FUNCTIONAL PARCELLATION OF MEMORY RELATED BRAIN NETWORKS BY SPECTRAL CLUSTERING OF EEG DATA

by

Çağatay Aydın

B.S. in Electronics Engineering, Işık University, 2009

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
2012

ACKNOWLEDGMENTS

First of all, I would like to thank my thesis advisor, Prof. Dr. Ahmet Ademoğlu, for his academic support and friendly supervision. I am deeply grateful to my co-advisor, Assist. Prof. Dr. Rifat Koray Çiftçi, for his guidance, support and suggestions. I am also grateful to TUBITAK for funding of this study (Project Number: 109E202).

I see myself obliged to thank Namık Kemal Univesity Biomedical Engineering Laboratory, for its comfortable environment during the analysis of this study.

I am indebted to Prof. Dr. Yorgo Istefanopulos and to Assist. Prof. Dr. Hakan Gürkan for their leading me to the field of Biomedical Engineering. Besides, I feel grateful to my friends Adem Umut Gunebakan for his great support since my undergraduate studies; Oytun Oktay, for his full time accommodation support in Çorlu; Barış Bulutlu, for his sincere behavior; Namık Kemal Univesity Biomedical Enigneering students, for their great patience during the experiments. Lastly, I want to thank my father, Turgut Aydın, and my mother, Emine Zeynep Aydın. They have brought up me to these days with love and confidence.

ABSTRACT

FUNCTIONAL PARCELLATION OF MEMORY RELATED BRAIN NETWORKS BY SPECTRAL CLUSTERING OF EEG DATA

The EEG signal and its oscillatory components may relate with temporal modulation of information processing of a sensory activation in a local electrical field and neural populations. In this study, we investigate the clustering information of alpha band brain networks during memory load task. For this purpose, short time memory experiment with a varying memory load combinations was designed. The functional coupling among EEG electrodes were quantified via mutual information in the time-frequency plane. A spectral clustering algorithm was used to parcellate memory related circuits in the brain in a load-dependent manner. The method was based on determining the eigenspectrum of the adjacency matrix of a graph and assigning nodes to clusters with respect to this spectrum. To be able to circumvent the problem of choosing the number of clusters beforehand a soft clustering approach was implemented. It is a novel method which allows to construct significant clusters without fixing their number and increases the inside cluster significance by normalized-cut value decomposition at each clustering level. In the N-cut clustering, clustered nodes which are projected on occipital and bilateral regions increase in number with respect to the memory load. In soft clustering, inter-cluster connections between left lateral and occipital clusters are decreasing in the second time interval which can be linked to the enhancement of posterior region due to an increase in the memory demand.

Keywords: Normalized Cut, EEG, Working Memory, Memory Load, Mutual Information, Information Theory

ÖZET

HAFIZA İLE İLİNTİLİ EEG BEYİN AĞLARININ İZGESEL KÜMELEME YÖNTEMİ İLE GRUPLANDIRILMASI

EEG sinyali ve onun salınımlarının, bilgi işlenmesi esnasında beyindeki zamana bağlı nöral topluluklar ve lokal elektriksel alanlar ile bir ilişki içerisinde bulunduğu düşünülmektedir. Bu çalışmamızda, hafıza yükleme deneyi esnasında, alfa frekans bandı temelli beyin ağları öbeklenmesi yapılmıştır. Deney taslağı, kısa süreli değişen hafıza koşulları ile tasarlanmıştır. EEG elektrotları arasındaki eş bağlantının hesaplanması adına, zaman ve frekans dağılımları hesaplanarak, karşılıklı ilişki ölçütleri hesaplanmıştır. Hafıza yüklemesi durumları hesaba katılarak, izgesel öbekleme işlemi ile beyin alt ağlarına ayrıştırılmıştır. İzgesel kümeleme yönteminin temeli, komşuluk matrisinden oluşturulan bir çizgenin, komşuluk matrisinin öz görüngesi hesaplanarak, bu öz vektörler ile öbeklenmesine dayanır. Beyin ağlarının kaç adet öbekten oluşması gerektiği problemini çözmek adına, hassas öbekleme yöntemi kullanılmıştır. Yeni bir yöntem olan hassas öbekleme algoritması, kesin bir öbek sayısı belirmeksizin, alt ağların içsel bağlantı güçlerinin arttırılması ile belirgin öbekler oluşturulmasına olanak sağlar. Düzgelenmiş kesim öbeklenmesi sonuçlarında, beyinin oksipital ve bilateral bölgeler üzerinde yansıtılmış öbek düğümlerinin hafıza ile ilintili artışı gözlemlenmiştir. Hassas öbekleme sonuçlarında ise ikinci zaman aralığında, oksipital ve sol lateral öbekleri arasındaki bağlantı sayılarında azalış, hafıza talebi artışında, art beyin ağlarının iyileşmesi ile ilişkilendirilebilir.

Keywords: Normalize Kesim, EEG, Çalışan Hafıza, Hafıza Yüklemesi, Müşterek Bilgi, Bilgi Teorisi

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS

G	Graph
E	Number of edges
V	Number of vertices
$a_{i,j}$	Relation value between node i and j
α	Description of α

LIST OF ABBREVIATIONS

AD	Alzheimer Disease
BCI	Brain Computer Interface
CNS	Central Nervous System
EEG	Electroencephalography
ERD	Event Related Desynchronization
ERP	Event Related Potentials
EROs	Event Related Oscillations
fMRI	Functional Magnetic Resonance Imaging
LFPs	Local Field Potentials
LTMS	Long-term Memory System
MCI	Mild Cognitive Impairment
MI	Mutual Information
PDF	Probability Density Function
SL	Synchronization Likelihood
STMS	Short-term Memory System
TFD	Time Frequency Distribution
WM	Working Memory

1. INTRODUCTION

1.1 General Background

Hans Berger was a psychiatrist who had the purpose of investigating the relation between mind and body. He was the first person to discover alpha fluctuations based on the electrical activity and found an intermediary device, the human Electroencephalogram (EEG) [2]. The electrical potential over the scalp is produced by the activities of brain cortex and recorded with the help of EEG electrodes. It is noninvasive and has a high temporal resolution within milliseconds. Temporal resolution is important for investigating the time dependent behavior of the brain under sensory and cognitive stimulation [1].

Figure 1.1 Hans Berger (Human EEG founder) [2]

The use of EEG moved into multidisciplinary areas such as Psychiatry and Neurology. The brain abnormalities such as epilepsy and structural lesions were some of the major fields in clinical applications.

There are several usages of EEG for studying cognitive and memory based brain dynamics. The effect of memory load is investigated with event related potentials (ERP) and event related oscillations (EROs). ERP is derived by averaging the post-stimulus of EEG over trials. Kok et al. [4] have discovered that the cognitive demand

Figure 1.2 A clinical EEG system [1]

caused by memory load was reflected by the P3 signal. Additionally, with the help of EROs, increased alpha band oscillations ($\sim 10Hz$) proportional to the memory demand were found in a working memory task. The finding indicated that the correlation of inhibitory task-irrelevant neural networks relates to the increase of alpha oscillations.

The EEG signal and its oscillatory activity are strongly related with temporal modulation of information processing. Task-related brain dynamics can be referred to an activation of neural populations in a local electrical field [5]. Generally, with the oscillations of alpha (8 Hz- 12 Hz) and gamma (+30 Hz) bands of the EEG rhythm, the amplitude of the signal changes in the memory and cognitive based tasks can be observed. On the other hand, performing shorter time scale analysis of the EEG signal, information theoretic measures may provide better results rather than investigating the amplitude change in the EEG oscillations. Furthermore, the mutual information (MI) analysis allows to construct adjacency matrices due to the similarity information between electrode pairs.

1.2 Motivation and Objectives

The main motivation of this thesis study is to implement a method to investigate and to observe the memory related brain dynamics. The brain dynamics can be considered as an activation behavior of an individual voxel or a group of voxels. Thus, to determine the information flow between electrode pairs, the normalized cut clustering algorithm may provide a suitable option. Furthermore, the experimental design should be selected carefully in order to induce the memory related brain circuits because task-related brain dynamics may be inhibited by the complexity of the experiment design.

We investigate the clustering information of memory related brain networks. For this purpose, short time memory task which includes memory load varieties is implemented on the subjects. To calculate mutual information, time and frequency information are both taken into consideration with the Cohen class time-frequency distribution (TFD) formulation. Cohen class mutual information helps us to integrate adjacency matrices based on the similarity information of individual electrode pairs. In addition, essential frequency bins are selected from the TFD with respect to the default alpha frequency ($8Hz-12Hz$) intervals.

Spectral clustering was first used by Donath et al. [6] to partition a graph by calculating eigenvectors of an adjacency matrix. With the help of graph Laplacian, a graph can be represented in a matrix formation. The second important discovery about spectral clustering was that the bi-partitions of the graph were closely related with the second eigenvector of the graph Laplacian [7]. This idea led to the concept that bi-partitioning with the help of a specific eigenvector was an adequate solution. The ongoing research activities about Spectral Clustering became widespread in the Machine Learning community with the popular works of Shi and Malik [8]. In the bi-partitioning problem, the main reason why the normalized cut spectral clustering technique is used is because the form of clusters are constructed without strong assumptions. Furthermore, enhancement of weighted kernel-k-means algorithm with spectral clustering allows for determining the relation between the smallest eigenvectors of the

graph Laplacian with the largest eigenvectors of the kernel matrices . Moreover, spectral clustering technique is used for investigating the cognitive demand and how it changes over the organization of the brain networks with the EEG and fMRI. Based on correlated BOLD signal, Shen et al. [9] proposed an fMRI technique to examine network properties in terms of clustering the brain networks into the functional subunits.

Since the clustering analysis has an intuitive goal of grouping interrelated nodes and excluding the non-related nodes, investigating network properties may provide detailed information about functional integration between brain regions [10]. Hence, previous studies have suggested that, an adequate network bi-partitioning solution could be obtained with the help of eigenvectors. In addition, the normalized cut algorithm provides flexible clustering approach which allows to generate clusters without making strong assumptions about the form of clusters [8]. On the other hand, the main problem for the clustering analysis is finding out optimum number of clusters. For this reason, an adequate solution to determine optimum number of clusters before implementing a clustering analysis is one of the main motivations for this thesis study.

Moreover, one of the popular graph based spectral clustering algorithms, the normalized cut clustering method is used to group memory related circuits in the brain. From the calculated adjacency matrices, the N-cut algorithm is used for node wise clustering between nodes. After node wise clustering information, subject wise clustering is applied with respect to the similarities of node information over all subjects. To provide information for the number of clusters from the adjacency matrices, soft clustering method is used.

1.3 Outline of this thesis

The general outline is summarized as follows; Chapter 2 gives detailed information about the EEG signal characteristics due to the specified EEG band intervals. In addition, cognitive brain networks related to memory processing and short time mem-

ory load are introduced. The information theoretic measures are briefly explained and time-frequency based (Cohen Class) mutual information is explained. At the end of the Chapter, the graph based spectral clustering algorithm is briefly explained in EEG signal processing. The detailed approximation of methods and our experimental design are explained in Chapter 3. The preprocessing procedure, THE information theoretic measures and their calculations, THE formulation of N-Cut clustering algorithm are presented. Moreover, the soft clustering method is introduced and presented in detail. In Chapter 4, behavioral, N-cut clustering and soft clustering results are presented. In Chapter 5 and 6, discussions and conclusion are given respectively.

2. EEG

2.1 EEG Signal Characteristics

The electroencephalogram (EEG) signal is obtained from the electrodes on human scalp that can be observed as cyclic fluctuations of the brain electrical potentials. The EEG signal offers us the chance to investigate distinct sleep stages, depth of anesthesia, seizures and other neurological disorders [11]. Further studies show that, cognitive process is associated with working memory, analytical thinking and attention [12, 13]. In addition, neocortical dynamic functions can be measured with the help of scalp EEG. However, head volume conduction causes the EEG electrodes to have reduced spatial information over the scalp. On the other hand, intracranial electrodes allow obtaining detailed local information. For example, an intracranial electrode can provide averaged synaptic action of approximately 100 million neurons.

Figure 2.1 64 Channel EEG electrodes

The EEG has a fine time resolution in milliseconds in contrast to its a poor

spatial resolution. In other words, the physical layer between electrodes and the brain current sources generally limits the spatial resolution. However, recent advances on electrode materials and electronic equipment have improved spatial resolution by providing systems having 256 recording channels.

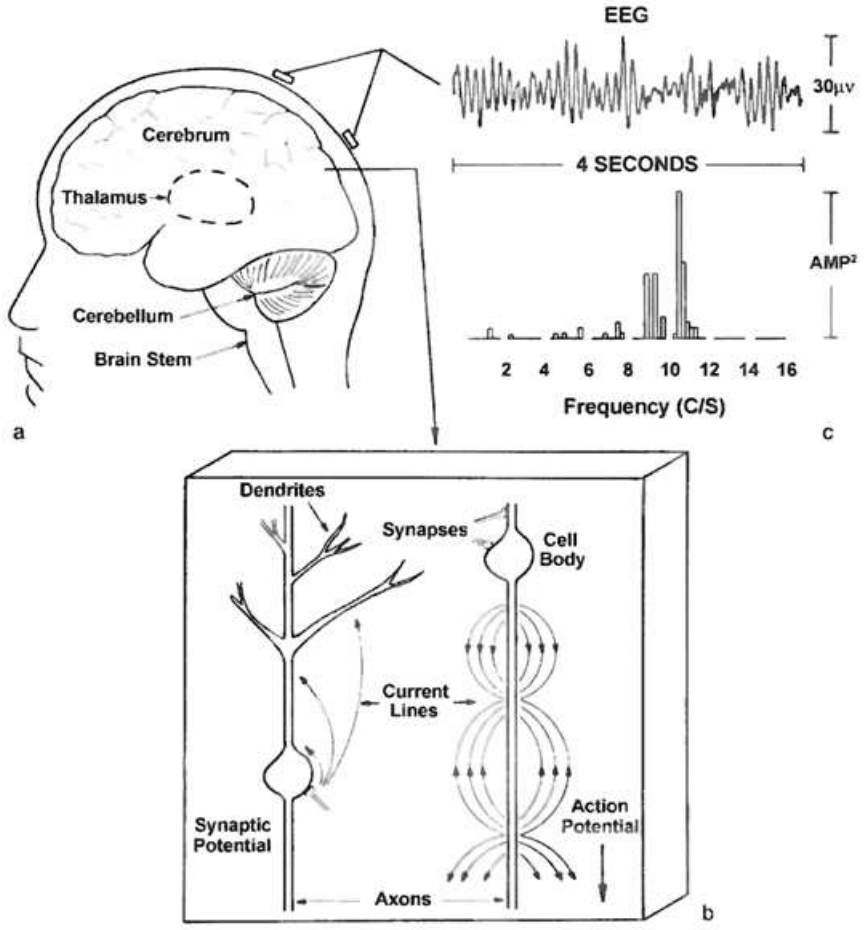

Figure 2.2 How EEG signal is generated [2]

The relation between afferent fibers and the cortical neurons causes the depolarization stage. The depolarization stage stands for generating the excitatory post synaptic potentials (EPSP) which may be reflected as a signal variability with respect to the amplitude and duration. Furthermore, sinusoidal fluctuations can be observed with the combination of periodic afferent neurons [2].

Standard EEG rhythms have a wide frequency range from low (0.1 Hz - 4 Hz) to high frequency oscillations (12 Hz - 30 Hz). They can be labeled as delta (0 Hz - 4 Hz),

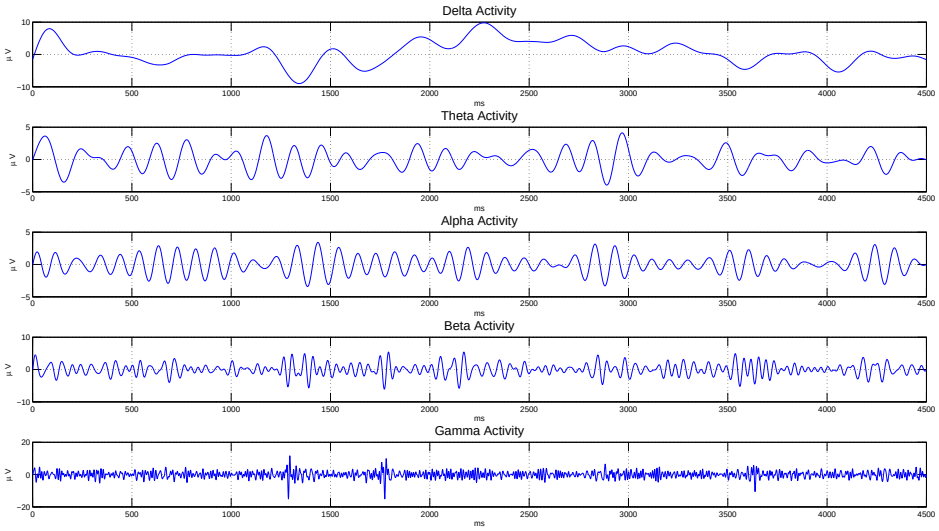

Figure 2.3 EEG frequencies

theta (4 Hz - 8 Hz), alpha (8 Hz - 12 Hz), beta (12 Hz - 30 Hz), gamma (above 30 Hz). Delta and theta rhythms are generally associated with drowsiness, deep sleep stage and can be observed in infancy and childhood. In the condition of relaxed wakefulness in eyes close stage, alpha fluctuations can be easily observed. Both alpha and gamma fluctuations are associated with the cognitive effort and memory based neural circuits of the brain. The beta rhythm is also seen during increased cognitive effort as well as during drowsiness and light sleep.

