A NOVEL APPROACH IN THE DETECTION OF OBSTRUCTIVE SLEEP APNEA FROM ELECTROCARDIOGRAM SIGNALS USING NEURAL NETWORK CLASSIFICATION OF TEXTURAL FEATURES EXTRACTED FROM TIME-FREQUENCY PLOTS

by

MOHAMMAD AHMAD AL-ABED

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN BIOENGINEERING

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2006

ACKNOWLEDGEMENTS

I would like to extend my deepest appreciation and gratitude to Dr. Khosrow Behbehani, Professor and Chair, Department of Bioengineering at UTA. He has been a mentor and role model for me, and his support, patience and guidance have a profound effect on me and have assisted me throughout this research.

I would like to also thank Dr. Michael Manry, Professor and Director, Image Processing and Neural Networks Laboratory at the Department of Electrical Engineering. He has continually presented his support and advice to me and has been always willing to share his expertise and knowledge.

I would like to thank my research group partners at the Signal Processing Lab for their courteous understanding and continuous help whenever I needed it.

Mom, Dad, sisters and brothers: you have always been my inspiration. Your unlimited support has been a continuous motivation for me to move forward and fulfill your dream and mine. Thank you.

And I cannot thank you enough, my friend and partner, Nada. You believe in me, and it has been for you standing always by my side that this work came to life. WMLN.

November 14, 2005

ABSTRACT

A NOVEL APPROACH IN THE DETECTION OF OBSTRUCTIVE SLEEP APNEA FROM ELECTROCARDIOGRAM SIGNALS USING NEURAL NETWORK CLASSIFICATION OF TEXTURAL FEATURES EXTRACTED FROM TIME-FREQUENCY PLOTS

Publication No. ______

Mohammad Ahmad Al-Abed, M. S.

The University of Texas at Arlington, 2005

Supervising Professor: Khosrow Behbehani, Ph.D., P.E.

Sleep-Disordered Breathing (SDB) is estimated to have a prevalence of 5% in middle-aged population. The population is widely thought to be under diagnosed, since the present method to detect and diagnose SDB, Nocturnal Polysomnography (NPSG), is still expensive and not accessible by most. SDB has been shown to affect the productivity and degree of life of the patient, and to have a high correlation with obesity and cognitive heart failure (CHF). Cheap and accessible means to screen the population for SDB are greatly pursued. This work presents an automatic algorithm to detect obstructive sleep apnea (OSA) events in 15-minute clips. Data is collected from 12 normal subjects (6 males, 6 females; age 46.27±9.79 years, AHI 3.82±3.25) and 14 apneic subjects (8 males, 6 females; age 49.08±8.82 years; AHI 31.21±23.90). The algorithm uses textural features extracted from co-occurrence matrices of gray-level encoded images generated by short-time discrete Fourier transform (STDFT) of the heart rate variability (HRV). Seventeen selected features are used as inputs to a 3-layer multilayer perceptron (MLP), with 45 hidden units and 4200 training epochs. A 1000-run Monte-Carlo simulation of the algorithm gave the following results: mean training sensitivity, specificity and accuracy of 99.00%, 93.42%, and 96.42%, respectively. The mean testing sensitivity, specificity and accuracy are 94.42%, 85.40%, and 90.16%, respectively.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Figure Page

LIST OF TABLES

LIST OF ABBREVIATIONS

NOR ....Normal Subject Group

OSA...Obstructive Sleep Apnea Subject Group

SDB...Sleep Disorder Breathing

OSAHS ....Obstructive Sleep Apnea-Hypopnea Syndrome

ECG...Electrocadriogram

CHF...Congestive Heart Failure

AHI ....Apnea-Hypopnea Index

NPSG ....Nocturnal Polysomnography

HRV ....Heart Rate Variability

RWA ....R-Wave Attenuation

RPE ....R-Peaks Envelope

STDFT ....Short-Time Discrete Fourier Transform

GLCM..Gray-Level Co-occurrence Matrix

NGLCM.......................................Normalized Gray-Level Co-Occurrence Matrix

ENT...Entropy

ASM..Angular Second Moment

CON ....Contrast

COR ....Correlation

DIS ....Dissimilarity

IND ....Inverse Difference

IDM..Inverse Difference Moment

VAR ....Variance

INR...Inverse Recursivity

NN...Neural Network

PLN ....Piece-wise Linear Network

MLP ....Multilayer Perceptron

BP..Back-Propogation

SEM ....Standard Mean Error

FLS..Fuzzy Logic Systems

CHAPTER 1

INTRODUCTION

1.1 Definitions, Diagnosis, and Prevalence

1.1.1 Basic Definitions

Obstructive Sleep Apnea (OSA) is defined as the total cessation of breathing during sleep for 10 seconds or more [1]. This cessation is caused by an airway blockage in spite of a continuous respiratory nervous effort. This blockage results from an upper airway collapse at the level of the tongue and/or soft palate, due to a combination of anatomic factors [2], [3]. This event leads to a drop in blood oxygenation levels and sleep interruptions [4]. Obstructive Sleep Hypopnea (OSH) is defined as the partial cessation (50% or more) of breathing during sleep for 10 seconds or more [1]. OSA and OSH combined are referred to as Obstructive Sleep Apnea-Hypopnea Syndrome (OSAHS). The number of OSAHS events that are recorded for a patient during sleep, divided by the number of sleep hours is defined as the Apnea-Hypopnea Index, or AHI. This is the measure [1] adopted by the American Academy of Sleep Medicine task force as the criteria on which to describe the level of disorder intensity. Patients with AHI of 5 hr^{-1} or more may need medical intervention.

1.1.2 Prevalence and Diagnosis

The standard Sleep Disordered Breathing (SDB) diagnosis test [1] that includes diagnosis of OSAHS is the Nocturnal Polysomnography (NPSG). This test is comprised of overnight sleep recording of different physiological markers, in which a trained sleep expert will mark the stages of sleep and the type of SDB, if present. Young et al [3], [5] have published two studies in which they show an astonishing prevalence of OSAHS of a 4% in men and 2% in women in their first study [3], then in [5] show a higher population prevalence of 5%. However, due to the fact that NPSG is very expensive and inaccessible by most patients, it is widely believed that OSAHS is under-diagnosed and actual population prevalence is higher [2].

OSAHS has been linked to different pathological abnormalities [5] and remains a notable cause of lowered quality of sleep, decreased productivity, and a cause to many auto-vehicle accidents in the middle-age group. OSAHS has been shown to have a high correlation with obesity, especially in men, a 50% occurrence with congestive heart failure (CHF) patients, and 50-60% having systemic high blood pressure [6]. High blood pressure is directly linked to the chronic and continuous decrease in the blood saturation levels, causing an increase in sympathetic nerve activity, and an eventual increase in blood pressure [6].

1.2 Alternative Methods of Diagnosis

1.2.1 ECG as a Marker

In 1984, Guilleminault et al have showed a cyclical variation in electrocardiogram (ECG) that is closely correlated with events of OSAHS. This variation is characterized by a progressive decrease in heart rate (bradycardia) simultaneous with an OSAHS event, followed by an abrupt increase in heart rate (tachycardia) at arousal and onset of breath resumption [7].

This finding started an ongoing research to find alternative, cheap, and practical screening tools that physicians can use. Ever since, different features were extracted from ECG recordings with reliable detection accuracy for screening purposes.

1.2.2 ECG Features

Different research group have concentrated on different features of the ECG overnight recordings [2], [8], [13]. Some groups concentrated on time-domain characteristics of the ECG signal, such as Angle of Mean Electrical Axis [9], [10], [11], Heart Rate Variability (HRV) [2], ECG-Derived Respiration (EDR) [2], [12], and R-Peak Envelope (RPE) [8]. Other groups looked at frequency-domain parameters such as PSD [8], [13] and time-frequency plots [2], [14]. However, time-frequency plots have only been studied in a qualitative manner, in which some noted differences between normal subjects and OSAHS patients have been described, but not quantitatively studied.

1.3 Study Overview and Organization

1.3.1 Overview

This investigation is as a continuation of efforts done by [8], [15] within the UTA-OSA research group. It is an effort to bridge the time-frequency plots investigation described by [14] and provide a quantitative study of these plots. The aim is to improve the overall detection accuracy and provide a reliable and practical screening method for OSAHS candidate patients. The study concentrates on extracting gray-level pictorial plots from HRV 15-minute clips, and performs reliable statistical image processing and classification schemes to distinguish event clips from normal ones.

1.3.2 Study Organization

The second chapter of this study gives a detailed structure of the methodology of the automatic detection algorithm proposed by this study. Chapter Three presents the results of applying the available sleep ECG record to the algorithm. Chapter Four discusses these results and their significance. Chapter Five, concludes with what this research has answered and new directions for future studies.

CHAPTER 2

AUTOMATIC DETECTION ALGORITHM DESIGN

The first two sections of this chapter summarize the previous efforts by the UTA SDB-research group to gather and process the data used in this study. Section 2.3 describes the process in which pictorial images are produced from the selected data. Section 2.4 details the image processing and feature extraction scheme used. Section 2.5 illustrates the detection methods used. Section 2.6 is a summary of the chapter that includes a graphic depiction of the over all process.

Figure 2.1 shows a block diagram the outlines and structure of this study. The reader can find the section that explains in detail the content and contribution of each block to the overall system.

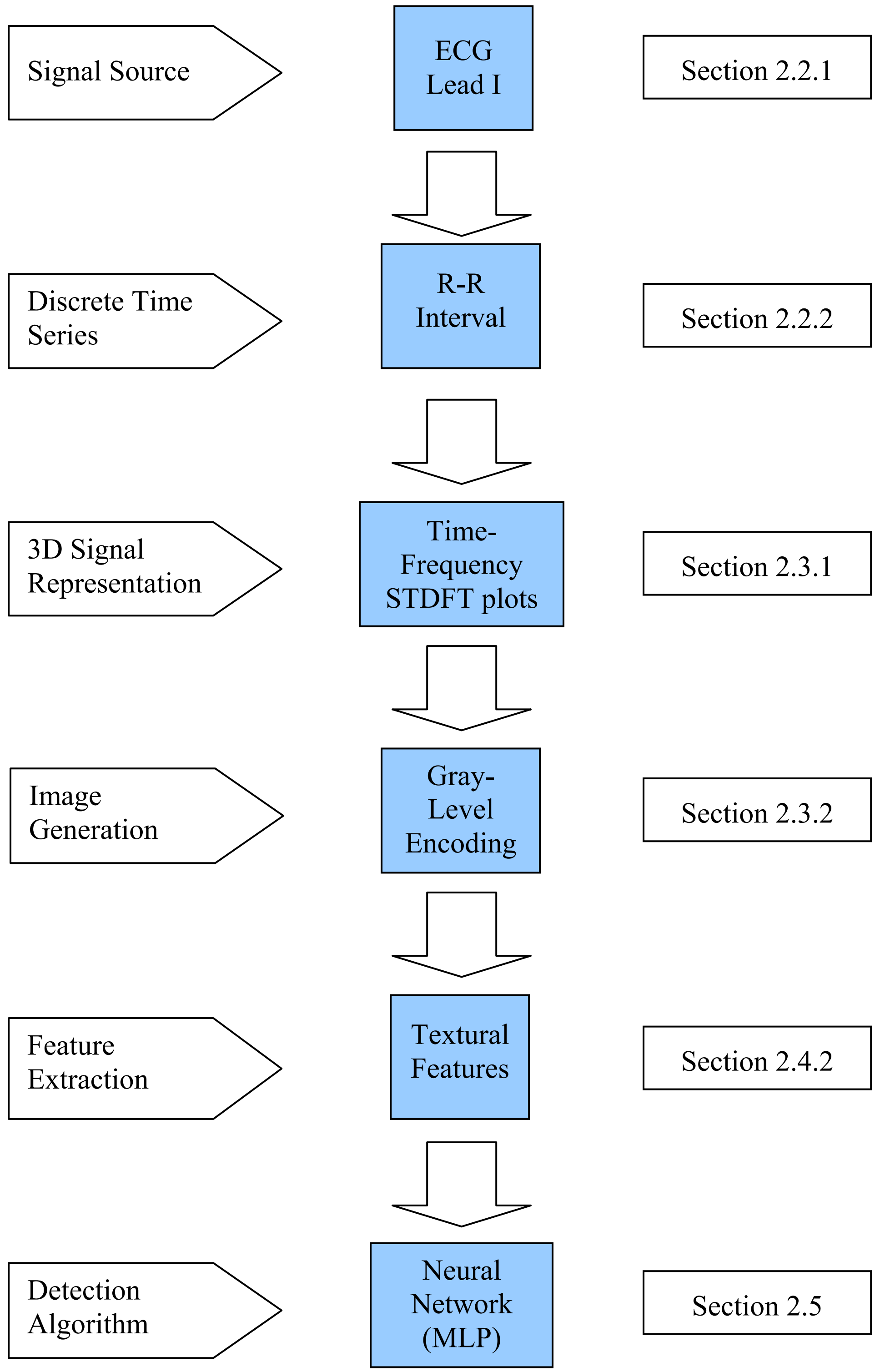

Figure 2.1 Block diagram of the proposed automatic detection algorithm. The section assignment of each stage is shown.

2.1 Subject Population and Data Gathering

This study is based on data collected and processed previously by the SDB-research group at UTA in collaboration with Sleep Consultants Inc., Fort Worth, TX.

2.1.1 Data Gathering and NPSG

The original experiment set-up called for an overnight recording of standard NPSG physiological parameters. A total of 18 channels were recorded; nine ECG channels, three EEG channels, EOG, chin EMG, chest and abdominal movements, nasal airflow, and percent oxygen saturation. All nine ECG channels (Leads I, II. III, and V1-V6), in addition to the nasal airflow, were acquired and sampled at a rate of 1024 samples/second. The remaining nine channels where acquired and sampled at different rates ranging from 25 to 100 samples/sec [8], [15].

2.1.2 Subject Population

As described by [8], [15], a volunteer population of sixteen normal (NOR) subjects was recruited for the purposes of the studies. The data from the subjects was used as a control group, where none of them had any known SDB history and none of them has had undergone any previous NPSG studies. Another group of fourteen previously diagnosed patients with OSAHS were selected for the study as well. This will be referred to as the (OSA) group.

The subject demographics of the NOR and OSA groups are shown in Table 2.1 and Table 2.2, respectively. The sleep expert scoring of their AHI is also included.

Table 2.1 NOR subject group demographics

Subject ID	Gender	Age (Years)	Weight (kg)	Height (m)	BMI (kg/m^2)	AHI
N01	M	43	87	1.85	25.4	3
N02	M	36	66	1.73	22.1	6
N03	F	58	64	1.6	25	0
N04	M	62	65	1.68	23	2
N05	M	49	95	1.75	31	4
N06	F	42	82	1.7	28.4	6
N07	F	40	61	1.6	23.8	2
N08	F	35	46	1.58	18.4	0
N09	M	38	68	1.65	25	6
N10	M	56	86	1.75	28.1	2
N11	F	54	57	1.6	22.3	3
N12	M	39	100	1.78	31.6	11
N13	F	36	81	1.68	28.7	2
N14	F	43	NA	NA	NA	20
N15	M	59	93	1.88	26.3	1
N16	F	42	78	1.65	28.7	14
Mean ± standard deviation		46.00 ± 9.38	73.08 ± 16.53	1.69 ± 0.09	25.34 ± 3.86	3.75 ± 3.11

Table 2.2 OSA subject group demographics

Subject ID	Gender	Age (Years)	Weight (kg)	Height (m)	BMI (kg/m^2)	AHI
N17	M	50	99	1.83	29.6	9
N18	M	38	91	1.88	25.7	4
N19	F	49	67	1.75	21.9	19
N20	M	39	157	1.90	43.5	70
N21	F	47	91	1.65	33.4	57
N22	M	37	64	1.63	24.1	8
N23	M	56	128	1.85	37.4	37
N24	F	44	89	1.70	30.8	20
N25	F	49	59	1.60	23.0	62
N26	M	49	100	1.80	30.9	14
N27	M	57	105	1.80	32.4	4
N28	F	54	92	1.52	39.8	30
N29	F	69	76	1.52	32.9	38
N30	M	66	95	1.75	33.2	65
Mean ± standard deviation		50.28 ± 9.60	93.79 ± 25.62	1.73 ± 0.13	31.33 ± 6.29	31.21 ± 23.89

Subjects N14 and N16 were excluded from the study for being recruited as NOR subjects, but turned out to have higher than normal AHI [15]. Subjects N13 and N15 were excluded from this study for having very noisy ECG LI recording, and the collected data was not suitable for further processing. Subjects N31 and N32 were excluded from the whole study for having amplifier saturated signal at all channels. The data extracted from them was deemed non-usable.

2.2 Clip Preparation and Selection

2.2.1 Clip Preparation

As described in Section 2.1.1, the data collected consisted of 18 channels. Previous time-domain and frequency-domain studies using this data [8], [15] showed that ECG Lead I has the most sensitivity of all ECG leads in detecting OSA events when R-Peak Envelope (RPE) is extracted, showing a sensitivity of 88.23% compared to a 70.00% for Lead II and 72.97% for Lead V6 [8]. Heart Rate Variability (HRV) showed limited sensitivity both in time and frequency domains. Time domain analysis reported in [8] of HRV showed a maximum of 78.57% sensitivity, whereas frequency domain tests, in the mentioned study, showed a sensitivity of 59.52%.

By contrast, this research combines both time and frequency domains to improve the detection sensitivity from HRV. Lead I will be used to extract the R-R interval, from which the HRV feature is found. The following is a summery of the steps used by [8] and [15] that explain the details of Lead I time-domain signal preprocessing.

2.2.1.1 Baseline Wander Removal

Caused by electrode movement and possibly the respiration, low-frequency drift in ECG is removed using a high-pass, linear phase, finite impulse response filter with cut-off frequency 0.8 Hz and length 200 [8].

2.2.1.2 R-Peak Detection

Finding the R-peaks in the ECG signal defines a discrete time series from which the HR can be calculated and used. In [8], [15], Hilbert-transform-based algorithm suggested by Benitez et al [16] to detect the R-peaks is used. The overall mean detection error for the acquired data (1.78 million beats) was 1% [8].

2.2.1.3 Manual Verification and Correction of the R-Peak Detection

Detection results using the automatic detection method mentioned above in Section 2.2.1.2 were manually verified. False detection was corrected. Clips with more than 15% detection error or more than 10% premature contraction were rejected.

2.2.2 R-R Interval Discrete Time Sequence

From the R-Peak detection, each R-peak can be associated with a single value that corresponds to the time interval between that peak and the peak before it. This clearly results in a discrete time series that is unevenly sampled. This time series is evenly resampled at 10 Hz with cubic spline interpolation, using MATLAB® function `spline`.

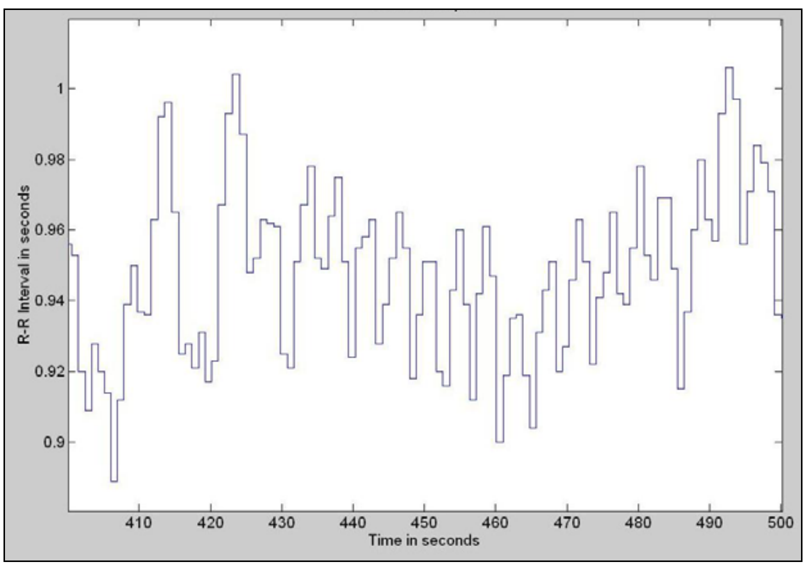

Figure 2.2 Typical R-R interval plot (a 100 seconds recording from NOR subject N10). HR in this clip varies from 59-66bpm.

2.2.3 Clip Selection

With significant spectral elements for studying and detecting OSA in HRV that are as low as 0.001Hz, a 900-second clip length has been recommended by [14] for frequency domain analysis of R-R interval time series [8]. During the study conducted by [15], all the available usable data was cut into 15-minute clips. As part of the study, these original data was visually scored by a certified sleep specialist, blind to the objective of the study.

For the purposes of this work, a clip that is extracted from a normal (NOR) subject and is free from any apnea or hypopnea episode was given a diagnostic value of zero (0). Any clip that is extracted from an apneic (OSA) subject and has an event of SDB was given a diagnostic value of one (1). Table 2.3 shows a break down of the number of clips used, per subject, for the purposes of this study.

Table 2.3 A break down of the number of the 900-second clips contributed by subject

NOR		OSA	
Subject ID	No. of Clips	Subject ID	No. of Clips
N01	7	N17	4
N02	6	N18	4
N03	8	N19	5
N04	12	N20	12
N05	11	N21	13
N06	4	N22	6
N07	7	N23	11
N08	11	N24	13
N09	9	N25	5
N10	5	N26	9
N11	6	N27	2
N12	6	N28	5
		N29	8
		N30	9
Total NOR	92	Total OSA	106

2.3 Time-Frequency Plots

2.3.1 Short-Time Discrete Fourier Transform (STDFT)

From Section 2.2.2, R-R interval time series was sampled at 10 Hz. Choosing 900-second clips from Section 2.2.3, the discrete time series develop into 9000-point clip for each of the 900-second clips.

STDFT is performed on these equally-spaced 9000-point R-R interval discrete time series. From [17], the STDFT is found using:

$$P(n,f) = \sum_{\tau=0}^{N-1} \{S(n+\tau) \cdot h_w(\tau)\} \cdot e^{-j2\pi\tau\frac{f}{F_s}}$$

where $P(n,f)$ is the complex Fourier Transform at discrete time n and discrete frequency f, $S(n)$ is the discrete R-R interval time signal, $h_w(n)$ is a Hanning window with of N = 300 point length, and F_s = 10Hz is the sampling rate.

In order to decrease the computational load, without effecting the resolution, the resulting matrix is sampled at every n = 4 points, and including the frequency range of 0-0.5 Hz at 0.004 Hz intervals, the STDFT complex matrix is reduced to 2250 discrete temporal columns by 125 discrete frequency rows.

2.3.2 Converting STDFT to Pictorial Images

198 125-rows × 2250-columns $P(n,f)$ STDFT complex-valued matrices are obtain. These matrices are handled as complex valued function of two variables i and j. This allows for conventional matrix processing methods to be applied, or for these matrices to be processed as pictorial images, after converting the complex values of the matrices to color or gray-level encoded values. This conversion is considered as one of

the major contributions of this investigation, where validated image recognition and classification schemes can be applied on of STDFT matrices extracted from temporal signal.

Performing a magnitude of the complex-valued matrices is equivalent to finding the signal power at each time-frequency point. When plotted with MATLAB® using the `mesh` function, the plot appears to have a smooth continuous surface. This function is used to visualize the resulting STDFT power distribution of the time-frequency plots. Different power distribution trends can be compared and contrasted between plots produced from NOR clips and OSA clips.

The following are four color-coded figures of chosen clips from the study that show the different trends in NOR and OSA clips. The coloring scheme is hot-cold; in which red signifies higher end of the power spectrum, and blue signifies the lower end. In Figures 2.3 and 2.4, there are visual distinctions between the NOR and OSA clips. However, in Figures 2.5 and 2.6, this distinction is not as clear. A reliable method of image detection and classification is needed.

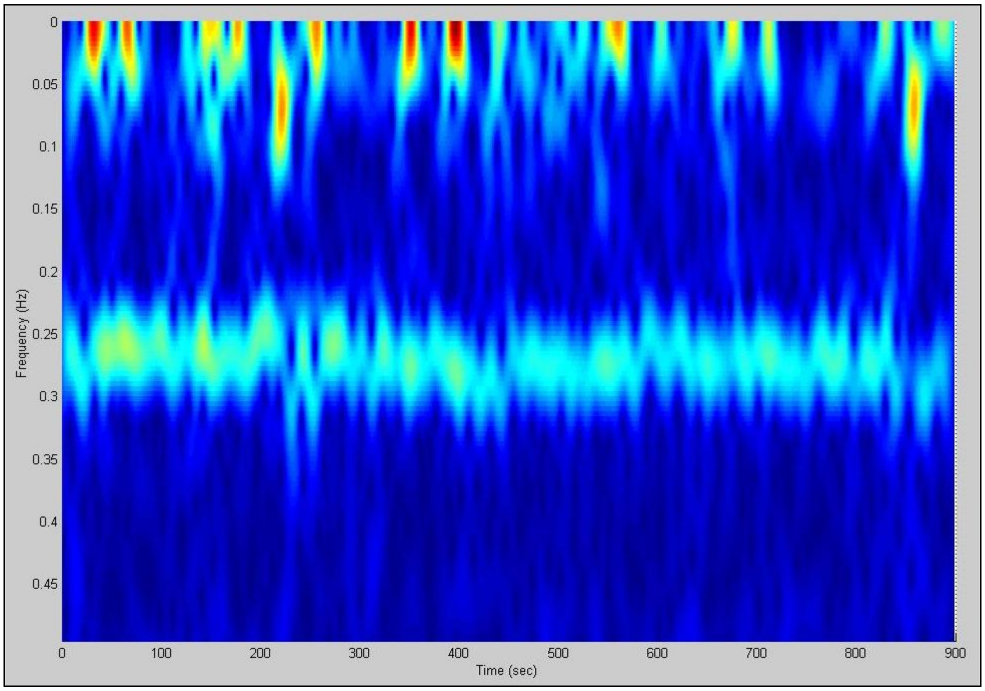

Figure 2.3 Color-code illustration of a 15-minute clip from a normal subject (N10). The coloring is done using MATLAB mesh command.

