ÇUKUROVA UNIVERSITY
INSTITUTE OF BASIC AND APPLED SCIENCES

PhD THESIS

Nuri Fırat İNCE

ANALYSIS AND CLASSIFICATION OF EEG WITH ADAPTED WAVELETS AND LOCAL DISCRIMINANT BASES

DEPARTMENT OF ELECTICAL AND ELECTRONICS ENGINEERING

ADANA, 2005

ABSTRACT

PhD THESIS

ANALYSIS AND CLASSIFICATION OF EEG WITH ADAPTED WAVELETS AND LOCAL DISCRIMINANT BASES

Nuri Fırat İNCE

DEPARTMENT OF.ELECTRICAL AND ELECTRONICS ENGINEERING.
INSTITUTE OF BASIC AND APPLIED SCIENCES
UNIVERSITY OF ÇUKUROVA

Supervisor: Yrd. Doç. Dr. Sami ARICA
2nd **Supervisor:** Prof. Dr. Pekcan UNGAN
Year: 2005**, Pages:** 99
Jury: Prof. Dr. Yakup SARICA
Doç. Dr. Caner ÖZDEMİR
Doç. Dr. Turgut İKİZ
Yrd. Doç. Dr. Ali KOKANGÜL

Electroencephalogram (EEG) can be used as a strategic tool in establishing communication and control between handicapped people and their environment. The so called “Brain Computer Interface” (BCI) is constructed by analysis and classification of specific patterns in the ongoing EEG which are induced without any need of muscular act. Motor Imagery has similar behavior and can be used as a strategy in the construction of BCI. Motor Imagery induced EEG patterns also have strong relationship to the real performance of the event. Several methods such as band power and autoregressive model parameters were used to analyze and classify the single trial EEG for a BCI task. Most of these methods used fixed time points or frequency indexes. However the movement EEG is non-stationary and contains subject specific patterns. Therefore it is crucial to extract local and subject depended information in an automated manner. In this work an adaptive time-frequency approach is investigated to analyze and classify real and imagery hand movement EEGs. At first movement EEG is adaptively divided in time axis by using the Best Base (BB) approach which is obtained from the Local Cosine Packets by entropy minimization. BB has captured time varying properties of the signal by adjusting analysis segments where it is assumed they correspond to physiological states. In the latter case a modified version of Best Base algorithm, “Local Discriminant Bases” (LDB) were used to extract subject specific time-frequency features in an automated manner for classification of left and right hand movement imagery. Unfortunately this method suffers from the lack of translation invariance and causes high dimensionality. Therefore several feature extraction and dimension reduction methods such as modified mel-scale, principal component analysis and spin cycle procedures are applied to improve the classification performance. One of the interesting results is the difference of the adaptive segmentations and feature characteristic of both hemispheres. The algorithm did not only adapt to time and frequency but also to space. As a further step the number of electrodes is increased to benefit from the different cortical areas. Accordingly our algorithm has promising success for the BCI technology.
Key Words: EEG, Brain Computer Interface, Movement Imagery, Adaptive Time-Frequency Analysis.

ÖZ

DOKTORA TEZİ

UYARLANABİLİR DALGACIKLAR VE YEREL AYRIMCI TABANLARLA EEG ANALİZİ VE SINIFLAMASI

Nuri Fırat İNCE

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ BÖLÜMÜ

Danışman : Yrd. Doç.Dr. Sami ARICA
İkinci Danışman : Prof. Dr. Pekcan UNGAN
Yıl : 2005**, Sayfa:** 99
Juri : Prof. Dr. Yakup SARICA
Doç.Dr.. Caner ÖZDEMİR
Doç. Dr. Turgut İKİZ
Yrd. Doç. Dr. Ali KOKANGÜL

Elektroensefalogram (EEG) son yıllarda engelli kişilerin çevreleri ile iletişim ve kontrol kurmalarında önemli bir araç haline gelmiştir. Beyin Bilgisayar arabirimi (Brain Computer Interface, BCI) olarak adlandırılan, EEG de herhangi bir kas işlevi gerektirmeden oluşan belirli birtakım imgelerin analizi ve sınıflaması ile oluşturulmaktadır. Hareketin hayali de böylesi bir davranışa sahiptir ve BCI için bir strateji olarak kullanılabilir. Hareketin hayali ile EEG de oluşan imgeler aynı zamanda gerçek hareketle de güçlü bir bağ içerir. Tek bir kayıtta EEG nin BCI amaçlı analizi ve sınıflanması için bir çok metod kullanılmıştır. Genellikle bu metodlar sabit zaman ve frekans noktalarını kullanmaktadırlar. Halbuki EEG durağan değildir ve kişiye özgü imgeler içerir. Bu nedenle lokal ve kişiye bağlı bilginin otomatik olarak çıkarılması büyük önem taşır. Bu tezde gerçek ve hayal edilen el hareketlerine ait EEG kayıtlarını analiz etmek ve sınıflayabilmek için uyarlanabilir bir zaman-frekans yaklaşımı araştırılmıştır. İlk olarak gerçek hareket EEG si, Yerel Kosinüs Paketlerinden entropinin minimize edilmesi ile elde edilen, en iyi tabanlar (Best Bases, BB) algoritması ile zamanda uyarlanabilir olarak bölütlenmiştir. BB analiz bölütlerini ayarlayarak zamanla değişen özellikleri yakalamıştır ki bu bölütlerin fizyolojik duruma karşılık geldiğine inanılmaktadır. Bir sonraki adımda BB nin değiştirilmiş bir versiyonu olan Yerel Ayırımcı Tabanları (Local Discriminant Bases, LDB), sağ ve sol el hareket hayallerini sınıflama için, kişiye özgü nitelikleri zaman ve frekans bölgesinden otomatik olarak çıkarmıştır. Ne yazık ki bu metodun kaydırma karşı değişmezliği yoktur ve yüksek boyutlu uzaya neden olmaktadır. Bu nedenle Mel-skalası, temel bileşen analizi (Principal Component Analysis, PCA) ve devirli kaydırma (Spin Cycle) gibi özellik çıkarıcı ve boyut azaltıcı işlemler sınıflama başarısını artırmak için kullanılmıştır. İlginç sonuçlardan biri uyarlanabilir bölütlemenin ve özelliklerin kişiler arasında ve de aynı kişide her iki hemisferde farklı olmasıdır. Algoritma sadece zaman ve frekansa değil aynı anda uzaya da uyarlanmıştır. Daha ileri bir adım olarak elektrodların sayısı diğer kortikal bölgelerden yararlanmak için artırılmıştır. Sonuç olarak algoritma BCI teknolojisi için umut verici başarıya sahiptir.

Anahtar Kelimeler: EEG, Beyin Bilgisayar Arabirimi, Hareketin Hayali, Uyarlanabilir Zaman-Sıklık Analizi.

CONTENTS **PAGE NUMBER**

ACKNOWLEDGEMENTS

First of all I would like to thank my advisor Assit. Prof. Dr. Sami Arica for his support, useful suggestions, self-sacrifice and friendship.

Since my thesis was supported by The National Scientific Research Council of Turkey (TUBITAK), I had the chance to work with many institutes around world. The introductory part of this thesis is developed at Hacettepe University Medical School, Department of Biophsics,Turkey and Graz Technical University Department of Medical Informatics. Therefore I would like to thank my second advisor Prof. Dr. Pekcan Ungan and Prof Dr. Gert Pfurtscheller and their staff for their help and support.

I am thankful to Prof. Dr. Erdal Panayirci who has co-operated with us in order to start this international project and who has never grudged his support from me.

A big portion of this thesis is completed at the University of Minnesota USA under the guidance of Prof. Dr. Ahmed Tewfik. I want to thank him for his support, encouragement and useful discussions. I was influenced by his philosophy. I was very fortunate to have an opportunity to work with him.

Many thanks to Prof. Dr. Hamit Serbest who has always been there for me. Without his guidance and support we would not have been able to start this international cooperate project.

On personal side, especially I would like to thank to my wife Meryem İnce. I could not have completed this thesis without her support.

LIST OF TABLES **PAGE NUMBER**

LIST OF FIGURES **PAGE NUMBER**

ABREVIATIONS

AAR	Adaptive Autoregressive
ALS	Amyotrophic Lateral Sclerosis
AR	Autoregressive
BB	Best Base
BCI	Brain Computer Interface
BF	Band Filtering
BP	Band Powers
CART	Classification and Regulation Trees
CDF	Cumulative Distribution
CP	Cosine packets
CS	Class Separability
CSP	Common Spatial Patterns
D	Euclidean Distance
DA	Decision Algorithm
DCT	Discrete Cosine Transform
DP	Discrimination Power
DSLVQ	Distinction sensitive learning vector quantization
ECoG	Electrocorticogram
EEG	Electroencephalogram
EMG	Electromyogram
EP	Evoked Potential
ERD	Event Related Desynchronization
ERP	Event Related Potentials
ERS	Event Related Synchronization
F	The Fisher Criterion
G1	Hand Movement Imagery Subject 1
H	Hellinger Distance
J	Kullback Leibler Distance
k-NN	k^{th}-Nearest Neighbor

KLT	Karhunen Loeve Transform
LCP	Local Cosine Packets
LDA	Linear Discriminant Analysis
LDB	Local Discriminant bases
LMS	Least Mean Squares
LVQ	Learning Vector Quantization
M1	Finger Movement Subject 1
M2	Finger Movement Subject 2
M3	Finger Movement Subject 3
ME	Movement Execution
MI	Movement Imagery
MLP	Multilayer Perceptron
NC	Number of Coefficients
PCA	Principal Component Analysis
PDF	Probability Densities
RLS	Recursive Least squares
S1-9	Imagery Subjects 1 through 9
SCPs	Slow Cortical Potentials
SMA	Supplementary Motor Area
SNR	Signal to Noise Ratio
Sp3- BF	3 spin cycle BF
STD	Individual Cosine Packets
STFT	Short time Fourier transform
T-F	Time-Frequency
TFW	Time Frequency Weighting
WP	Wavelet Packet

1. INTRODUCTION

Electroencephalogram (EEG) is defined as the electrical potential measured on scalp and it reflects the brain activity. This noninvasive signal has helped in understanding the brain functions and is used widely in diagnosis and during therapy. Many times a stimulus is used to induce amplitude changes in the EEG which are called Event Related Potentials (ERP). Since these potentials are phase locked they can be averaged across trials to remove the random variations. Evoked potentials are very good examples of the phase locked family of potentials. The neural potentials do not always occur in a phase locked manner. As known, the human EEG can be decomposed into 5 different bands. The delta (δ: 0-4Hz), theta (θ:4-8Hz), alpha (α: 8-13Hz), beta (β: 14-30Hz) and gamma (γ>30Hz) bands. These rhythmic components can be modulated due to an event in a time locked manner. First, Berger reported such changes as the decrease of alpha (α) activity upon opening eyes (Berger 1939). When amplitude decrease of rhythmic activity is short lasting it is called "Event Related Desynchronization" (ERD) (Lopes 1999; Neuper 2001). Inversely, an increase in rhythmic EEG activity is called "Event Related Synchronization" (ERS) (Lopes 1999; Neuper01). It is assumed that the ERD and ERS reflect the activation of the underlying neural circuit on the measurement space. When the neural circuit is activated then the synchrony between neurons is decreased where it is reflected as ERD. In the opposite case where the neural circuit is deactivated the neurons start to have a coherent activity, which in turn induces ERS (Neuper 2001).

1.1 Brain Computer Interface and Motor Imagery

Besides diagnosis and exploring brain functions, it has been shown that the brain electrical activity can be used as a tool to help handicapped people to interact with the outer world. The so called "Brain Computer Interface" (BCI) is constructed to translate the brain electrical activity for communication and control (Wolpaw 2000). This translation is accomplished without the need of any muscular act. Since this new communication channel does not depend on peripheral nerves or muscles, it can be used by people with severe motor disabilities. BCIs can allow patients who are totally paralyzed (or 'locked in') by amyotrophic lateral sclerosis (ALS), brainstem stroke or other neuromuscular diseases to express their wishes to the outside world.

A BCI basically can be decomposed into Data Acquisition, Signal Processing - Classification and Feedback/Actuator units (see Figure 1.1.). The feedback unit can be a robotic arm, a neuroprosthesis or a monitor which allows selecting words to answer questions. There are different brain signals which can be used as input to a BCI: The signals can be split into two groups as being non-invasive, such as evoked potentials (EP) (Sutter 1990; Farwell 1988, McMillan 1995; Polikoff 1995), slow cortical potentials (SCPs) (Birbaumer 1981; Kotchoubey 1997; Kuebler 1998), oscillatory EEG components (Pfurtscheller 1994; Pfurtscheller 1999; Wolpaw 1994; Anderson 1998; Penny 1999) or invasive, such as ECoG (Levine 2001; Leuthardt 2004), single neuron activities (Chapin 1999; Nicoleilis 2001; Wessberg 2002) and local field potentials (Scherberger 2002).

Noninvasive BCIs use EEG activity recorded from the scalp. They are convenient, safe and inexpensive, but they have relatively low spatial resolution, are susceptible to artifacts such as electromyographic (EMG) signals, and often require extensive user training. Invasive BCIs have higher spatial resolution and might provide control signals with many degrees of freedom. However, BCIs that depend on electrodes within cortex, face substantial problems in achieving and maintaining stable long-term recordings. See (Thakor 2004) for a review of recording advances. The small, high-impedance recording sites make penetrating electrodes susceptible to signal degradation due to encapsulation. Also, small displacements of the tiny penetrating electrodes can move the recording sites away from the cortical layers. Besides these invasive methods require removal of the skull, which has several risks. This issue is a crucial obstacle that currently prohibits their clinical use in humans. Therefore almost always EEG is preferred.

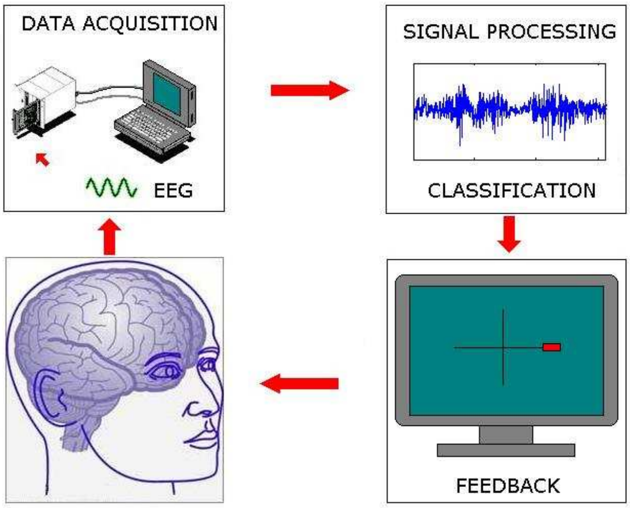


Figure 1.1. Block Diagram of a simple BCI system

As stated before, several types of EEG based BCIs exist. EP based systems use brain potential changes elicited by specific stimuli and identified by latency, amplitude and location. For example in a P300-based BCI (Farwell 1988) the authors presented a 6×6 matrix of letters in which each column and row was flashed in a sequence. The subject looked at the letter he wanted to select. When the desired row and column flashed, a P300 component was detected and the letter was selected.

In (Kotchoubey 1997) it has been demonstrated that healthy subjects and patients can learn to reliably control their slow cortical potential (SCP) amplitudes at the vertex (electrode position C_z, according to the international 10-20 EEG system). SCPs enable a patient to operate an electronic spelling device (Kuebler 1998). See (Guger 1999) for a detailed description of available systems.

Especially the Graz group uses Motor Imagery (MI) based oscillatory changes (Pfurtscheller 2001). For MI tasks the subjects are instructed to imagine themselves performing a specific motor action without overt motor output. The underlying basis of using the MI as a BCI strategy is that unilateral hand movement imagery results in a contra-lateral ERD close to the primary motor areas and, in certain cases, in a simultaneous ipsi-lateral ERS of sensorimotor rhythms (Pfurtscheller 2001; Pfurtscheller 1996; Pfurtscheller 1997; Pfurtscheller 1998; Neuper 1999). This behavior is also observed during the preparation and planning stages of the real movements. Therefore in the last years analysis and classification of real and imagery movements have gained significant interest. Positron emission tomography and functional magnetic resonance imaging studies suggest that, cortical sensorimotor systems are activated during MI (Jeannerod 1994). Other experiments have demonstrated that, the supplementary motor area (SMA), prefrontal area, premotor cortex, cerebellum, and basal ganglia are activated during both movement execution and imagery (Ersland 1996; Decety 1990). Several EEG studies also further confirm the notion that MI can activate primary sensorimotor areas (Beisteiner 1995; Lang 1996).

As a result it has been found that planning and execution of movement leads to a short-lasting and circumscribed attenuation/enhancement in the mu (8–12Hz) and the central beta (13–28Hz) rhythm known as event-related (de-)synchronization (ERD/ERS) which have played an important role in BCI studies (Pfurtscheller01). Such simultaneously attenuated and enhanced EEG rhythms can be used to classify brain states related to the planning or even imagination of different types of limb movements. In this thesis, the use of ERD /ERS structures, induced by movement execution (ME) and MI will be investigated as well.

1.2 Movement ERD & ERS

The functional role of EEG rhythms and motor behavior is explored in detail. It has been shown that a motor task can induce ERD and ERS on motor cortex. Figure 1.2. shows the EEG sweeps accompanying finger movements as recorded from C3 electrode, from a standard 10-20 system. In a motor task, such as a hand/finger movement, rhythmic components show interesting patterns. The following results are reported:

- ERD starts around 2sec. prior to the onset of self-paced movement on the contra-lateral hemisphere (Lopes 1999; Neuper 2001).
- ERD occurs shortly after on the ipsi-lateral hemisphere.
- During movement, ERD becomes almost symmetric on both sides.
- EEG recovers after movement offset, slowly, in a few seconds, in the α (7-13Hz) band and relatively fast, in the β (14 – 30Hz) band. The characteristic

of β band activity is not only a fast recovery, but also a burst activity (Lopes 1999; Neuper 2001; Pfurtscheller 1996).

Also three different functional levels can be distinguished in the ongoing EEG during a voluntary movement (Lopes 1999).

a) Preparation (planning and programming) of movement: Prior to Electromyogram (EMG) onset contra-lateral α and β ERD.
b) Execution of movement: Bilateral symmetric α and β ERD.
c) Recovery from movement after EMG onset: Contra-lateral preponderant β ERS.

All of these findings indicate that both ERD and ERS can be characterized with their magnitude, latency, reactive frequency band and spatial localization (Lopes 1999).

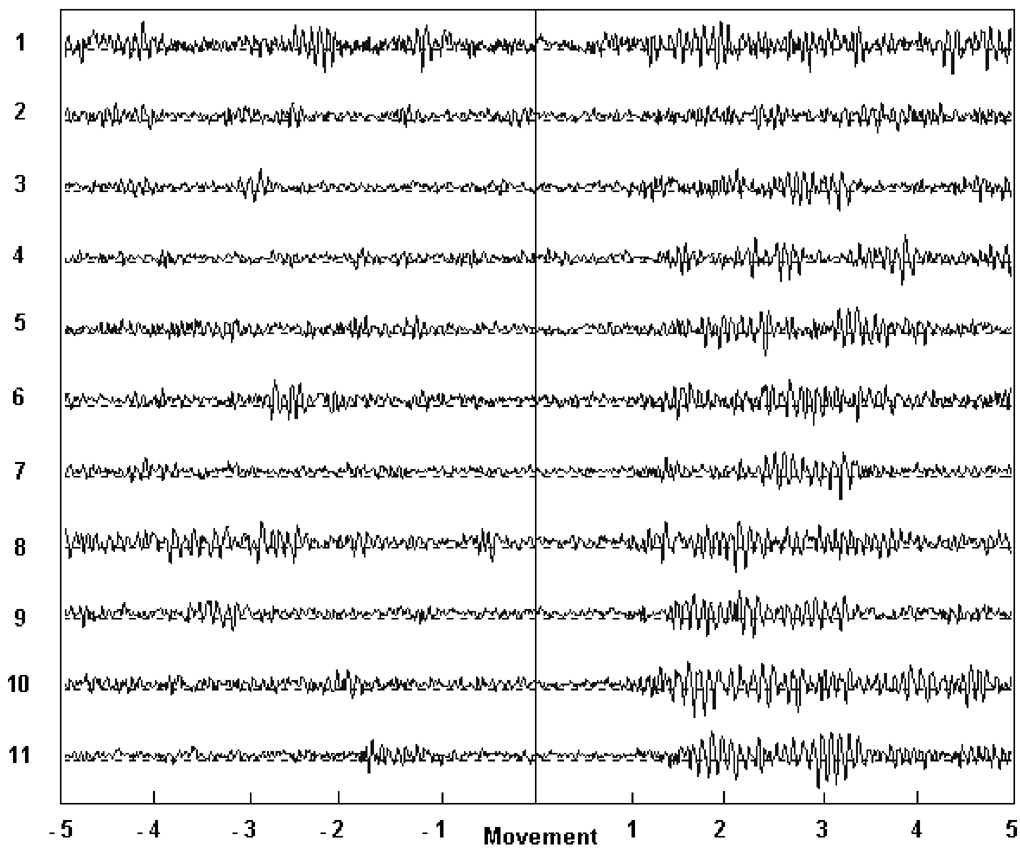


Figure 1.2. EEG sweeps of C3 electrode recorded during right hand movement. The vertical line indicates the onset of the movement. Notice the absence of oscillations right before, during and after the onset

1.3 Signal Representation, Feature Extraction and Classification

In nature the signals which we see such as “images” and hear such as “sound, voice”, as well as EEG, have high dimensionality and contain temporal information. Our brain accomplishes structural information processing to extract characteristic properties of the observed system and maps them to pre-known categories where these steps correspond to feature extraction and classification. Therefore Pattern Recognition can be seen as a combination of these two important steps. Instead of concentrating on the details of the original space the signal is reduced to a small subset that represents the vital information (*dimension reduction*). Also in practical pattern recognition the dimension reduction is a crucial step to improve the generalization of the classifier, which is the capability of mapping unseen dataset successfully to the desired values by using the knowledge, which is obtained from the *training/learning* set.

If we look at the literature several methods have been proposed such as: Fast Fourier Transform (Duhamel 1990) to explore the harmonic components of the signal, Principal Component Analysis and Discrete Cosine Transform (Akansu 1993; Rao 1990) for signal compression, Linear Discriminant, Neural Networks and Decision Trees (Duda 1973; Bishop 1995; Thedoridis 2003) for classification. None of these methods, alone, can accomplish the important steps of pattern recognition, which were defined previously. Therefore signal processing can be seen as a large toolbox and one needs to find the right tool(s) to solve the problem at hand. Many times the combination of different approaches can provide the desired solution.

Besides this, in the case of the important features appearing as a transient phenomenon, one will face the inability to capture local information with any of the previous methods. In order to overcome this problem, it is a good choice to analyze the signal in time-frequency plane (T-F). The nature of EEG contains local structures and the investigation of EEG in T-F plane can be efficient. As indicated previously the ERD/ERS are characterized with latency (time), reactive band (frequency) and electrode location (space). These tree parameters are very important in the quantification and classification of MI/ME EEG. In this thesis a series of methods will be applied and combined for local feature extraction, dimension reduction and classification to be able to use the oscillatory activities of the ME and MI EEG data for a BCI task.

The use of oscillatory components in the ongoing EEG induced by MI for a BCI task is hard to accomplish due to low signal to noise ratio (SNR), high dimensionality and sweep to sweep variations of EEG. Also, obtaining mentally induced EEG patterns requires extensive training. Therefore, as a natural choice the analysis of real movement is a good introductory step since it has correlations with the imagery and ERD/ERS patterns and is easily obtained from the EEG during ME.

Based on the physiological findings which were reported previously, besides conventional methods, first the movement EEG recorded from 3 subjects will be investigated by time-frequency approaches. It will be shown that the adaptive signal representation algorithms, such as best basis approach obtained from Local Cosine Packets, can efficiently extract subject depended features. Then in the following chapters the algorithms will be updated to classify single sweep MI EEG recordings for a BCI task on a publicly available dataset by using Local Discriminant Bases algorithm. How the developed methods can be used to classify and visualize discriminant structures in MI and ME tasks will also be shown. The Local Discriminant Bases algorithm provides an adaptive T-F segmentation to maximize the discrimination between left and right ME/MI tasks. This approach will be extended to space-time-frequency analysis for classification and topographical mapping of brain activity. As a summary this thesis will explore the adaptation in time, frequency and at last in space to quantify and classify MI/ME EEG data.

2. PREVIOUS STUDIES

2.1 Quantification of ERD and ERS

Since ERD/ERS are non-phase locked activities, EEG data can not be averaged directly to reduce the effect of random variations. A number of different methods are offered for the quantification of ERD and ERS. These methods are well summarized in (Pfurtscheller 1999c). The most preferred ones are:

- Band Pass Filtering and Squaring,
- Autoregressive (AR) Modeling,
- Hilbert Transform.

The Band Pass filtering and the AR parameter approaches are the most used ones due to their simplicity and ease of implementation. The Hilbert transform is the improved version of Band Pass filtering method where the ripples of the envelopes of the reactive bands are reduced. In the following section these methods will be summarized shortly.

2.1.1 Band Pass Filtering and Squaring

Band pass filtering and squaring is the most commonly used method in quantification of synchronization and desynchronization in EEG. In order to quantify ERD and ERS at least 30 trials are needed to satisfy the statistical reliability. And the length of each trial should be at least several seconds long. However this method requires prior knowledge of the reactive bands. The reactive bands are generally selected by comparing two 1 second spectra before and after movement onset. These segments should correspond to the physiological events of the motor task as defined previously. The bands which show significant difference in the comparison stage are selected for the filtering process. The block diagram of the band power method is given in Figure 2.1.

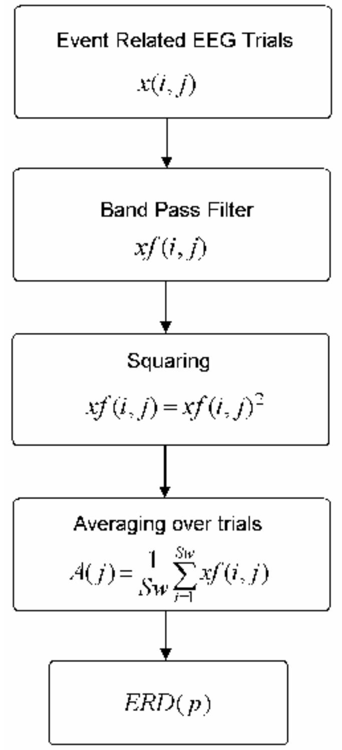


Figure 2.1. The Block Diagram of "Band-Pass Filtering" Approach

Here $x(i,j)$ is the jth sample of the ith trial. $ERD(p)$ is the operation to calculate the percentage ERD with respect to a reference. Here p stands for percentage. The reference R is calculated as the average power of an interval before the trigger signal. The interval should be at least a few seconds in order to satisfy the statistical stability.

$$R=\frac{1}{k}\sum_{j=n_0}^{n_0+k} A(j) \qquad (2.1)$$

where R is the average power, n_0 is the start index of reference interval, k is the length of reference interval and $A(j)$ is the power at the jth sample.

$$ERD(j)=\frac{A(j)-R}{R}\times 100\% \qquad (2.2)$$

Here a negative percentage value indicates ERD and a positive one ERS.

2.1.2 ERD Quantification by Hilbert Transform

The envelope of the band pass filtered signal, $s_f(t)$, is obtained by using the Hilbert transform,

$$h(t) = s_f(t) * \frac{1}{\pi t} \ , \tag{2.3}$$

where $*$ is the convolution operator. This can be calculated as

$$H(f) = S_f(f)(-j)sign(f) \ . \tag{2.4}$$

The envelope is given by

$$m(t) = \sqrt{s_f^2(t) + h^2(t)} \ . \tag{2.5}$$

Figure 2.2 shows the steps for determining the envelope by using the Hilbert transform.

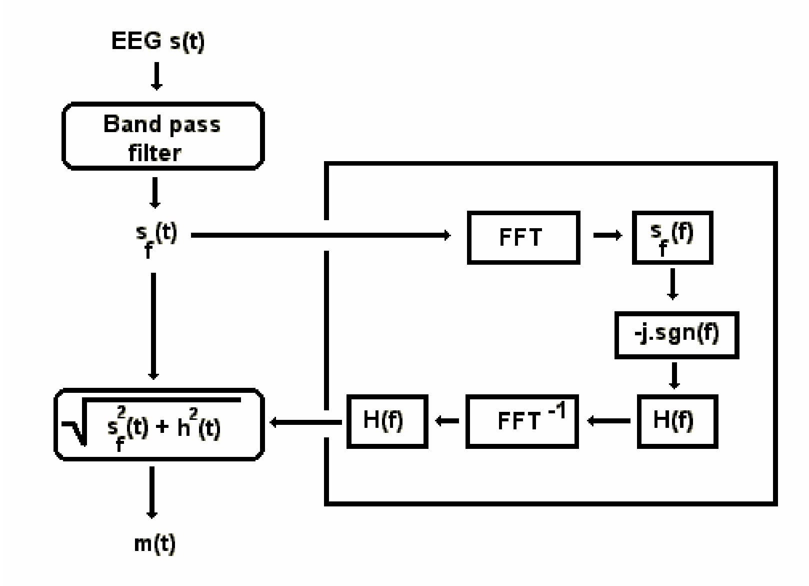


Figure 2.2. The Block Diagram of Hilbert Transform Approach

As stated previously except for the AR model all quantification methods are based on band pass filtering. Therefore, it is important to determine upper and lower limits of frequency bands. In order to find these upper and lower limits, comparison of two spectra is used. The most widely used method to analyze the underlying spectrum is the Fourier transform.

$$x(f) = \int_{-\infty}^{\infty} x(t) e^{-2j\pi ft} dt \tag{2.6}$$

The coefficients of $|x(f)|^2$ represent the energy distribution in the frequency of the entire signal.

