# 國立交通大學

## 生醫工程研究所

## 碩 士 論 文

利 用 視 覺 誘 發 腦 電 波 之 身 份 辨 識

Person Identification using Electroencephalographic
Signals Evoked by Visual Stimuli

研 究 生：林家萍

指導教授：陳永昇　博士

中 華 民 國　一百 年 八 月

# 利用視覺誘發腦電波之身份辨識
# Person Identification using Electroencephalographic Signals Evoked by Visual Stimuli

研 究 生：林家萍　　Student：Jia-Ping Lin

指導教授：陳永昇　　Advisor：Yong-Sheng Chen

國 立 交 通 大 學
生 醫 工 程 研 究 所
碩 士 論 文

A Thesis
Submitted to Institute of Biomedical Engineering
College of Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in

Computer Science

August 2011

Hsinchu, Taiwan, Republic of China

中華民國一百年八月

# Person Identification using Electroencephalographic Signals Evoked by Visual Stimuli

A thesis presented

by

Jia-Ping Lin

to

Institute of Biomedical Engineering

College of Computer Science

in partial fulfillment of the requirements

for the degree of

Master

in the subject of

Computer Science

National Chiao Tung University

Hsinchu, Taiwan

2011

**Person Identification using Electroencephalographic Signals Evoked by Visual Stimuli**



by

Jia-Ping Lin

# 摘　要

近年來利用生物特徵的方式來進行身份辨識越來越普遍，其原因是由於生物特徵具有難以遭到破解或竊取的優點。然而，隨著科技的進步目前的生物特徵(例如：指紋、虹膜等)已有被複製的風險。由於腦電波具有個體間的差異，因此在本研究中我們利用視覺刺激誘發的腦電波為分析訊號來發展身份辨識系統，實驗於安靜無干擾的房間進行 ，讓受測者接受事件相關的視覺刺激(oddball paradigm)，利用刺激材料出現頻率的不同誘發出腦波的事件相關電位。辨識的步驟主要分為類別與確認兩大部分，並利用支援向量機作為分類器。類別的部分，原始訊號經過特徵擷取後藉由一個多種類的分類器會得到一個一對多的分類結果；而接著在確認部分，由類別步驟所得到的最佳分類結果經由此部分二元的分類器進行確認。特徵擷取方面，包含了降維，時域以及頻域的分析方法，能將具有代表性的資訊保留。此外，我們嘗試利用重複確認步驟的二元分類器將前一步驟(類別)分類錯誤的資料進行修正，修正的準則是依照支援向量機中的信賴評估為指標。

我們利用 18 位受測者的辨識結果得到 97.25%的準確率，並且再經由確認的步驟能達到 98.89%的正確接受率，這樣的結果顯示腦電波訊號具有的個體差異性足夠用於進行身份辨識且利用類別和確認兩部分的結合能達到一個好的準確率，且辨識的可信度提升。而更深入的討論訊號間的差異，我們發現不同受測者的訊號相關性低於同一受測者不同天的受測資料，這個發現符合腦電波具有低個體內差異性以及高個體間差異性，且隨著時間的變化同一人的訊號是恆定的。相關性的高低也解釋了某些受測者容易被錯誤分類的情況，也就是他們和其他人的訊號具有高度的相關性。總結我們系統所得出結果顯示，結合未來硬體發展更趨成熟腦波能成為一個新的生物特徵以發展成一套更安全的辨識系統。

**Abstract**

The biometrics contains emerging methods for human identification. As advances in technology, conventional techniques using fingerprint or iris have the risk of being duplicated. In this work we utilize the inter-subject differences in the electroencephalographic (EEG) signals evoked by visual stimuli for person identification. The identification procedure is divided into classification and verification phases. For our classification system, it is based on the supervised classification method with support vector machine. During the classification phase, we extract the representative information from the EEG signals of each subject and construct a multi-class classifier. The best-matching candidate is further confirmed in the verification phase by using a binary classifier. The methods of feature extraction include dimension reduction and time-frequency analysis. Moreover, we try to correct those misclassified data through the iterative verification that depends on the confidence values of SVM classifier, which is a confidence level of classification. According to our experiments in which 18 subjects were recruited, the proposed method can achieve 97.25% identification rate. The results revealed that EEG data with individual differences can reach a high accuracy in person identification. Combining classification with verification, the reliability of the system can be increased. The correlation values of EEG signals between different subjects is lower than those of EEG signals acquired at different days for the same subject. This finding suggests that the characteristics of EEG has low intra-subject variability but high inter-subject variability and it is stable over time. The correlation values may also explain why some subjects apt to be misclassified when they have high correlation values to others. Our experimental results demonstrated that the proposed methods have great potentials for identifying individuals in daily life applications.

# 致　謝

兩年的研究生活首要感謝我的指導教授－陳永昇老師與陳麗芬老師的傾囊相授，兩位老師無論在研究熱誠以及待人接物上都是我學習的楷模，十分感謝老師們這些日子以來的協助與指導，讓我從中收穫許多；此外，感謝王振興教授與王才沛教授對於論文的審視，以及於口試時所提供的指導與建議。接著感謝所有在我實驗中參予的受測者：電鍋、sheep、大頭、陳麒宇、小乖的朋友們、實驗室的同學與學弟妹們，有了你們的幫忙我的研究才得以順利進行。

很幸運能和所有 618 的成員一起經歷我最後的學生生涯，蓮霧、國維、乙宛、小可，能和你們當同學真的非常棒！希望大家接下來也都能有很好的發展。謝謝實驗室學長姐小白和慧伶從我剛進實驗室就不吝提供各方面的指導，祝福你們研究順利，最後要感謝我的朋友以及總是支持我的家人，有了你們的關懷與打氣是我堅持下去的動力。

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

In the first chapter we introduce some relevant background knowledge. Firstly, we briefly describe the motivation of the proposed work at the first section. In Section 1.2 we give some introduction to Electroencephalography (EEG), so called brainwave. The measurement methods, some basic analyses and researches are presented. In Section 1.3 we describe about the application of ERP waveform and give a description of the person identification based on ERP signals. Finally, Section 1.4 gives the overview of the thesis.

## 1.1 Motivation

As advances in technology we need more efficient methods for person identification in order to prevent occurrence of information leaks or cheats. The general person identification methods such as passwords and smart cards which are convenient to use; however, they do not have a high degree of security protection because the risk of forgetting, loss and might be stolen. As more and more criminal activities of identity fraud, identity fraud has become a critical and global issue. Several approaches have been applied in order to overcome this problem.

To improve these disadvantages the biometrics is nowadays widely studied because of the highly reliable results. The biometrics is to authenticate person by physical characteristics: fingerprint, iris, voice, gait, and palm. However, these typical biometrics may be dissolved by physical damage (dry skin, scar, sound damage or loss); in fact, about 2-3% of the population lose the features, not to mention these features could be duplicated or imitated by imposters. The new type of biometrics EEG that is brought up for person identification [21, 22]. EEG has low intra-subject variability and high inter-subject variability. Moreover it is stable over time. For now brainwave can not be stolen or duplicated that is more favorable for person identification.

We attempt to develop a person identification system based EEG signal that evoked by visual stimuli, and used efficient feature extraction methods to transform high signal to noise ratio (SNR) data to discriminative information. In addition to high accuracy of classification we try to improve the results; therefore, the main framework is divided into classification and verification that the results will be more reliable. By the security and accuracy of EEG biometrics the new method can be highly applied with future advances in hardware.

## 1.2 Electroencephalography

### 1.2.1 Introduction to Electroencephalography

To obtain brainwave periods there are already many measurements. Among all the measurements which are non-invasive such as magnetoencephalography (MEG), electroencephalography (EEG), and functional magnetic resonance imaging (fMRI) specially suit to use for extensive research. The reason why many research use EEG signal as their analysis data is, it is portable, easy to operate, and it costs low prices compared with the other measurements. EEG is used to measure the electrical activity of the brain. This activity is generated by billions of nerve cells which called neurons. Each neuron is connected with thousands of other neurons, and the neurons send action potentials to other neurons when they are communicating. For EEG measurement, we actually measure the combined electrical activity of millions of neurons on the cerebral cortex because the potential of a single neuron is too small to be measured.

A typical EEG measuring device consists of several components, including EEG electrode cap that receives the electrical activity from the scalp, EEG amplifier that processes the signal amplification, computers that record the data, and monitors that present the visual stimulus. The devices are shown in Fig. 1.1.

The EEG signal has high temporal resolution but relatively poor spatial resolution, which depends on the electrode number of an EEG electrode cap. The electrode layout on an EEG electrode cap has a international standard called the international 10-20 system, as Figure. 1.2 shows. While measuring EEG we often put some single electrodes surrounding the eye. This is used to measure the electrical activity of eye movement and eye blinking, which is called EOG. This EOG contaminates the EEG signal badly, so by measuring it we can remove the trials that was affected. This processing is called EOG rejection.

During the measurement of EEG we have to fill each electrode with the electrolyte gel


Figure 1.1: EEG measuring devices. From left to right is the EEG amplifier and the electrode cap.

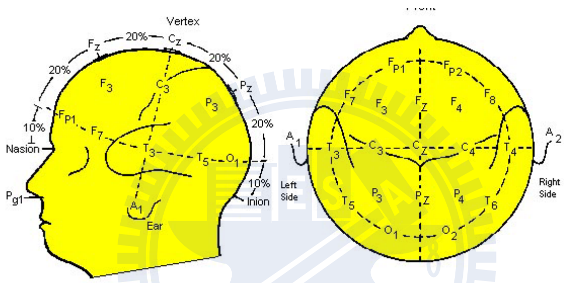


Figure 1.2: The international 10-20 system. The 10 and 20 refer to the 10% and 20% interelectrode distance of the skull. The F, C, P, O, T stand for frontal, central, parietal, occipital, temporal lobe. The odd number is placed in the left side and the even number refer to electrode positions on the right side [9].

using a blunt needle; and further, we must ensure that the electrolyte gel is exposed to the scalp completely. This makes the electrodes contact the scalp and lower the impedance. In an EEG experiment we often wait until all the electrodes have an impedance lower 5 k $\Omega$ before we start the signal acquisition.

### 1.2.2 Basic Analysis of Electroencephalography

For increasing SNR of EEG, there are some basic EEG analyses. In the subsection we mainly described time domain and frequency domain analysis.

#### Time domain analysis

The common time domain analysis is to observe an Event-related Potential (ERP). An ERP is a potential change in the EEG when a particular event or stimulus occurs. The potential change is time-locked and phase-locked, it is a very small potential change and can not be easily observed in a single trial. So we have to average a few trials to observe it. Because of the time-locked and phase-locked characteristic, by the averaging process we can eliminate the random noise and enhance the SNR. Another technique that is often used to separate these signals from background activity and noise is low-pass or bandpass filter. It is reasonable because most of the energy of ERP is concentrated at low frequencies. Some well-known ERP include P100 in the Visual-evoked Potential (VEP), N170 which reflects the structural encoding process, P300, N400 which reflects access to person identity node and semantic processing, and Audio-evoked Potential (AEP) [15].

#### Frequency domain analysis

As the name suggests frequency domain analysis it is used to observe the changes in oscillatory activity. Such changed can be evoked by presentation of stimulus or by concentration of the subject on a particular mental task. Usually, the phase of oscillatory activity is not time-locked to the stimulus or to mental task of subject. Therefore, time domain analysis technique cannot be used. Instead, we need frequency domain analysis to observe the oscillatory activity. For instance, SSVEP have band power in the harmonics of the visual stimulation frequency at occipital cortex. Fast Fourier transform can be used to estimate the band power as features. Another example in systems based on motor imagery, the band power in the mu and beta rhythm over the sensorimotor cortex is used as features. Moreover, the band power in alpha rhythm is widely used (fatigue detection, concentration) for it evoked while eye closing.

