PERSONAL IDENTIFICATION BY RECOGNITION OF EEG POWER SPECTROGRAM HAVING SHORT PROCESSING TIME

Mr. Chesada Kaewwit

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Computer Science and
Information Technology
Department of Mathematics and Computer Science
Faculty of Science
Chulalongkorn University
Academic Year 2015

การระบุบุคคลโดยการรู้จำของอีอีจีเพาเวอร์สเปกโทรแกรมโดยใช้เวลาสั้นในการประมวลผล

นายเจษฎา แก้ววิทย์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2558

Thesis Title	PERSONAL IDENTIFICATION BY RECOGNITION OF EEG POWER SPECTROGRAM HAVING SHORT PROCESSING TIME
By	Mr. Chesada Kaewwit
Field of Study	Computer Science and Information Technology
Thesis Advisor	Associate Professor Peraphon Sophatsathit, Ph.D.
Thesis Co-Advisor	Professor Chidchanok Lursinsap, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

....Dean of the Faculty of Science

(Associate Professor Polkit Sangvanich, Ph.D.)

THESIS COMMITTEE

....Chairman

(Saichon Jaiyen, Ph.D.)

....Thesis Advisor

(Associate Professor Peraphon Sophatsathit, Ph.D.)

....Thesis Co-Advisor

(Professor Chidchanok Lursinsap, Ph.D.)

....Examiner

(Assistant Professor Suphakant Phimoltares, Ph.D.)

....Examiner

(Assistant Professor Saranya Maneeroj, Ph.D.)

....External Examiner

(Associate Professor Sartra Wongthanavasu, Ph.D.)

เจษฎา แก้ววิทย์ : การระบุบุคคลโดยการรู้จำของอีอีจีเพาเวอร์สเปกโทรแกรมโดยใช้เวลาสั้นในการประมวลผล (PERSONAL IDENTIFICATION BY RECOGNITION OF EEG POWER SPECTROGRAM HAVING SHORT PROCESSING TIME) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: รศ. ดร. พีระพนธ์ โสพัศสถิตย์, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม: ศ. ดร. ชิดชนก เหลือสินทรัพย์, 70 หน้า.

การวัดสัญญาณกระแสไฟฟ้าในสมองหรือ อีอีจี ได้ถูกนำมาใช้ในด้านไบโอเมตริกซ์ เนื่องจากตัวสัญญาณมีคุณสมบัติที่ยากในการสร้างหรือปลอมแปลงได้ ในงานวิจัยนี้ได้นำสัญญาณอีอีจี มาใช้จำแนกบุคคล และเสนอการสกัดลักษณะโดยเทคนิคเพาเวอร์สเปกโตรแกรมร่วมกับเอสวีดี เพื่อให้ได้เวลาในการจำแนกบุคคลได้เร็วที่สุดและช่องสัญญาณน้อยที่สุด โดยใช้เวลาต่ำกว่า 1 วินาที ในการจำแนกบุคคลได้แบ่งกลุ่มภายในและภายนอกโดยผลของการจำแนก มีความถูกต้อง 100 เปอร์เซ็นต์ ทั้งบุคคลภายใน และบุคคลภายนอก นอกจากนี้ยังแสดงให้เห็นว่าพื้นที่ในส่วนของสมองส่วนหน้า เป็นส่วนที่สำคัญในการจำแนกบุคคล

ภาควิชา	คณิตศาสตร์และวิทยาการคอมพิวเตอร์	ลายมือชื่อนิสิต ....
		ลายมือชื่อ อ.ที่ปรึกษาหลัก ....
สาขาวิชา	วิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ	ลายมือชื่อ อ.ที่ปรึกษาร่วม ....
ปีการศึกษา	2558	

5373899623 : MAJOR COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
KEYWORDS: ELECTROENCEPHALOGRAPHY (EEG) / BIOMETRICS / POWER SPECTROGRAM / SINGULAR DECOMPOSITION VALUE / PERSON AUTHENTICATION

CHESADA KAEWWIT: PERSONAL IDENTIFICATION BY RECOGNITION OF EEG POWER SPECTROGRAM HAVING SHORT PROCESSING TIME. ADVISOR: ASSOC. PROF. PERAPHON SOPHATSATHIT, Ph.D., CO-ADVISOR: PROF. CHIDCHANOK LURSINSAP, Ph.D., 70 pp.

Electroencephalography (EEG) is the electrical measurement of brain activity. EEG is an interesting signal in biometrics on account of individuality and arduous to imitate the signal patterns. This research concentrates on person identification using power spectrogram without composite EEG cleaning. The underlying technique encompasses feature extraction by using spectrogram to determine dominant values from spectrogram with Singular Decomposition Value (SVD). The proposed method can separate a person from an outsider group yielding the highest accuracy at 100 percent classification in less than 1 second. The outcomes also reveal that the frontal area is an important location of scalp electrode placement in person identification.

Department:	Mathematics and Computer Science	Student's Signature	....
		Advisor's Signature	....
Field of Study:	Computer Science and Information Technology	Co-Advisor's Signature	....
Academic Year:	2015		

ACKNOWLEDGEMENTS

The author would like to thank Dr. Preecha Tangkraingkij at Sripathum University and Tayard Desudchit, M.D. at The Chulalongkorn Comprehensive Epilepsy Program (CCEP), Faculty of Medicine, Chulalongkorn University for the EEG data used in this research.

CONTENTS

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

The general applications of biometric functions can be categorized into two types, namely, identification and authentication. The former is a verification and validation of an individual person from a database or a group of persons. The latter is an acceptation of individual person. Biometrics is a person authentication and identification research area related to several real organs. Typically, it can be divided into 2 groups, namely, physiology and behavior of human beings. Examples of physiology group are fingerprints, face shape, hand geometry, ear shape, DNA, and eye irises. Examples of behavior group are gait, keystroke, running patterns, speech, and signature. Recently, both classes of biometrics are likely to be imitated due to medical advances and development of information technology such as plastic surgery, high resolution devices, and high technology digital tools. Such operations undermine the biometric identification process. Meanwhile, physiological signals are considered as the other biometric trait since it is generated from human body. The signal from each person is distinct and arduous to imitate. Thus, it is a good candidate for biometric trait. One interesting signal is brain wave. Currently, there are several noninvasive measures or capturing the brain activities that produce different types of signals. Magnetoencephalography (MEG) records magnetic signals of the brain activity. A function Magnetic Resonance Imaging (fMRI) is used for capturing brain metabolic activity. An electroencephalography (EEG) is a medical method that records scalp electrical activity of millions neurons from the same position generated by human brain [1]. Both MEG and fMRI require complex machines to measure and take a long time of recording which make them unsuitable for real time processing. Fortunately, EEG signal can be efficiently applied to daily life research and application such as neuroPhone [2], neural sensing healthcare [3], and monitoring driving car [4]. It has been said that EEG signals can be applied to many sectors such as medical center, online security, biometric mobile, and so on.

Due to noninvasive, inexpensive, and portable EEG device, many researches focus on using EEG as biometrics such as Poulos [5] using Autoregressive (AR) model and computational geometry method for feature extraction and Learning Vector Quantization (LVQ) as the main classifiers. Paranjape et al. [6] used AR and variance/covariance matrix with statistical tools for modeling of EEG signal from a single channel.

Additionally, EEG is able to diagnose the function of the brain. The functional area in human brain can be categorized into four areas, namely, frontal, parietal, temporal, and occipital lobe. The frontal lobe is responsible for conscious thoughts. The parietal lobe integrates sensory information from various senses. The temporal lobe performs auditory sense, while sense of sight is processed by the occipital lobe. Moreover, EEG can be grouped into five different rhythms based on their frequencies: Delta rhythm (1-4 Hz) is seen during deep sleep in adults and in infants as an unusual activity; Theta rhythm (4-8 Hz) occurs in drowsiness in adults and during waking up of children; Alpha rhythm (8-12 Hz) is seen normally during relaxed with eye closed; Beta rhythm (12-30 Hz) is associated with anxious thinking and active concentration; and Gamma rhythm (30-100 Hz) is associated with certain cognitive senses. There have been many researchers focusing on person identification based on EEG with grouped of rhythms. Palaniappan's research [7] [8] used energy feature of gamma rhythm within Visual Evoked Potential (VEP) to identify individuals, K-Nearest Neighbors, and Elman Neural Network (ENN) as classifiers. Rocca et al. [9] used three electrodes with sub-band of EEG signal while the subjects were in resting and closed eyes. AR stochastic and polynomial approach were used for feature extraction.

There are many feature extraction methods of EEG proposed in biometrics field. Most methods required long data length and involved sophisticated techniques. Therefore, those methods are time-consuming and impractical for real-time security applications. The issue is a significant problem which is essential to be further explored. Since the minimum identification time is vital to system performance, this study proposes an EEG biometric method for real life and real time application by time and frequency techniques called power spectrogram and Singular Decomposition Value

(SVD) to extract main features from the EEG signal. The next problem is how many channels are the highest accuracy of identification? This study will determine the minimized EEG data length and fewer channels to identify subjects by practical implementation, such as minimized data length. Afterward, the shorter minimized channels will be revealed by brute force experiment to obtain the significant EEG channels for biometric. Some subject data sets are conducted in resting state because subjects are not engaged in any activities which affect the EEG signal detection [13]. It has been estimates that the EEG signals are rather smooth with less artifacts. To confirm the experimental results, the experimental data length and channels are performed on unknown outsiders with the proposed classification algorithm.

Objectives

The objectives of this dissertation are to develop techniques:

1. To classify an individual with a minimum EEG identification time.
2. To reduce the number of location of scalp electrodes or channel and use only the necessary channels for personal identification.
3. To identify insiders and an outsiders by using EEG power spectrogram.

Scope of Work

1. The feature extraction of EEG signal using power spectrogram and SVD is proposed to obtain singular value of the signal. These features are used to classify a person with minimum identification time.
2. The experiments with combining channels are implemented to minimize the significant location of scalp electrodes or channels leading to accurate identification. Moreover, comparison with Preecha's work [10] is performed.
3. The limitations of this work are (1) the data set obtained from the research data set of Preecha's work [10]. There were 40 subjects with 16 EEG channels. The subjects were in resting state with opened eyes. They were asked not to perform any task or movement during recording EEG signal. Processed 3,000 data points (15 seconds recording) of each channel are selected. (2) a published

data set [11, 12] which consists of 109 subjects performing in resting state 1 minute with opened eyes, 1 minute closed eyes, three 2 minutes running of 4 tasks (2 motor and 2 motor-imagery), 64 EEG channels, international 10-10 system, are recorded at a sampling rate of 160 Hz.

Research Contribution

The significant outcomes of this research that are conducive toward the area of person identification are as follows: (1) application of EEG to correctly identify a person, (2) important location of scalp to the application of scalp to the application, and (3) minimum identification time obtained by the proposed algorithm.

Dissertation Organization

This dissertation is organized as follows. Chapter 2 discusses the literature review. Some general backgrounds are described in Chapter 3. Chapter 4 proposes the methodology. The experiment and results are given in Chapter 5. Some important findings are discussed in Chapter 6, while the final remarks and future work are concluded in Chapter 7.

CHAPTER 2

LITERATURE REVIEW

There are studies involving EEG signal in person identification. Subasi and Gursoy [14] used signal processing and EEG analysis to predict whether each patient had epileptic seizure or not. The signals were divided into the frequency sub-bands using DWT and a set of statistical features was extracted from the sub-bands to represent the distribution of wavelet coefficients. Principal Components Analysis (PCA), Independent Components Analysis (ICA), and Linear Discriminant Analysis (LDA) were used to reduce the dimension of data. Then these features were used as inputs to a Support Vector Machine (SVM) with two discrete outputs: epileptic seizure or not.

Muthuswamya and Thakor [15] reviewed and revealed some drawbacks and limitations of EEG analysis techniques such as Fast Fourier Transform, Autoregressive, and Wavelet. They used EEG signals recorded in animals during hypoxic-asphyxic injury to brain for classification.

Mustafa et al. [16] used time-frequency method to analyze EEG signal. Generally, EEG signal is analyzed by many methods such as time based, frequency based, time-frequency based, and wavelet. They used Gray Level Co-occurrence Matrix to extract the texture features from the spectrogram image. Then, k-NN and PCA were used to classify spectrogram image with recognition accuracy 70.83%.