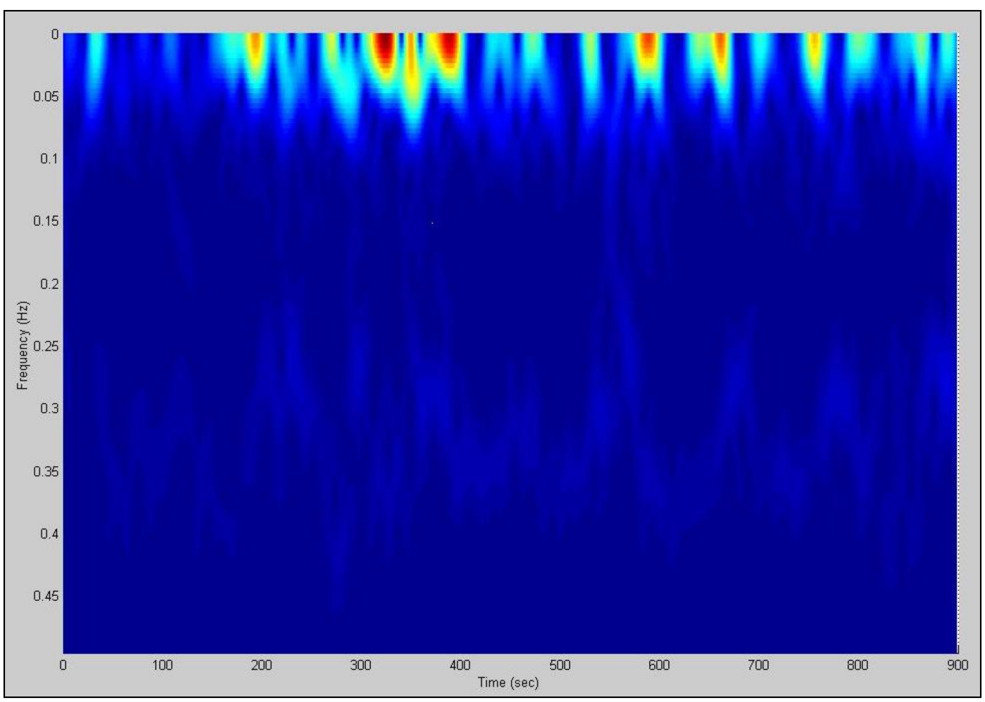

Figure 2.4 Color-code illustration of a 15-minute clip from an apneic subject (N20). The coloring is done using MATLAB mesh command.

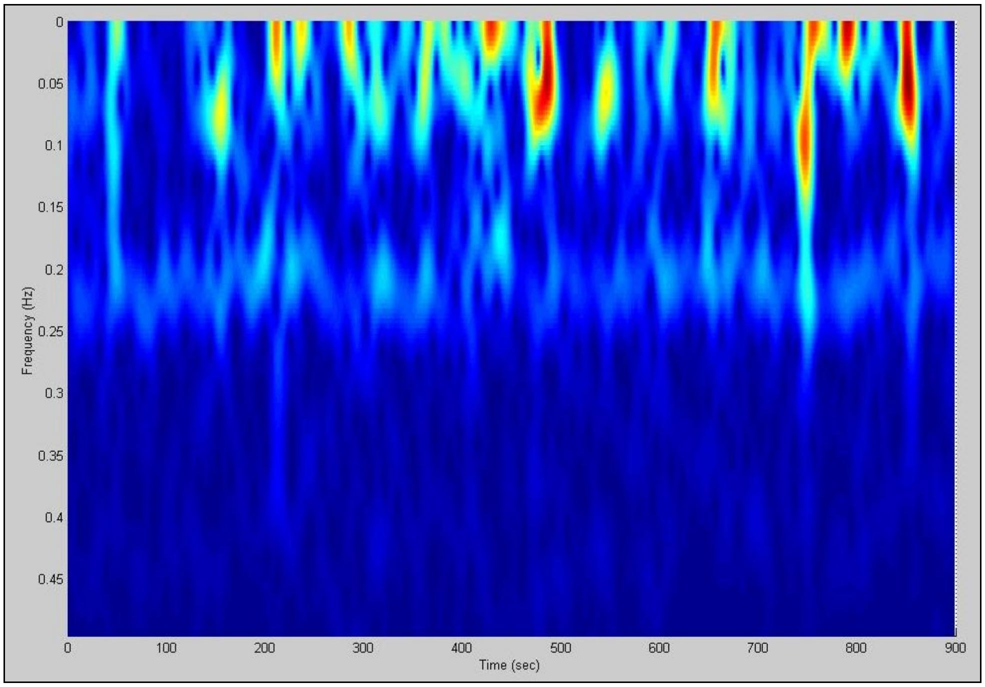

Figure 2.5 Color-code illustration of a 15-minute clip from a normal subject (N11). The coloring is done using MATLAB mesh command.

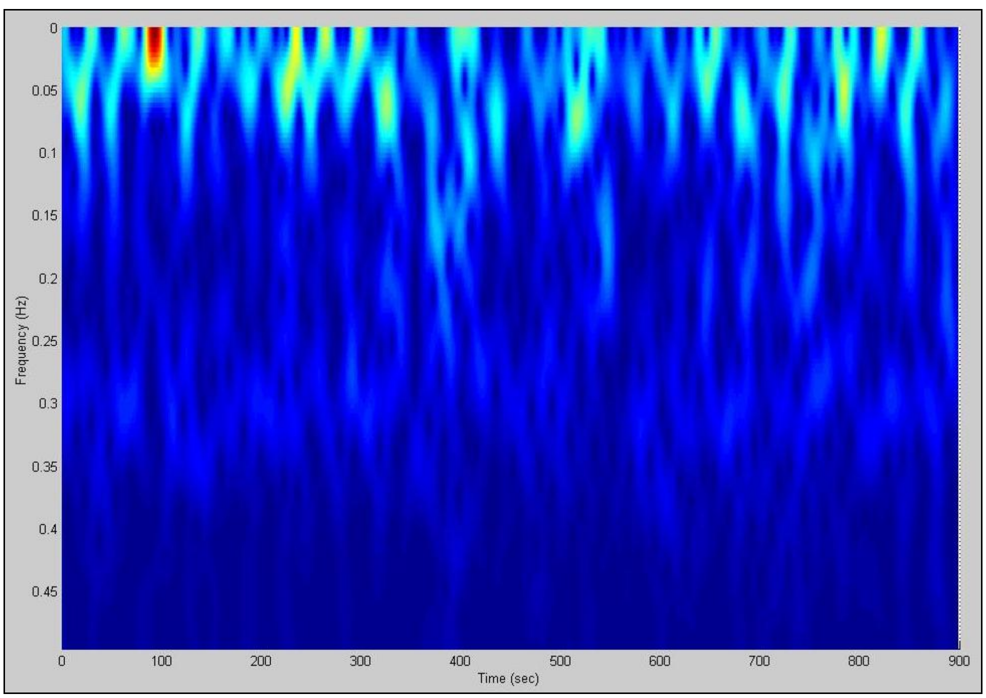

Figure 2.6 Color-code illustration of a 15-minute clip from an apneic subject (N20). The coloring is done using MATLAB mesh command.

However, for the purposes of this investigation, only gray-level encoding is considered. Choosing gray-level encoding simplifies the detection schemes. In a gray-level encoding, N_g levels are usually used, where N_g is equal to 2^N, N=1, 2, 3, etc. The value of each level corresponds to a shade of gray in which level 0 is black and level N_g-1 is white. The magnitude of each complex matrix $P(n,f)$ is used to produce real-valued matrix $I_{125\times 2250}$.

The following four sections describe four characteristically different encoding schemes that can be applied to generate four different images per each 900-sec clip. Each of the following encoding schemes will be trailed with example plots representing the encoding scheme applied on the clips shown in figures 2.3-6.

2.3.2.1 Magnitude encoding with 16 Gray levels ($N_g = 16$)

For encoding purposes, the global maximum and minimum magnitude values of the entire matrix $I_{125\times 2250}$ were found. The entire matrix was normalized with respect to $I(m,n)_{\max}$. Then all values were quantized to 16 equally-spaced bins, where each bin's length is equal to $\frac{I(m,n)_{\max} - I(m,n)_{\min}}{16}$. The quantized 16-gray-level matrix, $I_{g16}(m,n)$, was found using the following rule:

$$If\left(\left[i\times\frac{I(m,n)_{\max}-I(m,n)_{\min}}{16}\right]\le I(m,n)<\left[(i+1)\times\frac{I(m,n)_{\max}-I(m,n)_{\min}}{16}\right]\right)\Rightarrow I_{g16}(m,n)=i$$

for $0\le i\le 15$, $1\le m\le 125$, $1\le n\le 2250$. The following are plots of the same clips shown in figures 2.3-6 encoded in this scheme.

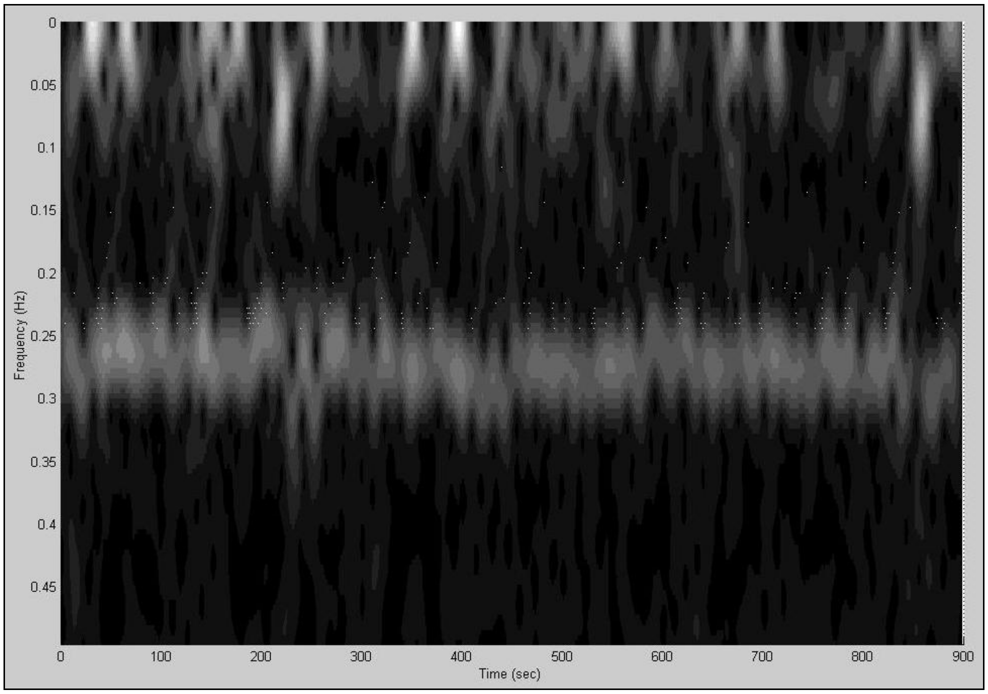

Figure 2.7 Gray-level encoding of a 15-minute clip from a normal subject (N10). The encoding is done using absolute value with Ng = 16.

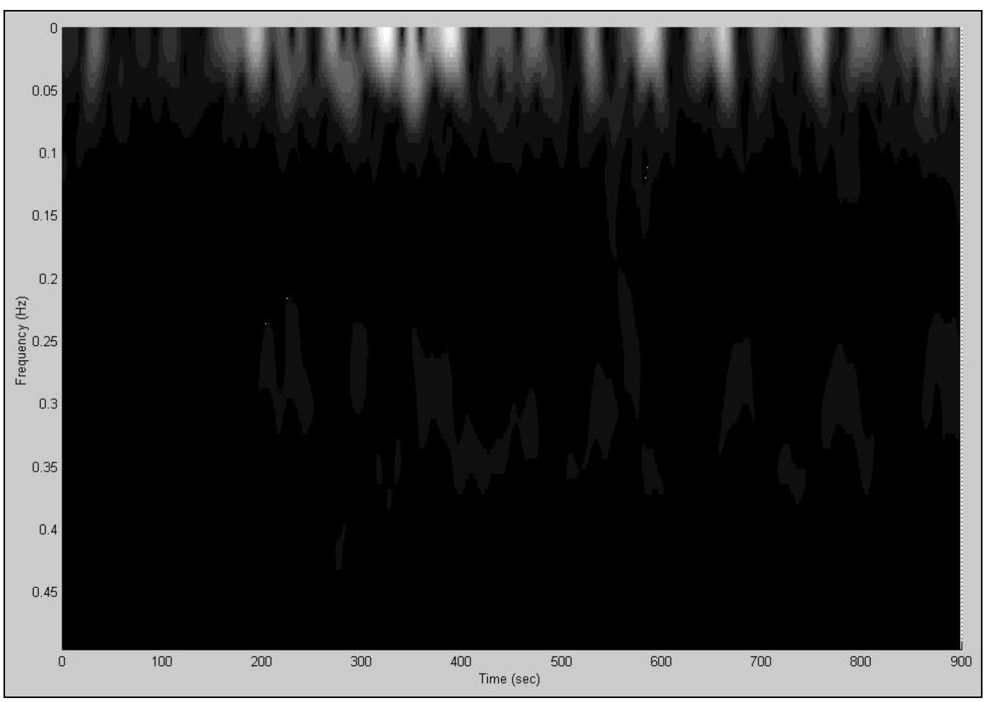

Figure 2.8 Gray-level encoding of a 15-minute clip from an apneic subject (N20). The encoding is done using absolute value with Ng = 16.

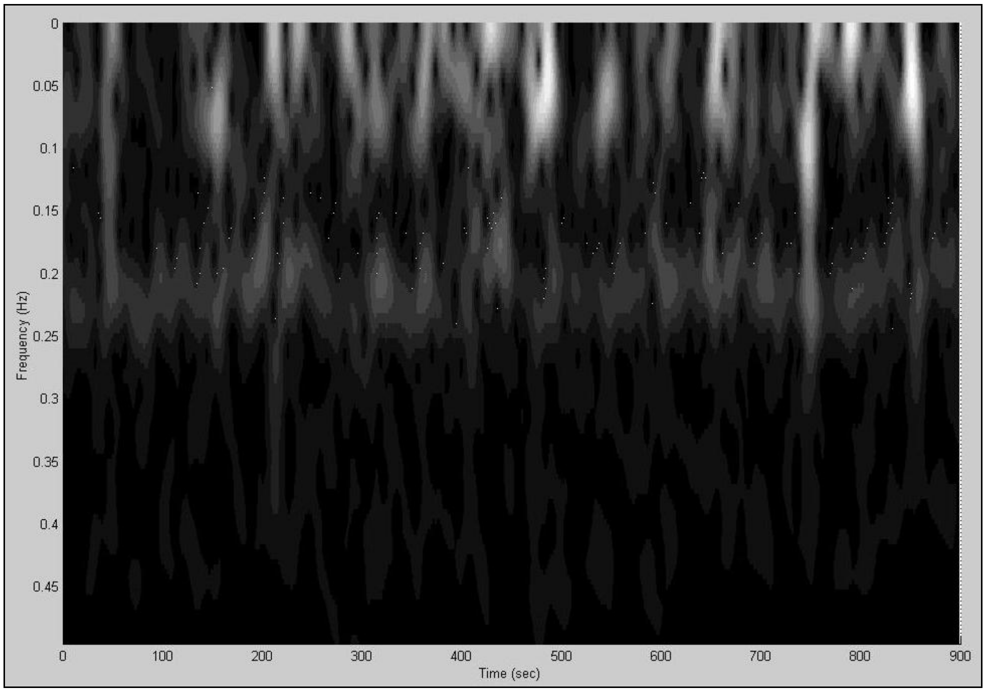

Figure 2.9 Gray-level encoding of a 15-minute clip from a normal subject (N11). The encoding is done using absolute value with Ng = 16.

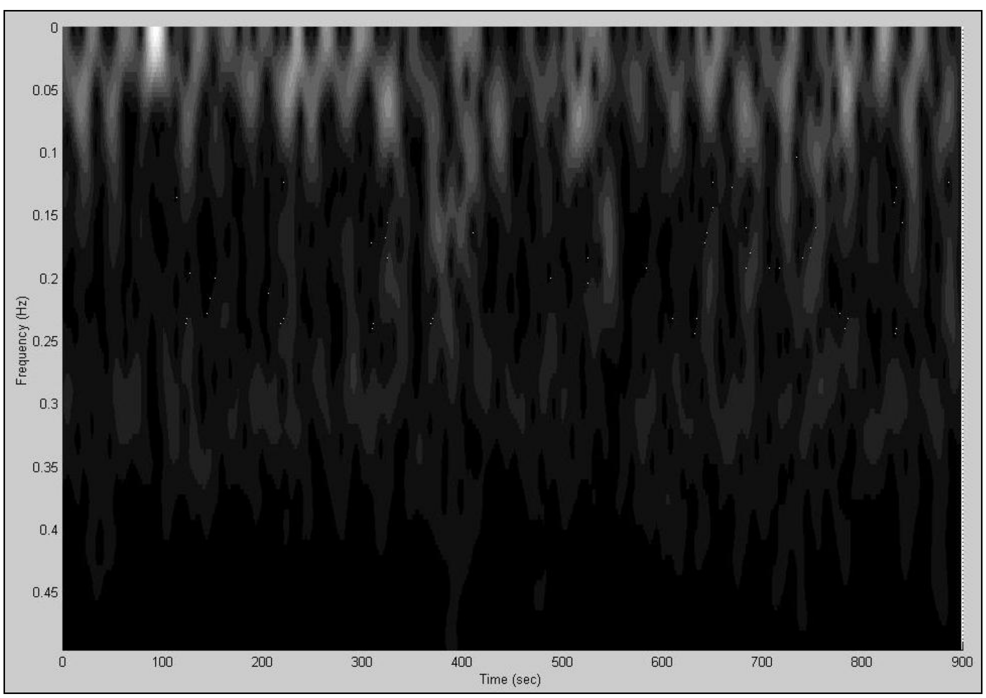

Figure 2.10 Gray-level encoding of a 15-minute clip from an apneic subject (N21). The encoding is done using absolute value with Ng = 16.

2.3.2.2 Magnitude encoding with 32 Gray levels ($N_g = 32$)

Higher resolution images, with smaller quantization bins, and lower quantization error, can be achieved by using 32 gray levels, rather than 16. As in Section 2.3.2.1, the global maximum and minimum values of the entire matrix $I_{125\times2250}$ were found. The entire matrix was normalized with respect to $I(m,n)_{\max}$. Then all values were quantized to 32 equally-spaced bins, where each bin's length is equal to $\frac{I(m,n)_{\max} - I(m,n)_{\min}}{32}$. The quantized 32-gray-level matrix, $I_{g32}(m,n)$, was found using the following rule:

$$If\left(\left[i \times \frac{I(m,n)_{\max} - I(m,n)_{\min}}{32}\right] \le I(m,n) < \left[(i+1) \times \frac{I(m,n)_{\max} - I(m,n)_{\min}}{32}\right]\right) \Rightarrow I_{g32}(m,n) = i$$

for $0 \le i \le 31$, $1 \le m \le 125$, $1 \le n \le 2250$. The following are plots of the same clips shown in figures 2.3-6 encoded in this scheme.

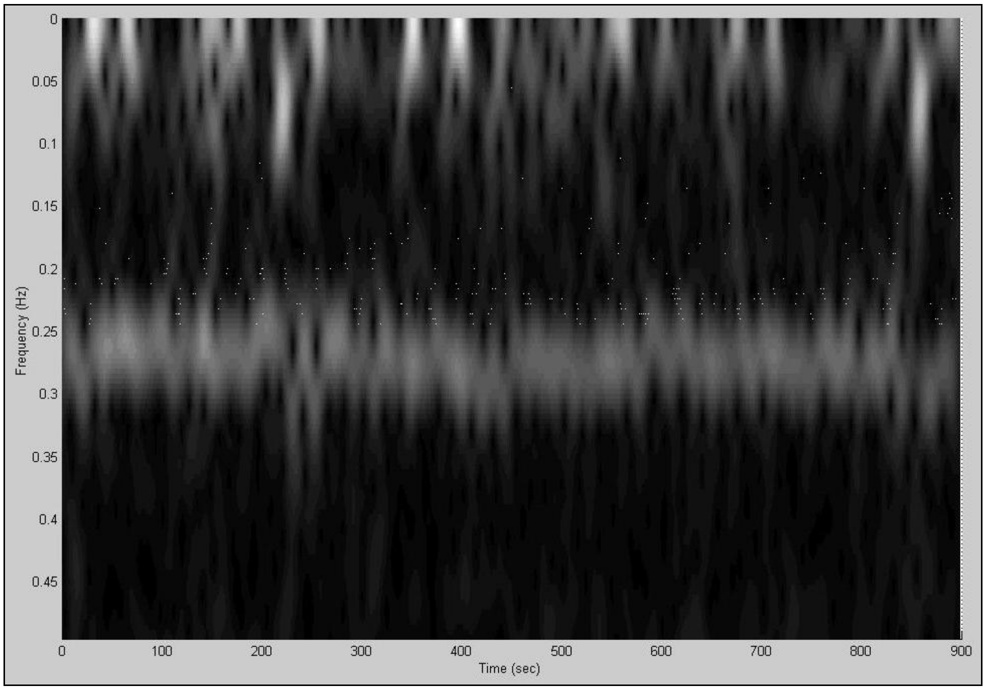

Figure 2.11 Gray-level encoding of a 15-minute clip from a normal subject (N10). The encoding is done using absolute value with Ng = 32.

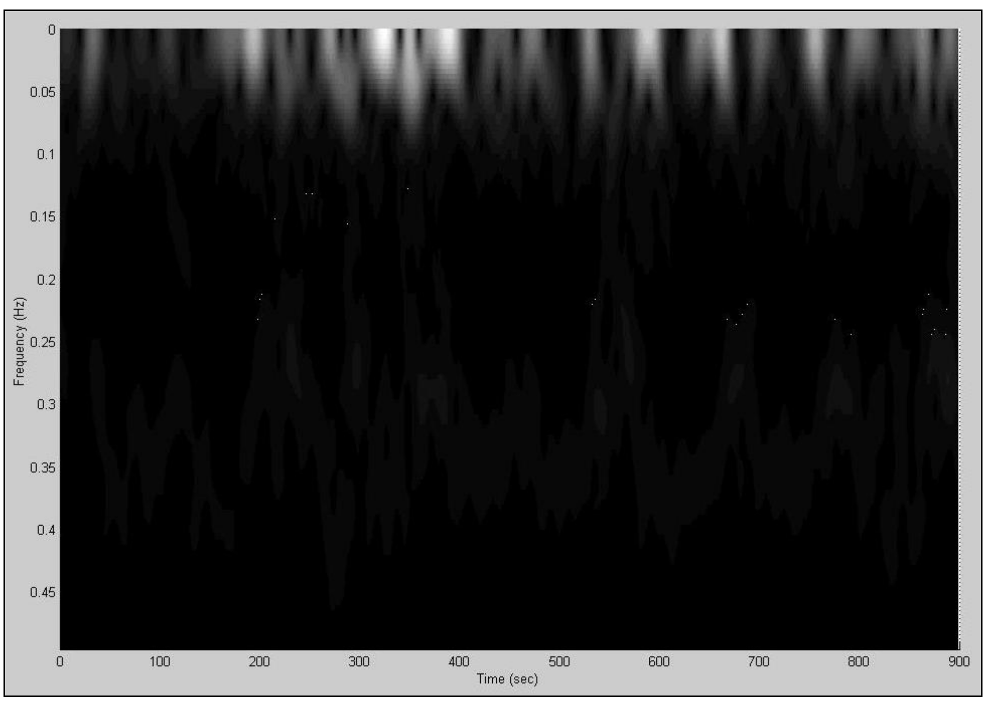

Figure 2.12 Gray-level encoding of a 15-minute clip from an apneic subject (N20). The encoding is done using absolute value with Ng = 32.

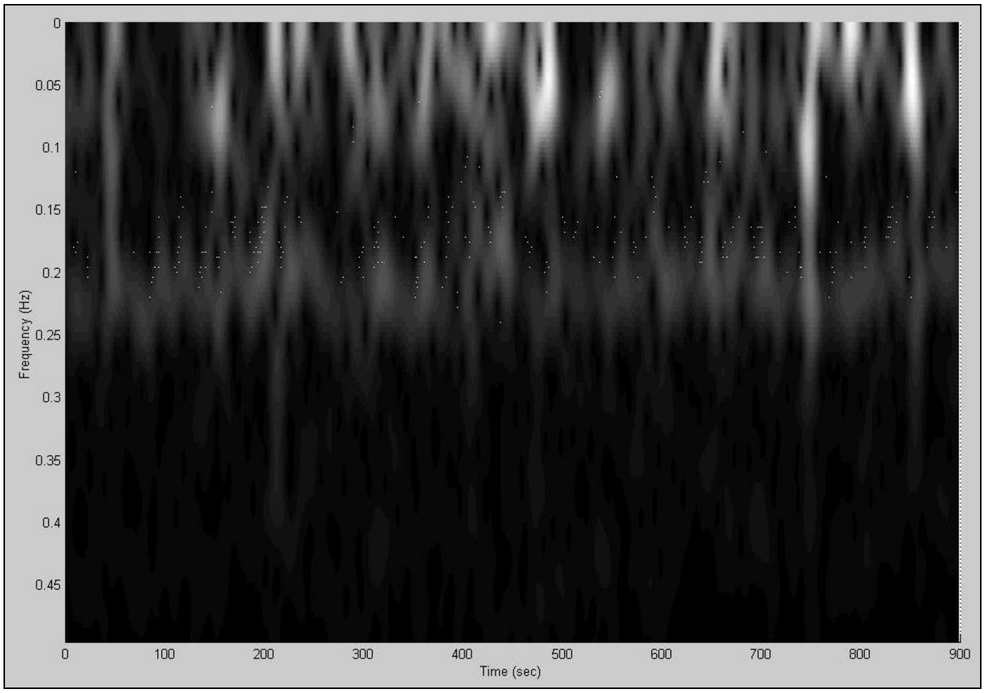

Figure 2.13 Gray-level encoding of a 15-minute clip from a normal subject (N11). The encoding is done using absolute value with Ng = 32.