2.1.3 Autoregressive Modeling

The EEG records accompanying event related trials have non-stationary characteristics. In order to investigate their time varying properties they are divided into segments where local stationarity is assumed. Then the segments are analyzed with an autoregressive model of order p. The model is,

$$x_t = a_1 x_{t-1} + ... + a_p x_{t-p} + e_t \ , \tag{2.7}$$

where x is the output sequence and e_t is the white noise. Then regression coefficients of the equation are found by solving the Yule-Walker equations (Hayes 1996 and Marple 1989). The estimated coefficients define the power spectrum of the EEG.

2.2 Attempts to Classify Single Trial EEG for a BCI Task

Since MI shares similar structures to ME previously mentioned approaches are already suitable to extract features. Initially band powers (BP) of reactive bands are used in combination with learning vector quantization (LVQ). These features are obtained by filtering and squaring the EEG in fixed windows. These windows were selected manually where it is assumed that they correspond to reactive time locations (Kalcher 1993). Besides representing the signals, it is important to find frequency and time indexes, which may yield better classification. This problem was solved by distinction sensitive learning vector quantization (DSLVQ), which is an improved learning vector quantization (LVQ) (Prezenger 1999). This method allowed selecting the most relevant frequency bands in an automated manner by using fixed time windows. It has been reported that these bands are subject specific (see Figure 2.3.). Auto regressive (AR) modeling is used as another approach which attempts to define the dynamic spectral information by its parameters (Anderson 1995; McFarland 1997). In general, the signal is segmented into short intervals where local stationarity is assumed. Then in each segment AR parameters are calculated and fed to a classifier. In (Schlögl 1997) adaptive autoregressive (AAR) modeling is used in combination with a linear discriminant to select the best time point for classification. The AR parameters are calculated for each time point by using either Least Mean Square (LMS) or Recursive Least Square (RLS) algorithm. AAR is also found to be suitable for online application where no buffering of EEG data is required. Searching for the optimal time point for classification is a challenge in this approach. The author reported that there exist different time points for each subject, where minimum error is obtained (see Figure 2.4.).

Both DSLVQ and AAR algorithms have attempted to find the frequency bands or time points, which have maximum influence on the detection of ERD/ERS. The two methods emphasize the importance of selecting the optimal T-F points and the subject dependency of this selection. Therefore an algorithm which classifies the mental states of human subjects should be able to set parameters adaptively through training. Following sections will be based to solve this problem by introducing an adapted time frequency approach which can extact subject specific features.

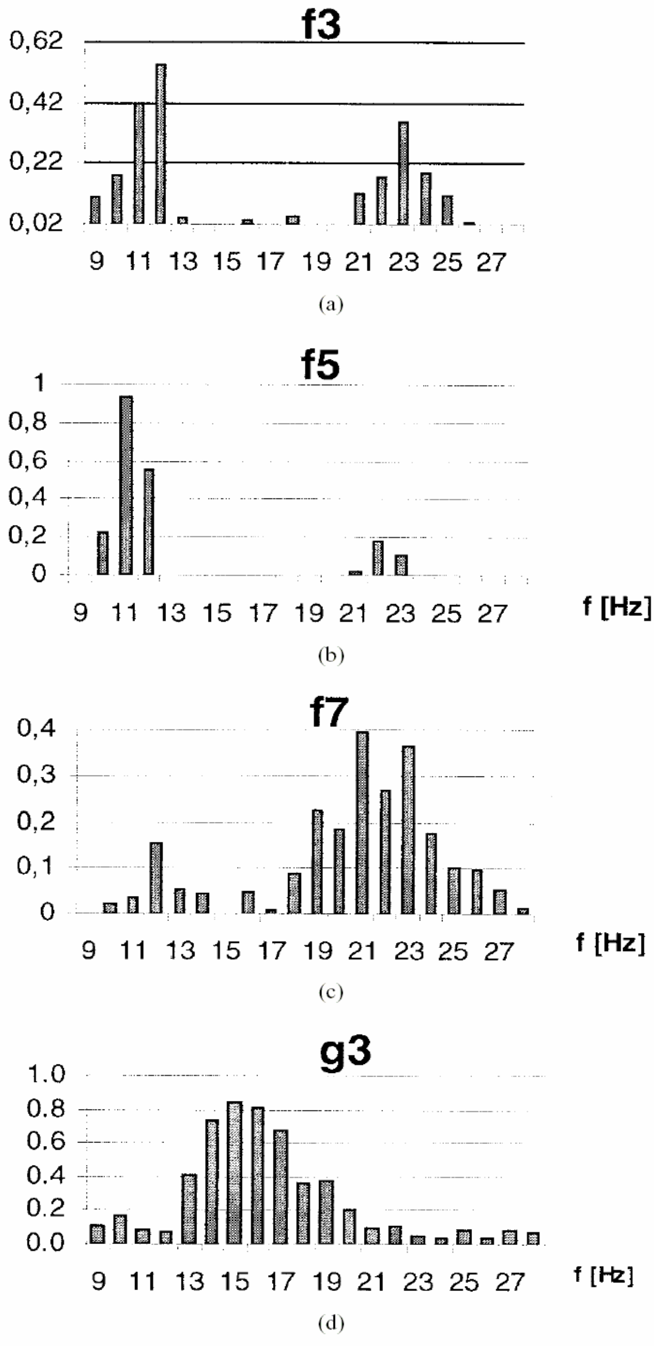


Figure 2.3. The relevant 1Hz frequency bands of 3 subjects obtained via DSLVQ algorithm. For subject f3, highest relevance of two frequency bands (9–13Hz and 21–26Hz) can be seen; for subject f5 slightly different bands (10–12Hz and 21–23Hz) are most relevant. For subject f7, a broad beta band from 18 to 26Hz shows highest relevance; less pronounced is a narrow band around 12 Hz. For subject g3 a wide frequency band centered at 15 Hz is of highest relevance (13–19Hz). (Modified from (Pfurtscheller 1998a))

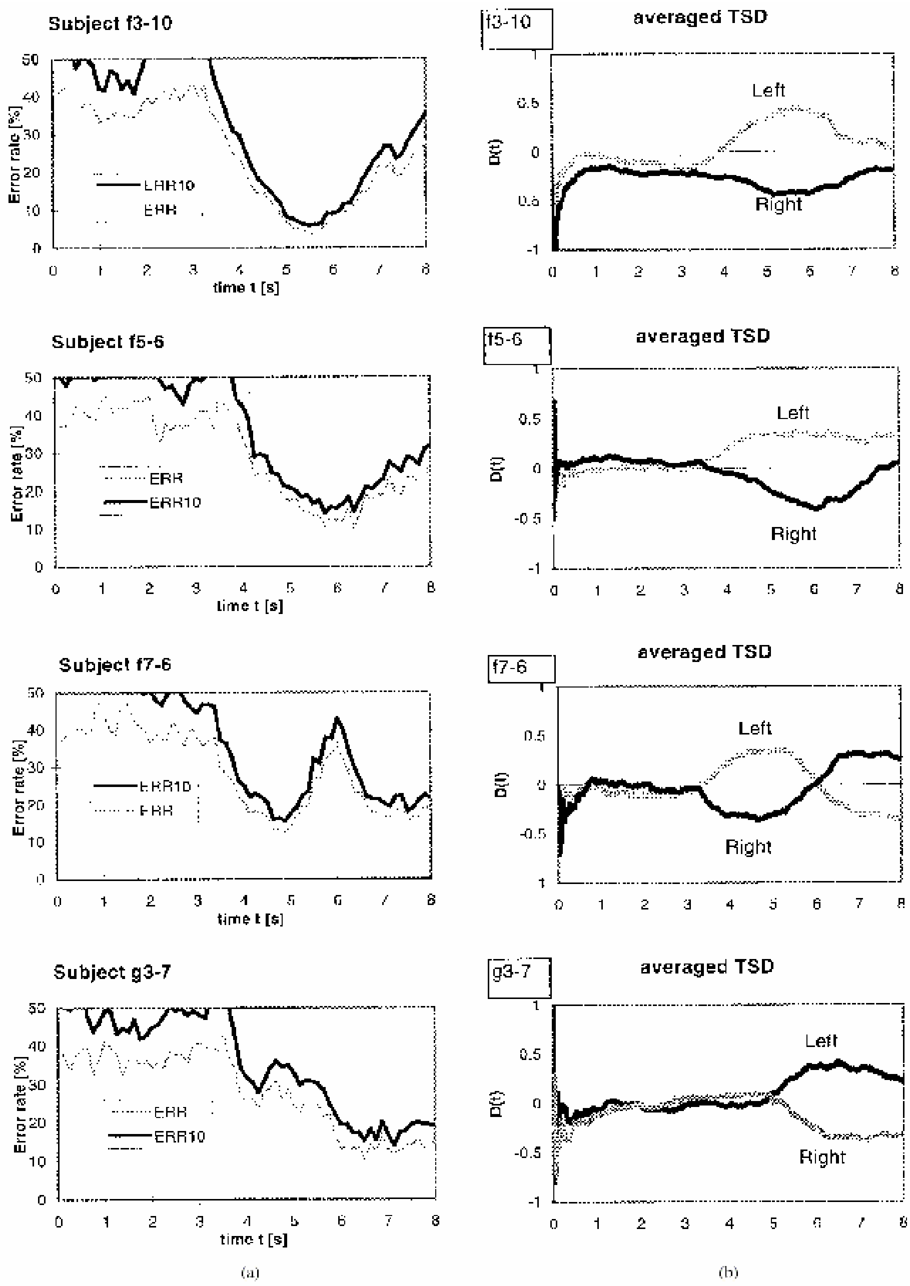


Figure 2.4. The time courses of the error rate (left) and averaged time-varying distance right. The AAR(6) parameters were used from both C3/C4 electrode sites (Modified from (Pfurtscheller 1998a))

3. MATERIALS AND ANALYSIS METHODS

3.1 Real Movement EEG Recordings

In order to investigate the movement ERD and ERS structure, the EEG from 3 healthy subjects during finger movements were recorded at Hacettepe University Medical Faculty Biophysics Department. The subjects sat on an arm chair in a dim room. Consecutive movements of finger were performed with a time separation of no less than 10 seconds. These movements were detected via a micro-switch. EEG was digitized from 19 electrodes selected from the extended 10-20 system with Nuamps amplifier, Compumedics Neuroscan™. The electrode skin impedance was kept below 10 kΩ. The signal was sampled at 250 Hz. In this study the EEG data from only the center electrodes C3 and C4 and their 4 surrounding electrodes are used to get the Hjorth derivation (Figure 3.1.). These channels are converted to Hjorth/Surface Laplacian derivation in order to get the local activity (Hjorth 1975).

$$C_i^H = C_i - \sum_{j \in S_i} g_j C_j \tag{3.1}$$

Where C_i^H is the Hjorth derivation of the center electrode C_i. Here S_i is the index of 4 surrounding electrodes and $g_j = 0.25$ are the weights of each neighbor electrode (Figure 3.1.). The transformed data is filtered between 0.5-40 Hz. There are 240, 120 and 100 sweeps for each movement from each subject respectively.

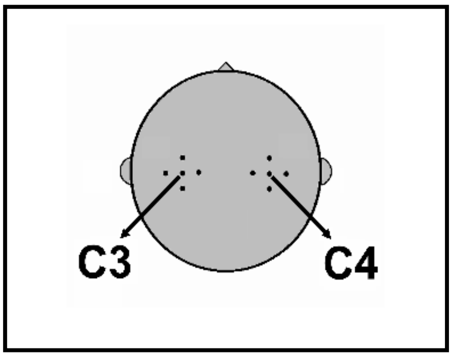


Figure 3.1. C3/C4 electrodes and surrounding 4 electrodes are used when calculating Hjorth (Surface Laplacian) derivation

3.2 Motor Imagery EEG Recordings

3.2.1 Competition 2002 Dataset

The dataset from BCI competition 2002 (Osman 2002), which was provided by Dr. Allen Osman from University of Pennsylvania, is used for investigation.

The imagery EEG data was collected from 9 (S1-9) subjects. The task of each subject was to synchronize an indicated response with a highly predictable timed cue. The subjects were well trained until their responses were consistent within 100ms of the synchronization signal. Each trial began with a blank screen and was designated as a time when it was acceptable for the subject to blink. This blank screen lasted exactly 2 seconds after which it was replaced by a fixation point on the screen telling the subject that the trial has begun. The fixation point lasted for 500ms. The fixation point was replaced by the letter 'E' or 'I', which instructed the subject to perform either explicit or imagined movement. This letter remained on the screen for 250ms and was then replaced again by the fixation point. 1250ms, after the onset of the 'E' or 'I', the fixation point was replaced by the letter 'L', 'R', 'B', or 'N', instructing the subject to act with left index finger, right index finger, both index fingers, or nothing at all, respectively. This letter remained on the screen for 250ms and was then replaced by the fixation point. 1250ms after the letter indicating which finger to use, an 'X' appears for 50ms, which is the synchronization cue and indicates it is time to make the requested response. After the 'X' disappeared, the fixation point stayed on the screen for 950ms and then was replaced by the blank screen, indicating the beginning of the next trial (see Figure 3.2. for timing diagram). The eight different trial types were randomly mixed within a 7 minute 12 second block. Each block consisted of 72 trials. Therefore, nine of each trial type was performed in each block. EEG was recorded from 59 electrodes placed on site corresponding to the International 10/20 System and referenced to the left mastoid. In this study the EEG data from only the C3 and C4 electrodes are used. These channels are converted to Hjorth derivation in order to get the local activity.

3.2.2 Competition 2003 Dataset

This dataset was recorded from a normal subject (G1-female, 25y) during a feedback session and used as an introductory step to classify MI data for a BCI task. The subject sat in a relaxing chair with armrests. The task was to control a feedback bar by means of imagery left or right hand movements. The order of left and right cues was random. The experiment consists of 7 runs with 40 trials each. All runs were conducted on the same day with several minutes break in between. Given are 280 trials of 9s length. The first 2s was quite, at t=2s an acoustic stimulus indicates the beginning of the trial, the trigger channel (#4) went from low to high, and a cross "+" was displayed for 1s; then at t=3s, an arrow (left or right) was displayed as cue. At the same time the subject was asked to move a bar into the direction of a cue. The feedback was based on AAR parameters of channel #1 (C3) and #3 (C4), the AAR parameters were combined with a discriminant analysis into one output parameter. The recording was made using a G.tec amplifier and Ag/AgCl electrodes. Three bipolar EEG channels (anterior '+', posterior '-') were measured over C3, Cz and C4. The EEG was sampled with 128Hz; it was filtered between 0.5 and 30Hz. The trials for training and testing were randomly selected. This should prevent any systematic effect due to the feedback.

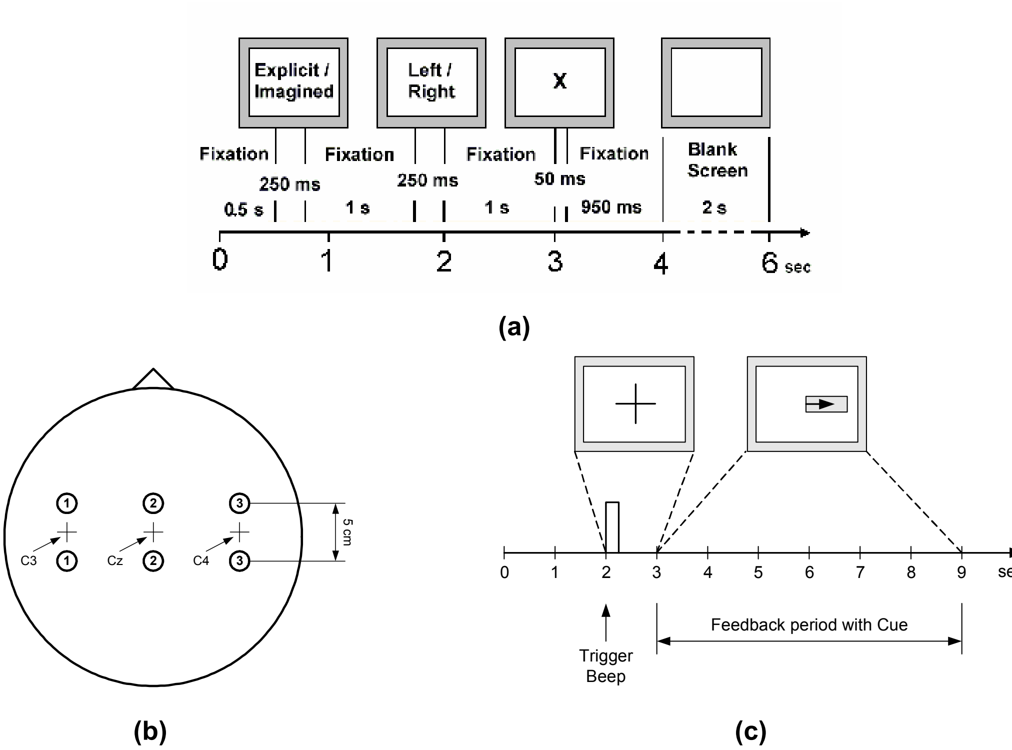


Figure 3.2. (a) The timing diagram of 2002 Data set. (b) Electrode positions of Graz data and (c) timing scheme

3.3 ERD and ERS Maps

As an introductory step in this section the EEG data recorded from 3 subjects (M1, M2 and M3) during real finger movements will be investigated with the Band Pass - Hilbert Transform approach. As stated previously the EEG is decomposed in 5 bands: delta, theta, alpha, beta and gamma. Due to the low pass filtering effect of the scull the gamma band is rarely obtained. This band also overlaps the power line interference, which is one of the biggest obstacles in recording the EEG and it is removed by the filtering in the recording stage. Therefore the gamma band is neglected in the analysis of movement EEG. In literature, generally, the alpha and beta bands are investigated in detail. In this study, these bands will be investigated as well. However, because the limits of these bands are subject depended, a preprocessing is needed to define the upper and lower limits. This step is accomplished by using the standard spectrum comparison approach followed by a filtering process. The Figure 3.3. represents the 1sec. EEG spectrum just before, during and after the movement onset.

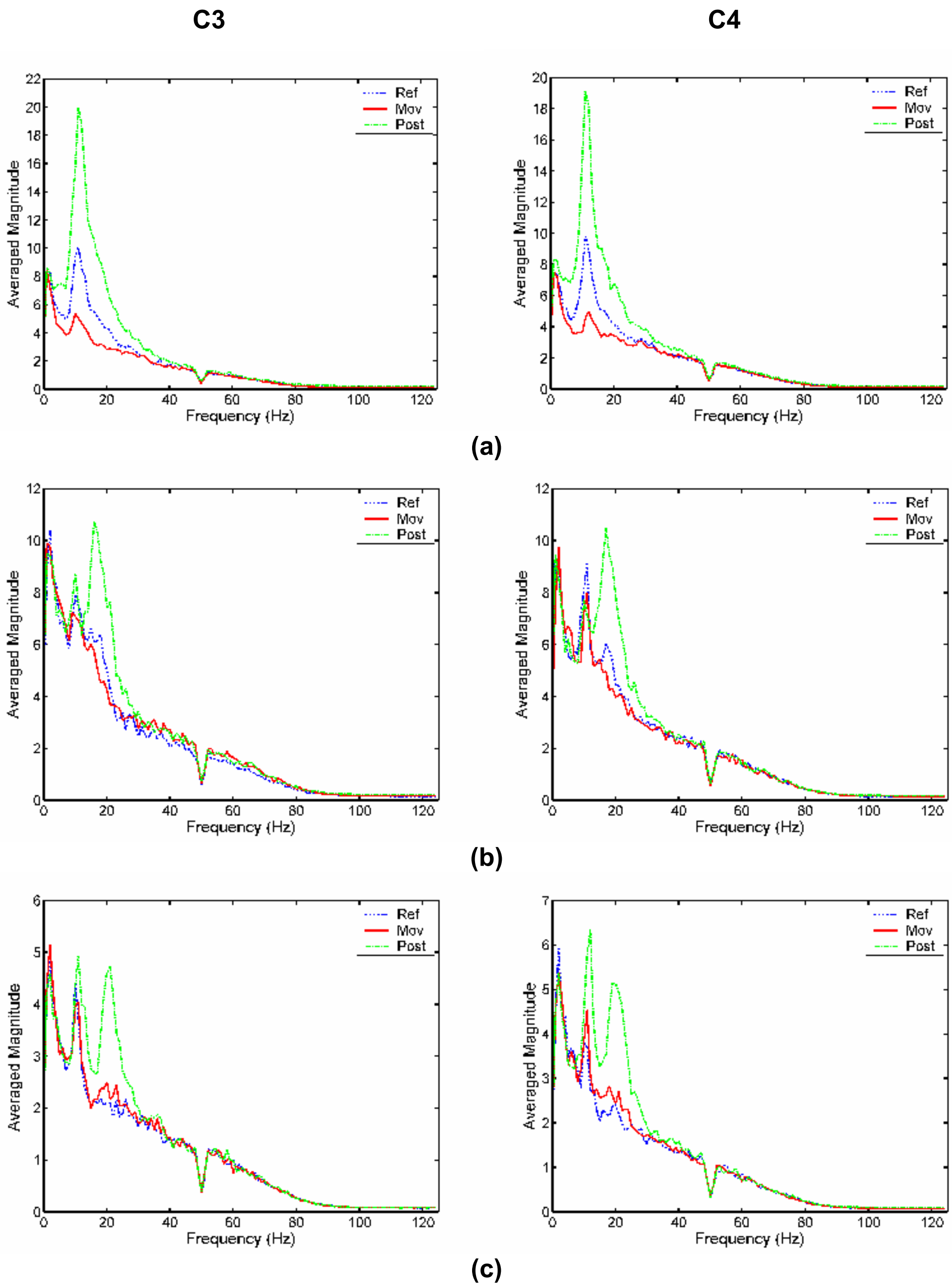
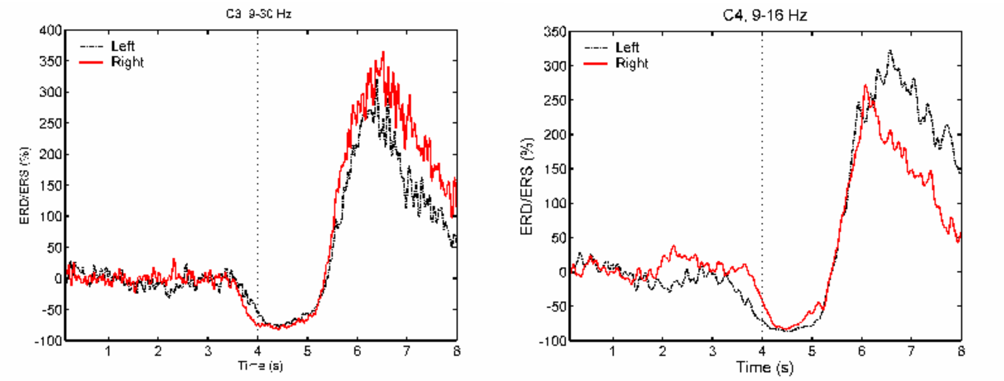
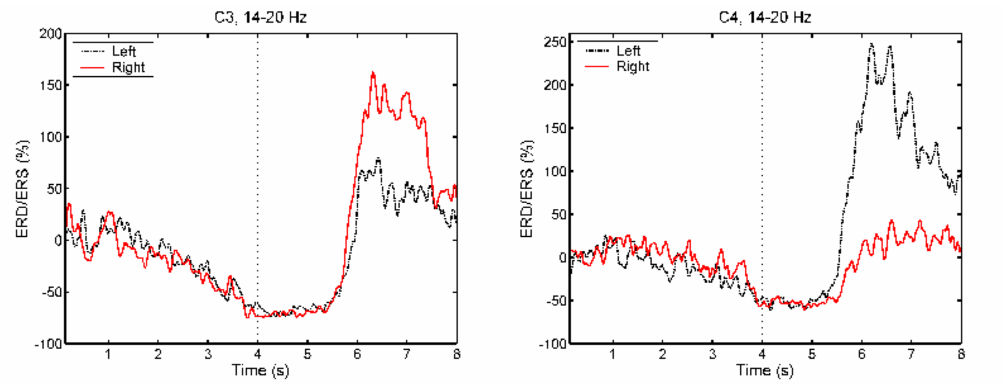


Figure 3.3. EEG spectrum just before and after the movement onset for (a) subject1 (M1) (b) subject2 (M2) and (c) subject3 (M3)

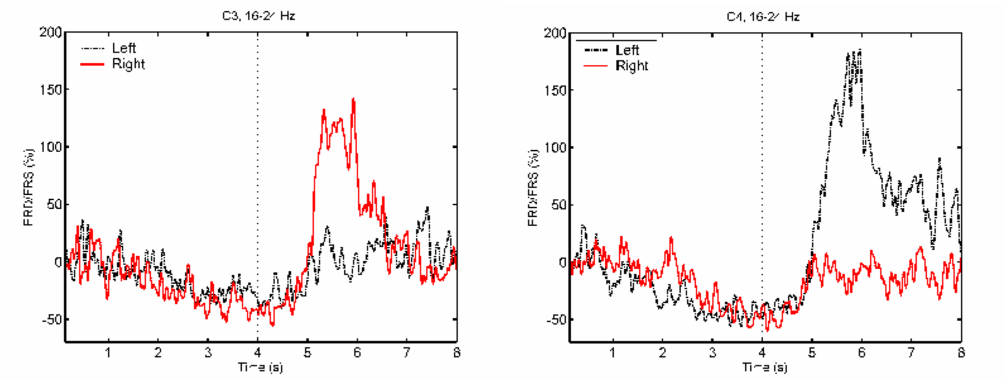
The spectra in Figure 3.3. indicates that for each subject the EEG has different reactive bands. For example for M1 there is a large amplitude modulation nearly ranging from 5 to 30Hz. A visible β peak does not exist. On the other hand for M2 and M3 a sharp power increase in the β band does exist. Also for M3 on the C4 electrode location there is an amplitude enhancement in both α and β bands.



(a)



(b)



(c)

Figure 3.4. The ERD and ERS maps of M1-M3 obtained with Hilbert method for (a) M1 (b) M2 and (c) M3, moving their left or right hand index finger

The ERD/ERS time courses given in Figure 3.4. shows that for M1 the amplitude enhancement takes several seconds to return to the baseline. However, for M2 and M3 the amplitude enhancement in β band is sharp and returns faster to the baseline than α. For M2 and M3 the amplitude of the reactive bands starts to decrease a few seconds prior to the movement onset. However for M1, the ERD starts right before movement onset.

3.4 Time– Frequency Analysis

Obtained ERD/ERS maps, DSLVQ and AAR based studies indicate that subject specific time and frequency domain EEG signal features do exist. However none of these methods consider features from multiple time and frequency indexes. They ignore the possibility that subjects may have physio-anatomical differences and/or different imagery strategies to induce ERD and ERS patterns. In addition, there is a rich literature which indicates that the time and frequency characteristics of the alpha and beta components can vary widely where the beta band shows burst activity whereas the alpha band changes take seconds to attenuate and recover (Lopes 1999; Neuper 2001; Pfurtscheller 1996; Pfurtscheller 1997). This makes the time frequency methods suitable to analyze movement EEG.

The Short time Fourier transform (STFT) is traditionally used to analyze the local frequency content of a signal. In general the signal is segmented into overlapping intervals and in each interval Fourier transform is applied.

$$Stft[m,l] = \sum_{n=0}^{N-1} x[n]g[n-m]e^{\frac{-2j\pi nl}{N}} \qquad (3.2)$$

For each $0 \le m \le N$, $Stft[m,l]$ is calculated for $0 \le l < N$ with discrete Fourier transform of $x[n]g[n-m]$. Here the STFT is applied to single sweeps (see Figure 3.5.).