### 1.2.3 Event-related Potentials

Several kinds of internally or externally paced events will result in time-locked and phase-locked brain signals. Almost all the evoked activities have a more or less fixed time-delay to the stimulus. These time-locked and phase-locked called event-related potentials (ERPs) or evoked potentials (EPs). ERP can be viewed as potential changes of the neurons when our brain deal with mental tasks. Usually the brain activities of mental task is smaller than the spontaneous brain signals, thus concealed in the irregular and noisy spontaneous brain signals. In order to extract the ERPs, synchronous averaging are performed, implying we have to repeat the same mental tasks more than once, after applying synchronous averaging, most of the noise will be eliminated, therefore enhancing the SNR and obtaining the time-locked and phase-locked signals, ERPs.

For the EEG of our experiment evoked by visual stimuli; therefore, we particularly introduce visual evoked potential (VEP) from widespread ERP.

### VEP

Visual evoked potential (VEP) is induced when the users eyes are stimulated by looking at a test pattern which often is a flashing pattern. The well-known VEP is P100 (Fig. 1.3) which are part of early components called exogenous, because they require a stimulus. Such early components can be modulated by sustained attention and top-down cognitive control processes [13]. To measure VEPs, the recording electrodes are placed over the visual cortex.

The other VEP which related to oscillatory activity is SSVEP that elicited by a brief visual stimulus modulated at a specific frequency. The visual stimulus flick at different frequencies lead to brain oscillation at the same frequency and at harmonics and subharmonics of the stimulation frequency.

### P300

A longer latency component, the high amplitude P300 appearing approximately 300 ms after the presentation, is called endogenous, because it can be present even in response to an expected stimulus (however, it is missing actually). P300 is elicited in oddball tasks in response to task-relevant, salient infrequent targets, with higher amplitude over posterior (parietal) scalp. Many different stimulus modalities can be used to evoked the P300, such as visual, auditory, sense of touch, gustatory or olfactory. In other words, P300 reflects

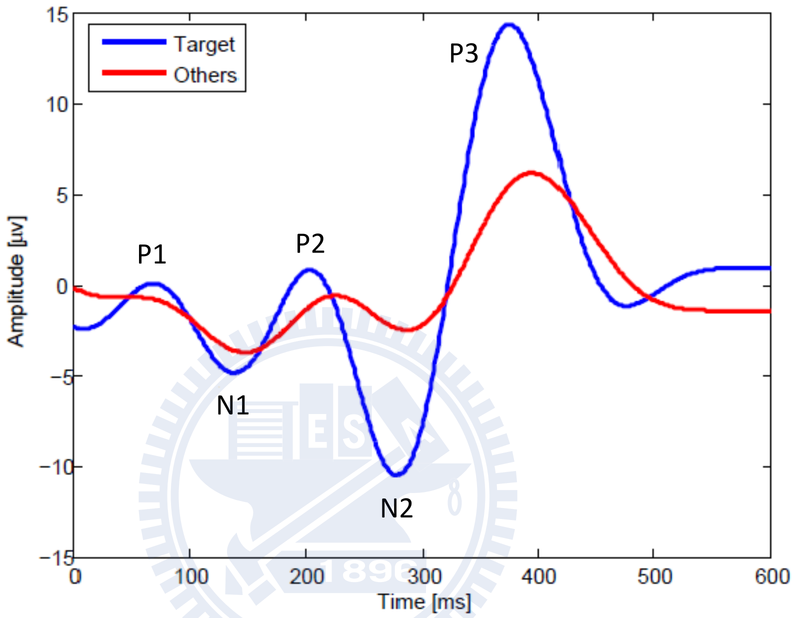


Figure 1.3: A typical ERP waveform. The P100 is a positive deflection in the EEG, which appears approximately 100 ms after presentation of a visual stimulus. The P300 (P3) is a positive deflection in the EEG, which appears approximately 300 ms after presentation of a rare or meaningful stimulus. A series of negative and positive components (N1, P2, N2) proceed the P3. The P3 reflect high-level processing of stimuli such as cognitive processing, the earlier components reflect low-level, automatic processing of stimulus.

higher-order cognitive processes. The P300 is a popular topic because it can be reliably measured and the contribution of characteristics waveform. For instance the latency and amplitude can be influenced by various factors. Some important factors influencing the P300 are listed below.

- Interstimulus interval

  The interstimulus interval is the temporal interval between the offset of one stimulus to the onset of another. In recent study, the ISI are positively related to the P300 amplitude. The longer the ISI be used, the larger the P300 amplitude are presented.

- Target frequency

  The amplitude of P300 is inversely related to the frequency of the stimulus. The frequency of the target stimulus is low means that it is more significant for subject. To elicited a stable P300 response it is helpful using low frequency of the stimulus; however, this requires a longer measurement time.

- Concentration

  The concentration of subject play an important role for signal quality. The amplitude of the P300 depend on how the subjects focus on the stimulus. In an oddball paradigm, the P300 cannot be elicited while the subject is absent from the target stimulus.

The typical paradigm used in P300 experiments is oddball, in which irregular relevant stimuli (target) that require a specific cognitive response. For inducing the P300 it have to be detected in a sequence of frequent irrelevant non-target stimuli.

## 1.3 Application of ERP

Because of the obvious features of ERP there are a lot of application covering various aspects. Some research focus on the self-relevant task to induce ERP by the stimuli which is meaningful to the subjects, and some research interest in more in-depth cognitives issues such as mind [2, 6, 8] The P300 is widely used in criminal detection due to it is relative to high-order cognitive processes that when people receive event-related information. In addition to applications of ERP other studies attempt to find effective methods of feature extraction that make data more practical.

### 1.3.1 Person identification based ERP

Since the EEG signals with discriminative individual differences the researchers begin to study that how to use EEG as a biometric. The ideal conception is to promote brainwaves as new keys for safer person identification systems. The resting data of EEG are used and then some research investigated the task-related EEG signals [26] to induce more specific waveform such as ERP. The details of person identification based EEG will introduce in next chapter.

## 1.4 Thesis overview

Chapter 2 provides the overview of person identification systems, including the basic components and key-issues. We also illustrate the disadvantages and limitations of the present-day systems. The paradigm of data acquisition and experimental procedures will be introduced in Chapter 3. Chapter 4 provides the main structure of the system and the methods of feature extraction. The classification to evaluate our system also describe in this chapter. In Chapter 5, the results of our classification. Finally we summarize this work with the results and explain some possible reasons of misclassification of our methods in Chapter 6 and give some conclusions in Chapter 7.

# Chapter 2

# Related works

## 2.1 Introduction to person identification systems

There are a lot of systems require reliable personal recognition schemes to either confirm or determine the identity of an individual requesting their services. The purpose of such schemes is to ensure that the supplied services only can be accessed by a allowed user and no one else. Examples of such applications include secure access to buildings, computer systems, laptops, cellular phones and ATMs [12].

Conventional person identification methods include passwords, smart cards, and a variety of biometric techniques. Passwords and smart cards are widely-used because of the advantage of convenience. However, smart cards might be stolen, simple passwords might be deciphered, and complicated passwords might be forgotten. Biometric recognition, refers to the automatic recognition of individuals based on their physiological or behavioral characteristics is popular recently and considered more secure way.

**Biometric recognition**

By using biometrics it is to confirm or establish an individual's identity based on "who he/she is", rather than by "what he/she possesses" (e.g., a smart card) or "what she remembers" (e.g., a password). What qualities need to have a biometric can be used in identity? Any human physiological or behavioral characteristic can be used as a biometric characteristic on condition that it satisfies the following requirements:

- Universality: anyone have this characteristic;
- Distinctiveness: the characteristics of any two people should have sufficient differences to separate different people;
- Permanence: the characteristic should be sufficiently invariant with time (correspond to the matching criterion);
- Collectability: the characteristic can be measured quantitatively.

In addition to the above requirements, a practical biometric system should be provided with sufficient accuracy and speed, be user-friendly, and it can expand the number of users. It is also necessary to prevent the impostors by robust anti-theft mechanism.

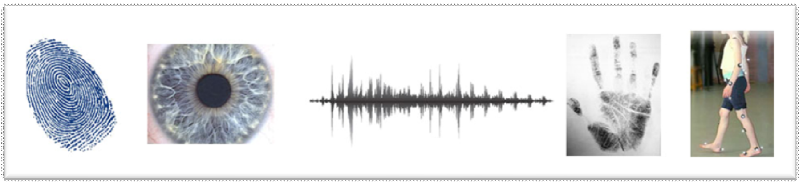

Figure 2.1: Examples of biometric characteristics. From left to right are fingerprint, iris, voice, palm and gait.

### Biometric system

A biometric system is essentially a pattern recognition system that works by biometric data acquired from an individual. After feature extracting from the acquired data the features are compared to the template set in the database.

The biometric system can be divided into two mode depending on the application. One is verification mode that the system validates a person's identity by comparing the acquired biometric data with his/her own biometric template stored in database. People who want to be approved by the system need to claim an identity via a PIN (Personal Identification Number), a user name or a smart card. The system conducts a one-to- one comparison to determine whether the claim is true or not (e.g., *user*: I am John. *system*: Whether this biometric data belong to John.). The other is recognition mode that the system recognizes an individual by searching the templates of all the users in the database for a match. A one-to-many comparison to establish an individuals identity consequently conducted in this system with out a declaration of subjects; however, it will fail if the subject is not enrolled in the database (e.g., *system*: Whose biometric data is this?).

## 2.2 Categories of biometrics

### Fingerprint

In $19^{th}$ century, Alphonse Bertillon who is a chief of the criminal identification division conceived and then practiced the idea of using a number of body measurements to identify criminals [23]. After this, a more significant and practical discovery of the distinctiveness of the human fingerprints became clear and soon the fingerprints used for criminal identification. While law enforcement agencies were the earliest adopters of the fingerprint identification technology, it is being increasingly used later because more identity fraud has created.

#### Iris

According to human eye micro vascular, infinite variety of combinations, the complex iris texture carries very distinctive information useful for personal recognition. The combination of iris blood vessels will not change basically even with the age increasing (except for severe diabetes, glaucoma). It is difficult to surgically tamper because of the living conditions (such as vascular blood flow). In view of those characteristics of iris it become popular in person recognition.

#### Voice

Unlike fingerprint and iris, the biometric of voice is using signal characteristics. Voice is a combination of physiological and behavioral biometrics. The features of an individual voice are based on the shape, amplitude and frequency. The physiological characteristics of human speech are invariant for an individual. Speaker recognition is most appropriate in phone-based applications such as voice dialling. However, there are some disadvantages lead to difficulties in the application of person identification (we will detail in next subsection).

### 2.2.1 Disadvantages of present-day biometric systems

There are other biometrics such as face, palm and gait except above described. Although biometrics are more reliable and secure method of identity, it still has some disadvantages. For fingerprint, people long-term need to work by hands might lose their fingerprint (e.g., repair worker). For iris recognition, it is necessary to use infra-red scanning eye that may cause safety concerns. Voice is not very distinctive and it can be imitated by training. The behavioral part of the speech of a person might change over time, and is also influenced by emotion and different physical states (such as common cold). Furthermore, collection of voice depend on the quality of microphone and is sensitive to background noise.

In addition to the disadvantages above mentioned of different biometrics, present-day biometrics can be stolen, duplicated, or even provided under violent threats. Therefore, a more robust person identification system is necessary.

## 2.3 EEG-based person identification systems

To address the disadvantages of existing biometrics some researchers brought up using brain signal as a biometric [21, 22]. To evaluate the uniqueness and consistency of the characteristics in EEG signal, the work in [18] confirmed that the inter-subject variation of EEG spectra where different subjects administered the same task was larger than the intra-subject variation where the EEG signals of the same subject were repeatedly acquired for several times. The characteristics in EEG signal fit well with the requirements of biometric system.