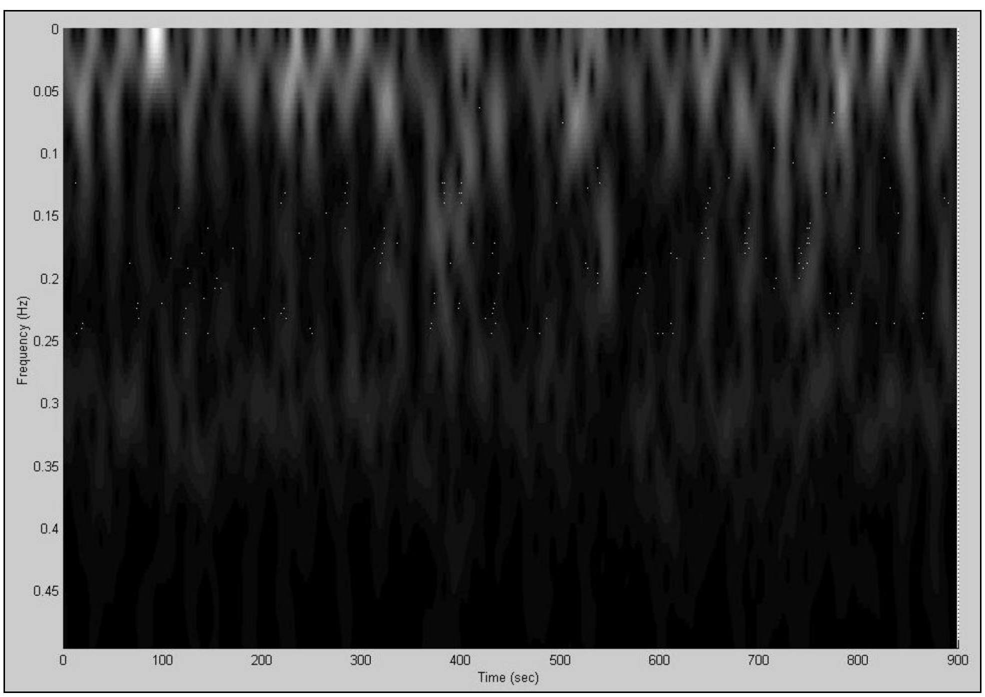

Figure 2.14 Gray-level encoding of a 15-minute clip from an apneic subject (N21). The encoding is done using absolute value with Ng = 32.

2.3.2.3 *Ln(*Magnitude*)* encoding with 16 Gray levels ($N_g = 16$)

The encoding process described in Sections 2.3.2.1-2 is very sensitive to the global minimum and maximum values of $I(m,n)$, and particularly sensitive to the maximum value. For images that showed a localized high peak in magnitude, all other local maxima are dwarfed and data is potentially lost during the process of quantization.

Taking the natural logarithm of the magnitude of $I(m,n)$ is a method that will suppress the global maximum, and give rise to the details of the local maxima. Then as in Section 2.3.2.1, the global maximum and minimum intensity values of the entire matrix $I_{125\times2250}$ were found. The entire matrix was normalized with respect to $I(m,n)_{\max}$. Then all values were quantized to 16 equally-spaced bins, where each bin's length is equal to $\frac{I(m,n)_{\max} - I(m,n)_{\min}}{16}$.

The quantized 16-gray-level matrix, $I_{gLn16}(m,n)$, was found using the following rule:

$$If\left(\left[i \times \frac{I(m,n)_{\max} - I(m,n)_{\min}}{16}\right] \le I(m,n) < \left[(i+1) \times \frac{I(m,n)_{\max} - I(m,n)_{\min}}{16}\right]\right) \Rightarrow I_{gLn16}(m,n) = i$$

for $0 \le i \le 15$, $1 \le m \le 125$, $1 \le n \le 2250$. The following are plots of the same clips shown in figures 2.3-6 encoded in this scheme.

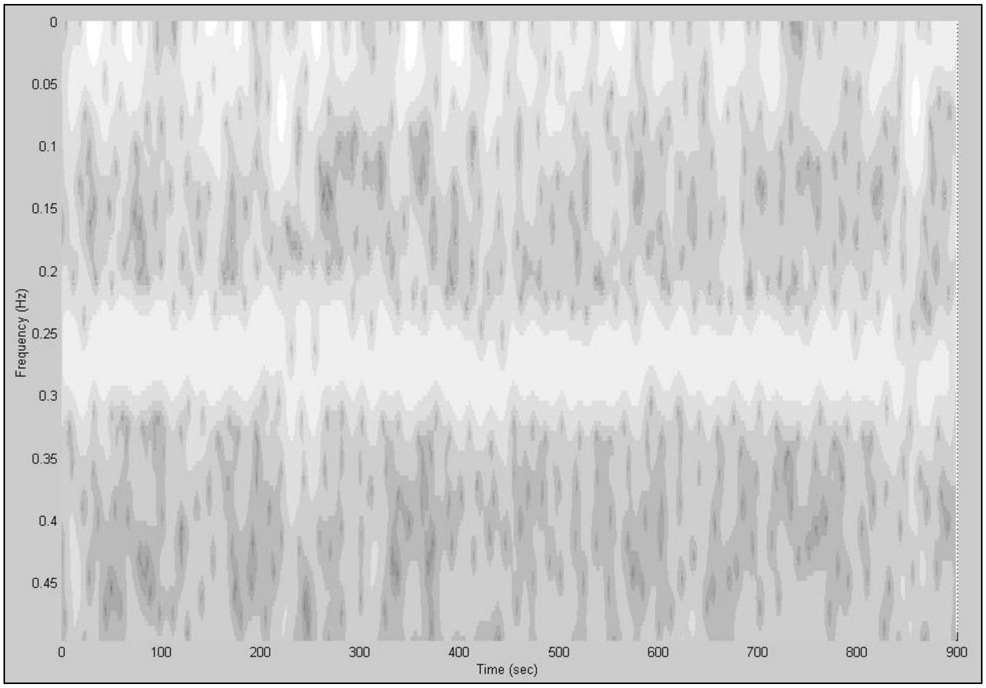

Figure 2.15 Gray-level encoding of a 15-minute clip from a normal subject (N10). The encoding is done using ln of absolute value with Ng = 16.

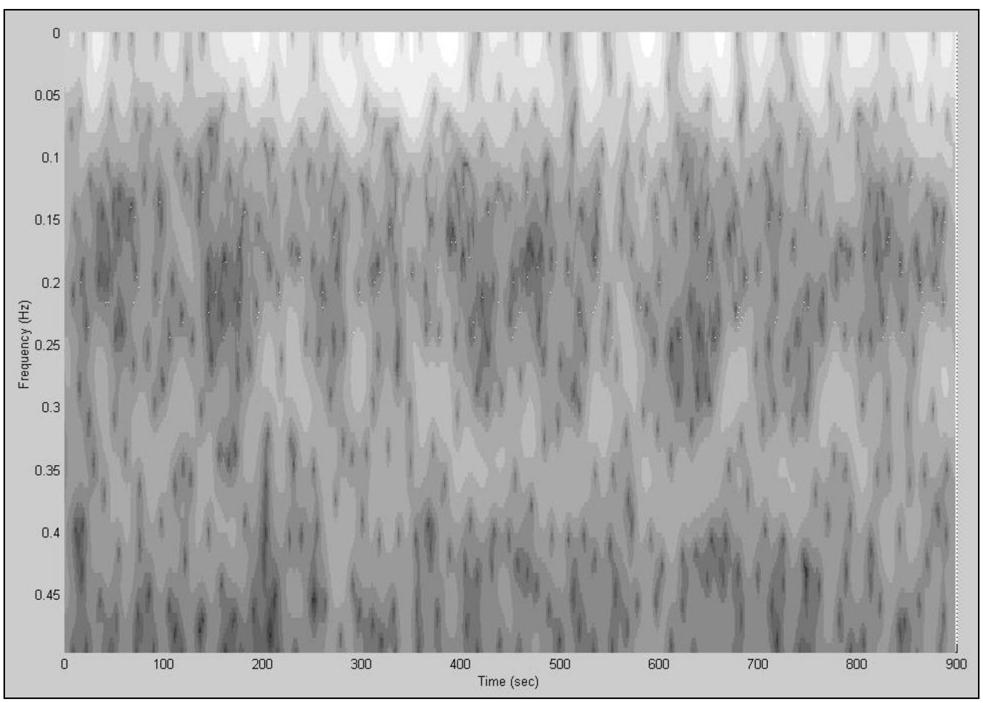

Figure 2.16 Gray-level encoding of a 15-minute clip from an apneic subject (N20). The encoding is done using ln of absolute value with Ng = 16.

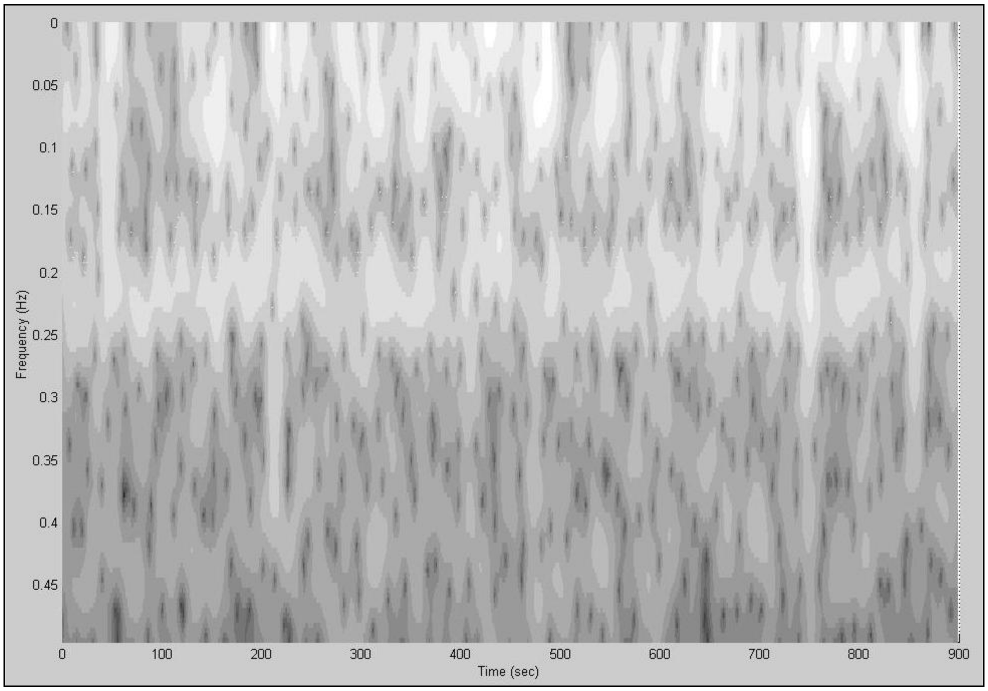

Figure 2.17 Gray-level encoding of a 15-minute clip from a normal subject (N11). The encoding is done using ln of absolute value with Ng = 16.

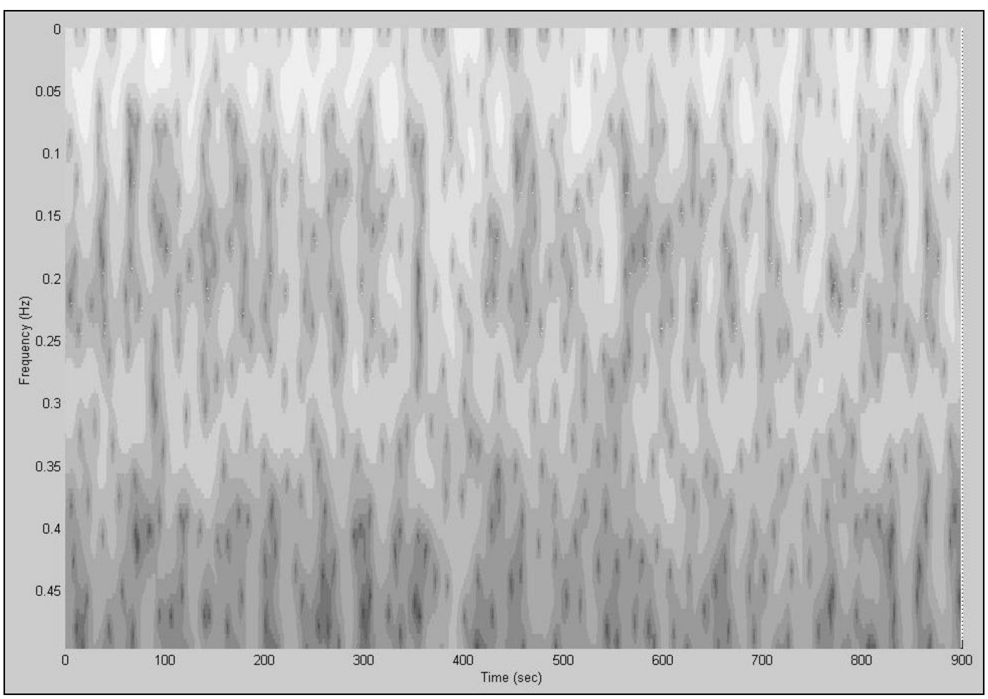

Figure 2.18 Gray-level encoding of a 15-minute clip from an apneic subject (N21). The encoding is done using ln of absolute value with Ng = 16.

2.3.2.4 *Histogram(*Magnitude*)* encoding with 16 Gray levels ($N_g = 16$)

Based on the fact that the previous three images have gray-level bins that were equally spaced, a fourth image was introduced. This image is based on quantizing the image into *un*equally spaced bins, where the length of the bin is inversely proportional to the histogramic distribution of the *I(m,n)* magnitude values. In other words, the higher the occurrence of a certain range of values in the matrix *I(m,n)* is, the smaller the bin size that will contain them.

Hence, values with similar characteristics with be grouped together, and the effect of extremely high or extremely low values will be nullified. This means that all N_g gray levels will have equal distribution in the image. Figure 2.19 illustrates this encoding scheme with $N_g = 4$. It can be noticed that the unequal bin length is due to the unequal intensity distribution, which is a function of the image itself.

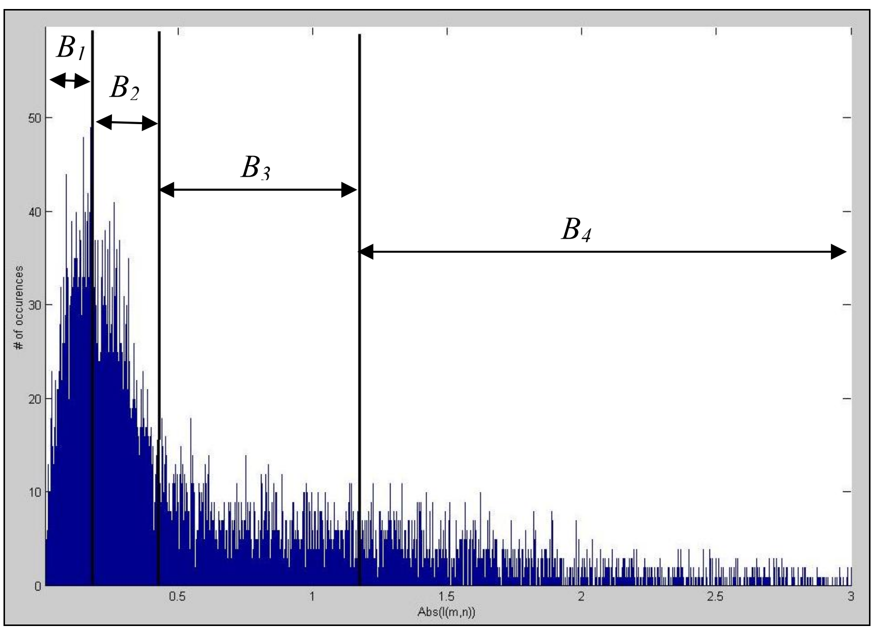

Figure 2.19 Histogram distribution of absolute values of clip in N10. Notice the unequal sized bins used to generate a 4-gray level image.

It is observed that the bin thresholds are chosen so that ratio of value occurrences within the bin to the overall value occurrences for all bins is fixed. This value is equal to $\frac{1}{N_g}$, and for N_g = 16, this value is 0.0625. No normalization is required for this encoding method.

The quantized 16-gray-level matrix, $I_{gHist16}(m,n)$, was found using the following rule: $If\left(I(m,n) \subset bin\#i\right) \Rightarrow I_{gHist16}(m,n) = i$, for $0 \le i \le 15$, $1 \le m \le 125$, $1 \le n \le 2250$.

For example, in Figure 2.19, an example intensity image histogram is shown. Choosing N_g = 4, the values falling within bin B_1 will have a quantized value of i = 0 (Black). The values falling within bin B_2 will have a quantized value of i = 1, which corresponds to a 66%-Gray shade. The values falling within bin B_3 will have a quantized value of i = 2 (33%-Gray), and finally the trailing tail of intensity values falling within bin B_4 will have a quantized value of i = 3 (White). The following are plots of the same clips shown in figures 2.3-6 encoded in this scheme.

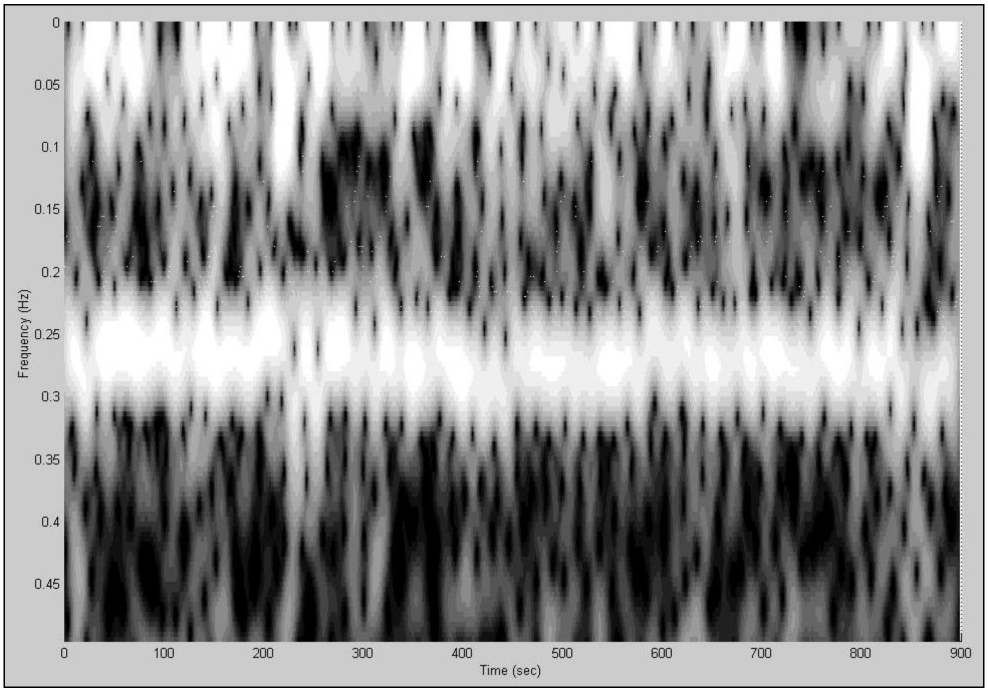

Figure 2.20 Gray-level encoding of a 15-minute clip from a normal subject (N10). The encoding is done using histogram of absolute value with Ng = 16.

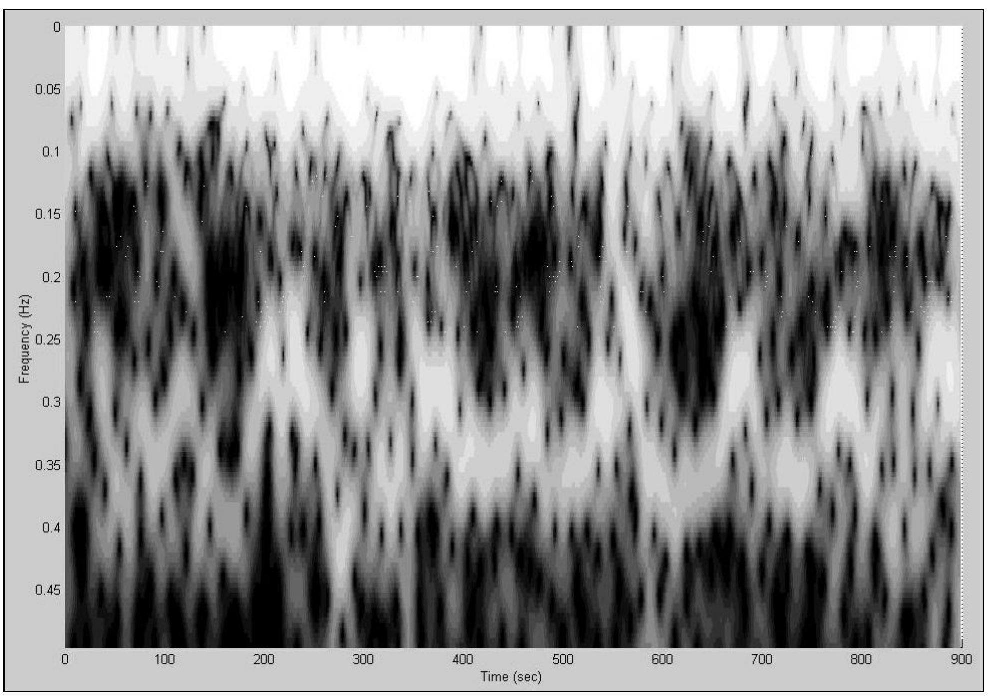

Figure 2.21 Gray-level encoding of a 15-minute clip from an apneic subject (N20). The encoding is done using histogram of absolute value with Ng = 16.

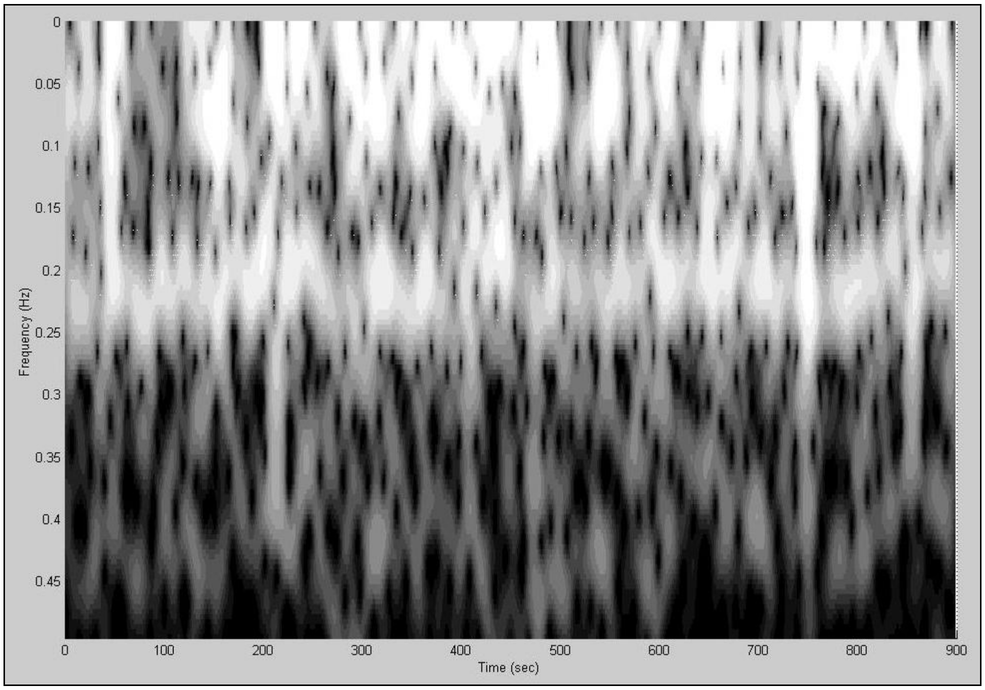

Figure 2.22 Gray-level encoding of a 15-minute clip from a normal subject (N11). The encoding is done using histogram of absolute value with Ng = 16.