For security approach, there are many works using EEG as biometric. This is a novel and challenging study that integrates many fields to achieve high identification accuracy. Different kinds of feature extractions and classification approaches were revealed. For example, the study of Marcos et al. [17] reviewed EEG subject identification. There were various feature extraction and classification techniques that yielded high accuracy. They contended that EEG biometrics research encompassed variables that affected the accuracy such as time, frequency, space, recording, and algorithms.

Poulos et al. [5] presented EEG signals as biometrics using AR and computational geometry methods for feature extraction. Classification of accuracy percentage ranged from 80 to 95 with 4 subjects. The work of Poulos et al. [18] improved accuracy of person identification with respect to non-linear model and Learning Vector Quantization (LVQ) classifier. The accuracy percentage increased to 99.5 with 4 subjects.

Palaniappan's research [7] [8] [8] used energy feature of gamma band within Visual Evoked Potential (VEP) to identify individuals. They used a large group of 102 subjects and a high number of VEP signals. They chose k-Nearest Neighbors and Elman Neural Network (ENN) as classifiers. The results showed that the maximum accuracy percentage of ENN and k-NN was about 98.12 and 96.13. For this reason, they indicated that the significant potential of brain electrical activity could be considered as biometrics from the experimental results.

Paranjap et al. [6] used AR and variance/covariance matrix with statistical tools for modeling of EEG signal from a single channel. There were 40 subjects in resting state with eyes opened and eyes closed. The accuracy percentage was 80.

Marcel et al. [19] employed Surface Laplacian, Power Spectral Density (PSD), and statistical tools to extract EEG signals as significant features. In their paper, there were 9 subjects with HTER, ((False Acceptance + False Rejection)/2), 19.3-42.6 classification accuracy.

Shedeed [20] applied Discrete Fourier Transform and Wavelet packet decomposition for feature extraction and applied Multi-layer perceptron to classify each subject with 4 channels C3, C4, P3, and P4. There were 3 subjects in resting and eye closed in the experiment. The result reached 100 percent.

Cempirek and Stastny [21] used sub-band of EEG frequency transformed into spectra by using Fourier Transform. They proposed LVQ neural network to classify the EEG spectra on 8 subjects. The subjects sat and closed eyes in the experiment. The results showed that EEG spectra analysis based on Euclidean distance was unable to classify some subjects. In addition, an average accuracy was about 80 percent. Besides, the analysis of the segment length influenced on the subject identification success. They revealed that they could lower the frequency resolution to 1/90 Hz or 1/126 Hz for higher classification score without significant impact on the classification. However, the frequency resolution was not suitable for daily life applications.

The research of Rocca et al. [9] used three electrodes with sub-band of EEG signal to classify 45 persons in resting and closed eyes. Recognition rate was about 98.73 percentage. Autoregressive stochastic and polynomial approach were used for feature extraction.

Yazdani et al. [22] employed AR and the peak of power spectral density (PSD) for feature extraction and used LDA for dimension reduction. The k-NN was used as a classifier. The accuracy percentage reached 100 with 20 subjects.

Riera et al. [23] introduced feature extraction using five models combinations: AR, Fourier Transform, Mutual Information, Coherence, and Cross Correlation. Fisher's Discriminant Analysis was applied as the classifier of this study. There were 51 subjects and 36 intruders. The accuracy percentage was between 87.5 and 98.1.

Abdullah [24] reported features to be extracted using AR model to obtain the feature set. Results showed that data from eyes open and eyes closed using 4 channels gave good classification rates of 96% and 97%, respectively. The above studies can be summarized in Table 1.

Table 1. Summarized related works.

Research	Feature	#Subjects	% Accuracy	Classifier
Poulos	AR, geometry	4	80-95	LVQ
Paranjap	AR,VAR/COV	40	80	Statistics
Palaniappan	AR, gamma	102,20,40	95-99	k-NN, ENN, LVQ
Marcel	PSD	9	(HTER)19.3-42.6	GMM
Yazdani	AR PSD	20	100	LDA
Riera	AR+PSD+MI+COH+ correlation	51	87.5-98.1	Linear Classifier
Abdullah	AR	10	96-97	ANN
Shedeed	DFT WPD	3	100	ANN
Rocca	AR	45	98.73	Polynomial
Maiorana	AR	50	>90	k-NN

Table 1 shows that most feature extraction methods of EEG signals are based on signal processing methods such as AR, PSD, and DFT. Each classification approach depends on properties of extracted features.

The study of DelPozo et al. [25] applied Short Time Fourier Transform (STFT) to 6 publicly available EEG databases to obtain high identification accuracy. The maximum number of subjects in that experiment was 20. The accuracy was from 92.5 to 95 percent at 4 to 6 seconds. Rocca et al [26] proposed the spectral coherence-based connectivity as feature extraction of EEG signals for person identification. The work used the data set from EEG Motor Movement/Imagery Dataset [12]. There were 108 subjects with opened and closed eyes in resting state. They divided the brain into 3 regions. The electrode pair in each region matched score fusion according to a forward/backward to improve the identification accuracy. The accuracy could reach 100 percent in frontal lobe. However, the main design problem of the biometric system was due to many electrodes to be placed on the scalp.

Maiorana's work [27] used AR model with reflection coefficient and k-NN as classifier with 50 subjects. The EEG data of each subject were recorded in resting state with longitudinal data. The results showed that AR method gave a discriminating capability higher than using PSD and COH as feature extraction methods. The identification time for a subject was less than a minute.

Preecha's works [28] [29] applied Independent Component Analysis (ICA) for signal cleaning and multilayer perceptron neural networks for signal classification. From the experimental results, 4 channels F7, C3, P3, and O1 could identify a group of 20 users with high accuracy. The insider and outsider person identification of Preecha’s et al [10] experiment revealed that position P4 was the significant location for identification. The identification time for a subject was 5 seconds.

However, shorter identification time of EEG are rarely published. It is an important point of security view if EEG data are applied to the real applications. It can be summarized that many researches involving EEG biometrics and EEG classification use complex method that are time-consuming for identification which is hard to implement for practical real time security applications. This study develops an EEG biometric method for real life and real time application by using time and frequency techniques called Short Time Fourier Transform (STFT) or spectrogram and Singular Decomposition Value (SVD) to extract EEG signals. This research focuses on person identification using power spectrogram. Furthermore, classification techniques are experimented to achieve high accuracy and to discover a proper classifier. The results are minimized EEG data signals length and shorter minimum channel combinations to correctly identify the subjects.

CHAPTER 3

GENERAL BACKGROUND

3.1 Short Time Fourier Transform (STFT)

A signal can be expressed in time domain. However, the frequency domain of the signal obtains significant frequency information providing available signal analysis. Fourier Transform is a method transforming time domain into frequency domain of signals to obtain the magnitude of frequency components. The Fourier Transform normally performs based on Discrete Fourier Transform (DFT) [30]. It can be defined as in Equation 1.

$$X(i)=\sum_{k=0}^{N-1} x(k)e^{-j2\pi/N} \tag{1}$$

where $x(k)$ denotes signal with N points, $X(i)$ denotes the discrete Fourier of $x(k)$, and $X(i)$ = $DFT(x(k))$. A constraint on Fourier Transform method is unsuitable for non-stationary signal analysis owing to the uncertainty of signal characteristics. Examples of non-stationary signal are speech, sound or music, sonar signals, EEG, and so on. Hence, Short Term Fourier Transform (STFT) is proposed to deal with the limitations of the signals. STFT is able to decompose non-stationary signal to time and frequency analysis producing a matrix called spectrogram. It has been considered an efficient method for appraising non-stationary signals. The STFT can be expressed in Equation 2.

$$\mathbf{G}(m,i)=|DFT(w(k)x_m(k))|^2, 0\leq m<2s-1, 0\leq i<l \tag{2}$$

Let $x(k)$ be an N point signal decomposed into 2s-1 overlapping sub-signals $x_m(k)$ of length l. That is $N = l*s$ for integer l and s, where $w(k)$ denotes a window function. **G**(m,i) is spectrogram of $x(k)$, m is time interval of each sub-signal, and i is frequency.

The window function is used for the frequency domain leakage reduction. The popular window functions are Hanning, Hamming, and Blackman. This work uses the Hamming window.

$$w(k) = [.54 - .46cos(\frac{2\pi k}{l})] \tag{3}$$

Normally, $w(k)$ is shorter than the length of data point signals. Various lengths of window function cause wide-band and narrow-band of signals. The smaller the window function length or wide-band is, the better time resolution is recognized. Meanwhile, STFT is calculated by using Discrete Fourier Transform on the input data multiplied by the length of window function. This computational complexity is $O(l^2)$ where l is the data points signals [31]. STFT is performed based on Fast Fourier Transform (FFT) algorithm because the number of data points and the window function length are defined in power of two. Thus, its computational complexity is $O(l\log l)$ [32]. The magnitude squared of STFT produces the power spectrogram.

3.2 Singular Value Decomposition (SVD)

SVD method is used to decompose a matrix to obtain the dominant value matrices consisting of three related matrices. The first matrix denotes columns of the matrix as singular vector. The second matrix holds singular values diagonally arranged in descending order. The last matrix denotes the rows of the matrix as the singular vector. The equation of SVD can be defined in Equation 4.

$$\mathbf{S} = \mathbf{U}\mathbf{X}\mathbf{V}^T \tag{4}$$

where $\mathbf{U} \in \mathbb{R}^{p \times p}$ and $\mathbf{V} \in \mathbb{R}^{q \times q}$ denote squared matrix. Both are mutually orthogonal property. $\mathbf{X} \in \mathbb{R}^{p \times q}$ is a diagonal matrix comprising singular values. The singular values represent the significant values of the matrix. The computational time of SVD is usually $O(pq \cdot \min(p,q))$ or $O(pq^2)$ [33].

3.3 Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is a particular technique of blind source separation (BSS) which divides multivariate signals into individual source signals or components that are statistically independent. ICA can be explained as in Equation 5.

$$\mathbf{O} = \mathbf{A}\mathbf{B} \tag{5}$$

where $\mathbf{O}$ denotes mixture of the source signals $\mathbf{B}$, $\mathbf{A}$ is the unknown mixture matrix. Let $\mathbf{O}$ be the EEG estimation from the electrodes on the scalp, $\mathbf{O}$ = [O_1(t)... O_n(t)]^T. Let $\mathbf{B}$ = [B_1(t) ... B_n(t)] be source EEG signals. An estimated source EEG signals can be expressed as

$$\mathbf{Y} = \mathbf{D}\mathbf{O} \tag{6}$$

where $\mathbf{Y}$ = [Y_1(t)... Y_n(t)]^T is an estimated source signals of $\mathbf{B}$. $\mathbf{O}$ denotes the inverse matrix of mixture matrix $\mathbf{A}$. There are many algorithms which estimate accurate value of matrix $\mathbf{D}$ such as JADE[34], SOBI [35] [36], SOBIRO [37], FastICA [38], ERICA [39] and so on. In this work, SOBIRO is selected for the ICA process.

3.4 Autogressive Model (AR)

The AR model of order z can be defined as in Equation 7.

$$X_l = \sum_{i=1}^{z} \varphi X_{(l-i)} + \varepsilon_l \tag{7}$$

where X_l is the signal at sample point l, φ is the real valued AR coefficient, and ε_l is white noise. The coefficients, φ, can be evaluated by Yule-Walker and can be solved repeatedly by performing the Levinson method to obtain the reflection coefficients of AR model. However, Burg algorithm can compute the reflection coefficients directly from the signal X_l [40].

3.5 Classification Approaches and Validation

To obtain the prediction accuracy of person identification, two classification approaches are used in this research, namely, Artificial Neural Networks (ANN) and k-Nearest Neighbor (k-NN).

3.5.1 Artificial Neural Networks (ANN)

ANN is a simple abstraction of biological neuron in the human brain. It consists of input vector of pattern ψ, weight of each input element or link of input vector W_j, activation function $g()$. Let $\mathbf{a}^{\psi}$ = [$\mathbf{a}_1^{\psi} \ldots \mathbf{a}_n^{\psi}$]^T be an input vector of pattern ψ. The class $\mathbf{b}^{\psi}$ of the input vector can be expressed as

$$\mathbf{b}^{\psi} = g(\sum w_j \Phi_j(\mathbf{a}^{\psi})) \tag{8}$$

where $\Phi_j(\bullet)$ is the j^{th} kernel function. w_j is a weight of link j. $g(\bullet)$ is the activation function of output. There are many kernel functions and many adjusted weight methods for ANN. In this work, three kernel functions, e.g., Multilayer Perceptron, Radial

Basis Function Network, and Probabilistic Neural Networks are selected to be candidate classifiers.

