AR Order (c) and NT (d) on the CR for subject 1, which was indicative of all the subjects. In each of the figures, five points around the optimum parameter setting were selected. As was shown in 4.10, given (LDS, SPD, AR Order, NT) as (0.3, 0.9, 11, 0.52), respectively, a maximum CR of 81.43% was achieved by Data Set 1 (tripolar) and 69.84% was achieved by Data Set 2 (virtual disc). However, even if only one of the parameters was changed, the CR changed greatly. For example, if LDS was changed (see Figure 4.10 (a)), the CR could decrease to as low as 58.1% (Data Set 1) and 58.9% (Data Set 2).

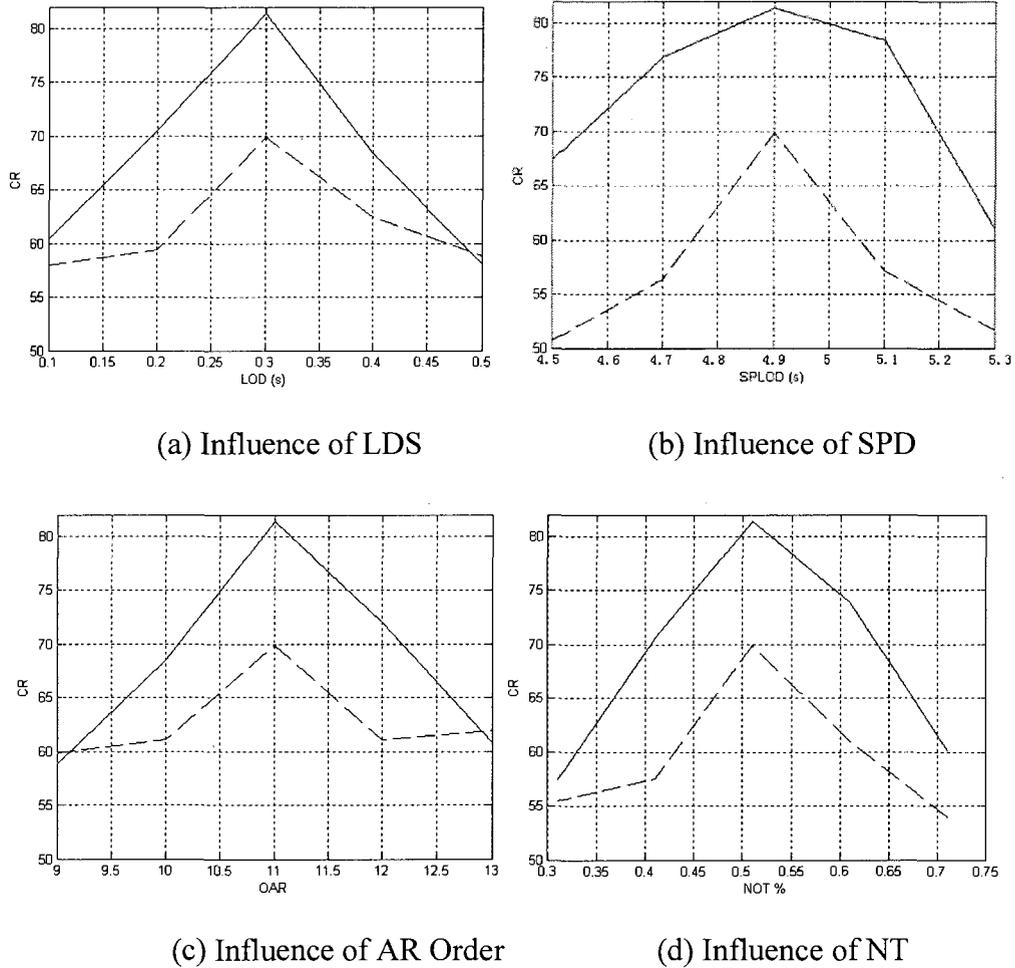


Figure 4.10. The influence of LDS, SPD, Ar Order and NT on CR. The solid traces were from Data Set 1 (tripolar) and the dashed traces were from Data Set 2 (virtual disc).

Another noticeable phenomenon in the pre-feature selection is the subject-dependency of parameters. As was shown in Table 4.4, each subject achieved their highest CR at different LDS, SPD, AR Order and NT, which suggests that the optimum parameters were subject-dependent. In Table 4.3, CR1 and CR2 were the CRs for Data Set 1 and Data Set 2, respectively. The relative variation (RV) for each of the parameters in Table 4.4 was defined by Equations (4.9)

$$RV = \frac{STD}{Mean} \times 100\% . \quad (4.9)$$

Table 4.4 The subject-dependency of the optimum parameter value

Sub.#	CR1	CR2	LDS	SPD	AR Order	NT
1	81.43	69.84	0.30	0.90	11	0.73
2	73.89	59.23	1.70	2.70	13	1.00
3	80.18	61.09	2.50	0.70	13	0.94
4	80.16	70.06	0.70	2.30	13	1.00
5	84.23	72.45	1.70	0.90	12	0.93
6	77.60	60.05	2.30	1.3	10	1.00
7	75.77	71.23	2.0	1.9	11	0.98
8	73.49	69.01	2.0	2.2	11	0.94
9	79.88	73.34	2.3	1.4	13	0.94
10	82.48	71.39	1.9	1.0	12	1.00
11	77.87	70.13	1.7	0.75	13	1.00
12	77.75	68.32	0.9	1.6	12	0.97
Mean of parameter			1.6667	1.47	12	0.95
STD. Of parameter			0.69	0.67	1.04	0.08
Relative Variation%			41.40	45.58	8.67	8.42

4.6.2 Influence of the Application of Tripolar Electrodes

For this work the CR of the two sets of signals from the tripolar concentric electrodes and virtual disc electrodes was compared and the auto search algorithm was performed for LOD, SPD and AR Order to find the factors that generated the highest CR for each data set.

From Table 4.3, the CR for Data Set 1 (tripolar) was $78.7 \pm 3.3\%$ and the CR for Data Set 2 (virtual disc) was $68.0 \pm 5.0\%$. There was a significant difference between the CR of Data Set 1 using the signals from tripolar concentric electrodes compared to the CR of Data Set 2 using signals from the virtual disc electrodes ($P = 2.9 \times 10^{-6}$). This

difference could also be seen from Figure 4.10, where the solid traces (from Data Set 1–tripolar electrodes data) were always higher than the dashed traces (from Data Set 2–virtual disc data).