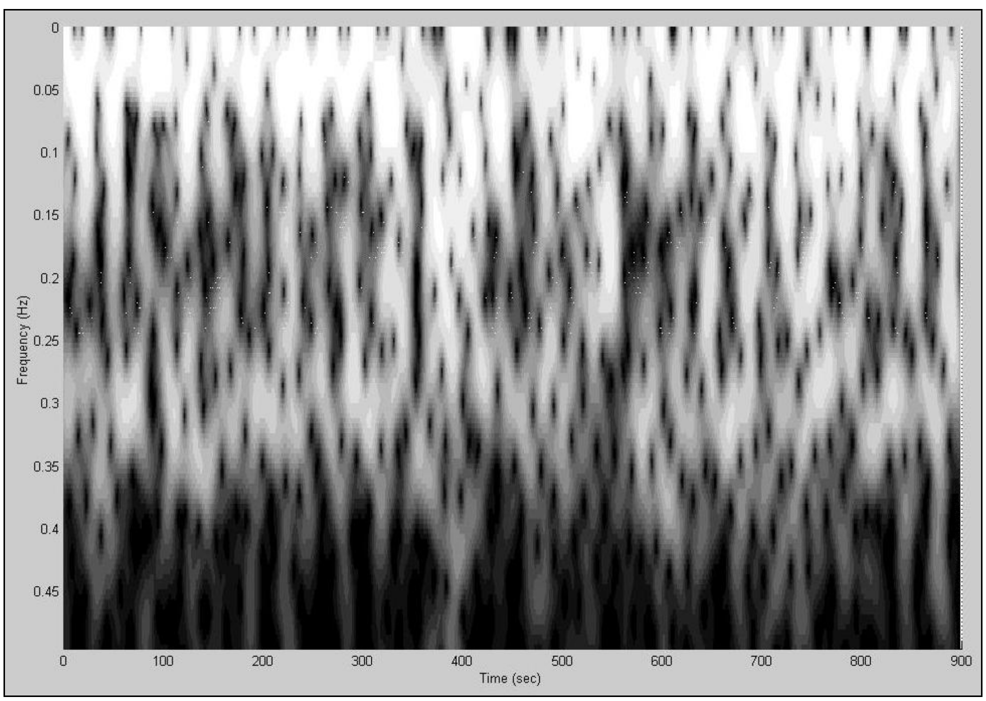

Figure 2.23 Gray-level encoding of a 15-minute clip from an apneic subject (N21). The encoding is done using histogram of absolute value with Ng = 16.

2.4 Feature Extraction

In previous sections, we have created four characteristically different gray-level images from each of the 900-second clips. Known and published image processing and classification techniques can be explored and used for the purposes of this research. Image classification based on their statistical properties is well-established and used in many industrial and medical applications. It is based on describing the image by a set of numerical features, such as Fourier, moment, Zernike, and Textural Features. Textural Features, first introduced by Haralick [18], are used in this investigation.

2.4.1 Co-occurrence Matrices

Described first by [18], Gray Level Co-occurrence Matrices (GLCM) are $N_g \times N_g$ symmetric matrices that contain the count of paired i and j gray levels separated by a certain distance, d, and angle, θ. So, a *single* image can produce numerous *different* GLCMs, depending on the choice of parameters, d and θ. The main aim of developing this method was the need for a faster, reliable and automatic method to sort the massive number of satellite images in the early 1970's [18].

For images with square pixels, there are four conventional values for θ that dictate the neighborhood relationship of any two paired pixels [18]. θ take on the values of 0 º, 45 º, 90 º, and 135 º. However, for the purposes of this research, only θ = 90 º values will be considered for the majority of GLCMs, for the reason explained in Section 2.3.2, and further discussed in Section 2.4.3. A θ = 0 º GLCM was used only once.

Calculating GLCMs has a high computational load, especially for images with large N_g and/or large dimensions. Different algorithms to efficiently calculate the GLCMs have been suggested [19]-[22].

In this work, a simple code of for-loop counters was used to calculate the GLCMs. The results of this code were manually verified and compared to the examples and results shown in the literature [18], [20], [23], and [24].

Other forms of co-occurrence matrices have been suggested, including Neighborhood Gray-tone Difference Matrices [25], [26], and Gray-Level Gradient-based Co-occurrence Matrices [27]. Also, the co-occurrence matrices definition has been generalized for multispecral images, where a single image is composed of different intensity, visible color, UV and/or IR sensors [28].

Normalized GLCMs (NGLCM) are GLCMs normalized to the total number of counted pairs, N_p. N_p is function of image size, the orientation θ, and the distance d. For an image with size $I_p \times I_p$ pixels, N_p is given as

$N_p = 2 \cdot (I_p \times (I_p - d))$ for $\theta = 0°$ and $90°$, and

$N_p = 2 \cdot ((I_p - d) \times (I_p - d))$ for $\theta = 45°$ and $135°$.

NGLCM's have been shown to consistently perform better than the un-normalized versions [29]. NGLCMs have been used for the purposes of this investigation.

The following is an example of calculating a GLCM and NGLCM from a simple 5-pixel by 5-pixel image, with only two gray levels ($N_g = 2$); black and white, for all four standard orientations ($\theta = 0°, 45°, 90°$, and $135°$), and $d = 1$.

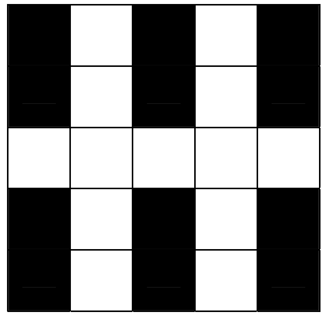

Figure 2.24 Example 2-tone gray level image. The image is 5-by-5 pixels. Its corresponding GLCMs and NGLCMs are found in Table 2.4.

Table 2.4 GLCMs and NGLCMs found from the 5-pixel by 5-pixel example image in Figure 2.24

	GLCM	NGLCM
$\theta = 0\,^{\text{o}}$, $N_p = 40$	$\begin{array}{c\|cc} i\downarrow, j\rightarrow & 0 & 1 \\ \hline 0 & 0 & 16 \\ 1 & 16 & 8 \end{array}$	$\begin{array}{c\|cc} i\downarrow, j\rightarrow & 0 & 1 \\ \hline 0 & 0 & 0.4 \\ 1 & 0.4 & 0.2 \end{array}$
$\theta = 45\,^{\text{o}}$, $N_p = 32$	$\begin{array}{c\|cc} i\downarrow, j\rightarrow & 0 & 1 \\ \hline 0 & 0 & 12 \\ 1 & 12 & 8 \end{array}$	$\begin{array}{c\|cc} i\downarrow, j\rightarrow & 0 & 1 \\ \hline 0 & 0 & 0.375 \\ 1 & 0.375 & 0.25 \end{array}$
$\theta = 90\,^{\text{o}}$, $N_p = 40$	$\begin{array}{c\|cc} i\downarrow, j\rightarrow & 0 & 1 \\ \hline 0 & 12 & 6 \\ 1 & 6 & 16 \end{array}$	$\begin{array}{c\|cc} i\downarrow, j\rightarrow & 0 & 1 \\ \hline 0 & 0.30 & 0.15 \\ 1 & 0.15 & 0.40 \end{array}$
$\theta = 135\,^{\text{o}}$, $N_p = 32$	$\begin{array}{c\|cc} i\downarrow, j\rightarrow & 0 & 1 \\ \hline 0 & 0 & 12 \\ 1 & 12 & 8 \end{array}$	$\begin{array}{c\|cc} i\downarrow, j\rightarrow & 0 & 1 \\ \hline 0 & 0 & 0.375 \\ 1 & 0.375 & 0.25 \end{array}$

2.4.2 Textural Features

In [18] definitions of statistical measures calculated from GLCM and later expanded to NGLCM are presented. These statistical measures are used as numerical descriptors of the textural features of the image from which the GLCM. More measures have been proposed in the literature [19], [30], and [31].

For all the subsequent measures, the following definitions are applicable:

M(i,j) = The Gray Level Co-occurrence Matrix *M* element located on the i^{th} row and j^{th} column. Since *M* is symmetric, *M(i,j)* = *M(j,i)* since.

N_g = Number of gray levels used in the image.

Matrix Mean, $\mu = \sum_{i}^{N_g} i \sum_{j}^{N_g} M(i,j)$

Matrix Variance, $\sigma^2 = \sum_{i}^{N_g} (i-\mu)^2 \sum_{j}^{N_g} M(i,j)$.

Nine textural features are chosen for the purpose of this research, and are defined herein.

Entropy (ENT) [18], [19], [31]: $ENT = -\sum_{i}^{N_g} \sum_{j}^{N_g} M(i,j) \cdot \log(M(i,j))$

Angular Second Moment (ASM) [18], [19], [31]: $ASM = \sum_{i}^{N_g} \sum_{j}^{N_g} [M(i,j)]^2$

Contrast (CON) [19], [31]: $CON = \sum_{i}^{N_g} \sum_{j}^{N_g} M(i,j) \cdot (i-j)^2$

Correlation (COR) [19], [30], [31]: $COR = \sum_{i}^{N_g} \sum_{j}^{N_g} \frac{M(i,j) \cdot (i-\mu) \cdot (j-\mu)}{\sigma^2}$

Dissimilarity (DIS) [19], [31]: $DIS = \sum_{i}^{N_g} \sum_{j}^{N_g} M(i,j) \cdot |i-j|$

Inverse Difference (IND) [19]: $IND = \sum_{i}^{N_g} \sum_{j}^{N_g} \frac{M(i,j)}{1+(i-j)}$

Inverse Difference Moment (IDM) [18], [19], [30]: $IDM = \sum_{i}^{N_g} \sum_{j}^{N_g} \frac{M(i,j)}{1+(i-j)^2}$

Variance (VAR) [18], [30]: $VAR = \sum_{i}^{N_g} \sum_{j}^{N_g} M(i,j) \cdot (i-\mu)^2$

Inverse Recursivity (INR) [30]:

$$INR = -\sum_{i}^{N_g} \sum_{j=i+1}^{N_g} 2 \cdot M(i,j) \cdot \log(2 \cdot M(i,j)) - \sum_{i}^{N_g} M(i,i) \cdot \log(M(i,i))$$

The following are the calculations of these features from the NGLCMs calculated in Table 2.4.

Table 2.5 Textural features calculated from NGLCMs found in Table 2.4

	NGLCM-0 º	NGLCM-45 º	NGLCM-90 º	NGLCM-135 º
ENT	+1.05462	+1.08190	+1.29644	+1.08190
ASM	+0.36000	+0.34375	+0.29500	+0.34375
CON	+0.80000	+0.75000	+0.30000	+0.75000
COR	-0.66667	-0.60000	+0.39394	-0.60000
DIS	+0.80000	+0.75000	+0.30000	+0.75000
IND	+0.60000	+0.62500	+0.85000	+0.62500
IDM	+0.84000	+0.85000	+0.94000	+0.85000
VAR	+0.24000	+0.23438	+0.24750	+0.23438
INR	+0.50020	+0.56214	+1.08860	+0.56214

2.4.3 Image Preparation

From Section 2.3.2, each of the 198 15-minute clips has been converted to a 125×2250 point complex-valued matrix using STDFT. Each of these matrices in turn produced four characteristically different gray-level images, as detailed in Sections 2.3.2.1-4.

For the purposes of having square images during for processing, each one of the images is cut into 18 segments, each segment of size 125×125 (18×125=2250). So each segment represents 50 seconds in length. The last segment, segment-18, is disregarded.

All seventeen segments coming from one image clip will be processed separately, and their results will be averaged to give a mean value for that one image clip.

Studying the characteristics of these images, we noticed in Section 2.3.2 the following:

1. All these images have vertical grooves or striations (at $\theta = 90^{\circ}$).

2. Most of the clip energy and fluctuations are evident in the lower half of the frequency range (0-0.25Hz).

It has been shown that different GLCMs can be derived from any given image, depending on the choice of distance, d, and angle, θ. Based on the two observations, orientation of GLCM calculation was mostly along $\theta = 90^{\circ}$. One image was calculated in the $\theta = 0^{\circ}$ direction during the research experimentation phase. Also, a couple of images show a size of 64×125, to account for the observation that most of the energy in the signal was seen to be in the lower half of the frequency range.

Applying feature extraction techniques as in Section 2.4.2, ten different GLCMs were extracted from each segment as follows:

From the image constructed in Section 2.3.2.1, four GLCM's were found:

a. Full image 125×125, $d = 5$, $\theta = 90^{\circ}$ gives GLCM-1

b. Full image 125×125, $d = 1$, $\theta = 90^{\circ}$ gives GLCM-2

c. Half image 64×125, $d = 5$, $\theta = 90^{\circ}$ gives GLCM-3

d. Full image 125×125, $d = 5$, $\theta = 0^{\circ}$ gives GLCM-4

From the image constructed in Section 2.3.2.2, three GLCM's were found:

a. Full image 125×125, $d = 5$, $\theta = 90^{\circ}$ gives GLCM-5

b. Full image 125×125, $d = 3$, $\theta = 90^{\circ}$ gives GLCM-6

c. Full image 125×125, $d = 1$, $\theta = 90^{\circ}$ gives GLCM-7

From the image constructed in Section 2.3.2.3, two GLCM's were found:

a. Full image 125×125, $d = 5$, $\theta = 90^{\circ}$ gives GLCM-8

b. Half image 64×125, $d = 5$, $\theta = 90^{\circ}$ gives GLCM-9

From the image constructed in Section 2.3.2.4, one GLCM was found:

Full image 125×125, $d = 5$, $\theta = 90^{\circ}$ gives GLCM-10

In the following figure 2.25, the process of slicing the clip into 17 usable square segments is illustrated. Each segment is processed to produce 4 different gray-scale images. Each image in turn is sorted to produce 10 different GLCMs as described above. For each GLCM, 9 features are extracted, and in turn these features are averaged across all 17 segments to give 9 features per treated image. Having 10 treated images per clip, this adds up to 90 features per clip.

It should be pointed out that all GLCMs are modified to their normalized form [29], NGLCMs, as described in Section 2.4.1.

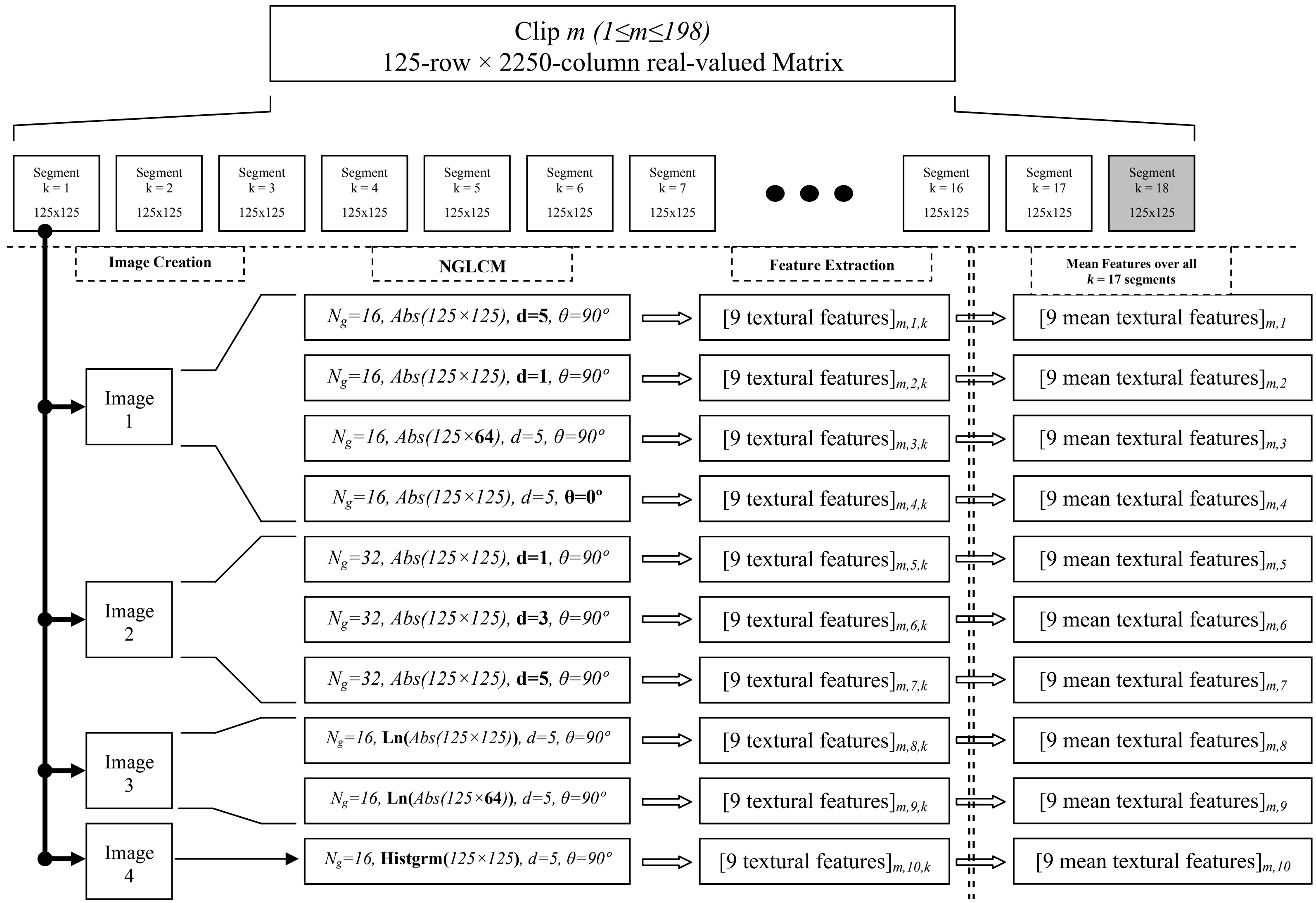

Figure 2.25 Block diagram of the overall clip encoding and feature extraction

2.4.4 Feature Organization

With each of the 198 diagnosed clips processed, ninety average features are found to describe each clip. These features were sorted to form a 198×92 matrix for data processing purposes; 198 rows representing each example clip (example vectors). Each example vector composed of 90 columns containing all 90 features, column 91 has the expert diagnosis of the clip (0 for NOR, 1 for OSA), and column 92 has the clip number, for clip tracking purposes.

2.4.5 Feature Selection

With 90 features making up every example vector, the issue of identifying the features with that will give an optimum detecting accuracy rises. Optimum feature selection is complex problem addressed extensively in the literature [33]-[36]. The goal is to find the subset of features that will result in the least classification error [36]. This is particularly important due to the fact that many of the features are redundant, noisy, or irrelevant to the classification problem at hand [35]. Also, the selection of a small subset of features from all candidate feature set can greatly reduce the computational load, especially in the case of using a neural network (NN) as a classifier.

The following is a description of the different feature selection schemes used in the classification stage of the research.

2.4.5.1 Complete Extracted Feature Set

Using all 90 features as inputs to the classifier was an available and legitimate option, even if it is computationally exhaustive, and most likely will not result in optimal results. This option can be considered as a control group to compare it with the subsequent feature selection methods and results.

2.4.5.2 Feature Selection Using Statistical Difference

At the first stages of this research, it was noticed that some features showed statistically significant difference between NOR clips and OSA clips within the general population of the clips. These features seemed to be trivial candidate to be used as inputs to the classifier. The features that showed best 1st degree statistical separation, i.e. features with largest mean gap between NOR and OSA and with minimum spread around each mean, were selected. These features were identified by minimum $\frac{\sigma}{\mu}$ value.

2.4.5.3 Feature Selection Using Piecewise Linear Network

Several feature selection methods have been developed in the quest of finding an optimal feature subset that has the least classification error. Inherently, this requires a user-define criterion to base the feature selection or rejection on [33]. Bottom-up or top-down methods are considered to be easy to construct, where a feature subset is growing or pruned, respectively. These methods are not optimal since they suffer from the nesting effect, where a feature cannot be discarded after being selected in the top-down method, or a feature cannot be selected if it was rejected early on in the bottom up method. The branch and bound method [32] has been shown to give the optimal selection of features subset. However, this method was shown impractical for data sets with more than 30 features [33]. Floating search algorithm [39] has been introduced to prevent the nesting effect, in which selection of the optimal feature subset is open to discarding chosen features based on conditional criteria.

Due to the inherent high non-linearity of the feature selection problem, Piecewise Linear Networks (PLN) has been shown to optimally perform in selecting the feature subset regardless of the size of feature set [33]. In [33], the authors have introduced feature selection algorithm that utilize piecewise linear orthonormal least square procedure. This algorithm has been used to find the optimal subset of the 90 available textural features in the available data. It has the advantages of giving optimal feature subset and being computationally efficient since it requires only one pass of the data. Data is divided in an appropriate number of clusters and auto- and cross-correlation matrices are calculated only once.

This algorithm is capable of finding combination of raw features that have high variance noise and discovering the useful ones. This is evident in choosing features beyond the 1st degree of statistical difference, and eliminating the noise.

2.5 Detection Method

2.5.1 Multilayer Perceptrons

A Multilayer Perceptron network (MLP) is a form of a feed-forward (FF) neural network. A feed-forward NN is one that does not have any of its frontward layer outputs feeding back into a previous layer's input. Its simple structure allows for relatively easy training, using conventional back-propagation (BP) tainting algorithms. It has been shown that MLP classifiers are very successful in image classification applications [32].

In this work, a three layer MLP was used; an input layer, hidden layer and an output layer. The input layer has a number of neurons (nodes) equal to the input vector length. The output layer consists of one neuron, accounting for a possibility of only 2 classes to be classified. The number of units in the hidden layer is adjustable, to achieve maximum classification accuracy.

Besides changing the number of hidden units, each layer of a MLP has two parameters that are trained to achieve maximum detection: node transfer function and weight vector. Both input and output layers use linear transfer functions (TF) for each neuron. This is achieved using MATLAB®'s linear TF `purelin`. The hidden layer, on the other hand, uses a hyperbolic tangent sigmoid function. MATLAB®'s implementation of this TF is $f(n) = \frac{2}{(1+e^{-2n})-1}$, and the built-in TF is `tansig`.

A MLP can be created using the `newff` function in the Neural Network toolbox in MATLAB®. Here, `ff` stands for "feed-forward". Three MLPs were created to classify the clips based on the three sets of input features described in Section 2.4.5.

2.5.2 Training the Multilayer Perceptrons

Training of a MLP is achieved by fitting the network parameters to the desired output using BP. By training the network using a training set, the network training accuracy is increases with the number of training epochs, N_{ep}. This is characteristic of MLPs, where it has been shown that they can approximate any polynomial [32]. However, the problem of over-fitting arises with increased number of training epochs, since the accuracy of a test set tends to have an optimum N_{ep}, after which it starts to decline [32].

In order to optimize the performance of the three MLPs we created, the selected feature set, used as input vectors to the MLP, are randomly assigned into two groups; training and testing, with a 2:1 ratio. With total of 198 clips, this is translated to 132 clips for training, and the remaining 66 for testing and validation.

The optimum number of hidden units and training epochs (N_h, N_{ep} pair) is found using a the following method:

For each possible N_h, the network is trained using the training set with N_{ep} epochs, and then the weights and biases parameters of the network are fixed in order to run the test set and calculate the accuracy of the network, since the network was blind to this set. Once the training and testing accuracies are stored for this N_{ep}, the MLP is then trained for the next N_{ep} increment, and so on. Once the maximum N_{ep} is reached, the training and testing for this N_h is repeated another 50 times, using different training and testing sets. This allows for studying the average performance of the MLP at that given N_h.

After experimenting with different back-propagation training algorithms, the one-step-secant back propagation method was used for training the MLP due to its fast convergence compared to other methods. It is implemented using the MATLAB® built-in `trainoss` algorithm.

2.5.3 Monte-Carlo Simulations and Network Performance

The goal in the network optimization is to find the optimum pair of hidden units and training epochs that will result in the maximum detection rate for the test set. Once these pairs are chosen for each of the three MLPs at hand, 1000-run Monte-Carlo simulations are performed to study the performance of the detector MLPs.

The performance of the network is measured using the following criteria:

$$\text{Sensitivity} = \frac{OSA_c}{Total\ OSA\ clips\ tested} \times 100\%$$

$$\text{Specificity} = \frac{NOR_c}{Total\ NOR\ clips\ tested} \times 100\%$$

$$\text{Accuracy} = \frac{OSA_c + NOR_c}{Total\ NOR\ \&\ OSA\ clips\ tested} \times 100\%$$

where OSA_c is the number of correctly detected OSA clips and NOR_c is the number of correctly detected NOR clips [8], [37].

Sensitivity refers to the probability that a diagnostic test is positive, given that the subject is apneic. Specificity refers to the probability that a diagnostic test is negative, given that the subject is normal. Accuracy refers to the probability that the diagnostic test is performed correctly [37].

CHAPTER 3

RESULTS AND ANALYSIS

The first section of this chapter summarizes the statistics of the study group. Section 3.2 reports the statistical results obtained from the textural features extracted from the clips' different images. Section 3.3 summarizes the results obtained from the different feature selections schemes described in the Chapter 2. Finally, Section 3.4 reports the training and testing performance of the detection algorithm suggested by this investigation.

3.1 Apnea-Hypopnea Index (AHI) Statistics

The AHI is the parameter used to distinguish between the NOR (N = 12) and OSA (N = 14) groups. Figure 3.1 shows the mean and Standard Error Mean (SEM) for the AHI of the two groups. SEM is calculated as $\frac{\sigma}{\sqrt{N}}$, where σ is the standard deviation of the group, and N is the number of subjects or examples in the group.