2.1.1 Introduction to Memory Processing

In terms of obtaining the memory related information from the brain, the memory processing has to be accepted as a dynamic process. Fuster et al. [14] claimed that, it was not possible to separate the perception, recognition, language, planning, problem solving and decision making apart from memory processing. From his findings, memory was initiated as distributed property of a cortical system. Furthermore, Rose et al. [15] initiated the idea that, investigation of the memory dynamics in a single specific region was inadequate. On the other hand, Damasio et al. [16] considered that various levels of neural groups could be able to settle up the memory processing by the

integration of several brain systems.

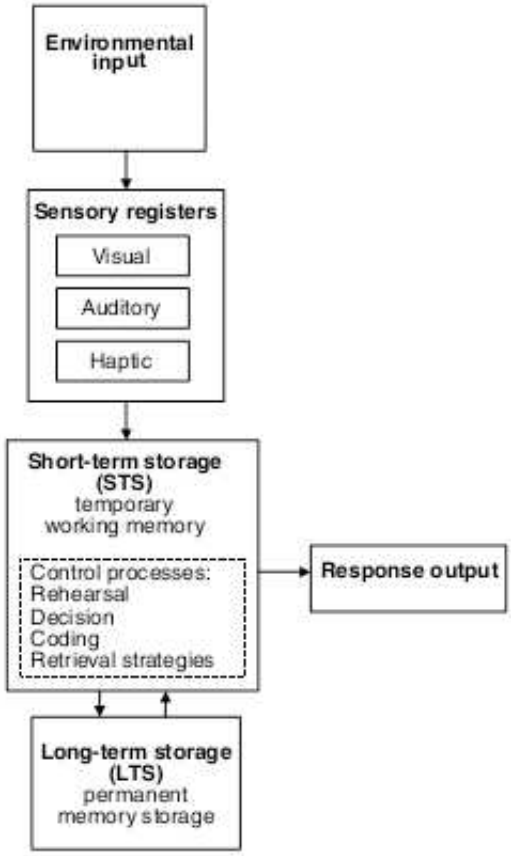

Figure 2.4 Structure of memory [3]

Examining everyday routines in terms of cognitive process in the brain, such as recognition of a familiar object, the working memory system (WMS) and long-term memory system (LTMS) are important concepts. The brief mechanism of the memory process was formed by Shiffrin et al. [3] who assumed that the bottom-up memory feedback loop was started from the sensory input and was ended with the comparison between sensory input and previously stored semantic information in LTM. The short term memory (STM) code was created by the positive outcome of this comparison. Furthermore, in the case of speaking and thinking which referred to complex cognitive processes, similar STM and LTMS interaction could exist. However, in the complex cognitive mechanism, sensory inputs could not be adequate enough to generate and compare with a STM code in simultaneous applications such as planning a speech and

talking at the same time. Thus, the recent idea of WM concept relies on the enhanced mechanism which comprises the attention control system, the central executive system and the assistant system [17]. In other words, WMS and LTMS are important memory features to explore the changes in cognitive performance and memory span.

Several strategies have to be discussed about investigating the memory processing on the brain. To improve the credibility of the strategy, Klimesh et al. [18] claimed that, a researcher might take both physiological and mental factors into consideration. As such, some physiological factors which could be listed as thickness of the skull or the volume of cerebrospinal fluid and some technical factors which could be listed as inter-electrode distance or montage variety were vital inconsistency factors. In addition, developing cognitive based experimental design might have uncertain affects on subjects due to the extraction of complex results which might consist both related and non-related cognitive circuits together. Moreover, the task performance might differ between subjects. To sum up, Klimesh et al. [18] proposed that, previously mentioned variety had to be carefully structured by the researcher.

Başar et al. [19] suggested that the alpha band and its oscillatory components could be a bridge between thalamus and cortex which was affected by the synchronized neural activity. Furthermore, Klimesh et al. [18] came with the assumption of investigating memory related brain dynamics with respect to dynamic property of memory processing. From their assumption, if the memory processing is using the longitudinal pathways (feedback loops) which link the thalamic nuclei with the cortex, one of the most predominant rhythm (alpha 8 Hz - 12 Hz) can be used to extract the memory information in these pathways. In addition, they found significant difference with respect to the frequency analysis of the alpha rhythm between subjects with respect to their memory performances. The relation between the alpha band and the memory related brain dynamics were investigated by using the amplitude analysis of the alpha signal such as the overall power distribution of the signal and shifts in the power. Furthermore, the shifts in the alpha power came with the event-related desynchronization (ERD) term which was first found by Pfurtscheller et al. [20] The main idea of ERD relies on the difference in percentage which can be determined as an increase or de-

crease in the band power during a test interval with respect to the reference interval. During retrieval period of the memory task, poor memory performers showed a lack of significant desynchronization in lower alpha (8 Hz-10 Hz).

Several clinical techniques are also used as a pre-diagnostic tool for memory related disorders in the brain. Pijnenburg et al. [21] measured the linear and non-linear statistical dependencies of the EEG signal in Alzheimer's disease (AD) and mild cognitive impairment (MCI) patients. The synchronization likelihood (SL) analysis which is similar to mutual information (MI) calculation, is used to measure statistical dependencies. However, the SL calculation is differentiated from MI with an additional normalization procedure. They found decreased functional interactions in AD patients, but they found increased functional interactions in MCI patients in lower alpha band (8 Hz-10 Hz). In a recent study, Haenschel et al. [22] investigated the WM difference between schizophrenia patients and healthy subjects in terms of event related potentials (ERP).

2.1.2 Short Term Memory and Memory Load

Particular studies were done to research brain circuitry during memory load. Jensen et al. [23] used modified STM paradigm (Stenberg task). In addition, their experimental design allowed to separate the encoding, the retention and the recognition states of the memory processing. They found increased peak in the alpha oscillations due to increased memory load. They assumed that power increase in the alpha oscillations reflected by the synchronization across multiple brain regions to achieve active inhibition. Fink et al. [24] selected individual alpha frequencies (IAF: Lower 1 6 Hz - 8 Hz, lower 2 8 Hz - 10 Hz, upper alpha 10 Hz - 12 Hz) during various types of memory related experimental designs to compare the IAF with default alpha frequency (8 Hz-12 Hz). The relation between the memory load and memory related circuit demands in the brain was determined by the ERD of the upper alpha band. It was observed as a decrease when the experiment became more complex and difficult. On the other hand, during the classical retention task and speed of the information processing task, the

interrelation between ERD patterns in the upper alpha band becomes stronger. Furthermore, during the increased task demands, a decreased ERD correlation between lower 1 and upper alpha bands was observed. Fink et al. [24] stated that upper alpha band was reflected by the retention specific task demands. However, lower 1 was generally related to the attention specific task demands. To find out working memory related brain circuits Sauseng et al. [5] used an experimental design which investigated the differences between retention and manipulation organizations in the brain. They found a strong prefrontal alpha power increase and occipital alpha suppression during visuospatial information processing. Furthermore, in the memory scanning task, an alpha band specific event related synchronization (ERS) was observed during the retention interval of the memory load based experiment [17]. In other words, retention interval can be referred to the time spend after certain encoded items were seen by subject and kept in the mind as the encoded information. In the experiment, they observed an increase in the alpha band power by the increase in the number of represented items.

One of the recent studies has proposed model of local field potentials (LPFs) by using the attractor network model [25]. During the memory load in the simulated working memory task, they observed an increase in the theta and gamma band power. In addition, they found a simultaneous decrease in the ratio od alpha/beta power.

2.1.3 Information Theoretic Measures Based on EEG

For estimating the time delays between EEG electrodes, the mutual information calculation (MI) was used by Moddemeijer et al. [26]. At that time, the analysis which was performed to observe statistical properties of the EEG signal, was quite rare. Thus, they first generated the probability density function from a pair of electrodes with respect to their time delay information. Then, Shannon entropy based mutual information was computed.

The entry into a new millennium has been accompanied by investigating the phase information of the EEG signal. Morrman et al. [27] used mean phase coherence

to diagnose temporal lobe epilepsy. They assumed that spatial and temporal changes in the phase synchronization might have promising relation in terms of detecting the epileptic activity. They indicated that the phase synchronization of an EEG signal might be used as an intermediary tool for performing functional analysis about cognitive circuits in the brain such as memory. Rosso et al. [28] used wavelet-based information theoretic measures to examine both time and frequency components of the EEG signal. They used the orthogonal wavelet transform (ODWT) in order to overcome the non-stationary characteristics of an EEG signal. In several BCI (brain computer interface) studies, information theoretic measures were also applied such as mutual information analysis to compute the information transfer within the brain regions [29]. To investigate the classification accuracy, they used an entropy difference method of stochastic process which can be considered as a mutual information computation.

In order to analyze both time and frequency components of an EEG signal by using information theoretic measures , time-frequency distributions (TDF) are conventional. The dependencies between signal pairs on time-frequency plane can be investigated within the signal characterization and classification method like mutual information (MI). The MI generally represents a measure of independence between random variables. Hence, to generate the statistical difference between signal pairs, the individual TFDs and a joint TDF of a signal pair are adequate enough to calculate the time-frequency based MI [30, 31]. Lu et al. [32] used the time-frequency cross mutual information analysis. They inquired the functional connectivity in alpha and the beta bands during resting, preparing, movement onset and movement offset states.

There are several clinical studies based on information theoretic measures which are used as a pre-diagnostic tools for AD and MCI patients. Dauwels et al. [33] enquired different synchrony measures such as correlation coefficient, mean-square, phase coherence, Granger causality, phase synchrony indices, information-theoretic divergence, state space based measures, stochastic event synchrony measures. Generally, for pre-diagnosing MCI patients, Granger causality and stochastic event synchrony methods offered successful results. The differences between local and global synchrony measures were emphasized. Local synchrony measures were offered by a relationship

between pairs of the signal. On the other hand, global measures assisted them to investigate all EEG signals over the scalp. In addition, decreased synchrony measures between the signal pairs were identified by functional disconnection of the neocortex areas.

2.2 Spectral Clustering

Several scientific fields have similar purposes to separate their data in groups which consist of interrelated components. Spectral clustering has several advantages among varieties of other clustering methods due to the implementation ease, calculation efficiency and enhancement of traditional clustering algorithms such as k-means. In the spectral clustering algorithm, the Laplacian matrix is an important measure. The spectral graph theory is formed with the help of Laplacian matrix properties. Mohar et al. [8] inferred that, the eigenvalues and the eigenvectors of the un-normalized Laplacian matrix could be used to distinguish properties of graphs. Furthermore, second eigenvector of the Laplacian matrix has an important role on bi-partitions of the graph. Apart from an un-normalized Laplacian matrix, a normalized Laplacian matrix is commonly used in bio-informatics studies. The normalized cut algorithm is one of the popular clustering algorithms which was founded by Shi and Malik [34]. The grouping algorithm is mainly about, extracting global impressions in a given image and generate perceptional grouping. It was referred to an unbiased measure which allows for establish a sub-group of a graph by minimizing normalized cut with maximizing the similarity within the cluster.

Investigating the use of normalized cut algorithm in brain research, Shen et al. [9] used graph theory based clustering algorithm to parcellate the resting state functional subunits of the brain by using fMRI. Normalized cut clustering was used to examine network properties in the brain. They parcellated the brain into the functional clusters based on a correlated BOLD signal. This study offers us to observe brain regions in terms of both cortical functioning and psychologically functioning clusters rather than to observe the brain only with psychological clustering information.

In addition, another study by Chen et al. [35], was made on the purpose of the brain tumor separation by using fMRI images. They found that graph cut based clustering algorithm was promising for the tumor exploration and its separation for anatomic variability.

3. METHODS

3.1 Experimental Procedure and Subjects

In this experiment, the EEG data was recorded over 17 (12 male, 5 female mean age of 23) healthy volunteers. Each participant gave an informed consent about the experimental procedure. First of all, a baseline recording was performed with 3 minutes eyes-open and 3 minutes eyes-closed resting condition. The visuo-spatial short time memory task was run afterwards. The locations of one, three or five targets on four by four box model were shown to the subjects for 500 milliseconds. During each stimulus, one of three box combinations (one, three or five box) was represented randomly. Due to the box combination model, their positions had to be kept in memory for $2,500$ milliseconds and had to be compared with an answer stimulus which was presented in gray color. A comparison should be made between the probe and the answer stimulus to decide on whether the target positions were exactly the same as answer positions or not. After $2,500$ milliseconds retention interval, if it was true, subjects had to answer with the right mouse button press or vice versa in 1000 milliseconds.

To be able to preserve the event synchronization, the clock in the data recording computer and the clock in the experiment computer were synchronized periodically after each trial.

The segmentation period for this experiment was 4000 ms. Each segment was started from 500 ms after the probe (green) stimulus to 1000 ms after the answer stimulus (gray). The significant time interval for each segment which was 2500 ms long between the probe and the response, was used in analysis. Furthermore, the main reason for initializing longer time interval in segmentation was to secure the signal from fluttering in the filtering of preprocessing analysis. Each answer, each reaction time, and the overall mean reaction time for each box model was calculated and written into a log file.

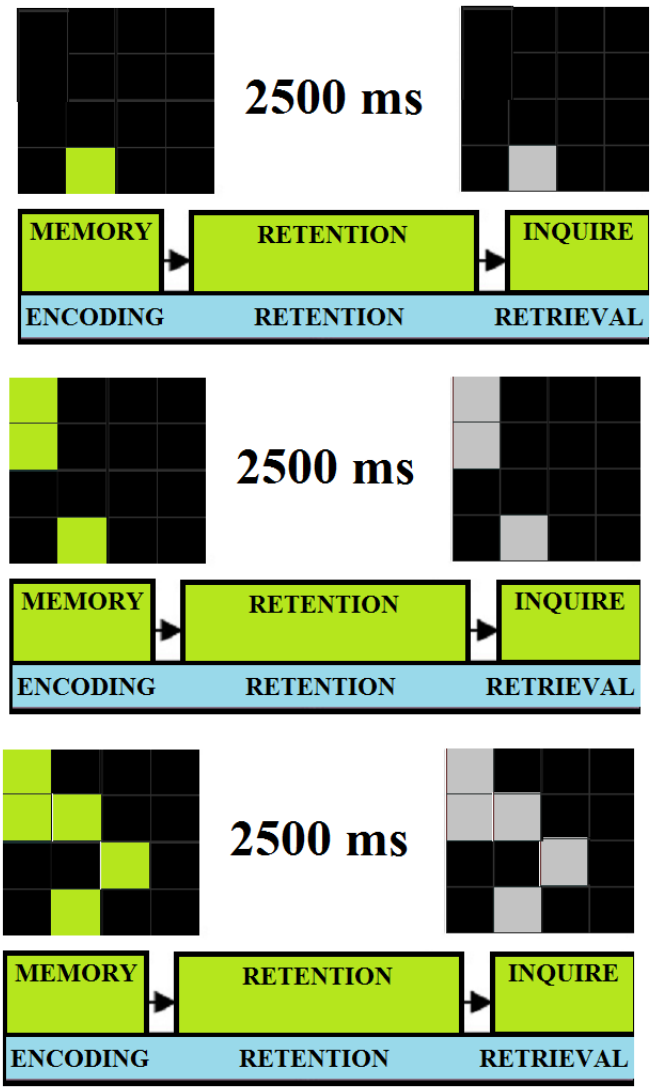

Figure 3.1 The experimental design

There were 100 stimuli for each box combinations and there were equal number of true and false probes. In total, 300 stimuli were represented during the experiment. Only correct answers were taken into consideration. The experimental design was made by the Psyctoolbox, Matlab software.

3.2 Data Gathering

EEG recordings were done by 64 channel EGI HydroCel amplifier and the ongoing recordings were stored on Macintosh Workstation computer. To establish the impedance levels which were set as 50 $k\Omega$, sponge electrodes were soaked in potassium chloride and alcohol free shampoo solution for about 10 min before the cap placement.

Figure 3.2 64 channel EGI HydroCel and experiment room

3.3 Preprocessing

The preprocessing steps can be listed as; 0.1 Hz first order high pass filtering, 100 Hz low pass filtering, 50 Hz Notch filtering, segmentation of correct segments, artifact detection, ocular artifact removal, bad channel replacement and file export. The preprocessing was done with EGI Net Station software tools.

The initial EEG recording was sampled at 1 kHz. In addition, filtering tools were used as a standard preprocessing steps to analyze the default EEG bands. There were important details in the filtering application such as selecting the filter roll-off to 2 Hz. The initial part and the end part of the signal was fluttered because of filter

roll-off property. The fluttered signal was observed as 506 ms shorter than the original signal. Fluttering problem was observed in the 0.1 Hz high pass filtering and 100 Hz low pass filtering process too. However, longer segment periods were used in order to prevent the signal from suffering the fluttering effect and to secure the preprocessed segments.