Multilayer Perceptron (MLP)

MLP is one of neural network approaches for solving classification issues. This method consists of 3 layers, namely, an input layer, an output layer, and a hidden layer. The input layer comprises the number of neurons equals to the number of features of input vector. The output layer has the same number of neurons as the output vectors or the classes. The hidden layer is able to define the number of neurons by trial and error. Each neuron unit of the hidden layer is a weight coefficient which can be conducted from its input vectors. Each input vector is multiplied by a weight matrix and is aggregated together to determine the output vector. Moreover, training of neural network leads to adjusting a weight value to minimize errors of the output. These methods are called back-propagation (BP) algorithm. There are many back-propagation methods to speed up training and increase accuracy. This work focuses on Levenberg Marquardt (LM) and Bayesian Regularization (BR) approaches on account of different functions of the MLP-BP.

Radial Basis Function Network (RBFN)

RBFN is one of neural network approaches like MLP. Generally, RBFN approach consists of three-layers, namely, an input layer, hidden layer, and output layer. The hidden layer nodes are determined by a parameter vector and a scalar, applying Gaussian density function as an activation function. The weight matrix is between the hidden and output layers. RBFN performs classification based on Euclidean distance between the test data set and trained data set.

Probabilistic Neural Network (PNN)

PNN is a radial basis network which is appropriate for classification problem. PNN consists of four layers such as input layer, pattern layer, summation layer, and decision layer. The input layer function will receive inputs and send them to the neurons in the pattern layer to compute the output. The summation layer neurons compute the maximum output from the previous layers to the decision layer, resulting in either correct or incorrect decision. If the PNN classifier has a problem for output decision, Bayes' decision will be performed [41].

3.5.2 k-Nearest Neighbor (k-NN)

k-Nearest Neighbor is a simple classifier with respect to calculating the distance to all trained data sets. Each test data is predicted as belonging to class which is the trained set having the shortest distance. Euclidean distance is applied. Computation of Euclidean distance can be expressed as

$$D_{Euc}(H,K) = \sqrt{\sum_{i=1}^{q_n}(H_i - K_i)^2} \tag{9}$$

where q_n is the number of features. *H* and *K* are patterns in specific data. There are many distance algorithms used by nearest neighbor such as Cosine, Hamming, Jaccard, Manhattan, Mahalanobis, and so on.

k-NN is a fundamental machine learning approach for object classification based on the shortest distance in the sample with k number of nearest patterns. Each object or class is classified by an absolute majority of its neighbors.

3.5.3 k-fold Cross Validation

k-fold cross-validation technique is an effective measurement of accurate predictive model based on repetitively executions. It is used for evaluating data to reliably generalize an independent data set. Data are randomly divided into k folds or data sets. The technique runs k times with defining index to k-1 folds as trained data set and the remaining fold as the test data set for each run. All the results from each fold are averaged to obtain the final result. For example, 10-fold cross validation is randomly divided into 10 folds. Nine of ten data are trained data and the remaining part is used as test data for calculating the mean accuracy. This process repeats 10 times until all data sets are trained and evaluated on ten different individual data sets.

CHAPTER 4

METHODOLOGY

The methodology is composed of 3 main sequential processes, namely, data segmentation, proposed feature extraction, and classification. The data segmentation process aims to separate the EEG signals to four data point lengths: 32, 64, 128, and 256. The segmented data point affects the identification time and consuming process. The proposed feature extraction is performed on the segmented data points by using STFT and SVD to extract dominant values of each segmented data point. The classification process is performed to classify individual person. However, the appropriate classifier of this work is experimented with 3,000 data points of 16 channels and 20 subjects. Then, the appropriate classifier is used to classify individual person to obtain the highest accuracy. Figure 1 highlights the input, the three processes, and the output of the methodology.

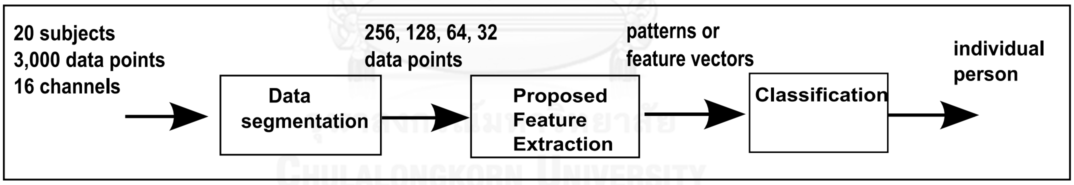

Figure 1. Methodology.

4.1 Data Segmentation

Initially, each EEG signal of each channel is divided equally into many fixed data point length in power of two, namely, 32, 64, 128 and 256. Furthermore, half of each data segment is overlapped to increase the number of data segments. Each fixed data point length is used to determine the minimum identification time. The time of each fixed data segment is shown in Table 2.

Table 2. The minimum time of each fixed data segment.

Data segment	256	128	64	32
Time (seconds)	1.28	0.64	0.32	0.16
Total Data Segments per channel	22	45	92	186

If the fixed data point length are 32, 64, 128 and 256, the number of segmented data per fixed length are 186, 92, 45, and 22 data segments per channel, respectively, as shown in Table 2. For extracted feature, each data segment is transformed into power spectrogram using STFT and SVD.

4.2 Proposed Feature Extraction

The proposed feature extraction process is shown in Figure2.

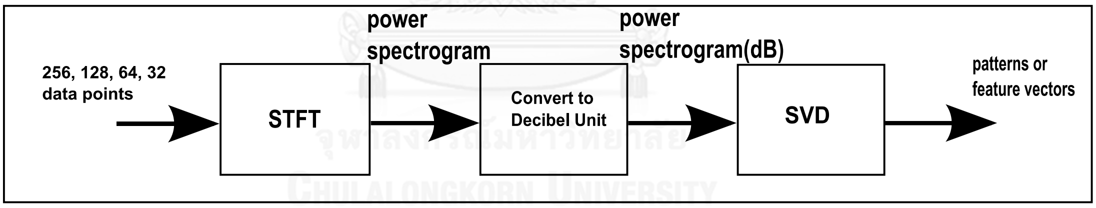

Figure 2. Proposed Feature Extraction.

The first step is to segment EEG data points by STFT and Hamming window with 50% overlapping. The result will return a matrix calculating from their magnitude. Each value of the matrix represents power intensity of the signal. The matrix is called power spectrogram. All values of power spectrogram are converted into decibel unit by multiplying with $10\log_{10}$. Finally, each power spectrogram is decomposed to take a singular value by SVD. The singular value of each channel is combined as a pattern or feature vector of each subject.

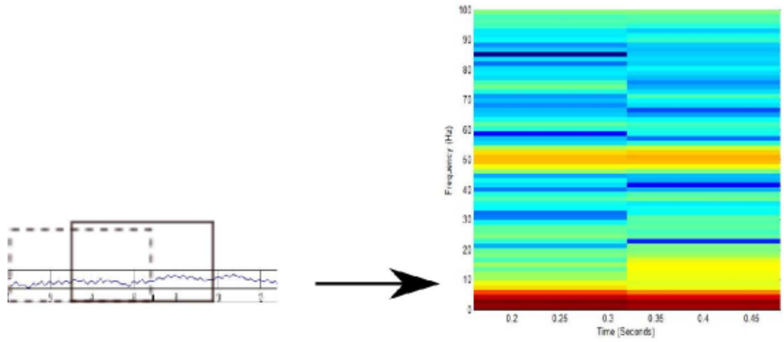

Figure 3. Transform EEG signal to power spectrogram.

From Figure 3, each segmented data is transformed into a power spectrogram matrix, each of which is a 2 dimension matrix.

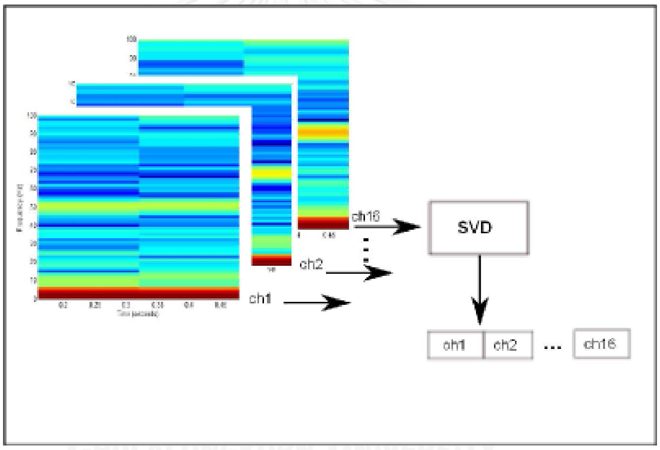

Figure 4. Decompose power spectrogram of each channel by SVD.

The matrix is extracted and converted to a row feature vector by using SVD. The number of features of each feature vector is equal to the number of columns of the power spectrogram matrix for each channel. Each feature vector of each channel is concatenated to form a sample of each subject. The process is illustrated in Figure 4. Subsequently, each feature vector of each channel is combined to form a row matrix which is pattern of each subject.

4.3 Classification and Evaluation

To predict the individual person, the classification approach is used to determine each class whose pattern belongs to high accuracy percentage. Each classifier is implemented by using 10-fold cross validation to increase the result reliability. The patterns of 20 subjects are split randomly into 10-fold cross validation. Trained data sets are allocated nine folds, while test data set takes up the remaining fold. The experiment runs 10 times. The result is shown in Figure 5.

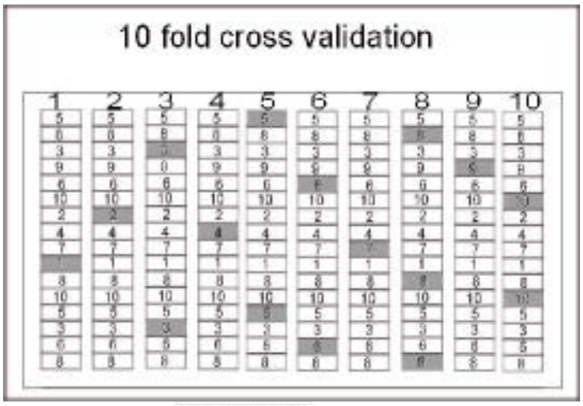

Figure 5. 10-fold cross validation.

In Figure 5, each fold cross validation has the same indexing. The highlighted box of each fold is the test pattern. The remaining patterns of each fold are train patterns. To measure performance of the experimental results, each fold cross validation runs the number of correct patterns [42] or the correct recognition rate (CRR) as defined in Equation 10 and 11.

$$CRR_f = \left(\frac{1}{N_{class}} \sum_{i=1}^{N_{class}} n_{(i)} \right) \times 100 \qquad (10)$$

where f denotes number of folds cross validation. N_{class} is the number of persons or classes. $n_{(i)}$ is the number of correct recognitions of each person or class, that is the number of the true positive of each class. Eventually, each CRR_f is averaged to CRR.

$$CRR = \frac{\sum_{f=1}^{10} CRR_f}{10} \tag{11}$$

This process selects five kinds of candidate classifiers using different classification methods, namely, probabilistic neural networks (PNN), Multilayer Perceptron (MLP), LM and BR, Radial Basis Function (RBF), and k-Nearest Neighbors (k-NN). The process is described in the next section.

4.4 Experimental Setup

First of all, appropriate classifier to achieve the high accuracy must be determined as illustrated in Figure 6.

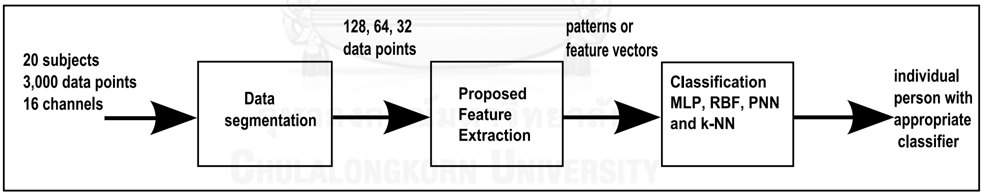

Figure 6. Experiment of the proposed feature extraction and classifier.