4.7 Conclusion and Discussion

4.7.1 Improved BCI Classification by Pre-Feature Selection

Just as Millan, *et al.* reported that there was a set of relevant EEG features that best differentiate spontaneous motor-related mental tasks (Millan, et al., 2002a), the present study showed that there was a set of data-segment-related parameters that achieved the best CR for each subject (Cao, Besio and Jones, 2009b). Figure 4.10 shows that the data-segment-related parameters LDS, SPD, AR Order and NT strongly influenced CR. A slight change of these parameter values might generate a great change in CR. Accordingly, proper selection of the four segment-related parameter values generated significantly higher CR (Table 4.2). The average CR generated from Data Set 1 with the individualization of the four data-segment-related parameters was $78.73 \pm 3.3\%$, which was comparable to the improved ARX model which was $79.10 \pm 3.9\%$ (Burke, et al., 2005). However, this significant improvement was achieved without increasing the complexity of the AR model or classification algorithms. It should still be possible to increase the complexity of the signal processing algorithms and further improve the CR.

Furthermore, Table 4.4 shows that the parameters concerned with segments of data were best for the recognition were subject-dependent. This dependence may be caused by variations in reaction rates, concentration and motor imaging capabilities from subject to subject. In Table 4.4, the RV of LDS and SPD were very high, which suggests that those

two parameters were strongly subject-dependent. The RV of the AR Order was relatively small, which may be due to the AR Order partial dependence on the complexity of the waveform (Burke, et al., 2005). Experiments also showed that the AR Order may depend on the number of trials and the length of the data segment that were used to construct the model. Usually, the larger the number of trials and segment length, the higher the AR model order. The low RV of NT can be explained in the following ways. The optimum value of NT depends on two factors. First, how many trials were available and second, how many suitable trials were available. Due to the limited number of trials in the model construction (only 240 for each subject), the selective-search algorithm has difficulty finding the upper limit of NT. However, for seven of the subjects, less than the maximum trials available (1.00) were used, suggesting that the selection of NT was necessary.

Since the value of LDS, SPD, AR Order and NT were all within a limited range, the selective-search algorithm approach was suitable to find the best data-segment-related parameter values for the BCI classification system. However, the segments of the data selected were continuous. If the selected data segments were discontinuous, there might have been a better selection of features and a further affect on the CR.

In this work, we also tested the features of BP (Pfurtscheller, et al., 1997) and PSD (Millan and Mourino, 2003). The results using those features in this work gave similar or even slightly lower CR (averaged 75.5 and 68.0 for Data Set 1 (tripolar) and Data Set 2 (disc) respectively) compared with the AR model results. Therefore, it appears that using AR model coefficients as features was more appropriate for our data. However, no matter which method we used for feature extraction, the data segment selection still greatly influenced the CR.

In conclusion it was found that the LDS, SPD, AR Order and NT had significant influence on CR for the EEG data recorded from the 12 subjects. On this data, the optimal values of these parameters to generate the highest CR in EEG-based BCI did exist. An auto selective searching algorithm was suitable for finding the optimum values. Our results showed that the optimum values of these four parameters were subject-dependent, which suggests that when constructing a model for BCI EEG analysis, subject variances should be considered and the parameters should be individually customized.

Though the conclusions were based on the model analyzed for this work, all the parameters discussed were subject specific. Therefore, our conclusions may be suitable for other models as well. The study also shows that signals from concentric tripolar electrodes generate significantly higher CR than signals from conventional disc electrodes. More subjects should be analyzed in the future to see if the same conclusions were valid. Nonconsecutive data segments should also be analyzed.

Due to time constraints, the length of the segments of data to search was limited. If the incremental step were decreased further the results may change. Moreover, the order in which the data were recorded from the subjects was always the same. Four 120-trial recordings were performed first with the tripolar concentric electrodes and then repeated with the virtual disc electrodes. For future work, signals from more subjects will be recorded with the order in which the electrodes were used randomized.

4.7.2 Improved BCI Classification by Application of Tripolar Electrodes

A noticeable phenomenon was that while using the same processing methods, the CRs of Data Set 1 were significantly higher than those from Data Set 2 (Table 4.2). This

difference may be due to the high signal-to-noise ratio and better spatial resolution of the tripolar electrodes that were used for the EEG signal acquisition of Data Set 1 (Besio et al., 2006b).

Tripolar concentric electrodes have been shown to possess significantly higher signal-to-noise ratio, spatial resolution and less mutual information than conventional electrodes and can realize the Laplacian method from each single electrode (Besio et al., 2006b; Koka and Besio, 2007; Besio, Cao and Zhou, 2008; Cao, Besio and Jones, 2009a, 2009b). It was beneficial that there was a significant improvement in CR between the two data sets – the difference was $10.7 \pm 4.7\%$. The CR of the signals from the tripolar concentric electrodes was significantly better than the signals from the virtual disc electrodes ($P = 2.9 \times 10^{-6}$).

Using the Laplacian of the potentials has been shown to be effective in EEG classification (Babiloni, et al., 2000; McFarland, et al., 1997). To carry out the surface Laplacian, interpolation must be performed on the scalp surface potentials and then the second spatial derivative of the interpolated potentials must be calculated. Performing the interpolation of the potentials and the second spatial derivative of the potentials may be taxing for real-time processing of EEG for BCI applications. Since the tripolar concentric electrodes directly acquire Laplacian potentials and were easily combined with simple math (Eq. 3.9) they may be suitable for use in real-time BCI applications. However, twice as many amplifiers were needed for tripolar concentric electrodes than for disc electrodes.

For this work, only two sensing electrodes were used to acquire the EEG. With what might be termed basic signal processing, CRs comparable to those produced with more complex signal processing were achieved (Penny, et al., 2000). Because signal sources

for imagery were primarily localized to the sensorimotor cortex, clustering concentric electrodes around those areas may produce more useful features and higher CR. It was also possible to perform more complex signal processing on the signals from the tripolar concentric electrodes to increase the CR.