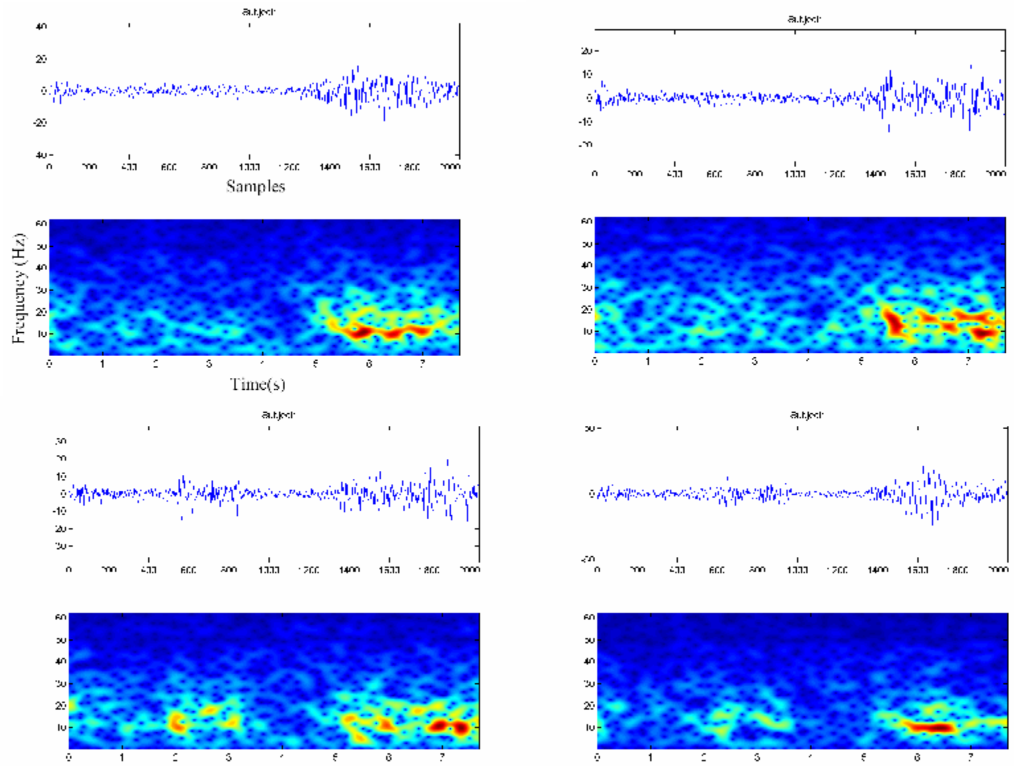


Figure 3.5. Single sweep EEG maps for M1

However it is hard to observe visually the structures in the EEG on a single sweep basis therefore the STFT maps for each channel are averaged for each movement side over the sweeps. This boosts the Signal to Noise Ratio (SNR) as in the analysis of EP and emphasizes common structures (see Figure 3.6.). A similar approach is utilized in (Ginter 2001) to visualize the ERD/ERS microstructure during a finger movement task by using Matching Pursuit analysis.

$$Sf_{av} = \frac{1}{Sw}\sum_{i=1}^{Sw}\left|Stft_i[m,l]\right|^2 \qquad Sw = \text{Number of Sweeps} \qquad (3.3)$$

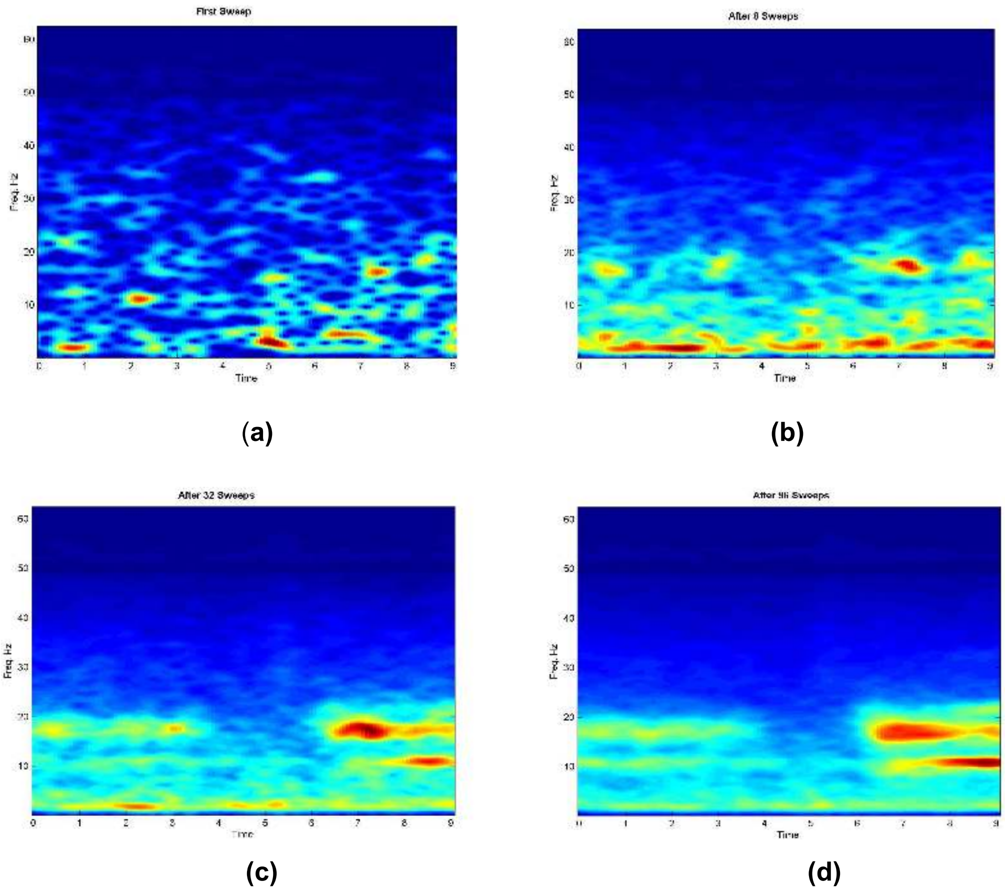


Figure 3.6. The obtained averaged STFT maps for M2 of C3 electrode during right hand movement. (a) First sweep (b) 8 sweeps (c) 32 sweeps and (d) after 96 sweeps for Subject2

The above T-F maps represent the ERD and ERS in an efficient manner. As seen from these maps, averaging T-F maps over sweeps emphasizes common structures. Now this algorithm will be applied on all 3 subject's movement EEGs (M1, M2 and M3) (see figures 3.7., 3.8., 3.9.).

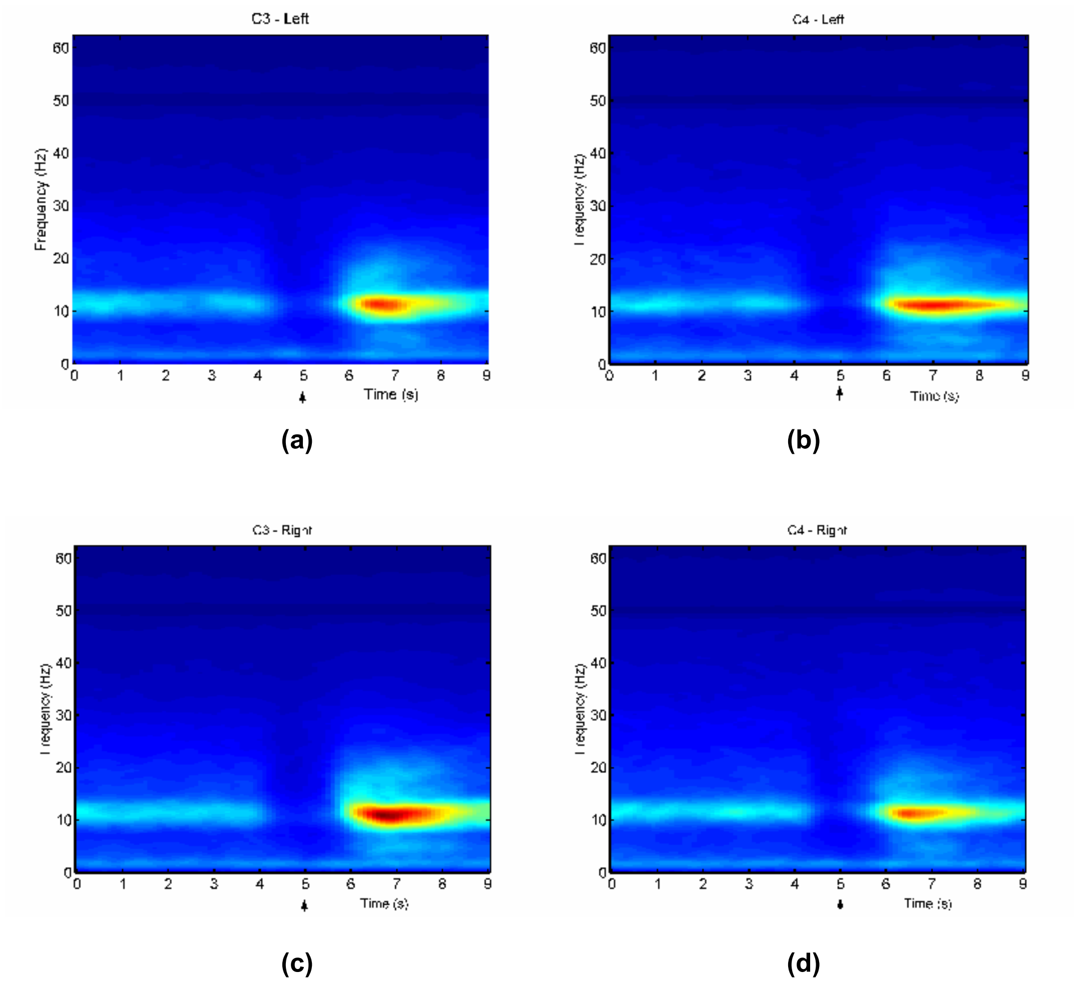


Figure 3.7. The averaged STFT maps of M1. (a), (c) and (b), (d) represent left and right hand sides for C3 and C4 electrode positions respectively for Subject1. The arrow indicates the onset of the movement

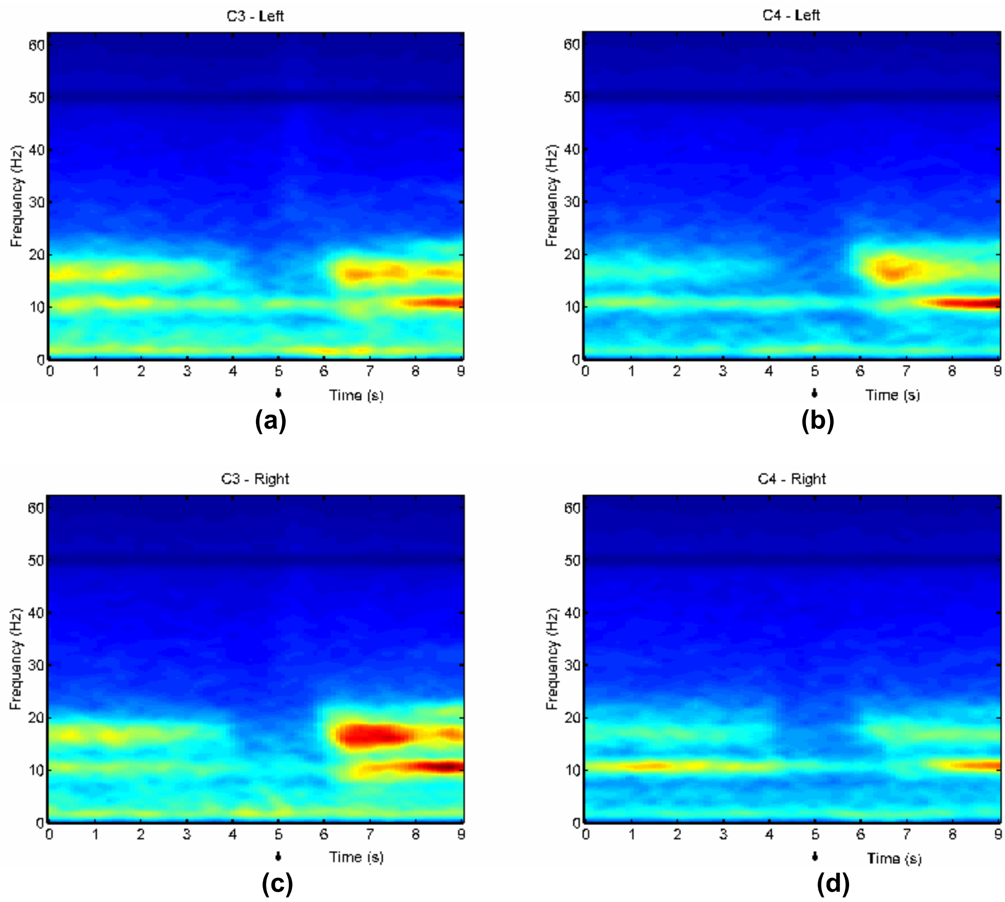


Figure 3.8. The averaged STFT maps of M2. (a), (c) and (b), (d) represent left and right hand sides for C3 and C4 electrode positions respectively for Subject2. The arrow indicates the onset of the movement

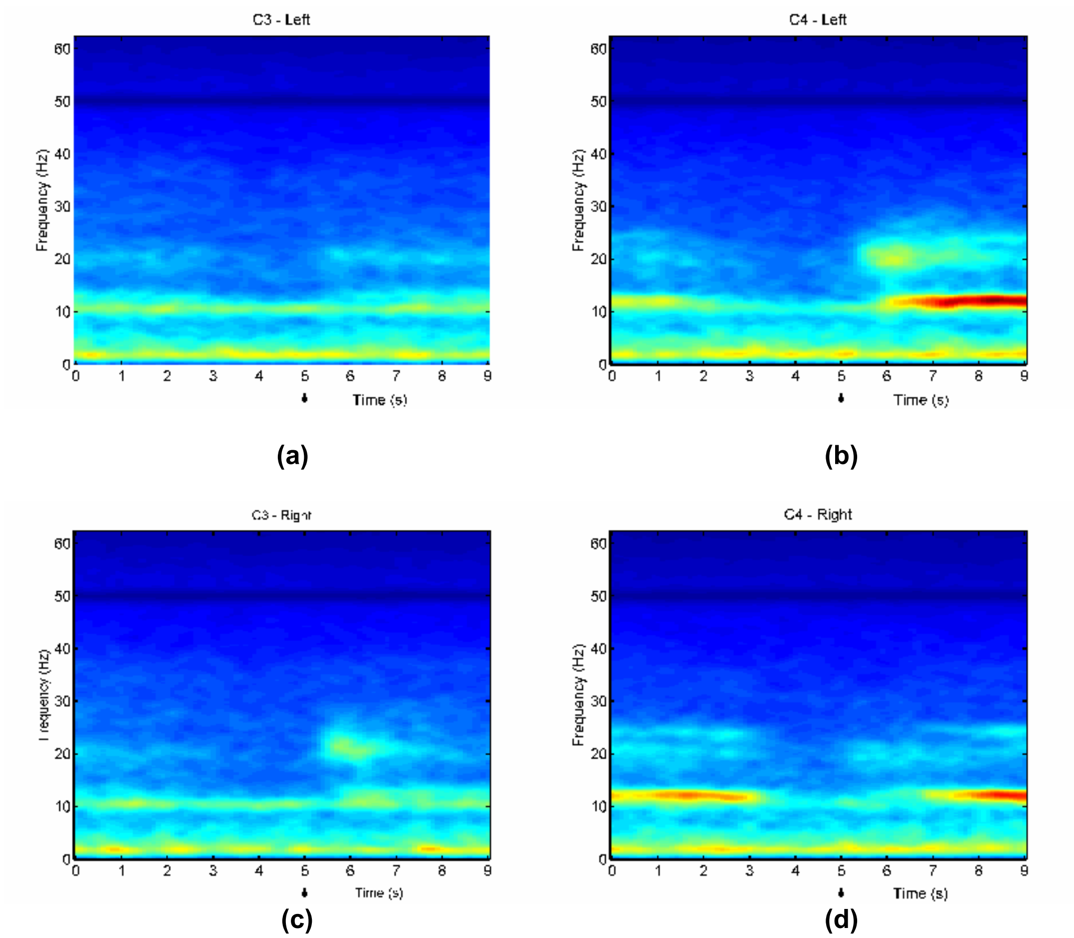


Figure 3.9. The averaged STFT maps of M3. (a), (c) and (b), (d) represent left and right hand sides for C3 and C4 electrode positions respectively. The arrow indicates the onset of the movement

The diverging time-frequency maps significantly emphasize common structures in the movement EEG. For M1 there is an ERD during movement and it is followed by α synchronization. Also the ERD becomes bilateral during the movement for M1, but not for M2 and M3. On M2 and M3 besides the α there is a visible β activity as well. For both subjects this band recovers earlier than the α component. The α ERS occurs several seconds later than the movement offset.

3.5 Adaptive Time Frequency Methods

As shown in the previous section the EEG can be decomposed into several frequency bands and each of them have their own behavior which is directly connected to the underlying physio-anatomical structure. In a movement task these bands form several types of complex structures which using predefined resolution may not capture. This motivates the use of time–frequency analysis methods which can partition the time frequency plane by adapting to the signal characteristic.

Wavelet Transform uses long windows on low and short windows on high frequencies. This provides good frequency resolution in low and good time resolution in high frequencies. Let $x(t)$ be the input signal then,

$$X_{CWT}(a,b) = \int_{-\infty}^{\infty} x(t)\psi^*_{a,b}(t)dt \tag{3.4}$$

is defined as the continuous Wavelet Transform. Here the $\psi^*_{a,b}(t)$ term is derived from the mother wavelet $\psi(t)$ by scaling and shifting it,

$$\psi^*_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \qquad a,b \in R^+ \tag{3.5}$$

where a is the scale and b is the time shift.

In (Mallat 1989a; Mallat 1989b) and (Daubechies 1988) the link between the wavelets and filter banks has been shown. It has been introduced that the wavelet series expansion can be implemented by using quadrature mirror filter where the signal is first filtered into low and high frequency bands and followed by a sub-sampling process. This iteration is repeated on the low frequency band until a certain tree depth and the high band is kept intact (see Figure 3.10.). The generalized version of wavelets, orthonormal "Wavelet Packet" (WP) bases, use quadrature mirror filters to divide the frequency axis in various sizes (Wickerhauser 1994; Mallat 2000; Vetterli 1995). A discrete signal is decomposed into wavelet packet bases by using a fast filter bank algorithm in both low and high frequency bands. The potential of WP bases lies in their capacity to offer a rich menu which allows one to select the best one by using a particular criterion. This is accomplished by decomposing the signal into various frequency bands over a tree structure. Such a structure enables one to have a multiresolution look to the frequency content of the signal (see Figure 3.11.). However, if the signal properties change over time Wavelet Packets may fail to capture such information. In this case Local Cosine Packets (LCP) are constructed, where they partition the time axis. In (Saito 2002; Wesfreid 1999a) LCP were found useful in the analysis of time locked oscillatory geo-acoustic waveforms and speech. Therefore in this work Local Cosine Packets are used as well. The details are given in the following section. The interested reader can find excellent descriptions of the algorithms described in (Mallat 2000; Vetterli 1995 and Kovacevic 1995).

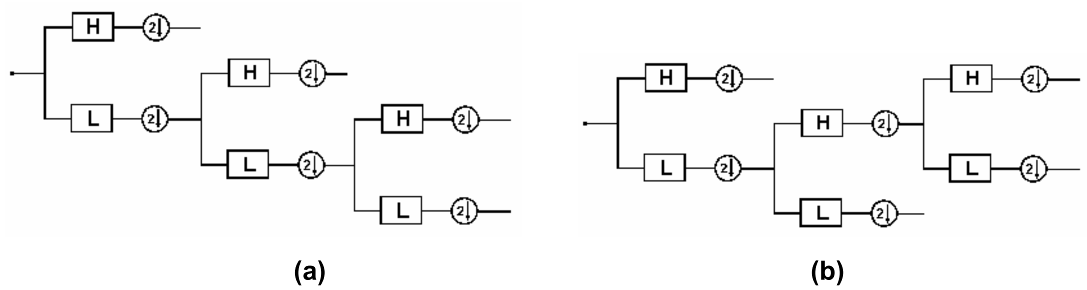


Figure 3.10. (a) Wavelet Tree (b) WP tree obtained by pruning full tree with a cost function where the details are given in the next section

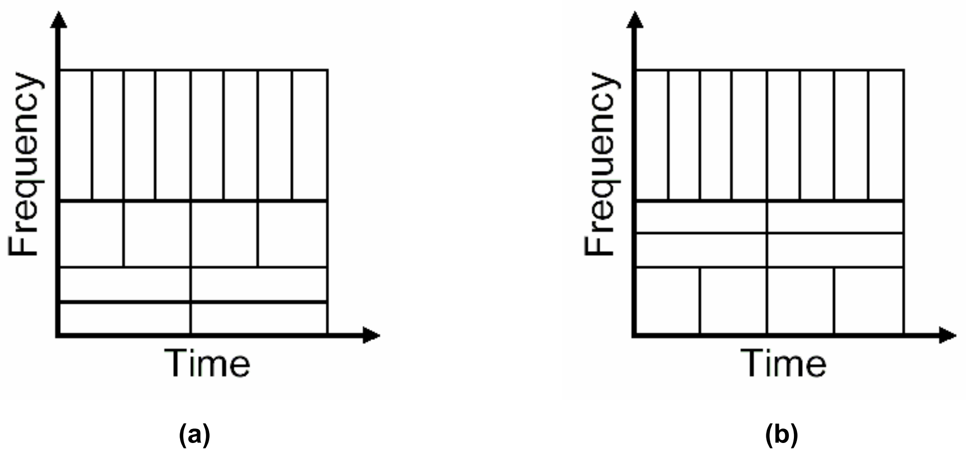


Figure 3.11. The time-frequency tiling (a) Wavelet basis (b) Best Wavelet Packet Basis

3.5.1 Local Cosine Packets

To represent the time varying spectrum, the signal is divided into disjoint rectangular windows and analyzed with Fourier bases. Multiplying a signal with a rectangular window causes side lobe artifacts, which extend to high frequencies. To avoid these artifacts, smooth windows are preferred. However, when smooth windows are used, then orthogonality is lost. It has been shown that one can construct orthogonal bases by using smooth windows modulated by cosine IV basis. This transform is also named as the *lapped orthogonal transform*. The windows are not only smooth but they also overlap their neighbor interval.

They are constructed using cutoff functions such as,

$$r(t) = \begin{cases} 0, & \text{if } t \leq \text{-1}, \\ \sin\left[\frac{\pi}{4}\left(1+\sin\left(\frac{\pi t}{2}\right)\right)\right], & \text{if } \text{-1} < t < 1, \\ 1, & \text{if } t \geq 1 \end{cases} \tag{3.6}$$

where it satisfies,

$$\left|r(t)^2\right| + \left|r(-t)^2\right| = 1 \quad \text{for all } t \in R; \quad r(t) = \begin{cases} 0, & \text{if } t \leq -1, \\ 1, & \text{if } t \geq 1. \end{cases} \tag{3.7}$$

in order to preserve the orthogonality. Consider, partitioning of the time axis.

$$R = \bigcup_{j \in Z} l_j \tag{3.8}$$

An interval $\left[a_j, a_{j+1}\right]$ describes a partition with $l_j = a_{j+1} - a_j$. Here $\gamma \leq \frac{l_j}{2}$ is the length of the overlapping part. The smooth window function is defined as

$$w(t) = \begin{cases} r\left(\dfrac{t-a_j}{\gamma}\right) & t \in \left[a_j - \gamma, a_j + \gamma\right] \\ 1 & t \in \left[a_j + \gamma, a_{j+1} - \gamma\right] \\ r\left(\dfrac{a_{j+1}-t}{\gamma}\right) & t \in \left[a_{j+1} - \gamma, a_{j+1} + \gamma\right] \end{cases} \tag{3.9}$$

$w_j(t)$ is a smooth window, which overlaps its neighborhood intervals and satisfies

$$w^2{}_{j-1}(t) + w^2{}_j(t) + w^2{}_{j+1}(t) = 1 \tag{3.10}$$

for $t \in \left[a_j - \gamma, a_{j+1} + \gamma\right]$ (see Figure 3.12.).

The set of functions

$$\psi_k^j(t) = w_j(t)\frac{\sqrt{2}}{\sqrt{l_j}}\cos\frac{\pi}{l_j}(k+\frac{1}{2})(t-a_j), \quad j \in Z,\ k \in N, \tag{3.11}$$

being an orthonormal basis for $L^2(R)$. A signal $s(t) \in L^2(R)$ can be written in terms of the functions $\psi_k^j(t)$ as

$$s(t) = \sum_{j \in Z, k \in N} c_k^j \psi_k^j(t) \tag{3.12}$$

where

$$c_k^j = \left\langle s(t).\psi_k^j(t)\right\rangle \tag{3.13}$$

The coefficients, c_k^j can be computed with the fast discrete cosine transform (type IV), after a preliminary folding step.

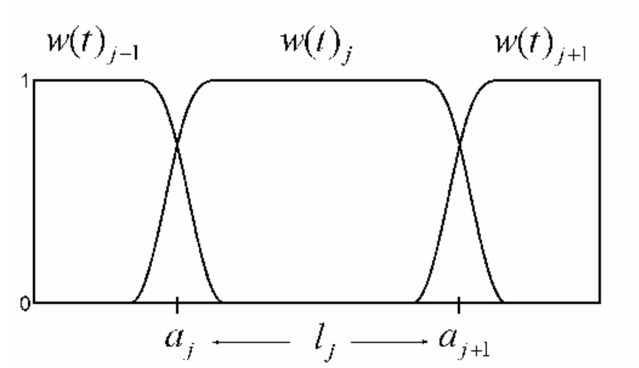


Figure 3.12. The smooth window which overlaps its neighbor intervals

3.5.1.1 Folding Operator

In order to calculate the inner product with a fast DCT-IV a folding procedure is introduced by (Wickerhauser 1994). The support of the smooth window in adjacent intervals is already defined. Now the overlapping part of the window is folded into the center segment with the following rules

$$O_j = \left[a_j - \gamma, a_{j+1} + \gamma\right] \text{ and } C_j = \left[a_j + \gamma, a_{j+1} - \gamma\right] \tag{3.14}$$

where O_j, C_j represent the overlapping and center segments respectively. The folded version f of x is:

$$f[n] = \begin{cases} w[n]x[n] + w[2a_j - n]x[2a_j - n] & \text{if } n \in O_j \\ x[n] & \text{if } n \in C_j \\ w[n]x[n] - w[2a_{j+1} - n]x[2a_{j+1} - n] & \text{if } n \in O_{j+1} \end{cases} \tag{3.15}$$

3.5.1.2 Fast Discrete Cosine IV

The discrete cosine IV basis is calculated by extending the original signal of length N into a signal with a length $4N$. In (Duhamel 1991) a fast approach is introduced where the DCT-IV basis can be calculated on a complex signal of length $\frac{N}{2}$. The complex signal is obtained by

$$\begin{aligned} a[n] &= f[2n] \\ b[n] &= f[N-1-2n] \\ c[n] &= (a[n]+jb[n])e^{-j\left(n+\frac{1}{4}\right)\frac{\pi}{N}} \end{aligned} \quad (3.16)$$

where f is the folded signal. The DCT-IV coefficients are computed from the DFT of $c[n]$.

$$c_f[k] = FFT(c[n]) \quad (3.17)$$

$$f_{IV}[2k] = \text{Real}\left\{ e^{\frac{-j\pi k}{N}} c_f[k] \right\} \sqrt{\frac{2}{N}} \quad (3.18)$$

$$f_{IV}[N-2k-1] = \text{-Im}\left\{ e^{\frac{-j\pi k}{N}} c_f[k] \right\} \sqrt{\frac{2}{N}} \quad (3.19)$$

The DCT-IV coefficients of the original signal $x[n]$ are obtained from $\frac{N}{2}$ length FFT plus a multiplication with $\sqrt{\frac{2}{N}}$ for normalization. Therefore Local Cosine Packets obtained from fast DCT-IV are very suitable for real-time applications.

Since BCI includes online analysis and classification of EEG patterns, current BCI systems are constructed by a rapid prototyping environment, which allows the developer to implement the algorithms in a fast and efficient manner on a real time kernel (Guger 1999). Such systems can run even on a personal PC. Therefore the computational complexity is a crucial step and the LCP algorithm can be implemented without any additional cost.

3.5.2 Best Base Algorithm

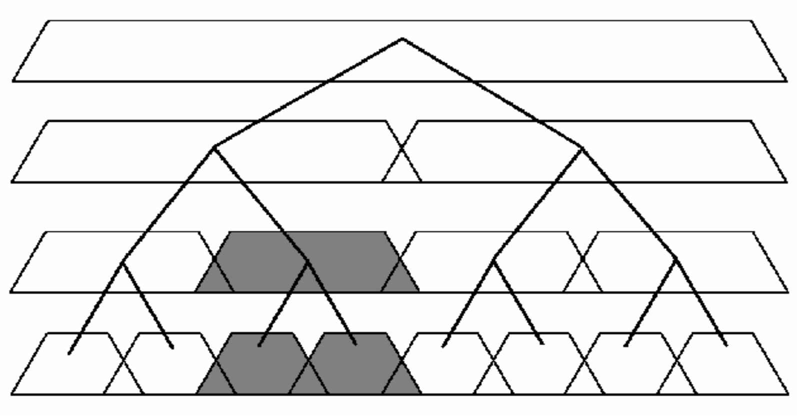
The segmentation of the line can be done smoothly by using Local Cosine Packets by using adjacent fixed windows. Now consider a tree structure which is constructed on the line (Figure 3.13). Each node of the tree represents a segment and it is orthonormal to its children. In each sub segment one can calculate the Local Cosine Packets coefficients in a fast manner. This gives the multiresolution zoom capability to the properties of the signal. However, such a signal representation will be redundant. In (Wickerhauser 1992) Best Base algorithm is introduced to prune the binary tree in a fast manner by optimizing an information cost function, such as entropy. The Best Base algorithm can be summarized as follows.