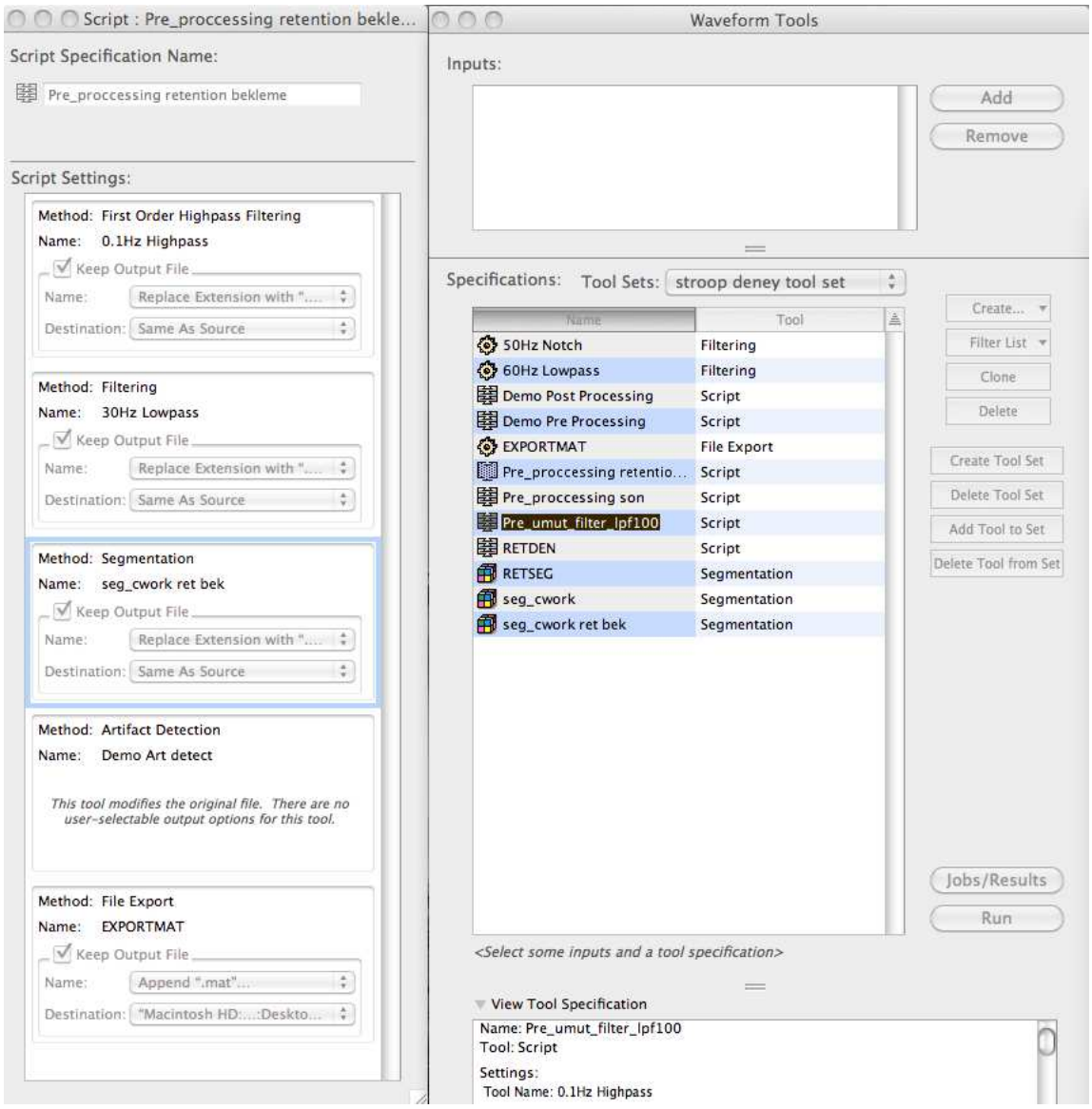

Figure 3.3 Netstation preprocessing software

The power line noise was filtered in the last part of the EEG filtering procedure. Narrow-band electrical field was generated by various electrical equipment which were located in the lab environment. In order to preserve the signal from 50 Hz power line interference, 50 Hz notch filter was used.

In the segmentation, only segments which consisted of correct responses, were generated as 4000 ms long time intervals with respect to the one, three and five box combination models.

The fourth preprocessing routine was defined as an artifact detection tool. It can exclude bad channels caused by eye blinks, eye movements, eye-blink segments. Detected bad channels were marked as unused channels for the entire recording. The bad channels were sorted out from the entire segments with the threshold level of ± 200 μV. 80 ms moving average window was applied to correct the bad segments. In addition, eye blinks were detected from 640 ms time windows with the threshold level of ± 140 μV and they were corrected with 80 ms moving average window. Eye movements were detected as with the same routine as an eye blink detection process except for a change of threshold level to ± 55 μV. The detected and refined bad channels due to the default characteristics of EEG amplitude were fluttered with the help of ocular artifact removal tool. There were specific face and eye electrodes to detect the eye blink. Detected eye blinks were flattened due the blink slope threshold level of 14 μV/ms with the references of Gratton et al. [36].

In bad channel replacement procedure, the bad channel was replaced with the interpolation of remaining channels. This process was used in case of a bad channel existence after the artifact detection and the ocular artifact removal process. Finally, for further analysis, the data was converted into a "mat." file to access in Matlab environment [10].

3.4 Generating Adjacency Matrices

3.4.1 Information Theoretical Measures on Time-frequency plane

To calculate both time and the frequency components of the signal, Cohen class distributions can be calculated. The main formula of Cohen's class $C(t, f)$ can be expressed as;

$$C(t,f) = \int\int\int \phi(\theta,\tau)s(u+\frac{\tau}{2})s^*(u-\frac{\tau}{2})e^{j(\theta u-\theta t-2\pi\tau f)}dud\theta d\tau \tag{3.1}$$

The s represents the signal, s^* represents for complex conjugate of the given signal and the $\phi(\theta,\tau)$ stands for the kernel function. The energy preservation and marginals which are the properties of time frequency distributions (TDF) are satisfied when $\phi(\theta,0) = \phi(\theta,\tau) = 1\forall\tau,\theta$

The analogy between probability density functions (PDFs) and time frequency distributions (TFDs) [30] of a two dimensional random variable can be presented as Equation 3.2 and 3.3. The conversion of information theoretic measures such as entropy into TFDs representation can be accomplished by Equation 3.3.

$$\int\int C(t,f)dtdf = \int |s(t)^2|dt = \int |S(f)^2|df \tag{3.2}$$

$$\int C(t,f)df = |s(t)^2|, \int C(t,f)dt = |S(f)^2| \tag{3.3}$$

PDFs and TFDs are different from each other for that, TFDs are not always positive like as PDFs. Hence, in this thesis spectrograms which are always positive, are used for generating Cohen class TDFs. Moreover, before the implementation for information theoretical based measures, TFD has to be normalized by its energy distribution.

3.4.2 Time Frequency Plane Mutual Information

For a brief approximation of one dimensional mutual information (MI) calculation, assume that two random variables X and Y have mutual information which can be expressed as;

$$I(X;Y) = \sum_{x}\sum_{y} p(x,y) log \frac{p(x,y)}{p(x)p(y)} \tag{3.4}$$

The $p(x,y)$ is a joint probability density function and, $p(x)$ and $p(y)$ are marginal PDFs of X and Y. If X and Y are independent from each other, MI is determined as minimum and equal zero.

To calculate Cohen class TFD with the help of Cohen class distribution, energy density functions are taken into account instead of PDFs. If we replace the marginal densities $p(x), p(y)$ with the individual energy densities $C_x(t,f), C_y(t,f)$ and the joint PDFs $p(x,y)$ with the joint energy distributions $C_{xy}(t,f)$ the main Equation 3.4 can be formed.

The MI equation 3.4 can be computed by;

$$I(C_x, C_y) = \int\int |C_{xy}(t,f)| log \frac{|C_{xy}(t,f)|}{C_x(t,f)C_y(t,f)} dtdf \tag{3.5}$$

Simplifying the Equation 3.5, the cross energy distribution can be stated as a combination of two individual TDF in root, equation 3.6 has a similar formulation such as Jensen Renyi divergence;

$$\sqrt{C_x, C_y}(t,f) = \sqrt{C_x(t,f)C_y(t,f)} \tag{3.6}$$

Thus, the MI calculation in Equation 3.6 can be transformed into Equation 3.8;

$$|C_{xy}(t,f)| = \sqrt{|C_x(t,f)|}\sqrt{|C_y(t,f)|} \tag{3.7}$$

In the end without calculating the joint energy distribution the MI Equation can be stated as 3.8;

$$I(C_x, C_y) = \int\int \sqrt{|C_x(t,f)|}\sqrt{|C_y(t,f)|} log \frac{1}{\sqrt{|C_x(t,f)|}\sqrt{|C_y(t,f)|}} dtdf \tag{3.8}$$

3.5 Graph Based Adjacency Matrices

A network is represented with the help of nodes or vertices and the edges between these nodes. The directed and undirected graph can be represented as follows;

the vertices and edges can be represented as $V = \{v_1, v_2, ..., v_n\}$, if the edges are connected in terms of the MI results $e(i,j) = 1$ otherwise $e(i,j) = 0$. The graph can be represented as $G(V,E)$ [9].

To determine the relationship between the nodes, the number of edges between each pair of nodes is used to create adjacency matrices. If the adjacency matrix consists of binary elements (either 1 or 0), it is called un-weighted adjacency matrix. It means, if there is an edge between pair of nodes, the value of adjacency is equal to 1 otherwise

it is equal to 0. In this thesis study, weighted graphs are generated by normalized mutual information of electrode pairs. Generated weighted graphs consist of the MI values which indicate the similarity bond between all node pairs [37].

3.5.1 Graph Based Clustering Parameters and Types

If the graph is considered as a weighted graph, the adjacency matrix can be represented by $W = ((w_{ij})i, j = 1, ..., n)$, then the $w_{ij} = 0$ means that there is not a connection between v_i, v_j otherwise, it carries non-negative value $w_{ij} \geq 0$.

The degree of a matrix can be defined as;

$$d_i = \sum_{j=1}^{n} w_{ij} \tag{3.9}$$

In this particular notation, this sum indicates that sum of all v_j vertices are adjacent to v_i. Degree matrix D is defined as a diagonal matrix which includes the sum of edges attached to given vertices $d_1, ..., d_n$ [8].

If a graph $G(x_i, x_j)$ consists of two partitions such as A and $\overline{A}$ where $A \cup \overline{A} = V$ and $A \cap \overline{A} = \oslash$, indicator vector of first partition can be defined as;

$$\begin{aligned} x_i &= 1 \quad if \quad v_i \epsilon A \\ x_i &= 0 \quad if \quad v_i \epsilon \overline{A} \end{aligned} \tag{3.10}$$

The number of vertices is used to measure the size of partition A. On the other hand, $vol(A)$ measures the size in terms of summation over the weights of all edges

which are attached to vertices in A.The size of the partition can be described in two ways;

$$\begin{aligned} |A| = &\ \ the\ \ number\ \ of\ \ vertices\ \ in\ \ A \\ vol(A) &= \sum_{i \epsilon A} d_i \end{aligned} \tag{3.11}$$

With the combination of weights and degree matrices, Laplacian matrices can be calculated. The graph Laplacian can be described as;

$$\begin{aligned} L(i,j) &= d_i \quad if \quad i = j \\ L(i,j) &= -w_{i,j} \quad if \quad e(i,j) = 1 \\ 0 &\qquad elsewhere \end{aligned} \tag{3.12}$$

$$L = \begin{pmatrix} L_1 & & & \\ & L_2 & & \\ & & \ddots & \\ & & & L_k \end{pmatrix}$$

The unnormalized graph Laplacian can be represented as following Equation 3.13;

$$L = D - W \tag{3.13}$$

The graph Laplacian can be normalized in two ways;

$$L_{sym} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}} = I - D^{-\frac{1}{2}} W D^{-\frac{1}{2}}$$
$$L_{rw} = D^{-\frac{1}{2}} L = I - D^{-\frac{1}{2}} W \tag{3.14}$$

The L_{sym} is denoted as symmetric matrix and L_{rw} is obtained by random walk normalization algorithm.

3.6 N-Cut Clustering Algorithm

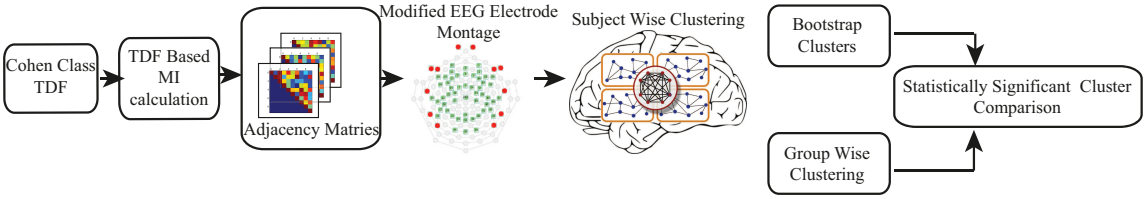

Figure 3.4 Normalized Cut clustering flow diagram: TFDs, MI adjacency matrix calculation, modified montage, subject-wise N-cut clustering, group-wise clustering, statistical comparison

N-cut clustering is an algorithm which computes an unbiased measure that allows to establish a sub-group of a graph by minimizing the normalized cut and maximizing the similarity within the cluster. In contrast to other graph clustering methods, it does not make strong assumptions on outliers which can be considered as an important advantage.

3.6.1 Subject-wise N-Cut Clustering

N-cut algorithm can be described as;

$$Ncut(A, \overline{A}) = \frac{cut(A, \overline{A})}{Vol(A)} + \frac{cut(A, \overline{A})}{Vol(\overline{A})} \tag{3.15}$$

$$cut(A, \overline{A}) = \sum_{v_i=A, v_j=\overline{A}} w(i,j) \tag{3.16}$$

According to Shi and Malik [34], normalized spectral clustering algorithm can be listed as follows;

- the number of clusters k clusters is selected,
- The un-normalized Laplacian is computed by $L = D - W$,
- Due to number of k clusters, $u_1, \ldots, u_k$ eigenvectors are computed with the help of generalized formula $Lu = \lambda Du$ and those eigenvectors corresponding to the k largest eigenvalues are selected,
- The $U \in R^{n \times k}$ matrix is generated by using $\mathbf{u_1}, \ldots, \mathbf{u_k}$ eigenvectors as columns,
- For $i = 1, \ldots, n$, the $y_i \in R^k$ vector is obtained as the i-th row of U matrix,
- The y_i in R^k, $i = 1, \ldots, n$ are clustered in with the $k - means$ algorithm into clusters $C_1, \ldots, C_k$ [8],
- The clusters $A_1, \ldots, A_k$ with $A_i = \{j | y_j \in C_i\}$ are generated,

3.6.2 Group-wise implementation of N-Cut Clustering

- The cluster vectors $A_1, \ldots, A_k$ of a single subject are merged by multiplying each A_i with its cluster index i to form the vector $N_A = [1.A_1 + ... + k.A_k]$,
- The subject-wise clustered matrix $S_A \in R^{s \times n}$ for 17 subjects is generated as $S_A = [(\mathbf{N_A})_\mathbf{1}, \ldots, (\mathbf{N_A})_\mathbf{s}]$,
- The 3-D prone matrix $P_A \in R^{n \times n \times s}$ is generated by taking each column of S_A and mapping its indices as

$$P_A(i,j,s) = \begin{cases} 1 & if \quad S_A(s,i) = S_A(s,j) \\ 0 & else \end{cases} \quad s = 1,\ldots,17 \quad i,j = 1,\ldots,n \tag{3.17}$$

- Generalized N-Cut clustering algorithm is applied to the mean of the 2D slabs of the prone matrix P_A to form the final cluster vectors $A_1,\ldots,A_k$ from which the N_A clustering matrix is obtained and mapped on the scalp surface.

This group-wise analysis is repeated over three task conditions i.e. 1, 3 and 5 box combination and clustering information is topologically plotted over modified electrode locations to investigate the comparison between different box models.

3.6.3 Bootstrap Statistical Test

In order to identify statistically significant electrodes from group-wise clusters, bootstrap technique is employed [38]. The 2D slabs of the P_A matrix are shuffled with the repetion allowed manner to obtain its surrogate versions. 1000 trials of P_A are obtained from a sequence of indices uniformly distributed between $[1-17]$ which give the shuffled indices of the slabs. The indices allow for a slab to appear more than once.

The shuffled P_A is fed into generalized N-cut algorithm 1000 times and the average cluster matrix is thresholded by %95 of its maximum value and the entries exceeding it are plotted on the scalp as statistically significant nodes.