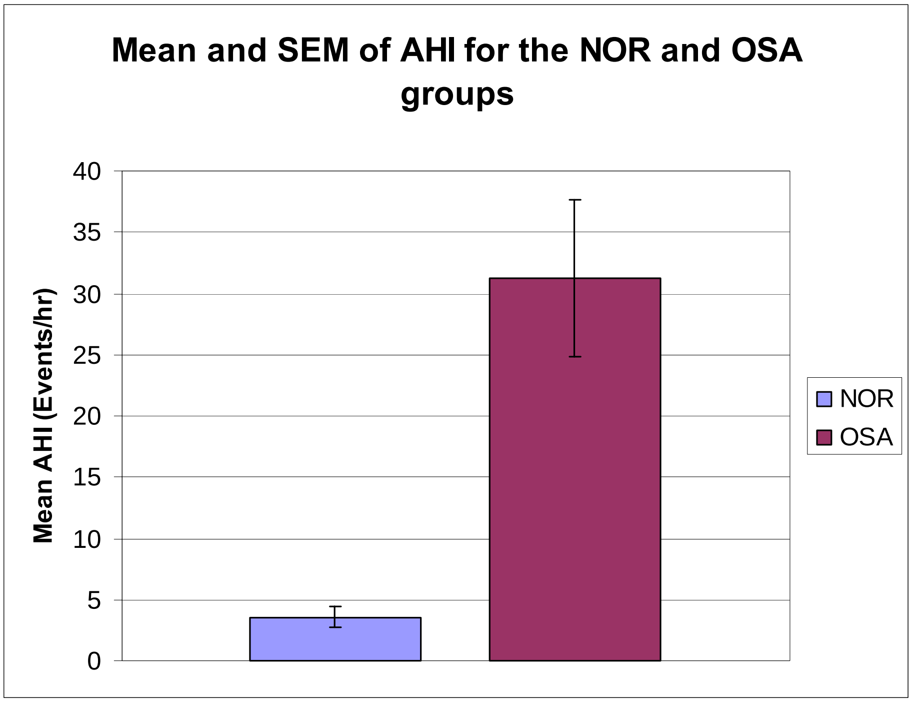

Figure 3.1 A comparison between the Mean of AHI for the NOR and OSA groups. The error bars represent the SEM. The two groups have p-value < 0.05.

3.2 Statistical Analysis of the Textural Features Extracted from the Images

In section 2.4.3, ten different GLCMs have been proposed to describe each of the 198 study clips. Each of these GLCMs was normalized to produce normalized GLCMs (NGLCMs). In order to study the difference between NOR and OSA clips, nine textural features were extracted from each of NGLCMs. These textural features are entropy (ENT), angular second moment (ASM), contrast (CON), correlation (COR), dissimilarity (DIS), inverse difference (IND), inverse difference moment (IDM), variance (VAR), and inverse recursivibility (INR). This collective group of ninety features per clip is the input to the proposed detection algorithm.

In the following ten subsections, a 1st degree statistical analysis of the nine features produced by each NGLCM is presented. The mean (μ), standard deviation (σ), and the percent standard deviation to the mean ratio ($\frac{\sigma}{\mu}$) are calculated. These values were initially used as discriminators to find the best feature subset to be used as inputs to the detection algorithm. Features with larger $|\mu_{NOR} - \mu_{OSA}|$ and least ($\frac{\sigma}{\mu}$) percentage were considered prime candidate for the 1st degree statistical difference feature subset.

3.2.1 Features Extracted from NGLCM-1

NGLCM-1 were normalized GLCM extracted from the gray-scale image scheme described in Section 2.3.2.1. This image has N_g=16, representing the absolute magnitude of the power spectrum of the time frequency plot of the clip. The full scale of the frequency was used (125 points). In calculating the GLCM, pairing of the pixels was at distance d=5 and orientation θ=90˚. Table 3.1 shows a tabulation of the statistics of the features extracted from NGLCM-1 for all clips extracted from the NOR and OSA groups. Figure 3.2 shows a graphical representation of these results.

Table 3.1 Mean, standard deviation, and mean to standard deviation ratio for all nine features for NGLCM-1 for both NOR and OSA groups

	NOR (N = 92)			OSA (N = 106)		
	Mean (μ)	St Dev (σ)	(σ/μ) %	Mean (μ)	St Dev (σ)	(σ/μ) %
ENT_1	2.38584	0.55879	23.42	1.82793	0.51032	27.92
ASM_1	0.52323	0.21250	40.61	0.32373	0.14266	44.07
CON_1	0.85136	0.49757	58.44	0.50537	0.24761	48.99
COR_1	0.80902	0.03738	4.62	0.82276	0.04468	5.43
DIS_1	0.52323	0.21250	40.61	0.33628	0.13404	39.86
IND_1	0.78190	0.07186	9.19	0.85367	0.05373	6.29
IDM_1	0.99675	0.00188	0.19	0.99807	0.00094	0.09
VAR_1	2.60765	1.31884	50.58	1.77622	0.87370	49.19
INR_1	2.10968	0.47982	22.74	1.63869	0.44409	27.10

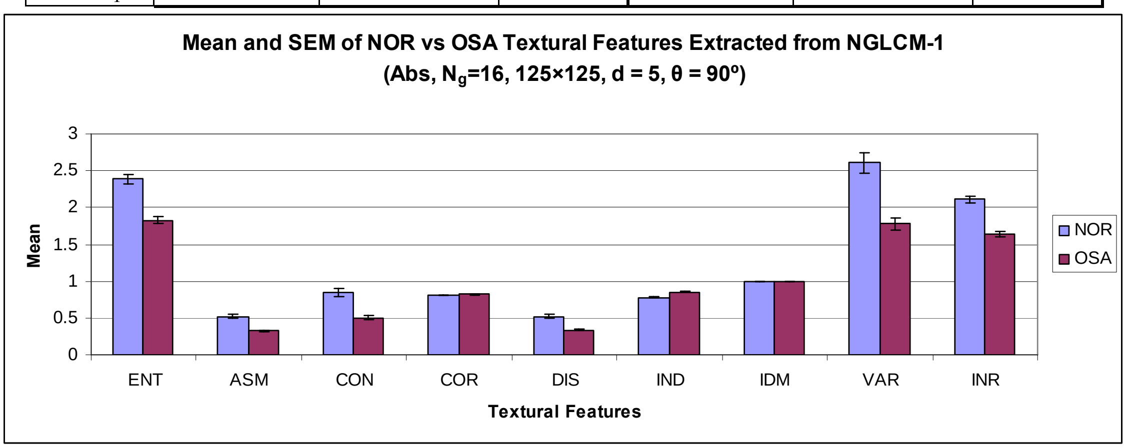

Figure 3.2 Comparison between the mean of the features extracted from NGLCM-1 between NOR and OSA groups. The error bars represent the SEM.

3.2.2 Features Extracted from NGLCM-2

NGLCM-2 were normalized GLCM extracted from the gray-scale image scheme described in Section 2.3.2.1. This image has N_g=16, representing the absolute magnitude of the power spectrum of the time frequency plot of the clip. The full scale of the frequency was used (125 points). In calculating the GLCM, pairing of the pixels was at distance d=1 and orientation θ=90˚. Table 3.2 shows a tabulation of the statistics of the features extracted from NGLCM-2 for all clips extracted from the NOR and OSA groups. Figure 3.3 shows a graphical representation of these results.

Table 3.2 Mean, standard deviation, and mean to standard deviation ratio for all nine features for NGLCM-2 for both NOR and OSA groups

	NOR (N = 92)			OSA (N = 106)		
	Mean (μ)	St Dev (σ)	(σ/μ) %	Mean (μ)	St Dev (σ)	(σ/μ) %
ENT_2	1.82328	0.41382	22.70	1.41891	0.37904	26.71
ASM_2	0.27502	0.10280	37.38	0.39803	0.13068	32.83
CON_2	0.11066	0.04549	41.10	0.07095	0.02896	40.82
COR_2	0.97185	0.00874	0.90	0.97213	0.01061	1.09
DIS_2	0.11040	0.04539	41.12	0.07076	0.02888	40.82
IND_2	0.94485	0.02268	2.40	0.96465	0.01443	1.50
IDM_2	0.99957	0.00018	0.02	0.99972	0.00011	0.01
VAR_2	2.74772	1.32471	48.21	1.95360	0.90989	46.58
INR_2	1.74765	0.38381	21.96	1.37064	0.35985	26.25

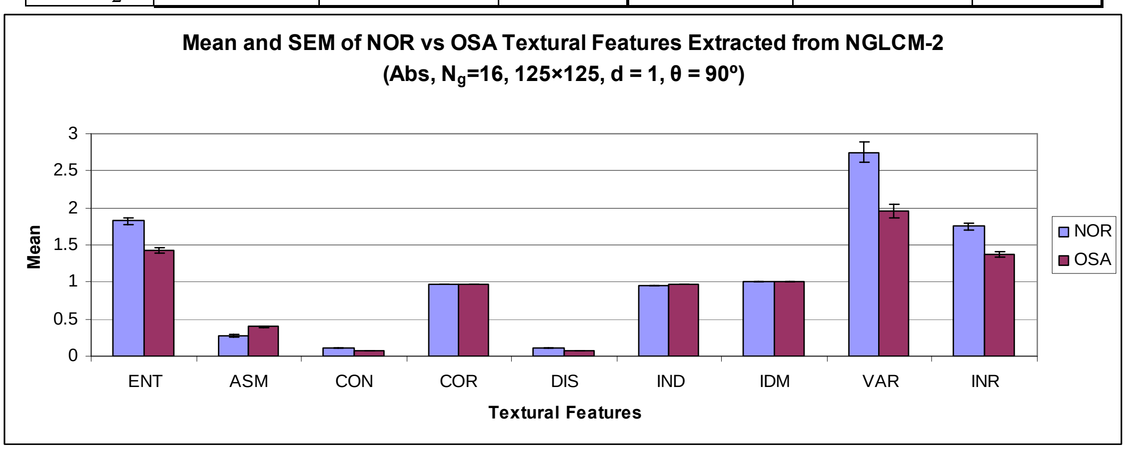

Figure 3.3 Comparison between the mean of the features extracted from NGLCM-2 between NOR and OSA groups. The error bars represent the SEM.

3.2.3 Features Extracted from NGLCM-3

NGLCM-3 were normalized GLCM extracted from the gray-scale image scheme described in Section 2.3.2.1. This image has N_g=16, representing the absolute magnitude of the power spectrum of the time frequency plot of the clip. Half the scale of the frequency was used (64 points). In calculating the GLCM, pairing of the pixels was at distance d=5 and orientation θ=90˚. Table 3.3 shows a tabulation of the statistics of the features extracted from NGLCM-3 for all clips extracted from the NOR and OSA groups. Figure 3.4 shows a graphical representation of these results.

Table 3.3 Mean, standard deviation, and mean to standard deviation ratio for all nine features for NGLCM-3 for both NOR and OSA groups

	NOR (N = 92)			OSA (N = 106)		
	Mean (μ)	St Dev (σ)	(σ/μ) %	Mean (μ)	St Dev (σ)	(σ/μ) %
ENT_3	2.36465	0.45421	19.21	2.02018	0.43726	21.64
ASM_3	0.17681	0.07608	43.03	0.25164	0.09900	39.34
CON_3	0.35623	0.13489	37.86	0.26736	0.11301	42.27
COR_3	0.92608	0.01989	2.15	0.92448	0.03126	3.38
DIS_3	0.30835	0.10159	32.95	0.23780	0.09115	38.33
IND_3	0.85332	0.04593	5.38	0.88573	0.04236	4.78
IDM_3	0.99862	0.00052	0.05	0.99896	0.00044	0.04
VAR_3	3.22542	1.44491	44.80	2.77618	1.20756	43.50
INR_3	2.16760	0.39570	18.26	1.86601	0.38404	20.58

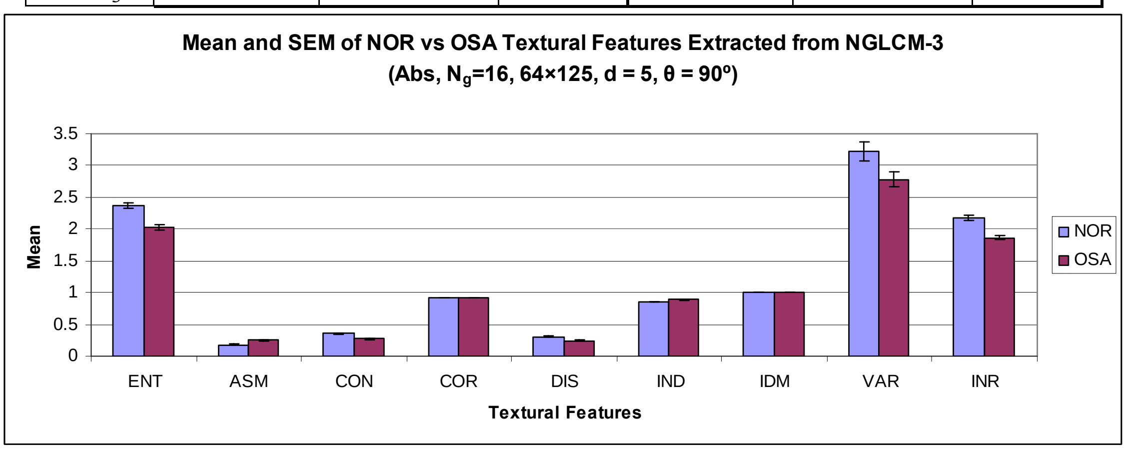

Figure 3.4 Comparison between the mean of the features extracted from NGLCM-3 between NOR and OSA groups. The error bars represent the SEM.

3.2.4 Features Extracted from NGLCM-4

NGLCM-4 were normalized GLCM extracted from the gray-scale image scheme described in Section 2.3.2.1. This image has N_g=16, representing the absolute magnitude of the power spectrum of the time frequency plot of the clip. The full scale of the frequency was used (125 points). In calculating the GLCM, pairing of the pixels was at distance d=5 and orientation θ=0˚. Table 3.4 shows a tabulation of the statistics of the features extracted from NGLCM-4 for all clips extracted from the NOR and OSA groups. Figure 3.5 shows a graphical representation of these results.

Table 3.4 Mean, standard deviation, and mean to standard deviation ratio for all nine features for NGLCM-4 for both NOR and OSA groups

	NOR (N = 92)			OSA (N = 106)		
	Mean (μ)	St Dev (σ)	(σ/μ) %	Mean (μ)	St Dev (σ)	(σ/μ) %
ENT_4	2.01444	0.45645	22.66	1.58810	0.43675	27.50
ASM_4	0.24840	0.10483	42.20	0.37128	0.13642	36.74
CON_4	0.23352	0.09304	39.85	0.16435	0.07620	46.36
COR_4	0.94740	0.01305	1.38	0.94143	0.02267	2.41
DIS_4	0.20657	0.07507	36.34	0.14905	0.06515	43.71
IND_4	0.90092	0.03500	3.89	0.92788	0.03094	3.33
IDM_4	0.99909	0.00036	0.04	0.99936	0.00030	0.03
VAR_4	2.78114	1.32986	47.82	1.99617	0.91865	46.02
INR_4	1.88117	0.41107	21.85	1.49084	0.39644	26.59

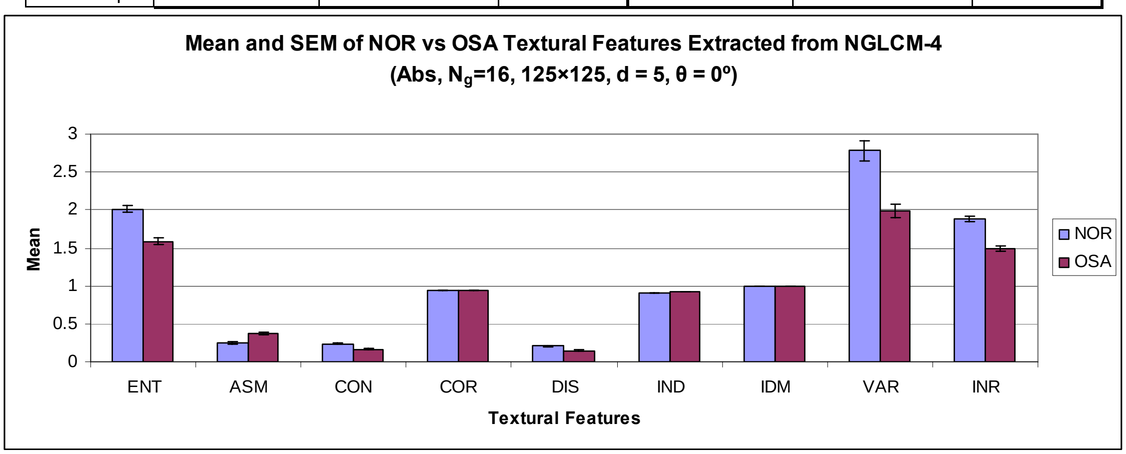

Figure 3.5 Comparison between the mean of the features extracted from NGLCM-4 between NOR and OSA groups. The error bars represent the SEM.

3.2.5 Features Extracted from NGLCM-5

NGLCM-5 were normalized GLCM extracted from the gray-scale image scheme described in Section 2.3.2.2. This image has N_g=32, representing the absolute magnitude of the power spectrum of the time frequency plot of the clip. The full scale of the frequency was used (125 points). In calculating the GLCM, pairing of the pixels was at distance d=5 and orientation θ=90˚. Table 3.5 shows a tabulation of the statistics of the features extracted from NGLCM-5 for all clips extracted from the NOR and OSA groups. Figure 3.6 shows a graphical representation of these results.

Table 3.5 Mean, standard deviation, and mean to standard deviation ratio for all nine features for NGLCM-5 for both NOR and OSA groups

	NOR (N = 92)			OSA (N = 106)		
	Mean (μ)	St Dev (σ)	(σ/μ) %	Mean (μ)	St Dev (σ)	(σ/μ) %
ENT_5	3.39839	0.65466	19.26	2.72565	0.61705	22.64
ASM_5	0.09054	0.05549	61.29	0.16475	0.08718	52.92
CON_5	3.14441	2.06239	65.59	1.77052	0.99209	56.03
COR_5	0.83279	0.03176	3.81	0.85376	0.02952	3.46
DIS_5	1.06320	0.42722	40.18	0.68797	0.26729	38.85
IND_5	0.66153	0.08570	12.96	0.75579	0.07345	9.72
IDM_5	0.99700	0.00195	0.20	0.99831	0.00094	0.09
VAR_5	10.56572	5.60065	53.01	7.08415	3.62526	51.17
INR_5	3.01224	0.58179	19.31	2.43304	0.54040	22.21

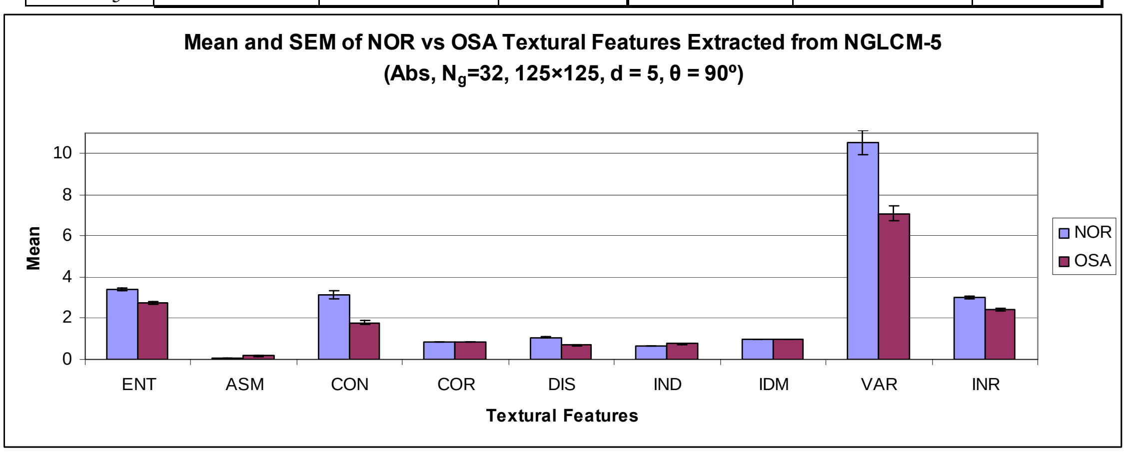

Figure 3.6 Comparison between the mean of the features extracted from NGLCM-5 between NOR and OSA groups. The error bars represent the SEM.

3.2.6 Features Extracted from NGLCM-6

NGLCM-6 were normalized GLCM extracted from the gray-scale image scheme described in Section 2.3.2.2. This image has N_g=32, representing the absolute magnitude of the power spectrum of the time frequency plot of the clip. The full scale of the frequency was used (125 points). In calculating the GLCM, pairing of the pixels was at distance d=3 and orientation θ=90˚. Table 3.6 shows a tabulation of the statistics of the features extracted from NGLCM-6 for all clips extracted from the NOR and OSA groups. Figure 3.7 shows a graphical representation of these results.

Table 3.6 Mean, standard deviation, and mean to standard deviation ratio for all nine features for NGLCM-6 for both NOR and OSA groups

	NOR (N = 92)			OSA (N = 106)		
	Mean (μ)	St Dev (σ)	(σ/μ) %	Mean (μ)	St Dev (σ)	(σ/μ) %
ENT_6	3.13069	0.61297	19.58	2.51151	0.56809	22.62
ASM_6	0.11126	0.06179	55.53	0.19188	0.09139	47.63
CON_6	1.30730	0.80438	61.53	0.76175	0.39854	52.32
COR_6	0.92846	0.01459	1.57	0.93373	0.01971	2.11
DIS_6	0.66388	0.26590	40.05	0.43027	0.16813	39.08
IND_6	0.74467	0.07568	10.16	0.82351	0.05963	7.24
IDM_6	0.99874	0.00077	0.08	0.99926	0.00038	0.04
VAR_6	10.84714	5.60850	51.70	7.44534	3.69733	49.66
INR_6	2.82015	0.53732	19.05	2.29028	0.49880	21.78

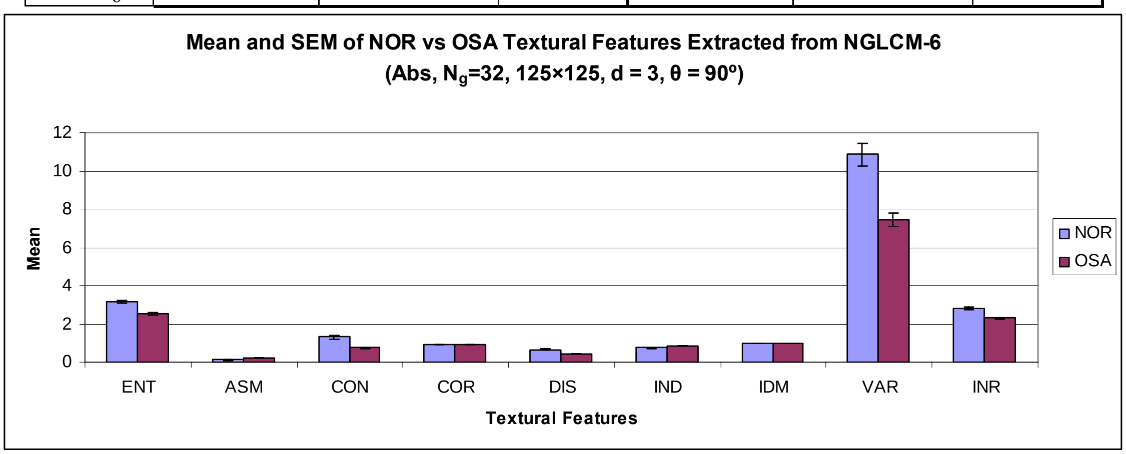

Figure 3.7 Comparison between the mean of the features extracted from NGLCM-6 between NOR and OSA groups. The error bars represent the SEM.

3.2.7 Features Extracted from NGLCM-7

NGLCM-7 were normalized GLCM extracted from the gray-scale image scheme described in Section 2.3.2.2. This image has N_g=32, representing the absolute magnitude of the power spectrum of the time frequency plot of the clip. The full scale of the frequency was used (125 points). In calculating the GLCM, pairing of the pixels was at distance d=1 and orientation θ=90˚. Table 3.7 shows a tabulation of the statistics of the features extracted from NGLCM-7 for all clips extracted from the NOR and OSA groups. Figure 3.8 shows a graphical representation of these results.

Table 3.7 Mean, standard deviation, and mean to standard deviation ratio for all nine features for NGLCM-7 for both NOR and OSA groups

	NOR (N = 92)			OSA (N = 106)		
	Mean (μ)	St Dev (σ)	(σ/μ) %	Mean (μ)	St Dev (σ)	(σ/μ) %
ENT_7	2.61686	0.49995	19.10	2.11759	0.46367	21.90
ASM_7	0.15445	0.06867	44.46	0.23860	0.09215	38.62
CON_7	0.24501	0.10922	44.58	0.15682	0.06403	40.83
COR_7	0.98408	0.00576	0.59	0.98350	0.00785	0.80
DIS_7	0.22534	0.09046	40.14	0.14588	0.05746	39.39
IND_7	0.89052	0.04234	4.76	0.92882	0.02774	2.99
IDM_7	0.99976	0.00011	0.01	0.99985	0.00006	0.01
VAR_7	11.15492	5.61780	50.36	7.83261	3.77899	48.25
INR_7	2.46960	0.44574	18.05	2.02210	0.42752	21.14

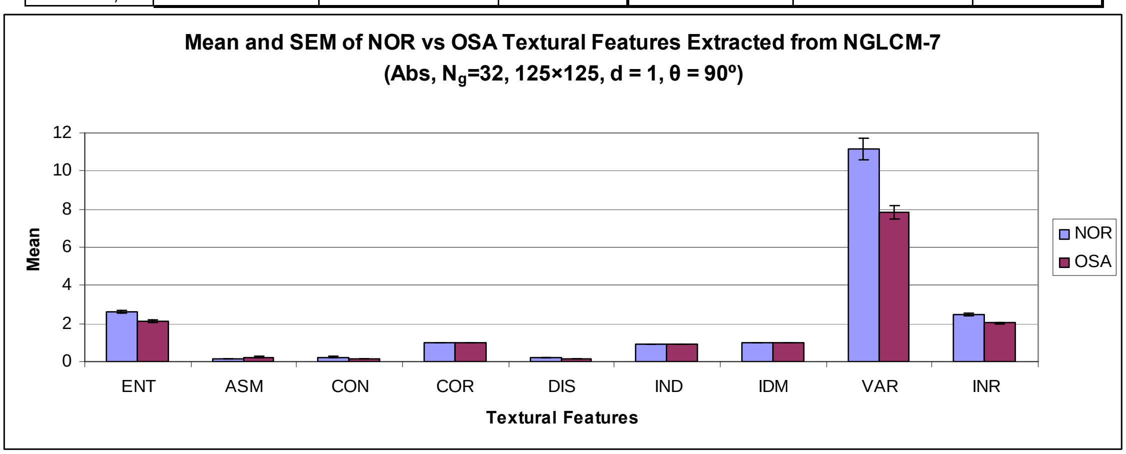

Figure 3.8 Comparison between the mean of the features extracted from NGLCM-7 between NOR and OSA groups. The error bars represent the SEM.