Step 1: Expand the signal into Cosine or Wavelet Packets over a tree structure.

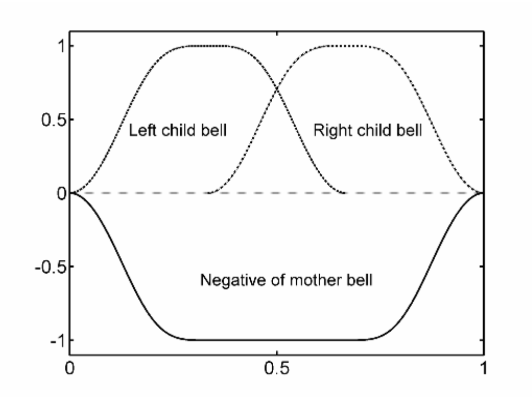
Step 2: In each coordinate calculate the cost function over the expansion coefficients.

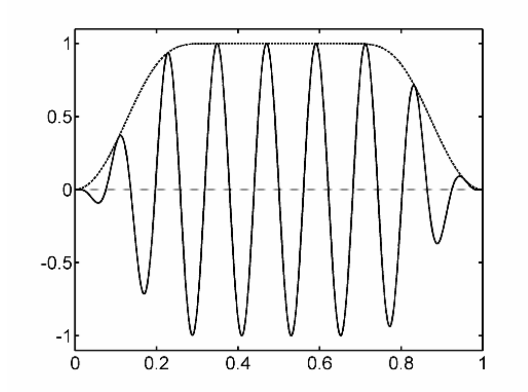
Step 3: Prune the tree from bottom to top by minimizing the cost function (entropy)

The pruning is completed from bottom to top by comparing the cost function of the mother with the children. Whenever the mother is lower information cost then the children are destroyed. In the opposite case the children are marked and the total costs of the children are copied to the mother node. And the comparison is repeated in the upper levels. If the lower levels have lower information then the mothers are destroyed.



(a)





(a) (c)

Figure 3.13. (a) The dyadic tree where each subspace is represented by smooth windows. (b) The children and parent window. They are also marked in (a). (c) The smooth cosine

Assume m, c_1 and c_2 represent the parent and its children subspaces over the dyadic tree respectively (also marked on Figure 3.13-a).

$$m = c_1 \oplus c_2 \tag{3.20}$$

Then for pruning the following rules are applied

$$\begin{aligned}
&if \quad I_m \leq \left(I_{c1} + I_{c2}\right) \\
&\qquad Keep \ m \\
&\qquad Destroy \ c1, c2
\end{aligned}$$

Else

$$\quad Keep\ c1, c2$$

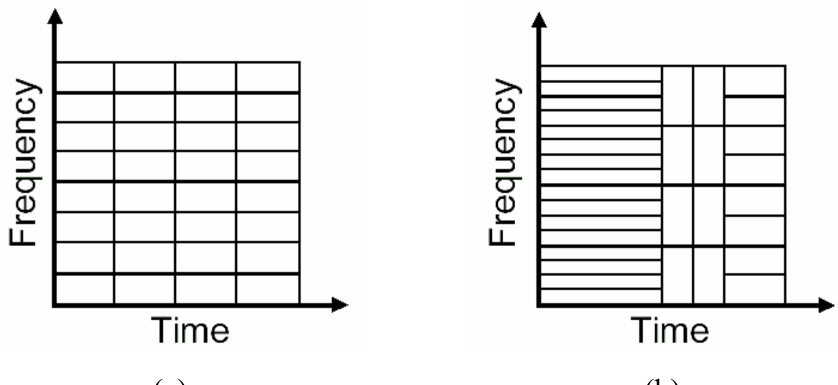
$$\quad I_m = \left(I_{c1} + I_{c2}\right)$$

where I_m, I_{c1} and I_{c2} represent the information cost, Shannon Entropy.

$$I = -\sum_k c_k^2 \ln(c_k^2) \tag{3.21}$$

As stated before whenever the parent node is higher of information cost, then the cumulative information of the children is copied to mother node and marked. In the opposite case the tree is cut at that level and the children are destroyed. Here the overlapping part is the half of deepest segment which preserves the orthogonality in the pruning stage.

In practice the entropy cost function measures the flatness of the spectrum of the segment. In case the children have similar spectrum then they are merged at the parent node otherwise the time varying spectrum is represented by two adjacent smooth windows. As a result the pruned tree is well adapted to the time varying characteristic of the signal (Figure 3.14.).



(a) (b)

Figure 3.14. (a) STFT Tiling, which is used to visualize the ERD/ERS structure smoothly (b) Best Local Cosine Basis Tiling which can adapt the time varying properties

The best base algorithm can be used on a single sweep basis. But such an approach makes it difficult to extract the ERD and ERS structure of a person due to the low SNR and random variations. The same obstacle arises when using STFT to extract global information. Instead of working on single sweeps the data set belonging to a class is first decomposed into a tree structure by using the cosine packets or wavelet packets. Then the squared expansion coefficients are averaged over the sweeps. On the averaged coefficients the entropy based divide and conquer algorithm was applied. The pruned subspace is named as "joint best base" (Wickerhauser 1994). It has been shown that the selected subspace via joint basis approximated the Karhunen Loeve Transform (KLT).

Now the movement EEGs of 3 subjects are analyzed with joint best base algorithm which is summarized as:

Step 1: Expand the signal into Cosine or Wavelet Packets over a tree structure.

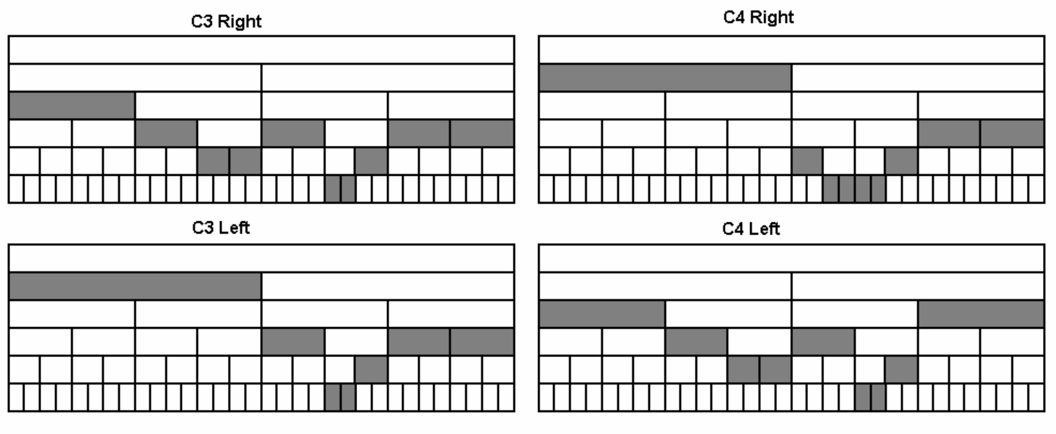
Step 2: In each subspace square the expansion coefficients and average over the sweeps.

Step 3: Prune the averaged tree from bottom to top by minimizing the cost function (entropy).

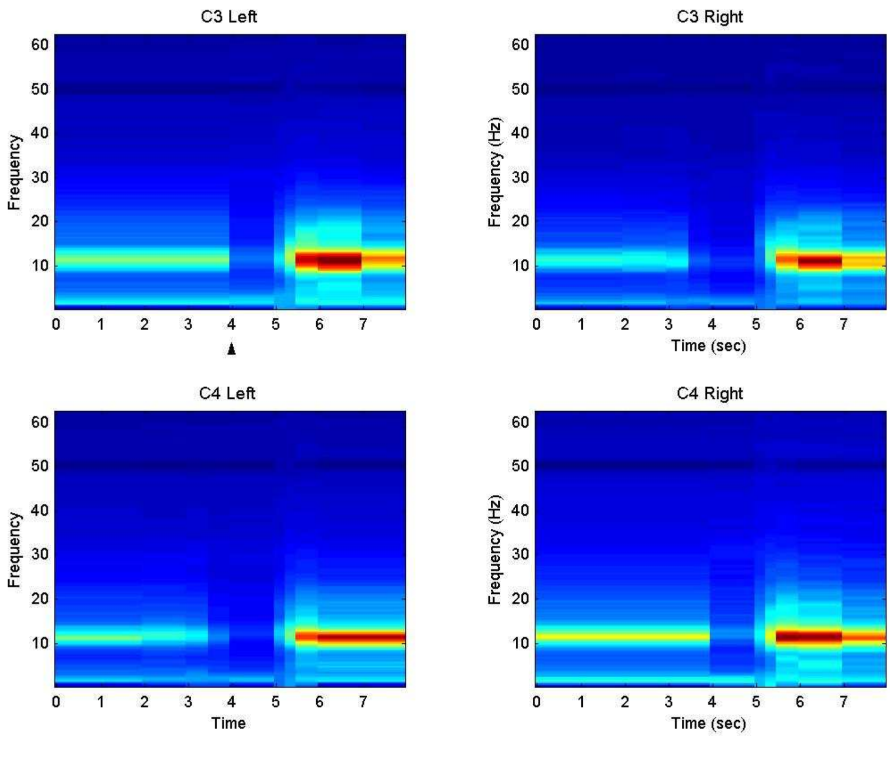
Here the selected subspaces correspond to the adaptive time segmentation of the movement EEG. This procedure is evaluated for each channel and for each side to extract the ERD/ERS structure of a person.

3.5.3 Adaptive Time Frequency Characteristic of ERD/ERS

The time segmentations and T-F maps obtained from the Adapted Local Cosine packets represent the time varying characteristic of alpha and beta bands. As seen from the figures 3.15, 3.16 and 3.17 each subject has its own individual frequency components. M1 has only alpha component. This component has a strong ERS right after the movement. However on M2 and M3 the alpha component recovers several seconds later. The characteristic of alpha band of M2 and M3 is in accordance with the literature. However the early recovery which is observed on M1 is interesting. It has a short ERD and recovers earlier with a burst characteristic. This behavior is reported by several authors.



(a)



(b)

Figure 3.15. (M1) With strong alpha activity. (a) The dyadic segmentation obtained from BB for C3/C4 electrodes. The premovement on the ipsilateral side is never segmented (b) The T-F maps. For C3 electrode during right finger movement there is a larger ERD region than the left hand side. The arrow indicates the onset of the movement

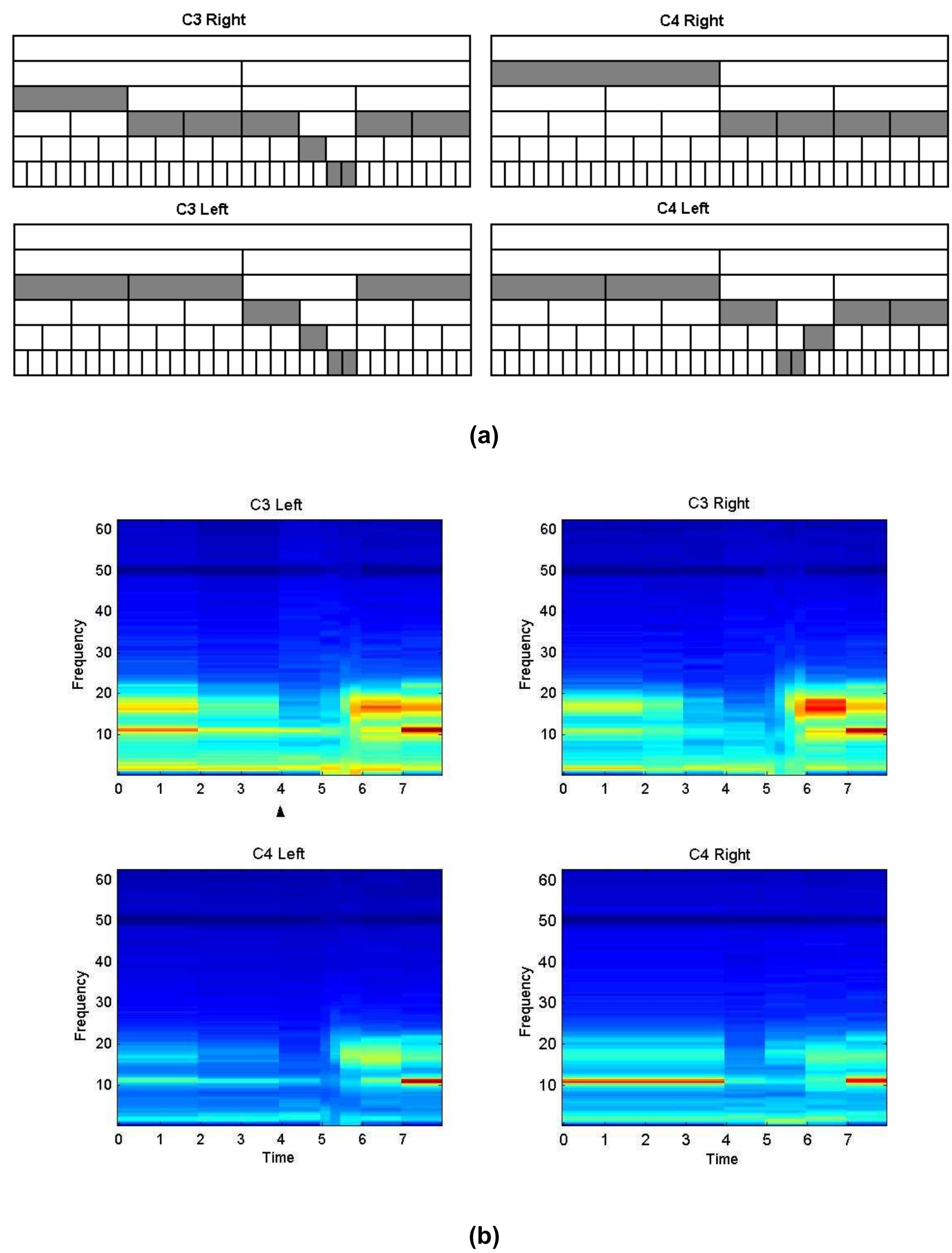


Figure 3.16. (M2) With both reactive alpha and beta bands. (a) The dyadic segmentation obtained from BB for C3/C4 electrodes. The pre-movement on the C4 electrode during right hand finger movement is not segmented. (b) The T-F maps of C3/C4 electrodes. On both sides C3 electrode location has reactive beta and alpha component. The arrow at second 4 indicates the onset of the movements

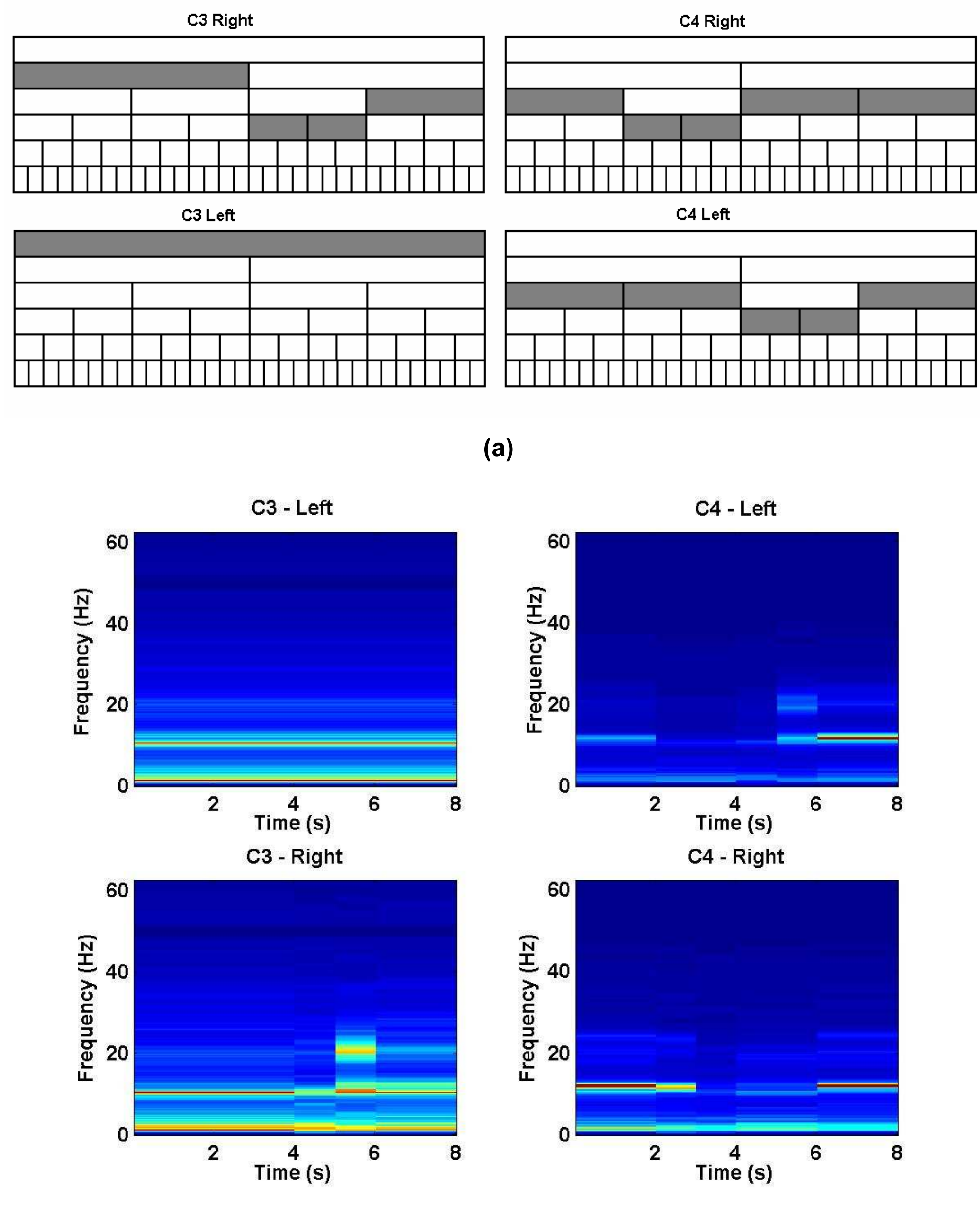


Figure 3.17. (M3) With reactive alpha and beta component. (a) The dyadic segmentation obtained from BB for C3/C4 electrodes. The time line of C3 electrode location is never segmented during left hand finger movement. (b) T-F maps. During the right hand finger movement there is a short lasting beta ERS right after movement. The movement onset is at second 4

The BB segmentation provided different tiling for each subject due to the variances between their EEG. Also the segmentation changed according to the side of the executed movement.

For example on S1 and S2 pre-movement is nearly never segmented on the ipsi-lateral side of the direction. The segmentation on the contra-lateral side is perfectly mirrored for these subjects. These findings agree with the characteristic of C3/C4 electrode locations. For S3 the EEG data from C3 electrode during left hand finger movement is never segmented. This means a bilateral ERD is not observable. It was expected that the C4 location would have the same behavior during the right hand finger movement. However for this subject the C4 electrode location is always activated no matter which hand is used. On this same subject a β ERS is observed between the 5^{th} and 6^{th} seconds. A similar ERS occurs between the 6^{th} and 7^{th} seconds for S2. These results strongly support inter-subject variability (Schlögl 1997).

4. CLASSIFICATION METHODS

The time varying characteristics and results obtained with BP and AR model based studies were already mentioned. Now for a representative subject in Figure 4.1., the left and right hand finger movement imagery EEG data is visualized by averaged STFT based method. As seen from the T-F maps the beta band shows early recovery and lasts shortly whereas the alpha band changes take seconds to attenuate and recover. This behavior is also observed in Chapter 2. These findings motivate the development of an algorithm, which can extract subject specific time – frequency features in an automated manner for discrimination between left and right hand tasks. In this chapter such an approach is introduced for the classification of single sweep ME and MI EEG recordings.

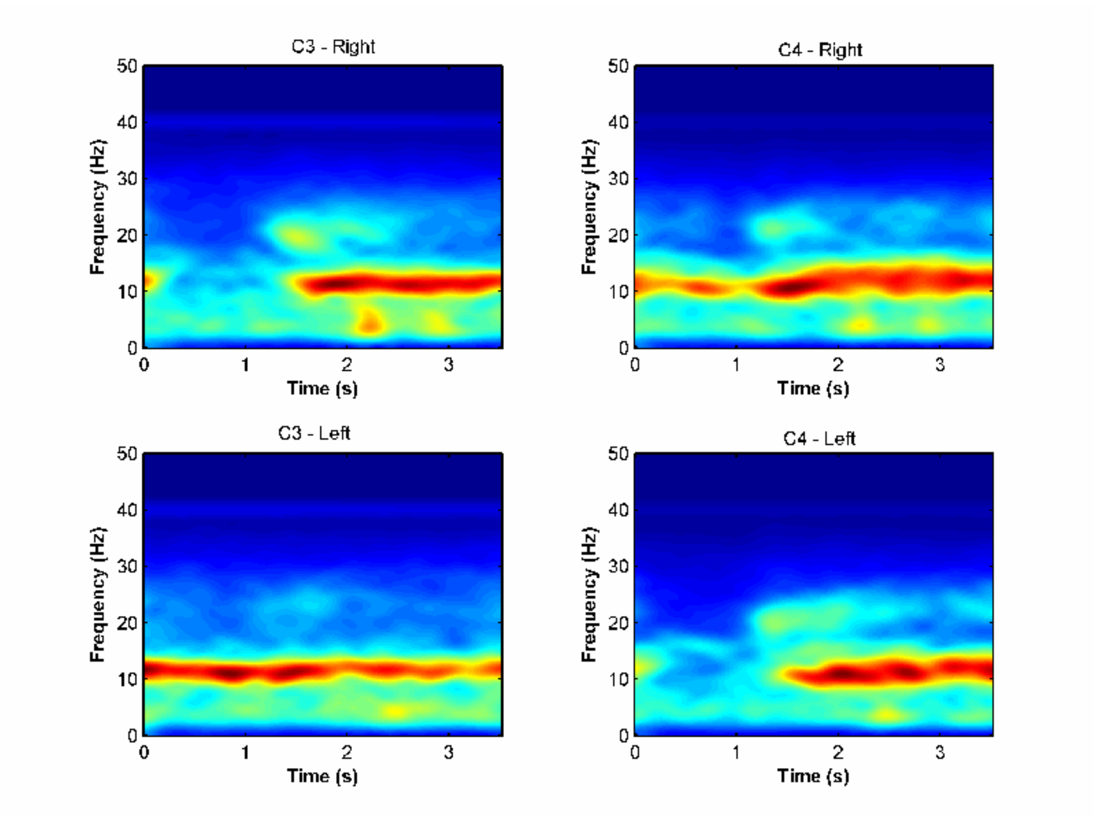


Figure 4.1. The T-F maps of C3 and C4 electrodes belonging to left and right hand motor imagery. Notice the time varying characteristic of the alpha and beta bands on each electrode location according to the side of movement

4.1 Local Discriminant Bases

The BB algorithm was developed to extract local information. It expands the signal into orthonormal bases by using wavelet packets or local trigonometric bases over a dyadic tree. This full tree is pruned to minimize a cost function, such as entropy, by a divide and conquer algorithm. The main concern of best-base method is signal representation. Since the selected cost function is entropy, such an algorithm ends up selecting the sub basis, which maximizes signal compression. However for classification, the discrimination power of the nodes must be measured. Therefore, the entropy criterion is replaced by another cost function, which can measure the distance of the nodes between the classes. By pruning the binary tree to maximize the selected distance criterion, local discriminant bases (LDB) are extracted (Saito 1994; Bishop 1995). As stated in the previous chapter Local Cosine Packets will be used in the estimation of Local Discriminant Bases. The LDB Algorithm can be summarized as follows:

Step 1: Expand each training signal Local Cosine Packets coefficients over the dyadic tree.

Step 2: In each subspace square the expansion coefficients and average over the sweeps.

Step 3: Calculate the distance between averaged expansion coefficients of each class and accumulate them in each subspace.

Step 4: Prune the tree from bottom to top via maximizing the cost function.

Step 5: Order the expansion coefficients from the pruned tree and select the top $k << n$ coefficients for classification. where n is the dimension of the training signal.

There are various choices for distance measures. Assume p, q are normalized energy distributions of signals belonging to class1 and class2 respectively.

The used measures are:

- The symmetric Kullback Leibler distance, which is also named as J-divergence,

$$I(p,q) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} \tag{4.1}$$

$$J(p,q) = I\left(p,q\right) + I\left(q,p\right) \tag{4.2}$$

- Euclidean distance

$$D\left(p,q\right) = \left\| p_i - q_i \right\|^2 = \sum_{i=1}^{n} \left(p_i - q_i\right)^2 \tag{4.3}$$

- and the Hellinger distance

$$H\left(p,q\right) = \sum_{i=1}^{n} \left(\sqrt{p_i} - \sqrt{q_i}\right)^2 \tag{4.4}$$

The Fisher criterion is also considered for feature selection.

$$F = \frac{\left(\mu_1 - \mu_2\right)}{\sigma_1^2 + \sigma_2^2} \tag{4.5}$$

where μ and σ are the mean and standard deviation of the feature it belongs to.

The dyadic tree is calculated for each class (see Figure 4.2.), and then a single tree is found, where both classes are well separated according to the selected criteria. From the resulting tree features are selected by using J, D, H and F criterion.

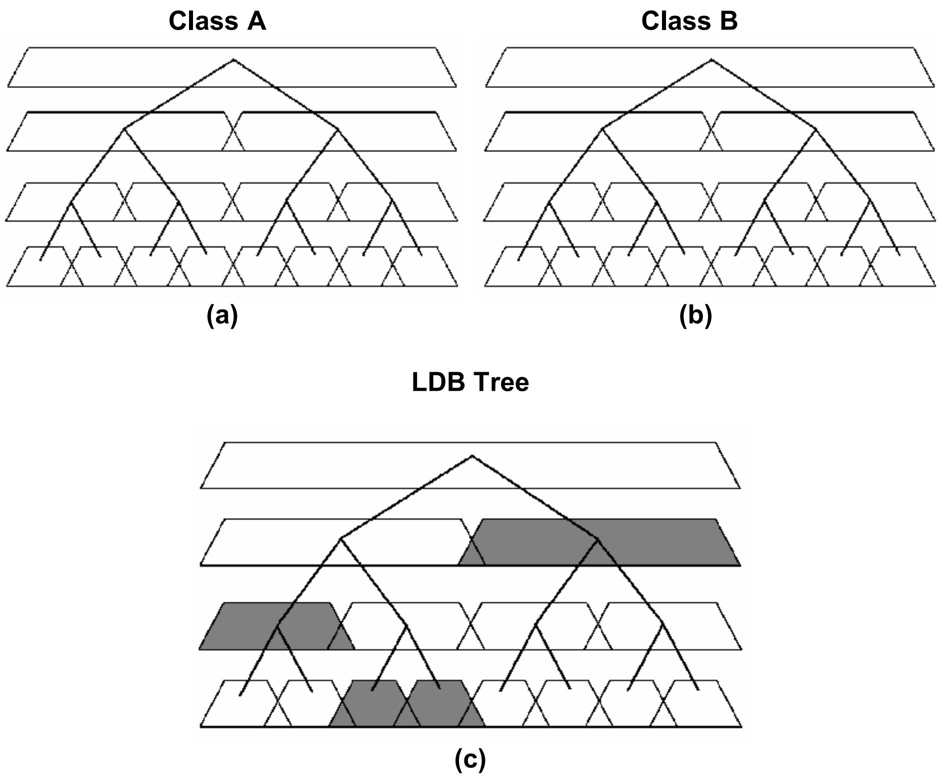


Figure 4.2. (a) dyadic tree for class A, (b) dyadic tree for class B. (c) selected LDB coordinates, which maximizes the distance between A and B

4.2 Band Features

The original LDB algorithm ends up with a feature space equal to the dimension of the original signal. High dimension feature space can cause decrease in the generalization capability where limited number of training samples exist - *curse of dimensionality* (Bishop 1995). In addition, it is hard to capture discriminant information with single CP coefficients unless the data is phase-locked. In order to reduce the dimensionality and the lack of shift invariance the expansion coefficients are grouped in α (8-13Hz) and β (14-30 Hz) bands. Such an approach has direct connections with Mel-scale approach used in acoustic signal processing and classification (Cetin 2004).