3.7 Soft Clustering

A significant drawback of existing clustering algorithms is determining the number of clusters beforehand. Most of the time it is not possible to define an optimum and this determination is done with heuristic methods. Additionally, when the task

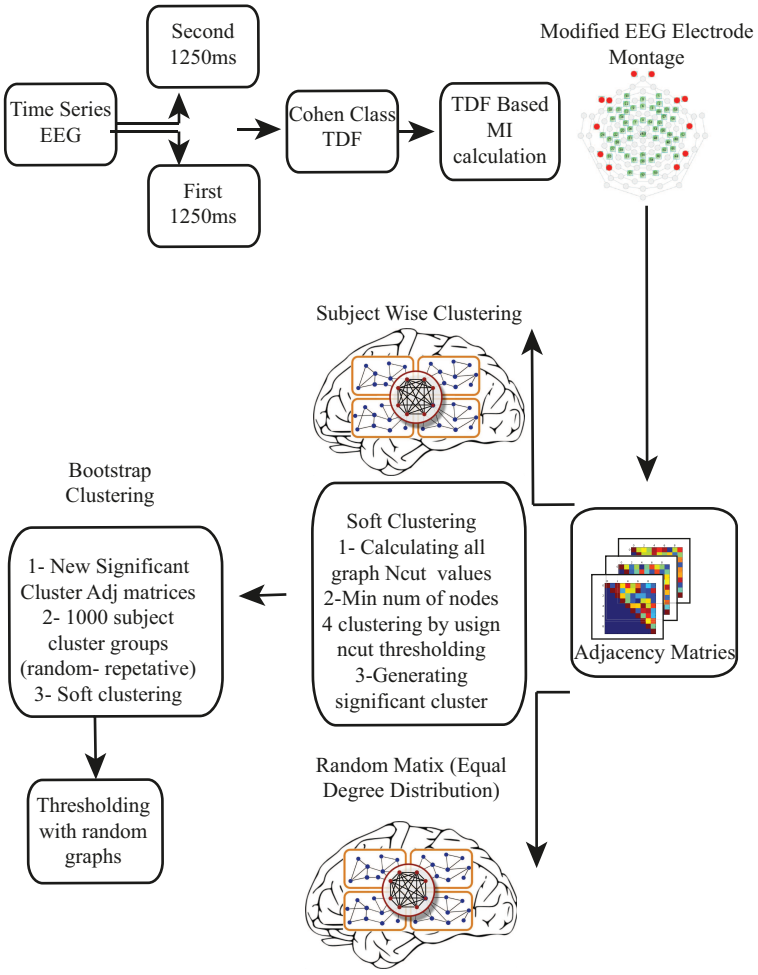

Figure 3.5 Normalized Cut clustering flow diagram: TFDs, MI adjacency matrix calculation, modified montage, subject-wise N-cut clustering, group-wise clustering, statistical comparison

is determining the group-wise clusters from individual subject level clusters, setting the cluster number to the same value for all subjects and the group may restrict the inference that might be obtained from this large group of data. Hence, to introduce some flexibility and circumvent the problem of setting a fixed cluster number, a soft clustering approach may be adopted. The method is based on the successive application of the clustering algorithm to the clusters that are obtained at earlier steps. The rationale is that if two nodes have a strong interconnection then they will be separated at later levels of these successive clusterings, but if they have a weak connection they will be segregated at earlier steps (Figure 3.5).

3.7.1 Soft Clustering Algorithm

The N-cut clustering algorithm is applied to the adjacency matrix W by choosing the symmetric version of the normalized L matrix, L_{sym}. At any particular level, existing clusters are partitioned individually and the global N-cut values for these candidate clusters are computed. Then, the cluster which gives the minimum N-cut value is partitioned. This procedure continues until all the clusters have at most 4 nodes. At each partitioning step a cluster identity vector is formed and written in a matrix form. After the termination of the procedure, these cluster identity matrices are averaged and a resultant matrix which represents clustering strength between any two nodes is obtained. This matrix has values between 0 and 1, and a value near 1 signifies that those two nodes are highly likely to be clustered together, whereas a value near 0 notifies that those two nodes are hardly to be clustered within the same cluster. As a result, this procedure gives us a new interaction matrix generated from the adjacency matrix, but this time the interaction strength between two nodes is computed from the probability of being clustered together.

In Figure 3.6, tree shaped decomposition represents the N-Cut value computation for the all possible end clusters in a given graph. In other words, a road map with respect to clustering information of the whole graph, is computed. In this thesis, subject-wise and group-wise adjacency matrices have low- sized matrix formation such as 52×52 which allow us to implement N-Cut analysis with computational ease.

- Using N-cut values $l_1, \ldots, l_l$ of N-Cut index I_N is used as a computation threshold for the generalized N-cut algorithm,
- The cluster vectors $A_1, \ldots, A_k$ of a specific N-cut value are merged by multiplying each A_i with its cluster index i to form the vector $N_S = [1.A_1 + ... + k.A_k]$,
- The N-cut value based clustered matrix $S_S \in R^{s \times n}$ for number of N-cut values is generated as $S_S = [(\mathbf{N_S})_\mathbf{1}, \ldots, (\mathbf{N_S})_\mathbf{s}]$,
- The 3-D prone matrix $P_S \in R^{n \times n \times L}$ is generated by taking each column of S_S

and mapping its indices as Equation 3.17

- The mean of 2-D prone matrix $P_S \in R^{n \times n}$ is generated to form the 2-D subject prone matrix.

The algorithm is calculated for all subjects ($s = 17$), two time intervals (0 ms - 1250 ms, 1250 ms -2500 ms) and 3 task conditions (1, 3 and 5 box combination).

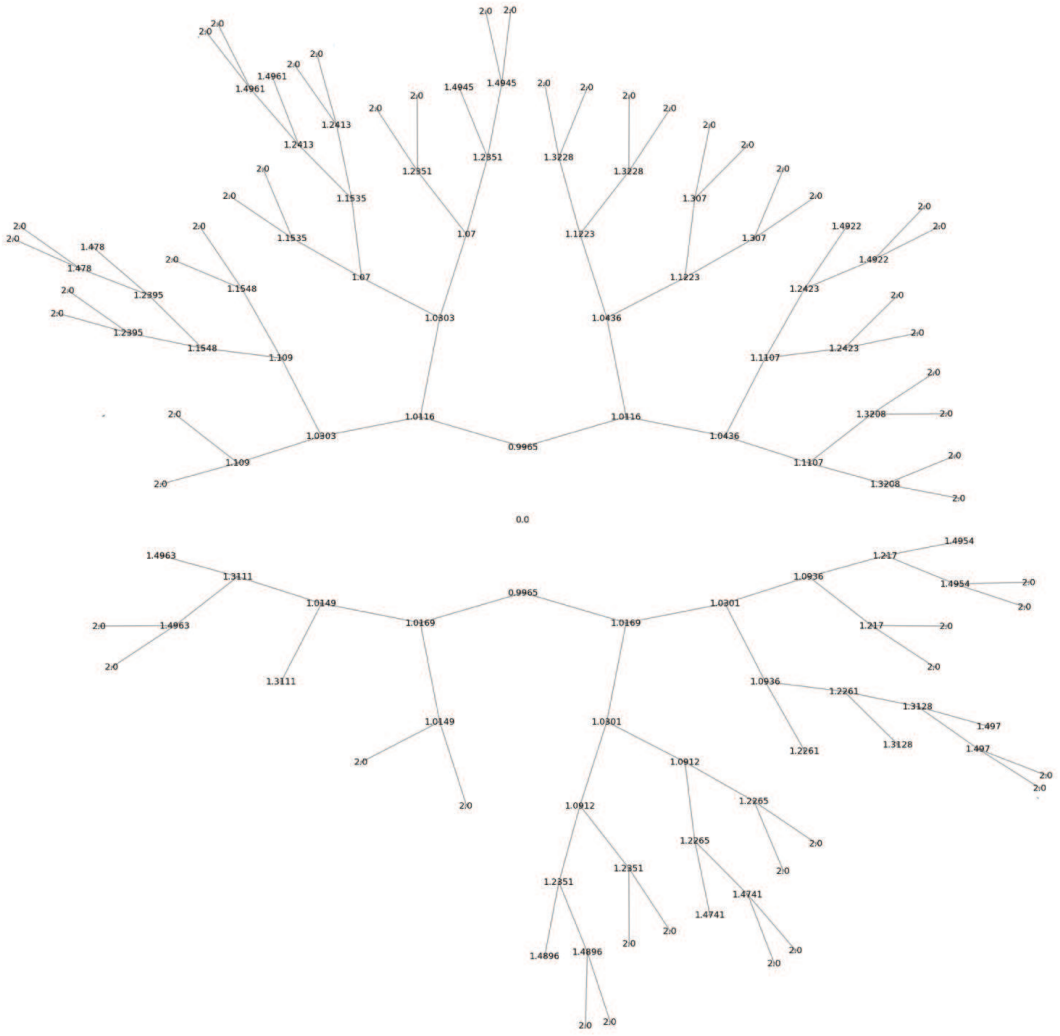

Figure 3.6 Soft clustering N-cut value decomposition of a graph

To implement group-wise anaysis over calculated subject-wise 2D slabs of the P_S matrix are shuffled with the repetition allowed manner to obtain its surrogate versions. 1000 trials of P_S are obtained from a sequence of indices uniformly distributed between

$[1-17]$ which give the shuffled indices of the slabs. The indices allow for a slab to appear more than once. (Figure 3.9).

Generalized soft clustering algorithm is applied to the mean of the 2D slabs of the prone matrix P_S to form the final cluster vectors $A_1, \ldots, A_k$ from which the N_S clustering matrix is obtained and mapped on the scalp surface. Hierarchical clusters are generated from prone matrixP_S, the function assigns a unique color to each group of nodes in the dendrogram where the linkage is less than threshold $t = (max(Z))$, upper triangle of $1 - P_S$ is Z vector (Figure 3.10),

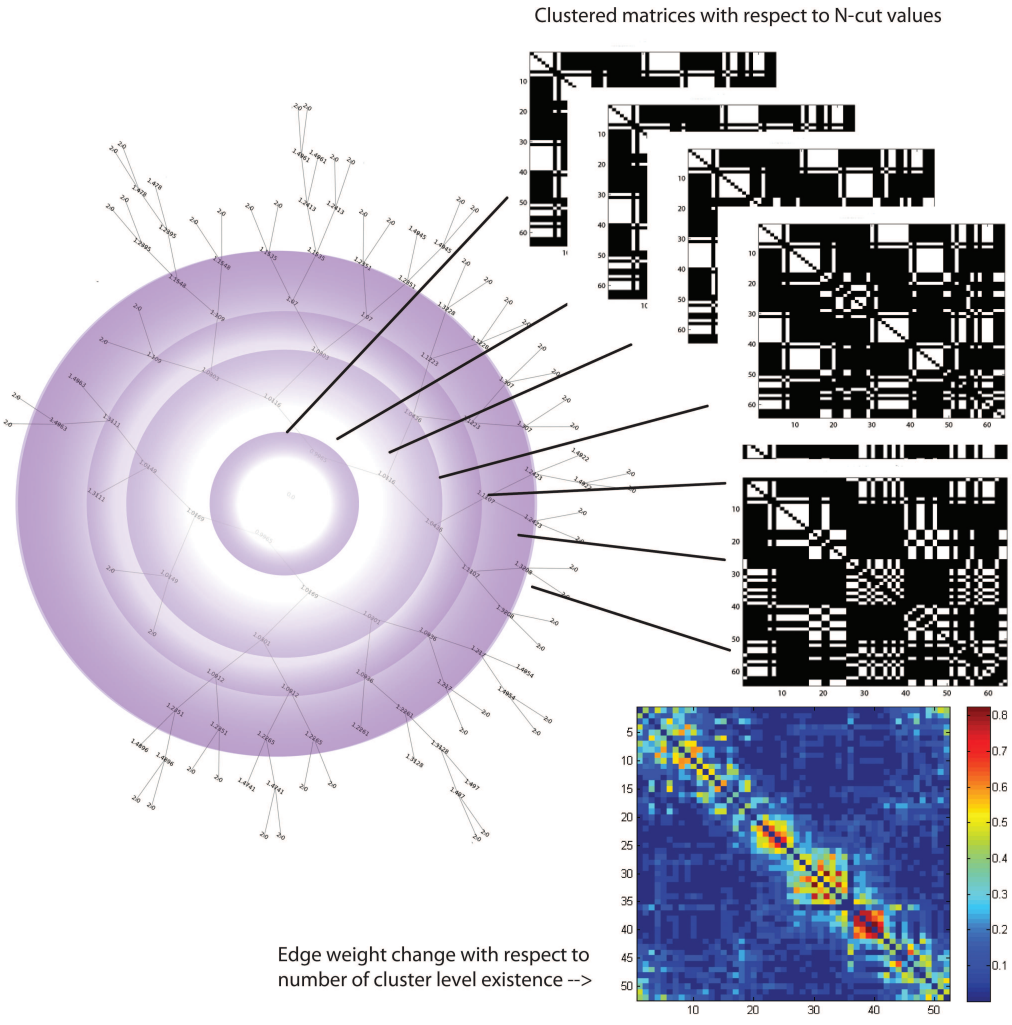

Figure 3.7 Soft clustering: generation of adjacency matrices with respect to node-cluster existence

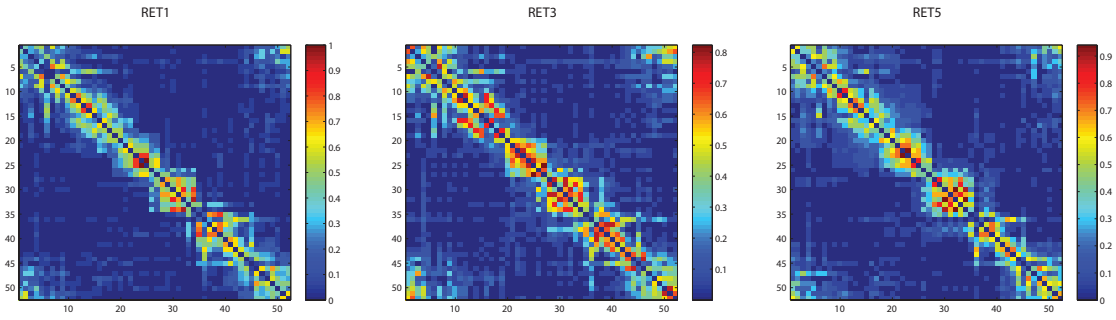

Figure 3.8 Soft Clustering: Subject-wise soft clustering matrices with respect to box combinations (first 1250ms time interval)

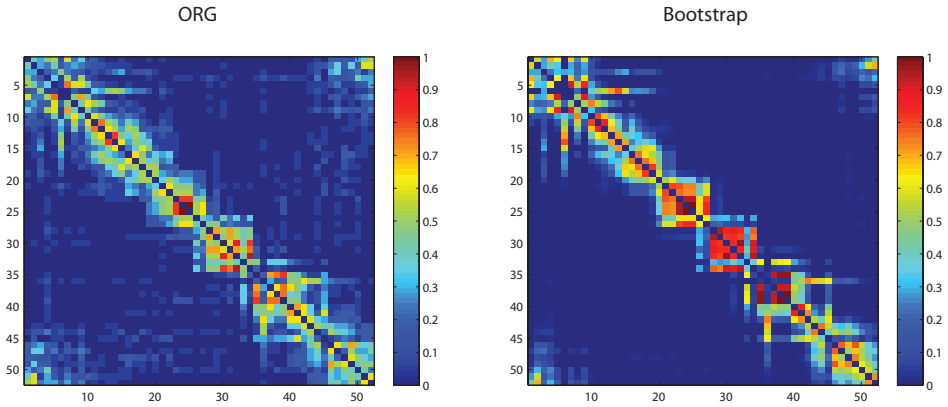

Figure 3.9 (Left) Original Group-wise Adjacency Matrix (Right) Bootstrap Adjacency Matrix

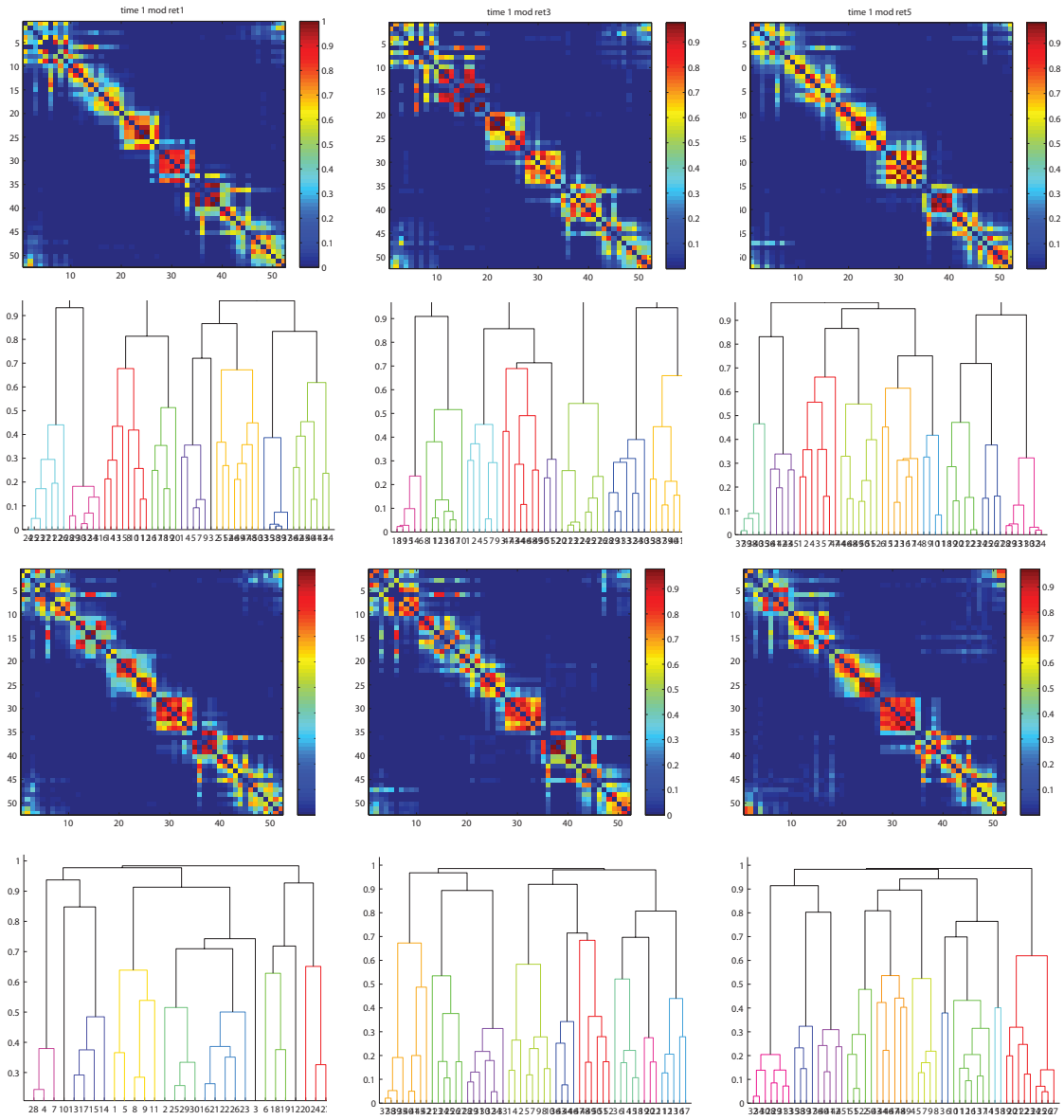

Figure 3.10 Columns: Box Combinations Ret1, Ret3, Ret5; Rows 1-3 : Group-wise soft cluster matrices, Row 1; first time interval 1250 ms, Row3; second time interval; Rows 2-4 Hierarchical Clusters with respect to above group-wise soft cluster matrices

4. RESULTS

4.1 Behavioral Results

During the experiment, the visuospatial stimulus information, subject response information and reaction times were recorded into a log file. In addition, the percentage of correct answers, the mean reaction time, fastest reaction time were calculated and added into log file. To compare the reaction time and error percentage between 3 different task conditions, mean reaction times and mean error rates of all subjects were computed (Figure 4.1).

