BRAIN COMPUTER INTERFACING (BCI) DATA ANALYSIS USING GRAPH SIGNAL PROCESSING

by

Sevde Büşra Bayrak

B.S., Mathematics, Mimar Sinan Fine Arts University, 2016

Submitted to the Institute of Biomedical Engineering
in partial fulfillment of the requirements
for the degree of
Master of Science
in
Biomedical Engineering

Boğaziçi University
2021

ACKNOWLEDGMENTS

Firstly, I am indebted to Prof. Dr Ahmet Ademoğlu, who believed in me from the beginning of my journey and gave me the freedom to discover the field. His passion and vision about asking the broader question made me a better researcher. I aspire to have your ability to discuss not only research subjects but also linguistics, literature, and history. I have grown by modelling his intellectualism and kindness. Thank you for encouraging me whenever I am stuck and for the endless support. I am also thankful to my jury members Assoc. Prof. Esin Öztürk Işık and Asst. Prof. İpek Şen.

Throughout the years, I sometimes come across with difficulties; but I was not alone. My family were always with me whenever I needed it. I want to thank my father, Sefer Bayrak, and my mother, Fatma Bayrak, for their faith in me. I could not manage to finish this task without their enduring supportive speech.

It would be remiss not to thank my lovely big brother, Rıfat Bayrak. He is not just a brother but also a best friend to me. Without our long hour phone calls, I would be devastated. I am the luckiest person for having a brother like you. Thank you for the laughs, gossips and motivational speeches. In addition, I want to thank my sister in law Emine and my nephew Reyyan Meva Bayrak for their love.

My mother sister, Hanife Başkan Kalafat, inspire my heart to become green with her words throughout the graduate life. Her encouragement and wisely recommendations help me to endure in both academic and social life. I sincerely thank you since you are the pillar of support for me. I also thank her husband Selçuk Kalafat for his valuable and unforgettable help. My lovely cousins Meryem Kalafat and Halil Íbrahim Kalafat, thank you for your love and willingness to spare time with me.

Thanks you to a wingless angel Elmas Par Lynch who helps me to realize my potential and take it out by asking the right questions. I appreciate your help me to find my way. My graduate school friends Başak Dalbayrak, Alamira Hajjar, and Özde Zeynep Güner, my labmates Ayşe Akgün Demir and Ebru Ayyürek, thank you all for help and understanding.

My dearest, funniest, and kindness friends Merve Yüce and Ayça Fırtın, what could I do if you were not with me. Without our laughs, adventures, and gossips, life would be colourless. Lastly, I personally feel gratified for believing in myself to do something that requires hard work, patience and persistence.

ACADEMIC ETHICS AND INTEGRITY STATEMENT

I, Sevde Büşra Bayrak, hereby certify that I am aware of the Academic Ethics and Integrity Policy issued by the Council of Higher Education (YÖK) and I fully acknowledge all the consequences due to its violation by plagiarism or any other way.

Name : ______________________________ Signature: ______________________________

Date: ______________________________

ABSTRACT

BRAIN COMPUTER INTERFACING (BCI) DATA ANALYSIS USING GRAPH SIGNAL PROCESSING

Data have been growing enormously in various domains including society, economics, industry, security, transportation, and medicine. The high dimensional structure of these data requires new techniques that employ their underlying connectivity structure. Graph signal processing (GSP) has emerged as a processing tool for high dimensional datasets as an extension of classical signal processing performed in the Euclidean space. In this thesis, electroencephalography (EEG) data collected for brain computer interfacing (BCI) are used for classification using GSP as a preprocessing tool. Two EEG datasets, one during emotion detection, and one during motor imagery are used. Support vector machines (SVM) and $K-$nearest neighboring algorithms are used for classification. The underlying connectivity structure of the EEG data is obtained using the distance and neighboring information of the electrode locations on the scalp. The results show that the classification accuracy is significantly improved when the data are projected to the underlying graph subspace determined by the graph spectral eigenvectors followed by a temporal filtering determined by Fourier spectral eigenvectors as a preprocessing step before classification.

Keywords: Graph Signal Processing, EEG, SVM, KNN, Brain Computer Interfacing (BCI).

ÖZET

ÇİZGE SİNYAL İŞLEME İLE BEYİN BİLGİSAYAR ARAYÜZÜ VERİLERİNİN ANALİZİ

Veriler, toplum, ekonomi, endüstri, güvenlik, ulaşım ve tıp dahil olmak üzere çeşitli alanlarda muazzam bir şekilde büyüyor. Bu verilerin çok boyutlu yapısı, alttlarında yatan bağlantı yapılarını kullanan yeni teknikler gerektirir. Çizge sinyal işleme, Öklid uzayında gerçekleştirilen klasik sinyal işlemenin bir uzantısı olarak yüksek boyutlu veri kümeleri için bir işleme aracı olarak ortaya çıkmıştır. Bu tezde, beyin bilgisayar arayüzü (BCI) için toplanan elektroensefalografi (EEG) verileri, ön işleme aracı olarak çizge sinyal işleme yöntemi kullanılarak sınıflandırma için kullanılmıştır. Biri duygu algılama ve diğeri motor imgeleme sırasında olmak üzere iki EEG veri kümesi kullanılmıştır. Destek vektör makineleri (SVM) ve K$-$ en yakın komşu algoritmaları sınıflandırma için kullanılmıştır. EEG verilerinin altında yatan bağlantısallık yapısı, kafatasına konumlandırılan elektrotların birbirleri ile uzaklığı ve komşu bilgileri kullanılarak elde edilmiştir. Sonuçlar, data Fourier spektral özvektörleri tarafından belirlenen zamansal frekansın takip ettiği graf spektral özvektörleri tarafından belirlenen graph alt uzayına sınıflandırma işleminden önce bir ön işlem olarak iz düşürüldüğünde, sınıflandırma doğruluğunun önemli ölçüde arttığını göstermektedir.

Anahtar Sözcükler: Çizge Sinyal İşleme, EEG, SVM, KNN, Beyin Bilgisayar Arayüzü

.

To my family

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

G	Graph
V	Set of Vertex
E	Set of Edge
v_i	Vertex i
N	Adjacent Vertices of v_i
A	Adjacency Matrix
D	Degree Matrix
L	Laplacian Matrix
U	Eigenvector Matrix
Λ	Eigenvalue Matrix
λ_i	ith Eigenvalue
s	Graph signal
δ	Graph Laplacian Spectrum
$\hat{f}(\lambda_i)$	Graph Fourier Transform
$\hat{f}(\xi)$	Fourier Transform
$f(j)$	Inverse Graph Fourier Transform
$\hat{f}_{out}(\lambda_i)$	Graph Filtered Signal
$\hat{h}_{Lr}$	Lowpass Graph Filter
$\hat{h}_{Hr}$	Highpass Graph Filter

LIST OF ABBREVIATIONS

ADHD	Attention Deficit Hyperactivity Disorder
ASD	Autism Spectrum Disorder
ASP	Algebraic Signal Processing
BCI	Brain Computer Interface
BMI	Brain Machine Interface
CT	Computer Tomography
EEG	Electroencephalography
EMD	Emprical Mode Decomposition
ERD	Event Related Desynchronization
ERP	Event Related Potentials
ERS	Event Related Synchronization
fMRI	Functional MRI
GAD	Generalized Anxiety Disorder
GSP	Graph Signal Processing
ICA	Independent Component Analysis
KNN	K-Nearest Neighbors
LLD	Late-Life Depression
MEG	Magnetoencephalography
MRI	Magnetoencephalography
MS	Multiple Sclerosis
PCA	Principal Component Analysis
PET	Positron Emission Tomography
RNA	Ribonucleic Acis
SPECT	Single Positron Emission Computed Tomography
STM	Short Term Memory
SVM	Support Vector Machines
WM	Working Memory

1. INTRODUCTION

The central goal of neuroscience is to investigate the human brain, which is known to be the most complex organ of the nervous system. The brain comprises billions of neurons that communicate with each other through electrical signals and chemical transmitters. This central organ organizes many different activities of the body. It receives all kinds of information from the body and makes decisions to maintain essential instructions. Scientists from various disciplines try to understand the brain processes principles to address many neurological diseases and behavioral problems.

There are various medical imaging techniques in modern technology to examine the human brain directly or indirectly. Anatomical brain imaging techniques as magnetic resonance imaging (MRI) and computer tomography (CT) provide us information about its structure. On the other hand, functional brain imaging techniques as positron emission tomography (PET), single photon emission computed tomography (SPECT), electro- and magneto-encephalography (EEG/MEG) enable us to know about the physiological status of the brain.

Of the above methods, EEG is the earliest technique which registers the brain electrical activity from the scalp electrodes. It is a powerful modality that measures the electrical activity resulting from ionic current changes in neurons. EEG has become one of the most widely techniques over the years for diagnosing neurological disorders such as epilepsy, sleep disorders, stroke and cognitive processes such as attention, memory, and emotion. Studying the networks in the brain using neuroimaging techniques help scientist discover how it functions.

EEG data is recorded by multiple electrodes placed on the scalp according to a standardization called the 10-20 electrode placement system, adopted by the International Federation in Electroencephalography and Clinical Neurophysiology [7]. Once EEG data has been recorded, there are numerous signal processing techniques used to extract relevant information. A mathematical branch called graph theory provides an ideal case for brain networks since it is not confined with the limitations of the Euclidean space as it represents a network as a graph consisting of nodes and edges.

Graph signal processing (GSP) is developed to identify and exploit the structural and functional relations of signals on graphs using classical signal processing techniques such as Fourier transfrom, filtering, and sampling. Two essential perspectives lie in graph signal processing. The first framework is rooted in the algebraic signal processing theory applicable to directed graphs. The second one is rooted in spectral graph theory that only uses undirected graph Laplacian matrix [8].

Algebraic signal processing, introduced in [9], leads to the use of weighted graph adjacency matrix as a shift matrix whereas spectral graph theory uses the Laplacian matrix of the underlying graph as its building block [6]. Since Laplacian is matrix symmetric and positive semi-definite, the Laplacian-based approach can only be applied to undirected graphs. Although both methods define traditional signal processing frameworks, difference in their origins gives rise to different techniques for signal processing on graphs.

The goal of this thesis is to implement GSP algorithms for preprocessing EEG signals for brain computer interfacing (BCI) and emotion estimation. We define graph Fourier transform, graph frequency, and filtering to classify graph signals recorded from motor/imagery and emotional EEG signals. The signals are first projected to temporal frequency subspace to extract information in different EEG bands, then to a subspace of eigenvectors of graph Laplacian before they are classified using support vector Machines (SVM) and K-nearest neighbors (KNN).

The organization of the thesis as follows: in Chapter 2, the EEG technique and the BCI systems are explained. Chapter 3 focuses on fundamental concepts in graph theory and its basic applications. In Chapter 4, graph signal processing tools and a brief revies of literature on GSP applications in various areas are given. Datasets used in the thesis are described and the results presented in Chapter 5. Discussions and

Conclusion are presented in Chapters 6 and 7, respectively.

2. ELECTROENCEPHALOGRAPY (EEG)

The complex network structure in human brain plays an essential role in processing motor or sensory stimuli can be observed with various imaging modalities such as functional MRI fMRI, CT, PET, MEG. Among these modalities the EEG has the advantage of having a high temporal resolution and inexpensive use of hardware compared with other measurement techniques in order to discover the underlying networks in the brain which active during the mental processes.

EEG measures the electrical potentials resulting from post-synaptic potentials in the neurons with the help of a number of electrodes on the scalp during neuronal excitations. EEG as a non-invasive and painless technique, is widely used to diagnose clinical problems such as epilepsy, sleep disorder, tumors, and stroke, as well as to assess cognitive processes such as perception, attention, memory, and emotion. In Figure 2.1, an example of an EEG recording can be seen.