There was evidence that not all of the imagery signals come from a single area (Roland, Larsen and Lassen, 1980). Recently Wang, et al. reported methodology that included coactivated areas of the brain during imagery (Wang, Hong and Gao, 2007). They found that a three conventional electrode configuration over C3, FCz and C4 outperformed a conventional 30 electrode system. They suggest that the signals at FCz act as a reference to derive stronger differences in the left and right signals from C3 and C4. The coactivated areas may have been one reason why McFarland, et al. found that a larger Hjorth-type Laplacian performed better than a smaller configuration. The coactivated area may have been outside of the surface area of the smaller Laplacian configuration. Tripolar concentric electrodes could also be placed over the coactivated areas, as was performed by Wang, et al. to acquire signals from coactivated areas

To sum up, the application of tripolar electrodes in BCI has following influence:

1. The CR, using tripolar concentric electrodes signals, was significantly better than that from virtual disc electrode signals.
2. Improvements in CR comparable to those obtained by Burke, et al. were achieved without performing complex feature extraction and classification algorithms.
3. Each individual had a specific LOD, SPD and AR Order, which gave the best classification accuracy.

4. When building the BCI model for analysis of EEG, it may be beneficial to consider subject variances, with the factors individually customized before feature extraction.

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Conclusions of the Work

5.1.1 Application of Tripolar Electrodes

Tripolar concentric ring electrodes were found to have the following fundamental advantages, compared to disc electrodes

1. Higher signal to noise ratio (SNR);
2. In particular, higher common mode noise rejection;
3. Higher spatial sensitivity;
4. Less mutual information.

Tripolar Electrodes were found to have better separation results over disc electrodes when applied to ICA for EEG signal separation, in terms of both the fidelity of the signals recovered and the CR. This advantage may be a result of several factors:

1. The higher spatial resolution causes the ICA results to be more sensitive to the source spatial distribution and it provides more uncorrelated/less mutual information signals for ICA.

2. The higher signal to noise ratio (SNR), especially higher common mode noise rejection of tripolar electrodes, compared to disc electrodes, reduces the effects of noise.

For a more detailed description, see section 3.4.2, 'Why tripolar electrodes give better ICA'.

The application of tripolar electrodes to BCI for EEG classification demonstrated the following results and advantages over disc electrodes:

1. The CRs from tripolar concentric electrodes signals were significantly larger than those from virtual disc electrode signals.
2. Improvements in CR comparable to those achieved by Burke, et al. were obtained, without performing complex feature extraction and classification algorithms.

For a more detailed description, please see section 4.7.2 'Improved BCI Classification by Application of Tripolar Electrodes.'

5.1.2 Conclusions for ICA

The ICA studies revealed several guidelines for the selection of the number of electrodes when using EEG signals for signal separation:

1. The number of independent sources recoverable by the ICA algorithm developed in this chapter was no more than the number of electrodes used in ICA.
2. More electrodes will slightly improve the separation results when the number of electrodes was greater than or equal to the number of sources, thus the work suggested that the number of electrodes should be one greater than the expected number of independent source signals when processing ICA.
3. The number of electrodes should not be less than the number of sources.

4. Because the number of independent components was not known *a priori*, the number of electrodes should be increased if the number of independent components was equal to the number of electrodes. Otherwise, it was not possible to know whether the number of independent components obtained by the ICA was limited by the number of electrodes or by the number of components in the EEG.
5. The number of electrodes should be slightly larger than the number of independent sources.

The ICA studies revealed that the separated components had the following features:

1. Source signals that function independently (not being controlled by other source signals) will be considered as independent source signals and will be separated as ICA components, even if they were the same kind of signals with the same frequency and amplitude.
2. Signals with exactly the same time series wave pattern, differing only in amplitude, will be considered as identical signal patterns, regardless of how many were present and where they were located.
3. Signals that depend on other sources (i.e. that were linear combinations of other sources) will not be found.
4. Sources with relatively strong signal/amplitude were more readily separated by ICA.

The ICA results demonstrated the following factors related to the effect of signal-to-noise ratio:

1. Common mode noises could not be separated, since they do not differ

spatially. Thus, they have the strongest negative impact on ICA separation.

2. Noises from a single source can be seen as an independent source. Thus the ICA process does not depend on the pathway through which the signal was produced on the electrodes.

5.1.3 Improved Four-Layer head model

In this work, an improved four-layer anisotropic concentric head model was developed. The following was a summary of this four-layer head model:

1. It can calculate the dipole-simulated sources with electrodes (tripolar or disc) placed at any position on the surface of the head.
2. It can calculate the head surface EEG and Laplacian EEG (LEEG) with high speed.
3. The dipole sources could be placed at arbitrary positions within the inner most layer of the head model.
4. The vector moments of the dipole sources can be oriented in any direction with arbitrary amplitude.
5. The radial and tangential conductivity of each layer can be set to adjust for different in vivo conditions.

5.1.4 Pre-feature Selection Method for BCI

Pre-feature selection was the data-segment related parameter selection. The study of this work showed that the pre-feature selection was as important as feature extraction and feature selection. The following conclusions can be made about pre-feature selection:

1. The LDS, SPD, AR Order and NT significantly influenced CR for the EEG data recorded from the 12 subjects;

2. For these data, optimal values of these parameters to generate the highest CR in EEG based BCI did exist;
3. Since all the data-segment-related parameters were within a limited range, an auto selective searching algorithm was suitable for finding the optimum values;
4. Our results showed that the optimum values of these four parameters were subject-dependent;
5. As a consequence of conclusion 4, when constructing a model for BCI EEG analysis, subject variances should be considered and the parameters should be individually customized;
6. Although the conclusions were based on the model analyzed for this work, all the parameters discussed were subject-specific. Therefore, when building the BCI model for analysis of EEG, it may be beneficial to consider subject variances, with the factors individually customized before feature extraction.

5.2 Future Work

This dissertation lays a foundation for the application of tripolar electrodes to a BCI. Further work will be necessary to implement this application, to further study the effects of different electrode types and to further improve the overall functionality. Some of this work was described in this section.

5.2.1 Improvement of Tripolar Electrodes

The electrode size and ring distribution could be better designed. As was shown in Figure 4.3, the diameter of the tripolar electrodes used in the study of BCI classification was 10 mm, while the width of each ring of the electrodes was 1 mm. In the development

of the tripolar rings (Besio, et al., 2006a, 2006b), (Koka and Besio, 2007), the outer concentric ring ranged from 5 to 36 mm in diameter, with the disc and middle ring sized proportionally from 0.4 to 5mm and 2.5 to 10 mm, respectively. These were tested and shown to yield different results. Thus, if more electrode sizes and different ring distributions were tried, the resulting EEG classification might generate better results.