4.3 Principal Component Analysis (PCA)

The problem which can arise in high dimensionality is already discussed. Another problem in classification is the correlation between coefficients. One can select the features by using a class separability criterion. Like in LDB these features can be sorted according to their discrimination power. However this approach does not guarantee that the selected features are orthogonal. The same information during the sorting process can be repeated. Therefore it is desirable to select discriminant features which represent different type of information. In (Englehart 1999) PCA was used on the wavelet packet coefficients for dimension reduction to classify electromyogram signal. In this work the PCA is also used to reduce the dimensionality and create a decorrelated feature set. The PCA seeks a coordinate system which best explains the variance of the data set. While projecting the data set on the new coordinate system the linear dependencies are grouped and represented by one of the principal axes. Each eigenvector is one of the principal axes. In practice the PCA can be computed as follows (Bishop 1995). First the covariance matrix of the *d* dimensional feature set is calculated.

$$\overline{x} = \frac{1}{N}\sum_{n=1}^{N} x^n \tag{4.6}$$

$$\Sigma = \sum_{n}\left(x^n - \overline{x}\right)\left(x^n - \overline{x}\right)^T \tag{4.7}$$

Then the eigenvalues and corresponding eigenvectors are found. The eigenvectors corresponding to the largest eigenvalues are retained and used to project the original feature set;

$$\Sigma u_k = \lambda_k u_k \tag{4.8}$$

$$W_k = u_k^T x \tag{4.9}$$

where W_k is the projected vector, u_k is the kth eigenvector of the covariance matrix of the feature set and x^n is the d dimensional nth feature vector. Here the error introduced by dimension reduction is directly related to corresponding eigenvalues.

4.4 Linear Discriminant Analysis (LDA)

The reduced feature set for classification is used as an input to LDA. LDA is one of the most commonly used statistical classifier due to its ease of implementation and trainability. It can be seen as a method for identifying best discriminating hyper-plane in an n dimensional space. A detailed description can be found in (Duda 1973). The weight vector used in LDA is

$$v = (\Sigma_1 + \Sigma_2)^{-1}(m_1 - m_2) \tag{4.10}$$

where Σ, m are the covariance matrix and mean of class features respectively. The distance of a feature vector to the discriminating hyper-plane is calculated as

$$d = v^T x \tag{4.11}$$

where x is the feature vector.

Of course the classifier family is not limited with LDA. Possible choices are Multi Layer Perceptron (MLP), instance based classifier such as kth Nearest Neighbor (k-NN) or Classification and Regression Trees (CART). One can also use one of the listed classifier. Their success in BCI research is also compared and discussed. Although MLP can deal with nonlinear problems it has to be trained with a large set to prevent over-learning. Also it has many free parameters such as number of hidden neurons and number of training epochs. On the other hand k-NN has same capabilities but it is subject to high dimensionality. CART is a very effective classifier which uses consecutive hypothesis testing to create splits. Since in each split one variable is used it is hard to separate classes where the discriminant information is hidden in combinations.

5. FINDINGS, IMPROVEMENTS AND DISCUSSIONS

In order to find out whether the developed algorithm is capable of distinguishing between left and right hand finger movement, initially EEG records accompanying real executions are classified. The same data set was also analyzed with BB algorithm previously. Besides that the EEG dataset of the BCI competition 2003 recorded from one subject is used to test the algorithm on an imagery record. Both ME and MI EEGs are analyzed with LDB algorithm with a tree depth of 5. Both ME and MI data were 8 seconds long. ME covers 4 seconds pre and 4 seconds post movement region. The MI data starts from the second zero (see experimental paradigm). In the constructed time segmentation the Cosine Packet coefficients are grouped in alpha and beta bands. Obtained band features are sorted according to the J, D, H and F criterion. Besides this the entire feature set is projected onto its principal components and sorted according to the corresponding eigenvalues in descending order. On both feature sets the classification error is calculated (see (Ince 2005) for details). To compare whether the adaptive time segmentation is capable of improving the classification accuracy uniform/fixed time segmentation is also implemented. The length of each uniform segment was 0.5 seconds. Also in order to compare the classification accuracy with standard methods an AR model and LCP band power based classification procedure is implemented. The AR model order is selected as $p = 6$ where the model represents 3 peaks in the spectrum. Both model parameters and LCP band powers are calculated in a 1 second window for C3 and C4 electrodes. This resulted in a feature vector of dimension 12 for AR model based approach. The window is moved along the EEG signal, one sample at a time. For each shift, the LCP coefficients and AR model parameters are calculated. Then the LCP coefficients are grouped in alpha and beta frequency bins as described in the LDB procedure to capture band powers and reduce dimensionality. The LCP based line search is motivated by the belief that the band features in a smooth window Cosine Packets (CP) analysis will capture the ERD and ERS events by searching the time axis in a flexible manner. As a last step for each approach LDA is used for classification. Furthermore, on the ME data for each time point 10 fold cross validation is implemented to estimate the classification error. For the MI case the 5 fold cross validation is implemented on the given training set. Then the combination of features which give minimal error is used on the test set provided by competition organizers. The results of the methods discussed are given in tables 5.1. and 5.2.

Table 5.1. Results with minimum error for each subject are highlighted. Besides classification, the number of coefficients (NC) for minimal error is also given. ME stands for Movement Execution and MI for Movement Imagination. Fixed represents uniform time segmentation

ME			**CS**		**PCA**		**Fixed**
			Err	**NC**	**Err**	**NC**	**Err**
M1	**J**	**J**	19.7	37	22.3	37	21.5
		Fisher	19.7	33			
	D	**D**	19.6	40	22.1	40	
		Fisher	**18**	**40**			
	H	**H**	20.1	37	22.2	37	
		Fisher	19.7	32			
M2	**J**	**J**	9.5	23	11.7	15	12.4
		Fisher	9	21			
	D	**D**	**8.3**	**15**	11	26	
		Fisher	9.17	15			
	H	**H**	9.3	23	10.8	29	
		Fisher	9.17	21			
M3	**J**	**J**	12.8	17	12.6	30	20
		Fisher	12.5	18			
	D	**D**	12.3	13	12.5	30	
		Fisher	**12**	**13**			
	H	**H**	12.8	17	12.6	30	
		Fisher	12.5	18			
MI			**CS**		**PCA**		**Fixed**
			Err	**NC**	**Err**	**NC**	**Err**
G1	**J**	**J**	15	8	15	9	23.6
		Fisher	15	9			
	D	**D**	19.2	6	21	4	
		Fisher	22.1	8			
	H	**H**	**14.2**	**7**	15	9	
		Fisher	14.2	12			

Table 5.2. The classification error of three movement subjects obtained from AR modeling, LCP Band Powers (LCP-BP) and LDB

Subj.	AR(p=6)	LCP-BP	LDB
1	27	24	18
2	17	14	8
3	18	13	12
Avg	20.6	16.8	**12.6**

In this study LDB-CS has outperformed all other approaches. Especially grouping the coefficient in alpha and beta bands has improved the classification accuracy nearly 10%. Although CS and PCA based feature sorting have very similar results, the best results were obtained with the CS method. The PCA uses more features than CS to lower the classification error. The initial features of PCA resulted in poor classification error (see Figure 5.1.). This can be explained by the fact that the first eigenvectors may be in the direction of the most energetic features, which are alpha activities, or some high amplitude oscillations, before the movement with large variability (see Figure 5.2.). In average *F* based feature sorting provided the best results. It was observed that the primary features selected by *F*, are most of the time, in post movement beta synchronization. These features are followed by alpha oscillations, which belong to the planning stage. Post movement alpha synchronization is also selected as a primary feature. The locations of these features are a few seconds after the movement, whereas beta synchronization occurs right after the movement (see next sections for visualization). It has been noticed that adding more PCA features for classification is more robust than using CS features. A reason for this can be the fact that PCA features have a nature of being uncorrelated.

Interestingly the algorithm constructed different tiling for each channel for some subjects, where normally a mirroring was expected. The adaptive segmentation of C3 and C4 electrodes, implemented to discriminate left and right MI, was very close. Here a perfect mirroring did not exist either. Both CP and BP are used separately for classification. There was an improvement from CP to BP, but not as large as it was with ME data. The main reasons for this are that the subject has only alpha activity and that the only discriminant components occur in a short interval just after the cue in the alpha band. So both methods use a few coefficients to represent this discriminative information. The features are selected from the 3-6 and 4-6 sec intervals on C3 and C4 electrodes respectively (Figure 5.3.). The locations of the features were not identical on either channel. Interestingly, it has been observed that a few of the selected features occur before presentation of the cue. It is believed that these intervals are the mental states where the subject is producing imagery activity. Because the original LDB is limited to using intervals of dyadic lengths, two 8 seconds subintervals (the 0-8 and 1-9 seconds intervals) were examined from the total record length, which were 9sec. In both cases the algorithm chose the same intervals for feature selection.

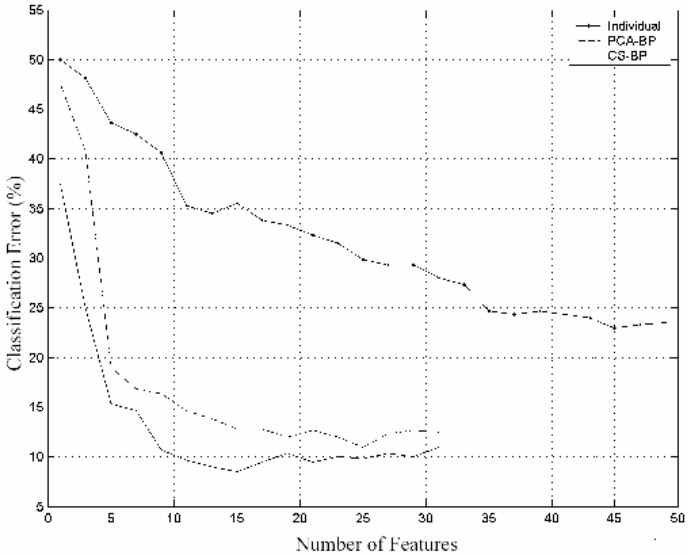


Figure 5.1. The classification error obtained by CS ordered individual coefficients, CS and PCA ordered band power coefficients of Subject 2 (CS = D)

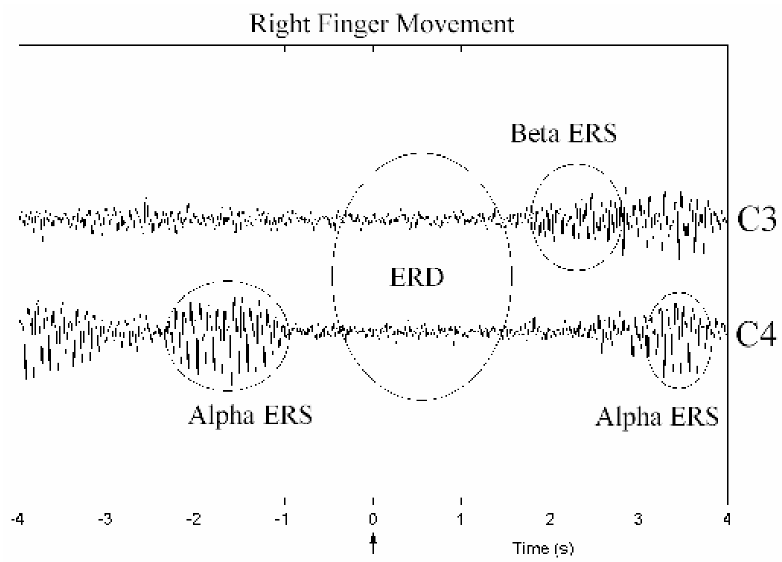


Figure 5.2. Single EEG sweeps of C3 and C4 electrodes during right hand finger movement. There is alpha ERS on electrode C4 during planning phase. Right after movement the beta ERS starts at electrode C3 and it is followed by a strong alpha ERS on C4. The arrow indicates the onset of the movement

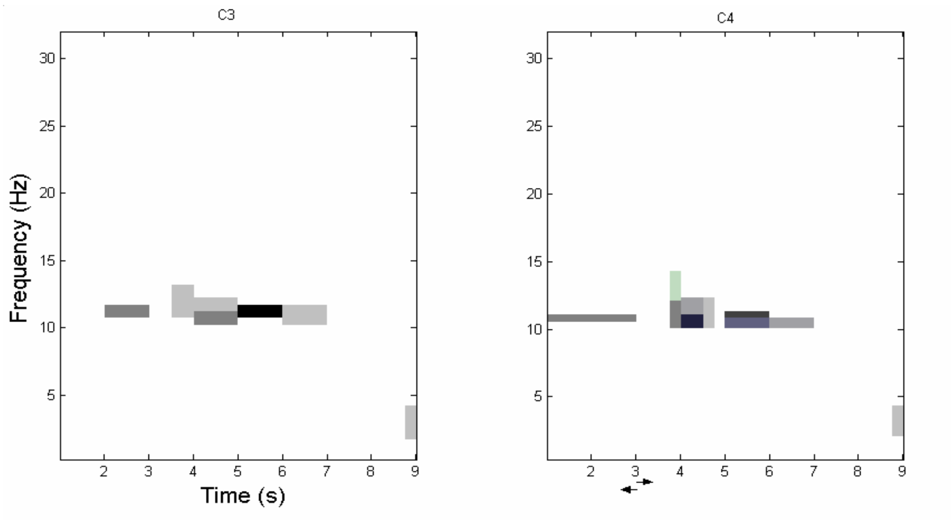


Figure 5.3. LDB features of imagery data. At sec 3 an arrow is displayed on the screen and the subject was asked to imagine left or right hand finger movement according to the side. The darker features have the more discrimination power

5.1 Discriminant Structures in Movement EEG

As stated in chapter 2, several time domain methods were used to visualize the ERD/ERS structures by taking the energy differences on C3/C4 electrodes. Specific bands were filtered and averaged over trials. Now let's have a look at the discriminant structures in a movement task which are extracted by LDB. The right and left hand finger movement EEG recordings are analyzed by LDB method with a tree depth of 5. Here the shortest segment is about 250ms. Because the *J, H, D* criteria were all calculated on the energy mean difference in the original LDB algorithm; the *F* criterion is also included in the feature sorting stage. Since *F* can consider the standard deviation of the selected indexes, this enables us to see the effect of different criteria. In addition *F* based sorting has provided better results in the classification which was mentioned previously. After constructing the segmentations the frequency indexes were grouped in 4Hz bins. The tree pruning process was also updated with a threshold. The mother space is kept only if it captures 95% of discriminant information of its subspaces. This has reduced the tree instability which may be caused due to small DP changes (Ince 2005b). Figures 5.4-6. represent the obtained T-F structures from 3 different subjects.

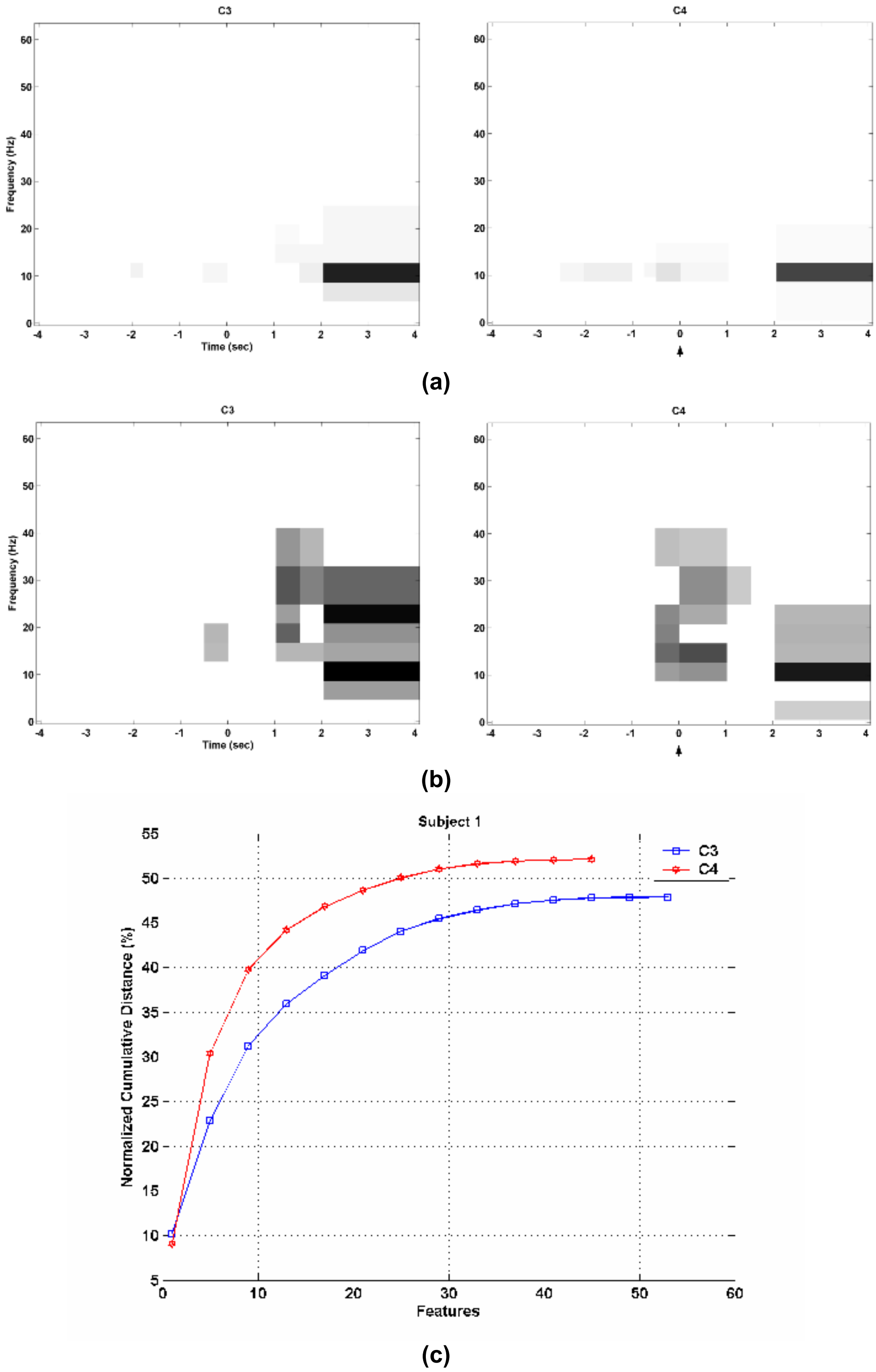


Figure 5.4. T-F maps of M1 obtained by (a) J (b) F criterion. Notice the change of the selected features. When F is used a wide band is selected during the movement. (c) The normalized cumulative distance of C3 and C4 electrodes when F criterion is used. The arrow indicates the onset of the movement

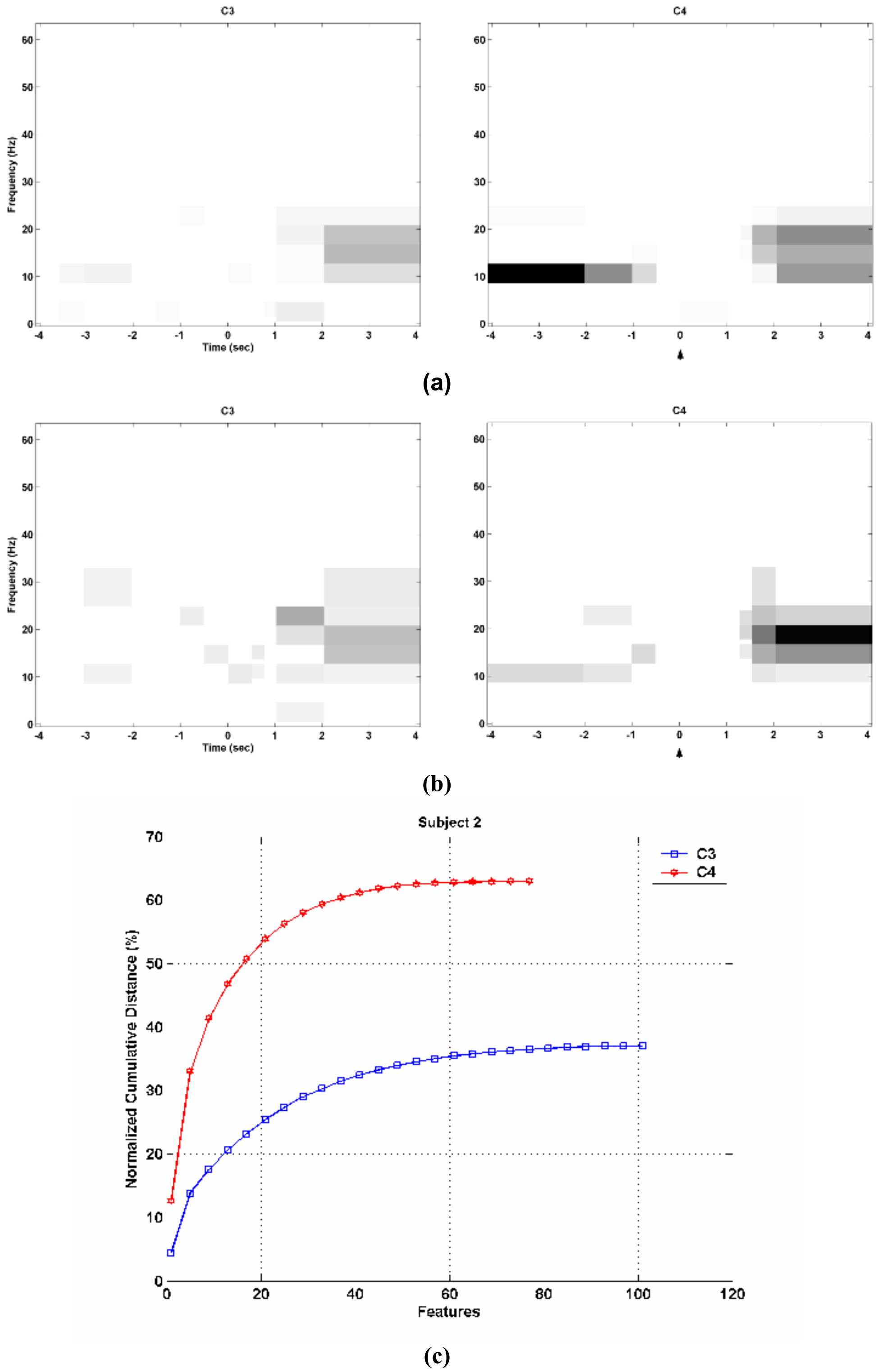


Figure 5.5. T-F maps of M2 obtained by (a) J (b) F criterion. Notice the change of the selected features. When F is used the post movement beta ERS is selected as the most discriminant features on both sides. (c) The normalized cumulative distance of C3 and C4 electrodes when F criterion is used. The arrow indicates the onset of the movement

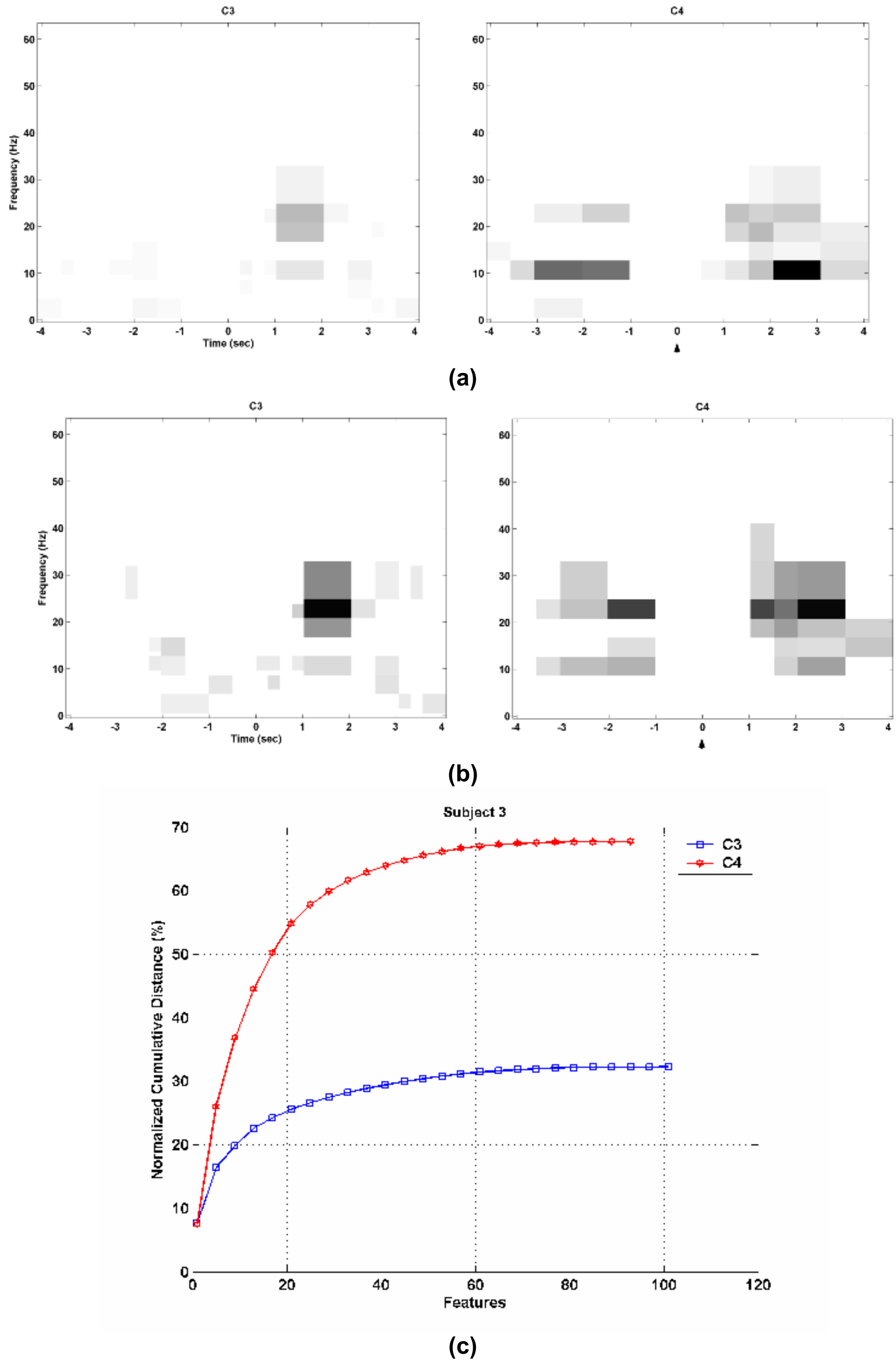


Figure 5.6. T-F maps of M3 obtained by (a) J (b) F criterion. Notice the change of the selected features. When F is used the post movement beta ERS is selected as the most discriminant features on both sides as in M2. (c) The normalized cumulative distance of C3 and C4 electrodes when F criterion is used. The arrow indicates the onset of the movement

Based on the obtained results, not only the time segmentation obtained, but also the selected features are subject depended. It is also observed that the top discriminative features come from different time locations (Figure 5.4., 5.5. and 5.6). With the J or D distance functions, most discriminative components are located in pre and post movement alpha-beta bands. When the feature selection criterion is replaced with the Fisher (F) criterion, the coefficients with the most discriminating power move to post movement beta synchronization. Furthermore, for subject S1 a wide frequency range is selected during the movement. The main characteristic of this region is the bilateral ERD. Due to ERD the energy mean is also low. Since the F criterion does not take only the distance between the means into consideration, but also the variance of the distributions, as well, it is able to catch low energy differences, such as ERD in S1. On the other hand, features with large variances are pushed back in the sorted coefficient queue. Since the alpha components have large variances, their relative discriminant power is reduced when the F criterion is used.

It is also noticed that there is an asymmetry between C3 and C4 features. For subjects 2 and 3 the DP of C4 is radically greater than C3's. Independent of the side of the finger, the left hemisphere is always involved in producing the movement. The right hemisphere seems to be involved only when the subject executes a left-hand finger movement.

5.2 C3-C4 Asymmetry

It has been reported that the left and right hand finger movements cause ERD and ERS events on the left and right hemisphere (Lopes 1999, Neuper 2001). The occurrences of these activities are mirrored on both hemispheres according to the side of the movement. It has been stated that the ERD starts earlier on the contra-lateral side during planning and becomes symmetric on both sides during the movement (Lopes 1999). In the post movement state, ERS occurs on both sides but with higher amplitude on contra-lateral side (Neuper 2001). Similar patterns are observed on these data sets. But the adaptive segmentation of C3 and C4 electrodes were different on all subjects. A large difference exists on beta ERS energy between left and right hand finger movements on subject M2. The energy of beta ERS of C4 for left-hand finger movement was much higher than the right-hand finger movement beta ERS. On the other hand the C3 electrode always shows similar ERS activity on both movement tasks. Another important difference was the pre movement alpha oscillations on the C4 side before right hand finger movement (see Figure 5.5). The ipsilateral ERS is already reported in MI studies (Neuper 1999; Neuper 2001). Here in the pre-movement stage same observations are obtained. The similarity between planning and preparation with the MI can be a reason for this asymmetry.