### 2.3.1 Basic components of EEG-based person identification

#### Signal pre-processing

Typical procedures include amplification, filtering and artificial rejection in order to improve the signal-to-noise ratio. In general, the brain activity is obscure and difficult to detect, a sufficient amplification consequently required. The bandpass filter is usually applied for filtering to cover high pass and low pas filter. In addition, a notch filter is also used to suppress the 60 Hz power line interference. For artificial rejection, the electro-oculographic (EOG) and electromyographic (EMG) are excluded detected by a predened threshold.

#### Feature extraction

The original signal of brainwave is chaotic that lead to difficulties in the practical application. In addition to common amplitude and latency information are used in analysis various feature extraction methods have been studied to extract more discriminative features, such as discrete wavelet transform, continuous wavelet transform, autoregression model (AR), power spectrum and so on.

**Classification**

The results of classification is to perform how suitable the feature used for distinguishing between different people. Many classification methods have been proposed in pattern recognition field. The classifier in a person identification can be anything from a simple linear model to a complex non-linear or a machine learning models. The acquired data are divided into training phase and testing phase. The training phase consists of a repetitive process of tasks to train a classifier and then use the testing phase to evaluate the performance.

### 2.3.2 EEG signals: resting data

The acquired signal of brainwave are mainly divided into two types, resting data and task dependent data. The resting data has the advantages of easy operation and when people in resting state the brain will generate the alpha rhythm that can be used as a waveform characteristic of each subject. In 2002, Poulos proposed a bilinear model to find the non-linear components in the EEG and the identification rate ranged from 72 to 85% [20]. By a lot of methods of feature extraction such as autoregressive (AR) coefficients, coherence and cross-correlation, the performance analysis of the system that Riera proposed obtained true acceptance rate of 96.6% [24]. For simplicity and practicability, the work [16] classified subjects simply by thresholding the EEG power spectrum.

### 2.3.3 EEG signals: task dependent data

Compared to resting data, the task-induced EEG signals are more specific that will reduce spontaneous effects of signal analysis. In 2003, Palaniappan and Ravi investigated the task-related EEG signals [19], the stimuli used in their work were standard image database [27]. By extracting features(channel wise power spectral density) from visual

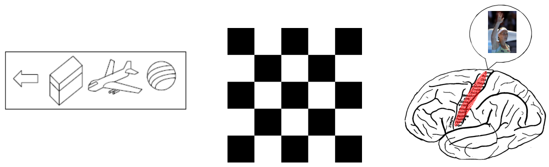

Figure 2.2: Task dependent experiment. To induce EEG of task dependent data there are three examples: From left to right are standard image database, checkerboard and motor imagery.

evoked potentials (VEPs), the identification accuracy was improved to be larger than $93\%$. A novel peak matching algorithm proposed by Singhal that only relied on recording from single channel gave 78% accuracy [26], they used a checkerboard pattern to induce steady-state visually evoked potentials (SSVEP). Moreover, Marcel [14] devised more appropriate mental tasks which contained imagination of repetitive self-paced hand movements and generation of words to perform their research. A statistical framework based on Gaussian mixture models and maximum a posteriori model adaptation successfully applied to person authentication.

## 2.4 Limitations of biometric systems

### Noise

A fingerprint with a scar, or a voice altered by cold are examples of noisy data. Similarly, the brainwave might be affected by emotions; however, this disadvantage may be an advantage for person identification that prevent impostors from threats. The old measuring instruments or unfavorable ambient conditions such as poor illumination in face recognition system will reduce the accuracy of discrimination.

### Intra-class variations

Although the biometric is considered stationary, it may be very different from the data used to generate the template. This variation is typically caused by a user who is incorrectly interacting with the sensor. For instance, the different angle used in face recognition.

Sensor replacement might cause different results of the same subject. The varying psychological makeup of an individual might result in different behavioral characteristic at various time measurements.

**Moral issues**

While the biometric is a more reliable and secure method of identity, some people may be reluctant to provide part of their body to as a biometric that they probably feel that privacy has been violated. It is important that biometrics acquiring needs to solicit for user's consent. To a large extent, the human factor dominates the success of the biometric-based identification system. Therefore, the biometric system must be user-friendly and accepted by users.

## 2.5 Thesis scope

After reviewing related researches for person identification, we attempt to adopt the method of training the EEG of subjects as a biometric to identify different persons. We designed a simple task that formed by different-sized disc and presented in different proportions of occurrences to induce the VEPs and ERPs. The feature extraction of raw EEG data which correspond with the classifier and can achieve the better classification is the major part of our work. The proposed EEG-based person identification system will test by 18 subjects to verify its reliability. The methods of feature extraction described in the Chapter 3, we used the features to train a accurate classifier which contained classification phase and verification phase.

The preliminary classification obtained by a multi-class classifier. The best-matching candidate of each classification is further verified by using a binary classifier to exclude the impostors in verification phase. For the performance evaluation, the accuracy rate and the error rate are were used. Besides, we tried to correct the false classified data by the confidence value of the SVM classifier. Summarizing the methods described above, we developed a reliable and accurate EEG-based person identification system.

# Chapter 3

# The proposed methods for person identification

## 3.1 System Overview

In this study we present a person identification system using EEG signals. Because resting state is prone to be more fluctuating, we adopt task-related EEG signals evoked by visual stimuli in this work. Representative information is extracted from the EEG signals of subjects and are used to train a one-to-many classifier for person classification. The best-matching candidate of each classification is further verified by using a binary classifier to exclude the impostor. The main structure of our system shows in Fig. 3.1.

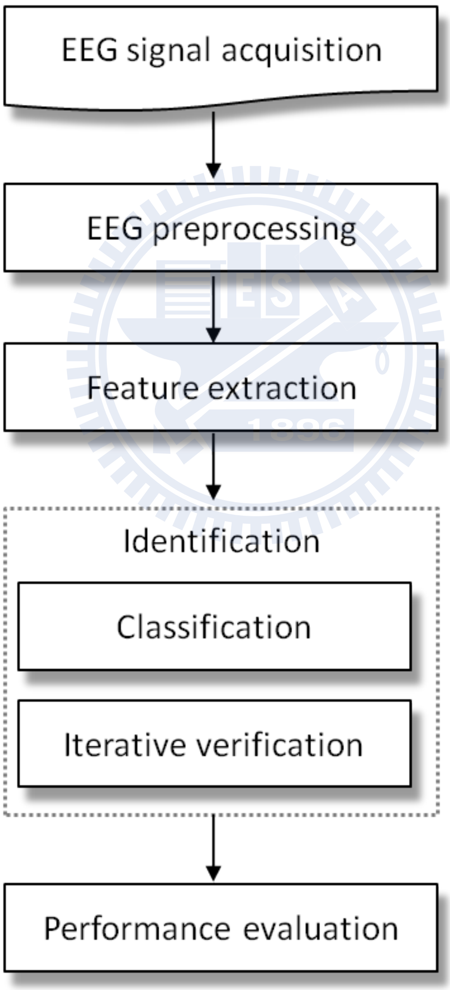


Figure 3.1: Flow chart of EEG-based person identification system.

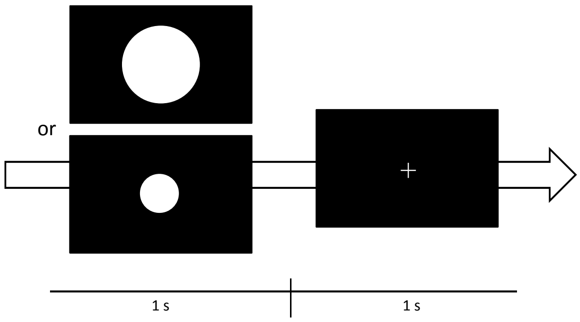


Figure 3.2: A trial consists of one-second stimulus, an image containing either a small disk or a large one, and one-second fixation.

## 3.2 Materials

### 3.2.1 Experiments

#### Participants

Eighteen subjects participated in this study (age ranges from 21 to 33 years with mean 24 years, twelve males). Informed consent was obtained from all the participants. All the subjects have normal or corrected-to-normal visions. For five participants among all the subjects, EEG data were acquired two times with an interval of more than one week.

#### Stimuli

The oddball paradigm of data acquisition in this study is shown in Fig. 3.2 and presented under the visual angle in human. The subject was seated comfortably in a silent room where no other signal interference except our experimental equipments and was asked to watch a monitor screen. The visual stimulus, an image containing either a small disk or a large one (ten times larger than the small one) and the visual angle is $6.1°$ and $1.3°$ separately, was presented for one second followed by another second of fixation image of a cross using Presentation 0.71 software. The order of the small disk or the large disk were randomize. During the fixation cross eye blink was allowable. The frequency ratio between the stimulus images is one (large disk) to three (small disk). Around 250 trials were acquired for each participant which contain at least 50 trials of large disk for analysis.

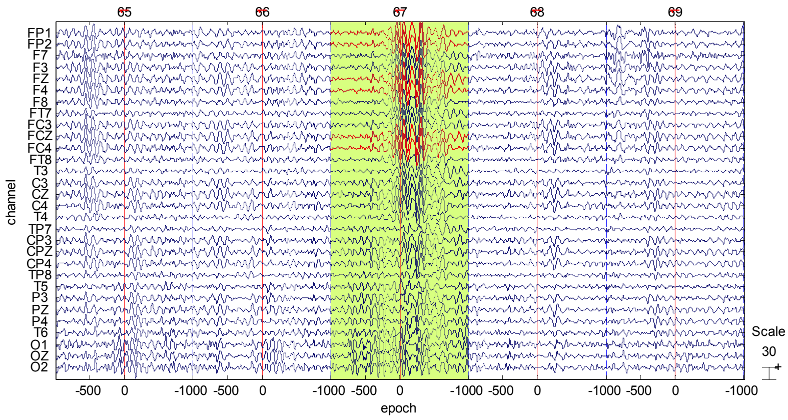


Figure 3.3: The epoch with eye movement in the green zone was rejected.

**EEG recording**

Thirty-two standard scalp electrodes were placed according to the International 10-20 System of Electrode Placement and the EEG data were recorded with Scan 4.3 software. The sampling rate for data acquisition was 500 Hz with a 16-bit A/D conversions. The earlobe electrodes A1 and A2 provided the reference. Impedance was kept below 5 k$\Omega$. Signals were digitally filtered within the 5-30 Hz band.

## 3.2.2 Signal preprocessing

We used EEGLAB 9.0 [4] to perform the following signal preprocessing procedure. The EEG data were first segmented into epochs starting from one second before the stimulus onset to one second after stimulus onset. The baseline correction picked the interval that relatively smooth compared to the activity regions was applied to remove the DC drift. Epochs with burst activities during the post-stimulus period were rejected (with the threshold values -50$\mu$V and 50$\mu$V) (Fig. 3.3).

The trials evoked by the large disk events were used in the following person identification analysis because it reflected the VEPs and ERPs.

## 3.3 Dimension reduction

The features we extracted from time domain series that the raw data have 500 sample points and it be calculated separately for all channels. A high-dimensional data need more time to analyze and it might include noise. In addition, it is difficult to interpret the character of the high-dimensional data. For this reason, how to reduce the dimension with an efficient method is important. In our work we reduced the time domain dimensions from 500 to 50.

### Principal components analysis

Principal component analysis (PCA) is a method for reducing feature dimension [10]. Its main idea is to find a set of basis, usually with a much smaller dimension, to represent the original data set while preserving as much as information measured by the variance of data distribution. For a N-dimension training data $\{\mathbf{x}_1, \mathbf{x}_2, ... \mathbf{x}_N\}$, PCA aim to find a linear transform matrix $\mathbf{P}$ which transform the $\mathbf{x}_k$ into M-dimension (M $\leqslant$ N). The $\mathbf{z}_k$ are more representative.

$$\mathbf{z}_k = \mathbf{P}^T \mathbf{x}_N, k = 1, 2, 3, ... N \quad , \tag{3.1}$$

The scatter matrix is the matrix of eigenvectors of $\mathbf{X}^T\mathbf{X}$, we need to find the $\mathbf{P}_{opt}$ to maximize the transformed $\mathbf{S}_T$.

$$\mathbf{S}_T = \sum_{k=1}^{N} (\mathbf{x}_k - \bar{\mathbf{x}})(\mathbf{x}_k - \bar{\mathbf{x}})^T \quad , \tag{3.2}$$

The original data are transformed by matrix of eigenvectors corresponding to the multiple eigenvalue of $\mathbf{S}_T$ will obtain $\mathbf{z}_k$ with low-dimension and maximum scatter matrix.