To find the minimum data points and shorter significant channels, the possible combinations of channels are deployed in experiment setup. Consequently, the experimental results are compared to using the methods described earlier in related work. This experiment is explained in Figure 7.

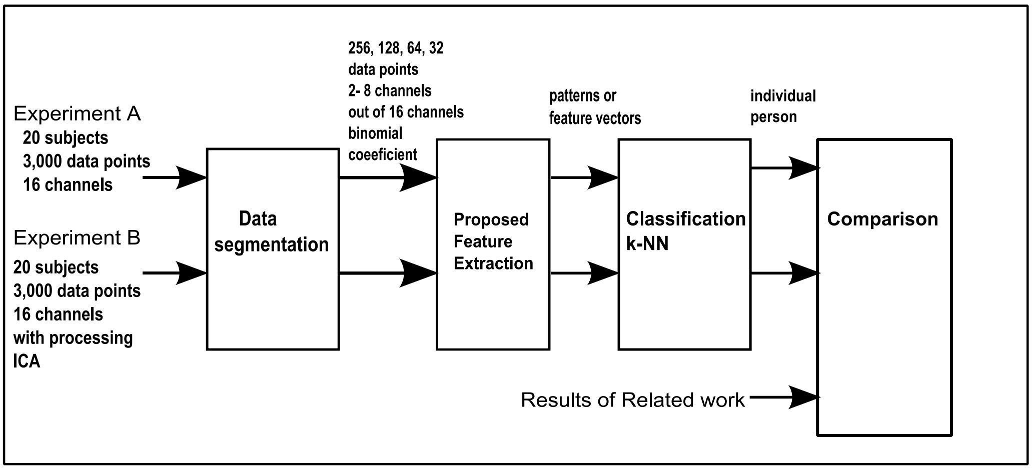

Figure 7. Experiment on finding minimum data length and significant channels in comparison to the related work.

It is obvious that the highest classification accuracy of minimum important channels with possible minimum identification time is obtained. To confirm the shorter important channels and minimum identification time, the experimental results are tested with the outsiders by the proposed algorithm in Figure 8.

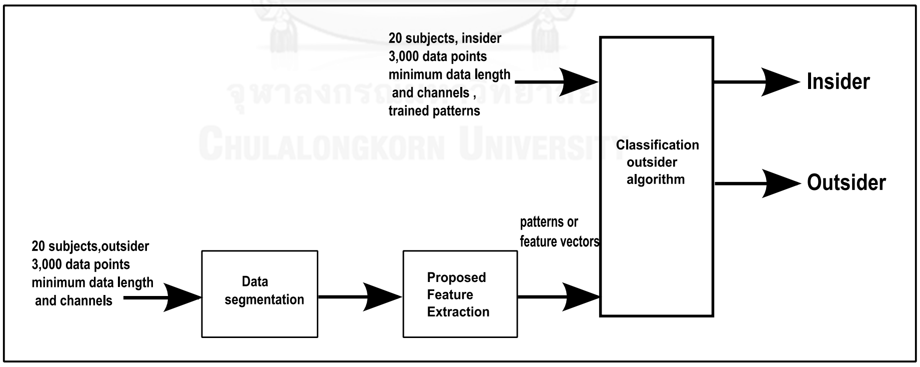

Figure 8. Experiment on separating a person from outsider with minimum time and channels.

The algorithm for classification between insider and outsider subjects

For determining the minimum number of channels and the minimum time for person identification, it is essential to classify the unknown person (outsider group) from the known person group (insider group). This experiment used the same data set as Preecha [10]. The data set consisting of 40 subjects was divided into 2 equal groups representing insider and outsider groups. The first 20 subjects are the insider group or known persons, and the remaining 20 subjects are the outsider group or unknown persons. The insider group is the trained set and the outsider group is the test set of this experiment.

The algorithm for insider and outsider subject classification is proposed as follows:

1. Finding and keeping three main values of each insider, mean, co-variance and max distance of each insider. The maximum distance, called a threshold value of each subject, is calculated by using mahalanobis distance. Let S_i^N be each sample of subject i. Let $\overline{S_i}$, S_i^{cov}, and $S_i^{maxdist}$ be the mean value, co-variance value, and max distance value of each subject i. This process is explained in Algorithm 1.
2. Each pattern of each outsider subject is subtracted by insider mean value. Next, the result is multiplied by the eigenvector of insider, calculating from co-variance. The result is divided by square root eigenvalue of the insider. If the result value is more than the insider threshold value, it is decided as outsider, otherwise it is an insider. Let S_j^N be each sample of outsider subject j. This process follows the Algorithm 2.

Algorithm 1 GetMeanCovarianceAndMaxdist

//comment S_i^N is an insider subject i and N is the number of samples from 1 to n//

1. procedure GetMeanCovarianceAndMaxdist (S_i^N)
2. $\overline{S_i} \leftarrow mean(S_i^N)$
3. $S_i^{cov} \leftarrow cov(S_i^N)$
4. $EigVectorS_i, EigenValueS_i \leftarrow S_i^{cov}$
5. $sort(EigVectorS_i)$
6. For S_i^N do
7. $Ssubtract_i^N \leftarrow (S_i^N - \overline{S_i})$
8. $Soutput_i^N \leftarrow (EigVectorS_i * Ssubtract_i^N)$
9. For $Soutput_i^N$ do
10. $Sdist_i^N \leftarrow sum((Soutput_i^N \,/\, sqrt(EigenValueS_i))^2)$
11. $smaxdist_i \leftarrow max(Sdist_i^N)$
12. return $\overline{S_i} S_i^{cov} smaxdist_i$

Algorithm 2 OutsiderSubjectDetection

//comment S_j^N is an outsider subject j and N is the number of samples 1 to n //

procedure Outsider Detection ($\overline{S_i}, S_i^{cov}, smaxdist_i, S_j^N$)

$EigVectorS_i, EigenValueS_i \leftarrow S_i^{cov}$

$sort(EigVectorS_i)$

For S_j^N do

$Ssubtract_j^N \leftarrow (S_j^N - \overline{S_i})$

$Soutput_j^N \leftarrow (EigVectorS_i * Ssubtract_j^N)$

For $Soutput_j^N$ do

$sdist_j \leftarrow sum((Soutput_j^N / sqrt(EigenValueS_i))^2)$

If $sdist_j > smaxdist_i$

$S_j^N.prediction \leftarrow outsider$

else

$S_j^N.prediction \leftarrow insider$

return $S_j^N.prediction$

Noise robustness of the proposed method is measured by adding two kinds of external noise affecting the EEG signals, namely, line noise and White Gaussian noise at SNR 5 and 20 dB, respectively. This experiment is demonstrated in Figure 9.

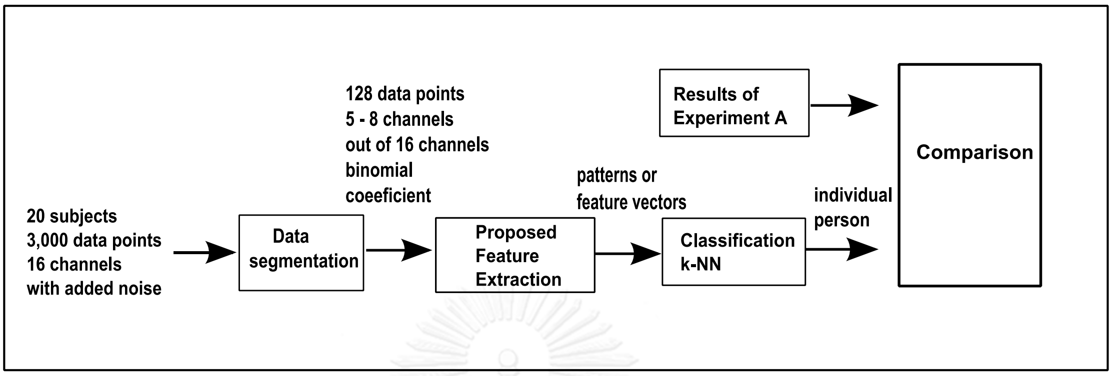

Figure 9. Experiment on noise robustness of the proposed method.

Finally, the experiment performs the proposed method on public database as mentioned. Result comparisons of the proposed method to other researches applying on the data set are shown in the next chapter.

CHAPTER 5

EXPERIMENTAL SETUP AND RESULTS

The identification process of this work comprises the following six processes:

5.1 Data collection

5.2 Experiment of the proposed feature extraction and classifiers

5.3 Experiment on determining minimum data, significant channels, and comparison

5.4 Experiment on separating a person from outsiders using minimum time and channels

5.5 Experiment on noise robustness of the proposed method

5.6 Experiment on implementing the proposed method on additional data set

5.1 Data Collection

Data of the EEG signals in this research were obtained from research data sets of Preecha's work [28] [29] [10] and PhysioBank [12]. These EEG signals of the first data set were recorded by Chulalongkorn Comprehensive Epilepsy Program (CCEP) of King Chulalongkorn Memorial Hospital. There are 40 normal subjects consisting of eighteen men and twenty two women. The age range of all subjects are between 12 and 40 years. EEG signals of the subjects were recorded based on international 10-20 system defining the location of scalp electrodes. Sixteen electrodes placement corresponding to 10-20 system were attached on the scalp of each subject. In addition, the subjects were motionless and performed no task during collecting the EEG signals. Recording of the EEG signals was performed at the following locations: Fp1, F7, T3, T5, Fp2, F8, T4, T6, F3, C3, P3, O1, F4, C4, P4, and O2. To easily represent EEG locations in this experiment, all of the above locations are labeled as ch1, ch2 ... ch16, respectively, as shown in Figure 10.

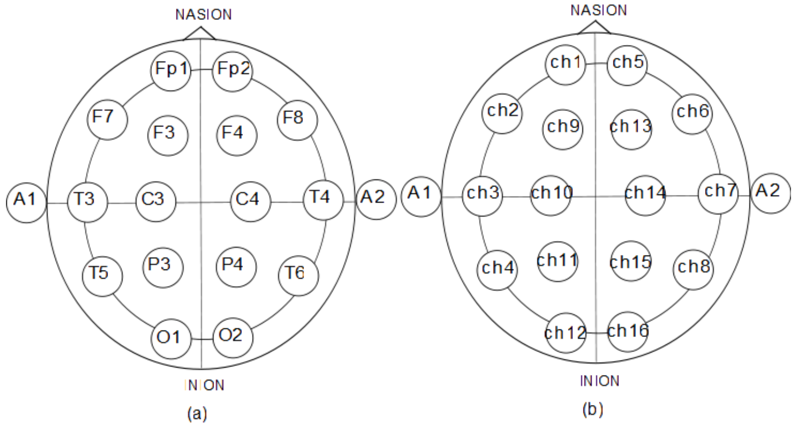

Figure 10. The 10-20 international system

In [28] [29] [10] recording process, Mastoid area A1 and A2 were electrically linked and used as reference with mono-polar montage. The EEG amplifier was Grass model 8 plus. The sampling rate of this EEG signals recording was 200 Hertz. The EEG signals were digitized and were notch filtered at 50 Hertz by BMSI board using Stellate Harmony EEG software. The digitized EEG data were exported as EDF (European Data Format). Electromyography (EMG) and Electrocardiogram (ECG) signals were primarily removed from EEG signals. For each channel of each subject, 3,000 data points were collected in 15 seconds.

Additional data set [11, 12] consisting of 109 subjects were also investigated for comparative purpose. The subjects were performed in resting state 1 minute with opened eyes, 1 minute closed eyes, three 2 minutes running of 4 tasks (2 motor and 2 motor-imagery). The EEG signals of each subject were recorded with 64 EEG channels at a sampling rate of 160 Hz, international 10-10 system. Due to more EEG channels in the data set, only 16 EEG channels were chosen to comply with 10-20 International system with closed and opened eyes.

5.2 Experiment on Determination of Proper Classifier for Classification

The aim of this experiment is to investigate an appropriate classifier to bestow high classification accuracy. This work used 3,000 data points for all 16 channels per subject. There were 20 subjects obtained in [28]. All signals in each channel were divided into 32, 64, and 128 data points for each subject. STFT was applied to transform the sample into a power spectrogram matrix by means of adjusting its various window lengths. The length of the data point signal and the length of STFT's window function were two important factors of the accuracy which needed to be precisely calculated. Basically, the window function length of STFT had to be smaller than the length of the data points expressed in power of two. The significant feature vector of each pattern was extracted from the matrix by means of SVD to a singular vector. Then classifiers were applied to measure the classification accuracy. For reliability of the experimental results, 10-fold cross validation was performed. Five classifiers were implemented in this experiment to achieve the general classification performance for high accuracy. They are MLP-Levenberg Marquardt, MLP-Bayesian regularization, RBF, Probabilistic Neural Networks, and k-NN.