Here Local Discriminant Bases obtained from LCP are used to select the most discriminative components of an EEG signal corresponding to left or right hand index finger movements. The features are selected in 4Hz bins from an adaptive time segmentation of the EEG signal from C3 and C4 electrodes. The coefficients with the most discriminating power were located in the alpha and beta bands, justifying the almost universal selection of these two EEG bands for basic research on cortical mechanisms of movement, as well as in the efforts towards a BCI. Although in the previous section only alpha and beta ranges were utilized for classification here it can be seen that some discriminant frequency indexes can extend up to 40Hz. Also in some time indexes lower bands are selected. This will be considered in the next sections. When a distance on energy mean is used for feature selection, the alpha components are selected as the primary discriminant features. Interestingly, except for subject M1, no or minimal, features are selected in the ERD region where the movement is executed. When Fisher criterion is used, the top coefficients are always located on post movement components. Especially for subjects M2 and M3 the beta ERS is selected as the primary feature. Various authors reported that the beta ERS is more correlated with the movement variables and hand dominance (Neuper 2001). This may be a reason why they are selected. For S1 some features are selected in beta band right after the movement, which agrees with the burst activity characteristic of that component (Pfurtscheller 1996). These features are detected in a very short interval right after the movement. This explains why adaptive algorithms are suitable for capturing such temporal activities (Ince 2004, Ginter 2001). The algorithm constructed different segmentations for each subject as expected from the well-known inter-subject variability of EEG. The asymmetry between the hemispheres suggests a more comprehensive systematic study in which hand preference and hemispheric dominance of the subjects could be taken into account. This finding also correlates with a recent study where the asymmetry of two hemispheres is modeled with a Hidden Markov Model and a genetic algorithm (Obermaier 2001). The authors reported a significant improvement in classification accuracy when such an asymmetric classifier is used.

In Table 5.2. AR model and LCP-BP based results are presented. The classification error obtained by using LDB features is lower than 1sec AR and LCP – BP feature based approach. Since the same number of features is used in AR and LCP-BP for each hemisphere, the previously mentioned hemispheric asymmetry would not be taken into account. Also the previous T-F maps provide strong evidence of time-frequency plane spread features. As a result estimating the features from single windows might not always capture the relevant information for discrimination.

In this study the LCP-BP approach resulted in lower error than AR model although they were both calculated in 1sec. windows. It is believed that the smooth window can be a good approximator for local oscillatory activity. In addition the AR model is very sensitive to noise. However both algorithms have the minimum error in the post movement region (see Figure 5.7.). This is in accordance with the top LDB features which were adaptively selected in the post movement stage for each subject when F criterion was used.

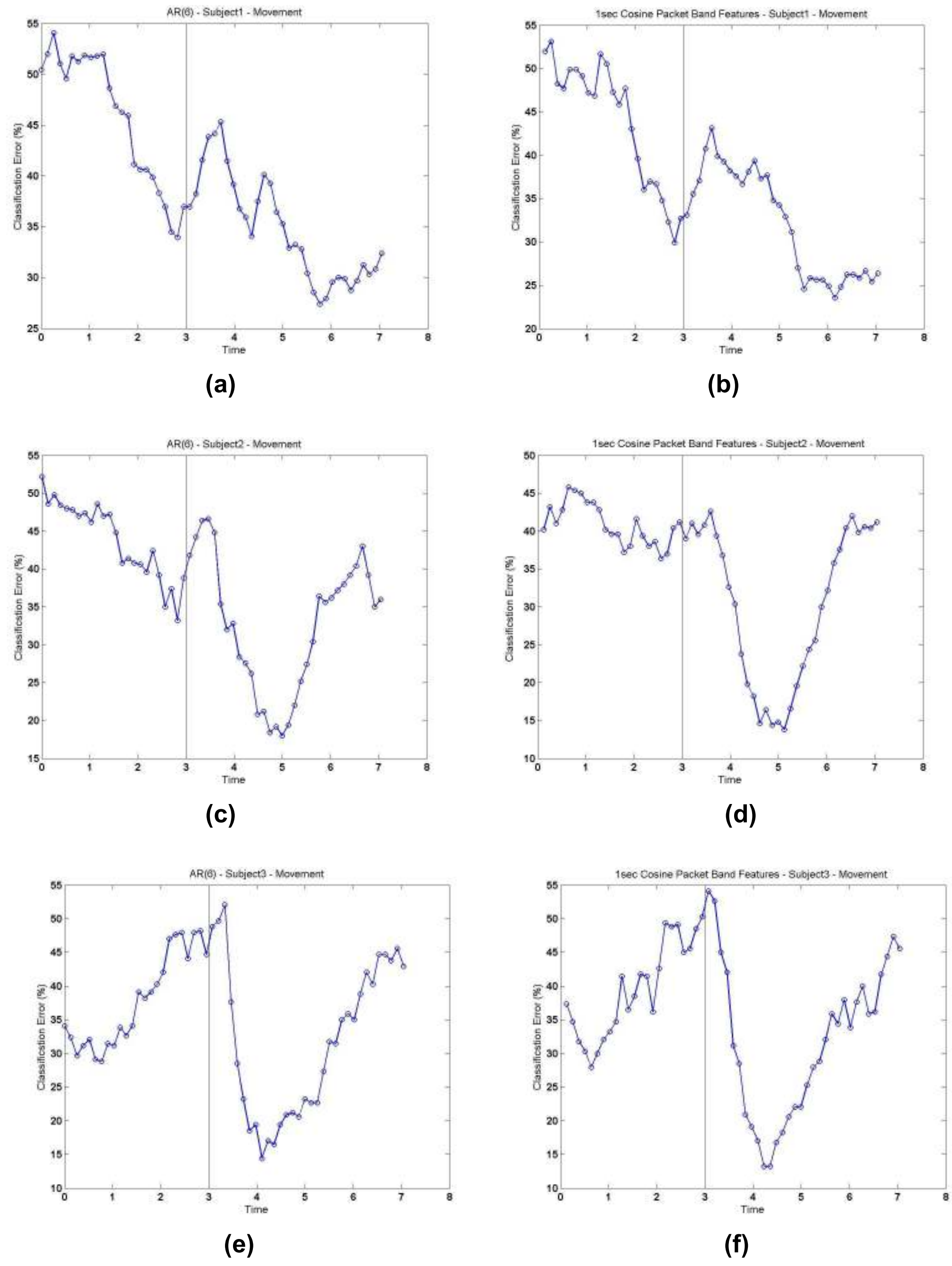


Figure 5.7. Time courses of classification errors (a), (c) and (e) AR and (b), (d) and (f) LCP Band Power classification curves. Notice that the lowest error is always obtained in the post movement region

5.3 Results of BCI Competition 2002 Subjects

Now the same algorithms, introduced in the previous section, are used to classify left and right hand finger movement imagery EEG data of BCI competition 2002. Table 5.2. represents the classification accuracy of the proposed methods for each of the nine subjects.

Table 5.3. Classification accuracy by using individual LDB coefficients, 4Hz bin LDB band powers, AR model with 6th order and fixed tiling band powers

	LDB	AR(6)	Fix-32sampl
Subjects	4-4Hz		4-4Hz
S1	80.00	81.1	80.00
S2	91.70	86.3	93.90
S3	66.40	69.9	70.00
S4	70.50	71.3	70.60
S5	67.80	74.7	73.90
S6	82.00	79.8	81.70
S7	82.20	79.2	81.70
S8	62.70	71.5	62.90
S9	78.40	76	77.80
Avg.:	75.7	76.6	**76.9**

The LDB algorithm based on the energy mean could not get better results than AR model and fixed window length based approaches. In the previous section it has been already shown that the LDB and PCA algorithms are very sensitive to energy variations. Since available dataset is used without rejecting any type of artifact, several sweeps with significant energy variations exist. Also on some subjects there is great energy difference between the rhythmic components of EEG. Therefore the most energetic ones have higher effect on the distance. This problem will be handled by so called “Improved LDB” in the next section. The effect of energetic sweeps/bands will be shown by an experiment.

5.4 Improved LDB

In (Saito 2002) several types of LDB were proposed over a dyadic tree by using wavelets or local trigonometric bases. The LDB-I procedure is based on the distance of the energy mean of the expansion coefficients of each class. However, sometimes the mean may not represent the real discriminant information. Therefore the so called Improved LDB (LDB-II) based on the distance between probability densities (PDF) of expansion coefficients was developed. Because PDF estimation via histogram is difficult one can also use cumulative distribution (CDF) instead (LDB-III). The LDB-III algorithm can be summarized as follows:

Step 1: Expand each training signal into Wavelet or Cosine Packets coefficients over the dyadic tree.

Step 2: For each expansion coefficient calculate the distance between CDF for each class and accumulate the distances of expansion coefficients in each subspace.

Step 3: Prune the tree from bottom to top via maximizing the cost function.

Step 4: Order the expansion coefficients from the pruned tree and select the top $k << n$ coefficients for classification where n is the dimension of the training signal.

Previously the sensitivity of LDB-I to artifacts and outliers was already mentioned. Now how improved LDB can overcome this problem will be shown with a simple experiment. Two synthetic classes (X, Y) are generated by filtering the white noise in 8-12Hz range. Each of them is 256 samples long. Then the amplitude of the classes are modulated to generate a region with lower energy. Both the energy mean and probability distance based LDB algorithms were run. See Figure 5.8. for generated classes and obtained segmentations.

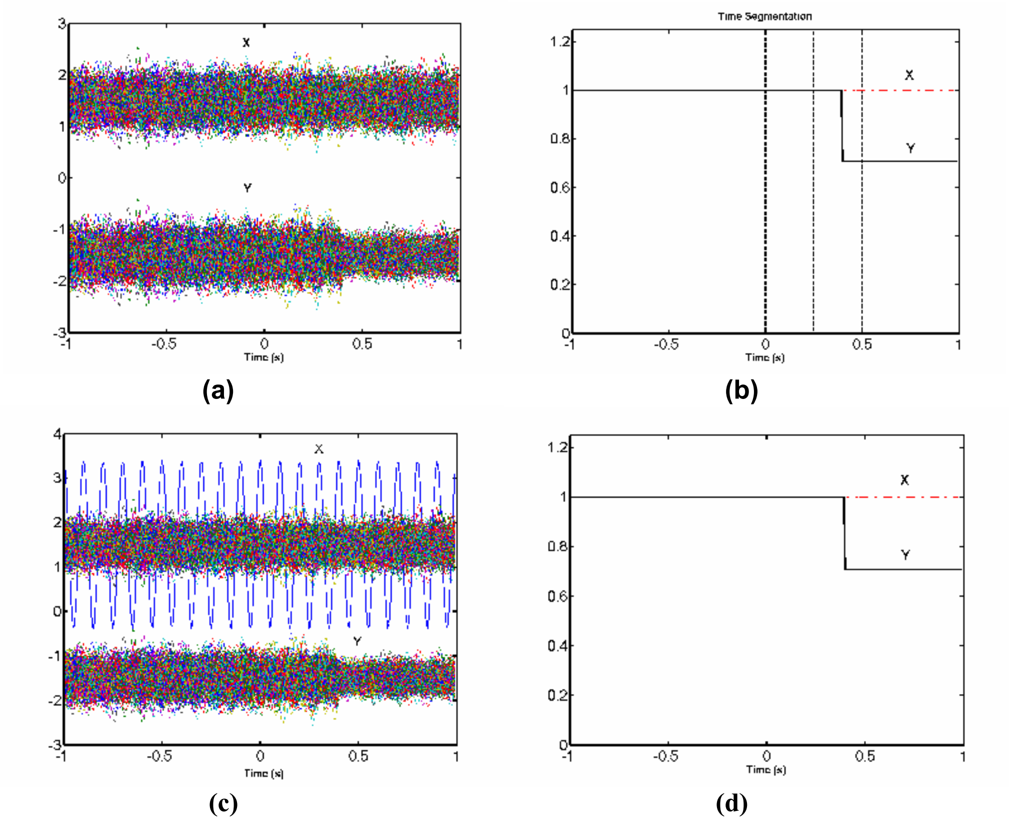


Figure 5.8. (a) Classes X and Y. (b) The modulating envelope of X, Y and LDB segmentations. Both algorithms resulted in the same segmentations. (c) A single sweep with high energy is included to class X (d) Energy mean based segmentation. The data is not divided. The top coordinate was selected

The LDB-I and LDB-II have yielded the same segmentation on the original feature set. However after including a single sweep with a striking energy difference caused LDB-I to fail. The LDB-II has resulted in the same segmentation despite the outlier. As such the probability distance based LDB-II- "improved LDB" - can deal with high energy variations and outliers and will be used from now on.

5.5 Spin Cycle Procedure

As mentioned previously WP and CP do not satisfy shift invariance property meaning that a shift in the original signal can cause changes in the expansion coefficients. This makes the WP and CP hard to use for pattern recognition. Thus it is preferable to make the expansion coefficients less sensitive to translations. For this the so called "spin cycle" procedure is used (Coifman95, Saito02). The Spin Cycle procedure expands the training set by generating its shifted versions in both directions in a circular manner. If the desired number of shift is τ then the training set is expanded to its, $(-\tau, -\tau+1, ..., \tau)$, $2\tau+1$ copied version.

5.6 A Decision Algorithm

During the implementation of the LDB-II algorithm it was observed that the small coefficients have significant effect on the discrimination power of each subspace. They mask the real discriminative coefficients and cause unstable trees. Furthermore, the dyadic grid also limits the applicability of the basic procedure. The subject specific features do not necessarily occur in dyadic segments. To decide whether the available training set is suitable for an LDB-II based analysis or not, a decision algorithm (DA) was developed and is summarized below.

Step 1: Prune the tree for DP minimization, (MinTree)
Step 2: Prune the tree for DP maximization, (MaxTree)
Step 3: Calculate the features for each pruned tree and sort them according to their DP in descending order.
Step 4: Select the top 5-10 features from each tree and calculate the gain (G).

$$\Delta g = \left(\sum_{i=1}^{K} dp_i^{\max} - \sum_{i=1}^{K} dp_i^{\min} \right) \tag{5.1}$$

$$G = \frac{\Delta g}{\sum_{i=1}^{K} dp_i^{\min}} = \frac{\sum_{i=1}^{K} dp_i^{\max}}{\sum_{i=1}^{K} dp_i^{\min}} - 1 \qquad K = 5,..,10 \tag{5.2}$$

where $dp_i^{\max}, dp_i^{\min}$ are the discrimination power of max and min tree features respectively.

Note that using the reverse pruning, where the distance criterion is minimized, yields a tree in which the classes overlap the most. The gain G has interesting connections to the split criteria of classification trees such as C4.5 (Quinlan93). Criterion C4.5 uses consecutive splits to separate classes that completely overlap at the top level. Each feature is evaluated by its entropy gain. Here G is used as a measure to decide whether to use a dyadic LDB tree (G>=0.25) or a Local Cosine Packet line search (G<0.25). Table 5.4. shows the gain for each subject in the data set used and represents the results obtained by the introduced algorithms.

The proposed procedure can be summarized now. The procedure is also described in the block diagram in Figure 5.9. The available training set is first processed by the decision algorithm in order to find out whether the data set should be analyzed using a dyadic LDB or an LCP line search. When G≥0.25 the EEG signal is analyzed with LDB then the band features are extracted. In order to reduce the dimension of the resulting feature set, the top subset of CS ordered features is projected onto its principal components. Finally, LDA is used on the reduced PCA features. When G<0.25 the EEG signal is analyzed with a smooth window Local Cosine Packet analysis. Next, band features are obtained and these features are sorted according to their DP. Finally, LDA is used on the ordered coefficients. An analysis window of 512 samples and a tree depth of 4 are selected where the deepest segment is 320ms. For PCA typically the number of features is in the range of 24 to 48 because most of the discrimination power is concentrated in these limits. 10 fold cross validation is used to achieve the average classification accuracy.

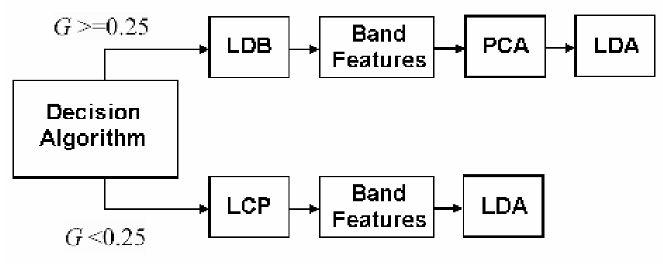


Figure 5.9. The block diagram of the complete system. The decision algorithm switches between LDB and LCP-Line Search method by checking the gain (G)

Table 5.4. The average classification rates and selected algorithm by DA for each subject. STD and BF represents the results obtained by using individual CP expansion coefficients and 4Hz bin band features respectively. Sp3-BF means 3 spin cycle BF sorted by CS=F

	AR	LDB	DA	
	(p=6)	STD	GAIN	ALG.
S1	81.1	74.2	2.05	LDB
S2	86.3	89.7	2.9	LDB
S3	69.9	62.2	**0.03**	**LCP**
S4	71.3	60.3	**-0.07**	**LCP**
S5	74.7	67.2	1.21	LDB
S6	79.8	79.2	0.61	LDB
S7	79.2	72.2	3.19	LDB
S8	71.5	67.8	**0.01**	**LCP**
S9	76	67.8	0.33	LDB
Avg.	76.6	71.2	-	-

	LDB	LDB	LDB	LCP	
	BF	Sp3-BF	Sp3-BF-PCA	Sp3-BF	OVERALL
S1	77	79.7	80.5	-	80.5
S2	93.6	91.6	91.4	-	91.4
S3	-	-	-	75.6	75.6
S4	-	-	-	71.4	71.4
S5	73	76.4	76.1	-	76.1
S6	84.5	87.5	87.5	-	87.5
S7	83.3	85.3	86.7	-	86.7
S8	-	-	-	71.3	71.3
S9	82.2	83	83.6	-	83.6
Avg.	82.3	83.9	84.3	72.8	**80.5**

The overall classification performance of the proposed algorithm outperformed the standard AR based approach. Note however that individual LDB coefficients do not represent the discrimination power due to the lack of shift invariance. Using this features without any preprocessing resulted in poor classification results. This was also observed in the classification of ME data. However by grouping the expansion coefficients in 4Hz bins the classification error was reduced by around 6%. It was empirically found that using 3 spins yields the best result. Increasing the spin to more than 3 did not improve the classification accuracy any further. For subjects S3-S5 the LDB based segmentations have a gain close to zero. During the iterations it was observed that DA selected random subspaces while maximizing the cost function for these subjects. A similar behavior was observed for LDB segmentations. This can be explained by the fact that the algorithm is trying to maximize the cost function by using all expansion coefficients rather than those ones which represent the rhythmic components of EEG. Therefore it is strongly recommended that the interested reader use an algorithm selection procedure such as the one described here before running improved-LDB. On subjects S3-S5 the algorithm switched to the LCP branch. Despite the switch the LDB is applied on these subjects to see how it performed. The results are given in Table 5.5. On these subjects the AR and LCP methods outperformed the LDB approach. Note that the LCP band features achieved the lowest error. One reason for this can be that the AR approach sometimes models the noise.

The LDB algorithm constructed different time segmentations for different subjects. One of the interesting points is that not only between subjects, but also on the same subject between the two hemispheres (C3, C4) they are represented with distinct segments and features. In Figure 5.10. the selected sub-dyadic tree, and the most discriminant time frequency features for two subjects are visualized. Note the differences between the feature characteristic on both electrodes and their discrimination power. It can be seen that the developed algorithm does not only adjust the T-F plane but also weights the space. With the other standard methods, equal numbers of features are selected from each hemisphere by assuming the activities are mirrored according to the side of the imagery. However obtained results indicate that both hemispheres do not contribute in the same manner to the imagery task. Since only two electrodes are used it is impossible to say anything about the behaviour on the posterior and frontal cortices which suggest a new study.

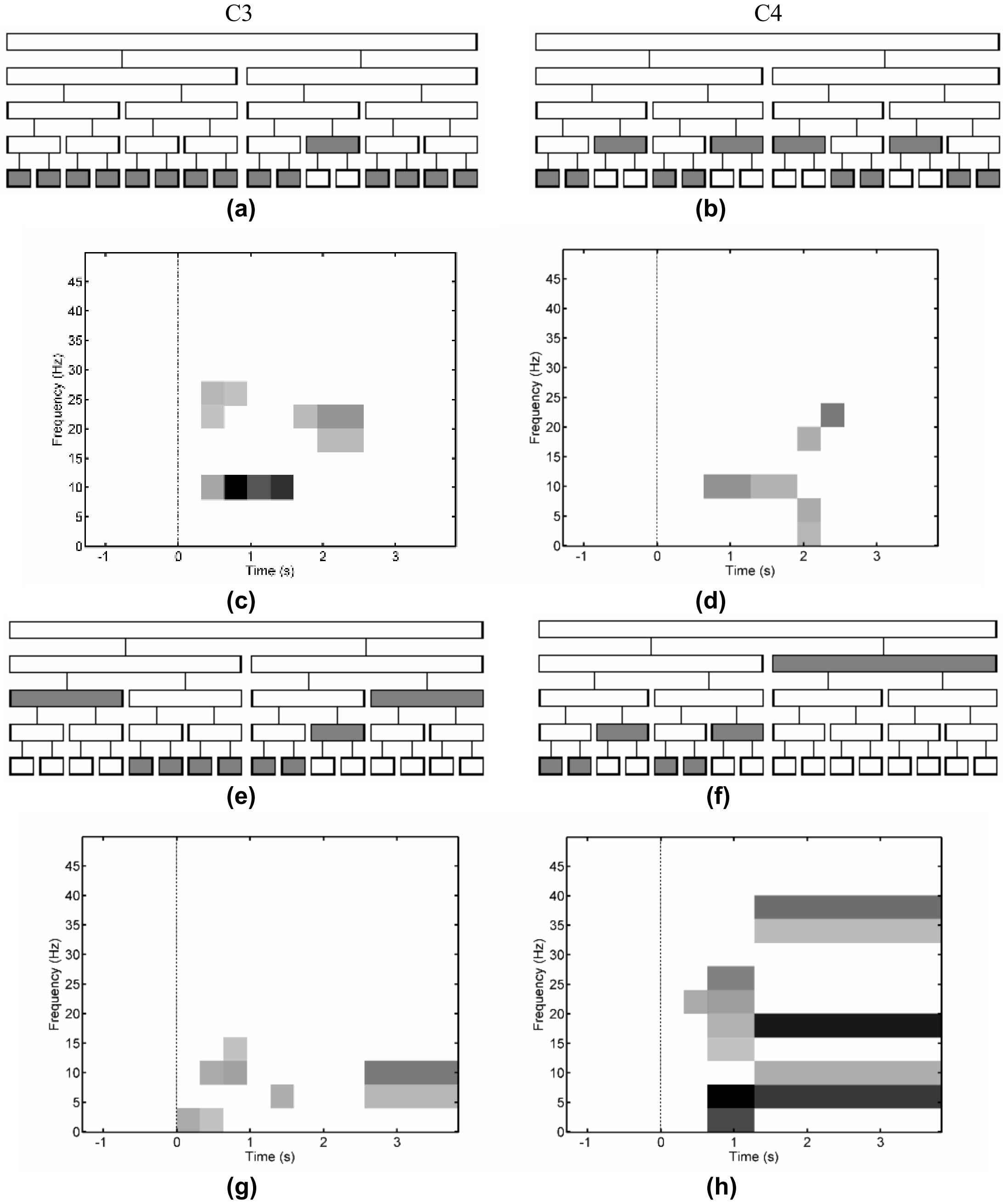


Figure 5.10. (a) (b) (e),(f) LDB tree of S2 and S6 for C3 (left column) and C4 (right column) electrodes respectively. (c), (d) (g), (h) the time-frequency features for the subjects S2 and S6. The dark features have the more discrimination power. Notice the difference of discrimination power and the feature characteristic of each electrode. Also for each subject different time and frequency points are weighted. The vertical line represents the time point of left-right queue

Although the main improvement in classification is accomplished by merging the expansion coefficients, it was observed that the error curve is affected when Spin Cycle and PCA features are used. For S6 and S7 the minimum classification error is obtained by using a small number of coefficients when PCA features are used. The decorrelated nature of PCA components can be a reason for this.

Also adding more features in the original feature space decreased the generalization capability more than when Spin Cycle ordered features are used. The Spin Cycle approach preserves its generalization capability in high dimension. Since each shift provides a new copy of the feature set, the classifier is immune to high dimensionality that may have been produced by translated versions of the feature set. The LDB algorithm is not without shortcomings. The most important of these is the instability of the constructed trees. On some subjects changing the training set resulted in different types of pruned trees. Similar results are obtained when classification trees are constructed.

In (Breiman 1984) the authors reported that although the classification tree structure was changed, they achieved similar error rates by using different paths. Similar behavior is observed. However such a behavior makes it hard to interpret the selected features that achieve a minimum error classification. Since each tree results in a different classification error curve, the min-error sometimes is canceled out by the cross validation procedure. When an online implementation is used, one can pick a single tree that results in the shortest path to minimal error. This in turn gives higher generalization capability.

Table 5.5. The classification accuracy of AR, LCP-BF and LDB-BF methods

	AR	LCP	LDB
	(p=6)	BF	BF
S3	69.9	74.5	71.1
S4	71.3	69.4	68.6
S8	71.5	71.4	61.4
Avg.	70.9	**71.8**	67.0

5.7 Nondyadic Segmentations – The Flexible LDB

The best base and LDB methods are selected over a dyadic tree by a divide and conquer algorithm. However there is no guarantee that the local patterns will occur in dyadic segments. In case the analyzed data does not match to the tree structure than both algorithms results unpredictable arbitrary trees. Therefore a more flexible segmentation algorithm is desired. In (Wesfreid 1999b) authors used the Fang algorithm to overcome the dyadic limitations. The signal was analyzed sample by sample by a mother and two children which satisfies the orthogonality condition given in the equation (3.10). For each time point Local Cosine Packets are calculated in both mother and children segments. On the resulting expansion coefficients the entropy change function $E_{ch}(i)$ is calculated as

$$E_{ch}(i) = E_M - (E_{C1} + E_{C2}) \tag{5.3}$$

where E_M, E_{C1} and E_{C2} represent the entropy of mother and children subspaces respectively. However authors have shown that this function oscillates due to the lack of shift variance. Therefore they used a post filtering and finally a local maximum detection algorithm to detect the segmentation point.

In (Wu 1999) an alternative approach is introduced to construct flexible segmentations for the compression of geo-acoustic waveforms. First a time cell is selected where its length defines the resolution of the current procedure. Then the signal is decomposed in LCP coefficients in this cell and as well as in its neighbor and in a window which is the union of both sub cell. This is exactly the same structure of parent and children smooth windows. In each cell and in their union the entropy is calculated. This is followed by comparison of the entropy of the mother and its children. In case the entropy is minimized in the mother space then the children are merged and the merged unit is used as the left child in the next step. As a result the proposed algorithm merges or divides the time line from left to right and vice versa (see Figure 5.11.). This approach can be easily applied to LDB procedure.

Let $\varsigma_i^M, \varsigma_i^C$ represent the distances between given two classes on the mother and children spaces.

$$\varsigma_i^M = D\left(p_M, q_M\right) \tag{5.4}$$

$$\varsigma_i^C = D\left(p_{c_i}, q_{c_i}\right) + D\left(p_{c_{i+1}}, q_{c_{i+1}}\right) \tag{5.5}$$

Then the following rules are applied

$If\ \varsigma_i^M \geq \varsigma_i^C$

Merge C_i, C_{i+1}

$C_i = M_i$

else

$Divide\ M_i$

$C_i = C_{i+1}$

Here there is no limitation for the length of the constructed segmentation. It is formed by the multiples of the cell. Compared to the algorithm in (Wesfied 1999b) the *Merge/Divide* approach has several advantages. First of all since the line is analyzed in discrete segments the algorithm is fast. In addition each time the comparison is implemented on the adjacent windows where the harm of the lack of translation invariance is reduced.

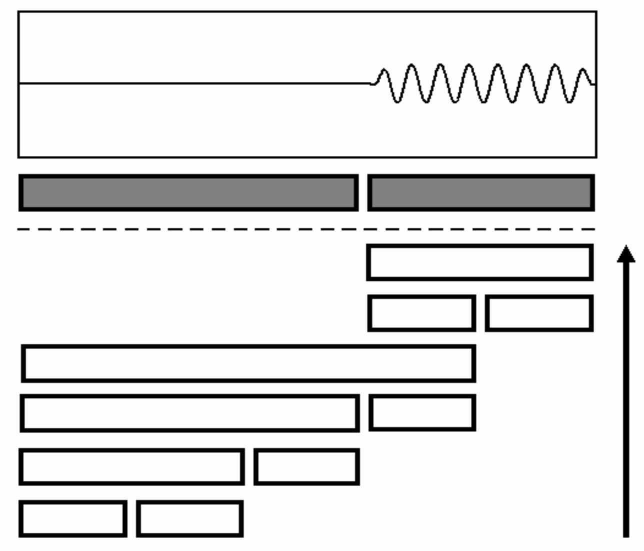


Figure 5.11. Merge and Divide algorithm is segmenting the signal from left to right via entropy minimization. The gray boxes represent the selected intervals

5.8 Frequency Band Clustering

It was already mentioned that LDB algorithm ends up with a feature space equal to the dimension of the original space, which results in high dimensionality. In order to reduce the dimensionality and the lack of shift invariance, the expansion coefficients were grouped in alpha or beta band ranges or in 4Hz frequency bins between 0-40 Hz in each segment. However the selected frequency bin may not be suitable so that the discriminant information is maximized in that selected scale. In order to construct partitions in frequency axis we reshape the Merge/Divide which in turn extracts relevant bands. A similar approach is implemented in (Akgul 2005) to find spectral subbands of near infrared spectroscopy recordings.