The procedure of PCA can be simplify as follows: 1. Compute the covariance matrix of the original input data. 2. Measure the eigenvalues of the covariance matrix in the order from large to small and then find the corresponding eigenvectors. 3. The normalized input data are multiplied by the eigenvectors obtaining the PCs. The data after PCA could represent the distribution of the original high-dimensional data in a low-dimensional space.

**Locally linear embedding**

LLE transforms the data to a low-dimensional space while preserving the relationships of relative distances between data points. If there is an embedded non-linear manifold lying in a high-dimensional space and the dimension of the manifold is relatively low, this manifold can be well represented in a low-dimensional space [25]. Therefore, we also applied the locally linear embedding (LLE) method to transform the data to a low-dimensional space while maintaining the manifold structure manifested in the original high-dimensional space. Firstly, we find a set of nearest neighbors for each data point $X_i$ in $D$-dimensional Euclidean space. Then we reconstruct, or represent, each data point by a linear combination of its neighbors $X_{ij}$ with weightings $W_{ij}$ as the contribution of the neighbor $X_{ij}$ to this linear combination for $X_i$. The reconstruction error is:

$$\mathbf{E}(W) = \sum_i |X_i - \sum_j W_{ij} X_{ij}|^2 \ , \tag{3.3}$$

where the sum of the weightings for each data point $X_i$ equals one. Like the previous cost function the data point $X_i$ can be mapped to the corresponding point $Y_i$ in a low-dimensional space as follow but here we fix the weights $W_{ij}$ while optimizing the coordinates $Y_i$.

$$\Phi(Y) = \sum_i |Y_i - \sum_j W_{ij} Y_{ij}|^2 \ , \tag{3.4}$$

The data point $X_i$ can be mapped to the corresponding point $Y_i$ in a low-dimensional space as:

$$\mathbf{Y}_i = \sum_j W_{ij} Y_{ij} \ , \tag{3.5}$$

where the point $Y_{ij}$ is the point in low-dimensional space corresponding to $X_{ij}$ in the original high-dimensional space. From neighbor k=1$\sim$n (n=30 that is the minimum number of trials), the classification accuracy is the highest when k=9. Therefore, we set the number of neighbor k to be 9.

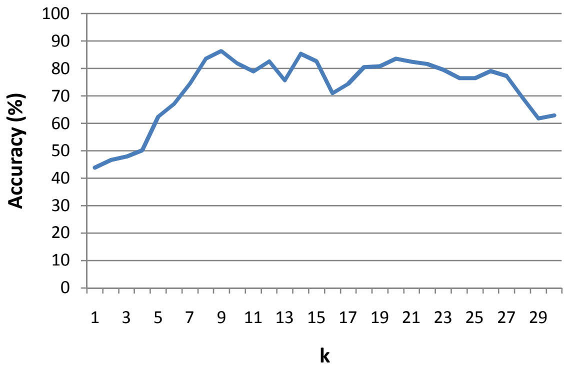


Figure 3.4: Classification of different number of neighbor (k). The classification accuracy is the highest when k=9.

## 3.4 Feature extraction

The signal-to-noise ratio (SNR) of brainwave is low. It will take a large amount of time to calculate the compact bases which represent the original signal; moreover, the number of obtained bases might be infinite. Through efficient feature extraction we can reduce the quantity of data and transform the original data into representative coefficients. Thus the computational load can be decrease and the acquired data will be more applicable.

### Morphological features

Evoked potentials are characteristic waveforms that are reproducible by time-locking EEG to a stimulus over repeated trials. Because of the characteristics and typical delay of EEG components the morphological features are calculated as follows [1].

$$t_{s_{max}} = \{t|s(t) = s_{max}\} \ , \tag{3.6}$$

The latency ($t_{s_{max}}$) and amplitude ($s_{max}$) of each EEG epoch were computed as the morphologic features which contain VEPs (with the time interval from 50 ms to 150 ms after stimulus onset) and ERPs (with the time interval from 250 ms to 400 ms after stimulus onset). Latency to amplitude ratios (LARs) also were measured as a morphologic features. Fig. 3.5 depicts morphological features of averaged data acquired from different subjects, it is discriminated despite being stimulated by the same task.

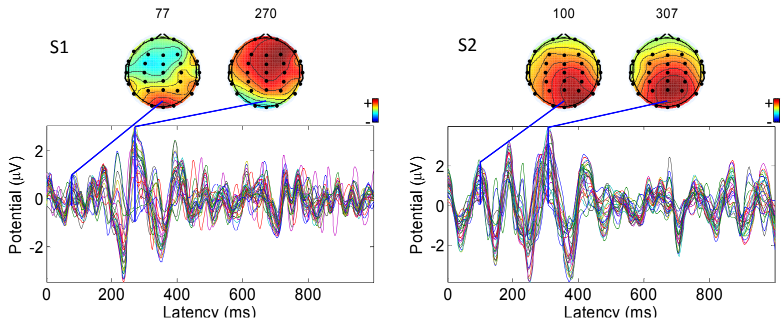


Figure 3.5: The latency and amplitude of VEP and ERP of different subjects.

### Frequency features

The spectral analysis has been used in signal processing of EEG for a few decade and it also applied to person identification [7, 21]. EEG is composed of different wave bands, often called as the rhythm of waves. The frequency spectrum mainly divide into five bands termed delta (2-4 Hz), theta (4-8 Hz), alpha (8-12 Hz) and beta (12-32 Hz). In this work we focus on the frequency band from 5 Hz to 30 Hz. Because rhythmic delta and theta activity related to sleep condition. The discrete Fourier transform (DFT) were used to compute the power spectrum for each epoch.

$$X(k) = \sum_{j=1}^{N} x(j) w_N^{(j-1)(k-1)} \quad , \tag{3.7}$$

where

$$w_N = e^{(-2\pi i)/N} \tag{3.8}$$

is an $N$th root of unity. In our case, $N$ is equal to 500 (500 Hz*1 second after onset).

### Stochastic modeling

Poulos and Rangoussi [22] have proposed to model the EEG signal by autoregressive (AR) models and the parameters of the AR model are used for identification. The presented work [17] utilized the coefficients of AR model as features and reached correct classification scores at the range of 80% to 100%. Considering the EEG signal as an AR process, we used the Yule-Walker equations to estimate the AR coefficients as the features. To fit a

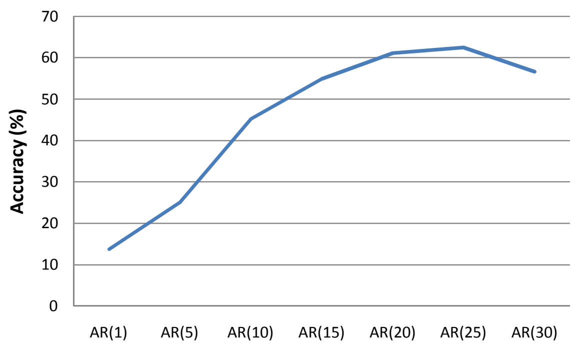


Figure 3.6: When the number of coefficients is 25 the result of classification achieve the best.

$p$th-order AR model to the EEG data $X(t)$, we minimize the following prediction error by using the least squares regression:

$$X(t) = \sum_{i=1}^{P} a(i)X(t-i) + e(t) \quad , \tag{3.9}$$

where $a(i)$ are the auto regression coefficients, $e(t)$ represents the white noise, and the time series can be estimated by a linear differential equation. Different number of parameters depended on the order of models were used to calculate the results of classification Fig. 3.6.

**Time-frequency model**

The wavelet transform uses a set of time-scale basis to represent the original signal. Fig. 3.7 shows the fundamental of discrete wavelet transformation. For one level of the transform, signal $S$ is divided in half which the approximation coefficients retain the representative information of $S$ and the detail coefficients include comparatively unimportant information such as noise. Here we applied the Daubechies wavelets to transform the time-domain EEG signals and obtained 250 coefficients as the time-frequency features.

## 3.5 Identification

For classifier design, we employed the support vector machine (SVM) and the k-nearest neighbor (kNN) search method (k=9). Through the results of classification we can deter-

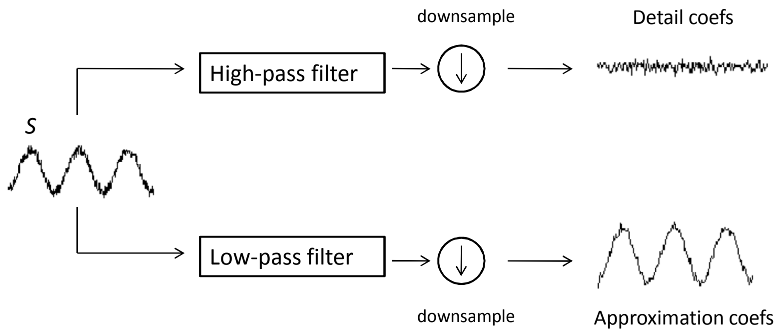


Figure 3.7: The samples are decomposed through a low-pass filter for approximation and high-pass filter for detail.

mine the feasibility of EEG-based person identification.

**SVM**

Support vector machine (SVM) [3] is a statistical-based classification method that finds a hyperplane to separate the two different sets of data. Using the term hyperplane is due to the data may be a high-dimensional information. Examples of two-dimensional data shows in Fig. 3.8 that we attempt to find a line that will separate the red point and blue point and distance of the border (margin) between these two sets is maximal. Assume there are a set of training data $\{x_i, y_i\}$, $i$=1,...,$n$ and $x_i \epsilon R^d$, $y_i$ denotes a known class label. The optimal separating hyperplane can be solve as follow:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2} \parallel w \parallel^2 \\ \text{subject to} \quad & y_i(w^T x_i - b) - 1 \geq 0 \quad \forall i \end{aligned} \tag{3.10}$$

We utilized the linear type SVM as the classifier to train multi-class classifier for classification and one-against-rest classifier for verification.

**kNN**

The other method of classification we compared to SVM is the k-nearest neighbor (kNN), it is widely used in pattern recognition because of uncomplicated basis of clas-

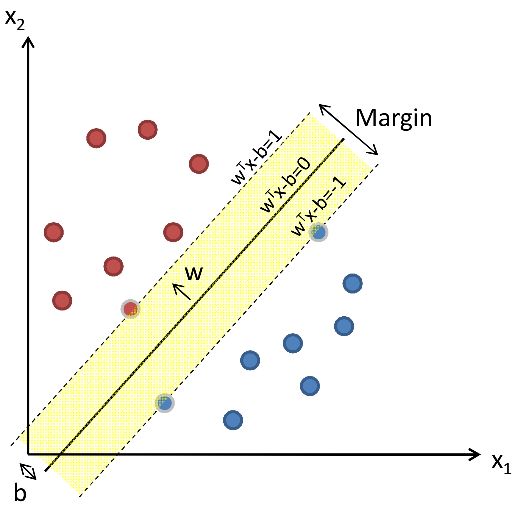


Figure 3.8: Two set of data we attempt to find a decision boundary (solid line) to separate and the margin (dashed line) is maximal. Data locates on the margin is defined as support vector.

sification. The training examples are vectors in a multidimensional feature space, each with a class label. The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples. The distance between instances is typically determined by Euclidean distance. The classification accuracy often can be improved with the larger value of "k" because it reduce the effect of noise on the classification; however, it will require a longer computing time. We set the value of k to be 9 (the same number of neighbor k for LLE).

### Cross-validation

Cross-validation, is a technical assessment of the results of statistical analysis will be extended to an independent data set. It is mainly used to fairly estimate how accurately a predictive model will perform.