For selecting parameter of Levenberg Marquardt (LM), the number of hidden neurons was set from 10 to 50 in step of 10. It was founded that the highest classification result of this approach was 99.56% with 40 hidden neurons and 64 data points having window function length of 8 data points. From the experiment, this method gave high accuracy results. However, it took a lot of time for the classification process.

Bayesian regularization (BR) function was selected on account of different pattern classification technique. The number of hidden neurons varied from 10 to 30 in step of 10. It was founded that the 20 hidden neurons led to the best classification accuracy. The result of this method gave an accuracy at 99.77% which was high achieved accuracy at 128 data points with window function length of 16 data points.

For RBFN classifier, the number of hidden neurons and the standard deviation parameter affected an average accuracy of classification. In this practice, the number of neurons was 20 and a spread parameter value varied from 100 to 300 in the step of 10. It was founded that the number of hidden neurons and spread parameters were 20 and 240 to achieve high classification accuracy at 98.67% at 128 data points having window function length of 16 data points.

For PNN classifier, the main parameter was the smoothing parameter which was used to adjust the classification to achieve accurate performance. The standard deviation values varied from 10 to 20 in step of 2. It was found that the best accuracy percentage was 99.83 when spread value was 12 unit at 64 data points having the window function length of 8 data points. The more spread value was, the less accuracy became.

For using k-NN classifier, the k value of k-NN approach was set to 1, 3, 5, 7 and 9 for accuracy of classification comparison. The classification accuracy was the highest if k value of k-NN was 1. It was founded that the best accuracy percentage was 99.83 when k was 1 at 64 data points with the window function length of 8 data points. The classification accuracy and experiment results are shown in Table 3.

Table 3. Determination of the optimal parameters and the suitable classifiers with using 10-fold cross validation.

Data Points	Window Length	MLP (LM)	MLP (BR)	RBF	PNN	k-NN
128	64	98.88	99.44	97.67	99.55	99.55
	32	98.44	99.55	98.00	99.55	99.55
	16	99.00	99.77	98.67	99.78	99.78
	8	97.21	98.88	90.57	99.77	99.77
64	32	98.64	99.51	96.73	99.45	99.29
	16	99.40	99.51	97.17	99.78	99.72
	8	99.56	99.29	93.15	99.83	99.83
32	16	98.11	99.00	95.08	98.81	98.49
	8	96.80	98.41	89.54	99.24	99.11

From Table 3, the highest accuracy is considered for classifier selection. It can be seen that PNN and k-NN obtain high average accuracy percentage which are 99.24 and 99.11, respectively, when the data point length equals 32 or about 0.16 seconds. Both give the highest average accuracy 99.83 for data points which equal 64 or 0.32 seconds. Furthermore, the data point length of 128 or 0.64 seconds is considered. Both classifiers bestow the same highest average accuracy of 99.78.

It can be considered that PNN and k-NN give the highest classification accuracy. It is evident that PNN and k-NN can be chosen as the suitable classifiers in this study. Due to its simplicity and less processing time, k-NN is used as the main classifier in the experiment. However, PNN may be used in real world applications.

The next step is to decide the suitable length of data point signals, the length of STFT window function, and the minimum channel for the highest accuracy using the proposed method. The window function length needs to be adjusted. Channel combination and the classifier perform in the same manner.

5.3 Experiment on Determining Minimum Data, Significant Channels, and Comparison

The objective is to investigate minimum channel combinations which give the highest classification accuracy. The same EEG data were used in this section. The experiment was divided into two parts. The first part called experiment A was implemented without cleaning the signals using ICA process. The second part called experiment B used the same data set incorporating ICA process following [10]. Both parts were divided into subgroups. Each subgroup was equally set the length of data point signals, namely, 32, 64, 128, and 256. All subgroups of each subject were conducted by selecting 2 to 8 channels out of 16 channels with binomial coefficient. Thus, each subgroup of each subject consisted of 120, 560, 1820, 4,368, 8008, 11,440, and 12,870 combinations for ${}^{16}C_2$, ${}^{16}C_3$, ${}^{16}C_4$, ${}^{16}C_5$, ${}^{16}C_6$, ${}^{16}C_7$, and ${}^{16}C_8$, respectively. In addition, each combination was extracted features with varied length of STFT's window function and classified by k-NN setting k to 1 and 10-fold cross validation was used.

Table 4. Classification accuracy of 20 subjects with 2 to 8 channel combinations of an experiment A and experiment B (with ICA).

Experiment A		Data points				Experiment B		Data points			
#CH	window	256	128	64	32	#CH	window	256	128	64	32
16	128	98.65				16	128	99.77			
	64	98.85	99.56				64	99.77	99.89		
	32	99.33	99.56	99.29			32	99.77	99.89	99.95	
	16	99.55	99.78	99.73	98.49		16	99.77	99.89	99.95	99.97
	8	99.31	99.78	99.84	99.11		8	99.77	100	100	100
8	128	99.33				8	128	99.77			
	64	99.55	99.67				64	99.77	99.89		
	32	99.55	99.78	99.19			32	99.77	99.89	100(139)	
	16	99.77	99.89	99.89	97.88		16	99.77	100(12)	100(1270)	
	8	99.78	100(104)	99.95	98.23		8	100(691)	100(1913)	100(7222)	100(2229)
7	128	99.33				7	128	99.77			
	64	99.55	99.56				64	99.77	99.89		
	32	99.55	99.78	98.91			32	99.77	100(1)	100(16)	
	16	99.77	99.89	99.89	97.34		16	99.77	100(18)	100(483)	100(617)
	8	100(1)	100(27)	99.84	98.04		8	100(482)	100(1154)	100(3639)	100(617)
6	128	99.11				6	128	99.77			
	64	99.33	99.34				64	99.77	99.89		
	32	99.55	99.78	98.69			32	99.77	100(1)	99.95	
	16	99.77	99.89	99.78	96.35		16	99.77	100(16)	100(58)	99.95
	8	99.77	100(1)	99.62	97.12		8	100(224)	100(474)	100(782)	100(77)
5	128	98.65				5	128	99.77			
	64	99.55	98.67				64	99.77	99.89		
	32	99.55	99.67	97.01			32	99.77	99.89	99.78	
	16	99.77	99.89	99.40	94.82		16	99.77	100(2)	99.95	99.54
	8	99.32	99.78	99.19	95.3		8	100(16)	100(92)	100(42)	100(6)
4	128	97.54				4	128	99.77			
	64	99.1	96.78				64	99.77	99.78		
	32	99.55	99.23	94.56			32	99.77	99.89	99.41	
	16	99.77	99.89	98.1	90.46		16	99.77	99.89	99.84	97.93
	8	98.87	99.22	97.77	91.77		8	100(12)	100(7)	99.95	99.79

Table 4. (continued).

Experiment A		Data points				Experiment B		Data points			
#CH	window	256	128	64	32	#CH	window	256	128	64	32
3	128	94.12				3	128	99.55			
	64	97.75	92.88				64	99.54	98.56		
	32	99.08	97.77	86.85			32	99.77	99.55	95.33	
	16	99.55	99.22	95.55	81.59		16	99.77	99.89	98.21	91.91
	8	97.68	98.09	94.84	84.33		8	99.78	99.89	99.62	98.42
2	128	84.21				2	128	95.21			
	64	93.41	79.01				64	96.1	88.31		
	32	95.66	87.33	66.46			32	97.7	92.01	80.22	
	16	97.25	94.91	82.39	61.54		16	98.88	96.69	87.56	73.34
	8	94.76	91.89	83.37	68.89		8	99.54	98.34	94.79	87.13

From Table 4, there are three main factors directly affecting the accuracy: the number of channel combinations, data point's length, and window function length of STFT. In other words, the figure in parentheses of 100 percentage means the number of combinations. The accuracy percentage of experiment A was 100 with 6 channel combinations which had only one combination. But for 4 channel combinations with window function length 8 of experiment B, it was founded that there were 12 possible combinations reaching 100 percent at 256 data points and 7 possible combinations reaching 100 percent at 128 data point signals. Each highlighted cell in experiment B corresponded to 100 percent of experiment A.

From the experiment, experiment B gave higher accuracy than experiment A. Nevertheless, experiment B took more time than experiment A. Experiment B actually used all channels and all data points for ICA processing. This issue will be discussed in Discussion Section.

Preecha's research [28] [29] [10] implemented ICA as the feature extraction and artificial neural networks as the main classifier for person identification. For 20 subjects, there were 16 EEG channels. Each channel consisted of 3000 data point signals. In order to investigate the minimal channel to obtain high average accuracy of person identification, 2 to 7 channel combinations were performed. According to experimental results of Preecha's works, the best identification accuracy percentages of each channel combination were 61.67, 90.52, 98.85, 99.87, 99.97, and 100. The parietal lobe of the brain was identified as the significant location of identification.

Table 5. Comparison accuracy percentage of this experimental results with those of Preecha.

	The number of channel combination						
	2	3	4	5	6	7	8
Preecha	61.67	90.52	98.85	99.87	99.97	100	N.A.
Experiment A (128, 8)	91.89	98.09	99.22	99.78	100	100	100
Experiment A (128, 16)	94.80	99.22	99.88	99.88	99.88	99.88	99.88
Experiment B (128, 8)	98.34	99.88	100	100	100	100	100
Experiment B (128, 16)	96.68	99.88	99.88	100	100	100	100

The results of the 2 to 7 channel combinations were compared with those of Preecha's experiments to evaluate the accuracy of the proposed approach. 128 data point length was used with window function length of 8 and 16 data points in both experiments. From Table 5, the accuracy percentage of the proposed experimental results was fairly low in 2 channel combinations. Nevertheless, these accuracy percentages were still higher than those experimental results of Preecha. The proposed experimental results could be summarized that identification of 20 subjects using 2 to 8 channel combinations had higher accuracy percentage than all Preecha experimental results. In addition, the identification time with deploying the proposed approach was less than that of Preecha [10] in all channel combinations.

Table 6. The accuracy percentage of 2 channel combinations at 128 data points with window function length of 8 data points using 10-fold cross validation.

2 channels		Average accuracy
3	4	91.89
1	11	91.00
1	12	90.67
4	16	90.44
1	4	90.34

Table 7. The accuracy percentage of 3 channel combinations at 128 data points with window function length of 8 data points using 10-fold cross validation.

3 channels			Average accuracy
1	3	11	98.098
1	4	9	98.018
3	4	6	97.893
3	4	16	97.772
1	9	11	97.674

Table 8. The accuracy percentage of 4 channel combinations at 128 data points with window function length of 8 data points using 10-fold cross validation.

4 channels				Average accuracy
1	4	11	16	99.22
1	4	9	12	99.11
1	4	12	16	99.00
1	9	11	13	99.00
1	4	9	13	99.00
1	9	11	12	99.00
1	9	10	11	99.00

Table 9. The accuracy percentage of 5 channel combinations at 128 data points with window function length of 8 data points using 10-fold cross validation.

5 channels					Average accuracy
1	6	9	10	12	99.78
1	4	6	9	11	99.78
1	6	9	11	16	99.77
3	6	9	10	16	99.77
1	3	5	11	13	99.77

Table 10. The accuracy percentage of 6 channel combinations at 128 data points with window function length of 8 data points using 10-fold cross validation.

6 channels						Average accuracy
1	5	6	9	11	16	100
1	3	4	6	9	10	99.891
1	3	6	8	9	12	99.891
1	6	8	9	10	12	99.891
1	6	9	10	12	15	99.891
1	6	9	10	12	16	99.891
3	6	9	10	15	16	99.891
1	3	6	9	10	16	99.89
1	3	6	9	11	16	99.89

Table 11. The accuracy percentage of 7 channel combinations at 128 data points with window function length of 8 data points using 10-fold cross validation.