The procedure can be summarized as follows:

Step 1: Construct adaptive time segmentation by using the Merge/Divide.

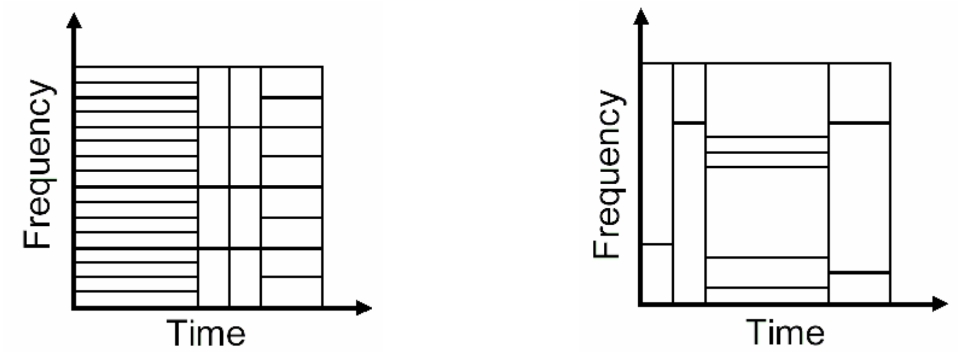
Step 2: Then in each time segment calculate the distance of the PDF of the consecutive frequency indexes and as well as in their union between each class.

Step 3: Merge those indexes where the union has larger distance then their selves.

Step 4: In the selected frequency segmentation calculate features and order them according to a given class separability criterion and select the top $k << n$ coefficients for classification.

Here consecutive frequency indexes are merged only if their union has larger discrimination power than their selves. The described procedure is basically a clustering approach obtained from cost function maximization, which in turns results in an adaptive frequency segmentation for discrimination. The resulting tiling is obtained by a combination of LDB and the clustering approach which is represented in Figure 5.12. The constructed time-frequency tiling is arbitrary and has interesting connections to (Herley 1993). In (Herley 1993) the authors used a double tree approach which was also limited to a dyadic grid.

Here this improved algorithm is applied on the same 9 subject data set. One more time 10 fold cross validation is implemented to achieve average classification accuracy. Tables 5.6 and 5.7 give the obtained results.



(a) **(b)**

Figure 5.12. The time frequency tilings (a) with dyadic tree (b) Non-dyadic time segmentation (Merge/Divide) followed by frequency axis clustering. Note the constructed segmentation in (b) can not be achieved by a tree structure

Table 5.6. The classification accuracy (%) of Dyadic and Flexible LDB. Also results using a 6[th] order AAR model is given

	Dyadic LDB	Flexible - LDB	AAR (p=6)
			mu=0.01
Subjects	Acc. (%)	Acc. (%)	Acc. (%)
S1	82.4	83.6	81.5
S2	91.2	92.6	88.5
S3	65.4	70	67.5
S4	66.2	70.7	71
S5	77.1	77.8	71
S6	81.9	87.2	78.5
S7	87.5	89.7	80.5
S8	62	70	67
S9	80.3	83.7	81.4
Avg	77.1	**80.6**	76.3

Table 5.7. The classification accuracy obtained by ordering the features with a CS (=*F*) criterion or PCA. Notice that PCA uses less features, but the average accuracy is nearly the same

	Flexible - LDB			
	CS		PCA	
Subjects	Acc. (%)	NoF	Acc. (%)	NoF
S1	83.9	13	83.6	12
S2	92.3	20	92.6	19
S3	72.4	8	70	12
S4	69.7	51	70.7	25
S5	77.4	46	77.8	43
S6	86.4	30	87.2	9
S7	88.5	34	89.7	10
S8	69.4	47	70	43
S9	84	9	83.7	7
Avg	80.4	29	**80.6**	**20**

The Non-Dyadic LDB updated with frequency axis clustering algorithm outperformed all other methods. It was observed that the algorithm has constructed different time-frequency tiling for different subjects. One more time not only subjects but both hemispheres are represented by distinct segmentations and features. The segmentations and the discrimination power of each electrode for representative subjects are visualized in figures 5.13., 5.14. and 5.15. By assuming that these electrodes represent the hemisphere where they are located, it can be concluded that both hemispheres are not involved symmetrically in the imagery task. In literature it has been always assumed that ERD/ERS will mirror on both hemispheres according to the side. However current findings do not support this. Since equal numbers of features from both channels are included, the standard AR /Band power methods may not be able to model this hemispheric asymmetry.

Another issue, which should be mentioned, is the selected features. As seen from figures 5.13., 5.14. and 5.15., most of the time wide bands are selected. Although subject S7 has very strong alpha activity on the C3 side, beta band is selected as the first feature. It seems not all of the oscillatory components carry the same or similar discrimination power. Since AR based methods will model the peaks on the spectrum, all energetic components will be directly included into the feature set.

Of course the LDB procedure is not with shortcomings. As mentioned previously the original LDB is limited to dyadic points. Here we overcome this limitation. However, both methods suffer from the effect of small expansion coefficients, which are mentioned in (Saito 2002). Many times the same information is repeated in the side lobes or in high frequency coefficients. The clustering procedure partly groups the correlated features. However, the effect of higher band is still present. One can apply a denoising procedure on the expansion coefficients. However, the imbalances between the energy levels of EEG bands make it hard to select a suitable threshold. An inappropriate threshold can remove a component which carries significant discriminant information.

Therefore to reduce the high dimensionality PCA is used prior to LDA. PCA can remove correlations between features. Therefore, same or better classification is obtained by using small number of features. On the other hand PCA is very sensitive to outliers. Especially the lower components show significant energy changes from sweep to sweep. This may eliminate the effect of PCA.

The occurrence of ERD and ERS events on the same electrode on different bands is previously reported by some authors (Pfurtscheller 2001). Here similar structures are observed. The algorithm successfully adjusted the tiling to the nature of each component on the same electrode location (see figures 5.13.-5.15.).

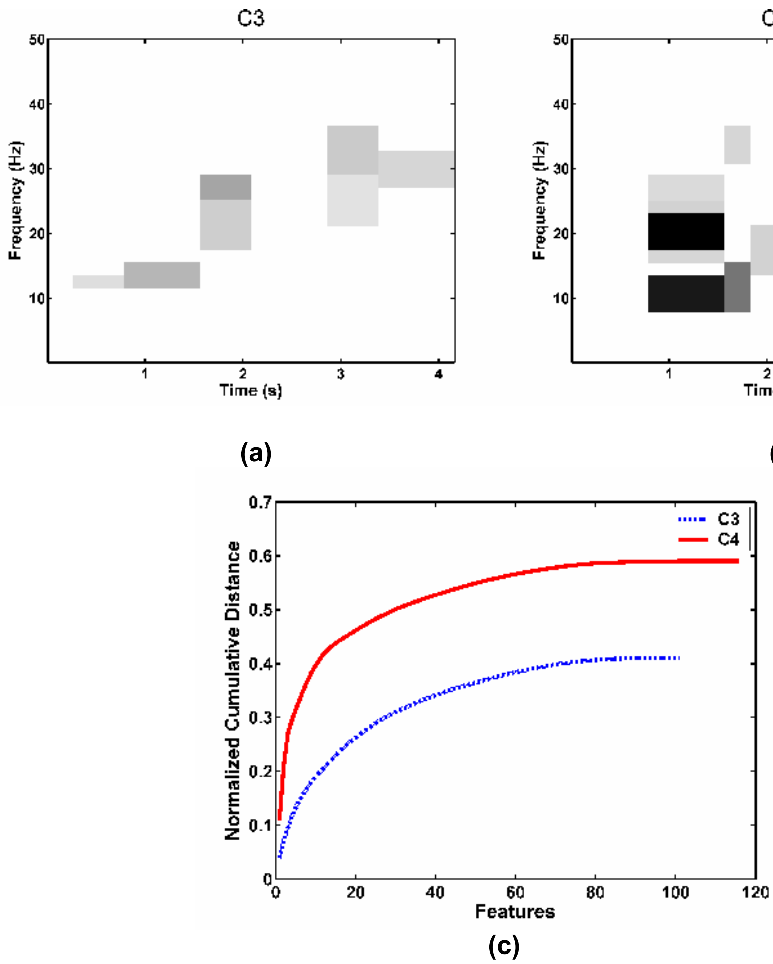


Figure 5.13. (a) and (b) The discriminant features of subject S1 from C3/C4 electrode locations. 416 samples are used in the analysis, starting from the Left/Right que which also includes the planning stage. (c) The cumulative discrimination power of C3/C4 electrodes. Notice the asymmetry between two hemispheres

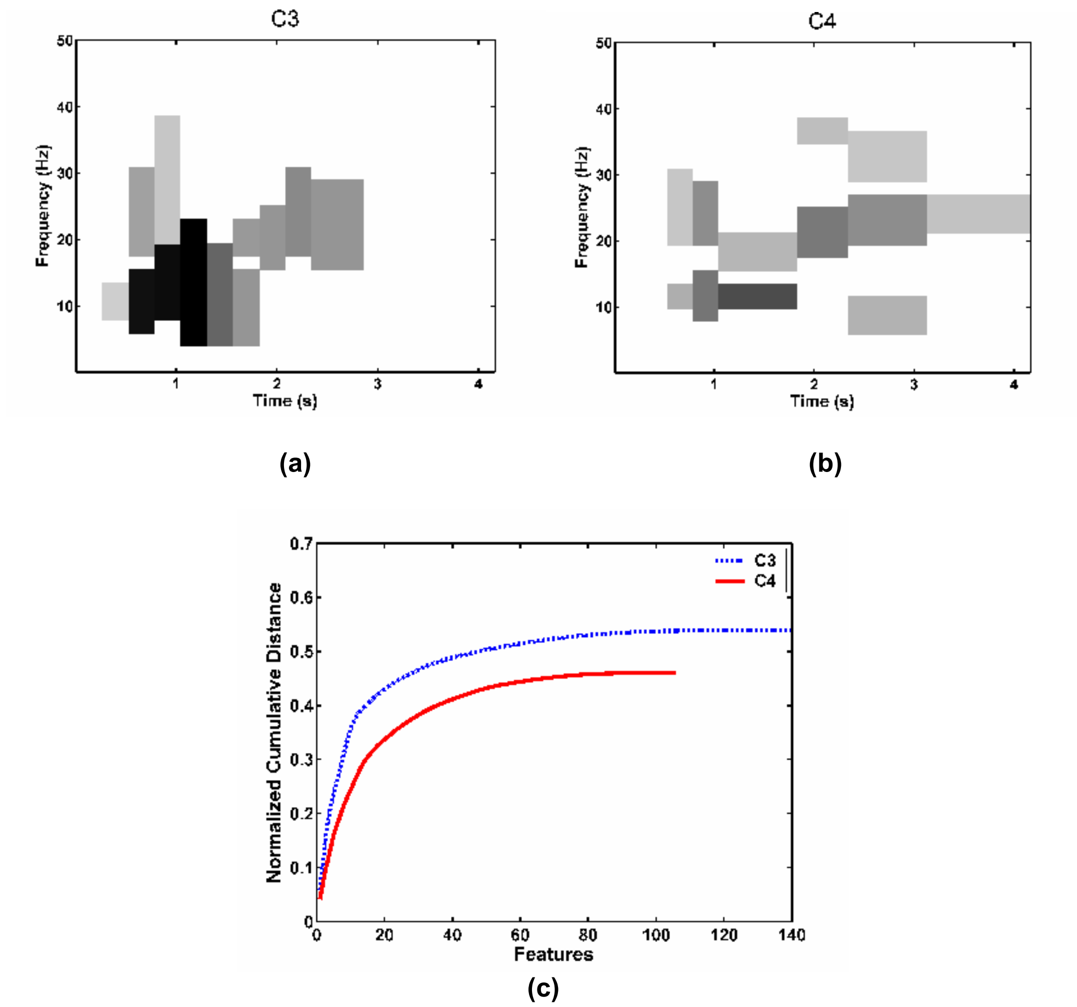


Figure 5.14. (a) and (b) The discriminant features of subject S2 from C3/C4 electrode locations . 416 samples are used in the analysis, starting from the Left/Right queue which also includes the planning stage. (c) The cumulative discrimination power of C3/C4 electrodes. Notice the asymmetry between two hemispheres

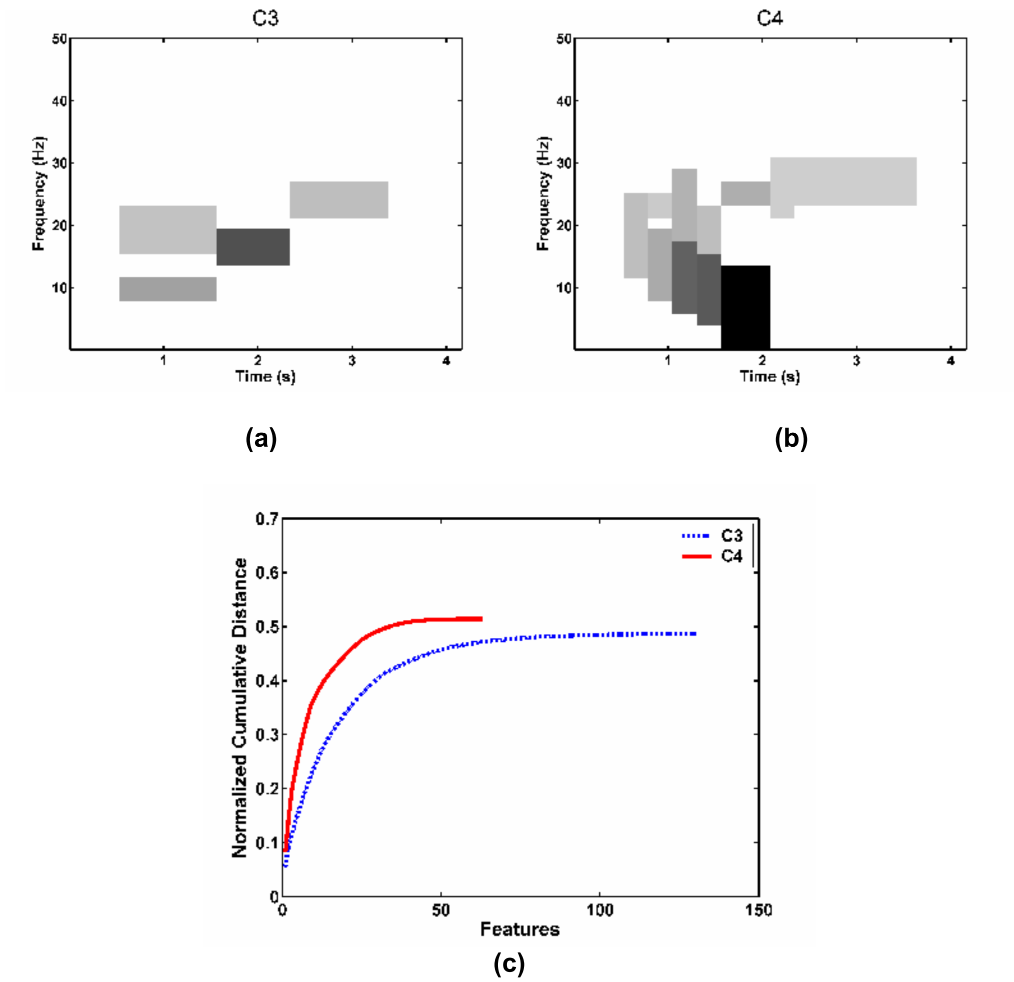


Figure 5.15. (a) and (b) The discriminant features of subject S7 for C3/C4 electrode locations . 416 samples are used in the analysis, starting from the Left/Right que which also includes the planning stage. (c) The cumulative discrimination power of C3/C4 electrodes. Notice the asymmetry between two hemispheres

The sections (a) and (b) in Figure 5.13. through Figure 5.15. represent the discrimination power of C3 and C4 electrode locations. The sections marked with (c) show the normalized cumulative sum of the discrimination power of the features from each channel. The darker features have the more DP. Note the differences between feature characteristics and their DP on each hemisphere. Features are constructed by arbitrary tilings where multiresolution representation is possible in each segment. Interestingly, in general, wide bands are selected.

5.9 Space-Time-Frequency Analysis

Obtained results indicate that the two hemispheres do not equally contribute to the classification task. This shows that space may contain additional information and can be used as another dimension to increase the classification accuracy. The Motor Imagery EEG recordings used in this study has been recorded from 59 electrodes which enable us to include different cortical areas for the analysis and classification. Therefore 21 electrodes indicated in Figure 5.16. were converted into surface Laplacian derivation and used for classification task. On each electrode the Flexible LDB algorithm was run and from the estimated adaptive segmentation the features are ordered according to their discrimination power as explained in the previous chapters. Since PCA has approached to lower error earlier than CS the top CS ordered feature set is projected onto its principal axes. Then the top features are fed to a LDA for classification. 10 fold cross validation is used to estimate the classification accuracy. Table 5.7. shows the obtained results.

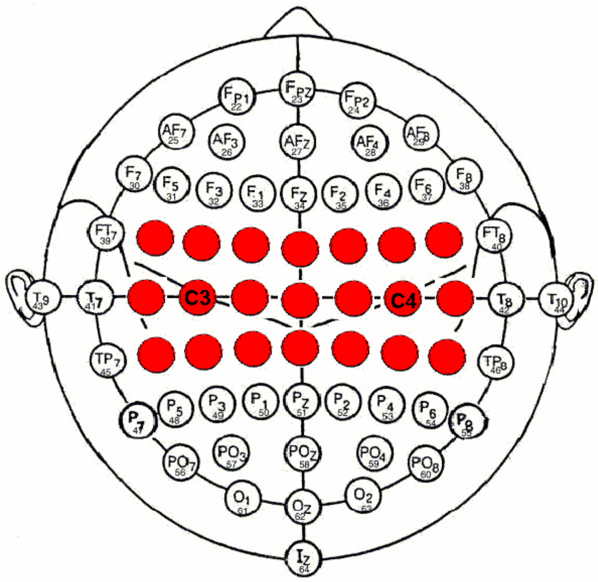


Figure 5.16. The electrode locations used in this study. Each one is converted into surface Laplacian derivation by using its 4 neighbor electrodes

Table 5.7. The results obtained from LDB based space-time-frequency analysis using 21 and just two electrodes are given. Best results are highlighted. In addition results obtained from a time-frequency based electrode weighting (TFW) algorithm is also presented. NoF stands for the number of features

	LDB		LDB		TFW
Subjects	21 electrodes	NoF	C3/C4	NoF	20 electrodes
S1	76.4	13	83.6	12	83
S2	94.3	12	92.6	29	91
S3	78.6	6	70	12	75
S4	81.6	14	70.7	25	78
S5	82	10	77.8	43	76
S6	88.4	10	87.2	9	77
S7	96.8	7	89.7	10	91
S8	80.1	6	70	43	71
S9	86.2	10	83.7	7	74
Avg	**84.9**	**9.8**	80.6	21	80

The Space-Time-Frequency analysis approach has improved the classification accuracy nearly 5%. The TFW column in Table 5.7. represents the results obtained by using an electrode weighting algorithm. Another popular approach is common spatial patterns (CSP) (Ramoser 2000). Same authors who developed TFW, reported an average classification accuracy of 76% on the same subject group by using CSP method. It is already reported that this approach is very sensitive to noise and outliers (Ramoser 2000). And also it requires prior knowledge of subject depended time point to calculate the features. Here the developed algorithm can weight the space automatically during the sorting of discriminant features before the classification stage. Prior to this step the LDB algorithm already adjusted the time and frequency for maximum discrimination. The main advantage of using the LDB on 21 electrodes is in latter case the electrodes with small or no discrimination power are discarded during the classification. This will enable to select the suitable ones in online experiments where the computational complexity will be reduced radically.

Besides classification the same algorithm is run on all available sweeps in order to visualize which electrodes are selected. This can also give information about the activation of different cortical areas and their contribution to discrimination (see Figure 5.17).

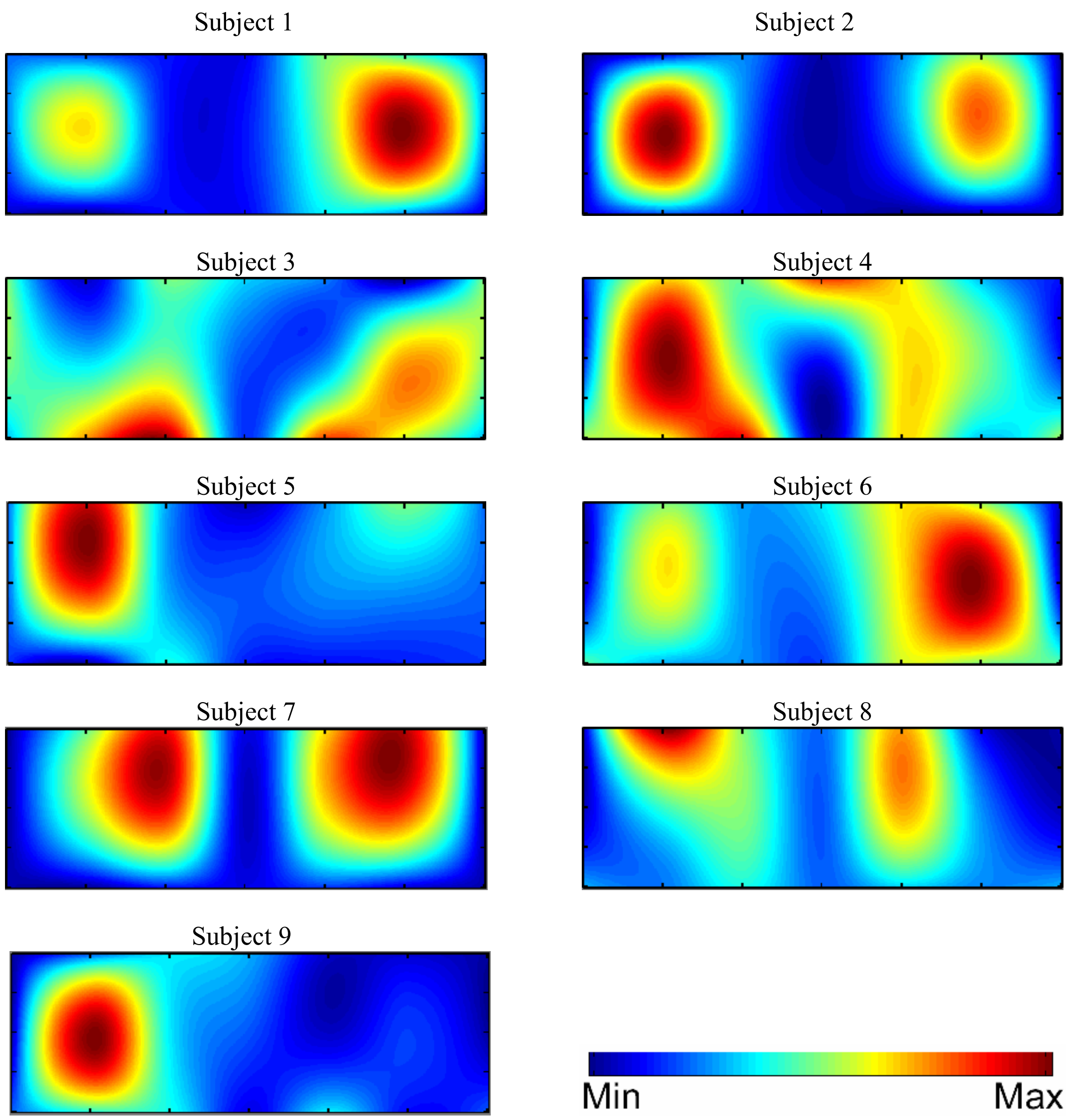
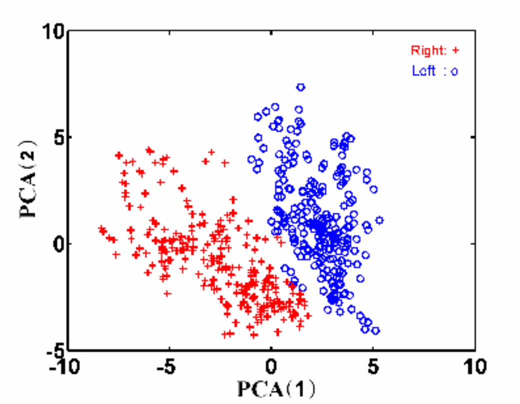


Figure 5.17. The topographical maps of 9 subjects estimated with the top 16 LDB features. Cubic interpolation was used to visualize the activities of 21 electrodes. Notice the differences between subjects

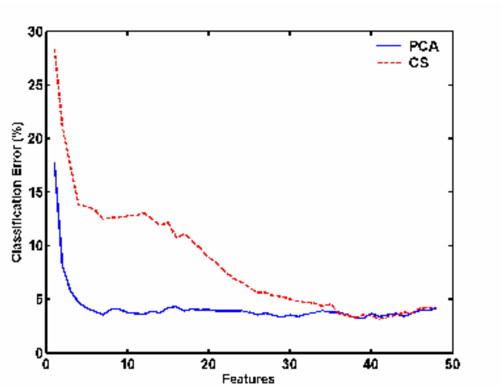
As seen from the topographical maps for some subjects the discriminant positions are not always located on C3/C4 area. Especially for Subject 4, Subject 7 and Subjects 8 the most discriminant locations were on the neighbor areas. Using features from these locations has improved the classification accuracy for each subject about 10%, 9% and 10% respectively. Also Subject 7 has the best classification result. Figure 5.18. represents the classification curve and distribution of the first principal components. The left and right classes are nearly linearly separable by the first 2 principal components.

However on the CS based error curve it can be seen it takes several features for the classifier to reach minimal error. This can be explained by the correlated feature characteristic of neighbor areas. Therefore PCA provides a fast convergence to minimal error by decorrelating the feature space.

Another important improvement is obtained with subject 4 and 8. For both subjects the selected locations were different than C3/C4. This explains the poor results obtained previously. Especially the selected features for subject 8 are quite interesting. For this subject Fc3 electrode location was the most discriminant one. In Figure 5.19. the selected LDB features and the energy mean differences obtained from STFT are visualized. For this subject the energy mean based distance can not represent discriminant structure. The improved LDB has extracted a wide T-F cell around 30 Hz which was invisible on the STFT map.

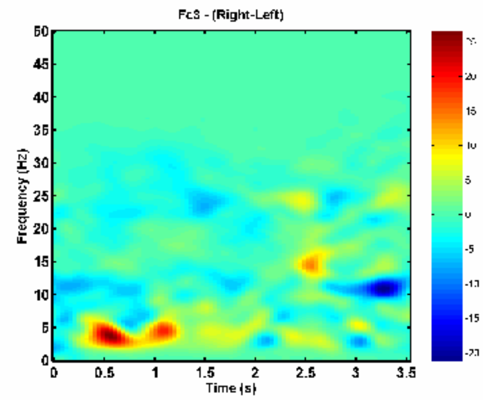


(a)

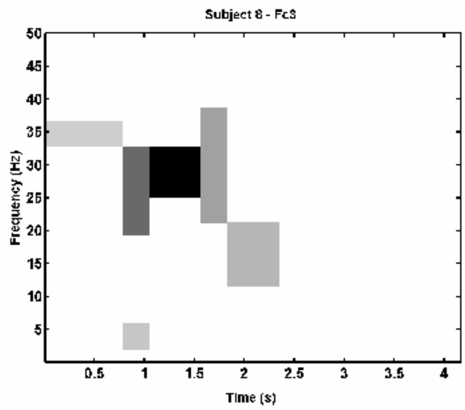


(b)

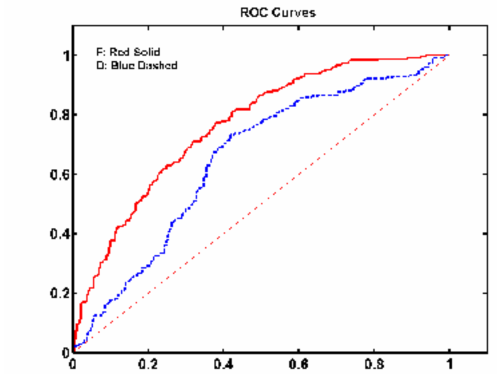
Figure 5.18. (a) The first two principal components of Left and Right classes. Notice that both classes are nearly perfectly separable by a linear discriminant, using just two principal axes. (b) The classification error of CS and PCA sorting. Notice that the PCA is applied on top CS sorted features achieves lowest error earlier than CS



(a)



(b)



(c)

Figure 5.19. The difference between the time-frequency maps of left and right hand imageries of Fc3 electrode location. (b) The T-F map obtained from Flexible LDB. The features are selected by using *F* criterion. The dark features have the more discrimination power. (c) The ROC curves of *F* and energy mean difference based top discriminant features. Notice the *F* selected feature has larger area and is not observable in (a)

6 RESULTS AND SUGGESTIONS

In this thesis an adaptive time-frequency approach is applied to analyze and classify EEG recordings accompanying real and imagery finger movements. It has been already emphasized that the motor imagery has strong relationship to the planning and programming stages of the movement. Therefore initially real movements are analyzed with short time Fourier transform to visualize ERD/ERS structures in the time frequency domain. Obtained T-F maps were smooth and provided general information of the characteristic behavior of various components of EEG. Then the so called Best Basis algorithm is applied to get adaptive T-F partitioning by entropy minimization. Resulting segmentations provided functional information about the behavior of both hemispheres according to the side of the movement. Especially the capability of adjusting windows according to time varying characteristic of EEG makes BB and LCP algorithms very suitable to analyze time locked oscillatory changes. However it was hard to get functional meaning from single sweep segmentations. Therefore the so called joint best basis, which is obtained by averaging expansion coefficients of each sweep, was used and it provided better results. Such an approach can only be successful, if the expansion coefficients are Gaussian distributed in the selected subspace. Therefore artifact free EEG recordings are needed and outliers must be removed.