The placement of the electrode positions could also be better designed. In the BCI study in this work, only two electrodes were placed at positions C3 and C4, as shown in Figure 4.3. (b). However, not all the internal activities of the brain were reflected in those two positions, thus, more electrodes with reasonable positions could be tested that may generate better EEG classification.

5.2.2 More on BCI

More feature extraction methods could be studied and employed. In the BCI study in this work, AR model parameters, PSD values and BP values for feature selection were studied. However, many other features have been studied for BCI design, such as amplitude values of the EEG signals, adaptive autoregressive (AAR) parameters, time frequency features and inverse model-based features. A subset of these features may be able to generate better EEG-based BCI classification results.

More classification methods should be tried. Since only two classification tasks were considered in this BCI study, linear discriminant analysis (LDA) and support vector machine (SVM) method and Mahalanobis distance-based non-linear classifiers were used. Many other classifiers could be considered, such as neural networks, nonlinear Bayesian classifiers, nearest neighbor classifiers and combinations of classifiers. Those methods all have their specific advantages over each other. For example, the neural networks were

more suitable for the non-linear complicated EEG signals. Thus, more classification methods should be tried for multi-tasking EEG based BCIs.

Hardware of BCI application should be designed. As described in the beginning of Chapter 4, BCI functions were divided into three parts, signal acquisition, signal processing and classification/translation and the BCI application. This dissertation focuses mainly on the first two parts of the Hardware and Software description BCI. To fully test the BCI, the hardware of the BCI application should also be developed and tested in combination with the former two functional parts.

5.2.3 Improvement of Pre-Feature Selection

Firstly, the pre-feature selection could be applied in other parameter selections. In the study of this BCI concept, it was shown that the pre-feature selection was as important as feature extraction and feature selection. However, due to the feature extraction method and classification algorithm used in this EEG-based BCI, only four data segment related parameters were tested: LDS, SPD, NT and AR Order. If other feature extraction methods and classification algorithms were used, it may be useful to incorporate new data segment-related parameters into the pre-feature selection.

Secondly, the search algorithm for pre-features selection could be improved. Though the auto-searching algorithm developed in this work can automatically search for optimum parameters for the BCI and was easy to use, the length of the searching steps and searching speed were inversely related. Thus, a shorter length of searching steps and high searching speed could not be achieved simultaneously. Better and more complicated algorithms could be developed to achieve faster searching speed and higher accuracy.

APPENDIX

MATLAB CODE FOR ICA AND COMPUTER HEAD MODEL

```
• function [Cnew,rz,ry]=ChangeCoordinates(Cold,Cc,bCold,bCc,bInverse)
%this function rotates the coordinates system around z axis by positive theta=Cc(1),
%then continues to rotate the system around y axis positive phi=Cc(2),
%by doing those two steps changing the original coordinates to the new
%coordinates such that the Z axis pass the point Cc.
%bInverse do the rotation mentioned above inversely, which change the new
%coordinates back to the old ones.
%Cold and Cc could be in spheric or
%cartesian coordinates, which was defined by bCold and bCc respectively.
coordinate(=1,default) or in Cartesian coordinate(=0)
%Cold=[n*3] were the coordinates of the points in the original system
%Cnew=[n*3] were the coordinates of the points in the new system
%rz,ry were the rotation matrix around z and y direction separately.
if nargin<5
    bInverse=0;%the default was in the forward rotation, that was, from old to new
coordinates
end
if nargin<4
```

```

    bCc=1;

end

if nargin<3

    bCold=1;

end

if bCold==1

    Cold(:,2)=pi/2-Cold(:,2);% PHI was the elevation angle from the xy plane, change it to
    from z axis

    [Cold(:,1),Cold(:,2),Cold(:,3)]=sph2cart(Cold(:,1),Cold(:,2),Cold(:,3));

end

if bInverse % see if it was the backward rotation

    %set the change matrix for the rotation of z axis

    ra=-Cc(1);

    rz=[cos(ra) -sin(ra) 0
        sin(ra) cos(ra) 0
        0    0    1];

    %set the change matrix for the rotation of y axis

    ra=-Cc(2);

    ry=[cos(ra) 0 sin(ra)
        0    1    0
        -sin(ra) 0 cos(ra)];

```

```

n=size(Cold,1);

Cnew=zeros(n,3);

for i=1:n

    %rotation around y axis

    Cnew(i,:)=(ry*(Cold(i,:)))';

    %rotation around z axis

    Cnew(i,:)=(rz*(Cnew(i,:)))';

end

else% if bInverse

    %set the change matrix for the rotation of z axis

    ra=Cc(1);

    rz=[cos(ra) -sin(ra) 0
        sin(ra) cos(ra) 0
        0    0    1];

    %set the change matrix for the rotation of y axis

    ra=Cc(2);

    ry=[cos(ra) 0 sin(ra)
        0    1    0
        -sin(ra) 0 cos(ra)];

n=size(Cold,1);

```

```

Cnew=zeros(n,3);

for i=1:n

    %rotation around z axis

    Cnew(i,:)=(rz*(Cold(i,:)'))';

    %rotation around y axis

    Cnew(i,:)=(ry*(Cnew(i,:)'))';

end

end% if bInverse

%%if the to be changed points were given in spheric coordinates, then the
%%return points should also in spheric coordinates

if bCold==1

    [Cnew(:,1),Cnew(:,2),Cnew(:,3)]=cart2sph(Cnew(:,1),Cnew(:,2),Cnew(:,3));

    Cnew(:,2)=pi/2-Cnew(:,2);% PHI was the elevation angle from the xy plane, change it
to from z axis

end

% %set the thita in (-pi,pi)

% n=size(Cnew,1);

% for i=1:n

%   thita=Cnew(i,1);