7 channels							Average accuracy
1	2	3	4	6	10	16	100
1	2	3	6	10	11	16	100
1	2	3	6	10	15	16	100
1	2	5	6	9	10	12	100
1	3	4	5	6	10	15	100
1	3	4	6	9	10	16	100
1	3	5	6	9	10	16	100
1	3	5	6	9	11	12	100
1	3	5	6	11	13	15	100
1	3	6	8	9	12	16	100
1	3	6	9	10	11	12	100
1	3	6	9	10	11	16	100
1	3	6	9	10	12	16	100
1	3	6	9	10	14	16	100
1	3	6	9	11	12	16	100
1	3	6	9	11	13	16	100
1	3	6	9	12	14	16	100
1	4	5	6	12	15	16	100
1	4	5	9	10	11	12	100
1	5	6	9	10	11	16	100

Table 11. (continued).

7 channels							Average accuracy
1	5	6	9	10	13	16	100
1	5	6	9	11	13	16	100
1	5	6	10	12	13	15	100
1	6	9	10	11	13	16	100
1	6	9	10	13	15	16	100
3	5	6	9	10	11	12	100
3	5	6	9	10	11	16	100

Table 12. The accuracy percentage of 8 channel combinations at 128 data points with window function length of 8 data points and using 10-fold cross validation.

8 channels								Average accuracy
1	2	3	4	5	6	10	12	100
1	2	3	4	6	9	10	12	100
1	2	3	4	6	9	10	16	100
1	2	3	4	6	9	11	12	100
1	2	3	4	6	9	12	16	100
1	2	3	4	6	10	13	16	100
1	2	3	5	6	8	10	16	100
1	2	3	5	6	8	12	16	100
1	2	3	5	6	9	10	12	100
1	2	3	5	6	9	11	12	100
1	2	3	5	6	10	13	16	100
1	2	3	6	9	10	11	16	100
1	2	3	6	9	10	12	16	100
1	2	3	6	9	11	12	15	100
1	2	3	6	10	11	13	16	100
1	2	5	6	8	9	10	12	100
1	2	5	6	8	12	15	16	100
1	2	5	6	9	10	11	12	100
1	2	5	6	9	10	12	13	100
1	2	5	6	9	12	13	16	100
1	2	5	6	10	12	15	16	100
1	2	5	6	10	13	15	16	100
1	2	5	6	10	13	15	16	100
1	2	6	9	10	13	14	16	100

Table *12. (Continued).*

8 channels								Average accuracy
1	2	6	9	10	13	15	16	100
1	3	4	5	6	7	10	16	100
1	3	4	5	6	8	11	16	100
1	3	4	5	6	9	10	11	100
1	3	4	5	6	9	10	15	100
1	3	4	5	6	9	10	16	100
1	3	4	5	6	9	11	13	100
1	3	4	5	6	9	12	16	100
1	3	4	5	6	10	11	16	100
1	3	4	5	6	10	13	15	100
1	3	4	5	6	10	15	16	100
1	3	4	5	6	12	15	16	100
1	3	4	5	9	10	11	12	100
1	3	4	6	9	10	14	16	100
1	3	5	6	7	9	11	12	100
1	3	5	6	8	9	10	12	100
1	3	5	6	8	9	10	16	100
1	3	5	6	8	9	12	13	100
1	3	5	6	8	9	12	16	100
1	3	5	6	8	10	15	16	100
1	3	5	6	8	11	15	16	100
1	3	5	6	9	10	11	12	100
1	3	5	6	9	10	11	16	100
1	3	5	6	9	10	13	16	100
1	3	5	6	9	10	15	16	100
1	3	5	6	9	11	12	13	100

Table 12. (Continued).

8 channels								Average accuracy
1	3	5	6	9	11	12	15	100
1	3	5	6	10	11	13	16	100
1	3	5	6	10	11	15	16	100
1	3	5	6	10	13	15	16	100
1	3	5	8	10	12	13	16	100
1	3	6	7	8	9	12	16	100
1	3	6	7	9	10	11	16	100
1	3	6	7	9	10	12	16	100
1	3	6	8	9	10	12	16	100
1	3	6	8	9	11	12	16	100
1	3	6	8	9	12	13	16	100
1	3	6	8	9	12	14	16	100
1	3	6	8	9	14	15	16	100
1	3	6	9	10	11	12	16	100
1	3	6	9	10	11	14	16	100
1	3	6	9	10	12	14	16	100
1	3	6	9	10	14	15	16	100
1	3	6	9	11	12	14	16	100
1	3	6	9	11	14	15	16	100
1	4	5	6	9	10	11	16	100
1	4	5	6	9	10	13	16	100
1	4	5	6	11	12	15	16	100
1	4	5	6	12	13	15	16	100
1	4	5	6	12	14	15	16	100
1	5	6	7	9	10	12	16	100
1	5	6	8	9	10	12	15	100

Table 12. (Continued).

8 channels								Average accuracy
1	5	6	8	9	10	13	16	100
1	5	6	8	9	11	13	16	100
1	5	6	8	9	12	13	16	100
1	5	6	8	11	12	15	16	100
1	5	6	9	10	11	12	13	100
1	5	6	9	10	11	12	15	100
1	5	6	9	10	11	13	16	100
1	5	6	9	10	11	15	16	100
1	5	6	9	10	12	13	15	100
1	5	6	9	10	12	13	16	100
1	5	6	9	10	13	14	16	100
1	5	6	9	10	13	15	16	100
1	5	6	9	10	14	15	16	100
1	5	6	9	11	12	13	15	100
1	5	6	9	11	12	13	16	100
1	5	6	9	11	13	14	16	100
1	5	6	9	11	13	15	16	100
1	5	6	10	11	12	15	16	100
1	5	6	10	11	13	15	16	100
1	5	6	10	12	13	15	16	100
1	5	6	11	12	13	15	16	100
1	5	6	11	12	14	15	16	100
2	5	6	8	9	10	12	15	100
2	5	6	8	12	13	15	16	100
3	4	5	6	9	10	11	12	100
3	4	5	6	9	10	12	16	100

Table 12. (Continued).

8 channels								Average accuracy
3	5	6	8	9	10	11	16	100
3	5	6	9	10	11	12	15	100
3	5	6	10	11	13	15	16	100

Table 6 to Table 12 show the best accuracy percentage of 2, 3, 4, 5, 6, 7 and 8 channel combinations with 10-fold cross validation and 128 data points using k-NN, k = 1 having the highest accuracy percentage of 98.09, 99.22, 99.78, 100, and 100, respectively. For 8 channel combinations, there are 104 combinations that achieve the average accuracy at 100 percent. The results are summarized in Table 13.

Table 13. The highest result of each channel combination in the 20 subject insider.

Channel combination	Channel number	Highest accuracy %
8	(104 combinations)	100
7	(27 combinations)	100
6	1, 5, 6, 9, 11, 16	100
5	1, 6, 9, 10, 12	99.78
4	1, 4, 11, 16	99.22
3	1, 3, 11	98.00
2	3, 4	91.80

From Table 13, the highest accuracy of each channel combination adopted the reflection coefficient of AR approach of Maiorana's study. The results of Maiorana were compared with the proposed method and Preecha's work as shown in Table 14.

Table 14. Comparison of the highest result from each channel combinations in the 20 subject insider group with 128 data points 10-fold cross validation using AR method in Maiorana [27] and ICA method in Preecha [10].

Channel combination	Channel number	Highest percentage accuracy		
		Proposed % (SD)	Maiorana %/order th (SD)	Preecha %
7	(27 combinations)	100 (0)	100/36 (0) (2 combinations) (ch 1,3,6,9,11,12,16 order 42) (ch3,5,6,9,10,11,16 order 36)	100
6	1, 5, 6, 9, 11, 16	100 (0)	99.77/47 (0.47)	99.97
5	1, 6, 9, 10, 12	99.78 (0.68)	98.88/46 (0.74)	99.87
4	1, 4, 11, 16	99.22 (0.74)	99.22/34 (1.17)	98.85
3	1, 3, 11	98.00 (1.83)	98.89/44 (1.15)	90.52
2	3, 4	91.80 (1.53)	94.76/44/(2.23)	61.67

Table 14 shows the result accuracy percentage of 3 channel combinations of the proposed method that are slightly different from using AR approach, but are higher than those of Preecha's work. Although the 4 channel combination results of the proposed method are equal to AR, the standard deviation is less than AR. In 6 channel combinations, the results yield the highest accuracy at 100 percent classification. For 7 channel combinations, all three approaches reach 100 percent.

Table 15. The time consumption of the proposed method with 100 % and Maiorana [27] method for 6 to 8 channel combinations for insider group using 128 data point.

Channel combination	Channel number	Proposed method (second)				Maiorana (AR method) (second)			
		128 data points	Average of proposed method time x channels	classification	Total time (second)	Filter time with 3000 data points	Average of AR method time (128 data points) x channels	classification	Total time
8	(104 combinations)	0.64	0.003x8	0.0003	0.6643	15 (s)	0.0044x8	0.004	15.0392
7	(27 combinations)	0.64	0.003x7	0.0003	0.6613	15 (s)	0.0044x7	0.004	15.0348
6	1, 5, 6, 9, 11, 16	0.64	0.003x6	0.0003	0.6583	15 (s)	0.0044x6	0.004	15.0304

Table 15 shows the time consumption calculated from recording, feature extraction, and classification of test pattern. The minimum 6, 7, 8 channel combinations attain the identification time of 0.65, 0.66, 0.66 seconds, respectively, which are less than AR method.

5.4 Experiment on Separating a Person from Outsiders Using Minimum Time and Channels

The insider group was deployed as the trained set and the outsider group was the test set. Each sample size of the data set and the test set had 128 data points having the length of window function of 8. Algorithm 1 and 2 were used to classify each outsider person one by one from each insider person. The experimental results showed the accuracy percentage of the outsider samples belonging to each insider person. For instance, subject 21 in Table 16 with 3 channel combinations (19)100 means wrongly identified subject 21 who was predicted as subject 19 with 100 percentage of the outsider's samples. The number in parentheses is the incorrect subject identified.

Consider the previous experimental results, the classification of 6 channel combinations were ch 1, ch 5, ch 6, ch 9, ch 11, and ch 16 reaching 100 percent accuracy in experiment A using 128 data points with the window function length of 8 data points. The 128 data points became the proper length of EEG signals that yielded less identification time for the best insider subject classification.

From the experimental results, it can be seen that using 2 and 3 channel combinations cannot classify each outsider person from the insider person. For example, using 2 channel combinations, subject number 21 was predicted as subject number 19 with accuracy percentage 100. The 3 channel combinations did not improve satisfactory classification results because some outsider persons were predicted as the insider persons since more than ten outsider persons who were completely classified as the insider person.

Hence, further investigation on classification was conducted. The combined results of 6, 7, and 8 channels having 100 percent indicated that all outsiders were completely classified from insider. Therefore, 4 and 5 channel combinations were concentrated in order to find the minimum channels to classify the outsiders. The results are summarized in Table 17 and Table 18.

Table 16. The accuracy percentage of outsider classification tested with the highest accuracy 2, 3, 4, 5, 6, 7, and 8 channel combinations.

Subject number	Channel combination						
	2	3	4	5	6	7	8
	[3,4]	[1,3,11]	[1,4,11,16]	[1,6,9,10,12]	[1,5,6,9,11, 16]	27 combina tions	104 combina tions
21	(19) 100	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
22	(19) 60.00	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
23	(5) 44.44	(17) 2.22	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
24	(5) 60.00	(15) 4.44	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
25	(17) 26.67	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
26	(17) 44.47	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
27	(5) 93.33	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
28	(5) 100	(17) 4.44	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
29	(17) 93.33	(17) 33.33	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
30	(5) 68.89	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
31	(5) 82.22	(15) 4.44	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
32	(14) 28.89	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
33	(19) 66.67	(19) 11.11	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
34	(19) 40.00	(19) 11.11	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
35	(19) 35.57	(19) 13.33	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
36	(18) 57.78	(19) 13.33	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
37	(17) 35.56	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
38	(17) 84.44	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
39	(5) 28.89	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
40	(5) 88.89	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0

Table 17. The accuracy percentage of outsider classification tested with the optimum number of 4 channel combinations.