3.2.8 Features Extracted from NGLCM-8

NGLCM-8 were normalized GLCM extracted from the gray-scale image scheme described in Section 2.3.2.3. This image has N_g=16, representing the log of the absolute magnitude of the power spectrum of the time frequency plot of the clip. The full scale of the frequency was used (125 points). In calculating the GLCM, pairing of the pixels was at distance d=5 and orientation θ=90˚. Table 3.8 shows a tabulation of the statistics of the features extracted from NGLCM-8 for all clips extracted from the NOR and OSA groups. Figure 3.9 shows a graphical representation of these results.

Table 3.8 Mean, standard deviation, and mean to standard deviation ratio for all nine features for NGLCM-8 for both NOR and OSA groups

	NOR (N = 92)			OSA (N = 106)		
	Mean (μ)	St Dev (σ)	(σ/μ) %	Mean (μ)	St Dev (σ)	(σ/μ) %
ENT_8	3.10340	0.19130	6.16	3.05394	0.18498	6.06
ASM_8	0.06564	0.01447	22.04	0.07053	0.01454	20.62
CON_8	1.06220	0.20693	19.48	0.92169	0.15636	16.96
COR_8	0.78911	0.04556	5.77	0.81818	0.02292	2.80
DIS_8	0.68506	0.08656	12.64	0.62603	0.07232	11.55
IND_8	0.70839	0.02863	4.04	0.72762	0.02568	3.53
IDM_8	0.99594	0.00078	0.08	0.99647	0.00059	0.06
VAR_8	2.68323	0.62907	23.44	2.65270	0.52775	19.89
INR_8	2.73089	0.16808	6.15	2.70256	0.15844	5.86

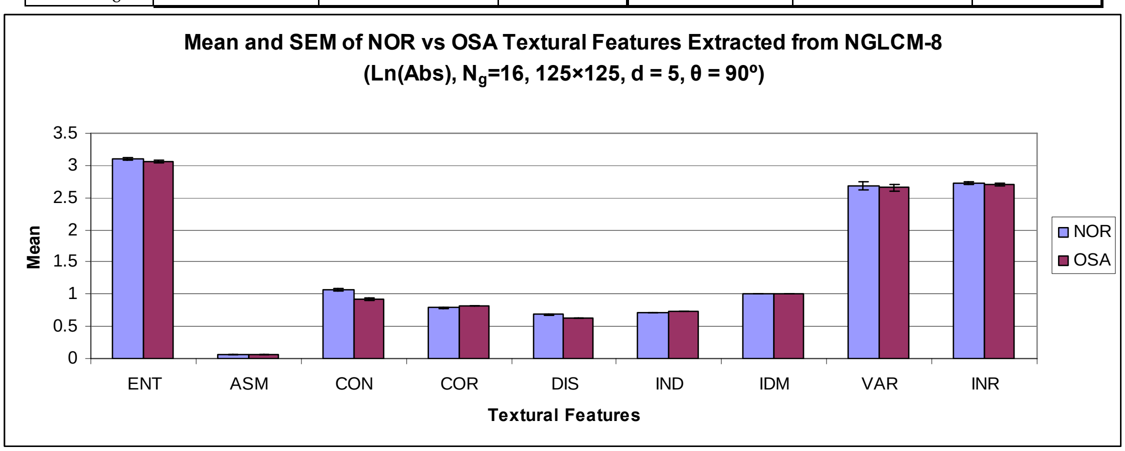

Figure 3.9 Comparison between the mean of the features extracted from NGLCM-8 between NOR and OSA groups. The error bars represent the SEM.

3.2.9 Features Extracted from NGLCM-9

NGLCM-9 were normalized GLCM extracted from the gray-scale image scheme described in Section 2.3.2.3. This image has N_g=16, representing the log of the absolute magnitude of the power spectrum of the time frequency plot of the clip. Half of the scale of the frequency was used (125 points). In calculating the GLCM, pairing of the pixels was at distance d=5 and orientation θ=90˚. Table 3.9 shows a tabulation of the statistics of the features extracted from NGLCM-9 for all clips extracted from the NOR and OSA groups. Figure 3.10 shows a graphical representation of these results.

Table 3.9 Mean, standard deviation, and mean to standard deviation ratio for all nine features for NGLCM-9 for both NOR and OSA groups

	NOR (N = 92)			OSA (N = 106)		
	Mean (μ)	St Dev (σ)	(σ/μ) %	Mean (μ)	St Dev (σ)	(σ/μ) %
ENT_9	2.35719	0.18907	8.02	2.45024	0.22939	9.36
ASM_9	0.15187	0.03184	20.97	0.13740	0.03541	25.77
CON_9	0.39263	0.05654	14.40	0.38551	0.05835	15.14
COR_9	0.87220	0.02253	2.58	0.88936	0.03607	4.06
DIS_9	0.31210	0.03495	11.20	0.31325	0.03713	11.85
IND_9	0.85455	0.01490	1.74	0.85306	0.01593	1.87
IDM_9	0.99849	0.00021	0.02	0.99852	0.00022	0.02
VAR_9	1.68962	0.45560	26.96	2.08339	0.79281	38.05
INR_9	2.16324	0.17309	8.00	2.25385	0.21669	9.61

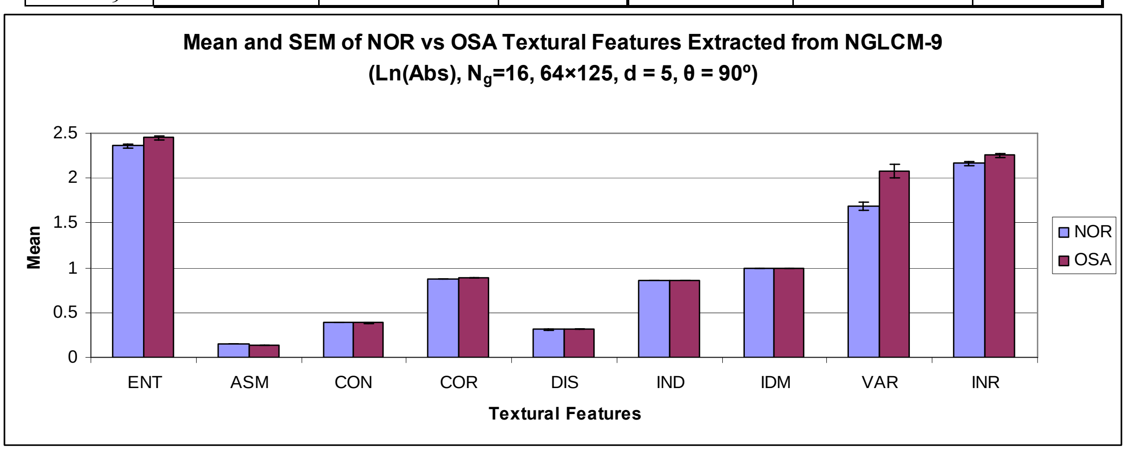

Figure 3.10 Comparison between the mean of the features extracted from NGLCM-9 between NOR and OSA groups. The error bars represent the SEM.

3.2.10 Features Extracted from NGLCM-10

NGLCM-10 were normalized GLCM extracted from the gray-scale image scheme described in Section 2.3.2.4. This image has N_g=16, representing the histogram of absolute magnitude of the power spectrum of the time frequency plot of the clip. The full scale of the frequency was used (125 points). In calculating the GLCM, pairing of the pixels was at distance d=5 and orientation θ=90˚. Table 3.10 shows a tabulation of the statistics of the features extracted from NGLCM-10 for all clips extracted from the NOR and OSA groups. Figure 3.11 shows a graphical representation of these results.

Table 3.10 Mean, standard deviation, and mean to standard deviation ratio for all nine features for NGLCM-10 for both NOR and OSA groups

	NOR			OSA		
	Mean (μ)	St Dev (σ)	(σ/μ) %	Mean (μ)	St Dev (σ)	(σ/μ) %
ENT_{10}	3.67206	0.10979	2.99	3.62396	0.06916	1.91
ASM_{10}	0.03628	0.00566	15.60	0.03808	0.00427	11.21
CON_{10}	0.53098	0.08413	15.84	0.48703	0.04576	9.40
COR_{10}	0.98661	0.00183	0.19	0.98772	0.00095	0.10
DIS_{10}	0.40910	0.05331	13.03	0.37571	0.02975	7.92
IND_{10}	0.81215	0.02238	2.76	0.82674	0.01307	1.58
IDM_{10}	0.99796	0.00032	0.03	0.99813	0.00017	0.02
VAR_{10}	20.19158	1.33592	6.62	20.20932	0.87440	4.33
INR_{10}	3.42489	0.08335	2.43	3.39575	0.05445	1.60

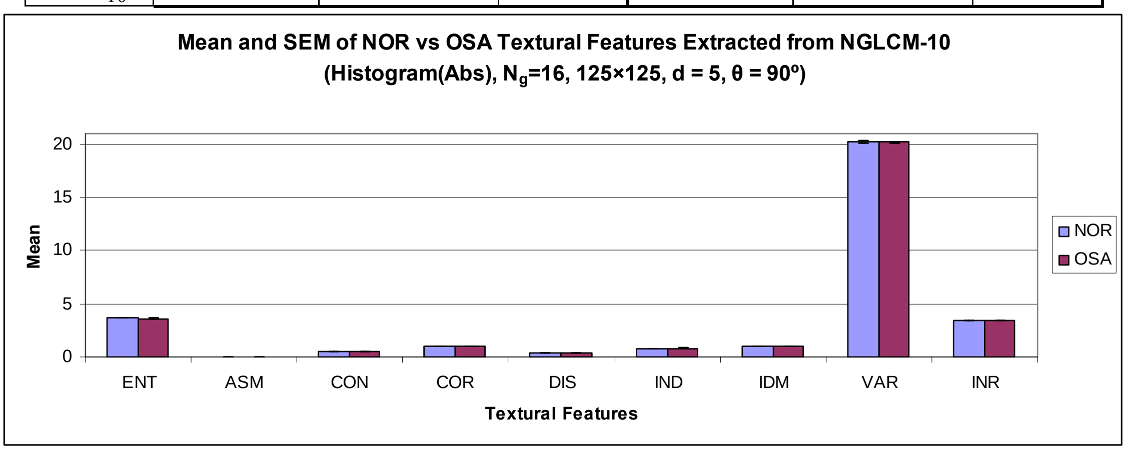

Figure 3.11 Comparison between the mean of the features extracted from NGLCM-10 between NOR and OSA groups. The error bars represent the SEM.

3.2.11 Comparison between Features with Varying Pairing Distances

It was noticed from the previous sections that varying the pairing distance have showed a distinct effect on the values of the features extracted from each image regardless of the clip classification. A study of the changes in values of the features ENT, CON, DIS, COR, and IND extracted from NGLCM-5, NGLCM-6 and NGLCM-7 with varying pairing distance, *d*, is shown in Figure 3.12. Each point represent the average value of the feature for NOR and OSA clips.

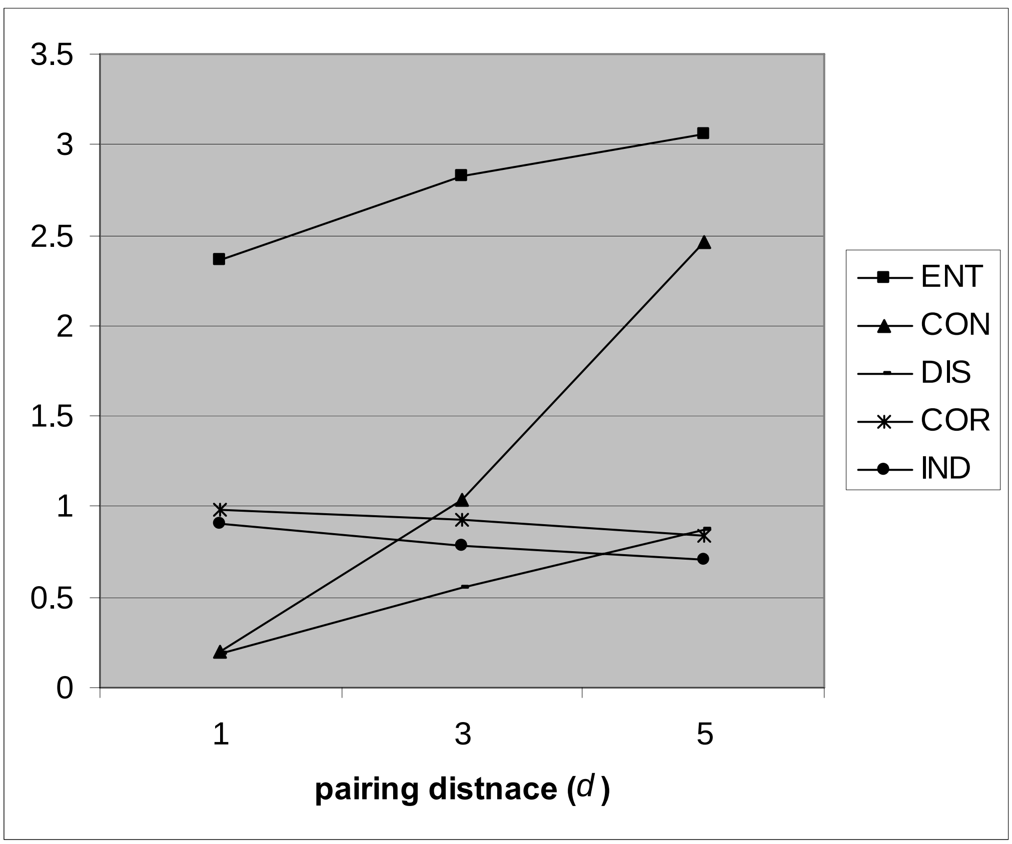

Figure 3.12 Comparison between different features with varying pairing distances.

3.3 Feature Selection

3.3.1 Feature Selection Using Statistical Difference between Features

In Section 2.4.5.2, it was proposed to use a 1st degree statistical feature selection by finding the features that least ($\frac{\sigma}{\mu}$) percentage. In Section 3.2, tabulations of the features statistical analysis were presented. It is noticed the COR, INV, and IDM features from all 10 NGLCM's had the lowest $\frac{\sigma}{\mu}$ ratio. This reduces the number of selected features per NGLCM from 9 to 3 (total of 30, down from 90). These thirty features will be the input to the detector and the results will be compared to these from using all ninety features and those from the features selected in the following section.

3.3.2 Feature Selection Using Piecewise Linear Network

In Section 2.4.5.3, the concept of using a Piecewise Linear Network in fitting the data and selecting an optimal feature subset using floating search algorithm was described. Using a PLN algorithm, the following seventeen features comprise the optimal feature subset. These features are:

COR_2	VAR_2	ASM_1	ENT_{10}	DIS_7	IND_2	IDM_{10}	INR_6
COR_3	VAR_3	ASM_2					
COR_5	VAR_8	ASM_{10}					
COR_6	VAR_{10}						
COR_9							

Table 3.11 summarizes these features by NGLCM number.

Table 3.11 a break down of feature contribution per NGLCM

	NGLCM number									
	1	2	3	4	5	6	7	8	9	10
ENT										×
ASM	×	×								×
CON										
COR		×	×		×	×			×	
DIS							×			
IND		×								
IDM										×
VAR		×	×					×		×
INR						×				

These seventeen features will be inputs to the NN detector and their results will be compared with the results from those from the control results and those from the features found in the previous section.

3.4 Detection Results

Based on Section 2.4.5, three sets of input features are introduced and are to be compared. This brings forth the construction of three different detection networks. The detection results of these three networks are to be compared. These three networks were separately optimized, trained, and performance results where found using a 1000-run Monte-Carlo simulation.

3.4.1 Network Optimization

As detailed in Section 2.5, the goal in the MLP network optimization phase is to find the optimum number of hidden units, and the optimum number of training epochs that will result in the best performance. The best performance is defined as the maximum accuracy the MLP network can find.

For each one of the three networks, the tested hidden units (N_h) were 2, 4, 6, 8, 10, 14, 18, 22, 26, 30, 35, 40, 45, and 50. At each hidden unit value, a 50-run Monte-Carlo simulation was performed, and the training and testing results were stored at fixed training epochs. For each run, a new and random assignment of the training and testing set were used.

The number of possible distinct arrangements that the example vectors can assume can be calculated using the general un-ordered arrangement equation:

$$\text{Number of arrangements} = \frac{m!}{(m-n)!\,n!} = \frac{198!}{(198-132)!132!} = 3.2553 \times 10^{53}$$

As it shows, this number is found to be very large; which means that it is practically impossible to have a train-test pair ran twice in any given Monte-Carlo simulation.

The following subsections illustrate the results found for the three networks optimization operations.

3.4.1.1 Optimization Using the Complete Feature Set

Table 3.12 shows the results of the average performance using all 90 features as inputs to the MLP network. For tabulation purposes, only the testing accuracy mean and standard deviation are shown.

Table 3.12 Testing accuracy mean and standard deviation (std) for 50-run Monte-Carlo Simulation. Values shown are per each N_h and N_{ep} pair for an MLP with 90 input features

N_h	Test Accuracy	Number of Training Epochs N_{ep}												
		200	600	1000	1400	1800	2200	2600	3000	3400	3800	4200	4600	5000
2	Mean	74.79	83.42	85.52	86.15	86.39	86.18	86.36	86.21	86.06	85.79	85.94	85.82	85.79
	Std	8.59	4.92	4.86	5.17	5.29	5.32	5.28	5.27	5.48	5.33	5.33	5.18	5.22
4	Mean	75.55	83.88	85.24	85.18	85.45	85.30	85.27	85.67	85.58	85.67	85.48	85.48	85.55
	Std	7.99	6.00	5.87	6.31	6.17	6.05	6.25	6.12	6.02	5.87	5.85	5.72	5.70
6	Mean	77.58	83.73	85.42	85.64	86.42	86.33	86.45	86.42	86.21	86.15	85.91	85.82	86.03
	Std	8.98	6.93	6.75	6.66	4.81	4.88	5.07	5.16	5.11	5.13	4.99	5.09	5.17
8	Mean	76.67	85.30	86.61	87.73	88.12	87.91	87.88	87.70	87.70	87.39	87.42	87.58	87.61
	Std	8.98	7.48	6.05	5.75	5.30	5.26	4.89	4.88	4.83	4.78	4.76	4.88	4.85
10	Mean	77.21	84.67	86.06	86.33	86.58	87.33	86.97	86.58	86.67	86.24	86.27	86.42	86.61
	Std	9.12	6.13	5.70	6.13	5.98	4.49	4.79	4.64	4.61	4.55	4.61	4.63	4.54
14	Mean	76.09	83.45	85.21	85.97	86.27	86.03	86.00	86.06	85.88	85.91	85.94	85.91	85.67
	Std	6.25	6.54	5.33	5.26	5.29	5.30	5.57	5.07	5.29	5.24	5.34	5.23	5.11
18	Mean	76.18	83.55	85.67	85.85	86.39	86.45	86.64	86.52	86.67	86.73	86.61	86.55	86.58
	Std	6.71	6.26	5.05	5.38	5.24	5.16	5.35	5.34	5.23	5.35	5.29	5.24	5.31
22	Mean	76.15	82.94	84.48	85.12	85.42	85.64	85.82	85.97	85.70	85.70	86.39	86.24	86.18
	Std	7.13	6.89	5.88	6.07	6.09	5.91	6.00	6.14	6.06	5.95	5.06	5.07	5.11
26	Mean	75.58	82.58	85.18	86.12	85.91	85.91	85.73	85.88	85.91	85.82	85.67	85.64	85.76
	Std	7.39	6.39	5.06	5.16	5.23	5.18	4.97	4.95	5.02	5.02	5.02	4.98	4.97
30	Mean	75.52	81.70	83.21	84.30	85.70	85.48	85.97	85.82	85.82	85.67	85.61	85.33	85.52
	Std	8.67	8.61	8.56	8.08	5.72	5.29	5.05	5.33	5.38	5.55	5.59	5.71	5.71
35	Mean	75.76	82.67	84.48	85.27	85.45	85.67	85.94	85.94	85.97	86.03	86.15	86.27	86.42
	Std	7.57	7.01	5.52	5.24	5.38	4.54	4.29	4.60	4.35	4.45	4.40	4.54	4.53
40	Mean	74.39	82.30	84.39	84.73	85.15	85.27	85.39	85.33	85.64	85.39	85.18	85.33	85.27
	Std	7.71	7.82	6.52	5.93	5.86	5.69	5.85	5.52	5.49	5.63	5.61	5.59	5.46
45	Mean	76.45	84.30	86.33	87.15	87.79	87.67	87.30	87.48	87.52	87.55	87.36	87.33	87.12
	Std	6.03	5.92	4.46	4.11	4.39	4.58	5.01	5.29	5.42	5.42	5.41	5.07	4.87
50	Mean	74.24	79.88	82.30	83.36	83.45	83.61	84.06	83.94	83.94	83.76	83.70	83.79	83.67
	Std	6.24	6.53	6.57	6.02	6.29	6.22	5.55	5.74	5.81	6.21	6.35	5.94	5.90

Figure 3.13 shows a graphical representation of the optimization results for five selected hidden unit values: 8, 14, 18, 26, and 45. This graph shows the mean training and testing curves for this MLP.

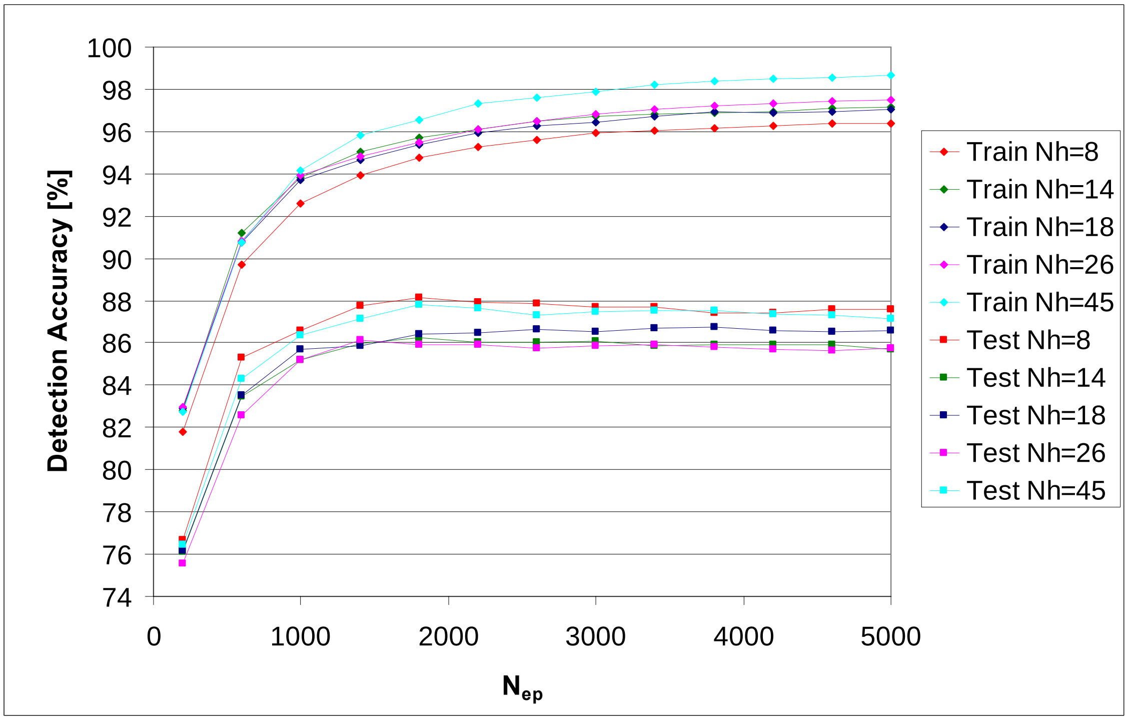

Figure 3.13 Training and testing accuracy for selected hidden units shown in Table 3.12. 90 features are used as inputs to the MLP.

3.4.1.2 Optimization Using 30 Selected Features

Table 3.13 shows the results of the average performance using 30 selected features as in Section 3.3.1 as inputs to the MLP network. For tabulation purposes, only the testing accuracy mean and standard deviation are shown.