The common type of the cross-validation are k-fold, repeated random sub-sampling and leave-one-out. The disadvantage of repeated random sub-sampling is that some observations may never be selected in the validation sub-sample, whereas others may be selected more than once. The leave-one-out cross-validation usually used in a small number of

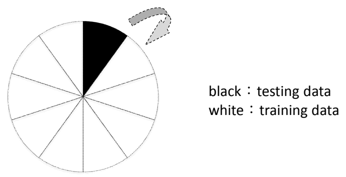


Figure 3.9: The 1/10 trials are picked as testing data, the other 9/10 trials are training data. Repeat this way and use the next part of testing data each time.

analysed data. For these reasons, we choose k-fold cross-validation to measure our classification. The k-fold method that divides all data into k parts then one part picked as the testing data. The remaining k-1 parts were trained by the classifier ( Fig. 3.9). The cross-validation process is then repeated K times (the folds), with each of the K sub-samples used exactly once as the validation data. We apply the 10-fold cross-validation to obtain the average classification accuracy for person identification.

### 3.5.1 Classification

In the classification phase, the system recognizes an individual by a multi-class classifier. A one-to-many comparison is established for individuals identity without the subject having to claim an identity. This part can be regarded as negative recognition application where the system establishes whether the person is who he/she denies to be. The purpose of negative recognition is to prevent users from single person to use multiple identities. The block diagrams of classification phase are depicted in Fig. 3.10.

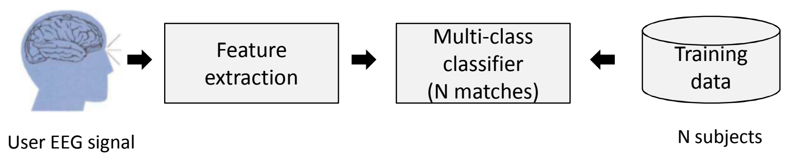


Figure 3.10: Block diagrams of classification. The multi-class classifier obtained a one-to-many comparison of testing data. For a N-class classifier, it will be N-classification results.

### 3.5.2 Iterative verification

The purpose of the verification procedure is to reconfirm the best-matching result of classification (Fig. 3.11).

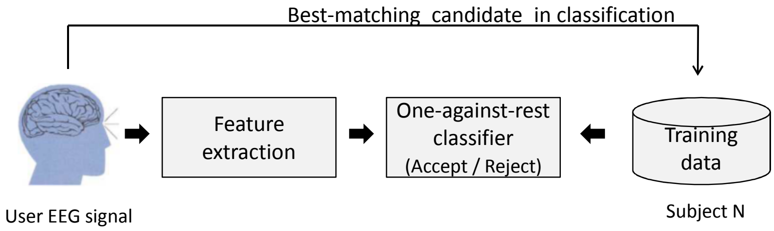


Figure 3.11: Block diagrams of verification. The best-matching result of classification was verified using the one-against-rest classifier.

For each of the eighteen subjects, we trained a binary classifier by using to groups of training data, EEG data of the targeted subject and those of all others. We evaluate the binary classifier for verification according to the accuracy rate and the error rate. The best-matching subject from the classification procedure is verified by the corresponding binary classifier.

In addition, we correct the false classified data in classification phase through iterative verification. The confidence of SVM classifier makes a criteria for classification and determines whether the data have chance of being corrected. Only the correctly verified results are approved whereas the failure results are regarded as impostors.

# Chapter 4

# Experimental results

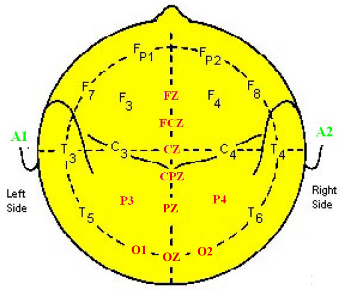


Figure 4.1: Ten channels were used in our analysis.

## 4.1 Results of classification

### 4.1.1 Temporal characteristics in the acquired signals

We first verified whether the resting EEG or ERP is better for distinguishing subjects' identities. By applying the SVM classifier to categorize the pre-stimulus (500 ms before onset) EEG signals among the fifteen subjects with whole brain information, the classification accuracy achieved 12.24%. When the post-stimulus (500 ms after onset) ERP signals were used for person identification, the classification accuracy achieved 25.26%. Therefore the ERP contains more information for person identification than resting EEG does and it truly can improve classification.

Furthermore, we picked the channels which are related to the visual field and significant activities of stimuli (frontal, frontal-central, parietal and occipital) [5]. The 10 electrodes we selected were Fz, FCz, Cz, CPz, P3, Pz, P4, O1, Oz, O2 (Fig. 4.1). This process will reduce the quantity of data and eliminate the activities which are not induced by the events. The classification of the whole brain 30 channels is lower than the selected 10 channels (29.28%) confirms the above statement.

### 4.1.2 The identification in classification phase

Because of the above result, the results of feature extraction (1000 ms post-stimulus signals, 10 channels) demonstrated by classifying single trial and averaging data (Table 4.1). For the total number of 1144 data, the best result of classification is using power spectrum as the features in SVM classifier and the data which after averaging is 1126 trials for smoothing noise is helpful to improve the classification. We increased the number of averaged data and obtained the classification accuracy: 90.25% (avg:3), 88.26% (avg:4), 89.55% (avg:5), and 88.19% (avg:10) by using the power spectrum as features. The results show that more average number of data does not improve the classification accuracy, and in practical application we attempt to use fewer data to achieve better identification efficiency.

Table 4.1: Results of classification with different features. The data of each subject acquired in the same experiment. The feature of power spectrum obtained the classification rate of 91.61%.

| | **SVM** | | **kNN** | |
|---|---|---|---|---|
| **Feature** | Single trial | Avg:2 | Single trial | Avg:2 |
| Raw data | 29.28% | 80.82% | 23.43% | 76.38% |
| LLE | 30.68% | 86.32% | 28.06% | 83.39% |
| PCA | 27.62% | 83.21% | 25.26% | 81.17% |
| Latency | 11.54% | 35.17% | 10.23% | 33.57% |
| Amplitude | 38.55% | 50.80% | 36.19% | 45.20% |
| LAR | 39.25% | 52.66% | 37.67% | 47.51% |
| Power spectrum | 72.03% | **91.56**% | 60.05% | 85.88% |
| AR | 53.50% | 62.52% | 50.96% | 60.57% |
| Wavelet | 27.27% | 85.35% | 22.90% | 77.26% |

Fig. 4.2 shows the power spectrum at the frequency band from 5 Hz to 30 Hz of different subjects. The spectrum is equivalent for each subjects while is different from another subject.

In order to cover information of complementary domain we combined different features. The features of different domains were normalized before combining with others.

$$x_i = \frac{x_i - x_{min}}{|x_{max} - x_{min}|} \times I \ \ ,1 \leq i \leq n \tag{4.1}$$

where $n$ is the number of trials and the normalized features rang between 0 and 1. The $I$ is a constant to prevent the variance presented in data is too small, we set the $I$=10 in our

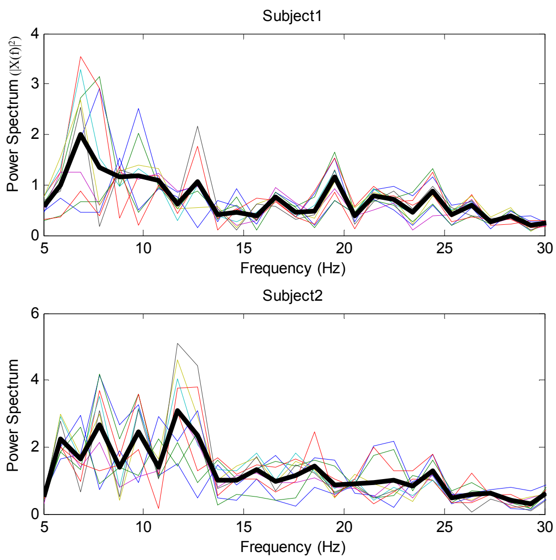


Figure 4.2: The power spectrum in different 10 trials of two subjects (thick black line represents the average value). Each trial show the averaging data in ten channels.

work. The results of combined features are shown in Table 4.2. It is improved compared to unilateral features.

Table 4.2: Results of classification with combined features.

| | SVM | |
|---|---|---|
| **Combined feature** | Single trial | Avg:2 |
| Spectrum + Latency | 73.43 | 91.47 |
| Spectrum + Amplitude | 73.69 | 91.65 |
| Spectrum + LAR | 78.32 | 92.10 |
| Spectrum + PCA | 82.17 | 96.00 |
| Spectrum + LLE | 85.31 | **96.36** |
| Spectrum + AR | 74.48 | 88.28 |

The best result of the features that combined spectrum and LLE that we chose to train our classifier. In predictive analytics, a table of confusion, also known as a confusion matrix to report the percentage of true negatives, false positives, false negatives, and true positives.

The relationship among terms are shown as Table 4.3.

Table 4.3: Relationships among terms. Definitions of True positive (TP), False negative (FN), False positive (FP) and True negative (TN).

| | | **Prediction outcome** | |
|---|---|---|---|
| | | **positive** | **negative** |
| **Actual value** | **positive'** | True Positive (TP) | False Negative (FN) |
| | **negative'** | False Positive (FP) | True Negative (TN) |

A good classification results on high proportion of true positive. Fig. 4.3 is the confusion matrix of 18 subjects using combined features of power spectrum and LLE in 10-fold cross-validation. The blue frames of the matrix is the positive predictive value gave 81-100% and it happens to the best-matching for each subjects.

| Actual \ Predicted | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **subject1** | 100% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| **subject2** | 0.00% | 100 % | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| **subject3** | 0.00% | 0.00% | 100% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| **subject4** | 0.00% | 0.00% | 0.00% | 97.50% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 2.50% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| **subject5** | 0.00% | 4.29% | 0.00% | 0.00% | 92.86% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 2.86% |
| **subject6** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| **subject7** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 98.00% | 0.00% | 0.00% | 2.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| **subject8** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| **subject9** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 3.33% | 0.00% | 0.00% | 93.33% | 0.00% | 0.00% | 0.00% | 3.33% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| **subject10** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| **subject11** | 1.43% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 1.43% | 0.00% | 97.14% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| **subject12** | 0.00% | 3.33% | 1.67% | 0.00% | 0.00% | 0.00% | 1.67% | 0.00% | 1.67% | 3.33% | 0.00% | 86.67% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 1.67% |
| **subject13** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 98.57% | 0.00% | 0.00% | 0.00% | 1.43% | 0.00% |
| **subject14** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100% | 0.00% | 0.00% | 0.00% | 0.00% |
| **subject15** | 1.67% | 0.00% | 0.00% | 0.00% | 0.00% | 1.67% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 96.67% | 0.00% | 0.00% | 0.00% |
| **subject16** | 0.00% | 0.00% | 1.67% | 0.00% | 0.00% | 0.00% | 6.67% | 0.00% | 0.00% | 1.67% | 0.00% | 0.00% | 0.00% | 0.00% | 5.00% | 81.67% | 1.67% | 1.67% |
| **subject17** | 2.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 98.00% | 0.00% |
| **subject18** | 6.67% | 3.33% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 1.67% | 0.00% | 0.00% | 0.00% | 0.00% | 88.33% |

Figure 4.3: Confusion matrix of 18 subjects. An element in row i and column j counts the percentage of subject i was classified as j. The blue frames of the matrix is the positive predictive value.

## 4.2 Results of verification

### 4.2.1 Accuracy determination in verification phase

Through the results of classification phase we used the binary classifier to verified the best-matching. It is aim to reconfirm the accuracy of classification and to exclude the impostor. The best situation is high accuracy rate and low error rate obtained from verification. Furthermore, we attempt to correct the false classified data in classification phase in order to increase the initial classification.