```

```

% if thita>pi
%     Cnew(i,1)=Cnew(i,1)-2*pi;
% end
% if thita<-pi
%     Cnew(i,1)=Cnew(i,1)+2*pi;
% end
% end

• function (W,S,Erro,iterN)=FixPointAlgor(x,tol,MaxIterN)
%for a random vector  $x=(x_1,x_2,\dots,x_n)'$ ,  $x_i$  were mutually uncorrelated and all have unit
variance==> $E(x*x')=I$ 
%The program was used to get  $W=(w_1,w_2,\dots,w_n)$ , which  $s_i=(w_i)'*x$  was one of the
seperated
%source, while  $W=(w_1,w_2,\dots,w_n)$  was an orthogonal matrix
%MaxIterN was presettet max number of iteration,default value was 100
%tol was the erro
%S=(s1,s2,...) were the seperated coponents

%get the size and initial w0 so that norm(w0)=1
(n,m)=size(x);
w0=zeros(n,1);
w0(1)=1;
%decide the size of W
W=zeros(n,n);

```

```
Erro=zeros(n,1);
iterN=zeros(n,1);
for i=1:n
    w0=zeros(n,1);
    w0(1)=1;
    B=W(:,1:i);
    if i>1%
        w0=w0-B*B'*w0;
    end
    while iterN(i)<MaxIterN
        w1=kurt(x,w0);
        if i>1&&iterN(i)<10
            w1=w1-B*B'*w1;
        end
        w1=w1/norm(w1);
        Erro(i)=abs(abs(w1'*w0)-1);
        if Erro(i)<tol
            break;
        end
        w0=w1;
        iterN(i)=iterN(i)+1;
    end% while iterN<MaxIterN
    W(:,i)=w1;
```

```
end%for i=1:n
```

```
S=W*x/2;
```

- function

```
(PL,PDisc,Pc,PMiddle,POut)=FourLayerHeadModel(DipPos,DipMPos,Ec,r,Layer,eps,yita,Er,bEc,np)
```

```
%This function get the Laplacian electrode and disc electrode Potential
```

```
%generated by a unit bipolar source inside the four layer concentric sphere
```

```
%head model.The dipole moment can be in any direction
```

```
%DipPos was the dipole center coordinates, in spheric
```

```
%coordinates(theta,phi,r)
```

```
%DipMPos was the dipole moment vector, it was from the center of the dipole to the positive of the dipole
```

```
%Ec was the center coordinates vector of the electrodes, was n*3
```

```
% r was the radius of the rings(rout,rmid), r=0 when it was a disc electrode.
```

```
% Layer=(75,71,65,63);%the thickness of each layer,from outer to inner
```

```
% eps=(0.33,0.0042,1,0.33);%the radial conductivity of each layer
```

```
% yita=(0.44,0.0084,1,0.33);%the tangential conductivity of each layer
```

```
%Er was the accuracy that needed
```

```
%bEc specify if the electrode position were given in pheric coordinate(=1,default) or in Cartesian coordinate(=0)
```

```
% np was the number of points that was going to be used for the points on each electrode,default n=60
```

```

% PL was the Laplacian Potential for tripolar electrodes,
% PDisc was potential for disc electrode.
% Pc,PMiddle,POut were the center, middle ring, outer ring potential
dm=(DipMPos(1)^2+DipMPos(2)^2+DipMPos(3)^2)^0.5;%Get the dipole mement
value,
%default n=60
if nargin<10
    np=60;
end
%default electrode position were given in pheric coordinate
if nargin<9
    bEc=1;
end
%default accuracy that needed
if nargin<8
    Er=10^(-3);
end
%default tangential conductivity of each layer
if nargin<7&&nargin>5
    yita=eps;
end
%default radial conductivity of each layer
if nargin<6

```

```

    eps=(0.33,0.0042,1,0.33);
    yita=eps;
end
if nargin<5
    Layer=(75,71,65,63);%the thickness of each layer
end
if nargin<4
    r=(0.9,0.5);%set the tripolar electrodes radius
end
%test if the electrode was tripolar or bipolar
rl=length(r);
if rl==2%tripolar
    r1=r(1);%outer ring
    r2=r(2);%middle ring
end
%get the number of the electrodes
nElectrode=size(Ec,1);

nTotalPoints=nElectrode*(2*np+1);
Pr=zeros(nTotalPoints,1);%the four layer scalp potential of dl dipole-moment values on
2*n+1 points
SurPos=zeros(nTotalPoints,3);%positions of 2*n+1 points
%get the coordinate of the points on the electrode rings

```

```

for i=1:nElectrode

    SurPos((2*np+1)*(i-1)+1:(2*np+1)*(i-
1)+np,:)=GetElectrodePoints(Ec(i,:),r1,1,np);%the ourter ring points

    SurPos((2*np+1)*(i-1)+np+1:(2*np+1)*(i-
1)+2*np,:)=GetElectrodePoints(Ec(i,:),r2,1,np);%the middle ring points

    SurPos((2*np+1)*(i-1)+2*np+1,:)=Ec(i,:);%the disc position

end

%change the coordinates of the points according to the dipole position

SurPos=ChangeCordinates(SurPos,DipPos,1,1);

%get the new coordinates of the dipole moment

DipPosCar=DipPos;%get the Cartician coordinates of the Dip

DipPosCar(:,2)=pi/2-DipPosCar(:,2);% PHI was the elevation angle from the xy plane,
change it to from z axis

(DipPosCar(1),DipPosCar(2),DipPosCar(3))=sph2cart(DipPosCar(1),DipPosCar(2),DipP
osCar(3));

DipMPos=DipMPos+DipPosCar;

(DipMPos(1),DipMPos(2),DipMPos(3))=cart2sph(DipMPos(1),DipMPos(2),DipMPos(3)
);

DipMPos(:,2)=pi/2-DipMPos(:,2);% PHI was the elevation angle from the xy plane,
change it to from z axis

DipMPos=ChangeCordinates(DipMPos,DipPos);

```

```

%change the coordinate system such that the dipole moment will be in the
%x=0 plan
RotAngle(1)=-(pi/2-DipMPos(1));%rotate the coordiantes system around z axis by
thita=RotAngle(1)
RotAngle(2)=0;%no rotation around y axis
RotAngle(3)=1;
SurPos=ChangeCordinates(SurPos,RotAngle);
DipMPos=ChangeCordinates(DipMPos,RotAngle);
%get the tangential and radial components of the dipole source
DipM(1)=DipMPos(3)*cos(DipMPos(2))-DipPos(3);%the radial component
DipM(2)=DipMPos(3)*sin(DipMPos(2));%the tangential component
% DipM(1)=1;
% DipM(2)=0;
% DipM=DipM
%set the position of the dipole to be on the z axis
DipPos(1)=0;
DipPos(2)=0;
%get the potential of the dipole on every points of the ring
Pr=GetFourLayerDP(DipPos,DipM,SurPos,Layer,eps,yita,Er);
for i=1:nElectrode
    POut(i)=mean(Pr((2*np+1)*(i-1)+1:(2*np+1)*(i-1)+np));%get the potential of the
outer ring