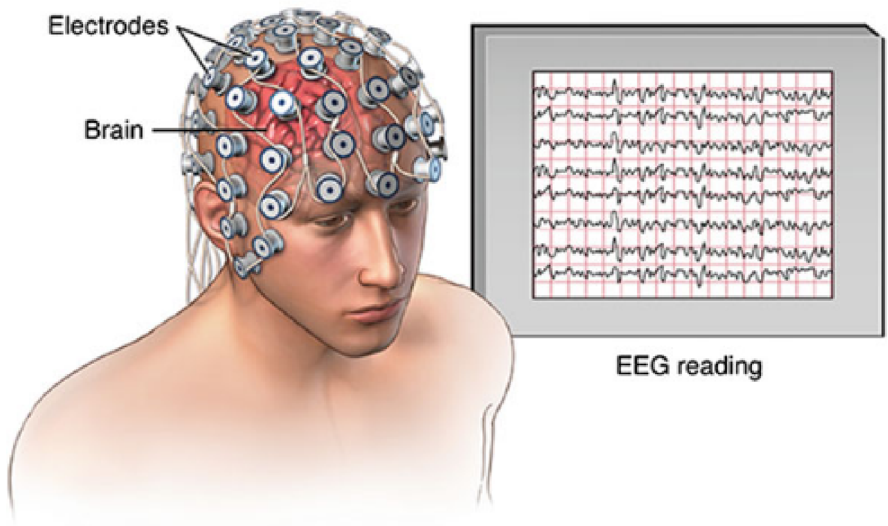

Figure 2.1 An example of EEG recording [1].

2.1 History

The first electrical activity was recorded in rabbits in 1875 by scientist Richard Caton (1842-1926) with the iad of a galvanometer's aid. Since then, the concepts of electro, encephalo, and graphy refer to the registration of brain electrical activities, r emit- ting the signals from the head), and drawing or writing, respectively. The terms were combined so that the term EEG was subsequently used to denote the brain's electrical neural activity [10]. In 1912, Vladimir Vladimirovich Prravdich-Neminsky, a Russian physiologist, published a study and showed the first EEG oscillations [11].

Hans Berger, a German neurologist, studied human EEGs with a string galvanometer, and he managed to record brain electrical activity from the human scalp. In 1929, this was the first report on human EEG recordings of one to three minutes duration on photographic paper. Berger used a bipolar recording technique with frontal-occipital leads for his one-channel EEG tracings, along with simultaneous electrocardiogram (ECG) recording and a time marker. He observed 8-12 Hz rhythm and named it alpha [12].

The gamma rhythm came in a paper in 1938 by Herbert Jasper [11]. Kornmüller recognized the potential of recordings by using a significant number of electrodes. In 1947, the first international EEG congress was held, which led to the American EEG society foundation. In the 1950s, sleep researches emerged up, and in 1958, REM sleep was described [12]. To sum up, the development of EEG studies started in early 19th century has been a continuous process that led to the improvement of clinical and computational studies in order to treat the number of neurological and physiological abnormalities of human beings.

2.2 Rhythm in Brain Network

The rhythms of the brain range from 0.1 Hz up to higher frequencies depending on wakefulness or sleep state which manifest waves of different characteristics. Since the electrical activity in different brain regions show a random behavior, it requires sophisticated methods to analyze the brain signals. These signals are distinguished into five prominent brain oscillations labeled in Greek letters for different frequency ranges. These oscillations are delta (0.1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-25/30 Hz), and gamma (30-100 Hz).

Delta wave, introduced by Walter in 1936 to designate all frequencies below alpha range [11], is the low-frequency signal seen in deep sleep especially in stages 3 and 4. Delta rhythm represents the frequency oscillations in the brain below 4Hz. With the help of EEG studies during sleep, this slow wave has been localized in many cortical areas such as the inferior frontal, medial prefrontal, precuneus, and posterior cingulate [13]. Delta rhythm is also important in case like brain injury or child atttention deficiency hyperactivity disorder (ADHD) [14].

Theta waves denote oscillations in the range from 4 to 8 Hz which can be observed while humans are in deep meditation, stress, or frustration. A theta rhythm is most prominent in childhood, but it can be seen in adults having pathological problems or depression as well. Theta rhythm enhances the production of human growth and serotonin hormone that increase relaxation [15]. Theta rhythms are related to various cognitive activities involving working memory and learning tasks [13].

Alpha rhythms are the first discovered EEG oscillations,also known as Berger's waves, and they lie between 8-12 Hz. They are observed in different cortical areas with different frequencies. The alpha activity is originated from the occipital lobe during rest while eyes closed, but it is also generated in other cortical areas such as the over frontal cortex, sensory-motor cortex, and the supplementary motor area. The amplitude of the alpha rhythm is enhanced during short-term (STM) and working memory (WM) tasks, whereas it is suppressed by visual stimuli [16].

Beta oscillations, vary within the range of 12-25 Hz, and have low amplitude and high frequency which can be observed in awake state. It is a waking rhythm that occurs in normal adults in conscious state, especially while humans focus on thinking, problem-solving, and decision-making [10]. A beta rhythm related to behavior and actions is more prominent in the frontal and parietal regions [16].

Gamma waves, also known as fast beta waves, lie within the range of 30-100 Hz and are related to perception and consciousness. Even though these waves rarely occur with low amplitude and high frequency, they can indicate mood disorders such as epilepsy and schizophrenia. A gamma wave can be detected in the somatosensory cortex, temporal and parietal cortical regions [15].

Although many EEG studies rely on five primary brain oscillation, there are other waveforms in the brain which can be categorized according to their morphological structure. These are Mu, K-complex, Lamda, V, spike waves, and spike spindles, known as sigma activity [15]. In this thesis, five brain waves are investigated that can be seen in the Figure 2.2.

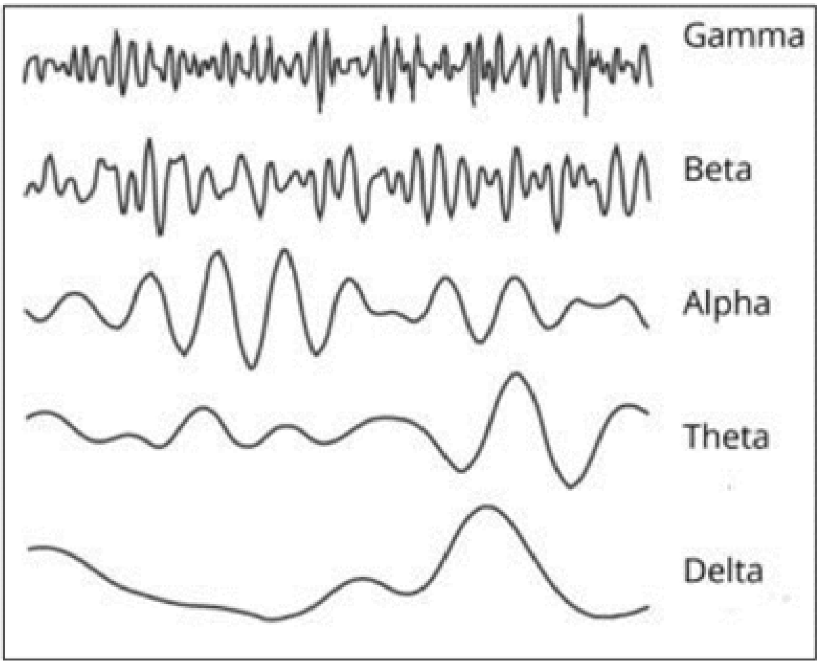

Figure 2.2 Five different brain waves.

2.3 Electrode Positioning

The electrode positioning is critical for recording high-quality and reliable data for interpretation. The International Federation of Societies for Electroencephalography and Clinical Neurophysiology adopted a 10-20 electrode placement system due to the reliability and comparability of EEG studies in 1958. (2). The "10" and "20" represent the proportion distance between electrodes. The electrode positions are shown, and the widely used 64 electrode system, according to the extended version of the International 10-20 system, is shown in Figure 2.3.

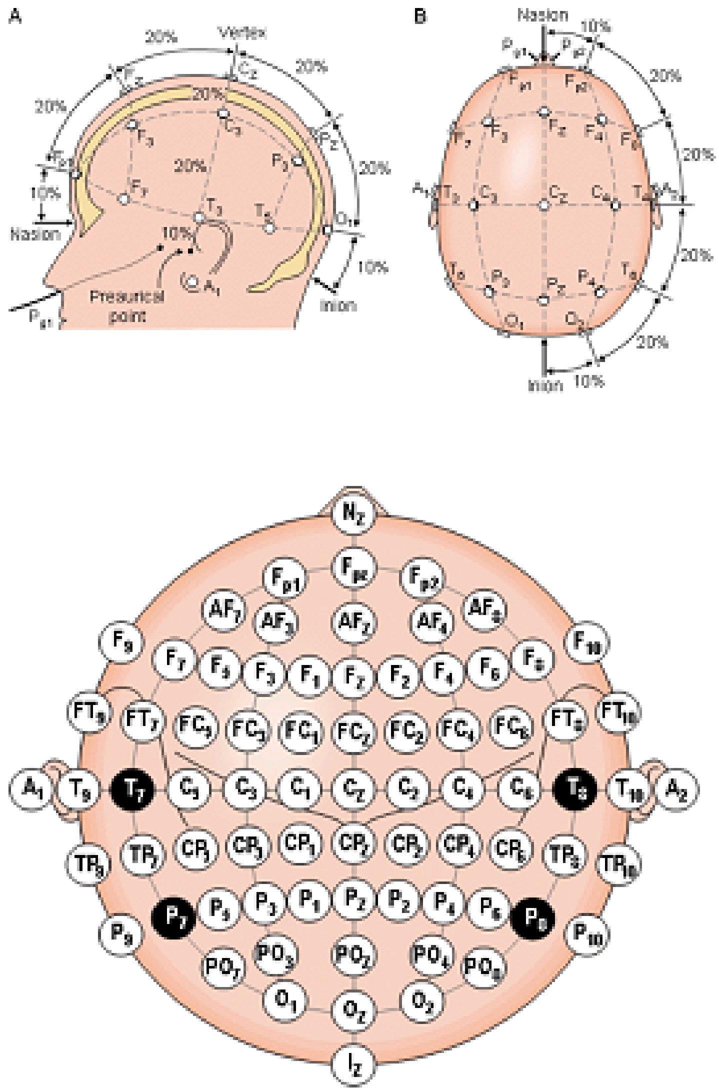

Figure 2.3 Electrodes positions according to International 10-20 system [2].

2.4 Abnormal EEG Patterns

Recognition of artifacts can prevent scientists from misdiagnosis. The raw EEG data may contain frequency components subjected to various contaminations, categorized into two chief categories: physiological (from a patient) and non-physiological (not from a patient) artifacts. Physiological artifacts originate from eye movement, hyperventilation, sweating, and cardiac activities, whereas non-physical artifacts, known as external artifacts, are caused by errors in the device and recording equipment generated from the subjects' environment.

2.4.1 Physiological Artifacts

The simultaneous recording of EEG consists of ocular artifacts such as eye blinking and eyeball movement. Since artifacts generate signals, one should recognize the nonsensical electrical activity to make an accurate interpretation. The eyeblink generates signals with a larger amplitude than cortical signals, and eye movement rotation causes significant external field variations that contaminate EEG readings [17]. Moreover, the peak amplitude of ocular artifacts is seen frontal region in humans. Hyperventilation or drowsiness may lead to affect the majority of electrodes and thus misdiagnosis.

The other causes of physiological artifacts are head movements, muscle movements, and cardiac activities [18] which increase the misreading of EEG signals. Muscle artifacts such as jaw and eyebrow movements cause artifacts spotted above 13 Hz [19]. In addition, cardiac activity may render EEG signals uninterpretable.