#### Accuracy rate

The accuracy rate measures the percentage of the best-matching candidate in classification that are accepted in binary classifier of verification. In other words, we verified the data which displayed in blue frames in Fig. 4.3. Table 4.4 shows the accuracy rate of 18 subjects, the average could reach 98.89% accuracy.

Table 4.4: Accuracy rate of verification phase. The percentage of the best-matching candidate in classification that are accepted in verification.

| **Subject** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
|---|---|---|---|---|---|---|---|---|---|
| Accuracy rate (%) | 98.57 | 98.75 | 98.75 | 100 | 100 | 96.67 | 93.88 | 100 | 96.43 |
| **Subject** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** |
| Accuracy rate (%) | 100 | 100 | 100 | 98.55 | 100 | 94.83 | 100 | 100 | 100 |

#### Error rate

The error rate measures the percentage of the false classified data in classification that are accepted in binary classifier of verification. In other words, we observed whether the data which displayed in white frames in Fig. 4.3 will be rejected in the verification phase. Table 4.5 shows the error rate of 18 subjects, all the false classified data were rejected in verification phase.

Table 4.5: Error rate of verification phase. The percentage of the false classified data in classification that are accepted in verification.

| **Subject** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
|---|---|---|---|---|---|---|---|---|---|
| Accepted/False classified | 0/7 | 0/7 | 0/2 | 0/0 | 0/0 | 0/2 | 0/5 | 0/0 | 0/2 |
| Error rate (%) | 0 | 0 | 0 | - | - | 0 | 0 | - | 0 |
| **Subject** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** |
| Accepted/False classified | 0/4 | 0/0 | 0/0 | 0/3 | 0/0 | 0/3 | 0/0 | 0/2 | 0/4 |
| Error rate (%) | 0 | - | - | 0 | - | 0 | - | 0 | 0 |

### Iterative verification

We attempt to correct the data that should be classified as subject i, but classified as subject j in classification by the iterative verification. The criterion complied with the confidence value when it higher than 80% compared to the maximum, if the data which false classified have chance of being classified correctly the accuracy will be improved. The confidence value of the SVM classifier makes a criteria for classification shown in Fig. 4.4 determined whether the data have chance of being corrected. For the true positive data in classification phase, the blue diamond is the maximum confidence value and it is significantly large. While the false classified data shown in green triangles are smaller compared to the true positive data and the confidence values of true classes (purple square) have little difference from maximum confidence value. Table 4.6 illustrates the results of iterative verification.

After iterative verification the overall accuracy of our system is 97.25% that is higher than non-iterative verification (95.29%).

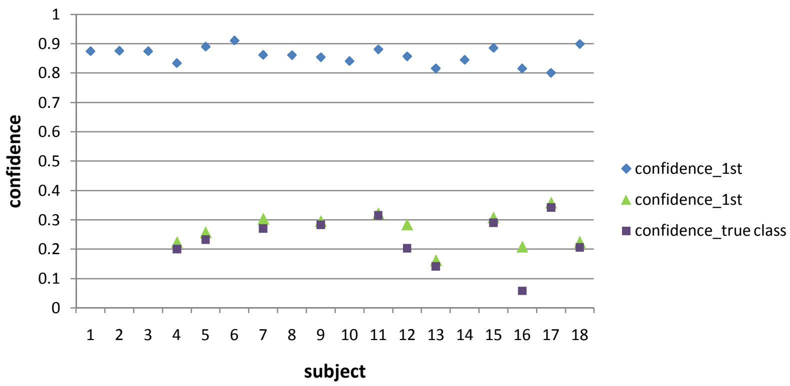


Figure 4.4: Confidence value of the classification mainly divide into true positive data (blue diamond) and false classified data (green triangle and purple square).

Table 4.6: Iterative verification. The row of False classified and Corrected indicate that the number of false classified data which are corrected in the binary classifier of iterative verification.

| **Subject** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
|---|---|---|---|---|---|---|---|---|---|
| Corrected/False classified | 0/0 | 0/0 | 0/0 | 1/1 | 5/5 | 0/0 | 1/1 | 0/0 | 0/2 |
| Original accuracy (%) | 100 | 100 | 100 | 97.50 | 92.86 | 100 | 98.00 | 100 | 93.33 |
| New accuracy (%) | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 93.33 |
| **Subject** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** |
| Corrected/False classified | 0/0 | 1/2 | 4/8 | 1/1 | 0/0 | 1/2 | 1/11 | 1/1 | 6/7 |
| Original accuracy (%) | 100 | 97.14 | 86.67 | 98.57 | 100 | 96.67 | 81.67 | 98.00 | 88.33 |
| New accuracy (%) | 100 | 98.57 | 93.33 | 100 | 100 | 98.33 | 83.33 | 100 | 98.33 |

## 4.3 Results of different days

For eight participants among all the subjects (three females and five males), EEG data were acquired two times with an interval of more than one week. Because we would like to know whether the EEG data is sufficiently constant even though the data were acquired in different day. The day1 signal were used as training data and the day2 signal as testing data for classification. We observed the power spectrum of different day (Fig. 4.5 to 5.12). The power spectrum in different 10 trials of two subjects (thick black line represents the average value). Each trial show the averaging data in ten channels. Although it is impossible that the features of power spectrum are exactly the same they have a certain degree of similarity.

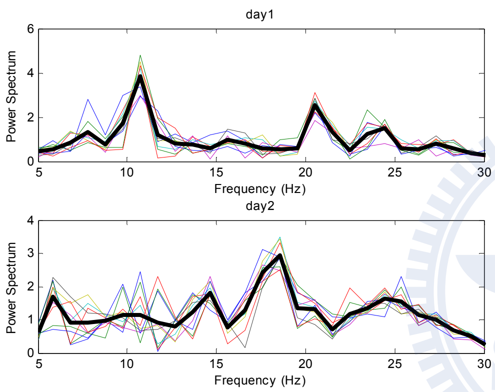


Figure 4.5: Power spectrum of subject3 acquired from different day

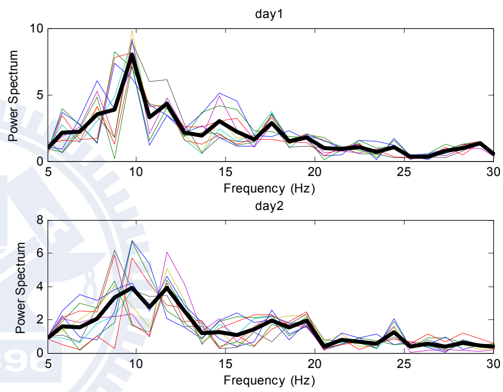


Figure 4.6: Power spectrum of subject5 acquired from different day

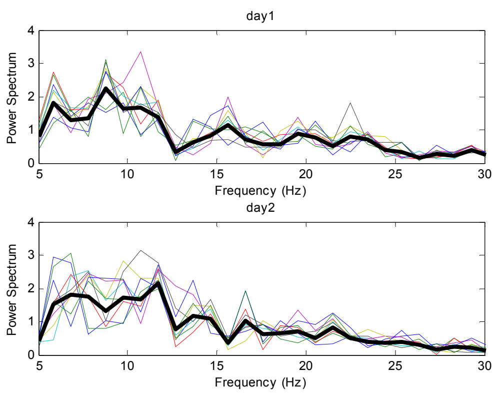


Figure 4.7: Power spectrum of subject8 acquired from different day

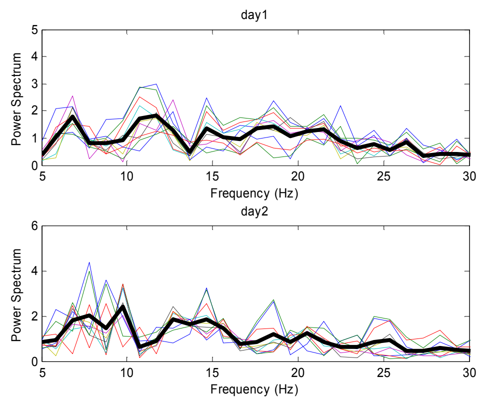


Figure 4.8: Power spectrum of subject9 acquired from different day

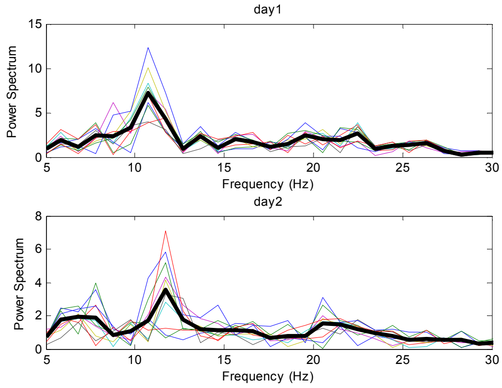


Figure 4.9: Power spectrum of subject12 acquired from different day

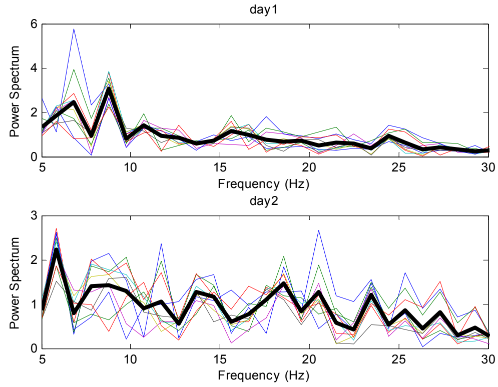


Figure 4.10: Power spectrum of subject13 acquired from different day

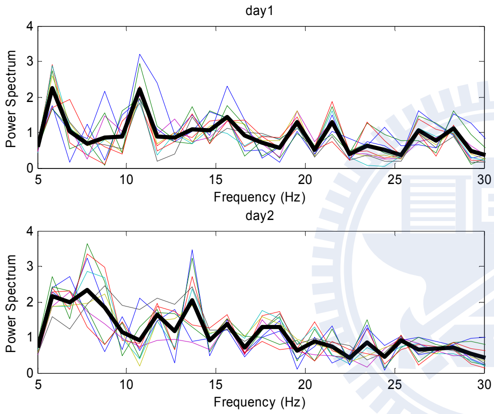


Figure 4.11: Power spectrum of subject17 acquired from different day

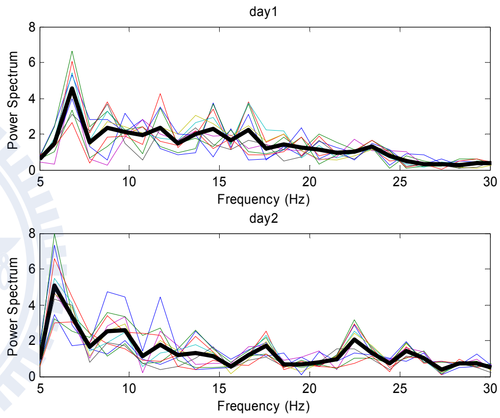


Figure 4.12: Power spectrum of subject18 acquired from different day

In addition to signal observation, we need to classify the data of different day by classification and verification. For classification phase, the confusion matrix of eight subjects (the number of 436 data) shown in Fig. 4.13 and the features we used are the same as preceding section (power spectrum combine with LLE). The inter-class accuracy is 63.08%. After iterative verification the overall accuracy of our system is 46.26% that is higher than non-iterative verification (41.36%). The detail data of accuracy rate, error rate and results of iterative verification are as follows. The results demonstrate that the performance degrades over days. We will discuss in detail in next chapter.

| Actual \ Predicted | 3 | 5 | 8 | 9 | 12 | 13 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|
| **subject3** | 81.40% | 2.33% | 0.00% | 0.00% | 2.33% | 0.00% | 0.00% | 13.95% |
| **subject5** | 0.00% | 97.30% | 0.00% | 1.35% | 0.00% | 0.00% | 0.00% | 1.35% |
| **subject8** | 0.00% | 1.67% | 75.00% | 0.00% | 0.00% | 0.00% | 18.33% | 5.00% |
| **subject9** | 0.00% | 0.00% | 0.00% | 0.00% | 7.69% | 23.08% | 60.00% | 9.23% |
| **subject12** | 0.00% | 5.66% | 0.00% | 0.00% | 88.68% | 1.89% | 1.89% | 1.89% |
| **subject13** | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 29.63% | 70.37% | 0.00% |
| **subject17** | 20.00% | 0.00% | 0.00% | 5.71% | 0.00% | 14.29% | 48.57% | 11.43% |
| **subject18** | 0.00% | 1.92% | 0.00% | 0.00% | 3.85% | 1.92% | 3.85% | 88.46% |

Figure 4.13: Confusion matrix of eight subjects with data of different day. An element in row i and column j counts the percentage of subject i was classified as j. The blue frames of the matrix is the positive predictive value.