Table 3.13 Testing accuracy mean and standard deviation (std) for 50-run Monte-Carlo Simulation. Values shown are per each N_h and N_{ep} pair for an MLP with 30 input features

N_h	Validation Accuracy	Number of Training Epochs N_{ep}												
		200	600	1000	1400	1800	2200	2600	3000	3400	3800	4200	4600	5000
2	Mean	78.91	79.55	78.73	77.91	77.94	77.64	77.24	77.15	76.94	76.79	76.64	76.36	76.64
	Std	3.91	4.33	4.58	4.34	4.04	3.83	4.08	3.92	4.11	4.12	4.51	4.29	4.20
4	Mean	78.61	78.18	77.70	77.24	76.97	76.58	76.45	76.09	76.12	75.97	76.09	76.06	75.94
	Std	4.36	4.36	4.52	4.36	4.49	4.12	3.86	3.80	3.74	3.82	3.85	3.75	3.47
6	Mean	78.24	77.39	77.21	76.52	76.33	75.85	75.67	75.79	75.36	75.21	75.18	75.21	75.24
	Std	3.93	5.07	5.08	4.82	4.40	3.97	4.19	3.96	3.68	3.70	3.54	3.47	3.53
8	Mean	77.73	77.21	77.00	76.45	75.52	74.70	74.64	74.42	74.39	74.42	74.24	74.18	74.33
	Std	4.11	4.26	4.00	3.90	3.70	3.78	3.70	3.57	3.76	3.58	3.45	3.24	3.13
10	Mean	78.18	78.82	77.73	77.06	76.55	76.45	75.91	75.64	75.48	75.61	75.30	75.24	75.21
	Std	4.54	3.35	3.47	3.93	3.53	3.51	3.38	2.90	2.98	2.89	2.78	2.70	2.70
14	Mean	78.76	77.91	77.24	76.82	76.70	76.06	75.73	75.36	75.33	75.21	75.21	75.18	75.09
	Std	3.89	4.45	3.91	3.90	4.04	3.54	3.60	3.29	3.31	3.29	3.29	3.24	3.06
18	Mean	78.58	77.64	76.48	75.85	75.73	75.52	75.39	74.88	75.06	74.94	75.03	74.94	74.82
	Std	4.25	4.11	3.98	3.63	3.62	3.73	3.64	3.40	3.31	3.15	3.13	3.08	3.15
22	Mean	78.48	78.15	77.36	77.00	76.52	76.09	75.85	75.36	75.27	75.27	75.15	75.18	75.00
	Std	4.68	4.55	4.03	3.96	3.97	3.57	3.85	3.55	3.56	3.56	3.46	3.48	3.43
26	Mean	78.97	77.94	76.61	76.45	76.12	75.58	75.45	75.18	74.97	74.88	74.67	74.82	74.64
	Std	3.59	3.69	4.15	3.98	3.85	3.35	3.60	3.50	3.40	3.30	3.18	3.12	3.10
30	Mean	79.67	78.61	77.85	77.00	76.45	76.33	76.00	75.82	75.61	75.58	75.45	75.39	75.21
	Std	4.54	3.69	3.99	4.10	3.56	3.15	3.30	3.37	3.24	3.25	3.33	3.30	3.45
35	Mean	78.55	77.64	76.55	76.52	75.97	75.48	75.18	75.24	75.21	75.24	75.18	75.15	75.06
	Std	4.42	3.87	3.60	3.93	3.53	3.31	3.41	3.55	3.67	3.46	3.25	3.14	2.97
40	Mean	78.55	77.33	76.61	75.94	75.09	74.91	74.82	74.76	74.76	74.79	74.73	74.67	74.48
	Std	4.15	4.49	3.74	4.00	3.48	3.17	3.30	3.26	3.18	3.16	3.05	2.98	2.96
45	Mean	79.42	78.21	77.82	77.36	76.67	76.12	75.70	75.64	75.61	75.42	75.42	75.45	75.42
	Std	4.65	4.34	4.27	4.32	3.81	3.63	3.57	3.35	3.34	2.99	3.02	3.08	3.14
50	Mean	79.18	78.94	77.97	77.06	76.55	75.97	75.64	75.52	75.33	75.18	74.85	74.82	74.91
	Std	4.32	3.76	3.87	3.71	3.65	3.09	2.93	3.13	3.31	3.14	3.03	3.04	3.08

Figure 3.14 shows a graphical representation of the optimization results for five selected hidden unit values: 8, 14, 18, 26, and 45. This graph shows the mean training and testing curves for this MLP.

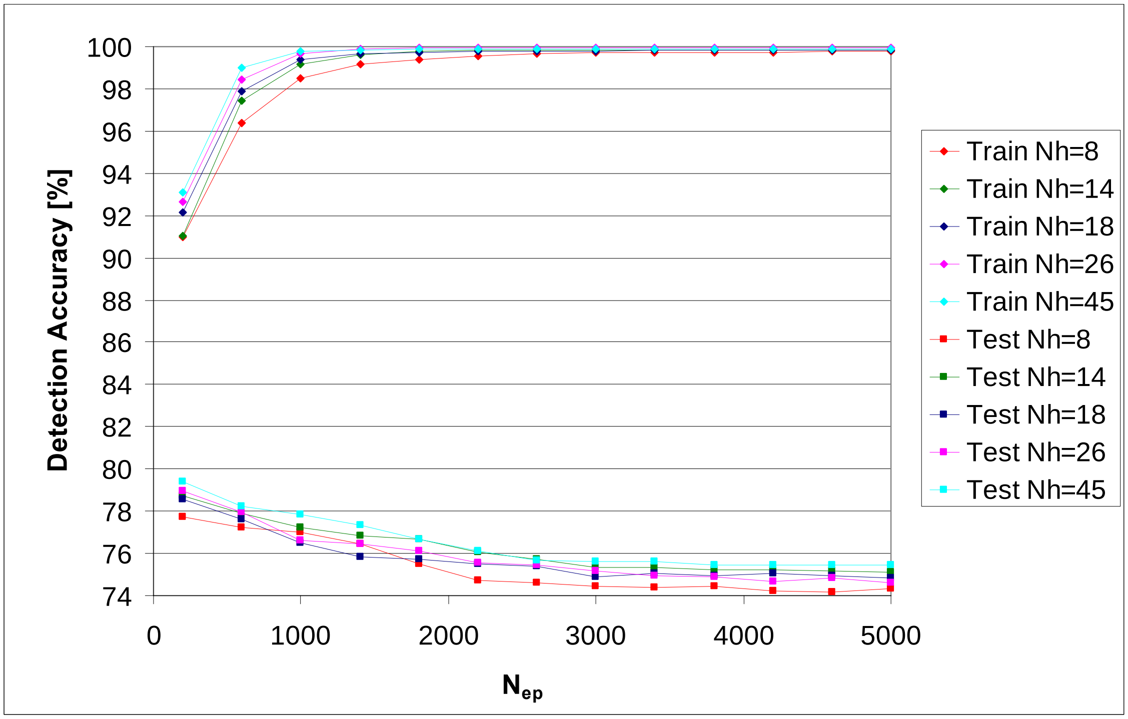

Figure 3.14 Training and testing accuracy for selected hidden units shown in Table 3.13. 30 selected features are used as inputs to the MLP.

3.4.1.2 Optimization Using 17 Selected Features

Table 3.14 shows the results of the average performance using 17 selected features as in Section 3.3.2 as inputs to the MLP network. For tabulation purposes, only the testing accuracy mean and standard deviation are shown.

Table 3.14 Testing accuracy mean and standard deviation (std) for 50-run Monte-Carlo Simulation. Values shown are per each N_h and N_{ep} pair for an MLP with 17 input features

N_h	Validation Accuracy	Number of Training Epochs N_{ep}												
		200	600	1000	1400	1800	2200	2600	3000	3400	3800	4200	4600	5000
2	Mean	81.94	86.88	87.70	88.24	88.18	88.30	88.46	88.58	88.42	88.39	88.46	88.33	88.42
	Std	5.77	4.54	3.92	4.00	4.25	4.52	4.32	4.48	4.69	4.75	4.56	4.60	4.58
4	Mean	81.82	85.03	87.42	87.52	87.79	87.73	87.67	87.79	87.88	87.94	87.64	87.76	87.82
	Std	6.41	6.72	4.44	3.87	4.16	4.40	4.54	4.74	4.85	4.65	4.89	4.68	4.75
6	Mean	80.24	84.73	86.36	86.79	87.09	87.67	87.73	87.91	88.33	88.18	87.97	88.00	88.03
	Std	5.24	4.25	4.47	4.47	4.64	4.16	4.24	4.19	4.19	4.29	4.27	4.42	4.62
8	Mean	81.06	84.97	86.48	88.00	87.94	89.09	89.30	89.30	89.33	89.27	89.39	89.39	89.30
	Std	8.52	7.06	7.05	7.01	7.35	4.37	4.34	4.20	4.15	4.28	4.27	4.46	4.46
10	Mean	80.27	85.33	87.88	88.73	88.76	88.36	88.76	88.73	88.70	88.85	88.73	89.21	89.21
	Std	8.66	5.14	4.49	4.09	4.02	3.96	3.90	4.14	4.24	4.10	4.21	3.89	4.33
14	Mean	79.36	85.85	88.15	89.09	89.94	90.48	90.39	90.79	90.73	91.18	90.94	91.24	91.33
	Std	9.81	6.92	5.57	4.60	4.56	3.67	4.19	3.83	3.96	3.83	4.01	3.72	3.91
18	Mean	81.52	86.15	88.15	89.55	90.03	90.30	90.52	90.70	90.82	90.64	90.64	90.61	90.64
	Std	6.36	5.47	4.08	4.06	4.34	4.16	3.76	3.76	3.72	3.93	3.87	3.94	4.06
22	Mean	79.82	84.58	86.48	87.85	88.61	88.94	89.15	88.91	89.30	89.21	89.55	89.79	89.85
	Std	9.88	4.24	4.68	4.10	4.21	3.96	4.44	4.91	4.65	4.76	4.96	4.79	4.77
26	Mean	79.48	85.27	87.88	89.09	89.76	90.18	90.79	90.70	90.55	90.67	90.64	90.64	90.52
	Std	7.78	6.03	4.97	5.36	5.23	4.88	4.58	4.65	4.86	5.15	4.93	5.03	5.04
30	Mean	79.91	85.09	86.94	87.79	88.61	89.42	89.55	89.58	89.58	89.70	89.67	89.97	89.82
	Std	7.48	5.08	4.78	4.96	5.32	5.17	5.10	5.21	5.35	5.13	5.05	4.91	4.85
35	Mean	80.00	84.30	85.36	86.27	87.12	87.58	88.18	88.39	88.55	88.61	88.64	88.73	88.55
	Std	6.33	4.99	5.12	5.17	5.10	4.95	4.87	4.94	4.73	4.66	4.85	4.60	4.76
40	Mean	79.24	84.58	86.79	87.85	89.03	89.36	89.79	89.82	90.06	89.97	89.91	89.94	90.15
	Std	10.07	6.37	5.45	5.63	4.96	4.88	4.82	5.05	4.52	4.62	4.91	5.04	4.77
45	Mean	80.39	85.88	87.61	88.48	89.27	89.73	89.85	90.27	90.39	90.30	90.42	90.36	90.45
	Std	8.30	3.95	4.07	4.37	3.86	3.84	3.93	3.93	4.06	4.08	3.97	3.93	4.02
50	Mean	78.06	84.09	86.36	87.52	88.94	89.73	90.03	90.09	90.09	89.97	90.15	90.30	90.18
	Std	9.93	7.95	7.73	7.98	5.44	5.32	5.42	5.30	5.05	4.96	5.08	4.99	5.32

Figure 3.15 shows a graphical representation of the optimization results for five selected hidden unit values: 8, 14, 18, 26, and 45. This graph shows the mean training and testing curves for this MLP.

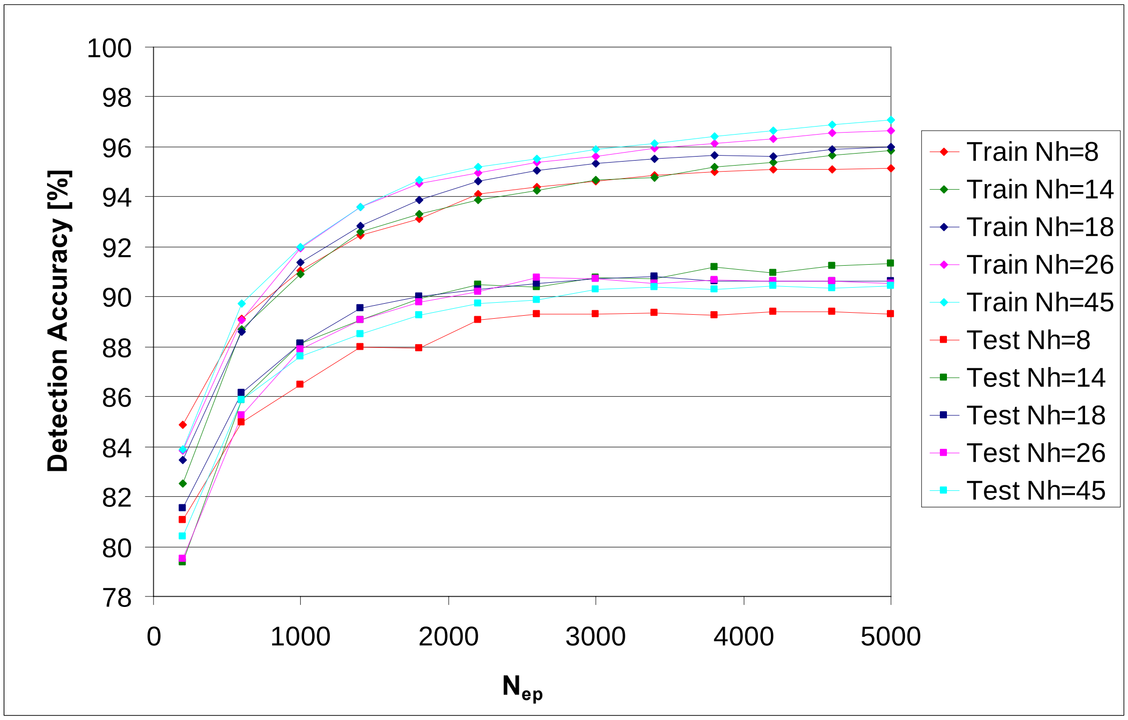

Figure 3.15 Training and testing accuracy for selected hidden units shown in Table 3.14. 17 selected features are used as inputs to the MLP.

3.4.2 Monte-Carlo Simulation

A 1000-run Monte-Carlo simulation is performed to find the overall performance and robustness of the suggested algorithm for the three discussed networks. A random assignment of training and testing vectors was performed, as described in Section 3.4, and the network's training and testing performance was calculated after each run. The choice of N_h and N_{ep} for each of the three Monte-Carlo simulations was found after examining all potential results in Section 3.4.1.

3.4.2.1 Monte-Carlo Simulation Using all 90 Features

After examining all potentially promising N_h and N_{ep} pairs in Table 3.12 that yielded the best validation accuracy, N_h = 8 with N_{ep} = 1800 were chosen.

The following Table 3.15 summarizes the results of the 1000-run Monte-Carlo Simulation.

Table 3.15 Sensitivity, specificity and accuracy mean and standard deviation (std) for training and testing sets after 1000-run Monte-Carlo simulation for a MLP of 90 inputs

	Training			Testing		
	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)
Mean ± std	97.06 ± 6.58	90.39 ± 5.93	93.96 ± 4.29	90.67 ± 8.29	81.42 ± 8.15	86.36 ± 5.46

Figure 3.16 is a graphical Representation of these results.

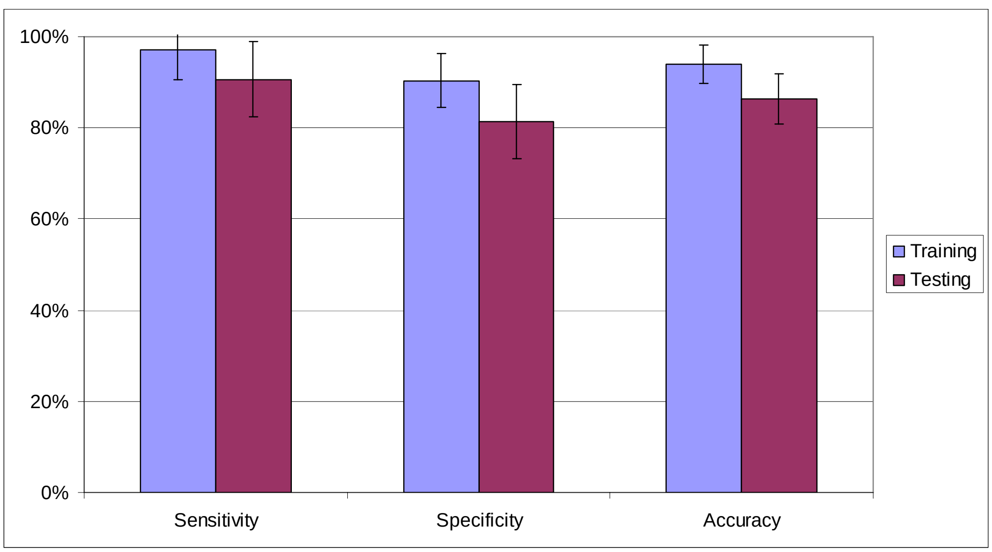

Figure 3.16 Graphical representation of the sensitivity, specificity and accuracy mean for training and testing sets after 1000-run Monte-Carlo simulation and MLP of 90 inputs. Error bars represent the standard deviation.

3.4.2.2 Monte-Carlo Simulation Using 30 Features

After examining all potentially promising N_h and N_{ep} pairs in Table 3.13 that yielded the best validation accuracy, $N_h = 45$ with $N_{ep} = 200$ were chosen.

The following Table 3.16 summarizes the results of the 1000-run Monte-Carlo Simulation.

Table 3.16 Sensitivity, specificity and accuracy mean and standard deviation (std) for training and testing sets after 1000-run Monte-Carlo simulation for a MLP of 30 inputs

	Training			Testing		
	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)
Mean ± std	81.91 ± 5.76	68.09 ± 4.86	75.70 ± 3.11	82.10 ± 7.92	67.93 ± 8.03	75.14 ± 3.28

Figure 3.17 is a graphical Representation of these results.

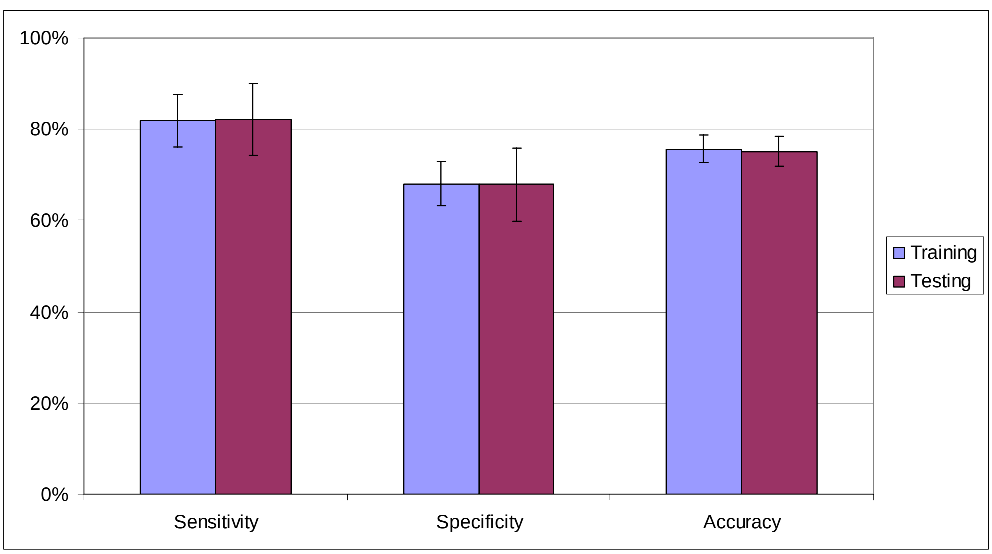

Figure 3.17 Graphical representation of the sensitivity, specificity and accuracy mean for training and testing sets after 1000-run Monte-Carlo simulation and MLP of 30 inputs. Error bars represent the standard deviation.

3.4.2.3 Monte-Carlo Simulation Using 17 Features

After examining all potentially promising N_h and N_{ep} pairs in Table 3.14 that yielded the best validation accuracy, $N_h = 8$ with $N_{ep} = 1800$ were chosen.

The following Table 3.17 summarizes the results of the 1000-run Monte-Carlo Simulation.

Table 3.17 Sensitivity, specificity and accuracy mean and standard deviation (std) for training and testing sets after 1000-run Monte-Carlo simulation for a MLP of 17 inputs

	Training			Testing		
	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)
Mean ± std	99.00 ± 1.84	93.41 ± 6.67	96.42 ± 3.50	94.42 ± 4.71	85.40 ± 8.49	90.16 ± 4.57

Figure 3.18 is a graphical Representation of these results.

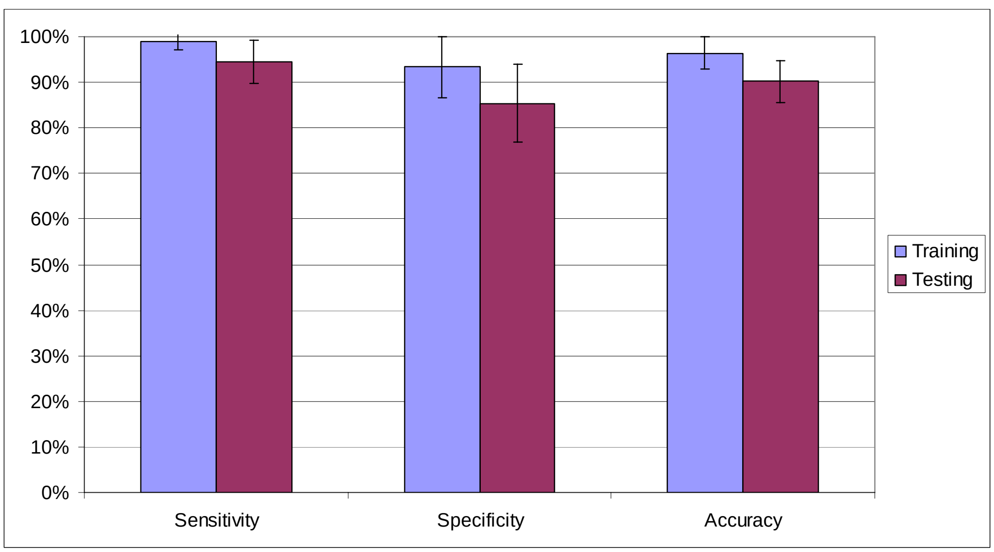

Figure 3.18 Graphical representation of the sensitivity, specificity and accuracy mean for training and testing sets after 1000-run Monte-Carlo simulation and MLP of 17 inputs. Error bars represent the standard deviation.

3.4.3 Comparison between the Detection Results

A summary of the testing sensitivity, specificity and accuracy results of the Monte-Carlo simulation for the three MLP networks is shown in the following table.

Table 3.18 Summery of the sensitivity, specificity and accuracy mean and standard deviation (std) for testing sets after 1000-run Monte-Carlo simulation for a the three discussed MLPs having input features of 90, 30 and 17

Number of input features to the MLP	Testing (mean ± std)		
	Sensitivity (%)	Specificity (%)	Accuracy (%)
90	90.67 ± 8.29	81.42 ± 8.15	86.36 ± 5.46
30	82.10 ± 7.92	67.93 ± 8.03	75.14 ± 3.28
17	94.42 ± 4.71	85.40 ± 8.49	90.16 ± 4.57

CHAPTER 4

DISCUSSION

This chapter contains the discussion of the different results obtained in this investigation. Section 4.1 discusses the rationale behind choosing the signal source. Section 4.2 elaborates on the findings of the clips visual illustrations. Section 4.3 discusses the statistical findings of the textural features when compared between the different images. Section 4.4 discusses the optimization and results of the automatic detection algorithm presented in this investigation.

4.1 Signal Source

4.1.1 Choice of R-R Interval as the Data Source

R-R interval is unique in the sense that it can be extracted from different markers such as ECG and possibly oximetery. This makes it particularly feasible to make relatively cheap and handy overnight recordings, in which the patient can sleep at their home rather at a sleep laboratory.

Furthermore, it has been shown that HRV is physiologically related to sleep events [14]. On the onset of an OSA event and breathing cessation, vagus activity increases, causing the heart to slow down (bradycardia), followed by arousal and marked increase in heart rate (tachycardia). This cyclic behavior in heart rate during an OSA event can be detected by an automated method.

Other studies [8], [14], [15] have shown a limited ability to use the HRV, represented by the R-R interval, as a marker to distinguish normal from eventful sleep clips. This investigation has concentrated on using this discrete index sequence by combining both time and frequency domains.

4.1.2 Choice of Time-Frequency Plots

This research is fundamentally different from the work of de Chazel, et al [2], [36]. Their work depends on extracting statistical features directly from the temporal R-R and EDR sequences (such features include mean, standard deviation, correlations, difference between adjacent features, etc).

This investigation, however, is based on statistical features from images derived from time-frequency plots. This intrinsically combines the properties of time and frequency domains together rather than studying them separately. This is a powerful tool in the sense that known and verified image detection and classifications schemes are readily used towards this application. Furthermore, the automated system developed in this investigation may potentially be applied as a universal tool in engineering and medical application in which a reliable offline classification of signals is needed. Previous studies, [13], [14], have demonstrated qualitative schemes to study these plots, but no quantitative schemes in previous studies were applied to these plots.

4.2 Clip Visualization

A comparison of visual differences in clip is shown in Figures 2.3 and 2.4. The reader, with the previous knowledge of the image classification, can identify differences between the two clips. However, that is the exception. Most time-frequency plots of clips generated [8], showed little visually distinguishable features between the two classifications as shown in Figures 2.5 and 2.6.

Using the visualization, some data is inherently lost using the coding scheme, whether it is colored scheme [28], [31] (Red-Blue-Green (RGB), or Intensity-Hue-Saturation (IHS)) or quantized gray-scale scheme. The data loss is due to use of linear or nonlinear transformations on the complex valued matrices that concentrate on the magnitude of the signal and does not include phase shift. Data loss also arises from the quantization process, in which the plots are quantized to a set number of gray scale levels, rather than having a continuous range of values.

The four different gray-level encoding schemes used in this research have allowed for different distribution of the gray-scale that represent each point in the spectrum. The aim at the beginning of this investigation was to search for a single encoding scheme that will show greatest separation between the two classes. Using the linear magnitude encoding in Section 2.3.2.1, it is noticed that the images produced are predominantly low power, and a lot of power variation within the band 0.2-0.3 Hz are lost due to quantization, especially if a local maximum eclipse them.