2.4.2 Non-Physiological Artifacts

In addition to physiological artifacts, non-physiological artifacts stem from both user environments and devices, which increases the difficulty of interpreting EEG recordings. These kind of contaminations may include electrode displacement, cable movement, and errors in recording equipment. The artifacts aggravate the problem of interpreting the signals and cleaning the EEG recordings requires to identify their characteristic features.

EEG signals are altered by artifacts that originate not only from physiological but also from some extrinsic artifacts. Artifacts contaminate the true EEG signals, so they can affect the results of classification accuracy that help diagnosis. Removal of these artifacts so as to obtain meaningful results needs appropriate procedure and planning. Different signal processing techniques such as ICA, PCA, regression-based method, and Empirical Mode Decomposition (EMD) analysis have been proposed to separate artifacts from valuable signals [20], [21].

2.5 Brain Computer Interface (BCI)

The brain-computer interface (BCI) is a system that interprets the commands from brain activity and sends commands to the physical world without controlling the peripheral nerves and muscles. BCI, also known as the brain machine interface (BMI), is an artificial intelligence system. The ultimate aim of the BCI is to allow people to interact with their physical environment using their brain activity which is essentially more important for those who are disabled.

BCI systems are inexpensive and user friendly. EEG is recorded from multiple sites over the scalp and information is analyzed using signal processing an machine learning algorithmss [22], [23]. The block diagram of a typical BCI system can be seen in Figure 2.4.

The BCI, which have various communication, entertainment, and medical applications, have portability, inexpensiveness, and ease of setting up. MEG and fMRI techniques are also used in BCI research but they are much more expensive and non transportable systems compared with EEG [24]. Although BCI is generally a non-invasive system, there are some invasive BCI applications for research purposes.

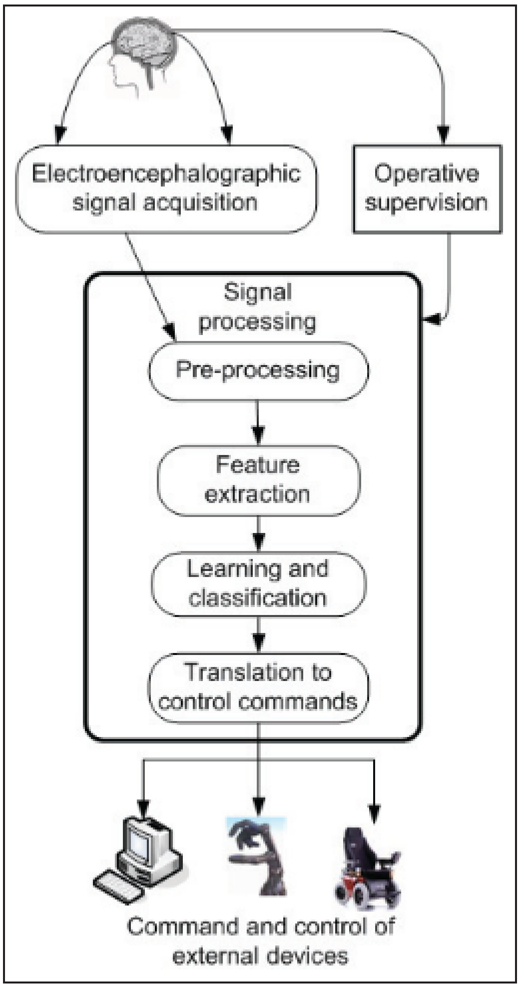

Figure 2.4 Typical block diagram of EEG based BCI [3].

Invasive BCI modalities include high risks due to their insertion of electrodes inside the grey matter [24]. Most of the research on invasive BCI was carried out on monkeys [10]. In this study [25], the monkey controlled prosthetic devices and modified motor functions using a small numbers of neurons. In another work [26], the monkey managed to control the robot arm.

Non invasive BCI generally uses EEG signals comprised of event-related desynchronization (ERD) and synchronization (ERS) events, as well as event related (ERP) and evoked potentials. ERD/ERS is a non-phased locked response, whereas ERP is a stimulus-locked response. P300 signal is the most widely used ERP, which can be visual, auditory, and somatosensory [10]. On the one hand, mu rhythm, recorded over the sensory motor cortex, which is given as an ERD example, is of particular interest in BCI. ERD starts over the contralateral rolandic region and, during the movement, becomes bilaterally symmetrical with the execution of an action [10].

As a communication tool, BCI guides those who do not use peripheral nerves and muscles so that they can deal with their needs without someone's help. For this reason, interactions must be effective between a user who sends electrical signals as inputs and the BCI systems which recognize inputs and makes inferences from them as output signals. As a tool, the BCI system can carry out particular functions; therefore, it has various applications in different areas such as neuromarketing, entertainment, detection, and diagnosis.

EEG-BCI systems have special features for detecting abnormalities associated with brain tumors, epilepsy, and sleep disorder. Brain tumors rise very rapidly, and if not treated, the possibility of survival can fall into danger. For this reason, early detection is vital for the sake of patients. Supervised methods such as Markov Random Field and Random Forest Classifiers can be used; however, deep learning methods are more robust [27] in analyzing tumor cells that have a strong relationship with high signal powers in the delta band [28].

An example of EEG-BCI application on detection, Epilepsy, defined as a hypersynchronous activity among neurons is a neurological disorder and affects %1 to %2 of the world population [27]. Although primary imaging techniques such as MEG and fMRI are used to detect seizures, EEG-based BCI is the most traditionally used approach for predicting epilepsy [10]. Various signal processing tools in the time or frequency domain and classification algorithms can be used for seizure detection. In this article [29], discrete wavelet transform with a classification algorithm called an artificial neural network is used to identify epilepsy. In [30], Support Vector Machines (SVM) were used detection of seizures.

Beyond medical applications, the BCIs system is used in entertainment-based applications, especially in video games such as Pinball, Pacman, and Pong [31]. In this study [32], subjects who used non-invasive BCI systems could fly a virtual helicopter to any point in 3D space. In another study [33], randomly, a labyrinth is created as a game called Pacman for non-invasive BCIs.

BCI signals can be used to trigger and enforce customers so that they persist in buying given products. The possibility of analyzing brain signals as responses to TV commercials are known as Neuromarketing and Neuropolicy and it plays a vital role for business and companies [31], [34], [35].

3. GRAPH THEORY

Complex systems whose specific structure is not clear to model show unpredictable characteristics among its' properties and environment. Such systems comprise many scientific disciplines ranging from economics, biology, and physics. Network science was asserted to explore the behavior and function of complex systems. As an academic field, network science uses nodes for representing the distinct properties lying under the complex networks and edges for expressing a relationship between systems elements. Network science takes benefit from one of the mathematical structures defined by graphs to understand the underlying mechanism of various chaotic systems.

The foundation of graph theory based on seven bridges of the Königsberg problem was introduced by Leonhard Euler (1707-1785). In 1735, a Swiss mathematician came up with the issue: whether it was possible to walk around via a route that crossed each bridge once and only once across the river Pregel which includes two large islands in the city of Königsberg, Prussia [4]. Euler simplified the problem as a binary graph and discovered the solution that, at most, two nodes must have an odd number of degrees.

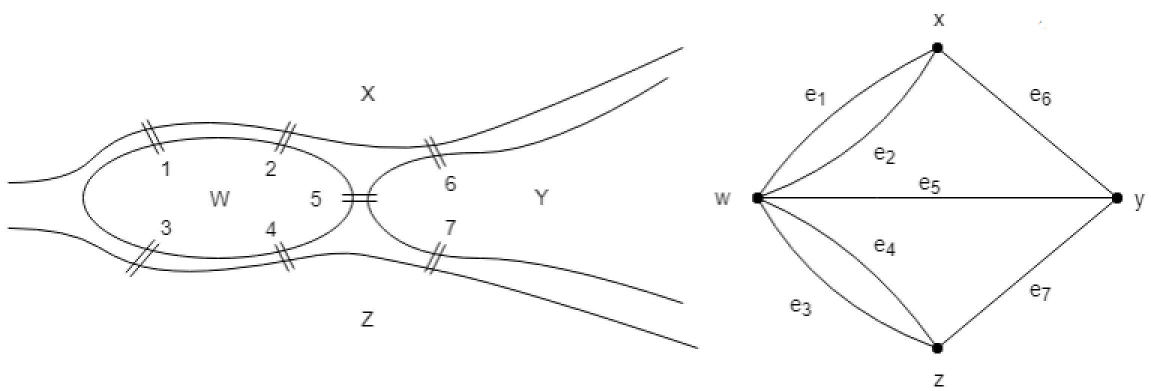

Figure 3.1 Euler's graphical representation of Königsberg bridge problem [4].

A graph, G, is a collection of a set of nodes (vertices) that are connected via edges. An edge defines a connection that can be weighted or directed between two nodes. Therefore, graphs are beneficial for representing data in various domains such as social networks, neuronal and sensor networks, and machine learning [6], [36]. They are also helpful for the representation and analysis of images and videos [37].

Graph G, of size N, consists of a non-empty and finite set of vertices $\boldsymbol{V} = \{v_1, v_2, ..., v_N\}$ and edge set $\boldsymbol{E} = \{e_1, e_2, ..., e_N\}$ and it is represented by $\boldsymbol{G} = (\boldsymbol{V},\boldsymbol{E})$. A graph can be weighted if its' all branches have weights, in contrast, it is called unweighted when the edges do not have a numerical value.

If a graph consists of a finite number of vertices, it is called a finite graph, and the size of a finite graph is the number of nodes. A graph can be called a directed: edges have a direction that specifies a specific route from node to node and undirected when there is no direction. An edge between (i, j) and (j, i) represent the same link between two adjacent vertices v_i and v_j.

We will be interested in undirected and connected graphs in this thesis. If there is an edge between two vertices in a graph, they are said to be adjacent. In this case, adjacent vertices are also called neighbors of each other, and we can define all adjacent vertices for given vertex $\boldsymbol{i}$ as :

$$N(i) = \{j \in \boldsymbol{V} \mid \{i, j\} \in E\} \tag{3.1}$$

A graph $\boldsymbol{G} = (\boldsymbol{V},\boldsymbol{E})$ is called connected if a path exists for any two vertices i≠j. That means there exists a subset of vertices $\{i_l\}_{l=0}^{k} \subseteq V$with $i_0 = i$ and $i_k = j$, such that $\{i_l - 1, i_l\} \in E$, $l = 1, ..., k$. An example of undirected, unweighted, and connected graph can be seen in Figure 3.2.

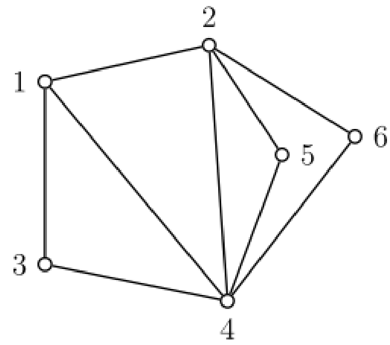

Figure 3.2 Example of an undirected and connected graph.

3.1 Graph Metrics

A graph can also be represented by an adjacency matrix known as incidence matrix. Adjacency matrix elements show whether a pair of nodes are neighbors or not. In the particular case of undirected graphs, the adjacency matrix is symmetric, and its components are comprised of 0 and 1. The adjacency matrix of graph G is defined. $\boldsymbol{G} = (\boldsymbol{V}, \boldsymbol{E})$ is defined as:

$$A_{ij} = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are adjacent.} \\ 0, & \text{otherwise} \end{cases} \tag{3.2}$$

The adjacency matrix of a given graph is n by n square matrix. Its vertices connection is crucial to represent the adjacency matrix, whereas positioning of node is not. The adjacency matrix of undirected and unweighted graph in Figure 3.2 is

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{pmatrix} \tag{3.3}$$

A degree matrix is a diagonal matrix whose elements represents the sum of adjacent edges of each vertex, such that

$$D_i = \sum_j A_{ij} \tag{3.4}$$

A degree matrix is a diagonal matrix whose elements represent the sum of adjacent edges of each vertex. A distance matrix, also called weighted adjacency matrix, is a square matrix whose elements represent the strength of the relation or distance between two nodes. If two vertices, i and j are neighbors, W_{ij} shows the weight of this edge.