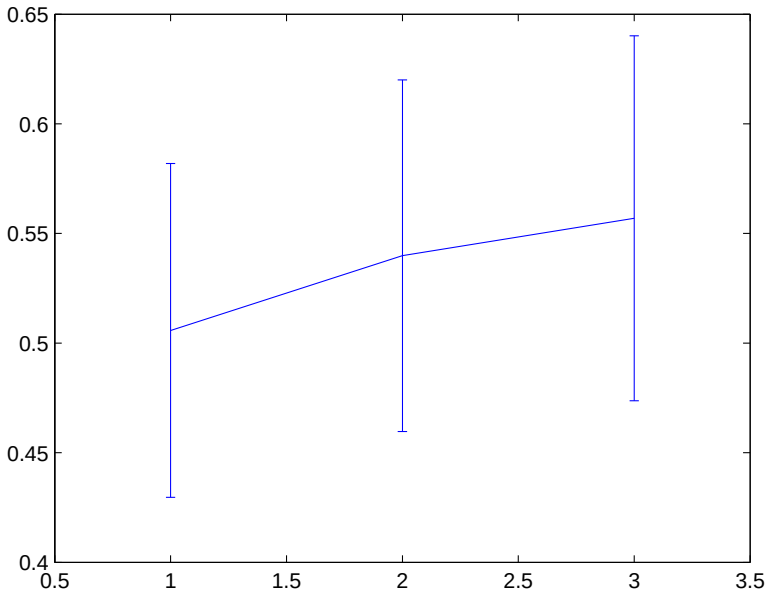

Figure 4.1 3 Box model mean reaction time comparison

There was no statistically significant difference in correct responses between 3 box and 5 box model ($\%92,9 \pm 4,6$) ($p = 0,811$). However, there were significant differences between 1 box ($\%94,8 \pm 4,0$) and 3 box model ($\%93,13 \pm 5,3$) ($p = 0,050$) and between 1 box and 5 box model ($p = 0,0030$).

Directly proportional reaction time increase was observed by the number of

represented boxes. The mean reaction time due to box models can be listed as follows; 1 box model $0,50 \pm 0,07$, 3 box model $0,53 \pm 0,08$, 5 box model $0,55 \pm 0,08$ seconds. There were significant differences between 1 box and 3 box model ($p = 0,00047$), between 1 box and 5 box model ($p = 0,000046$), 2 box and 3 box model ($p = 0,0067$).

4.2 Experimental Results

4.2.1 N-Cut Clustering Results

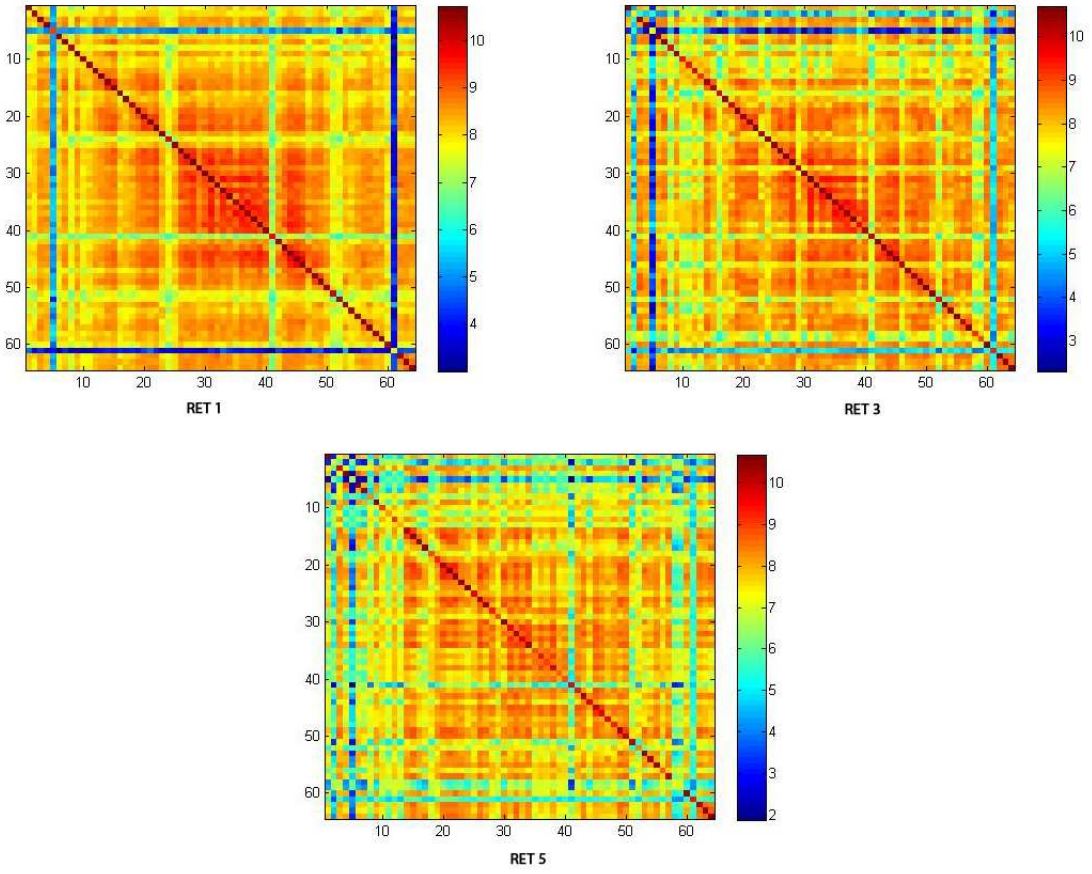

Figure 4.2 Mutual information based adjacency matrices

Preprocessing procedure is implemented by Netstation software. Preprocessed EEG segments which are selected as only good segments, are converted into .mat files to determine the ongoing procedures in Matlab software. Each segmented signal was converted into time-frequency distribution components with Cohen class distributions

[3.1]. To separate the frequency interval of the signal, associated frequency bins are selected. Mutual information is calculated by using Equation 3.8 equation to generate adjacency matrices. The representation of the MI based adjacency matrices are given in Figure 4.2.

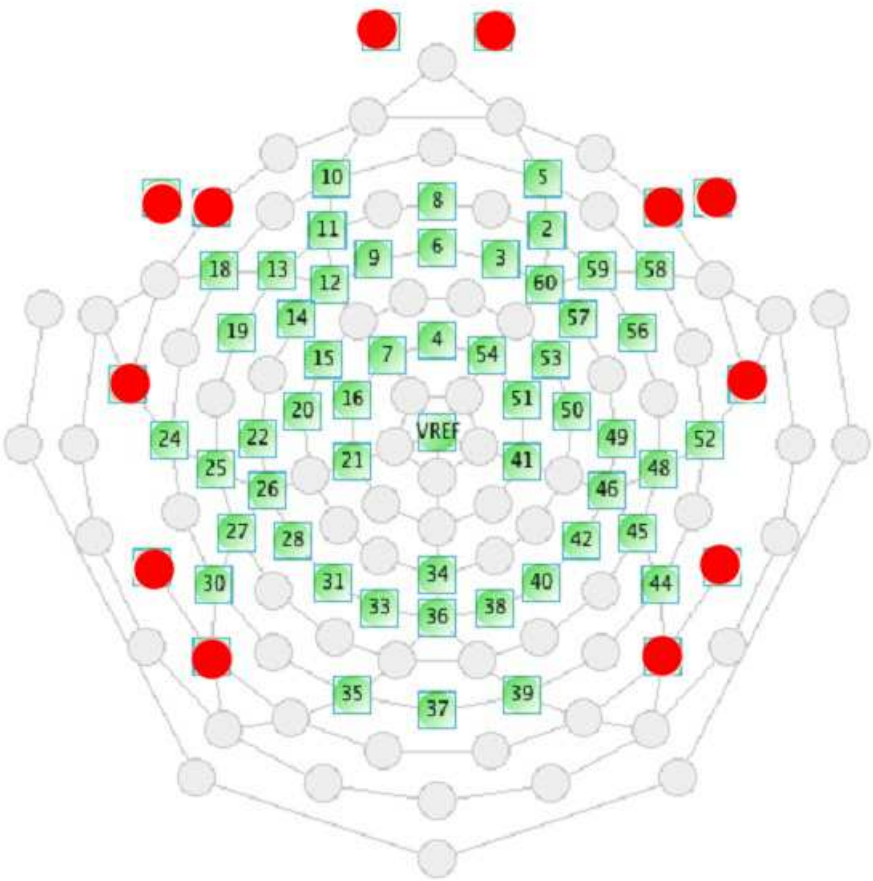

Figure 4.3 Eliminated electrode channel view

From the EGI EEG amplifier default electrode locations, the red dotted electrodes are eliminated (Figure 4.3). In other words, to increase the clustering significance, eye and ear electrodes are removed and the N-Cut clustering algorithm is applied both on normal clusters and bootstrap clusters. Furthermore, again 1000 pseudo clusters are generated with respect to the clustering information of each subject clusters to generate bootstrapped clusters. The same clustering algorithm is applied to one thousand clusters as well.

In the N-Cut clustering process adjacency matrices are transformed into sparse matrix formation. Then, with respect to the similarities of each subject clusters, new adjacency matrices of all subjects are calculated due to the similar clustering information of separate subjects. Furthermore, the procedure of N-Cut clustering is repeated over all subjects based adjacency matrix. The clustering information is topologically plotted over electrode locations to investigate the comparison between various box models. Only 4 cluster based, parcellation analysis is taken into consideration. When the number of boxes are increased, the topological parcellation is revealed with respect to the memory related areas on the brain. The main clusters are observed over posterior, bilateral and prefrontal areas of the brain. The number of nodes in the each cluster for determined brain regions are used to observe the memory related brain dynamics.

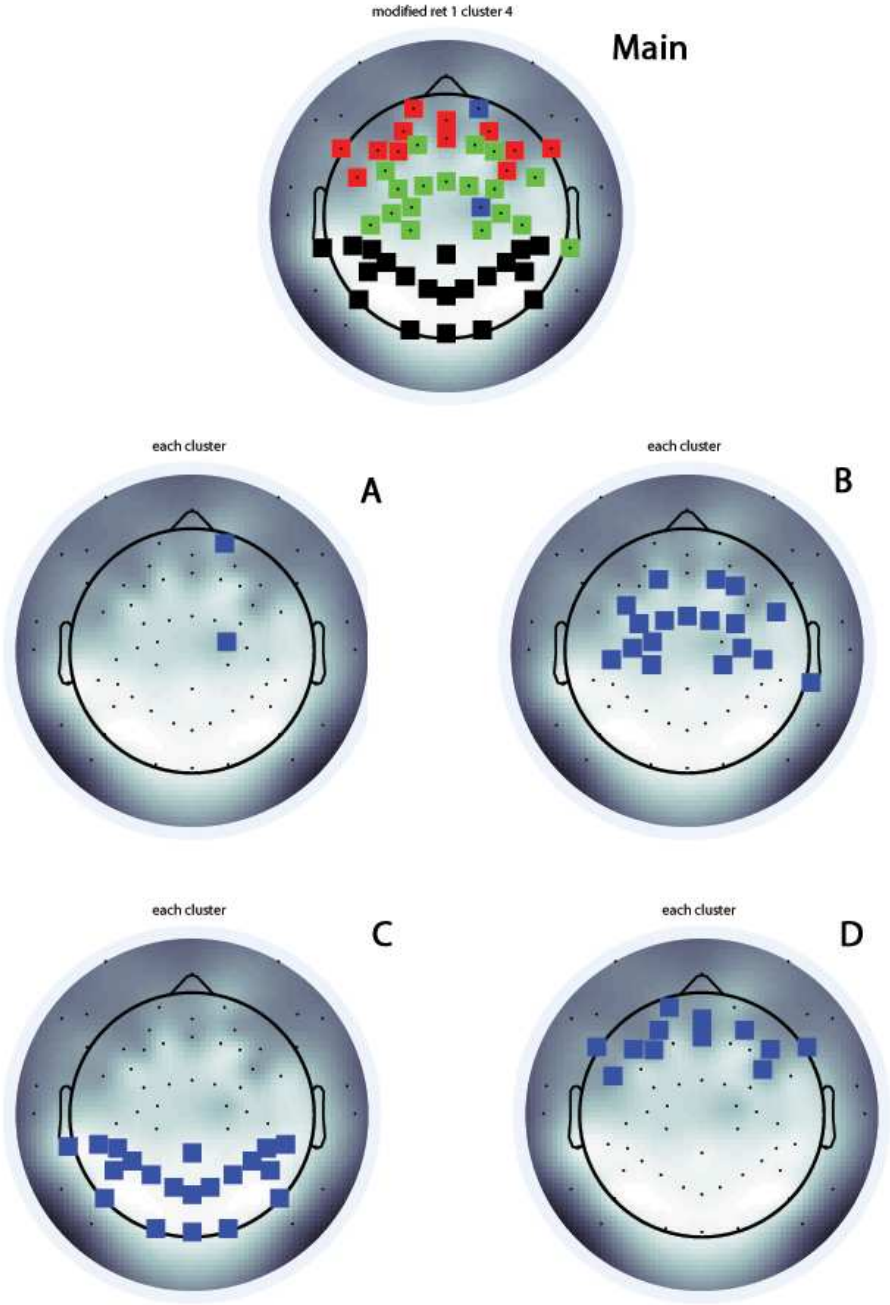

Figure 4.4 RET 1 box model all clusters and separate clusters

In Figure 4.4, 4.5 and 4.6 the reduced electrode locations of the 1, 3 and 5 box model clusters are presented. There were 3 main clusters in the 1 box model. In the cluster B, there are clusters which are well projected on the parietal regions of the brain. In addition, cluster C is positioned over the occipital region. Cluster D is mainly grouped over prefrontal region.