Table 4.7: Accuracy rate of different day. The percentage of the best-matching candidate in classification that are accepted in verification. The average value is 68.52% accuracy.

| **Subject** | **3** | **5** | **8** | **9** | **12** | **13** | **17** | **18** |
|---|---|---|---|---|---|---|---|---|
| Accuracy rate (%) | 37.14 | 97.22 | 100 | 0 | 80.85 | 25 | 17.65 | 26.09 |

Table 4.8: Error rate of different day. The percentage of the false classified data in classification that are accepted in verification. The average is 46.84%.

| **Subject** | **3** | **5** | **8** | **9** | **12** | **13** | **17** | **18** |
|---|---|---|---|---|---|---|---|---|
| Accepted/False classified | 0/7 | 4/6 | 0/0 | 0/3 | 3/8 | 3/22 | 64/91 | 0/21 |
| Error rate (%) | 0 | 66.67 | - | 0 | 37.5 | 13.64 | 70.33 | 0 |

Table 4.9: Iterative verification. The row of False classified and Corrected indicate that the number of false classified data which are corrected in the binary classifier of verification.

| **Subject** | **3** | **5** | **8** | **9** | **12** | **13** | **17** | **18** |
|---|---|---|---|---|---|---|---|---|
| Corrected/False classified | 0/8 | 0/2 | 6/15 | 0/65 | 0/6 | 14/38 | 1/18 | 0/6 |
| Original accuracy (%) | 81.40 | 97.30 | 75 | 0 | 88.68 | 29.63 | 48.57 | 88.46 |
| New accuracy (%) | 81.40 | 97.30 | 85 | 0 | 88.68 | 55.55 | 51.43 | 88.46 |

## 4.4 Summary

The theoretical value of transfer time in our system is:

$$\{\sum_{n=0}^{\infty} \frac{1}{16} \times (\frac{3}{4})^n (n+2)(n+1)\} \times N \tag{4.2}$$

where $N$ is two for average of two trials obtains 16 s for one identification. While the real transfer time in our experiments is 15.68 s.

**Classification phase**

- By the observation of original signal we confirmed that the task dependent data is advantageous in classification.
- The inter-class accuracies of confusion matrix range from 81 to 100% by using combined feature of power spectrum and locally linear embedding (LLE) which is a method of dimension reduction.

**Verification phase**

- The accuracy rate attain 98% that measures the percentage of the best-matching candidate in classification that are accepted in binary classifier of verification.
- In addition to a high accuracy rate, the error is 0% indicates that the system has a good effect against impostors. Through iterative verification the data false classified in classification phase could be corrected and the overall accuracy of our system could reach 97.25%.

# Chapter 5

# Discussion

### Feature selection

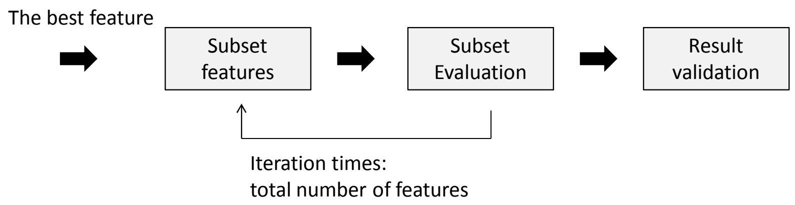


Figure 5.1: Steps of forward selection. The process of feature combination in order to find the best one.

For the feature selection we used the forward and backward selection to find the best combination of different features [11]. A forward selection method first finds the best feature among all features, then the rest of the features combined with it sequentially (Fig. 5.1). In our work, we firstly combined power spectrum with other features separately and found the best combination. After the combination of two features the other features added sequentially according to the classification results. Through this process we can find that which combination is helpful to improve the classification and which feature will depress the accuracy.

The result of forward selection shown in Fig. 5.2, the most obvious finding is that the feature of AR depress the accuracy. Therefore, we inferred that AR is not suitable for combining with other features. Furthermore, the results also reveal LLE dominates an important role of the combined features. Summing up the above results and considering computing time, we chose the combined feature of power spectrum and LLE to train the classifier.

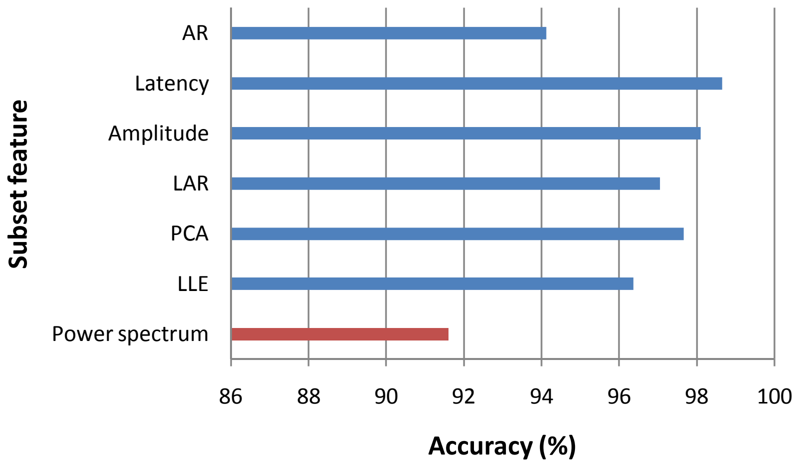


Figure 5.2: Results of forward selection. Different features are combined with the feature of power spectrum according to the classification results.

### Reason for bad classification

Although the performance of our person identification system can reach more than 95% accuracy, some misclassification of cases still exist. The reasons which will impact the accuracy of classification include external and internal interference. The external interference is caused by signal quality, during different EEG acquisitions the results of the instrument measurement will have a range of error and the environmental factors such like noise will impact on signal quality. The internal interference is difficult to determine and eliminate, the main factor is the state of subject. Different degree of concentration, tension or tiredness will impact on signal quality. However those cognitive factors are difficult to quantify we just can make a reasonable conjecture according the results.

By confusion matrix of 18 subjects(Fig. 4.3) we can clearly realize the classification of each one. We used the correlation to discuss those classifications. The Fig. 5.3 shows the detail of correlation between different subject and the Fig. 5.4 is the averaging data. Those subject have high correlation with the others are usually misclassified. For instance, many people were wrongly assigned to subject1 and its correlation is high indeed (like a common face). The subject5, subject12 and subject18 which with high correlation are more often misclassified as others.

| Correlation | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 | S11 | S12 | S13 | S14 | S15 | S16 | S17 | S18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S1 | 1.000 | 0.374 | 0.495 | 0.393 | 0.505 | 0.433 | 0.382 | 0.391 | 0.495 | 0.399 | 0.513 | 0.445 | 0.338 | 0.316 | 0.541 | 0.493 | 0.360 | 0.590 |
| S2 | | 1.000 | 0.347 | 0.342 | 0.404 | 0.361 | 0.532 | 0.387 | 0.305 | 0.490 | 0.343 | 0.610 | 0.378 | 0.261 | 0.133 | 0.293 | 0.385 | 0.477 |
| S3 | | | 1.000 | 0.374 | 0.425 | 0.367 | 0.399 | 0.316 | 0.338 | 0.359 | 0.393 | 0.432 | 0.200 | 0.168 | 0.114 | 0.304 | 0.301 | 0.390 |
| S4 | | | | 1.000 | 0.432 | 0.425 | 0.410 | 0.361 | 0.396 | 0.347 | 0.336 | 0.348 | 0.291 | 0.154 | 0.151 | 0.279 | 0.305 | 0.472 |
| S5 | | | | | 1.000 | 0.476 | 0.448 | 0.414 | 0.433 | 0.454 | 0.405 | 0.546 | 0.458 | 0.377 | 0.395 | 0.306 | 0.348 | 0.616 |
| S6 | | | | | | 1.000 | 0.439 | 0.185 | 0.390 | 0.412 | 0.415 | 0.319 | 0.332 | 0.158 | 0.231 | 0.423 | 0.346 | 0.507 |
| S7 | | | | | | | 1.000 | 0.341 | 0.509 | 0.603 | 0.437 | 0.540 | 0.469 | 0.371 | 0.223 | 0.480 | 0.368 | 0.436 |
| S8 | | | | | | | | 1.000 | 0.398 | 0.293 | 0.366 | 0.459 | 0.315 | 0.263 | 0.130 | 0.209 | 0.289 | 0.366 |
| S9 | | | | | | | | | 1.000 | 0.405 | 0.527 | 0.416 | 0.292 | 0.265 | 0.212 | 0.243 | 0.299 | 0.547 |
| S10 | | | | | | | | | | 1.000 | 0.422 | 0.548 | 0.370 | 0.269 | 0.160 | 0.277 | 0.366 | 0.530 |
| S11 | | | | | | | | | | | 1.000 | 0.433 | 0.438 | 0.486 | 0.448 | 0.428 | 0.515 | 0.575 |
| S12 | | | | | | | | | | | | 1.000 | 0.364 | 0.403 | 0.349 | 0.357 | 0.388 | 0.472 |
| S13 | | | | | | | | | | | | | 1.000 | 0.328 | 0.271 | 0.283 | 0.328 | 0.438 |
| S14 | | | | | | | | | | | | | | 1.000 | 0.522 | 0.447 | 0.295 | 0.371 |
| S15 | | | | | | | | | | | | | | | 1.000 | 0.495 | 0.441 | 0.355 |
| S16 | | | | | | | | | | | | | | | | 1.000 | 0.484 | 0.441 |
| S17 | | | | | | | | | | | | | | | | | 1.000 | 0.386 |
| S18 | | | | | | | | | | | | | | | | | | 1.000 |

Figure 5.3: Correlation between different subjects. We measured the correlation of the raw data of 18 subjects, each data of subject is the averaged value of total trials.

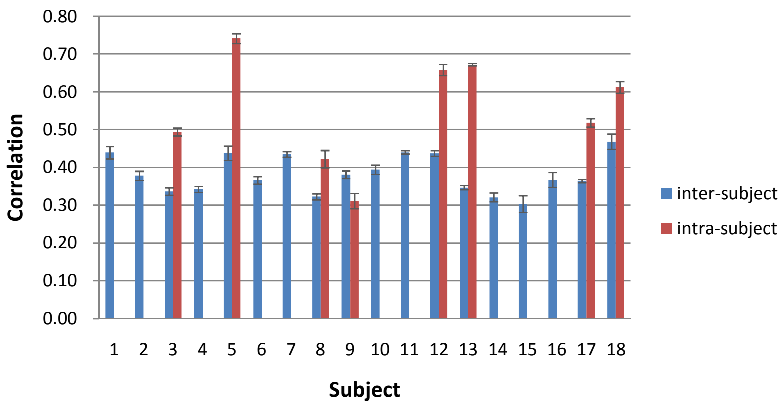


Figure 5.4: The averaging data of correlation between different subjects.

In addition, we interest in the correlation between data of different day. Compared to the correlation between subject it is higher (the red square shown in Fig. 5.4 and the detail in Table 5.1). The results illustrate that the EEG data of same individual is steady as opposed to data of different individuals.