```

```

    PMiddle(i)=mean(Pr((2*np+1)*(i-1)+np+1:(2*np+1)*(i-1)+2*np));%get the potential
of the middle ring

```

```

    Pc(i)=Pr((2*np+1)*(i-1)+2*np+1);%get the potential of the center
end

```

```

%calculate the Laplacian potential

```

```

for i=1:nElectrode

```

```

    PL(i)=16*(PMiddle(i)-Pc(i))-(POut(i)-Pc(i));

```

```

    PDisc(i)=mean(Pr((2*np+1)*(i-1)+1:(2*np+1)*i));

```

```

end

```

- function (S,Y,covY,RankS,w0,w1,w2,erro)=ICA(x,TolC,TolF,NF)

```

%this function separate  $x(n,m)=(x_1,x_2,\dots)'$  to  $S(l,m)=(s_1,s_2,\dots)'$ 

```

```

%where n was the number of signals, m was the length of each signal

```

```

%x must be uncorrelated, contain at most one Gauss signal and linearly

```

```

%synthesization of  $s_i$ 

```

```

%TolC was the tolerance of the PCA to determine the rank of the S

```

```

%TolF and NF were the tolerance and max iteration number use in the FixPoint

```

```

algorithm

```

```

%Y was the PCA of x

```

```

%covY was the covariance of Y

```

```

%RankS was the number of necessary principle components

```

```

%w0,w1,w2,Y,RankS were used in  $x=w_0*w_1*(S*w_2, Y(RnkS+1:n,:))'$ ;

```

```

%erro was the erro returned in the ICA calculation, if too big, method failed

```

```
if nargin<4
    NF=100;
end
if nargin<3
    TolF=10^(-5);
end
if nargin<2
    TolC=1;%default was all possible numbers
end
copn=size(x,1);%the number of the signals
meanx=mean(x');%The mean of the received signal has to be zero.
for i=1:copn
    x(i,:)=x(i,:)-meanx(i);
end
%use PCA to get Y
(w0,z,lt)=princomp(x');%x here was x=(x1,x2,...,xn)';
Y=w0'*x;
covY=cov(Y');
DCY=diag(covY);
wc=length(w0);
w1=zeros(wc);
for i=1:wc
```

```
w1(i,i)=1/sqrt(DCY(i));  
  
end  
  
Y=w1*Y;%variance of Y should be 1, but mean should not be zero.  
  
% w1=w1  
% varY=var(Y')  
% covY=cov(Y')  
% meanY=mean(Y')  
  
%calculate the rank of the S  
RankS=0;  
Pvalue=0;  
TotalValue=TolC*sum(DCY);  
while Pvalue<TotalValue  
    Pvalue=Pvalue+DCY(RankS+1);  
    RankS=RankS+1;  
    if RankS==copn  
        Pvalue=Pvalue;  
        break;  
    end  
end  
  
end  
  
Y1=Y(1:RankS,:);
```

```
(w2,S,erro,N)=FixPointAlgor(Y1,TolF,NF);  
  
S=-S;  
  
% %calculate the transform matrix A, S=Ax  
  
% copn=copn  
  
% ERanks=zeros(RankS,copn);  
  
% for i=1:RankS  
  
%   ERanks(i,i)=1  
  
% end  
  
% RankS=RankS  
  
% w0=w0  
  
% w1=w1  
  
% w2=w2  
  
%  
  
% A=w2'*ERanks*w1*w0'
```

REFERENCES

- Aapo, H. and Erkki, O., 1997. A fast fixed-point algorithm for independent component analysis. *Neural Computation* 9, 1483-1492.
- Akaike H., 1974. A new look at the statistical model identification. *IEEE Trans. on Automatic Control* 19, 716-723.
- Aufrichtig, R. and Pedersen, S.B., 1992. Order estimation and model verification in autoregressive modeling on EEG sleep recordings. *Med. and Bio. Soc.* 14, 2653-2564.
- Babiloni F., et al., 1995. Performances of surface Laplacian estimators: a study of simulated and real scalp potential distributions. *Brain Topography* 8, 35-45.
- Babiloni, F., Babiloni, C. and Carducci, F., 1996. Spline Laplacian estimate of EEG potentials over a realistic magnet resonance constructed scalp surface model. *EEG Clin. Neurophysiology* 98, 363-373.
- Babiloni, F., et al., 2000. Linear classification of low-resolution EEG patterns produced by imagined hand movements. *IEEE Transactions on Rehabilitation Engineering* 8, 186-188.
- Ben, H., Bourne, R. and James, W., 1981. Autoregressive estimation of short segment spectra for computerized EEG analysis. *IEEE Transactions on Biomedical Engineering* 28, 630-638.
- Bennett, K. P. and Campbell, C., 2001. Support vector machines: hype or hallelujah. *Sigkdd Explorations* 2, 1-13.
- Besio, W., Cao, H. and Zhou, P., 2008. Application of tripolar concentric electrodes and pre-feature selection algorithm for brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 16, 191-194.
- Besio, W., et al., 2006a. Development of a tri-polar concentric ring electrode for acquiring accurate Laplacian body surface potentials. *Annals of BME* 34, 426-435.
- Besio, W., et al., 2006b. Tri-polar concentric ring electrode development for Laplacian electroencephalography. *IEEE Trans. BME* 53, 926-933.