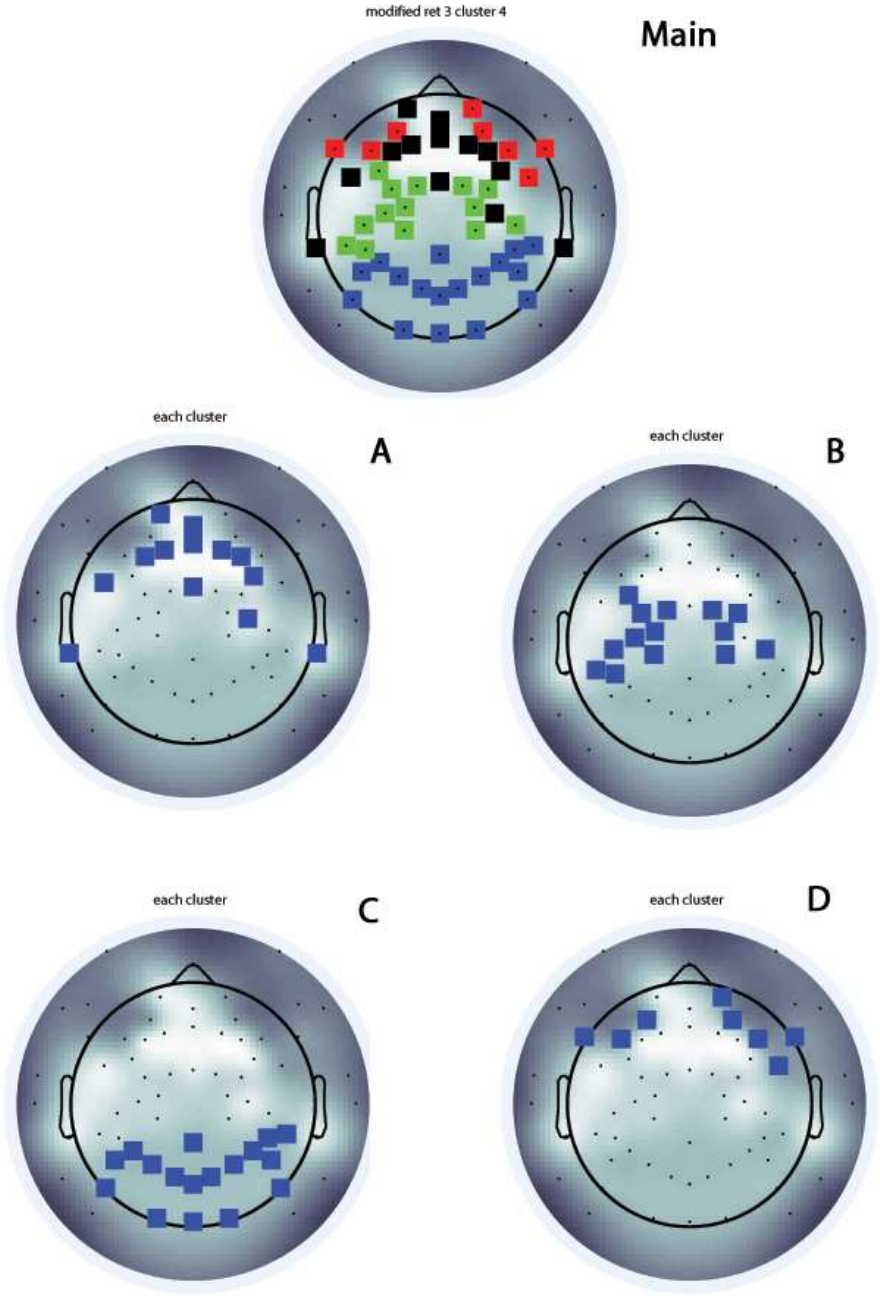

Figure 4.5 RET 3 box model all clusters and separate clusters

In Figure 4.5, there were 4 main clusters in the 3 box model. In the clusters A and D, there are clusters which are projected on the prefrontal regions of the brain. In addition, cluster C is grouped over occipital region. Cluster B is mainly grouped over parietal region.

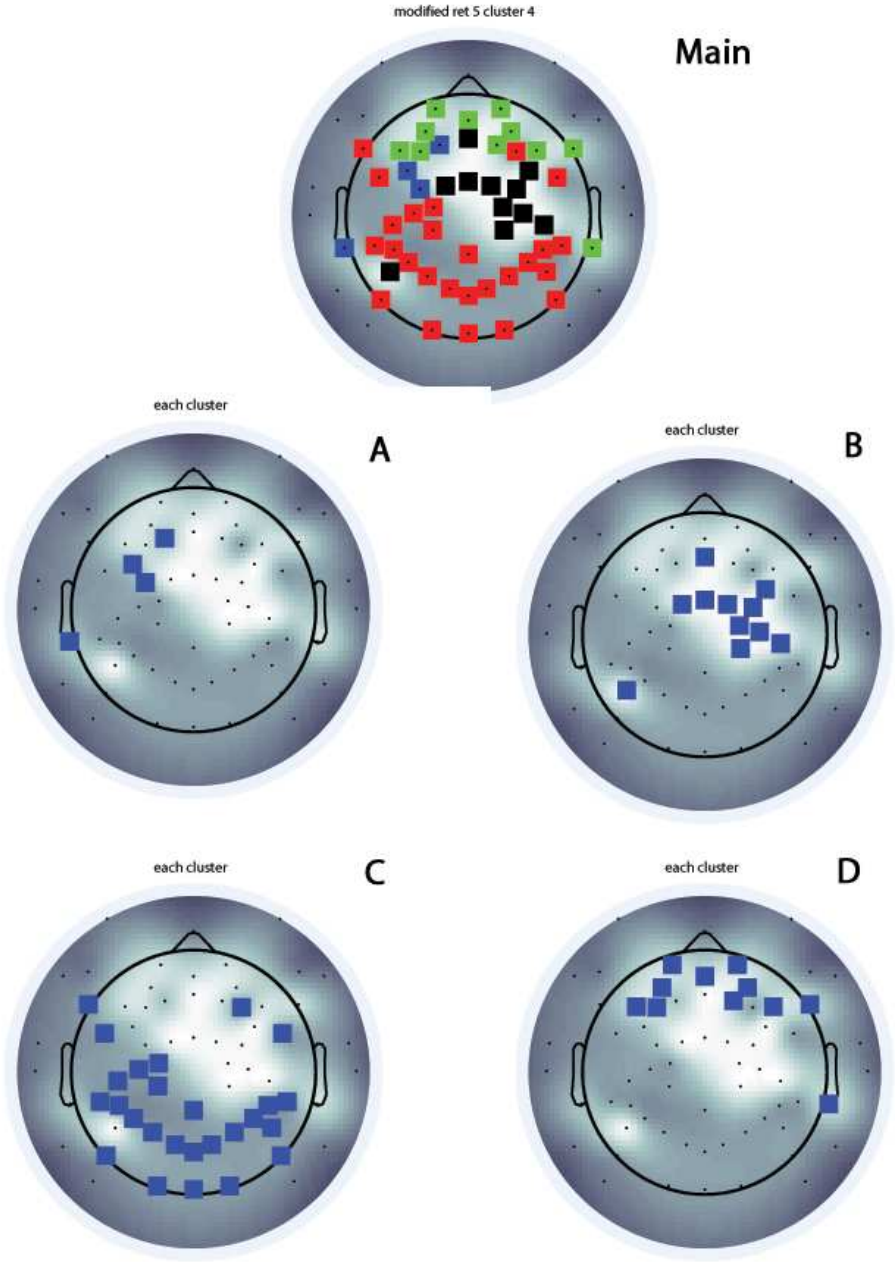

Figure 4.6 Modified RET 5 Box Model All Clusters and Separate Clusters

In Figure 4.6, there were 3 main clusters in the 5 box model. In the cluster D, there are nodes which are projected on the prefrontal regions of the brain. In addition, cluster C consists of increased group of nodes which can be observed as a combination of left parietal and occipital regions. Cluster B is mainly grouped over right parietal region.

Statistically significant nodes are represented as topological plots over different memory box models to investigate the significance of the cluster credibility and the relation between memory and functional clustering in terms of cognitive functioning.

In Figure 4.7, modified electrode locations which are generated by removing face and ear electrodes from default electrode locations are shown. Their clustering information and the statistically significant electrode node locations within their clustering information are plotted over topographic maps. In 1 box model, several prefrontal and left parietal electrodes are observed within two separate cluster groups in the statistically significant map. In addition, there are occipital electrodes which are generally grouped around the left occipital region of the brain. In the 3 box model, statistically significant electrodes are located over the left parietal region and the occipital cluster remained as same as 1 box model. However, in the 5 box model, statistically significant electrodes are located as a slightly huge cluster which can be observed as a combination of the left parietal and the occipital region of the brain due to the increased memory load. Furthermore, unrelated significant nodes and small group of clusters are located in right parietal and prefrontal regions.

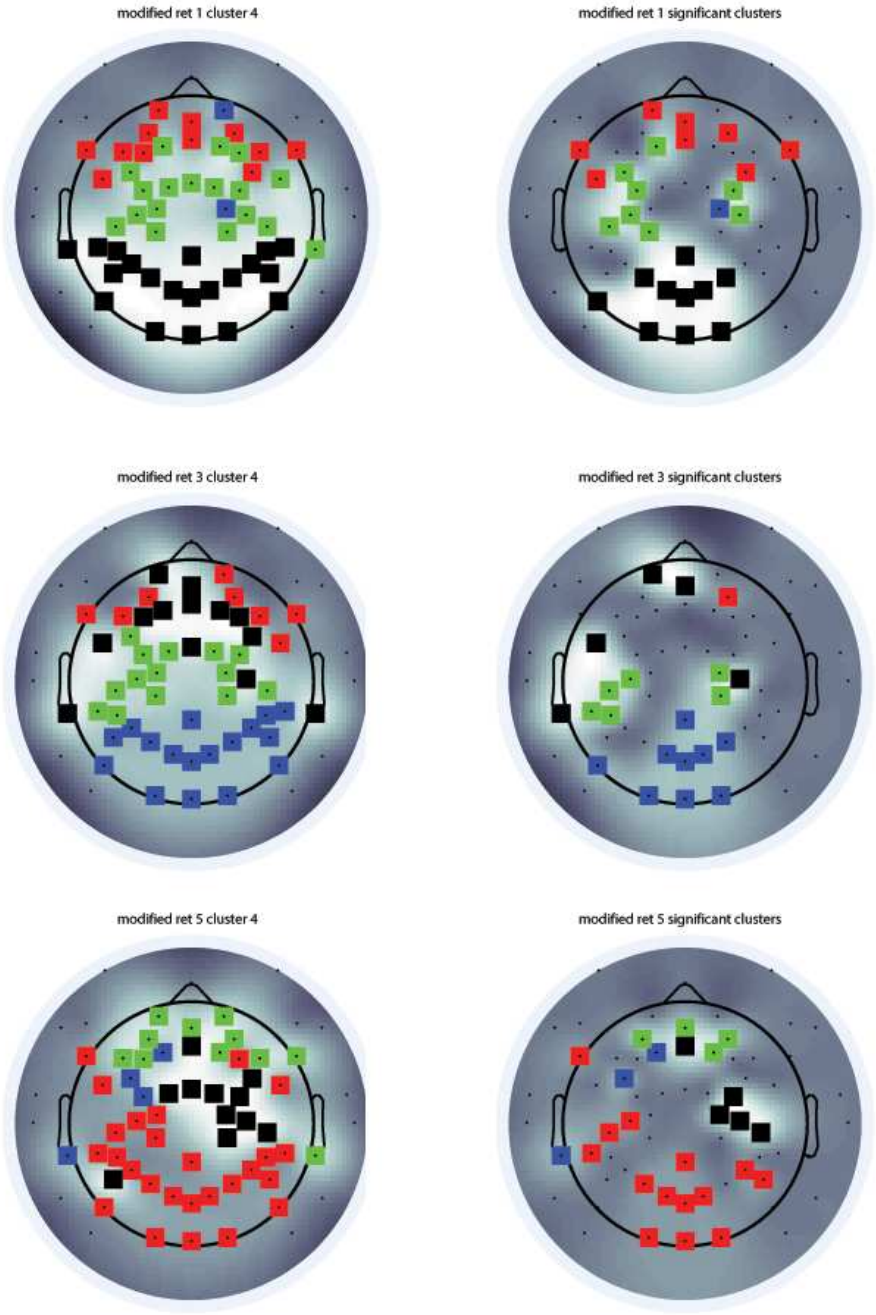

Figure 4.7 Modified Electrodes RET 1,3,5 Box Model Significant Nodes and Their Clusters

4.2.2 Soft Clustering Results

As in the N-Cut clustering, Preprocessing procedure is implemented by Netstation software. Preprocessed EEG segments which are selected as only good segments, are converted into .mat files to apply following procedures in Matlab software. Each segmented signal was converted into a time-frequency distribution with the help of Cohen class distributions [3.1]. To implement coherent analysis due to the temporal organization of the memory related brain networks, the 2500 ms retention interval is divided into two 1250 ms long time intervals. To separate the frequency interval of the signal, specified frequency bins which are directly related with default alpha band frequencies (8 Hz-12 Hz) are taken into consideration. Mutual information is calculated by Equation 3.8 to generate adjacency matrices.

In the soft clustering approach, only modified electrode montage [Figure 4.3] is used to determine mutual information based adjacency matrices. With respect to modified electrode locations, Soft Clustering algorithm is implemented over subjects with respect to the change in two time intervals and three task conditions. Using Equation 3.15 Laplacian vectors are computed to determine normalized cut values of the graph. Hence, the N-Cut values which are used to cluster the graph into sub-graphs are computed for all clustering levels using Equation 3.16.For each subject adjacency matrix is repetitively clustered for all previously calculated N-Cut values which can be considered as using N-Cut value to threshold the level of cluster computation. In addition, clustered adjacency matrices are saved to determine subject-wise soft clustering matrix. Subject-wise soft clustering matrix is calculated from the number of node existence within a specific cluster. Thus, if a node is eager to exists in the same cluster at various clustering levels, edge weights are supposed to be high valued. The mentioned procedure is repeated over each experiment condition and two time intervals.

Over 17 subject-wise soft cluster matrices, 1000 pseudo subject-wise soft cluster matrix combinations are generated. Pseudo combinations which can comprise repetitive sequences are randomly distributed. After the generation of simulation database, overall mean is calculated over subject-wise soft clusters ($1000 \times 17 \times 52 \times 52$), and

for further soft clustering analysis 1000 group clusters ($1000 \times 52 \times 52$) are generated. Similarly, to implement group-wise soft clustering analysis, from each subject-wise soft clustered matrix in simulation database, N-Cut values are computed. Minimum number of nodes inside the end clusters are set to four and group-wise soft clusters are generated with the help of previously computed N-Cut thresholds.

In Figure 4.8, modified electrode locations which are generated by removing face and ear electrodes from default electrode locations can be observed. After the procedure of subject-wise and group-wise soft clustering approach, clustering information are plotted over maps with respect to two different time intervals and three different task conditions. In addition, after group-wise soft clustering, all connections between various clusters can be observed. In Figure 4.8, maps A, C, E are representing first 1250 ms time interval in the retention period for $1, 3$ and 5 box combinations, respectively. First of all, with respect to the memory load, the number of soft clusters are increasing inversely proportional to the number of boxes which are represented in the probe stimulus. In addition, the form of the occipital and left bi-lateral clusters due to the memory load are not changing in the first time interval. However, the number of clusters located on prefrontal and left-bilateral regions are decreasing with respect to the number of boxes. In maps A and C, left prefrontal regions and left-parietal regions have several one-node clusters (for A; cluster 9, for C; cluster 8). However, in the map E, it can be observed as more stable clusters with respect to memory load. On the other hand, maps B, D, F are representing the second 1250 ms time interval in the retention period for $1, 3$ and 5 box combinations. Without thresholding with random matrices, increased inter-cluster connections (connections between different cluster pairs except intra-cluster connections) due to the different temporal organization can be observed from networks in the three different task condition. In addition, the occipital and prefrontal clusters, are not changing in the second time interval. However, node sizes are changing due to the degree variations which is obtained by using Equation 3.9. In the D map, the cluster which is located on the left parietal region consist of nodes that have high degrees with respect to number of inter-cluster and intra-cluster connections.

In Figure 4.9, apart from the previous maps, the final adjacency matrices are

thresholded with random adjacency matrices which are generated by preserving degree distribution with original adjacency matrices and soft clustered. In addition, node sizes are changing due to the degree variations which are obtained by using Equation 3.9.In Figure 4.9, maps A, C, E are representing first 1250 ms time interval in the retention period for $1, 3$ and 5 box combinations. The results which are obtained from Figure 4.8, have similar outcomes with Figure 4.8 except the differences between degrees of the each node. For instance, in the first time interval, A map has dominant nodes which are located on occipital, left lateral regions. On the other hand, in C map, there are several high degree nodes located on prefrontal region but, overall degree distribution can be observed as stable. In E map the over all degree distribution of each node increased with respect to the increase in intra and inter cluster connections. On the other hand, maps B, D, F are representing the second 1250 ms time interval in the retention period for $1, 3$ and 5 box combinations with random graph thresholding. As same as original network topologies which can be observed in Figure 4.8, increased inter-cluster connections directly proportional to the increased memory demand can be observed from networks.

In Figure 4.10, the soft clustered and random thresholded matrices are fixed to the number of six clusters in order to observe coherent results in the form of clusters. In the first time interval, the organization of bilateral and occipital clusters are observed as the same formation. However, their degree distribution (node sizes) changes over memory load. For instance, in E map, left lateral cluster is representing increased inter-cluster connections with occipital cluster. However, except for the C map, prefrontal cluster organization consists of 3 main clusters in both A and E maps. In the second time interval, all prefrontal, occipital and bilateral clusters, are representing the similar behavior and formation. The inter-cluster connections between left lateral and occipital clusters are obtained from the B and C maps. Although, in the F map, the linkage between left lateral and occipital cluster can be observed with respect to the memory load. In addition, several high degree nodes can be observed from the clusters located on left lateral and parietal regions.