A Laplacian matrix for simple graphs with n vertices is represented as:

$$L = D - A \tag{3.5}$$

where D is the diagonal degree matrix and A is the adjacency matrix of graph G. Laplacian matrix is symmetric and positive semi-definite matrix. [38]. Laplacian matrix is also defined as follows:

$$L(v_i, v_j) = \begin{cases} d, & \text{if } i = j. \\ 1, & \text{if } i \text{ and } j \text{ are adjacent.} \\ 0, & \text{otherwise} \end{cases} \tag{3.6}$$

Since the Laplacian matrix is symmetric, $L^T = L$, and positive semi-definite, it has orthogonal eigenvectors, and eigenvalues will be positive and real.Therefore, Singular Value Decomposition can be applied to the matrix L as:

$$L = Q\Lambda Q^T \tag{3.7}$$

where Λ is a diagonal matrix whose elements are positive and real eigenvalues, and U is the eigenvector matrix with the corresponding ordered eigenvalue $\lambda_0 \leq \lambda_1 \leq, ..., \leq \lambda_n$. If A is an m by m matrix, we know from linear algebra, it has m eigenvalues, which can be complex. On the other hand, if A is a symmetric matrix, then all m eigenvalues are real. For any complex number $z = a + ib$, where a and b are real numbers, its conjugate is $\bar{z} = a - ib$. With the same logic, for any vector v, its conjugate is defined as v^*. Therefore, we can write $vv^* = \|v^2\|$.

3.2 Graph Applications

Since the beginning of graph theory, graphs have been utilized to model relations and processes in numerous areas such as chemistry, computer science, biology, material science, computational neuroscience, and mathematics associated with real-world problems. Besides helping to demonstrate features of systems underlying these problems, graphs have also received attention in reducing complexity so that problems can be translated to an application that guides them with ease.

The combination of chemistry and graph theory constitutes the chemical graph theory applied to indicate the physical characteristics of molecules and their geometric structure. Graph vertices symbolize molecules, atoms, or electrons, whereas edges represent the covalent bonds. Graph theory can be used in chemistry in the phenomenon of isomerism [39], estimating the polarizability of a variety of polymers [40], predicting the ecotoxicity of chemical substances [41], and chemical kinetics [42].

Mathematical modeling has been subjected to interpretation for biological systems. Mathematical frameworks have been utilized in the study of physical problems, especially in the analysis of genes. RNA structure, RNA's recurring modular units, and the prediction of three-dimensional RNA folds have been investigated using mathematical graph modeling [43]. Graph theoretical approaches allow the representing RNA structure and discovering RNA topologies [44]. Graph theory has also been pioneered to capture critical elements in determining RNA secondary structures [45].

Graph theory provides parameters that help to construct networks so that organization of cognitive functions related to brain network structure can be used to detect brain dysfunction. Graphs are powerful tools to evaluate functional abnormalities of a human brain by examining both brain functional connectivity and structural connectivity investigated by EEG or fMRI analysis. Graph theory studies on brain diseases such as multiple sclerosis (MS) [46], autism spectrum disorder (ASD) [47], generalized anxiety disorder (GAD) [48], acute stroke [49], and late-life depression (LLD) [50], are used for characterizing brain organization.

The relevance of graph theory in the field of computer science can be seen in designing a database, representing networks of communication, image segmentation, software engineering, clustering [51]. For instance, planning routes for mail delivery, scheduling and assignments, and famous traveling salesman problems are addressed with graph theory applications in computer science [52].

Other application area of graph theory is in the electrical engineering to solve the electrical network problems. Current and voltage variable values can be solved using digraphs in a given network [53]. The other widespread application of digraphs is industrial engineering, including scheduling activities and different organizational problems.

Graph theory has been used in various fields due to its numerous features, such as balancing, modeling, decision-making ability, and showing the relationship among objects [51]. The other graph application studies include the following air pollution result from opencast mine [54], security services within IoT network [55], coding theory, X-ray crystallography [56], linguistics [52], traffic network design [51].

4. GRAPH SIGNAL PROCESSING

Data analysis has been growing in every aspect of human life, accompanied by different application domains, including social networks, transportation networks, financial and banking data, internet, economic network, and brain network. In many ways, these data are analyzed by traditional and ad-hoc methods, leading to superficial and non-systematic conclusions. There is thus an emerging field to cope with such high dimensional data that utilize graph theory, called graph signal processing (GSP) [6].

Signals on graphs require for extending the traditional signal processing concepts such as Fourier Transform, filtering, frequency analysis, smoothing, so that they would be utilized to construct graph models of various networks.

GSP which is the generalization of classical signal processing over graphs has two basic perspectives rooted in algebraic signal processing theory and spectral graph theory. The former approach provides us to use a digital signal processing framework by considering graph adjacency matrices as the graph shift operator [36], [5], [57]. In contrast, the latter led to using the undirected graph Laplacian matrix to represent data and its eigen spectrum for data processing [6], [58], [59].

In the algebraic signal processing (ASP) approach, concepts such as filtering, frequency response, Fourier Transform, and convolution are developed using z- transform.In the Laplacian-based framework, eigendecomposition of the graph Laplacian matrix constitutes graph spectrum. Both approaches utilize fundamental signal processing concepts.

This thesis focuses on the Laplacian-based approach that came out from manifold discovery research [60]. The graph Laplacian operator, the negative second-order derivative operator, helps construct a spectral projector-based graph signal by benefitting from relations between geodesic distance and graph distance on a manifold. In this case, the graph Laplacian operator is seen as a discrete approximation to the continuous Laplace-Beltrami operator on a manifold; consequently, a graph can be seen as a continuous manifold [61].

As for the high dimensional data processing on irregular data domains, the question of what is the most efficient way to extract information from these data comes to mind. In the following sections, some notations and traditional signal processing operations using graph Laplacian matrix are described.

4.1 Graph Signal

In traditional signal processing, data analysis includes operations as convolution, filtering, downsampling, Fourier transform. As for the data on irregular domains, how do we perform these operations, and what is the best strategy to extract information that we want to take. How do we translate a signal over a graph and what is the notation for it?

We are interested in analyzing high-dimensional datasets on undirected graphs, assuming that signals on a graph are indexed by the vertices of graph $\boldsymbol{G} = (\boldsymbol{V}, \boldsymbol{E})$. In other words, a graph signal will be represented as a function defined over the finite set of vertices with $\mid V \mid = N$

Using a graph G with n nodes, a graph signal or function s: $\boldsymbol{V} \longrightarrow \mathrm{I\!R}$ is defined as a map represented as a vector, where the i^{th} element s_i is indexed by node v_i of a given set of the vertex. Let write a graph signal as:

$$s = \Big[s_1, s_2, ..., s_N\Big]^T \in \mathrm{I\!R} \tag{4.1}$$

where each signal s_i ismorphic to $\mathrm{I\!R}^N$, and dimension of s is dim $S = N$. In Figure 4.1, graph signals in different datasets can be seen.

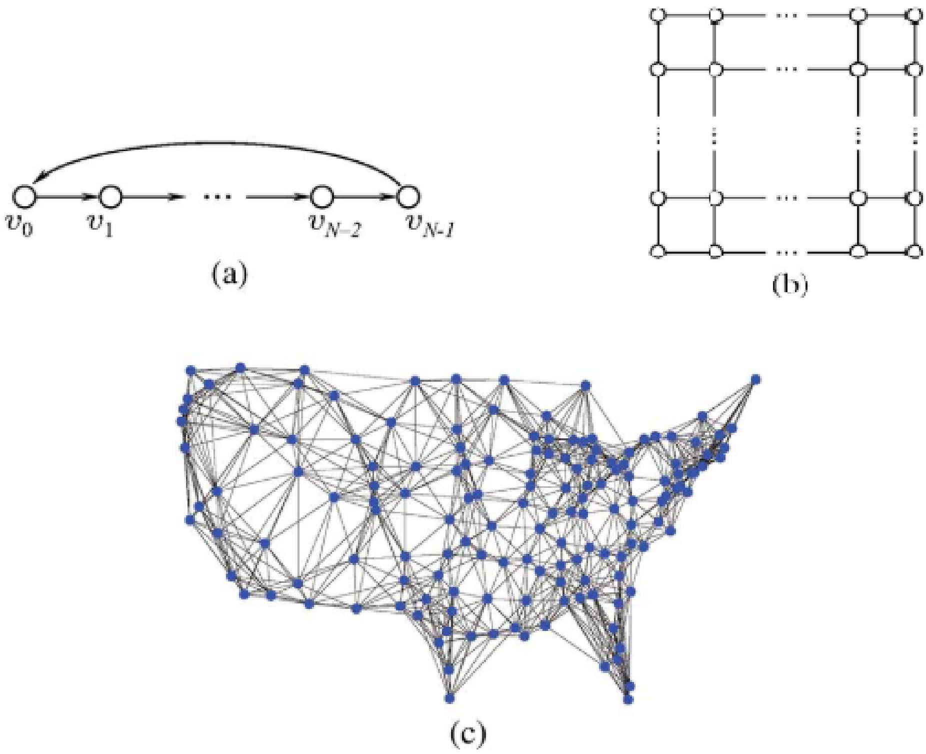

Figure 4.1 Examples of graph signals in different data sets [5].

4.2 Graph Laplacian

We consider undirected and unweighted graphs $\boldsymbol{G} = (\boldsymbol{V}, \boldsymbol{E}, \boldsymbol{A})$, where V is the set of vertices, E is a set of edges, and A is a graph adjacency matrix. The graph Laplacian, the difference operator, is defined as:

$$L(G) \stackrel{\text{def}}{=} D(G) - A(G) \tag{4.2}$$

where D is the diagonal matrix whose elements d_i represents geodesic distance between two points. As a result, Laplacian L is a symmetric and positive semidefinite matrix, and admits the eigendecomposition. The eigendecomposition of Laplacian matrix is defined as :

$$L = Q \Lambda Q^T \tag{4.3}$$

where $\lambda_1, ..., \lambda_n$ are the eigenvalues of given matrix. Because the eigenvalues of a symmetric matrix are real, we can order them as: $\lambda_0 \leq \lambda_1 \leq, ..., \leq \lambda_{n-1}$, so we can form the Laplacian spectrum by $\delta := \lambda_0, \lambda_1, ..., \lambda_{n-1}$.

4.3 Graph Fourier Transform and Notion of Frequency

How do we go from a combinatorial Laplacian matrix to a notion of frequency? Mathematically speaking, Fourier analysis is defined as the expansion of signals using an orthogonal set of functions $\sin(2\pi ft)$ and $\cos(2\pi ft)$. Fourier analysis has been one of the most beneficial methods in traditional signal processing to analyze signals and to decompose them. Just as the set of basis vectors captures the notion of frequency in time-domain the eigenfunctions of the Laplacian identify the frequency in the graph domain. A low-frequency graph signal varies slowly across its neighbors, whereas a high-frequency signal varies rapidly with respect to its neighbors.

Returning to the undirected graph Laplacian spectrum, the definition of graph fourier transform is given in Eq. 4.4. graph Fourier transform is the generalization of classical Fourier transform in the time domain given by Eq. 4.5. Moreover, the inverse graph Fourier transform is given in Eq. 4.6.

$$\hat{f}(\lambda_i) := \langle s, Q_i \rangle = \sum_{j=1}^{N} s(j) Q_i^T(j) \tag{4.4}$$

$$\hat{f}(\xi) := \langle s, e^{2\pi i \xi t} \rangle = \int_{\mathbb{R}} f(t) e^{-2\pi i \xi t} dt \tag{4.5}$$

$$f(j) = \sum_{i=1}^{N} \hat{f}(\lambda_i) Q_i(j) \tag{4.6}$$

The vector **s** that is defining a map represents the signal values at each vertex. Laplacian eigenvector associated with the smallest eigenvalue is a constant, and eigenvalues related to low frequencies correspond to smoother over the graph structure [8]. Furthermore, λ_i provides notion of frequency in the graph Laplacian spectrum.