- Blankertz, B., Curio, G. and Muller, K. R., 2002. Classifying single-trial EEG: towards brain computer interfacing. *Advances in Neural Information Processing Systems* 14, 157-164.
- Bostanov, V., 2004. BCI competition data sets ib and iib: feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram. *IEEE Transactions on Biomedical Engineering* 51, 1057-1061.
- Burges, C. J. C., 1998. A tutorial on support vector machines for pattern recognition. *Knowledge Discovery and Data Mining* 2, 121-167.
- Burke, D. P., et al., 2005. A parametric feature extraction and classification strategy for brain-computer interfacing. *IEEE Trans. Neural Systems and Rehab. Eng.* 13, 12-17.
- Cao, H., Besio W. and Jones, S., 2009a. Improved separability of dipole sources by tripolar versus conventional disk electrodes: a modeling study using independent component analysis. In *Proceeding of Bioengineering Conference, IEEE 35th Annual Northeast*. Cambridge, Massachusetts.
- Cao, H., Besio W. and Jones, S, 2009b. Individualization of data-segment-related parameters for improvement of EEG signal classification in brain-computer interface. Accepted by *Transactions of Tianjin University*.
- Cincotti, F., et al., 2002. Classification of EEG mental patterns by using two scalp electrodes and Mahalanobis distance-based classifiers. *Methods Inf. Med.* 41, 337-41.
- Cincotti, F., et al., 2003. Comparison of different feature classifiers for brain computer interfaces. *The 1st International IEEE EMBS Conference on Neural Engineering*. Capri, Italy
- Congedo, M., Lotte, F. and Lécuyer, A., 2006. Classification of movement intention by spatially filtered electromagnetic inverse solutions. *Phys. Med. Biol.* 51, 1971-89.
- De Munck, J.C., 1988. The potential distribution in a layered anisotropic spheroidal volume conductor. *J. Appl. Phys.* 64, 464-470.
- Delorme, A., et al., 2006. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Meth.* 134, 9-21.
- Desmedt, J.E., Chalklin, V. and Tomberg, C., 1990. Emulation of somatosensory evoked potential, SEP. components with the 3-shell head model and the problem of 'ghost potential fields' when using an average reference in brain mapping. *Electroenceph. Clin. Neurophysiology* 77, 243-258.

- Farino, D. and Cescon, C., 2001. Concentric-ring electrode systems for noninvasive detection of single motor unit activity. *IEEE Trans. Biomed. Eng.* 48, 1326-1334.
- Fattorus, V. and Tilmant, J., 1949. Exploration du champ électrique precordial a l'aide de deux electrodes circulaires, concentriques et rapprochees. *Arch. Mal du Coeur* 42, 452-455.
- Fukunaga, K., 1990. Introduction to statistical pattern recognition. Academic Press Professional, Inc., San Diego, CA, 311-321
- Garcia, G. N., Ebrahimi, T. and Vesin, J.M., 2003. Support vector EEG classification in the Fourier and time-frequency correlation domains. In Conference Proceedings of the First International IEEE EMBS Conference on Neural Engineering. Capri Island, Italy.
- Garrett, D., et al., 2003. Comparison of linear, nonlinear and feature selection methods for eeg signal classification. *IEEE Transactions on Neural System and Rehabilitation Engineering* 11, 141-144.
- Hannan, E.J. and Quinn, B.G., 1979. The determination of the order of an autoregression, *J. of the Royal Statistical Soc.* 41, 190-195.
- He, B., 1999. Brain electrical source imaging: scalp Laplacian mapping and cortical imaging, *Crit. Rev. Biomed. Eng.* 27, 149-188.
- He, B., Lian, J. and G. Li, 2001. High-resolution EEG: a new realistic geometry spline Laplacian estimation technique. *Clin. Neurophysiology* 112, 845-852.
- Hjorth, B., 1975. An on-line transformation of EEG scalp potentials into orthogonal source derivations. *EEG. Clin. Neurophysiology* 39, 526-530.
- Hill, N.J., et al., 2006. Classifying EEG and ECoG signals without subject training for fast BCI implementation: comparison of nonparalyzed and completely paralyzed subjects. *IEEE Trans Neural Syst Rehabil Eng.* 14, 183-186.
- Ince, N.F., Arica, S. and Tewfik, A., 2006. Classification of single trial motor imagery EEG recordings with subject adapted non-dyadic arbitrary time-frequency tilings. *J. Neural Eng.* 3, 235-244.
- Jain, A.K., Duin, R.P.W. and Mao, J., 2000. Statistical pattern recognition : A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 4-37.
- James, V. S., 2005. Independent component analysis. *Encyclopedia of Statistics in Behavioral Science* 2, 907-912.
- Jiruska, P., et al., 2005. Comparison of different methods of time shift measurement in EEG. *Physiol. Res.* 54, 459-465.

- Jonathan, R., et al., 2002. Brain-computer interfaces for communication and control. *Clinical Neurophysiology* 113, 767-791.
- Jung, T.P., et al., 1999. Analyzing and visualizing single-trial event-related potentials. *AdvNeural Inf Process Syst.* 11, 118-24.
- Jung, T.P., et al., 2000. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37, 163-168.
- Jung, T.P., et al., 2001. Analysis and visualization of single-trial event-related potentials. *Hum Brain Mapp* 14, 166-185.
- Kamoussi, B., Liu, Z. and He, B., 2005. Classification of motor imagery tasks for brain-computer interface applications by means of two equivalent dipoles analysis. *IEEE Trans. Neural Syst. Rehabil. Eng.* 13, 166-71.
- Kaper, M., et al. 2004. BCI competition, data set iib: support vector machines for the p300 speller paradigm. *IEEE Trans. Biomed. Eng.* 51, 1073-1076
- Kees, J. S., et al., 1994, Non-linear dynamical analysis of multichannel EEG: clinical applications in dementia and Parkinson's disease. *Brain Topography* 7, 141-150.
- Kennedy, P.R. and Bakay, R.A., 1998. Restoration of neural output from a paralyzed patient by a direct brain connection. *Neuroreport* 9, 1707-1711.
- Koka, K. and Besio, W. G., 2007. Improvement of spatial selectivity and decrease of mutual information of tri-polar concentric ring electrodes. *J Neurosci Methods* 165, 216-222.
- Lal, T. N., et al., 2005. Methods towards invasive human brain computer interfaces. *Advances in Neural Information Processing Systems* 17, 737-744.
- Law, S. K., Nunez, P. L. and Wijesinghe, R. S., 1993. High resolution EEG using spline generated surface Laplacians on spherical and ellipsoidal surfaces. *IEEE Trans Biomed Eng.* 40, 145-153.
- Leigh, R., et al., 2006. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature* 442, 164-171.
- Levine, S., et al., 2000. A direct brain interface based on event-related potentials. *IEEE Trans. Rehabil. Eng.* 8, 180-185.
- Lotte, F., et al., 2007. A review of classification algorithms for EEG-based brain-computer interfaces. *J. Neural Eng.* 4, 1-13.