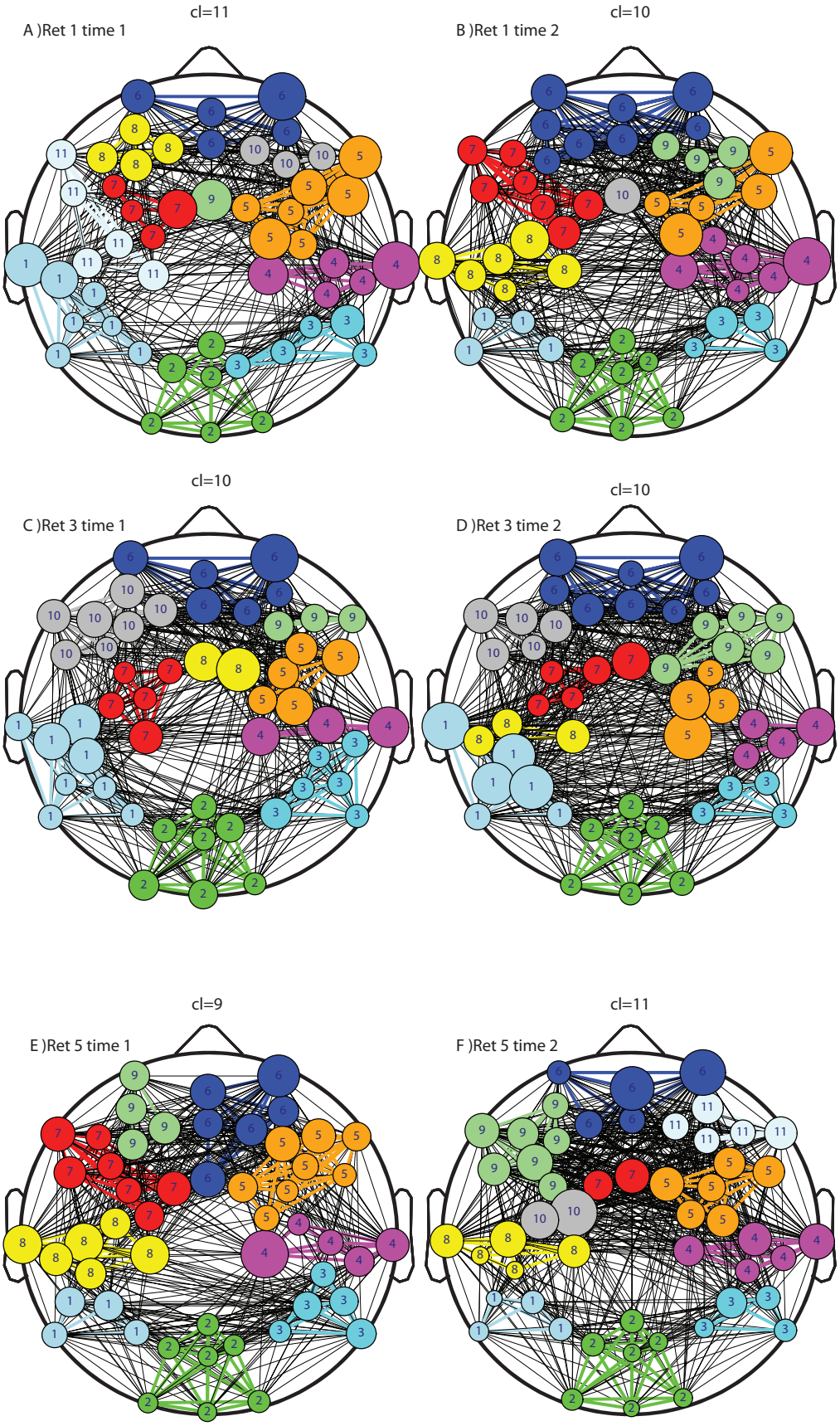

Figure 4.8 Soft Clusters with respect to color thresholds $A - C - D$ representing first time interval in 3 tasks (retention 1 box, retention 3 box, retention 5 box), $B - E - F$ representing second time interval in 3 tasks (node sizes are changing due to the degree variations which is obtained by using equation 3.9)

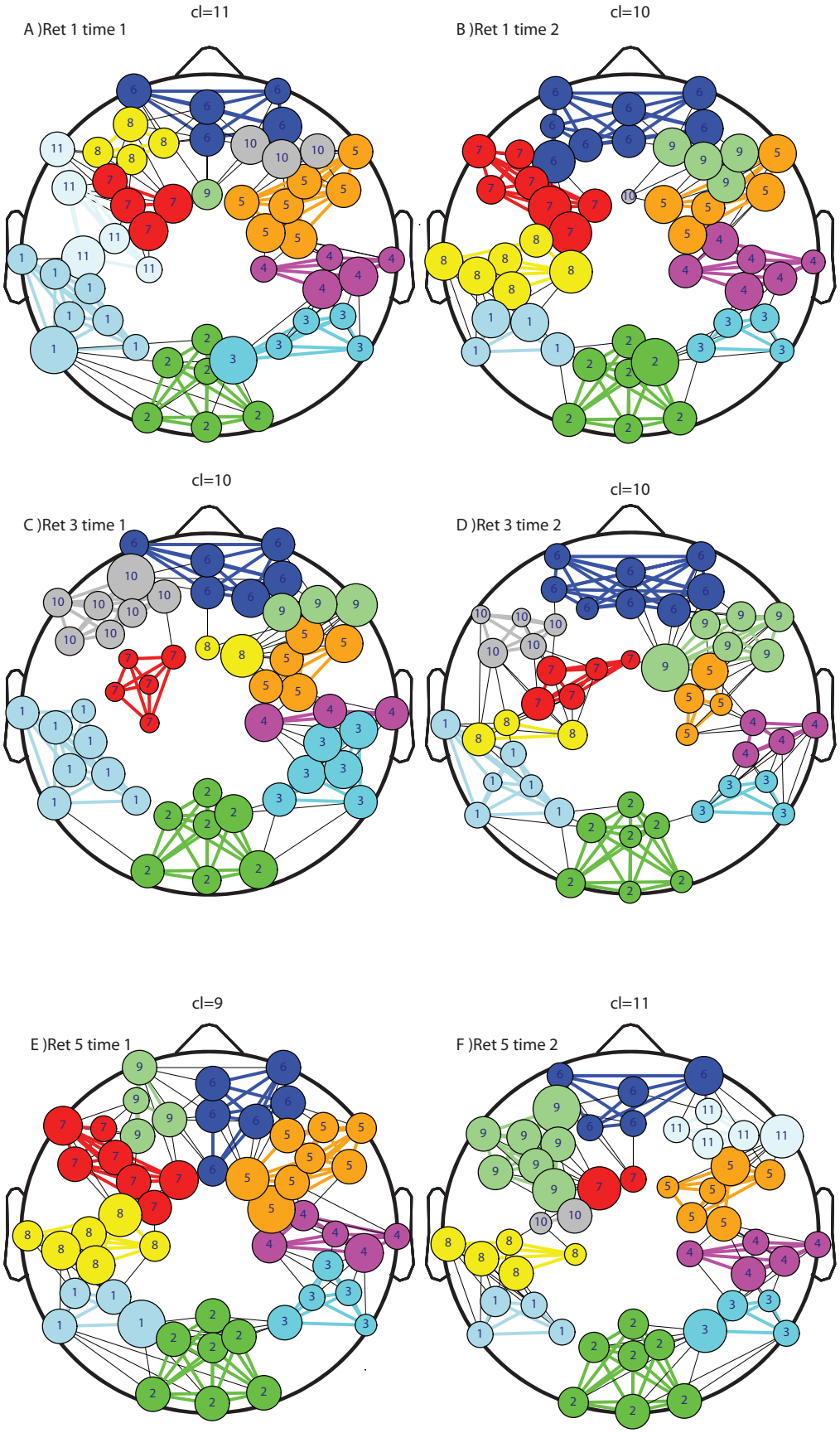

Figure 4.9 Soft Clusters with respect to random graph threshold, $A - C - D$ representing first time interval in 3 tasks (retention 1 box, retention 3 box, retention 5 box), $B - E - F$ representing second time interval in 3 tasks (node sizes are changing due to the degree variations which is obtained by using equation 3.9)

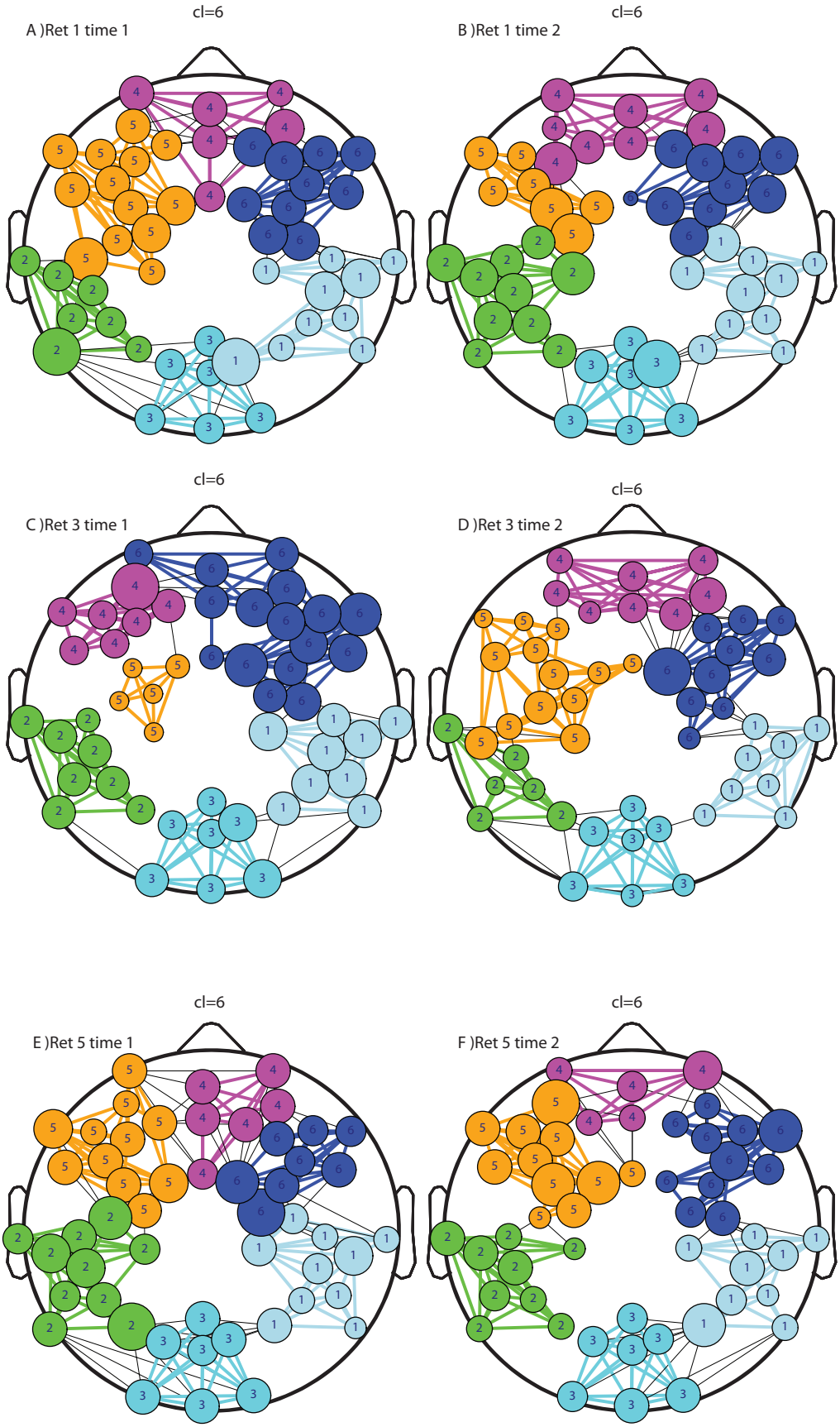

Figure 4.10 6 - Soft Clusters with respect to random graph threshold, $A - C - D$ representing first time interval in 3 tasks (retention 1 box, retention 3 box, retention 5 box), $B - E - F$ representing second time interval in 3 tasks (node sizes are changing due to the degree variations which is obtained by using equation 3.9)

5. DISCUSSION

5.1 Mean Reaction Time

In Figure 4.1, shows the difference between $1, 3, 5$ box model via mean reaction time values. Ole Jensen [23] assumed that difference in the mean reaction time can be considered as a function of memory load. Our experiment design revealed an increase of mean reaction time due to the increase in the number of boxes.

5.2 Modified Electrode Montage Design

There are several studies which analyzes the effect of the memory load on the alpha band. Tuladhar et al. [39] observed, a parametric increase in alpha band activity over posterior brain with increasing memory load. Jensen et al. [23] assumed that there were separate memory related sources which were located on the posterior and bilateral regions of the brain. In addition, they stated that, the memory load activity on the separate brain regions could reflect the degree of synchronization across multiple brain regions or number of regions involved. In Figure 4.7, modified electrode locations and their statistically significant nodes may provide an evidence for the previous studies. In other words, when the number of boxes increases, the number of clustered nodes which are projected on occipital and bilateral regions increases. Furthermore, in the left parietal and occipital clusters are observed separately in 1 and 3 box model and in the 5 box model, they are observed in a huge cluster which may be formed with respect to the increased memory load.

5.3 Memory load dynamics in optimum number of clusters

The parametric power increase in the first time interval with respect to the memory load was reported in the posterior regions of the brain [23]. This can be linked to the active inhibition of neural activity [40, 41]. There are several PET and fMRI experiments that have previously indicated that several prefrontal and parietal regions are involved in the working memory maintenance. In Figure 4.8 during the first time interval, number of clusters in the prefrontal region is decreasing with respect to the memory load which can be described as an inhibition of neural anterior activity [42]. In Figure 4.8, A, C and D maps, the number of clusters in the left lateral region are increasing (E map has two clusters but A and C maps have one cluster in the left lateral region). Furthermore, in Figure 4.9, E map is presenting that number of inter cluster nodes in the left lateral region are increasing which may provide evidence for the previously mentioned study [5].

There are several studies which stated that, memory load could produce enhancements in the posterior and lateral regions of the brain [23]. In Figure 4.10, B, D and F maps, the inter-cluster connections between the left lateral and occipital clusters are decreasing in the second time interval which can be linked to the enhancement of posterior region and fissure between the parietal - occipital region.

6. CONCLUSION AND FUTURE WORK

The EEG signal and its oscillatory activity are strongly related with temporal modulation of information processing. Thus, the brain has a complex structure and a complex functioning capability. Band and region specific studies are made to observe the overall brain functioning. Generally, oscillations of alpha (8 Hz - 12 Hz) band of the EEG rhythm and the amplitude of the signal changes in the memory and cognitive based tasks are observed. Furthermore, to obtain local and global interaction of memory processing, clustering analysis can be considered as important studies.

The sophisticated architecture of brain networks can be observed with the help of spectral clustering algorithm to create the dynamic interactions between regions like clusters and the directions of information flow from one cluster to another. In this thesis, a spectral clustering algorithm was used to parcellate memory related circuits in the brain in a load-dependent manner. To be able to circumvent the problem of choosing the number of clusters beforehand a soft clustering approach was implemented. To investigate both the spatial and the temporal change in terms of functional dynamics on the brain, EEG-fMRI fusion studies are incorporated into our ongoing schedule. Since the fMRI part of the future project, consists huge raw databases to cluster, our proposed method, the soft clustering algorithm, has to be updated with respect to computational ease and time.

In the end, the main future work will rely on developing an enhanced soft clustering algorithm which will be used as an intermediary tool to establish the link between the temporal and the spatial functional dynamics of the brain.

APPENDIX A. Code Descriptions

A.1 Cohen Class Time Frequency Distribution and Mutual Information Calculation

A.1.1 *pre_main.m* - Matlab Code

Description: Allocates the memory with respect to two time intervals and three task conditions. Stands for calculating the Cohen Class Based Time-Frequency distribution matrices from segmented time segments for each task condition and two time intervals (first $1250ms$, second $1250ms$).

Outputs: *mutual_ret*1 ($NumSubxNumNode \times NumNode$) TFD based Mutual Information adjacency matrices.

Inputs: *core_csegret*1 ($NumSub \times NumSeg \times TimeInt$) Time segments for each task condition and two time intervals.

Call: *mutual_adjacency.m*

Called by: Command line

A.1.2 *mutual_adjacency.m* - Matlab Code

Description: Re-samples the previously calculated segments (Preprocessing Netstation) from $1Hz$ to $250Hz$ and determines the Wigner-Ville TFD of alpha frequency bins ($8 - 12Hz$).

Outputs: *mutual_ret*1 ($NumSub \times NumNode \times NumNode$), TFD adjacency

matrix of 17 subjects with respect to each task condition.

Inputs: *core_csegret*1 ($NumSub \times NumSeg \times TimeInt$) Time segments for each task condition and two time intervals. Group of segments gathered from 17 subjects and each ($1, 3$ and 5) task condition.

Call: *resample.m*, *re_sizer.m*, *mutdene.m*

Called by: *pre_main.m*

A.1.3 *re_sample.m* - Matlab Code

Description: Uses the *resample.m* function from the signal processing toolbox if present. It down-samples the time segments from $1000Hz$ to $250Hz$ frequency interval. It also uses the *pop_resample.m* in EEGLAB Toolbox [43] of Matlab.

Outputs: *temp* ($NumSub \times NumSeg \times TimeInt$), down-sampled time segments.

Inputs: *segments* ($NumSub \times NumSeg \times TimeInt$), group of segments gathered from 17 subjects and each ($1, 3$ and 5) task condition.

Call: *resample.m*

Called by: *mutual_adjacency.m*

A.1.4 *re_sizer.m* - Matlab Code

Description: Reshapes the time segments from the format of $17x52x1250$ to $17x65000$ in order to obtain histogram based TFD calculation.

Outputs: *resized* ($NumSub \times MulSeg\_by\_TimeInt$), resized time segments.

Inputs: *temp* ($NumSub \times NumSeg \times TimeInt$), down-sampled time segments.