Doubling the number of gray levels from 16 to 32 in Section 2.3.2.2, smoothing the roughness of the plot texture was sought. Figures 2.11 and 2.12 are the same clips as in Figures 2.7 and 2.8. It is noticed that they are have a relatively smoother surface compared to their counterpart images in Figures 2.7 and 2.8 with 16 gray-levels. Same argument can be made for images 2.13 and 2.14 compared to their counterpart images in Figures 2.9 and 2.10.

Using the natural logarithm operation in Section 2.3.2.3 has an inherent disadvantage. The goal was to decrease the local maxima in order for the textures on the plot to be more defined. However, this operation is very sensitive to the local minima, where a value that is very close to zero will be magnified, causing the whole image values to be higher around it. This is apparent by the lighter gray shades of the images in Figures 2.15-18. This has produced characteristically different images than their predecessors.

The disadvantage in the given schemes so far is that the plot gray-scale distribution is very sensitive and dependant on the presence of high bursts in the power causing a local maximum. Examining the quantization equations in Sections 2.3.2.1-3, it is noticed that quantization levels are dependent on the maximum value. By using the unequal bin size method described in Section 2.3.2.4, that in the plots will have a textures of that have equal distribution of gray shades as can be seen Figures 2.19-22.

4.3 Textural Features

The first task in this work was finding different images that represent each clip. It is next sought to show if these images are automatically identifiable using the textural feature described in Section 2.4.

Using 1st degree statistical analysis, it is found that for all features, except for the variance feature (VAR), all images were described by statistically different features sets (t-test resulted in α value << 0.05). However, none of these features was adequate as a stand alone classifier. Using this method, it was not feasible to use threshold to distinguish between NOR and OSA clips due to the high spread of the calculated values.

4.3.1 Intra-Image Comparisons

It was initially noticed that for all images, the value of the inverse difference moment feature (IDM) is very close to 1, regardless of the classification (i.e. NOR or OSA), number of gray levels, distance, or orientation. This is an indication that the NGLCM's produced are nearly diagonal. A near diagonal NGLCM is produced when an image does not show great jumps in gray-levels between paired pixels (low contrast) [30]. It is also seen that there is a strong correlation between the ENT and INR features for all images. Correlation (COR) has showed promising results as distinguishing feature, since it showed a reasonable separation power between the means of the two classes with a relatively small spread.

4.3.2 Cross-Image Comparisons

Studying a single feature between different fixed images, with only one variable changed, has showed some noticeable findings. Studying the results in Sections 3.2.1, 3.2.8 and 3.2.10, it is noticed that the gray levels, $N_g = 16$, and distance $d = 5$, are fixed. However, these results represent different encoding schemes, namely *Abs*, *ln*, and *histogram* encodings. It is noticed that the entropy (ENT) of the latter two images increase (*ln* and *histogram*) indicating a visual increase in the roughness of the image. Since ENT is a parameter that measures the randomness of the image [30], we can visually see this effect when studying the images in Sections 2.3.2.1 (*Abs*), 2.3.2.3 (*ln*) and 2.3.2.4 (*histogram*).

For the same group of images, we observed a noticeable increase in the variance in the histogram encoding scheme compared to the other two. Having images expressed and spread on all gray levels equally is a very heterogeneous image, which the feature VAR readily measure [30]. It is worth mentioning that VAR was the most sensitive feature between all images whenever a parameter was changed. However, it showed no statistical difference between the two classes. This led us initially to believe that VAR cannot play a role in the detection algorithm. However, we found that this was untrue, as will be discussed in Section 4.4.

Studying Figure 3.12, it can be seen that changing the pairing distance, d, showed that the smaller the pairing distance (d = 1) the less the image contrast is and the more the homogeneity, compared to higher distance (d = 3 and d = 5). This is expressed by an increase in ENT, CON, and DIS, and decrease in COR and IND with the increase in pairing distance, d.

4.3.3 Bypassing Encoding

The introduction of image processing and textural features described above directly to the time-frequency complex-valued matrices generated in Section 2.3.1 is feasible. These feature values, however, will be complex, and theoretical understanding of the meaning of the features and their variation between the clips remain to be accomplished in the future studies.

4.4 Detection

4.4.1 Feature Selection

After initial trials with using the images separately for automatic detection, the study was later directed to combining all images produced in order to achieve higher sensitivity and specificity results, since each image may hold a certain level of distinguishability between the two classes across the nine features.

First, all ninety extracted features in the detection algorithm were used. Then, observing the results in Tables 3.1-10, it was seen that COR, IND, and IDM show the least $\frac{\sigma}{\mu}$ ratio for all ten NGLCMs. These comprise 30 selected features. Finally, using an optimum selective method in Section 2.4.5.2, it was that found 17 optimum features that yielded best results. The bulk of these features were COR (5 images) and VAR (4 images).

COR, as discussed in Section 4.3.1, was expected to be a keystone classifying feature. However, finding VAR as an important classifying feature gave noticeable improvements. We noticed that NGLCM-2 [Abs, N_g=16, 125×125, d=1, θ=90˚] and NGLCM-10 [Histogram(Abs), N_g=16, 125×125, d=5, θ=90˚] has showed the highest classification contribution, with 4 features each. NGLCM-4 [Abs, N_g=16, 125×125, d=5, θ=0˚] provided no classifying power. This can be contributed to its horizontal pairing, which is perpendicular to the striations originally noted on the images.

4.4.2 MLP Optimization

In Section 3.4.1, all N_h and N_{ep} pairs were calculated for training and testing sets for the three different feature permutations. While the training accuracy improves with N_{ep} for all N_h, testing starts at a much lower accuracy, and improves gradually before reaching an optimum value, leveling off, then declining. This declination is a result of over-fitting the network with the training set and narrowing the decision boundaries to confine smaller areas [32]. Figure 3.13 shows training and testing accuracy curves for selected N_h values.

In Section 3.4.1.1, using all 90 features as inputs to the MLP achieves a training accuracy less than 99% even after a 5000 training epochs, as apparent in Figure 3.13. On the other hand, using the selected 30 features would give a 100% training accuracy within 2500 training epochs, as noted in Figure 3.14. However, from this figure we observe that with an increasing training accuracy, the validation accuracy drops below 75%, which is in the less than acceptable range.

It is apparent from Section 3.4.1.3 that the optimum selection method of Piecewise Linear Network [33] gives an evident advantage over the last two feature permutations, allowing for the testing accuracy to jump over 90% in numerous N_h and N_{ep} pairs.

4.4.3 Monte-Carlo Simulation

A 1000-run Monte-Carlo simulations have been conducted for each of the three networks proposed in this research. As seen in Section 3.4, each network basically differed in the number of features as inputs to the first layer of the MLP. Consequently, different hidden units and training epochs where selected. This selection is based on the highest possible detection rate yielded by that specific network.

Using the complete set of extracted features, it was found that eight hidden units trained for 1800 epochs yielded the best results for this particular network. Section 3.4.2.1 summarizes the results of the 1000-run Monte-Carlo simulations for each of this N_h and N_{ep} pairs. These results are summarized in Table 3.15, and graphically shown in Figure 3.16. This network is considered as the control MLP since there was no attempt to minimize the number of input features. These results are higher than those reported in previous studies [8].

In a first attempt to capitalize on the results found above, and to reduce the calculation time, a reduction in the number of input features from 90 to 30 was implemented, base on statistical differences between the extracted features. It was found that 45 hidden units trained for 200 epochs yielded the best results for this particular network; however, it is obvious from Table 3.16 and Figure 3.17 that the test results are markedly much lower than those for the control network. This shows that using the 1st statistical measure of the $\frac{\sigma}{\mu}$ ratio is not adequate as feature selection criteria. It also shows that this feature subset is not an optimum set and yields validation accuracy below the desired range. Furthermore, the low number of training epochs accounted for the low and almost equal training and testing detection accuracies.

In the second attempt to reduce the input features, an optimum feature selection method was used, and the number of input features was reduced from 90 to 17.

Different 1000-run Monte-Carlo simulations were conducted for this network, due to the fact that multiple N_h and N_{ep} pairs yielded testing set accuracies above 90%, as noticed in Table 3.17. Any of these pairs are potentially the ones that will produce the optimum detection results. From these simulations, the 45 hidden units trained for 4200 epochs yielded the best results. These results are graphically shown in Figure 3.18.

These results are comparable to those reported by de Chazel et al [2], [38]. They reported results for combined features from HRV and R-Wave Attenuation (RWA). Their study shows a test sensitivity and specificity of 86.4% and 92.3% for combined HRV and RWA features. When using HRV temporal features only, the best results they achieved where test sensitivity and specificity of 80.0% and 90.3%, respectively [2].

It should be pointed out here that these reported results are based on minute-by-minute classification, whereas this investigation accounts for 15-minute clip classification, which is recommended to distinguish very low frequency (~0.001Hz) markers in HRV. Also, their described algorithm is based only on a set of time-domain features, whereas our proposed algorithm combines both time and frequency domains and quantitatively uses the time-domain plot to distinguish between normal and apneic clips.

Also, in [8], results of 88% sensitivity and 80% specificity were reported. In [15], an algorithm proposed to distinguish between patients with OSA events, Chyene-Stokes Respiration (CSR), and normal subjects resulted in detection rate accuracies of 70.3%, 91.6% and 94.8% in the training set and 71.8%, 90.1% and 77.1% in the test set, respectively. It is evident that our proposed algorithm has an enhanced performance over the previous studies.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

5.1 Conclusions

A new approach in detecting the presence of SDB in overnight ECG recording has been presented. This approach combines both time and frequency domains in generating pictorial images representing the spectral variation of the HRV over 900-second, 0-0.5 Hz clips. This approach showed very promising results than what has been thought to be a limited marker in previous studies.

Using co-occurrence matrices and their features showed a very promising approach to studying these time-frequency plots. However, calculating the GLCMs is computationally demanding.

We also found that combining different gray-level encoding schemes showed better results than using one scheme, where absolute encoding and histogram encoding with sixteen gray levels showed the most promising results. The correlation (COR), variance (VAR), and angular second moment (ASM), in this order, were the best distinguishing features between all nine textural features selected for this investigation.

The use of optimum feature subset selection algorithm, such as the Branch and Bound algorithm, is found to be crucial in finding the highest detection accuracy.

Relying directly on first-order statistical measures has showed little promise in identifying such optimum feature subsets.

The use of MLP is found very effective in the classification of the clips studied in this investigation. However, a more superior MLP algorithm that has a lower computational load and a faster training algorithm is needed.

5.2 Future Directions

This investigation has concentrated on improving the detection accuracy using a HRV marker, which was previously thought to have limited results. We have taken the extra step in studying the time-frequency plots quantitatively, rather than the qualitative descriptions seen in the literature [2], [13], [14].

Building on the block diagram in Figure 2.1, we can see in Figure 5.1 readily the areas in which this investigation can be expanded. HRV can be detected using any of the ECG leads or even an oximeter. However, continuing on the results of [8], [15], we can use time index sequence of the RPE as the data source. This sequence showed promising results in the previous work of this research group. Any or all of these leads can be used to further find better distinguishing features.

In this research, we used conventional STDFT to generate the pictorial clips. Future work can concentrate on wavelets for image generation. Wavelets have an advantage over STDFT in that they do not require periodicity of the time sequence or for it to be stationary. Also, the images generated from wavelets have greater time resolution and are computationally more efficient [14].

Different encoding schemes can be used for the matrices from above, including different gray-level quantization methods and generalized color encoding. This extends to the use of different possible feature extraction and optimum feature selection algorithms.

Due to representing the data in image format, a multitude of image processing and pattern recognition methods can be applied. The detection and classification algorithm can take different direction, including the different forms of neural network or fuzzy logic systems. Neural networks can be used using different node functions, single or multiple hidden layers. Fuzzy logic systems (FLS) are flexible and application-oriented, where the design can optimally benefit from the spread of the extracted features in type-I or type-II non-singleton rule-based fuzzy system.

The method presented in this investigation can be widely applied to different kinds of electrical, mechanical, chemical or photo signals in which offline classification of the signal source or nature is required. We believe that this generalization can be extended to different industrial and medical application, and is one of the main contributions of this investigation.

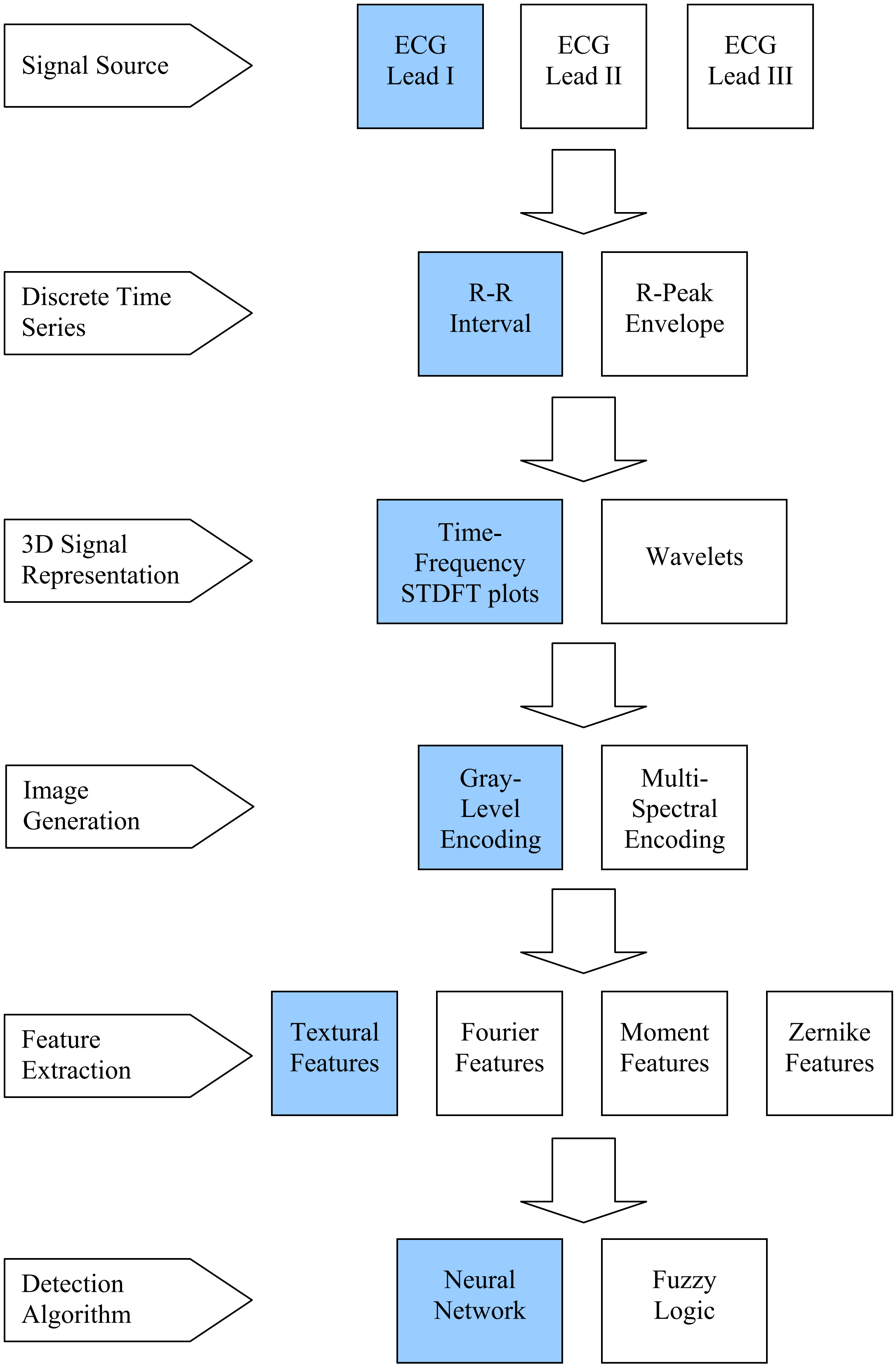

Figure 5.1 Block diagram of possible future directions at each level of research.

REFERENCES

[1] American Academy of Sleep Medicine Task Force, "Sleep-Related Breathing Disorders in Adults: Recommendations for Syndrome Definition and Measurement Techniques in Clinical Research", *Sleep*, 22 (5), pp. 667-689, 1999.

[2] P. de Chazel, C. Heneghan, E. Sheridan, R. Reilly, P. Nolan, and M. O'Malley, "Automated Processing of Single-Lead Electrocardiodiagram for the Detection of Obstructive Sleep Apnea", *IEEE Transactions on Biomedical Engineering*, 50 (6), pp. 686-696, 2003.

[3] T. Young, P.E. Peppard, and D.J. Gottlieb, "Epidemiology of Obstructive Sleep Apnea: A Population Health Perspective", *American Journal of Respiratory Critical Care Medicine*, 165, pp. 1217-1239, 2002.

[4] R. Wolk, T. Kara, V.K. Somers, "Sleep-Disordered Breathing and Cardiovascular Disease", *Circulation*, 108, pp. 9-12, 2003.

[5] T. Young, M. Palta, J. Dempsey, J. Skatrud, S. Weber, and S. Badr, "The Occurrence of Sleep-Disordered Breathing among Middle Adults", *New England Journal of Medicine*, 328, pp. 1230-1235, 1993.

[6] T.D. Bradly and J.S. Floras, "Sleep Apnea: Implications in Cardiovascular and Cerebrovascular Disease", *Lung Biology in Health and Disease*, 146, Mercel Dekker Inc., 2000.

[7] C. Guilleminault, S. Connolly, R. Winkle, K. Lelvin, and A. Tilkian, "Cyclical Variation of the Hear Rate in Sleep Apnea Syndrom", *Lancet*, 21, pp. 126-131, 1984.

[8] S. Vijendra, "An Investigation in the Detection of Sleep-Disordered Breathing using the Electrocardiogram", Master's Thesis, Dept of Biomedical Engineering, University of Texas at Arlington, Arlington, TX, USA, 2003.

[9] F. Pinciroli, R. Rossi, L. Vergani, P. Carnevali, S. Mantero, and O. Parigi, "Detection of Electrical Axis Variation for Extraction of Respiratory Information", *Computers in Cardiology*, 2, pp. 499-502, 1985.

[10] G.B. Moody, R.G. Mark, A. Zoccola, and S. Mantero, "Derivation of Respiratory Signals from Multi-Lead ECGs", Computers in Cardiology, 2, pp. 113-116, 1985.

[11] K. Behbehani, S. Vijendra, J.R. Burk, and E.A. Lucas, "An Investigation of the Mean Electrical Axis Angle and Respiration during Sleep", *Proceedings of the Annual IEEE EMBS-BMES Conference on Biomedical Engineering*, Houston, October 2002.

[12] A. Travaglini, C. Lamberti, J. DeBie, and M. Ferri, "Respiratory Signal Derived from Eight-Lead ECG", *Computers in Cardiology*, 25, pp. 65-68, 1998.

[13] T. Penzel, J. McNames, P. de Chazel, B. Raymond, A. Murray, and G. Moody, "Systematic Comparison of Different Algorithms for Apnoea Detection based on Electrocardiogram Recordings", *Medical and Biological Engineering and Computing*, 40, pp. 402-407, 2002.

[14] M.F. Hilton, R.A. Bates, K.R. Godfrey, M.J. Chappell, and R.M. Cayton, "Evaluation of Frequency and Time-Frequency Spectral Analysis of Heart Rate Variability as a Diagnostic Marker of Sleep Apnoea Syndrome", *Medical and Biological Engineering and Computing*, 37, pp. 760-769, 1999.

[15] R. Sanjee, “An Technique for the Detection of Cheyne Stokes Breathing and Obstructive Sleep Apnea using Electrocardiogram”, Master’s Thesis, Dept of Biomedical Engineering, University of Texas at Arlington, Arlington, TX, USA, 2005.

[16] D. Benitez, P.A. Gaydecki, A. Zaidi, and A.P. Fitzpatrick, “The Use of the Hilbert Transform in ECG Signal Analysis”, *Computers in Biology and Medicine*, 31, pp. 399-406, 2001.

[17] A.V. Oppenheim, *Discrete-Time Signal Processing*, Prentice Hall, 2nd Ed, 1999.

[18] R.M. Haralick, K. Shanmugam, and I. Dinstein, “Textrural Features for Image Classification”, *IEEE Transactions on Systems, Man, and Cybernetics*, 3 (6), pp.610-619, 1972.

[19] D. Clausi and Y. Zhao, “Grey Level Co-occurrence Integrated Algorithm (GLCIA): a Superior Computational Method to Rapidly Determine Co-occurrence Probability Texture Features”, *Computers & Geosciences*, 29, pp. 837-850, 2003.

[20] D. Clausi and Y. Zhao, “An Advanced Computational Method to Determine Co-occurrence Probability Texture Features”, *IEEE International Geosciences and Remote Sensing Symposium*, 4, pp. 2453-2455, 2002.

[21] D. Clausi and M Ed Jernigan, “An Fast Method to Determine Co-occurrence Texture Features”, *IEEE Transactions Geosciences and Remote Sensing*, 36, pp. 298-300, 1998.

[22] F. Argenti, L. Alparone, and G. Benelli, “Fast Algorithms for Texture Analysis Using Co-occurrence Matrices”, *IEE Proceedings for Radar and Signal Processing*, 137, pp. 443-448, 1990.

[23] D.I. Morales, M. Moctezuma, and F. Parmiggiani, “Urban and Non Urban Area Classification by Texture Characteretics and Data Fusion”, *IEEE International Geosciences and Remote Sensing Symposium*, 6, pp. 3504-3506, 2003.

[24] G. Lohmann, “Co-occurrence-based Analysis and Synthesis of Textures”, *IEEE International Conference on Pattern Recognition*, 1, pp. 449-453, 1994.

[25] D.G. Vince, K.J. Dixon, R.M. Cothren, and J.F. Cornhill, “Comparision of Texture Analysis Methods for the Characterization of the Coronary Plaques in Intravascular Ultrasound Images”, *Computerized Medical Imaging and Graphics*, 24, pp. 221-229, 2000.

[26] R. King and M. Amadasun, “Textural Features Corresponding to Textural Properties”, *IEEE Transactions on Systems, Man, and Cybernetics*, 19, pp. 1264-1274, 1989.

[27] S.W.-C. Lam, “Texture Feature Extraction Using Gray Level Gradient Based Co-occurrence Matrices”, *IEEE International Conference on Systems, Man, and Cybernetics,* 1, pp. 267-271, 1996.

[28] M. Hauta-Kasari, J. Parkkinen, T. Jaaskelainen, and R. Lenz, “Generalized Co-occurrence Matrix for Multispectral Texture Analysis”, *IEEE International Conference on Pattern Recognition*, 2, pp. 785-789, 1996.

[29] D. Clausi and R. Jobanputra “Texture Analysis Using Gaussian Weighed Gray Level Co-Occurrence Probabilities,” *First Canadian Conference on Computer and Robot Vision*, pp. 51-57, 2004.

[30] A. Baraldi and F. Parmiggiani, “An Investigation of the Textural Characteristics Associated with Gray Level Co-occurrence Matrix Statistical Parameters”, *IEEE Transactions on Geosciences and Remote Sensing*, 33, pp. 293-304, 1995.

[31] J. Shuttleworth, A. Todman, R. Naguib, B. Newman, “Colour Texture Analysis using Co-occurrence Matrices for Classification of Colon Cancer Images”, *IEEE Canadian Conference on Electrical and Computer Engineering*, 2, pp.1134-1139, 2002.

[32] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Boston: Academic Press, 1990.

[33] J. Li, M. T. Manry, P. L. Narasimha, and C. Yu, “Feature Selection Using a Piecewise Linear Network,” accepted by *IEEE Trans. on Neural Networks.*

[34] J. Li, J. Yao, R.M. Summers, N. Petrick, M. Manry, and A.K. Hara, “An Efficient Feature Selection Algorithm for Computer-Aided Polyp Detection,” special issue of the *International Journal on Artificial Intelligence Tools* (IJAIT) to be published in 2006.

[35] M.F. Redondo and C.H. Espinosa, “Comparison Among Feature Selection Methods Based on Trained Networks,” *Neural Network for Signal Processing – Proceedings for the IEEE Workshop*, pp. 205-214, 1999.

[36] A.K. Jain, R.P.W Duin, and J. Mao, “Statistical Pattern Recognition: A Review,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22 pp. 4-37, January 2000.

[37] D.Y. Tsai and Y. Lee, “Fuzzy-Reasoning-Based Diagnosis Scheme for Automated Classification of Heart Disease from Ultrasonic Images”, *IEEE International Conference on Fuzzy Systems*, 1, pp. 188-191, 2001.

[38] P. de Chazel, C. Heneghan, and E. Sheridan, “Apparatus for Detection Sleep Apnea Electrocardiogram Signals”, *US Patent Application Publication*, Pub No. US 2003/0055348 A1, March 20, 2003.

[39] P. Pudil, J. Novovicova, and J. Kittler, "Floating Search Method for Feature Selection," *Pattern Recognition Letters*, 15 pp. 1119-1125, November 1994.

BIOGRAPHICAL INFORMATION

Mohammad Ahmad Al-Abed was born in a small village 20 miles north of Amman, Jordan, on April 22nd, 1979, to a Royal Jordanian Air force pilot father and a math and sciences teacher mother. He is number three in a family of six: two sisters and three brothers. He received his high school education at the Jubilee School for Gifted and Talented Students in 1997, Amman, finishing in the top 0.1-percentile. He has graduated 3rd in his class from the Jordan University of Science and Technology with a B.Sc. in Communications and Electronics Engineering degree, 2003. His research interests include discrete signal processing, instrumentations and electronic systems design, optical imaging and detection, neural networks and fuzzy logic systems, and physiology and life sciences. He is currently working on improving algorithms for the detection of sleep disorders. He is planning to pursue a PhD degree in bioengineering.