- Mahalanobis, P.C., 1936. On the generalised distance in statistics. In Proceedings of the National Institute of Science of India 12, Bangalore, India.
- Makeig, S., et al., 2002. Dynamic brain sources of visual evoked responses. *Science* 295, 690-694.
- McFarland, D. J., et al., 1997. Spatial filter selection for EEG-based communication. *Electroenceph. Clin. Neurophysiology* 103, 386-394.
- Millan, J. R., et al., 2002a. A local neural classifier for the recognition of EEG patterns associated to mental tasks. *IEEE Trans. Neural Networks* 13, 678-686.
- Millan, J. R., et al., 2002b. Relevant EEG features for the classification of spontaneous motor-related tasks. *Biol. Cybern.* 86, 89-95.
- Millan, J. R. and Mourino, J., 2003. Asynchronous BCI and local neural classifiers: an overview of the adaptive brain interface project. *IEEE Trans. Neural Syst. Rehabil. Eng.* 11, 159-61.
- Nune, P. L., et al., 1994. A theoretical and experimental study of high resolution EEG based on surface Laplacians and cortical imaging. *EEG Clin. Neurophysiology* 90, 40-57.
- Palaniappan, R., 2006. Towards optimal model order selection for autoregressive spectral analysis of mental tasks using genetic algorithm. *International Journal of Computer Science and Network Security* 6, 153-162.
- Pardey, J., Roberts, S. and Tarassenko, L., 1996. A review of parametric modeling techniques for EEG analysis. *Med. Eng. Phys.* 18, 2-11.
- Parzen, E., 1974. Some recent advances in time series modeling. *IEEE Trans. on Automatic Control* 19, 723-730.
- Penny, W. D., et al., 2000. EEG-based communication: A pattern recognition approach. *IEEE Trans. Rehab. Eng.* 8, 214-216.
- Perri, F., Bertrand, O. and Pernier, J., 1987. Scalp current density mapping: value and estimation from potential data. *IEEE Trans. Biomed. Eng.* 34, 283-288.
- Pfurtscheller, G., et al., 1997. EEG-based discrimination between imagination of right and left hand movement *Electroencephalogr. Clin. Neurophysiology* 103, 642-51.
- Pfurtscheller, G. et al., 1998. Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters. *IEEE Trans. Rehabil. Eng.* 6, 316-25.

- Pfurtscheller, G. and Lopes, F.H., 1999. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin Neurophysiology* 110, 1842-57.
- Pfurtscheller, G., et al., 2000. Current Trends in Graz Brain-Computer Interface. *BCI Research, IEEE Trans. Rehab. Eng.* 8, 216-219.
- Qin, L., Ding, L. and He, B., 2004. Motor imagery classification by means of source analysis for brain-computer interface applications. *J. Neural Eng.* 1 135-141.
- Rissanen J., 1978. Modelling by shortest data description. *Automatica* 14, 465-471.
- Roland, P. E., Larsen, B. and Lassen, N. A., 1980. Supplementary motor area and other cortical areas in organization of voluntary movements in man. *J. Neurophysiology* 43, 118-136
- Scherer, R., et al., 2004. An asynchronously controlled eeg-based virtual keyboard: Improvement of the spelling rate. *IEEE Transactions on Biomedical Engineering* 51, 979-84.
- Schlogl, A., et al., 2005. Characterization of four-class motor imagery eeg data for the bci-competition 2005. *Journal of Neural Engineering* 2, 14-22.
- Schroder, M., Bogdan, M. and Rosenstiel, W., 2003. Automated EEG feature selection for Brain Computer Interfaces. *Proceedings of the 1st International IEEE EMBS conference on Neural Engineering. Capri Island, Italy.*
- Serruya, M.D. and Donoghue J., 2004. Design Principles of a Neuromotor Prosthetic Device. In: Horch, K. W., Dhillon, G. S. (Eds.), *Neuroprosthetics: Theory and Practice.* Singapore, 1158-1196.
- Stastny, J., Sovka, P. and Stancak A., 2003. EEG Signal Classification: Introduction to the Problem. *Radio Engineering* 12, 51-55.
- Vatta, F., Bruno, P. and Inchingolo, P., 2005. Multiregion bicentric spheres models of the head for the simulation of bioelectric phenomena. *IEEE Trans. Biomed. Eng.* 52, 384-389.
- Ventouras, E., et al., 2004. Independent component analysis applied to the P600 component of event-related potentials. In *Proceedings of the 26th Annual International Conference of the IEEE EMBS San Francisco. CA, USA.*
- Vidal, J., 1973. Toward direct brain-computer communication. In: Mullins, L.J. (Eds.), *Annual Review of Biophysics and Bioengineering.* Palo Alto, 157-180.
- Vidal, J., 1977. Real-time detection of brain events in EEG. *IEEE Proceedings* 65. New York, USA.

- Wang, T., Deng, J. and He, B., 2004. Classifying EEG-based motor imagery tasks by means of time-frequency synthesized spatial patterns. *Clin. Neurophysiology* 115, 2744-53.
- Wang, Y., Hong, B. and Gao, X., 2007. Design of electrode layout for motor imagery based brain-computer interface. *Electronics Letters* 43, 557-558.
- Wear, K.A., R.F. Wagner and Garra, B.S., 1995. A comparison of autoregressive spectral estimation algorithms and order determination methods in ultrasonic tissue characterization. *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control* 42, 709-716.
- Wolpaw J. R., et al., 2000. Brain-computer interface technology: A review of the first international meeting. *IEEE Trans. Rehab. Eng.* 8, 164-173.
- Wolpaw J. R., McFarland, D. J. and Vaughan, T. M., 2000. Brain-Computer Interface Research at the Wadsworth Center. *IEEE Trans. Rehab. Eng.* 8, 222-226.
- Wolpaw J. R., et al., 2002. Brain-computer interfaces for communication and control. *Clin. Neurophysiol* 113, 767-91.
- Zhou, H. and Van, O. A., 1992. Computation of the potential distribution in a four-layer anisotropic concentric spherical volume conductor. *IEEE Trans. Biomed. Eng.* 39, 154-158.