Subject number	4 Channel combination						
	[1,4, 11,16]	[1,4, 9,12]	[1,4, 12,16]	[1,9, 11,13]	[1,4,9 ,13]	[1,9, 11,12]	[1,9, 10,11]
21	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
22	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
23	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
24	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
25	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
26	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
27	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
28	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
29	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
30	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
31	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
32	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
33	(-) 0	(-) 0	(-) 0	(-) 0	(19) 2.22	(-) 0	(-) 0
34	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
35	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
36	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
37	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
38	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
39	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
40	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0

Table 18. The accuracy percentage of outsider classification tested with the optimum number of 5 channel combinations.

Subject number	5 Channel combination				
	[1,6,9 10,12]	[1,4,6, 9,11]	[1,6,9 11,16]	[3,6,9, 11,16]	[1,3,5 11,13]
21	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
22	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
23	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
24	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
25	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
26	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
27	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
28	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
29	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
30	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
31	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
32	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
33	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
34	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
35	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
36	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
37	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
38	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
39	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0
40	(-) 0	(-) 0	(-) 0	(-) 0	(-) 0

Table 17 demonstrates that the first four optimum 4 channel combinations achieve complete outsider classification. However, the channel combinations of ch 1, ch 4, ch 9, ch 13 give a slight misclassification of subject number 33. Table 18 demonstrates that the first five optimum 5 channel combinations achieve complete outsider classification.

5.5 Experiment on Noise Robustness of the Proposed Method

This experiment evaluated noise robustness of the proposed method. Two kinds of noise, line noise and White Gaussian noise, were simulated and added to the signals 128 data point length with 5 to 8 and 16 channel combinations. Most line noises occurred in general power supply devices and White Gaussian noise was found in external electronic devices. Both occurred inevitably in natural environment. The original signals were added with line noise to measure line noise robustness. The signals were then combined with white Gaussian noise having SNR 5 and 20 dB to measure the noise robustness. The combined signals were processed by the proposed method to obtain the highest classification accuracy of each person.

From Table 19, most classification accuracy of 16 channel combinations with the above white Gaussian noise are not different from the original accuracy. However, the accuracy of the added line noise signals drops slightly at window length 16 and 32. The accuracy of 5 to 8 channel combinations with adding Gaussian noise SNR 20 dB are not different from the original accuracy, while most accuracy of the same set up having SNR 5 dB drops slightly as compare with the original accuracy of most window lengths. Nevertheless, all accuracy of the added line noise signals for 5 to 8 channel combinations decrease in comparison with the original. The experimental results of adding white Gaussian noise show that the data point length 128 and window function length at least 8 yield a consistent classification accuracy. While the accuracy comparison of adding line noise with the original signal drops slightly at the window length of 16. It can be summarized that the proposed method is still consistently accurate with white Gaussian noise. However, appropriate parameters setting of the added line noise signal are attributive to robustness.

Table 19. The highest average accuracy percentage of noise robustness of the proposed method with 128 data point length from 10-fold cross validation.

#CH	window function	#128 data points (added noise)			
		(no noise)	(SNR 20dB)	(SNR 5dB)	(line noise)
16	64	99.558	99.558	98.78	99.33
	32	99.558	99.558	99.558	99.55
	16	99.78	99.78	99.78	99.56
	8	99.775	99.775	99.775	98.44
8	64	99.67	99.67	99.006	99.01
	32	99.779	99.779	99.67	99.22
	16	99.889	99.889	99.889	99.44
	8	100	100	100	98.34
7	64	99.562	99.562	98.681	99.11
	32	99.779	99.779	99.67	99.00
	16	99.889	99.889	99.889	99.22
	8	100	100	100	97.78
6	64	99.338	99.337	98.124	96.56
	32	99.779	99.779	99.559	97.22
	16	99.889	99.889	99.889	99.00
	8	100	100	100	92.79
5	64	98.671	98.569	96.681	97.45
	32	99.67	99.558	99.558	97.00
	16	99.889	99.889	99.889	97.00
	8	99.783	99.779	99.448	90.66

5.6 Experiment on Implementing the Proposed Method on the Other Data Set

This experiment implemented the proposed method on the data set [12] to obtain the short processing time and percentage accuracy. The opened and closed eyes task being resting state were chosen in this experiment. First of all, EEG signals of each subject were cleaned at 60 Hz. The data set was segmented at 6.4, 7, 8, and 9 seconds, namely, 1024, 1120, 1280, and 1440 data points, respectively. The window function length was set at 16, 32, 64, 128, and 256. 10-fold cross validation and k-NN were used to evaluate and classify the data set. The experimental results were summarized in Table 20 and Table 21.

Table 20. The percentage accuracy of 109 subjects closed eyes with 16 EEG channels and k=1 at 10-fold cross validation.

	Data point			
Window length	1440	1280	1120	1024
256	97.39(1.69)	98.10(0.89)	94.77(1.25)	97.83(0.78)
128	97.99(1.24)	**98.43(0.71)**	97.77(1.01)	**98.43(0.78)**
64	97.92(0.86)	97.97(0.91)	97.83(1.26)	97.89(1.02)
32	97.76(1.18)	97.39(1.01)	97.14(1.01)	97.35(1.29)
16	95.40(1.75)	96.34(1.06)	96.16(1.13)	96.17(1.52)

Table 21. The percentage accuracy of 109 subjects opened eyes with 16 EEG channels and k=1 at 10-fold cross validation.

	Data point			
Window length	1440	1280	1120	1024
256	97.69(1.06)	98.42(0.78)	95.85(0.84)	97.30(1.22)
128	98.68(0.82)	99.07(1.02)	98.07(0.99)	98.85(1.14)
64	98.83(1.17)	**99.20(0.68)**	98.58(0.86)	98.70(0.86)
32	98.62(0.93)	98.56(0.80)	98.06(0.73)	98.43(0.77)
16	96.70(1.06)	96.97(0.97)	96.63(0.81)	96.64(0.87)

From Table 20 and Table 21, the highest percentage accuracy of 109 subjects closed eyes task is 98.43 at 1024 data points. The highest accuracy of opened eyes task is 99.20 percent at 1280 data points. It is evident from the results of both experiments that the highest accuracy occurs when subjects are opened eyes. However, processing time is longer than closed eyes task. Both results give higher accuracy than the work of DelPozo et al. [25], having 91.60 and 94.72 percent using the same data set. In addition, the number of subjects in this experiment is more than DelPozo's et al work [25], having only 20 subjects. It appears from this experiment that the proposed method correctly identifies persons and obtains high accuracy using EEG biometrics.

CHAPTER 6

DISCUSSION

6.1 Analysis of Method Computational Complexity

AR parameter estimation in this study is based on Burg approach using least squares estimation to obtain the *z* parameter from *l* data samples. Marple's work [43] reviewed that the computational complexity of AR was $O(lz)$ in theory. Most order values of AR model, or *z*, are greater than 36 in this work due to more features in each pattern. In addition, *N* is 3000 data points recording in 15 seconds, which is the same as Maiorana's work [27] performing band pass filter on whole signal. Therefore, the computational time of this study is $N+O(lz)$.

Basically, ICA approach needs long data length to obtain the quality of data set decomposition. From practical consideration of Chen [44] stated that for the number of data samples each channel used, ICA had to perform $N = d \times M^2$, where *N* is the number of data samples, *d* is a constant more than or equal to 20, and *M* is the number of sources or channels. Thus, the longer data sample length is used for decomposition, the better outcome is performed for these approaches. Boscolo et al. [45] and Albera et al. [46] showed that $SOBI_{rob}$, SOBI, InfoMax, FastICA, JADE, ERICA and SIMBEC computational complexity were approximated to $O(N^3)$. Theoretically, computational complexity of multilayer-perceptron neural networks is $O(tH(o_p + i_p))$ for training time and $O(1)$ for testing time, where, i_p is input size, *H* is the hidden units, o_p is the outputs, and *t* is the number of training epochs.

Computational complexity of the nearest neighbor is $O(k_n p_n q_n)$ for both training and testing time, where k_n is the number of nearest neighbor, p_n is the number of patterns, q_n is the number of attributes of a pattern. In this study, k_n is 1 and q_n is constant. Therefore, the computational complexity of the nearest neighbor is $O(p_n)$. The computational complexities are summarized in Table 22.

Table 22. Time Analysis.

Time	Proposed method	Maiorana (AR) [27] method	Preecha's method
Input	$l < N$	$l < N$	$N = d \times M^2$
Feature extraction	$O(l\log l) + O(pq^2)$	Filter(Bandpass) time (15 s) $N + O(lz)$	$O(N^3)$
Classification (Testing time)	$O(p_n)$	$O(p_n)$	$O(1)$

In Table 22, both classifiers are in linear time complexity depending on their variables. The time complexity evaluation of Preecha's input took 3000 data points consuming about 15 seconds with ICA process. Owing to conditions of the experiment, all subjects need to be in resting state with opened eyes. Actually, the data samples should be at least 20 * 162 or about 5120 data points to corresponding to Onton [44]. Hence, these data points are enough to be used in the experiment to obtain high classification accuracy. However, if the number of data samples of ICA, $N=dXM^2$, is considered, the data samples will still be conformed to the condition where d is constant. The ICA time complexity is $O(N^3)$. However, this is not appropriate for real time applications.

The proposed method applies STFT and SVD to perform feature extraction. The first method is STFT whose computational complexity is $O(l\log l)$, where l is the length of segmented data points less than N. The second method is SVD whose computational complexity is $O(pq^2)$, where p is the number of rows is and q is the number columns of power spectrogram matrix. Each matrix has constant matrix size because of STFT. So, the computational complexity is $O(l\log l) + O(pq^2)$ which is less than that of Maiorana and Preecha's researches, i.e., $N+O(lz)$ and $O(N^3)$.

6.2 Relative Relevant Minimum Channels to Human Brain Lobes

To discuss the important of minimum channels for leading to high identification accuracy, the experimental results of 7 and 8 channel combinations accuracy at 100 percent are considered. Each channel needs be organized into the same function group. Del Pozo et al. [17] summarized that topological grouping of channels could be divided into five areas, namely, frontal, parietal, central, temporal, and occipital area. In addition, the temporal area is divided into left and right to easily segment their functions. The areas are illustrated in Figure 11. The experimental results are shown in Figure 12 and Figure 13.

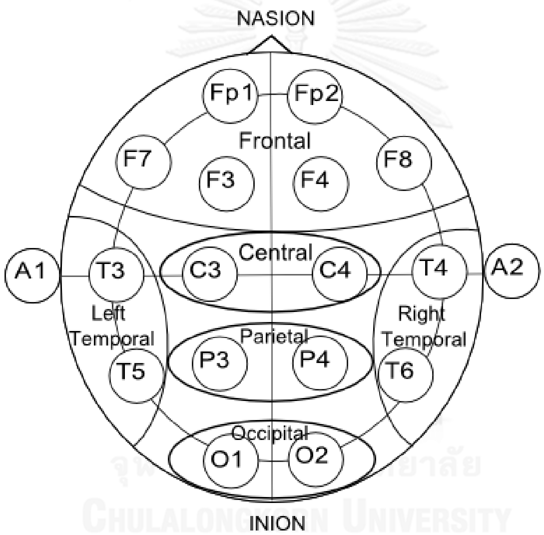

Figure 11. Topological grouping of channels on a human brain.

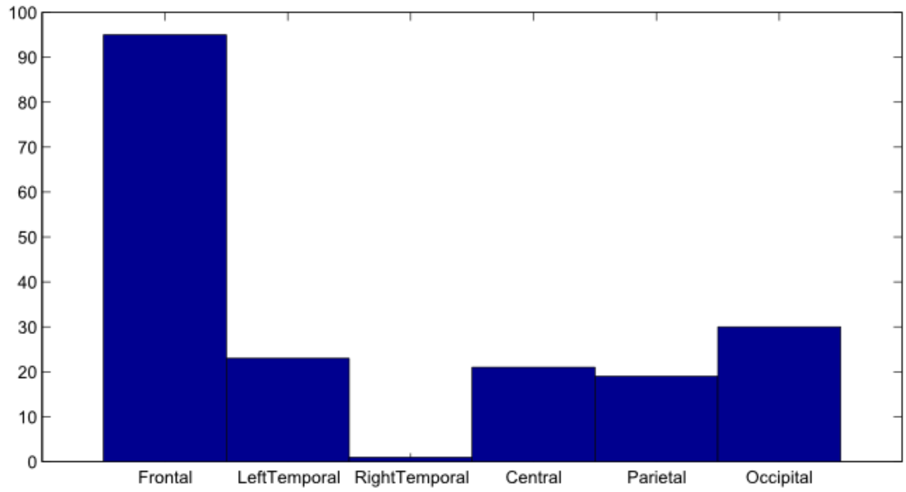

Figure 12. Histogram of relative relevant brain lobes with 7 channel combinations (128 data points and 8 data points of window function length).