Call: None

Called by: *mutual_adjacency.m*

A.1.5 *mutdene.m* - Matlab Code

Description: Calculates the TFD by Equation 3.3 and determines the TFD based Mutual information by Equation 3.8. It sets the default alpha band frequency ($8 - 12Hz$) bins and calculates the Mutual Information.

Outputs: *alpha_ret*1_*t*1 ($NumSub \times NumNode \times NumNode$) TFD based adjacency matrices for each subject, task condition and time interval.

Inputs: *resized* ($NumSub \times MulSeg\_by\_TimeInt$), resized time segments.

Call: $tfrwv.m$

Called by: *mutual_adjacency.m*

A.1.6 $tfrwv.m$ - Matlab Code to compute Wigner-Ville distribution

Description: Calculates the Wigner-Ville distribution based TFD by using 3.3. It uses the function in the Time Frequency Toolbox [44] of Matlab.

Outputs: $tfr1$ time frequency representations of given time segment, $t1$ frequency bins, $f1$ vector of normalized frequencies.

Inputs: X ($MulSeg\_by\_TimeInt$) Time segment, T time instant (length of time segment), N number of frequency bins (250 bin was selected).

Call: None

Called by: *mutdene.m*

A.2 N-Cut Data Clustering

A.2.1 *ncut_main.m* - Matlab Code

Description: Reallocates the memory for calculated clustered matrices. It sets the diagonal vector of the TFD adjacency matrix. It normalized the adjacency matrix before the clustering process. It converts the $64x64$ EEG electrode locations to $52x52$. It both computes the $Ncut$ clustering and bootstrap $Ncut$ clustering.

Outputs: mod_bret1_clu ($ModNode \times CluNumNodes$), adjacency matrix of $Ncut$ clusters for each task condition.

Inputs: $alpha\_ret1$ ($NumSub \times NumNode \times NumNode$), previously calculated TFD based Mutual Information adjacency matrices for each task condition.

Call: *all_cluster.m*, *all_bootsrt_ex.m*

Called by: Command Line

A.2.2 *all_cluster.m* - Matlab Code

Description: Computes the subject-wise $Ncut$ Clustering.

Outputs: cl_first ($NumSub \times ModNumNode \times ModNumNode$), subject-wise soft clustered adjacency matrices for each task condition

Inputs: $first$ ($NumSub \times ModNumNode \times ModNumNode$), normalized TFD Mutual Information adjacency matrices for each task condition, $num_cluster = 4$, predefined number of clusters, $num_ch = 52$, number of electrode channels.

Call: $cag\_shi\_mod.m$

Called by: $ncut\_main.m$

A.2.3 $all\_bootsrt\_ex.m$ - Matlab Code

Description: Computes the group-wise $Ncut$ Clustering. Implements bootstrap statistical test.

Outputs: cl_second ($NumSub \times ModNumNode \times ModNumNode$) Group-wise soft clustered adjacency matrices for each task condition

Inputs: cl_first ($NumSub \times ModNumNode \times ModNumNode$), Normalized TFD Mutual Information adjacency matrices for each task condition, $num_cluster = 4$, Predefined number of clusters, $num_ch = 52$, number of electrode channels.

Call: $cag\_shi\_mod.m$

Called by: $ncut\_main.m$

A.2.4 *cag_shi_mod.m* - Matlab Code

Description: Clusters the given subject-wise adjacency matrices by using the N-cut clustering algorithm [34] with predefined number of clusters. It converts the original adjacency matrix into sparse matrix formation before the clustering implementation.

Outputs: cl_second ($NumSub \times ModNumNode \times ModNumNode$) Group-wise soft clustered adjacency matrices for each task condition

Inputs: cl_first ($NumSub \times ModNumNode \times ModNumNode$), Normalized TFD Mutual Information adjacency matrices for each task condition, $num_cluster = 4$, Predefined number of clusters, $num_ch = 52$, number of electrode channels.

Call: *ncutW.m*

Called by: *ncut_main.m*

A.2.5 *ncutW.m* - Matlab Code

Description: Uses Shi Malik algorithm [34]. Calls *ncut.m* function to compute $NcutEigenvectors$ and $NcutEigenvalues$ of W with *nbcluster* clusters,computes continuous Ncut eigenvectors and computes discretize Ncut vectors.

Outputs: $NcutDiscrete$ ($NumSub \times ClusNumNode$) discretize $Ncut$ vectors, $NcutEigenvectors$ and $NcutEigenvalues$ of input W matrix with respect to predefined number of clusters.

Inputs: W ($ModNumNode \times ModNumNode$), normalized TFD Mutual Information adjacency matriX of individual subject, $nbcluster = 4$, Predefined number of clusters.

Call: *ncut.m*

Called by: *cag_shi_mod.m*

A.3 Soft Clustering

A.3.1 *post_main.m* - Matlab Code

Description: Reallocates the memory for calculated clustered matrices. It sets the diagonal vector of the TFD adjacency matrix. It normalized the adjacency matrix before the clustering process. It converts the $64x64$ EEG electrode locations to $52x52$. It both computes the first soft clustering and bootstrap soft clustering. It repeats the process for two time intervals.

Outputs: mod_bret1_clu ($CluNumNodes \times CluNumNodes$), adjacency matrix of bootstrap soft clusters for each task condition.

Inputs: $alpha\_ret1\_t1$ ($NumSub \times NumNode \times NumNode$), reviously calculated TFD based Mutual Information adjacency matrices for each task condition and time interval.

Call: *mainth.m*, *all_bootstr_exex.m*

Called by: *pre_main.m*

A.3.2 *mainth.m* - Matlab Code

Description: Gets the preprocessd TFD adjacency matrices and decomposes the adjacency matrix of an individual subject into all possible N-Cut value by using Equation 3.15. It repetitively clusters the adjacency matrix for all previously calculated

N-Cut values which can be considered as using N-Cut value to threshold the level of cluster computation. In addition, clustered adjacency matrices are saved to determine subject-wise soft clustering matrix. Subject-wise soft clustering matrix is calculated from the number of existing nodes within a specific cluster. Thus, if a node is eager to exists in the same cluster at various clustering levels, edge weights are supposed to be high valued.

Outputs: $m\_mainvec\_r1$ ($NumSub \times ModNumNode \times ModNumNode$), subject-wise soft clustered adjacency matrices for each task condition.

Inputs: mod_alpha_ret1 ($NumSub \times ModNumNode \times ModNumNode$), normalized TFD Mutual Information adjacency matrices for each task condition.

Call: $ncutter\_oycag.m$

Called by: $post\_main.m$

A.3.3 $ncutter\_oycag.m$ - Matlab Code

Description: Calculates the all possible N-Cut values of given adjacency matrix and clusters the matrix with respect to given N-Cut value.

Outputs: $ncutlist$, index of all possible N-Cut values, $mainvec$ ($NumNcut \times CluAdj$), adjacency matrix which is clustered with respect to given N-Cut value.

Inputs: W ($ModNumNode \times ModNumNode$), Adjacency matrix, $MinNosNodes$ (N=4), minimum number of nodes to exist as one cluster, threshold (TH=1.35 example), N-Cut value to cluster the adjacency matrix.

Call: None

Called by: *mainth.m*

A.3.4 *all_bootstr_exex.m* - Matlab Code

Description: Generates 1000 subject-wise soft cluster matrix combinations over 17 subject-wise soft cluster matrices. Pseudo combinations which can comprise repetitive sequences are randomly distributed. After the generation of simulation database, overall mean is calculated over subject-wise soft clusters ($1000x17x52x52$), and for the further soft clustering analysis 1000 group clusters ($1000x52x52$) are generated. It clusters the previously calculated group matrices due to soft clustering computation.

Outputs: *group_temp* ($NumModNode \times NumModNode$),group-wise clustered adjacency matrix.

Inputs: *cl_first* ($NumSub \times NumModNode \times NumModNode$), clustered subject wise adjacency matrices for each task condition.

Call: *mainth.m*

Called by: *post_main.m*

A.3.5 *random_yeni.m* - Matlab Code

Description: Generates random graphs by preserving degree distribution with original adjacency matrices, and clusters by soft clustering method. Implements subject-wise clustering and group-wise soft clustering analysis.

Outputs: *rand_mod_bret1_clu*, ($CluNumNodes \times CluNumNodes$), random adjacency matrix of bootstrap soft clusters for each task condition.

Inputs: $alpha\_ret1\_t1$ ($NumSub \times NumNode \times NumNode$), previously calculated TFD based Mutual Information adjacency matrices for each task condition and time interval.

Call: $mainth.m$, $all\_bootstr\_exex.m$

Called by: $pre\_main.m$

A.3.6 $all\_bootstr\_random.m$ - Matlab Code

Description: Computes the group-wise soft clustering of previously generated random adjacency matrices

Outputs: $group\_temp$ ($NumModNode \times NumModNode$), group-wise clustered adjacency matrix

Inputs: cl_first ($NumSub \times NumModNode \times NumModNode$) clustered subject wise adjacency matrices for each task condition

Call: $mainth.m$

Called by: $post\_main.m$

REFERENCES

1. Wang, Q., O. Sourina, and M. K. Nguyen, "EEG-based Serious Games Design for Medical Applications," *Dimension Contemporary German Arts And Letters*.

2. Boutros, N., and S. Galderisi, "Standard electroencephalography in clinical psychiatry: a practical handbook," *Acta Psychiatrica Scandinavica*, Vol. 124, no. 3, pp. 239–240, 2011.

3. Atkinson, R., and R. Shiffrin, "Human memory: A proposed system and its control processes," Vol. 2 of *Psychology of Learning and Motivation*, pp. 89 – 195, Academic Press, 1968.

4. Busch, N. A., and C. S. Herrmann, "Object-load and feature-load modulate EEG in a short-term memory task," *Memory*, Vol. 14, no. 13, pp. 15–18, 2003.

5. Sauseng, P., W. Klimesch, M. Doppelmayr, T. Pecherstorfer, R. Freunberger, and S. Hanslmayr, "EEG alpha synchronization and functional coupling during top-down processing in a working memory task.," *Human brain mapping*, Vol. 26, pp. 148–55, Oct. 2005.

6. Donath, W. E., and A. J. Hoffman, "Lower bounds for the partitioning of graphs," *IBM Journal of Research and Development*, Vol. 17, pp. 420 –425, sept. 1973.

7. Fiedler, M. *Czechoslovak Mathematical Journal*, Vol. 23, no. 2, pp. 298–305, 1973.

8. Luxburg, U., "A tutorial on spectral clustering," *Statistics and Computing*, Vol. 17, pp. 395–416, Aug. 2007.

9. Shen, X., X. Papademetris, and R. T. Constable, "Graph-theory based parcellation of functional subunits in the brain from resting-state fMRI data.," *NeuroImage*, Vol. 50, pp. 1027–35, Apr. 2010.

10. Manual, T., "Net Station Viewer Technical Manual,"

11. Nunez, P. L., *Electric fields of the brain : the neurophysics of EEG / Paul L. Nunez ; with contributions by Ron D. Katznelson*, Oxford University Press, New York :, 1981.

12. Kane, M. J., and R. W. Engle, "The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective.," *Psychonomic bulletin & review*, Vol. 9, pp. 637–71, Dec. 2002.

13. Stipacek, A., R. H. Grabner, C. Neuper, A. Fink, and A. C. Neubauer, "Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load," *Neuroscience Letters*, Vol. 353, pp. 193–196, 2003.

14. Fuster, J. M., *Cortex And Mind: Unifying Cognition*, Oxford University Press, Sept. 2005.

15. Rose, S., *The Making of Memory: From Molecules to Mind*, Anchor, Sept. 1993.

16. Başar, E., *Memory and brain dynamics: oscillations integrating attention, perception, learning, and memory*, Conceptual advances in brain research, CRC Press, 2004.

17. Klimesch, W., "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis.," *Brain research. Brain research reviews*, Vol. 29, pp. 169–95, Apr. 1999.

18. Klimesch, W., "EEG-alpha rhythms and memory processes.," *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, Vol. 26, pp. 319–40, June 1997.

19. Basar, E., "Memory as the whole brain work: A large-scale model based on oscillations in super-synergy," *International Journal of Psychophysiology*, Vol. 58, no. 23, pp. 199 – 226, 2005.

20. Pfurtscheller, G., "Functional brain imaging based on ERD/ERS.," *Vision research*, Vol. 41, pp. 1257–60, Jan. 2001.

21. Pijnenburg, Y. a. L., Y. v d Made, a. M. van Cappellen van Walsum, D. L. Knol, P. Scheltens, and C. J. Stam, "EEG synchronization likelihood in mild cognitive impairment and Alzheimer's disease during a working memory task.," *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, Vol. 115, pp. 1332–9, June 2004.

22. Haenschel, C., D. E. Linden, R. a. Bittner, W. Singer, and S. Hanslmayr, "Alpha phase locking predicts residual working memory performance in schizophrenia.," *Biological psychiatry*, Vol. 68, pp. 595–8, Oct. 2010.

23. Jensen, O., J. Gelfand, J. Kounios, and J. E. Lisman, "Oscillations in the alpha band (9-12 Hz) increase with memory load during retention in a short-term memory task.," *Cerebral cortex (New York, N.Y. : 1991)*, Vol. 12, pp. 877–82, Aug. 2002.

24. Fink, a., R. H. Grabner, C. Neuper, and a. C. Neubauer, "EEG alpha band dissociation with increasing task demands.," *Brain research. Cognitive brain research*, Vol. 24, pp. 252–9, July 2005.

25. Lundqvist, M., P. Herman, and A. Lansner, "Theta and Gamma Power Increases and Alpha/Beta Power Decreases with Memory Load in an Attractor Network Model," *Journal of Cognitive Neuroscience*, Vol. 23, no. 10, pp. 3008–3020, 2011.

26. Moddemeijer, R., "On estimation of entropy and mutual information of continuous distributions," *Signal Processing*, Vol. 16, no. 3, pp. 233–248, 1989.

27. Mormann, F., "Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients," *Physica D: Nonlinear Phenomena*, Vol. 144, pp. 358–369, Oct. 2000.

28. Rosso, O., M. Martin, and a. Plastino, "Brain electrical activity analysis using wavelet-based informational tools," *Physica A: Statistical Mechanics and its Applications*, Vol. 313, pp. 587–608, Oct. 2002.

29. Schlogl, a., C. Keinrath, R. Scherer, and P. Furtscheller, "Information transfer of an EEG-based brain computer interface," *First International IEEE EMBS Conference on Neural Engineering, 2003. Conference Proceedings.*, no. 1 0, pp. 641–644, 2003.

30. Aviyente, S., "Information-theoretic signal processing on the time-frequency plane and applications," in *Proc. of EUSIPCO*, pp. 4–8, Citeseer, 2005.

31. Aviyente, S., "A measure of mutual information on the time-frequency plane," in *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on*, Vol. 4, pp. iv–481, IEEE, 2005.

32. Lu, C.-F., S. Teng, C.-I. Hung, P.-J. Tseng, L.-T. Lin, P.-L. Lee, and Y.-T. Wu, "Reorganization of functional connectivity during the motor task using EEG time-frequency cross mutual information analysis.," *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, Vol. 122, pp. 1569–79, Aug. 2011.

33. Dauwels, J., F. Vialatte, T. Musha, and a. Cichocki, "A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG.," *NeuroImage*, Vol. 49, pp. 668–93, Jan. 2010.

34. Shi, J., and J. Malik, "Normalized cuts and image segmentation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 22, no. 8, pp. 888–905, 2000.

35. Chen, V., and S. Ruan, "Graph Cut Based Segmentation of Brain Tumor From MRI Images," Vol. 3, no. December, pp. 1054–1063, 2009.

36. Gratton, G., "Dealing with artifacts: The eog contamination of the event-related brain potential," *Behavior Research Methods*, Vol. 30, pp. 44–53, 1998. 10.3758/BF03209415.

37. Bullmore, E., and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems.," *Nature reviews. Neuroscience*, Vol. 10, pp. 186–98, Mar. 2009.

38. Bellec, P., V. Perlbarg, and A. C. Evans, "Bootstrap generation and evaluation of an fMRI simulation database.," *Magnetic resonance imaging*, Vol. 27, pp. 1382–96, Dec. 2009.

39. Tuladhar, A. M., N. ter Huurne, J.-M. Schoffelen, E. Maris, R. Oostenveld, and O. Jensen, "Parieto-occipital sources account for the increase in alpha activity with working memory load.," *Human brain mapping*, Vol. 28, pp. 785–92, Aug. 2007.

40. Jokisch, D., and O. Jensen, "Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream.," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, Vol. 27, pp. 3244–51, Mar. 2007.

41. Klimesch, W., P. Sauseng, and S. Hanslmayr, "EEG alpha oscillations : The inhibition-timing hypothesis," *Review Literature And Arts Of The Americas*, Vol. 3, 2006.

42. Scheeringa, R., K. M. Petersson, R. Oostenveld, D. G. Norris, P. Hagoort, and M. C. M. Bastiaansen, "Trial-by-trial coupling between EEG and BOLD identifies networks related to alpha and theta EEG power increases during working memory maintenance.," *NeuroImage*, Vol. 44, pp. 1224–38, Feb. 2009.

43. Delorme, A., "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, Vol. 134, pp. 9–21, Mar. 2004.

44. Auger, F., P. Flandrin, P. Gonçalvès, and O. Lemoine, "Time-frequency toolbox," *CNRS France-Rice University*, 1996.