Table 5.1: The correlation obtained from the eight subjects with data of different day.

| **Subject** | **3** | **5** | **8** | **9** | **12** | **13** | **17** | **18** |
|---|---|---|---|---|---|---|---|---|
| Correlation | 0.494 | 0.741 | 0.422 | 0.311 | 0.658 | 0.672 | 0.518 | 0.612 |

**Adaptive classifier comparison**

Although the EEG acquired from different day is steady from the results of correlation, we attempt to adapt the classifier using data of different day as training set. The classifier consequently more conform to each subjects. The inter-class accuracy of adaptive classifier upgrade from 63.08% to 97.66%, accuracy rate from 68.52% to 92.11%, and error rate decrease from 46.84% to 10% (Table 5.2). After iterative verification the overall accuracy of our system is 90.65% that is higher than non-iterative verification (89.95%). The results improve significantly using the adaptive classifier (46.26% $\rightarrow$ 90.65%).

| Actual \ Predicted | 3 | 5 | 8 | 9 | 12 | 13 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|
| subject3 | 100.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| subject5 | 0.00% | 97.14% | 0.00% | 1.43% | 0.00% | 0.00% | 0.00% | 1.43% |
| subject8 | 0.00% | 0.00% | 100.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| subject9 | 0.00% | 0.00% | 0.00% | 98.33% | 0.00% | 0.00% | 1.67% | 0.00% |
| subject12 | 2.00% | 0.00% | 0.00% | 2.00% | 90.00% | 0.00% | 4.00% | 2.00% |
| subject13 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100.00% | 0.00% | 0.00% |
| subject17 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100.00% | 0.00% |
| subject18 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 4.00% | 96.00% |

Figure 5.5: Confusion matrix of different day with adaptive classifier

Table 5.2: Accuracy rate, error rate and results of iterative verification

| Subject | 3 | 5 | 8 | 9 | 12 | 13 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|
| Accuracy rate (%) | 95 | 100 | 100 | 98.31 | 97.78 | 100 | 100 | 97.92 |
| Accepted/False classified | 0/1 | 0/0 | 0/0 | 0/2 | 0/0 | 0/0 | 1/5 | 0/2 |
| Error rate (%) | 0 | - | - | 0 | - | - | 20 | 0 |
| Corrected/False classified | 0/0 | 1/2 | 0/0 | 1/1 | 1/5 | 0/0 | 0/0 | 0/2 |
| Original accuracy (%) | 100 | 97.14 | 100 | 98.33 | 90 | 100 | 100 | 96 |
| New accuracy (%) | 100 | 98.57 | 100 | 100 | 92 | 100 | 100 | 96 |

### Robustness against impostor

A robust person identification system needs to prevent the invasion of imposter. In there we demonstrate the results of three impostors which have 148 EEG data. Fig. 5.6 shows that the impostors were classified as different in classification phase. The overall true rejection rate (TRR) of our system is 76.55% that is slightly lower than non-iterative verification (79.31%). It is high that means the data of impostors will be rejected after this person identification system.

| Actual \ Predicted | 11 | 15 | 16 | 17 |
|---|---|---|---|---|
| **impostor1** | 6.45% | 48.39% | 3.23% | 41.94% |

| Actual \ Predicted | 13 | 16 | 18 |
|---|---|---|---|
| **impostor2** | 4.55% | 78.79% | 16.67% |

| Actual \ Predicted | 1 | 3 | 11 | 13 | 15 | 17 | 18 |
|---|---|---|---|---|---|---|---|
| **impostor3** | 9.80% | 1.96% | 5.88% | 5.88% | 68.63% | 5.88% | 1.96% |

Figure 5.6: Confusion matrix of impostors

### Influence of ERP signal

The present studies of EEG-based person identification mostly use alpha band which evoked by eye closing and VEPs by visual stimuli. Our experiment designed for acquiring VEPs like the present studies while the acquired data contain P300 by the proportion of different stimulus. We compared the classification of small and large disk trials which correspond to pure VEPs and ERPs separately. For the raw data of different trial shown in Fig. 5.7 we observed that the signal of small disk (left) is dull and has no significant potential in the period of 300 ms compared to the large disk (right).

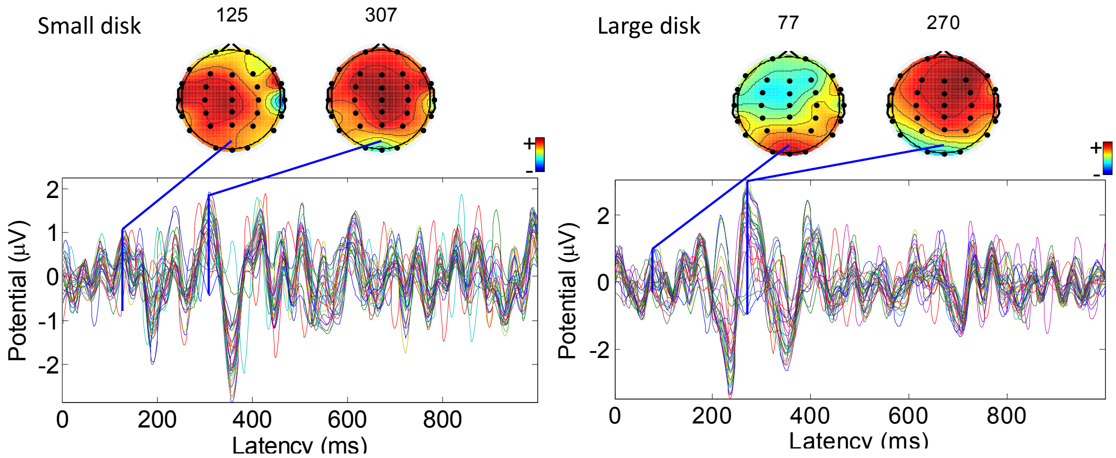


Figure 5.7: The raw data of small and large disk trials from subject1.

For the classification, we chose the same number of trials as the large one to train the classifier that the raw data of small disk obtained 71.31% accuracy which used the SVM classifier and average data with the combined feature of power spectrum and LLE obtained 80.20% accuracy. The classification accuracies are less than big disk (80.82% and 96.36% displayed in Table 4.1 and Table 4.2).

# Chapter 6

# Conclusion

We have proposed a person identification system using visual-evoked EEG signals. In experiments we used the visual stimuli to evoke VEP and ERP (P300) through the oddball paradigm. The methods of feature extraction include dimension reduction and time and frequency domain analysis. According to our experimental results, we concluded that the combination of power spectrum and LLE can extract informative features for distinguishing subjects. The person identification system contains the classification and verification phases. In the classification phase, we use a multi-class classifier to perform a one-to-many comparison for each acquired data. In the verification phase, the best-matching candidates are furthered verified sequentially by binary classifiers according to their matching levels. Moreover, we tried to correct those misclassified data through iterative verification. The results revealed that EEG data with individual differences can reach a high accuracy in person identification. Combining classification with verification the reliability of the system could be increased. In the discussion, compared with the inter-subject correlation the intra-subject correlation between EEG trials acquired at different times is higher. This result illustrate that EEG is beneficial used as a biometric.

Although there are some drawbacks in the practical application aspect, such as long preparation time required and the signal is susceptible to environmental impact. We can improve the situation by the advance in hardware facilities (e.g. dry electrode).

# Bibliography

[1] Vahid Abootalebi, Mohammad Hassan Moradi, and Mohammad Ali Khalilzadeh. A new approach for eeg feature extraction in p300-based lie detection. *Computer Methods and Programs in Biomedicine*, 94(1):48–57, 2009.

[2] I. Berlad and H. Pratt. P300 in response to the subject's own name. *Electroencephalogr Clin Neurophysiol*, 96(5):472–4, 1995.

[3] Chih-Chung Chang and Chih-Jen Lin. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3):1–27, 2011.

[4] A. Delorme and S. Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, 2004.

[5] E. Donchin and M. G. H. Coles. Is the p300 component a manifestation of context updating. *Behavioral and Brain Sciences*, 11(3):357–374, 1988.

[6] S. M. Fleming, R. B. Mars, T. E. Gladwin, and P. Haggard. When the brain changes its mind: Flexibility of action selection in instructed and free choices. *Cerebral Cortex*, 19(10):2352–2360, 2009.

[7] A. S. Gevins, C. L. Yeager, S. L. Diamond, J. Spire, G. M. Zeitlin, and A. H. Gevins. Automated analysis of the electrical activity of the human brain (eeg): A progress report. *Proceedings of the Ieee*, 63(10):1382–1399, 1975.

[8] Heather M. Gray, Nalini Ambady, William T. Lowenthal, and Patricia Deldin. P300

as an index of attention to self-relevant stimuli. *Journal of Experimental Social Psychology*, 40(2):216–224, 2004.

[9] Richard W. Homan, John Herman, and Phillip Purdy. Cerebral location of international 10-20 system electrode placement. *Electroencephalography and Clinical Neurophysiology*, 66(4):376–382, 1987.

[10] H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6):417–441, 1933.

[11] Guyon Isabelle, Andr, and Elisseeff. An introduction to variable and feature selection. *J. Mach. Learn. Res.*, 3:1157–1182, 2003.

[12] A. K. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. *Ieee Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, 2004.

[13] M. Liotti, S. R. Pliszka, R. Perez, D. Kothmann, and M. G. Woldorff. Abnormal brain activity related to performance monitoring and error detection in children with adhd. *Cortex*, 41(3):377–88, 2005.

[14] S. Marcel and J. D. R. Millan. Person authentication using brainwaves (eeg) and maximum a posteriori model adaptation. *Ieee Transactions on Pattern Analysis and Machine Intelligence*, 29(4):743–748, 2007.

[15] M. Miyakoshi, M. Nomura, and H. Ohira. An erp study on self-relevant object recognition. *Brain Cogn*, 63(2):182–9, 2007.

[16] C. Miyamoto, S. Baba, and I. Nakanishi. Biometric person authentication using new spectral features of electroencephalogram (eeg). *2008 International Symposium on Intelligent Signal Processing and Communications Systems (Ispacs 2008)*, pages 130–133 476, 2008.

[17] G. Moharnmadi, P. Shoushtari, B. M. Ardekani, and M. B. Shamsollahi. Person identification by using ar model for eeg signals. *Proceedings of World Academy of Science, Engineering and Technology, Vol 11*, 11:281–285 295, 2006.

[18] M. Napflin, M. Wildi, and J. Sarnthein. Test-retest reliability of eeg spectra during a working memory task. *Neuroimage*, 43(4):687–93, 2008.

[19] R. Palaniappan and K. V. R. Ravi. A new method to identify individuals using signals from the brain. volume 3, pages 1442–1445 vol.3, 2003.

[20] M. Poulos, M. Rangoussi, N. Alexandris, and A. Evangelou. Person identification from the eeg using nonlinear signal classification. *Methods of Information in Medicine*, 41(1):64–75, 2002.

[21] M. Poulos, M. Rangoussi, V. Chrissikopoulos, and A. Evangelou. Parametric person identification from the eeg using computational geometry. In *Electronics, Circuits and Systems, 1999. Proceedings of ICECS '99. The 6th IEEE International Conference on*, volume 2, pages 1005–1008 vol.2, 1999.

[22] M. Poulos, M. Rangoussi, V. Chrissikopoulos, and A. Evangelou. Person identification based on parametric processing of the eeg. In *Electronics, Circuits and Systems, 1999. Proceedings of ICECS '99. The 6th IEEE International Conference on*, volume 1, pages 283–286 vol.1, 1999.

[23] Henry Taylor Fowkes Rhodes. *Alphonse Bertillon, father of scientific detection*. Harrap, London,, 1956.

[24] A. Riera, A. Soria-Frisch, M. Caparrini, C. Grau, and G. Ruffini. Unobtrusive biometric system based on electroencephalogram analysis. *Eurasip Journal on Advances in Signal Processing*, pages –, 2008.

[25] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–6, 2000.

[26] G. K. Singhal and P. RamKumar. Person identification using evoked potentials and peak matching. *2007 Biometrics Symposium*, pages 156–161 167, 2007.

[27] Joan G. Snodgrass and Mary Vanderwart. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2):174–215, 1980.