Once a graph representation of data is built, eigenvectors of L allow us to define the frequency filtering of an input signal. In classical signal processing, we can filter the signal via the $\hat{h}$ transfer function defined in Eq. 4.7.

$$\hat{f}_{out}(\xi) = \hat{f}_{in}(\xi)\hat{h}(\xi), \tag{4.7}$$

We can generalize the filtering idea to graph spectral filtering as:

$$\hat{f}_{out}(\lambda_i) = \hat{f}_{in}(\lambda_i)\hat{h}(\lambda_i), \tag{4.8}$$

and taking an inverse graph Fourier transform we get Eq. 4.9

$$f_{out}(j) = \sum_{j=0}^{N-1} \hat{f}_{in}(\lambda_i)\hat{h}(\lambda_i)u_j(j) \tag{4.9}$$

In order to take a specific frequency window, low pass or high pass filter can be defined as:

$$\hat{h}_{Lr} = \mathbf{1}\ [r \leq R_L] \tag{4.10}$$

$$\hat{h}_{Hr} = \mathbf{1}\ [R_L \leq r] \tag{4.11}$$

where R_L corresponds to the cut-off frequency.

In summary, graph Laplacian matrix encoding connections among the nodes of the underlying graph is used to define the notion of frequency and graph Fourier transform. It is essential to keep in mind that smoothness and the spectral content of a graph are contingent on the underlying graph. For instance, in Figure 4.2, the same graph signal is represented with the same set of vertices, but with different edges (different undirected graphs) to show the significance of graph setting on eigenspectrum.

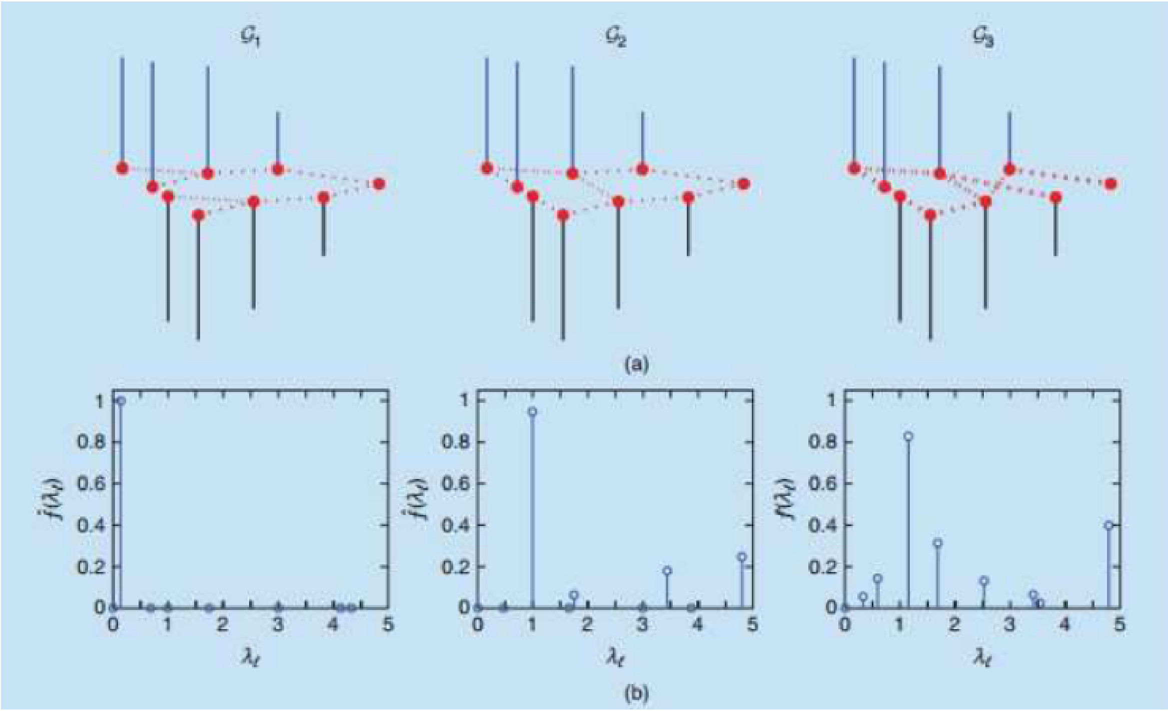

Figure 4.2 Importance of underlying graphs [6].

4.4 Graph Signal Processing Applications

In this section, we give some examples from set of application domains where GSP has been utilized. Additionally, we discuss both datasets that naturally exhibit irregular structures and more conventional datasets where traditional signal processing techniques are widely implemented.

One of the irregular data domains that are subjected to GSP is sensor networks. A graph helps to represent the positions of sensors via a set of nodes, and edges represent the distance between sensors. GSP application includes reconstruction, implementation, or compression of sensor data. In [62], [63], [64] one can find the graph-based analysis of sensor networks. Additionally, another scenario on graph-based data analysis is given by urban data processing. GSP can be used to analyze and model urban traffic problems [65], [66] air pollution [67], and power consumption [68].

Another popular application of GSP includes biological network studies such as the human brain. Graphs represent a map of human brain activity by indicating each node as a brain region while connections among nodes are determined by functional coherence [69]. Furthermore, it is shown that there is a relation between the spectral properties of brain networks and different tasks [70].

GSP is not only used for brain connectivity but also for the classification of brain signals [71]. It also offers promising solutions in the analysis of anomalies and some diseases, such as Alzheimer's. GSP is an excellent framework to cover the unknown connections in biological data. Moreover, gene regulatory [72] and protein interaction networks [73] can be addressed with the help of GSP tools.

Images and visual modalities are also subjected to GSP analysis. GSP tools provide more powerful operators the filtering [74] and graph representation of images and 3D meshes. For instance, graph Fourier transform can be used to adapt different types of image blocks and compressed smooth images [75].

Graphs have an indispensable position in machine learning applications because of their capability to represent the structure of a dataset. Therefore, GSP provides different signal processing operations that help to classify signals and construct architectures. Once the signal values are represented by graph labels, the graph signal processing techniques can be used to predict unknown labels [8], data clustering [76], and semi-supervised learning problems [77].

5. DATA ANALYSIS WITH GRAPH SIGNAL PROCESSING

5.1 EEG Emotion Dataset

EEG data were obtained from the Electrophysiology Laboratory at the Istanbul Medical School in Istanbul University. 13 healthy undergraduate/graduate students, whose mean age was 27.4 ($\pm$2.96), joined the experiment. All participants were informed and their written consent was taken for participation the experiment. ERP data were recorded from 30 channels with a 250 Hz sampling frequency. Approximately 280 pleasant and unpleasant pictures were chosen from the IAPS dataset with mean valence (7.13/2.96), and mean arousal level (4.99/5.02). Each image stayed on the monitor for one second, and the inter-stimulus interval was two seconds.

Figure 5.1 Example of pleasant and unpleasant pictures.

Emotion EEG data are recorded from 30 channels whose labels and locations in Figure 5.2.

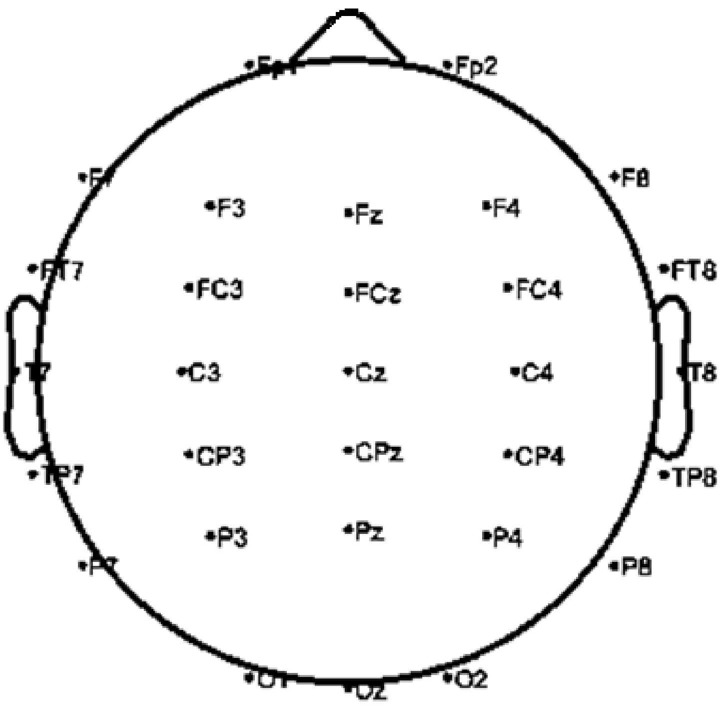

Figure 5.2 30 EEG channels.

5.1.1 Data Processing

Each subject data were decomposed into five brain waves using a temporal filter. After this preprocessing step, we separated data into train and test sets to measure classification rates using SVM, KNN, and GSP algorithms. For each subject, we chosed equal number of trials and we have $X_1 \in \mathbb{R}^{270x30x250}$ and $X_2 \in \mathbb{R}^{270x30x250}$ data structures for pleasant and unpleasant conditions, respectively. Data dimensions define *trial* $\times$ *electrode* $\times$ *time*. 150 trials for training and 50 trials for testing were chosen randomly from each data set to construct data for each experiment. Consequently, $X_{train} \in \mathbb{R}^{300x30x250}$ (%75) and $X_{test} \in \mathbb{R}^{100x30x250}$ (%25) portions of data became two groups whose binary output $Y \in \{1, 0\}^{100\times 1}$ were determined depending on which group they belonged.

Fourier transform based filtering into delta, theta, alpha, beta, and gamma bands project the data into temporal frequency subspaces. After filtering, SVM algorithm with radial basis function (RBF) as kernel and KNN algorithm with Euclidean and Chebyshev distances in the order of $k = 1$, $k = 3$, and $k = 5$ were applied. The data were also projected to a subspace determined by the eigenvalues and the graph eigenvectors of graph Laplacian such that a maximum separation between two conditions was achieved using the classification algorithms.

Before the GSP preprocessing analysis, the 3D electrode locations were projected into a two-dimensional subspace using their spherical coordinates. The underlying graph was constituted with the help of using the exact location of electrodes as a node. The adjacency matrix was created by assuming that a node has an edge with its closest neighbor. In addition, the spherical distance among each pair of nodes formed the degree matrix.

MATLAB 2019b was used for the analysis. MATLAB functions `fitcsvm` and `fitcknn` are used to train the data for the SVM and KNN, respectively. The randomization experiment was repeated 100 times for each subject and for each frequency band with the same proportions of trials, %75 for training and %25 for testing.

5.2 EEG Motor/Imagery Dataset

The second EEG dataset was obtained from PhysioNet [78], created by BCI developers [79]. 109 healthy subjects joined the experiment, and they performed 14 different motor/imagery experimental runs while 64 electrodes were recording via the BCI2000 system. Electrode locations can be seen in Figure 5.3. Each run sampled at 160 Hz frequency contained two one minute baseline runs (one with eyes open and the other one eyes closed), and four two-minute tasks, which are:

1. Opening and closing right or left fist with respect to target that appears on either right or left side of the screen and then relaxing when the target disappears.

2. Imagining to open and close right or left fist based on the target that appears on the corresponding side of the screen and then relaxing when the target disappears

3. Opening and closing both fists (if the target is on top) and both feet (if the target is on the bottom) based on the target that appears either the top or bottom of the screen and then relaxing when the target disappears.

4. Imagining to open and close both fists (if the target is on top) and feet (if the target is on the bottom) based on the target that appears either the top or bottom of the screen and then relaxing when the target disappears.