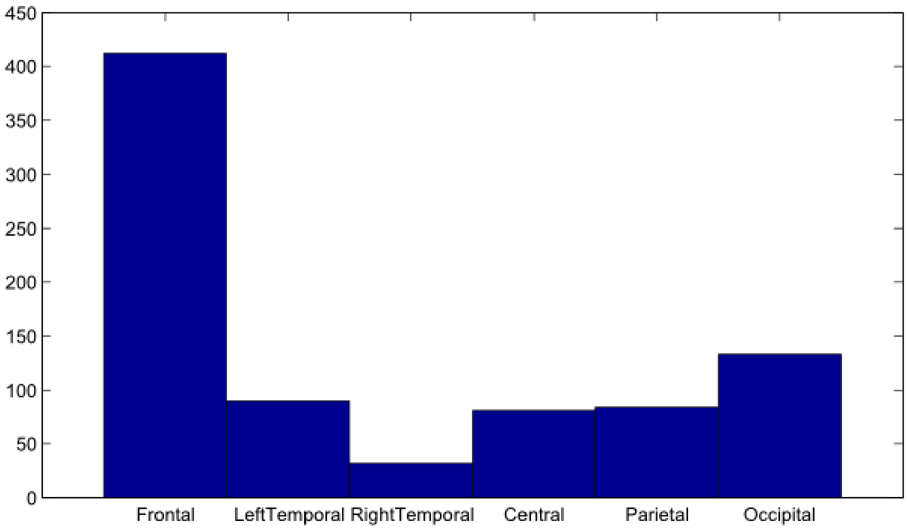

Figure 13. Histogram of relative relevant brain lobes with 8 channel combinations (128 data points and 8 data points of window function length).

From the histograms in Figure 12 and Figure 13, frontal and occipital areas perform better than other areas. Consider the number of channels in each area, it can be seen that frontal area will be used for person identification. Furthermore, this area complies with subject's brain network under resting state [13]. The frontal area processes many activities such as Attention, Judgment, Motor Planning, Emotional, and Verbal expression which are different for each individual. The accuracy percentage of using 6 channels in frontal area is 98.89. It is enough for identification. However, by combining the occipital area to improve higher accuracy, the accuracy percentage reaches 100. Therefore, the frontal area is the most meaningful area for person classification. The results are presented in Table 23.

Table 23. Accuracy percentage of using 6 channels 128 data points in Frontal area and combined Frontal and Occipital area.

Area	Channel Number	Accuracy percentage
Frontal	1,2,5,6,9,13	98.89
Frontal and Occipital	1,2,5,6,9,13,12,16	100

CHAPTER 7

CONCLUSION

In this work, feature extraction method based on STFT and SVD was proposed, along with the appropriate classifiers to achieve the highest person classification accuracy. In the experiment, each EEG signal was converted to power spectrogram matrix using STFT. Then, SVD was used to extract the singular values matrix from the STFT matrix, which were subsequently used as feature to classify a person. The appropriate length of EEG data points and the window function length of STFT were evaluated. Four kinds of classifiers were appraised, namely, MLP, RBF, PNN, and k-NN. The experiment indicated that PNN and k-NN were the suitable classifiers for person identification because both gave the average highest accuracy near 100% with 16 EEG channels. To investigate the minimum channel for subject identification, the experiment performed combinations of all channels on signals that were classified by using k-NN. The combinations were 2, 3, 4, 5, 6, 7, and 8 channels. Furthermore, the experimental results showed that between 2 and 7 channel combinations gave higher personal identification accuracy than the results of Maiorana [27] and Preecha [10, 28]. Moreover, the practical classification experiment using 6, 7, and 8 channel combinations yielded the accuracy of 100 percent of 20 subjects and completely classified unknown person who did not belong to the database with only identification time of 0.66 seconds. The proposed method was performed on the data set with only closed and opened eyes task in resting state. The average accuracy of this data set was 98 percent with 109 subjects. These results were higher than those of the previous work [25]. The success was attributive to frontal and occipital lobes which were likely to be significant for classification in real time. Thus, the proposed approach is very promising for person identification with high performance to be applied in real time and real life systems.

REFERENCES

1. Teplan, M., *Fundamentals of EEG measurement.* Measurement Science Review, 2002. **2**: p. 1-11.
2. Campbell, A., et al., *NeuroPhone: Brain-Mobile Phone Interface using a Wireless EEG Headset.* Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds - MobiHeld '10, 2010: p. 3-8.
3. Gong, F., et al., *NeuroGlasses: A neural sensing healthcare system for 3-D vision technology.* IEEE Transactions on Information Technology in Biomedicine, 2012. **16**(2): p. 198-204.
4. Yasui, Y., *A brainwave signal measurement and data processing technique for daily life applications.* Journal of physiological anthropology, 2009. **28**(3): p. 145-150.
5. Poulos, M., et al. *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*. 1999.
6. Paranjape, R.B., et al., *The electroencephalogram as a biometric.* Canadian Conference on Electrical and Computer Engineering 2001. Conference Proceedings (Cat. No.01TH8555), 2001. **2**: p. 1363-1366.
7. Palaniappan, R. *Method of identifying individuals using VEP signals and neural network*. in *Science, Measurement and Technology, IEE Proceedings-*. 2004. IET.
8. Palaniappan, R. and D.P. Mandic, *Biometrics from Brain Electrical Activity: A Machine Learning Approach.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007. **29**(4): p. 738-742.
9. Rocca, D.L., P. Campisi, and G. Scarano, *EEG Biometrics for Individual Recognition in Resting State with Closed Eyes.* International Conference of the Biometrics Special Interest Group (BIOSIG), 2012(Figure 1): p. 1-12.

10. Tangkraingkij, P., et al., *Insider and outsider person authentication with minimum number of brain wave signals by neural and homogeneous identity filtering.* Neural Computing and Applications, 2013. **22**(SUPPL.1): p. 463-476.
11. Schalk, G., et al., *BCI2000: a general-purpose brain-computer interface (BCI) system.* IEEE Transactions on biomedical engineering, 2004. **51**(6): p. 1034-1043.
12. Goldberger, A.L., et al., *PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals.* Circulation. **101**(23): p. e215--e220.
13. Rosazza, C. and L. Minati, *Resting-state brain networks: Literature review and clinical applications.* Neurological Sciences, 2011. **32**(5): p. 773-785.
14. Subasi, A. and M.I. Gursoy, *EEG signal classification using PCA, ICA, LDA and support vector machines.* Expert Systems with Applications, 2010. **37**(12): p. 8659-8666.
15. Muthuswamy, J. and N.V. Thakor, *Spectral analysis methods for neurological signals.* Journal of Neuroscience Methods, 1998. **83**(1): p. 1-14.
16. Mustafa, M., et al., *GLCM texture classification for EEG spectrogram image.* Proceedings of 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences, IECBES 2010, 2010(December): p. 373-376.
17. Del Pozo-Banos, M., et al., *Electroencephalogram subject identification: A review.* Expert Systems with Applications, 2014. **41**(15): p. 6537-6554.
18. Poulos, M., et al., *Person identification from the EEG using nonlinear signal classification.* Methods of information in medicine, 2002. **41**(1): p. 64-75.
19. Marcel, S. and J.d.R. Millan, *Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007. **29**(4): p. 743-748.
20. Shedeed, H.A., *A new method for person identification in a biometric security system based on brain EEG signal processing.* Proceedings of the 2011 World Congress on Information and Communication Technologies, WICT 2011, 2011: p. 1205-1210.
21. Cempírek, M. and J. Stastny, *The optimization of the EEG-based biometric classification.* Applied Electronics, 2007: p. 25-28.

22. Yazdani, a., et al., *Fisher linear discriminant based person identification using visual evoked potentials.* 2008 9th International Conference on Signal Processing, 2008: p. 1677-1680.

23. Riera, A., et al., *Unobtrusive biometric system based on electroencephalogram analysis.* Eurasip Journal on Advances in Signal Processing, 2008. **2008**.

24. Abdullah, M.K., et al., *Analysis of effective channel placement for an EEG-based biometric system.* Proceedings of 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences, IECBES 2010, 2010(December): p. 303-306.

25. DelPozo-Banos, M., et al., *EEG biometric identification: a thorough exploration of the time-frequency domain.* Journal of neural engineering, 2015. **12**(5): p. 056019-056019.

26. Rocca, D.L., et al., *Human brain distinctiveness based on EEG spectral coherence connectivity.* IEEE Transactions on Biomedical Engineering, 2014. **61**(9): p. 2406-2412.

27. Maiorana, E., et al., *On the Permanence of EEG Signals for Biometric Recognition.* 2016. **11**(1): p. 163-175.

28. Tangkraingkij, P., et al., *Selecting Relevant EEG Signal Locations for Personal Identification Problem Using ICA and Neural Network.* 2009 Eighth IEEEACIS International Conference on Computer and Information Science, 2009: p. 616-621.

29. Tangkraingkij, P., et al., *Personal identification by EEG using ICA and neural network.* Computational Science and Its Applications–ICCSA 2010, 2010: p. 419-430.

30. Schilling, R.J. and S.L. Harris, *Introduction to Digital Signal Processing Using MATLAB*. 2011: Cengage Learning.

31. Avargel, Y. and I. Cohen, *System identification in the short-time fourier transform domain with crossband filtering.* IEEE Transactions on Audio, Speech and Language Processing, 2007. **15**(4): p. 1305-1319.

32. Qaisar, S.M., L. Fesquet, and M. Renaudin, *An Adaptive Resolution Computationally Efficient Short-Time Fourier Transform.* Research Letters in Communications, 2008. **2008**: p. 1-5.

33. Brand, M., *Fast low-rank modifications of the thin singular value decomposition.* Linear Algebra and Its Applications, 2006. **415**(1): p. 20-30.

34. Cardoso, J.F. and a. Souloumiac, *Blind beamforming for non-gaussian signals*. 1993. p. 362-362.

35. Belouchrani, a., et al., *A blind source separation technique using second-order statistics.* IEEE Transactions on Signal Processing, 1997. **45**(2): p. 434-444.

36. Ziehe, A. and K.R. Müller, *TDSEP- an efficient algorithm for blind separation using time structure.* Proc. of the 8th International Conference on Artificial Neural Networks, ICANN'98, 1998: p. 675-680.

37. Belouchrani, A. and A. Cichocki, *Robust whitening procedure in blind source separation context.* Electronics Letters, 2000. **36**(24): p. 2050-2051.

38. Hyvärinen, a. and E. Oja, *Independent component analysis: algorithms and applications.* Neural networks : the official journal of the International Neural Network Society, 2000. **13**(4-5): p. 411-30.

39. Cruces, S., L. Castedo, and A. Cichocki, *Robust blind source separation algorithms using cumulants.* Neurocomputing, 2002. **49**: p. 87-118.

40. Anderson, C.W., E.A. Stolz, and S. Shamsunder, *Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks.* IEEE Transactions on Biomedical Engineering, 1998. **45**(3): p. 277-286.

41. Mao, K.Z., K.C. Tan, and W. Ser, *Probabilistic neural-network structure determination for pattern classification.* IEEE Transactions on Neural Networks, 2000. **11**(4): p. 1009-1016.

42. Chen, C.-h., L.-F. Pau, and P.S.-p. Wang, *Handbook of pattern recognition and computer vision*. Vol. 27. 2010: World Scientific.

43. Marple, L., *A New Autoregressive Spectrum Analysis Algorithm.* IEEE Transactions on Acoustics, Speech, and Signal Processing, 1980. **28**(4): p. 441-454.

44. Onton, J., et al., *Imaging human EEG dynamics using independent component analysis.* Neuroscience and Biobehavioral Reviews, 2006. **30**(6): p. 808-822.

45. Boscolo, R., H. Pan, and V.P. Roychowdhury, *Independent component analysis based on nonparametric density estimation.* IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council, 2004. **15**(1): p. 55-65.

46. Albera, L., et al., *ICA-Based EEG denoising: a comparative analysis of fifteen methods.* Bull Pol Ac: Tech, 2012. **60**(3): p. 407-418.

VITA