As a further process Real and Imagery EEG recordings are analyzed by using LDB which is equipped with discriminant information cost. The LDB features of real movement provided interesting results. The obtained segmentations and features were different for each subject. Not only subjects but also the hemispheres are represented by different T-F tiling. Obtained discrimination power of the top LDB indicates that there exists a hemispheric asymmetry. This could be due to handedness. These results suggest a new direction of research. The used criterion may reflect underlying physio-anatomical structures and perhaps can be used in clinical applications.

The direct application of the original LDB to classify imagery recordings has poor results. This is due to the lack of shift invariance of Local Cosine Packets and the single trial variability of EEG recordings. Since all EEG sweeps were used and no artifact rejection was applied, the mean is influenced by energetic sweeps. Also since the subjects are imagining the movement, it is hard to obtain perfect time locked recordings. Another important aspect is the energy difference between the rhythmic components of EEG. In general the lower bands contain more energy and have higher effect on the distance compared to the high frequency components. Lower band is also more affected due to eye movements and fluctuations in electrode skin impedances. Therefore the so called improved LDB algorithm is used for segmentation. The improved LDB uses the probability densities of expansion coefficients by calculating the histogram in each subspace. One can also use CDF instead which is less complex but needs more computation. The PDF is more sensitive to shift invariance and this problem is solved by using the spin cycle procedure, which expands the signal by using its translated versions. Also this approach makes it possible to estimate the variance of each feature. This expanded feature set makes the classifier immune to high dimensionality and increases the generalization capability.

The flexible LDB removed the limitations of dyadic grid. This same algorithm is modified to cluster consecutive frequency indexes. This was a crucial step for dimension reduction. Both grouping the coefficients in fixed or flexible frequency indexes provided significant improvements in classification.

The discriminant structures and their discrimination power are visualized with both dyadic and flexible LDB. As in the analysis of real movements, C3/C4 electrodes and subjects are represented by distinct T-F segmentation and features. The discrimination powers of the C3/C4 electrodes differ significantly. Several authors reported the mirroring of the activities according to the side of the movement. Therefore in general, same numbers of features are included from C3/C4 in the classification task. However the findings in this thesis do not support this. Obtained results strongly indicate an asymmetric behavior in both real and imagery tasks. This can be due to the physio-anatomical differences. The selected C3/C4 location may not represent the real active space. Therefore the number of electrodes is increased to benefit from different cortical areas. Using 9 with the best discrimination power of 21 electrodes covering the central region as indicated in Figure 5.16., has increased the classification accuracy nearly 5%. In order to visualize discriminant electrode positions topographical maps are constructed by using the top 16 discriminant features with cubic interpolations. Striking differences were found between subjects (see Figure 5.17.). For example for subject S1, S2, S6 and S9 the most discriminant location was on either C3 or C4 electrode locations. However for the remaining subjects the most discriminant information came either from frontal or posterior locations. Especially adjusting the space for S7 and S8 has provided significant improvements on the classification accuracy. For these subjects neither the C3 nor the C4 electrode positions were the most discriminant ones.

Besides classification the obtained topographical mappings are also found to be promising. Since the current analysis methods were based on either filtering reactive bands or fitting an AR model they are vulnerable to unpredictable energy differences, outliers and energetic components which do not carry any discriminant information. Therefore it is believed that the proposed method can explore the T-F plane for each location better than traditional approaches. Dr. Y. SARICA[1] strongly suggests using this algorithm in parallel with other imaging devices to assess its efficiency.

In this study EEG records were used for analysis and classification. In the last years there is growing interest in recording and classification of invasive neural activities such as spikes and local field potentials (LFP). Especially LFP has similar properties with EEG. It contains several rhythmic components, as well as gamma rhythm, which is rare to obtain on EEG. Recent results obtained by analyzing LFP show that gamma oscillations carry significant information about the variables of the executed movements. The developed algorithm can be applied directly to LFP. Besides LFP, recording of spikes activities have also become possible. It was found that getting the activities of a group of neurons gives important information for decoding the movement.

---

[1] Prof. Dr. Yakup SARICA, Cukurova University, Department of Neurology, School of Medicine, 01330, Balcali, Yuregir, Adana/TURKEY

Therefore exploration of the neural code has gained significant interest in the last year to develop neuroprosthetics.

In general spike counts of neurons of fixed intervals are estimated and used in decoding for a BCI task. Since the SNR and spatial resolution is larger than EEG this type of research enables one to predict the hand position in 3 dimensions. The developed algorithms can be modified to select time windows to minimize prediction error for a neuroprosthetics. However, long term recording of single neuron activities is still a big obstacle. Therefore LFP might provide a better solution which already reflects the activity of a group of neurons.

This study shows that the adapted wavelet analysis can be used to analyze and extract physiological states embedded in EEG. This outcome may be applied to determine and discriminate diseases, and track their prognosis. Already, the study is applicable to a Brain Computer Interface system which uses movement imagination as a strategy. Obtained improvements promise a better performance to the current BCI systems.

REFERENCES

AKANSU A. N., HADDAD R. A., 1993, Multiresolution Signal Decomposition. Academic Press, New York.

AKGÜL C. B., SANKUR B., AKIN A., 2005, Spectral Analysis of Event-Related Hemodynamic Responses in Functional Near Infrared Spectroscopy, Journ. Of Comp. Neuroscience, 18, pp. 67-83.

ANDERSON C., STOLZ E., and SHAMSUNDER S. 1995, Discriminating mental tasks using EEG represented by AR models, Proceedings of the 1995 IEEE Engineering in Medicine and Biology Annual Conference, Sept 20--23, Montreal, Canada.

ANDERSON C., STOLZ E., and SHAMSUNDER S., 1998, Multivariate autoregressive models for classification of spontaneous Electroencephalogram during mental tasks. IEEE Transactions on Biomedical Engineering, vol. 45, no. 3, pp. 277-286.

BEISTEINER R., HOLLINGER P., LINDINGER G., LANG W., BERTLOZ A., 1995, Mental Representations of Movements. Brain potentials associated with imagination of hand movements. Electroenceph. Clin. Neurophysiol., 96(2), pp. 183-193.

BERGER H., 1939, Über das Electroencephologramm des Menschen. Arch. Psychiat. Nerv.-Krankh., 87: 527-570.

BIRBAUMER N., ELBERT T., ROCKSTROH B. and LUTZENBERGER W., 1981, Biofeedback of event related slow potentials of the brain, North Holland Publishing Company, pp. 389-413.

BIRBAUMER N., ROBERTS L., LUTZENBERGER W., ROCKSTROH B., and ELBERT T., 1992, Area specific self-regulation of slow cortical potentials on the sagital midline and its effects on behavior, Electroenceph. Clin. Neurophysiol., vol. 84, pp. 353-361.

BIRBAUMER N., 1999, Slow cortical potentials: Plasticity, operant control, and behavioral effects, The Neuroscientist, vol. 5, pp. 74-78.

BIRBAUMER N., GHANAYIM N., HINTERBERGER T., IVERSEN I., KOTCHOUBEY B., KÜBLER A., PERELMOUTER J., TAUB E., and FLOR H., 1999, A spelling device for the paralysed, Nature, vol. 398, pp. 297-298.

BISHOP C. M., 1995, Neural Networks for Pattern Recognition. Oxford, U.K.: Clarendon.

BREAIMAN L., FRIEDMAN J., OLSHEN R. and STONE C., 1984, Classification and Regression Trees. Belmont, CA: Wadsworth.

CETIN A. E., PEARSON T.C., TEWFIK A. H., 2004, Classification of closed- and open-shell pistachio nuts using impact acoustical analysis, ICASSP.

CHAPIN J. K., MOXON K. A., MARKOWITZ R. S., and NICOLELIS M. L., 1999, Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex., Nature Neuroscience 2, pp. 664-670.

COIFMAN R. R. and DONOHO D., 1995, Translationinvariant de-noising, Wavelets and Statistics A. Antoniadis and G. Oppenheim, eds., Lecture Notes in Statistics, Springer-Verlag, pp. 125–150.

http://liinc.bme.columbia.edu/competition.htm

DAUBECHIES I., 1988, Orthonormal bases of compactly supported wavelets. Commun. On Pure and Apll. Math., 41, pp. 909-996.

DECETY J., SJOHOLM H., RYDINGNE., STENBERG G., INGVER D.H., 1990, The cerebellum participates in mental activity: tomographic measurements of regional cerebral blood flow, Brain Res., 535(2), pp. 313-317.

DUDA R., HART P., STORK D., 2000, Pattern Classification, Second Edition, John Wiley&Sons.

DUHAMEL P., and VETTERLI M., 1990, Fast Fourier Transforms: a tutorial review and a state of the art. Signal. Proc., 194, pp. 259-299.

DUHAMEL P., MAHIEUX Y., PETIT J. P., 1991, A fast algorithm for the implementation of filter banks based on time domain aliasing cancellation, Proceedings of IEEE ICASSP'91, Toronto Canada, pp. 2209-2212.

ENGELHART K., HUDGINS B., PARKER P.A., STEVENSON M., 1999, Classification of the Myoelectric Signal using Time-Frequency Based Representations, Medical Engineering and Physics on Intelligent Data Analysis in Electromyography and Electroneurography.

ERSLAND L., ROSEN G., LUNDERVOLD A., SMIEVOLL A., TILLUNG T., SUNDBERG H., HUDGAHL K., 1996, Phantom limb imagery finger tapping causes primary cortex activation: an FMRI study, Neuroreport, 8(1), pp. 207-210.

FARWELL L.A., and DONCHIN E., 1988, Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials, Electroencephalography and Clinical Neurophysiology, vol. 70, pp. 510-523.

GEORGOPOULOS A., SCHWARTZ A., and KETTNER R., 1986, Neural population coding of movement direction. Science, 233, pp. 1416–1419.

GINTER J JR, BLINOWSKA KJ, KAMINSKI M, DURKA PJ., 2001, Phase and Amplitude analysis in time-frequency space-application to voluntary finger movement. Journal of Neuroscience Methods 110, pp. 113-24.

GUGER C., 1999, Real-Time Data Processing Under Windows for an EEG-Based Brain Computer Interface, Dissertation.

HAYES M. H., 1996, Statistical Digital Signal Processing and Modeling, John Wiley & sons.

HERLEY C., KOVACEVIC J., RAMCHANDRAN K., and VETTERLI M., 1993, Tilings of the Time –Frequency plane: Construction arbitrary orthogonal bases and fast tiling algorithms, IEEE Trans. Signal Proc., special issue on Wavelets and Signal Processing, vol. 41, pp.3341-3359.

HJORTH B., 1975, An online transformation of EEG scalp potentials into orthogonal source derivations. Electroencephalography and Clinical Neurophysiology, 39, pp. 526-530.

INCE N. F., ARICA S., 2004, Analysis and Visualization of ERD and ERS with Adapted Local Cosine Transform, Proceedings of 26th Annual International Conference of IEEE Engineering in Medicine and Biology Society, EMBC 2004 IEEE, San Francisco.

INCE N. F., TEWFIK A., ARICA S., 2005, “Classification of movement EEG with Local Discriminant Bases”, 30th International Conference on Acoustics, Speech, and Signal Processing (ICASSP) Society, IEEE, Philadelphia

INCE N. F., TEWFIK A., ARICA S., YAGCIOGLU S., 2005, "Analysis and Visualization of Movement related EEG activities using Local Discriminant Bases" 26^{th} The 2nd International IEEE EMBS Conference on Neural Engineering, Washington D.C., USA

JEANNEORD, M., 1994, The representing brain: neural correlates of motor intention and imagery. Behav. Brain Sci., 17(2), pp. 187-245.

J. KALCHER, D. FLOTZINGER, G. PFURTSCHELLER, 1993, Brain-computer interface - A new communication device for handicapped persons. Journal of Microcomputer Applications, 16, pp. 293-299.

KOTCHOUBEY B., SCHLEICHERT H., LUTZENBERGER W., and BIRBAUMER N., 1997, A new method for self-regulation of slow cortical potentials a timed paradigm, Applied Psychophysiology and Biofeedback, vol. 22, pp. 77-93.

KÜBLER A., KOTCHOUBEY B., HINTERBERGER T., GHANAYIM N., PERELMOUTER J., SCHAUER M., FRITSCH C., TAUB E., and BIRBAUMER N., 1999, Thought Translation Device: a neurophysiological approach to communication in total motor paralysis, Exp. Brain Res., vol. 124, pp. 223-232.

KUEBLER A., KOTCHOUBEY B., SALZMANN H. P., GHANAYIM N., PERELMOUTER J., HÖRNBERG V., and BIRBAUMER N., 1998, Self-regulation of slow cortical potentials in completely paralyzed human patients, Neuroscience Letters, vol. 252, pp.171-174.

LANR, CHEYENE D., HOLLINGER P., GERSCHLAGER W., LINDINGER G., 1996, Electric and magnetic fields of the brain accompanying internal simulation of movement, Brain Res. Cogn., 3(2), pp. 125-129.

LEUTHARDT E. C., SCHALK G., WOLPAW J. R., OJEMANN J. G. and MORAN D. W., 2004, A brain–computer interface using electrocorticographic signals in humans, Journal of Neural Engineering, pp. 63–71.

LEVINE S. P., HUGGINS J. E., BEMENT S. L., KUSHWAHA R. K., SCHUH L. A., PASSARO E. A., ROHDE M. M., 1999, and ROSS D. A., Identification of electrocorticogram Patterns as the Basis for a Direct Brain Interface, Journal of Clinical Neurophysiology, pp. 439-447.

DA SILVA L., PFURTSCHELLER G., 1999, Basic Concepts on EEG synchronization and desynchronization. In: Pfurtscheller, G., Lopes da Silva, F.H., Event Related Desyncronization. Handbook of Electroencephalography and Clinical Neurophysiology, 6, Elsevier, Amsterdam.

MALLAT S., 2000, A Wavelet Tour of Signal Processing, Second edition, Academic Press

MALLAT S., 1989, Multifrequency channel decompositions of images and wavelet models. IEEE Trans. On Acous. Speech and Signal Proc., 37, pp. 3377-3396.

MALLAT S., 1989, A theory for multiresolution signal decompositions: wavelet representation. IEEE Trans. on Pattern Recog. Mach. Intell.,11, pp. 674-693

MARPLE S.L., 1987, Digital Spectral Analysis with Applications, Prentice Hall, Upper Saddle River, New Jersey

MCFARLAND D.J., NEAT G.W., READ R.F., and WOLPAW J.R., 1993, An EEG-based method for graded cursor control, Psychobiology, vol. 21 1, pp. 77-81.

MCFARLAND D.J., LEFKOWICZ A.T., and WOLPAW J.R., 1997, Design and operation of an EEG-based brain-computer interface with digital signal processing technology, Behavior Research Methods, Instruments & Computers, vol. 29 3, pp. 337-345.

MCMILLAN G.R. and CALHOUN G.L., et al., 1995, Direct brain interface utilizing self-regulation of steady-state visual evoke response, in Proceedings of RESNA, June 9-14, pp.693-695.

NEUPER, C., PFURTSCHELLER G., 2001, Event-related dynamics of cortical rhythms: frequency-specific features and functional correlates., Inter. Journal of Psychophysics, 43, pp. 41-58

NEUPER C. and PFURTSCHELLER G., 1999, Motor imagery and ERD. In: Pfurtscheller, G., Lopes da Silva, F.H. Eds., Event-Related Desynchronization. Handbook of Electroenceph. and Clin. Neurophysiol. Revised Edition Vol. 6. Elsevier, Amsterdam.

NICOLELIS M. A. L., 2001, Actions from thoughts, Nature, vol. 409, pp.403-407.

OBERMAIER B., MUNTEANU C., PFURTSCHELLER G., 2001, Asymmetric Hemisphere Modelling in an offline Brain Computer Interface, IEEE Transactions on Systems Man and Cybernetics PartC Applications and Reviews,4:31;536-540

OSMAN A., and ROBERT A., 2001, Time-course of cortical activation during overt and imagined movements, in Proc. Cognitive Neuroscience Annu. Meet., New York.

PENNY W.D. and ROBERTS S.J., 1998, Bayesian neural networks for detection of imagined finger movements from single-trial EEG, http://www.ee.ic.ac.uk/hp/staff/sjrob/pubs.html.

PENNY W. and ROBERTS S.J., 1999, Imagined hand movements identified from the EEG mu rhythm, Journal of Neuroscience Methods.

PFURTSCHELLER G., NEUPER C., 2001, Motor Imagery and Direct Brain-Computer Interface, Proceeding of IEEE, vol.89, pp. 1123-1134.

PFURTSCHELLER G., FLOTZINGER D., MOHL W., and PELTORANTA M., 1992, Prediction of the side of hand movements from single-trial multi-channel EEG data using neural networks, Electroenceph. Clin. Neurophysiol., vol. 82, pp. 313-315.

PFURTSCHELLER G., NEUPER C., FLOTZINGER D., and PREGENZER M., 1997, EEG-based discrimination between imagination of right and left hand movement, Electroenceph. Clin. Neurophysiol., vol. 103, pp. 642-651.

PFURTSCHELLER G., FLOTZINGER D., and KALCHER J., 1993, Brain-computer interface – a new communication device for handicapped persons, J. Microcomputer Appl., vol. 16, pp. 293-299.

PFURTSCHELLER G., NEUPER C., and BERGER J., 1994, Source localization using event related desynchronization ERD within the alpha band, Brain Topography., vol. 6/4, pp. 269-275.

PFURTSCHELLER G., KALCHER J., NEUPER C., FLOTZINGER D., and PREGENZER M., 1996, On-line EEG classification during externally-paced hand movements using a neural network-based classifier, Electroenceph. Clin. Neurophysiol., vol. 99, pp. 416-425.

PFURTSCHELLER G., STANACK JR. A. and NEUPER C., 1996, Post movement beta synchronization. A correlate of an idling motor area? Electroenceph. Clin. Neurophysiol., 98, pp. 281-293

PFURTSCHELLER G., NEUPER C., FLOTZINGER D., PREZENGER M., 1997, EEG based discrimination between imagination of right and left hand movement, Electroenceph. Clin. Neurophysiol., 103, pp. 642-651.

PFURTSCHELLER G., and NEUPER C., 1997, Motor imagery activates primary sensorimotor area in humans, Neurosci. Lett., vol. 239, pp. 65-68.

PFURTSCHELLER G., STANACK JR. A. and EDLINGER G., 1997, On the existence of different types of central beta rhythms below 30Hz. Electroenceph. Clin. Neurophysiol., 102, pp. 316-325

PFURTSCHELLER G., 1998, "Separability of Left and right hand movement imageries with adaptive autoregressive modeling, IEEE.

PFURTSCHELLER G., 1998, EEG event-related desynchronization ERD and event related synchronization ERS. In: E. Niedermeyer, F. H. Lopes da Silva, Eds., Electroencephalography: Basic Principals, Clinical Applications and Related Fields. 4th Edition, Williams and Wilkins, Baltimore, pp. 958-967.

PFURTSCHELLER G. and DA SILVA L., et.al., 1999, Event-Related Desynchronization, Handbook of Electroenceph. and Clin. Neurophysiol. Revised Edition Vol. 6. Elsevier, Amsterdam.

PFURTSCHELLER G., NEUPER C., GUGER C., OBERMAIER B., PREGENZER M., RAMOSER H., and SCHLÖGL A., 1999, Current trends in Graz brain-computer interface BCI research, IEEE Trans. Rehab. Eng.

PFURTSCHELLER G., 1999, Quantification of ERD and ERS in Time Domain, Handbook of Electroenceph. and Clin. Neurophysiol. Revised Edition Vol. 6. Elsevier, Amsterdam, pp. 89-105.

POLIKOFF J.B., BUNNELL H.T., and BORKOWSKI W.J. Jr., 1995, Toward a P300-based computer interface, in Proc. RESNA '95 Annual Conference, RESNAPRESS, Arlington Va. pp. 178-180.

PREGENZER M., PFURTSCHELLER G., and FLOTZINGER D., 1994, Selection of electrode positions for an EEG-based brain computer interface BCI, Biomed. Technik, vol. 39, pp. 264-269.

PREZENGER M. and PFURTSCHELLER G., 1999, Frequency component selection for an EEG-based brain computer interface, IEEE Trans. on Rehabil. Eng. 7, pp. 413-419.

QUINLAN J. R., 1993, C4.5: Programs for Machine Learning. San Francisco, CA: Morgan Kaufmann.

RAMOSER H., GERKIN J.M, PFURTSCHELLER G., 2000, Optimal spatial filtering of single trial EEG during Imagined Hand Movement, IEEE Trans. on Rehabil. Eng. vol.8, pp. 441-446.

RAO K. R.AND YIP P., 1990, Discrete Cosine Transform – Algorithms, Advantages and Applications, Academic Press.

SAITO N., et al., 2002, Discriminant feature extraction using empirical probability density and a local basis library, Pattern Recognition, vol.35, pp. 1842-1852

SAITO N., 1994, Local feature extraction using a library of bases, Dissertation, Yale University

SAITO N., Classification of Geophysical acoustic waveforms using time-frequency atoms. Am. Stat. Assoc. Statistical Computing Proceedings

SCHERBERGER H., MURRAY R.,J., ANDERSEN R. A., 2005, Cortical Local Field Potential Encodes Movement Intentions in the Posterior Parietal Cortex, Neuron, Vol. 46, pp. 347–354.

SCHLÖGL A., FLOTZINGER D., PFURTSCHELLER G., 1997, Adaptive autoregressive modeling used for single trial EEG classification, Biomed. Technik,42, pp. 162-167

SCHWARTZ A. B., 2004, Cortical Neuroprosthetics, Annu. Rev. Neuroscience, 27, pp. 487-507.

SUTTER E. and TRAN D., 1990, Communication through visually induced electrical brain responses, Computers for Handicapped Persons; Springer Verlag, pp. 279-288.

SUTTER E.E., 1992, The brain response interface: communication through visually-induced electrical brain responses, J. Microcomputer Applications, vol. 15, pp. 31-45.

THEODORIDIS S., KOUTROUMBAS K., 2003, Pattern Recogition Second, Academic Press

VETTERLI M., KOVACEVIC J., 1995, Wvalets and Subband Coding, Prentice Hall,Upper Saddle River, New Jersey

WANG T, and HE B., 2004, An efficient rhythmic component expression and weighting synthesis strategy for classifying motor imagery EEG in brain-computer interface. J Neural Eng,1, pp. 1–7.

WU R. S., and WANG Y., 1999, New Flexible Segmentation Technique in sesimic data compression using local cosine transform. Wavelet Applications in signal and image processing VII, Proc. SPIE, 3813, pp. 784-794.

WESFREID E., WICHERHAUSER M., 1993, Adapted Trigonometric Transform and Speech Processing, IEEE Trans.,Acoustic Speech Processing.

WESTFREID E., WICKERHAUSER V., 1999, Vocal command signal segmentation and phonomes classification, Wireless Networks, Available at http://www.math.wustl.edu/~victor/papers

WESSBERG J., STAMBAUGH C., KRALIK J., BECK P., LAUBACH M., CHAPIN J., KIM J., BIGGS S., SRINIVASAN M., and NICOLELIS M., 2002, Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. Nature, 408, pp. 361 365.

WICKERHAUSER M.V., COIFMAN R.R., 1992, Entropy based algorithms for best basis selection, IEEE Transactions on Information Theory.

WICKERHAUSER M.V., 1994, Adapted Wavelet Analysis from theory to software, A.K. Peters, Massachusetts.

WOLPAW J. R., BIRBAUMER N., WILLIAM J. HEETDERKS, DENNIS J. MCFARLAND, P. HUNTER PECKAM, GERWIN SCHALK, EMANUEL DONCHIN, LUIS A. QUATRANO, CHARLES J. ROBINSON and THRESA M. VAUGHAN, 2000, Brain-Computer Interface Technology: A review of the first international meeting, IEEE Trans. On Rehabil. Eng., 8, pp. 164-173.

WOLPAW J. R., MCFARLAND D.J., NEAT G.W., and FORNERIS C.A., 1991, An EEG-based brain computer interface for cursor control, Electroenceph. Clin. Neurophy., vol. 78, pp. 252-259.

WOLPAW J. R. and MCFARLAND D.J., 1994, Multichannel EEG-based brain-computer communication, Electroenceph. Clin. Neurophy., vol. 90, pp. 444-449.

CURRICULUM VITAE

Nuri Fırat İNCE was born in Adana, Turkey, February 9^{th}, 1976. He got his BS and MS from the Department of Electrical and Electronics Engineering, University of Cukurova. His Master thesis title is "*A Computer Based Data Acquisition and Signal Processing System for Measuring the Baroreceptor Sensitivity* ". It was conducted in cooperation with the Cardiology Department of Cukurova University Medical School - Supported by Prof. Ahmed BIRAND. After completing his Master's in 2002 he started with his PhD at the University of Cukurova, Department of Electrical and Electronics Engineering. During his PhD he was supported by an international joint PhD scholarship from the National Scientific Research Council of Turkey (TUBITAK); Including: University of Technology Graz Austria; Hacettepe University Ankara Turkey; Isik University Istanbul Turkey; University of Minnesota USA. His research areas are as follows:

- Biomedical Data Acquisition and Signal Processing
- Industrial Automation and Instrumentation
- System Level Programming and Interfacing
- Client Server Based Software Development

His publications are as follows:

- N.Firat INCE, Ahmed TEWFIK, Sami ARICA, Suha YAGCIOGLU, "Analysis and Visualization of Movement related EEG activities using Local Discriminant Bases" 26^{th} The 2nd International IEEE EMBS Conference on Neural Engineering, 2005, Washington D.C., USA
- N.Firat INCE, Ahmed TEWFIK, Sami ARICA, "Classification of Movement EEG with Local Discriminant Bases" 26^{th} ICASSP 2005, Philadelphia, USA
- N.Firat INCE, Sami ARICA, "Analysis and Visualization of ERD and ERS with Adapted Local Cosine Transform", EMBC 2004, San Francisco, USA

- N.Firat INCE, Sami ARICA, "Extraction Of ERD And ERS Patterns From EEG Using Adapted Local Trigonometric Transform", BIOSIGNAL 2004, Brno, Czech Republic
- N.Firat INCE, Sami ARICA, Ahmet BIRAND "Heart Rate Extraction & QRS Detection Using a Multi-Layer Perceptron in Comparison with D1-TH Algorithm", TAINN’03, Çanakkale, Turkey
- N.Firat INCE, Sami ARICA, Ahmet BIRAND "A PC Based Data Acquisition and Signal Processing System for Measuring the Baroreceptor Sensitivity", Proceedings of the IASTED International Conference on Signal Processing, Pattern Recognition, and Applications, Pages 93-97, 2003, Rhoedos, Greece
- N.Firat INCE, Sami ARICA, Ahmet BIRAND "Elektrokardiyogram ve Arterial Kan Basıncı Analizi ve Baroreseptör Duyarlılığının Ölçümü" SİU'2002 10. Sinyal İşleme ve İletişim Uygulamaları Kurultayı, cilt 1, sayfa 453-458, *in Turkish*
- N.Firat INCE, Sami ARICA, Ahmet BIRAND "PC Tabanlı Veri Toplama Sistemi: Kalp Ritmi ve Kan Basıncı İzleme ve Görüntüleme" TMMOB Elektrik Mühendisleri Odası, Elektrik-Elektronik-Bilgisayar 9. Ulusal Kongresi, cilt 2, sayfa 456-458, 2001, *in Turkish*