EEG motor imagery data was recorded using a 10-20 International Electrode system with 64 electrodes excluding Nz, F9, F10, FT9, FT10, A1, A2, TP9, TP10, P9, and P10 electrodes.

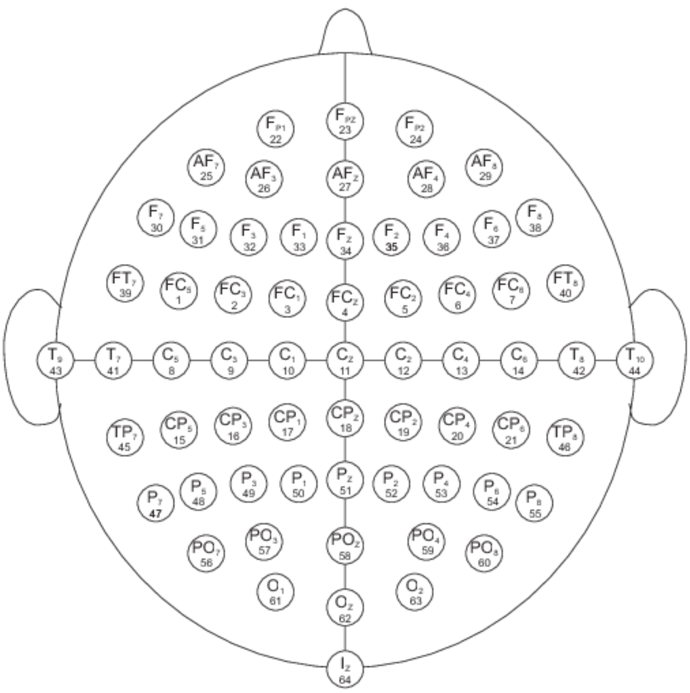

Figure 5.3 64 EEG electrode locations.

In our study, we worked with the right and left fist classification. In other words, task 1 and task 2 were compounded into one matrix. For each of the 109 subjects, the data were decomposed into five brain waves using a temporal filter. After this preprocessing step, we separated data into train and test sets to measure classification rates concerning SVM, KNN, and GSP algorithms. For each class, we had $X \in {\rm I\!R}^{90x64x656}$ data structures. Data dimensions define $trial \times electrode \times time$. For each subject, data ere split into $X_{train} \in {\rm I\!R}^{33x64x656}$ (%75) and $X_{test} \in {\rm I\!R}^{11x64x656}$ (%25) and they were associated with their binary outputs $Y \in \{1,0\}^{22\times1}$ before determining the classification algorithm performance.

The same data processing as described in Section 5.1.1 was repeated for the data motor imagery dataset.

6. RESULTS AND DISCUSSION

GSP framework for processing of brain network allows us to analyze signals from multiple sensors. Signal noise is discriminated by Graph Fourier Transform that depends on graph representation. In particular, GSP generalizes the elementary operator in signal processing. These operators such as Fourier transform, filtering, and sampling underlie the graph signal processing framework.

We reviewed Graph Fourier transform, frequency and filtering operations based on the Laplacian matrix. However, many applications on a graph, such as a vertex domain design [80], wavelet transform design [81], sampling [82], and denoising [83], have been introduced.

EEG emotion and EEG motor/ imagery data were classified in both graph domain and time domain. First of all, the emotion data was subjected to temporal filtering and divided into five frequency bands called delta, theta, alpha, and beta. Two of the most widely used machine learning algorithms i.e. the SVM and KNN used for classification. Results showed that the performance rate was around 60% for the delta band (0-4 Hz) and lower than 60% for other bands in SVM. Moreover, it was observed that the performance rate could not achieve a 60% accuracy for any band in KNN. In the second part of EEG emotion data analysis, data were subjected to a spatial filter defined by the underlying data. Results showed that classification accuracy increased to around 90% when the graph spectral decomposition was applied before classification.

In the analysis of EEG Motor/Imagery data analysis, data was first subjected to temporal filtering and divided into five frequency bands. Results showed that the performance rate was around 66% for the delta band and lower for other bands in SVM. Moreover, it was observed that the performance rate could not achieve a 60% accuracy for any band in KNN. In the second part of EEG motor/imager data analysis, data were subjected to a spatial filtering defined by the underlying graph. Results showed that the classification accuracy increased to 100% when the graph spectral decomposition was applied before classification. There is possibility that method may overfit due to limited data, that's why data will reinvestigate.

Table 6.1
The average results of EEG emotion data, taken from 13 subjects using SVM.

Method	0-4Hz	4-8Hz	8-12Hz	12-25Hz	25-30Hz
SVM	60.78%	58.15%	56.54%	55.25%	56.37%
GSP&SVM	56.27%	69.03%	84.61%	86.96%	85.40%

Table 6.2
The average results of EEG emotion data, taken from 13 subjects, using KNN with Euclidean and Chebychev distance with order of k=1, 3, and 5.

KNN	k=1			
	0-4	4-8	8-12	12-25
Euclidean Distance	55.96%	54.14%	53.17%	51.81%
Chebychev Distance	56.24%	51.05%	53.54%	49.69%

KNN	k=3			
	0-4	4-8	8-12	12-25
Euclidean Distance	54.98%	53.61%	53.55%	52.46%
Chebychev Distance	56.47%	52.36%	52.88%	47.92%

KNN	k=5			
	0-4	4-8	8-12	12-25
Euclidean Distance	56.02%	53.89%	54.08%	52.22%
Chebychev Distance	55.09%	52.49%	52.22%	47.54%

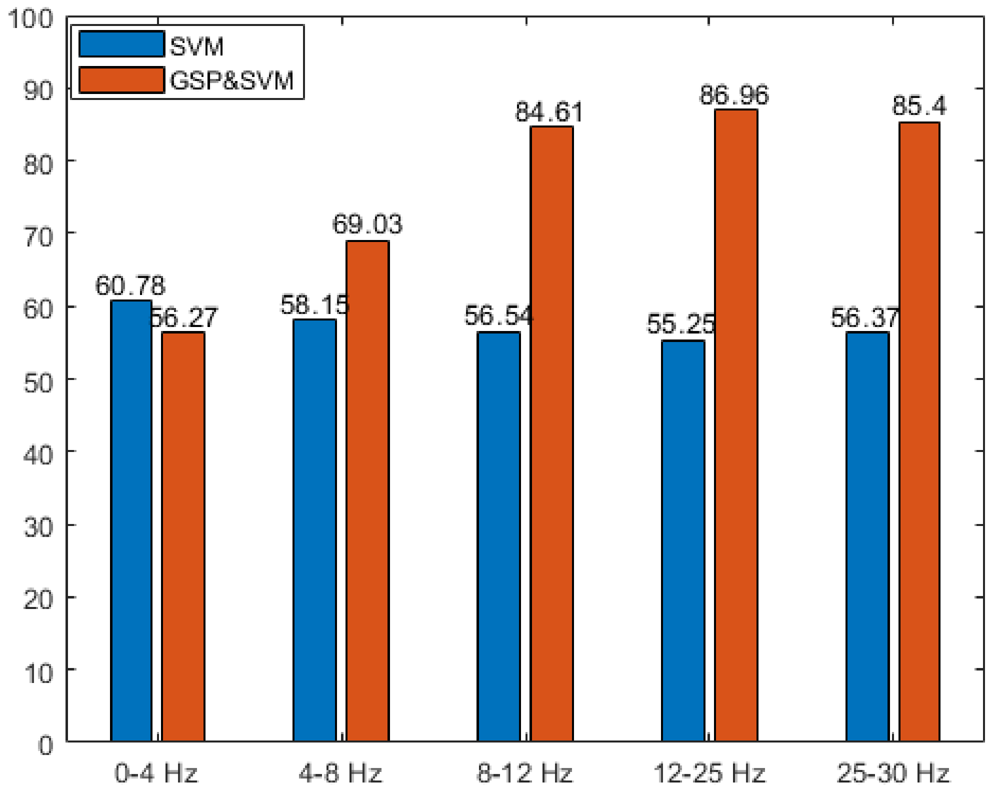

Figure 6.1 EEG Emotion data analysis results.

Table 6.3
The table shows that average results of EEG emotion data, taken from 13 subjects, using GSP & KNN with Euclidean and Chebychev distance with order of k=1, 3, and 5.

GSP & KNN	k=1			
	0-4	4-8	8-12	12-25
Euclidean Distance	56.08%	78.15%	80.96%	81.60%
Chebychev Distance	56.03%	61.22%	66.67%	66.29%

GSP & KNN	k=3			
	0-4	4-8	8-12	12-25
Euclidean Distance	56.92%	77.33%	81.71%	77.41%
Chebychev Distance	56.47%	62.03%	68.58%	68.12%

GSP & KNN	k=5			
	0-4	4-8	8-12	12-25
Euclidean Distance	54.73%	77.03%	80.8%	76.79%
Chebychev Distance	55.90%	63.14%	69.26%	67.29%

Table 6.4
The results of EEG emotion data, taken from 13 subjects, using GSP & SVM.

GSP & SVM Results				
Subjects	6-9 Hz	9-11 Hz	10-12 Hz	12-15 Hz
Subject 1	97.36 %	96.25 %	84.50 %	97.35 %
Subject 2	95.58 %	86.56 %	77.49 %	79.85 %
Subject 3	70.06 %	64.2 %	57.35 %	76.65 %
Subject 4	71.64 %	58.59 %	46.74 %	57.43 %
Subject 5	99.50 %	96.47 %	95.94 %	99.62 %
Subject 6	99.39 %	98.28 %	61.58 %	79.77 %
Subject 7	97.69 %	93.39 %	84.82 %	69.7 %
Subject 8	85.68 %	90.53 %	84.65 %	92.25 %
Subject 9	86.47 %	88.51 %	71.68 %	72.75 %
Subject 10	87.38 %	97.68 %	74.0 %	85.68 %
Subject 11	89.51 %	97.59 %	98.66 %	86.55 %
Subject 12	87.51 %	83.40 %	63.68 %	73.74 %
Subject 13	87.59 %	99.57 %	97.45 %	94.8 %
Average	86.81 %	88.51 %	76.81 %	86.02 %

Table 6.5
The results of EEG emotion data, taken from 13 subjects, using GSP & SVM.

GSP & SVM Results				
Subjects	15-17 Hz	17-20 Hz	18-20 Hz	19-21 Hz
Subject 1	99.39 %	99.25 %	96.16 %	96.67 %
Subject 2	97.27 %	82.18 %	75.63 %	95.58 %
Subject 3	88.91 %	84.28 %	81.54 %	53.15 %
Subject 4	54.32 %	86.66 %	65.44 %	50.09 %
Subject 5	96.25 %	94.59 %	97.81 %	97.45 %
Subject 6	99.50 %	99.40 %	96.97 %	99.75 %
Subject 7	77.46 %	92.91 %	94.37 %	98.72 %
Subject 8	82.73 %	83.78 %	85.35 %	91.38 %
Subject 9	79.20 %	72.42 %	53.79 %	74.00 %
Subject 10	93.31 %	99.75 %	85.08 %	88.78 %
Subject 11	86.67 %	95.96 %	92.74 %	92.73 %
Subject 12	95.37 %	92.36 %	73.70 %	91.9 %
Subject 13	96.79 %	89.93 %	98.42 %	97.78 %
Average	88.27 %	90.26 %	84.38 %	86.76 %

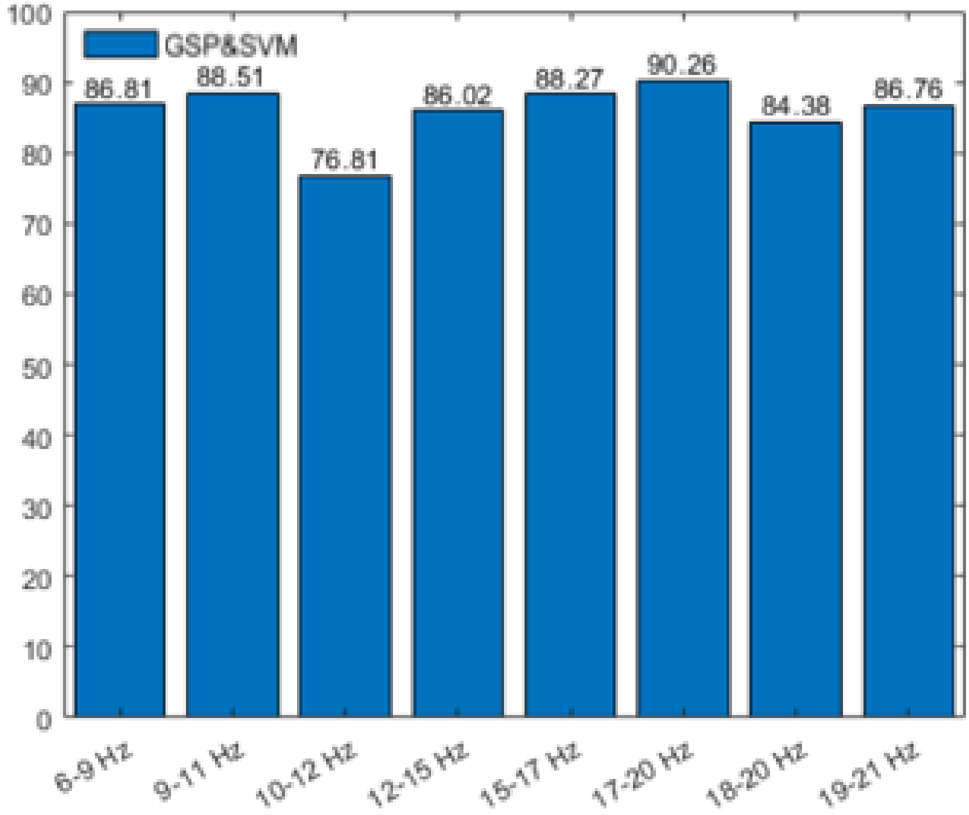

Figure 6.2 EEG Emotion data analysis results in various subspaces projection determined by graph spectral components.

Table 6.6
The results of EEG Motor/Imagery data, taken from 109 subjects,using KNN with Euclidean and Chebychev distance with order k=1, 3, and 5.

KNN	k=1			
	0-4	4-8	8-12	12-25
Euclidean Distance	50.25%	51.30%	51.36%	51.51%
Chebychev Distance	51.93%	51.04%	50.26%	49.86%

KNN	k=3			
	0-4	4-8	8-12	12-25
Euclidean Distance	55.37%	50.02%	51.78%	50.98%
Chebychev Distance	46.67%	46.18%	45.94%	45.49%

KNN	k=5			
	0-4	4-8	8-12	12-25
Euclidean Distance	57.01%	51.30%	52.05%	50.30%
Chebychev Distance	49.27%	47.85%	44.96%	49.94%

Table 6.7
The results of EEG Motor/Imagery data, taken from 109 subjects, using SVM.

Method	0-4Hz	4-8Hz	8-12Hz	12-25Hz
SVM	66.43%	51.66%	50.82%	43.05%
GSP&SVM	100%	100%	100%	100%

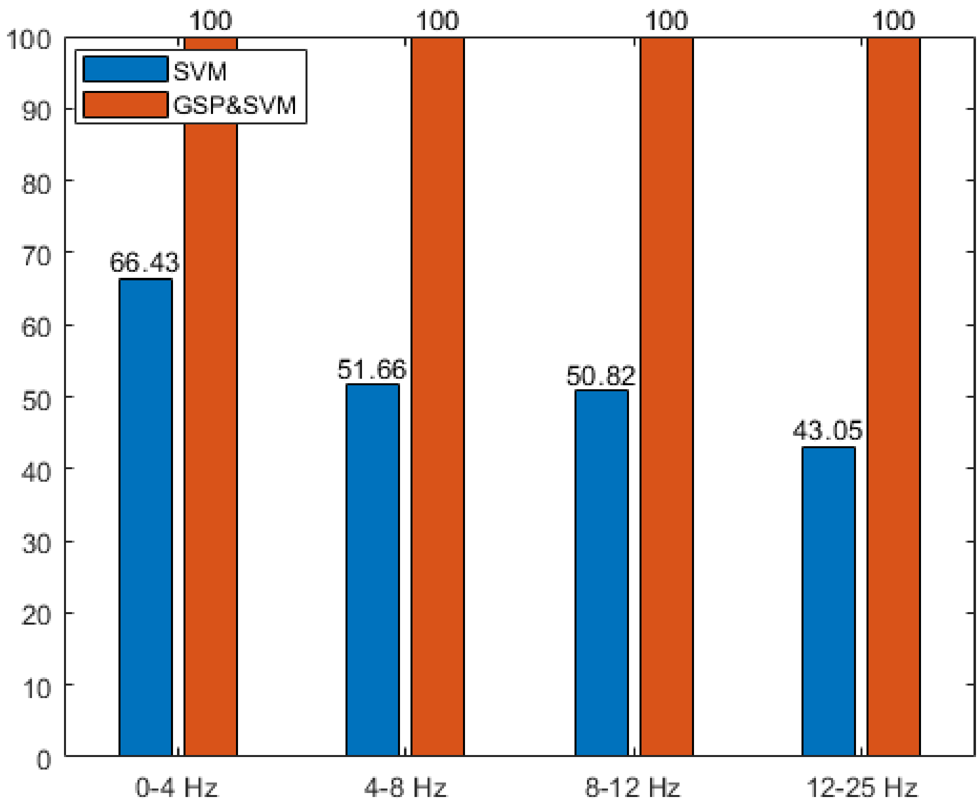

Figure 6.3 EEG Motor/Imagery data analysis results.

7. CONCLUSION AND FUTURE WORK

In this thesis, we presented some GSP notions in analysing high dimensional data by taking intuition from Euclidean spaces. We compared spatial and temporal brain network features using an undirected and unweighted graph. We proposed a spectral projector-based graph signal processing analysis for two different EEG data. We can conclude that graph spectral components carry essential information on data classification.

For further research, analyses can be done by using an algebraic signal processing framework. It can be determined whether the time shift matrix affects graph signal properties. Most recent works assumed that a graph is given; however, a graph can be constructed based on a statistical model. In addition, other data types, including fMRI, can be used to measure the performance of GSP algorithm.

REFERENCES

1. Tatum IV, W. O., *Handbook of EEG Interpretation*, Demos Medical Publishing, 2014.

2. Marcuse, L. V., M. C. Fields, and J. J. Yoo, *Rowan's Primer of EEG E-Book*, Elsevier Health Sciences, 2015.

3. Fazel-Rezai, R., *Recent Advances in Brain-Computer Interface Systems*, BoD–Books on Demand, 2011.

4. Euler, L., "Solutio problematis ad geometriam situs pertinentis," *Commentarii Academiae Scientiarum Petropolitanae*, pp. 128–140, 1741.

5. Sandryhaila, A., and J. M. Moura, "Discrete signal processing on graphs," *IEEE Transactions on Signal Processing*, Vol. 61, no. 7, pp. 1644–1656, 2013.

6. Shuman, D. I., S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Processing Magazine*, Vol. 30, no. 3, pp. 83–98, 2013.

7. Jasper, H. H., "The ten-twenty electrode system of the international federation," *Electroencephalogr. Clin. Neurophysiol.*, Vol. 10, pp. 370–375, 1958.

8. Sandryhaila, A., and J. M. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Transactions on Signal Processing*, Vol. 62, no. 12, pp. 3042–3054, 2014.

9. Püschel, M., and J. M. Moura, "Algebraic signal processing theory," *arXiv preprint cs/0612077*, 2006.

10. Sanei, S., and J. A. Chambers, *EEG signal processing*, John Wiley & Sons, 2013.

11. Ahmed, O. J., and S. S. Cash, "Finding synchrony in the desynchronized eeg: the history and interpretation of gamma rhythms," *Frontiers in Integrative Neuroscience*, Vol. 7, p. 58, 2013.

12. NiederMeyer, E., *Niedermeyer's Electroencephalography: Basic Principles, clinical applications, and related fields*, Lippincott Williams & Wilkins, 2011.

13. Ullsperger, M., and S. Debener, *Simultaneous EEG and fMRI: Recording, Analysis, and Application*, Oxford University Press, 2010.

14. Clarke, A. R., R. J. Barry, R. McCarthy, M. Selikowitz, and C. R. Brown, "Eeg evidence for a new conceptualisation of attention deficit hyperactivity disorder," *Clinical Neurophysiology*, Vol. 113, no. 7, pp. 1036–1044, 2002.

15. Kumar, J. S., and P. Bhuvaneswari, "Analysis of electroencephalography (eeg) signals and its categorization–a study," *Procedia Engineering*, Vol. 38, pp. 2525–2536, 2012.

16. Palva, S., and J. M. Palva, "New vistas for a-frequency band oscillations," *Trends in Neurosciences*, Vol. 30.

17. Shoker, L., S. Sanei, and J. Chambers, "Artifact removal from electroencephalograms using a hybrid bss-svm algorithm," *IEEE Signal Processing Letters*, Vol. 12, no. 10, pp. 721–724, 2005.

18. Kiamini, M., S. Alirezaee, B. Perseh, and M. Ahmadi, "Elimination of ocular artifacts from eeg signals using the wavelet transform and empirical mode decomposition," in *2009 6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, Vol. 2, pp. 1094–1097, IEEE, 2009.

19. Yong, X., R. K. Ward, and G. E. Birch, "Facial emg contamination of eeg signals: Characteristics and effects of spatial filtering," in *2008 3rd International Symposium on Communications, Control and Signal Processing*, pp. 729–734, IEEE, 2008.

20. Barua, S., and S. Begum, "A review on machine learning algorithms in handling eeg artifacts," in *The Swedish AI Society (SAIS) Workshop SAIS, 14, 22-23 May 2014, Stockholm, Sweden*, 2014.

21. Mannan, M. M. N., M. A. Kamran, and M. Y. Jeong, "Identification and removal of physiological artifacts from electroencephalogram signals: A review," *Ieee Access*, Vol. 6, pp. 30630–30652, 2018.

22. McFarland, D. J., and J. R. Wolpaw, "Sensorimotor rhythm-based brain-computer interface (bci): feature selection by regression improves performance," *IEEE Transactions on Neural Systems and Rehabilitation engineering*, Vol. 13, no. 3, pp. 372–379, 2005.

23. Penny, W. D., S. J. Roberts, E. A. Curran, and M. J. Stokes, "Eeg-based communication: a pattern recognition approach," *IEEE Transactions on Rehabilitation Engineering*, Vol. 8, no. 2, pp. 214–215, 2000.

24. Arafat, I., "Brain-computer interface: past, present & future," *International Islamic University Chittagong (IIUC), Chittagong, Bangladesh*, 2013.

25. Schmidt, E. M., "Single neuron recording from motor cortex as a possible source of signals for control of external devices," *Annals of Biomedical Engineering*, Vol. 8, no. 4, pp. 339–349, 1980.

26. Chapin, J. K., K. A. Moxon, R. S. Markowitz, and M. A. Nicolelis, "Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex," *Nature Neuroscience*, Vol. 2, no. 7, pp. 664–670, 1999.

27. Kulasuriya, K. H., and M. Perera, "Forecasting epileptic seizures using eeg signals, wavelet transform and artificial neural networks," in *2011 IEEE International Symposium on IT in Medicine and Education*, Vol. 1, pp. 557–562, IEEE, 2011.

28. Amin, J., M. Sharif, M. Yasmin, and S. L. Fernandes, "Big data analysis for brain tumor detection: Deep convolutional neural networks," *Future Generation Computer Systems*, Vol. 87, pp. 290–297, 2018.

29. Subasi, A., "Epileptic seizure detection using dynamic wavelet network," *Expert Systems with Applications*, Vol. 29, no. 2, pp. 343–355, 2005.

30. Gonzalez-Vellon, B., S. Sanei, and J. A. Chambers, "Support vector machines for seizure detection," in *Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology (IEEE Cat. No. 03EX795)*, pp. 126–129, IEEE, 2003.

31. Nicolas-Alonso, L. F., and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, Vol. 12, no. 2, pp. 1211–1279, 2012.

32. Royer, A. S., A. J. Doud, M. L. Rose, and B. He, "Eeg control of a virtual helicopter in 3-dimensional space using intelligent control strategies," *IEEE Transactions on Neural Systems and Rehabilitation engineering*, Vol. 18, no. 6, pp. 581–589, 2010.

33. Krepki, R., B. Blankertz, G. Curio, and K.-R. Müller, "The berlin brain-computer interface (bbci)–towards a new communication channel for online control in gaming applications," *Multimedia Tools and Applications*, Vol. 33, no. 1, pp. 73–90, 2007.

34. Ambler, T., S. Braeutigam, J. Stins, S. Rose, and S. Swithenby, "Salience and choice: neural correlates of shopping decisions," *Psychology & Marketing*, Vol. 21, no. 4, pp. 247–261, 2004.

35. Yoshioka, M., T. Inoue, and J. Ozawa, "Brain signal pattern of engrossed subjects using near infrared spectroscopy (nirs) and its application to tv commercial evaluation," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–6, IEEE, 2012.

36. Sandryhaila, A., and J. M. Moura, "Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure," *IEEE Signal Processing Magazine*, Vol. 31, no. 5, pp. 80–90, 2014.

37. Narang, S. K., Y. H. Chao, and A. Ortega, "Graph-wavelet filterbanks for edge-aware image processing," in *2012 IEEE Statistical Signal Processing Workshop (SSP)*, pp. 141–144, IEEE, 2012.

38. Von Luxburg, U., "A tutorial on spectral clustering," *Statistics and Computing*, Vol. 17, no. 4, pp. 395–416, 2007.

39. Balaban, A. T., "Applications of graph theory in chemistry," *Journal of Chemical Information and Computer Sciences*, Vol. 25, no. 3, pp. 334–343, 1985.

40. Polak, A. J., and R. C. Sundahl, "Application of chemical graph theory for the estimation of polymer dielectric properties," *Polymer*, Vol. 30, no. 7, pp. 1314–1318, 1989.

41. Takata, M., B.-L. Lin, M. Xue, Y. Zushi, A. Terada, and M. Hosomi, "Predicting the acute ecotoxicity of chemical substances by machine learning using graph theory," *Chemosphere*, Vol. 238, p. 124604, 2020.

42. Periyasamy, B., "Solution of linear differential equations in chemical kinetics through flow graph theory approach," *Journal of the Taiwan Institute of Chemical Engineers*, Vol. 57, pp. 9–17, 2015.

43. Schlick, T., "Adventures with rna graphs," *Methods*, Vol. 143, pp. 16–33, 2018.

44. Kim, N., K. N. Fuhr, and T. Schlick, "Graph applications to rna structure and function," in *Biophysics of RNA Folding*, pp. 23–51, Springer, 2013.

45. Benedetti, G., and S. Morosetti, "A graph-topological approach to recognition of pattern and similarity in rna secondary structures," *Biophysical Chemistry*, Vol. 59, no. 1-2, pp. 179–184, 1996.

46. Ashtiani, S. N. M., M. R. Daliri, H. Behnam, G.-A. Hossein-Zadeh, M. Mehrpour, M. R. Motamed, and F. Fadaie, "Altered topological properties of brain networks in the early ms patients revealed by cognitive task-related fmri and graph theory," *Biomedical Signal Processing and Control*, Vol. 40, pp. 385–395, 2018.

47. Fang, H., Q. Wu, Y. Li, Y. Ren, C. Li, X. Xiao, T. Xiao, K. Chu, and X. Ke, "Structural networks in children with autism spectrum disorder with regression: A graph theory study," *Behavioural Brain Research*, Vol. 378, p. 112262, 2020.

48. Makovac, E., M. Mancini, S. Fagioli, D. R. Watson, F. Meeten, C. L. Rae, H. D. Critchley, and C. Ottaviani, "Network abnormalities in generalized anxiety pervade beyond the amygdala-pre-frontal cortex circuit: Insights from graph theory," *Psychiatry Research: Neuroimaging*, Vol. 281, pp. 107–116, 2018.

49. Vecchio, F., C. Tomino, F. Miraglia, F. Iodice, C. Erra, R. Di Iorio, E. Judica, F. Alù, M. Fini, and P. M. Rossini, "Cortical connectivity from eeg data in acute stroke: a study via graph theory as a potential biomarker for functional recovery," *International Journal of Psychophysiology*, Vol. 146, pp. 133–138, 2019.

50. Ajilore, O., M. Lamar, A. Leow, A. Zhang, S. Yang, and A. Kumar, "Graph theory analysis of cortical-subcortical networks in late-life depression," *The American Journal of Geriatric Psychiatry*, Vol. 22, no. 2, pp. 195–206, 2014.

51. Singh, R. P., "Application of graph theory in computer science and engineering," *International Journal of Computer Applications*, Vol. 104, no. 1, 2014.

52. Riaz, F., and K. M. Ali, "Applications of graph theory in computer science," in *2011 Third International Conference on Computational Intelligence, Communication Systems and Networks*, pp. 142–145, IEEE, 2011.

53. Foulds, L. R., *Graph Theory Applications*, Springer Science & Business Media, 2012.

54. Gautam, S., J. Teraiya, and A. K. Patra, "Spatial statistics, spatial correlation and spatial graph theory in air pollution," *Environmental Technology & Innovation*, Vol. 11, pp. 384–389, 2018.

55. Godquin, T., M. Barbier, C. Gaber, J.-L. Grimault, and J.-M. Le Bars, "Applied graph theory to security: A qualitative placement of security solutions within iot networks," *Journal of Information Security and Applications*, Vol. 55, p. 102640, 2020.

56. Prathik, A., K. Uma, and J. Anuradha, "An overview of application of graph theory," *Int. J. ChemTech Res.*, Vol. 9, no. 2, pp. 242–248, 2016.

57. Püschel, M., and J. M. Moura, "Algebraic signal processing theory: 1-d space.," *IEEE Trans. Signal Process.*, Vol. 56, no. 8-1, pp. 3586–3599, 2008.

58. Agaskar, A., and Y. M. Lu, "A spectral graph uncertainty principle," *IEEE Transactions on Information Theory*, Vol. 59, no. 7, pp. 4338–4356, 2013.

59. Zhu, X., and M. Rabbat, "Approximating signals supported on graphs," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3921–3924, IEEE, 2012.

60. Tenenbaum, J. B., V. De Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, Vol. 290, no. 5500, pp. 2319–2323, 2000.

61. Chung, F. R., and F. C. Graham, *Spectral Graph Theory*, no. 92, American Mathematical Soc., 1997.

62. Wagner, R., H. Choi, R. Baraniuk, and V. Delouille, "Distributed wavelet transform for irregular sensor network grids," in *IEEE/SP 13th Workshop on Statistical Signal Processing, 2005*, pp. 1196–1201, IEEE, 2005.

63. Wagner, R. S., R. G. Baraniuk, S. Du, D. B. Johnson, and A. Cohen, “An architecture for distributed wavelet analysis and processing in sensor networks,” in *Proceedings of the 5th International Conference on Information processing in Sensor Networks*, pp. 243–250, 2006.

64. Ciancio, A., S. Pattem, A. Ortega, and B. Krishnamachari, “Energy-efficient data representation and routing for wireless sensor networks based on a distributed wavelet compression algorithm,” in *Proceedings of the 5th International Conference on Information Processing in Sensor Networks*, pp. 309–316, 2006.

65. Valdivia, P., F. Dias, F. Petronetto, C. T. Silva, and L. G. Nonato, “Wavelet-based visualization of time-varying data on graphs,” in *2015 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 1–8, IEEE, 2015.

66. Chen, S., Y. Yang, J. Moura, J. Kovačević, *et al.*, “Localization, decomposition, and dictionary learning of piecewise-constant signals on graphs,” *arXiv preprint arXiv:1607.01100*, 2016.

67. Jain, R. K., J. M. Moura, and C. E. Kontokosta, “Big data+ big cities: Graph signals of urban air pollution [exploratory sp],” *IEEE Signal Processing Magazine*, Vol. 31, no. 5, pp. 130–136, 2014.

68. He, K., L. Stankovic, J. Liao, and V. Stankovic, “Non-intrusive load disaggregation using graph signal processing,” *IEEE Transactions on Smart Grid*, Vol. 9, no. 3, pp. 1739–1747, 2016.

69. Sporns, O., *Networks of the Brain*, MIT Press, 2010.

70. Goldsberry, L., W. Huang, N. F. Wymbs, S. T. Grafton, D. S. Bassett, and A. Ribeiro, “Brain signal analytics from graph signal processing perspective,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 851–855, IEEE, 2017.

71. Ménoret, M., N. Farrugia, B. Pasdeloup, and V. Gripon, “Evaluating graph signal processing for neuroimaging through classification and dimensionality reduction,” in *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 618–622, IEEE, 2017.

72. Pirayre, A., C. Couprie, F. Bidard, L. Duval, and J.-C. Pesquet, “Brane cut: biologically-related a priori network enhancement with graph cuts for gene regulatory network inference,” *BMC Bioinformatics*, Vol. 16, no. 1, pp. 1–12, 2015.

73. Segarra, S., A. G. Marques, G. Mateos, and A. Ribeiro, “Network topology inference from spectral templates,” *IEEE Transactions on Signal and Information Processing over Networks*, Vol. 3, no. 3, pp. 467–483, 2017.

74. Milanfar, P., “A tour of modern image filtering: New insights and methods, both practical and theoretical,” *IEEE Signal Processing Magazine*, Vol. 30, no. 1, pp. 106–128, 2012.

75. Hu, W., G. Cheung, A. Ortega, and O. C. Au, “Multiresolution graph fourier transform for compression of piecewise smooth images,” *IEEE Transactions on Image Processing*, Vol. 24, no. 1, pp. 419–433, 2014.

76. Tremblay, N., and P. Borgnat, “Graph wavelets for multiscale community mining,” *IEEE Transactions on Signal Processing*, Vol. 62, no. 20, pp. 5227–5239, 2014.

77. Chen, S., F. Cerda, P. Rizzo, J. Bielak, J. H. Garrett, and J. Kovačević, "Semi-supervised multiresolution classification using adaptive graph filtering with application to indirect bridge structural health monitoring," *IEEE Transactions on Signal Processing*, Vol. 62, no. 11, pp. 2879–2893, 2014.

78. Goldberger, A. L., L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *Circulation*, Vol. 101, no. 23, pp. e215–e220, 2000.

79. Schalk, G., D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "Bci2000: a general-purpose brain-computer interface (bci) system," *IEEE Transactions on Biomedical Engineering*, Vol. 51, no. 6, pp. 1034–1043, 2004.

80. Crovella, M., and E. Kolaczyk, "Graph wavelets for spatial traffic analysis," in *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428)*, Vol. 3, pp. 1848–1857, IEEE, 2003.

81. Hammond, D. K., P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, Vol. 30, no. 2, pp. 129–150, 2011.

82. Pesenson, I., "Sampling in paley-wiener spaces on combinatorial graphs," *Transactions of the American Mathematical Society*, Vol. 360, no. 10, pp. 5603–5627, 2008.

83. Wagner, R., V. Delouille, and R. Baraniuk, "Distributed wavelet de-noising for sensor networks," in *Proceedings of the 45th IEEE Conference on Decision and Control*, pp. 373–379, IEEE, 2006.