A New Method to Detect Obstructive Sleep Apnea Using Fuzzy Classification of Time-Frequency Plots of the Heart Rate Variability

Mohammad Al-Abed[1], Khosrow Behbehani[1], John R. Burk[2], Edgar A. Lucas[2], and Michael Manry[3]
[1]Bioengineering Department, University of Texas at Arlington and University of Texas Southwestern Medical Center at Dallas. [2] Sleep Consultants, Inc. Fort Worth Texas. [3]Electrical Engineering Department, University of Texas at Arlington.

Abstract— **This paper presents a new method of analyzing time frequency plots of heart rate variability to detect sleep disordered breathing from nocturnal ECG. Data is collected from 12 normal subjects (7 males, 5 females; age 46 ± 9.38 years, AHI 3.75 ± 3.11) and 14 apneic subjects (8 males, 6 females; age 50.28 ± 9.60 years; AHI 31.21 ± 23.89). The proposed algorithm uses textural features extracted from normalized gray-level co-occurrence matrices (NGLCM) of images generated by short-time discrete Fourier transform (STDFT) of the HRV. Thirty selected features extracted from 10 different NGLCMs representing four characteristically different gray-level images are used as inputs to 10 Fuzzy Logic Systems (FLS) Classifiers. Each FLS is trained and their outputs are combined using a weighed majority rule method. The mean training detection sensitivity, specificity and accuracy are 86.87%, 71.72%, and 79.29%, respectively. The mean testing detection sensitivity, specificity and accuracy are 83.22%, 68.54%, and 75.88%, respectively.**

Keywords—**ECG, Heart Rate Variability, Sleep Disordered Breathing, Time-Frequency Plots, Co-occurrence Matrix, Textural Features, Fuzzy Logic Systems**

I. Introduction

Sleep-Disordered Breathing (SDB) is estimated to have a prevalence of 5% in middle-aged population [1]. 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 quality 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.

Previous studies [2] have showed increase interest in screening for OSA episodes using ECG overnight recordings. Different algorithms were developed to extract reliable markers from ECG signals including HRV, RWA, and power spectral analysis of particular frequency bands. Thus far, HRV showed the most promising results. One study has pointed to combining different temporal statistical features from HRV and RWA to improve the detection rate [3]. Different groups tried to qualitatively describe time-frequency plots of HRV during an international challenge for the detection of OSA using ECG signals [2].

In this research paper, a quantitative study of the HRV time-frequency plots is described. Image processing and classification schemes are applied to these plots to extract better representing markers and improved detection scheme.

II. Methodology

The data used in this research was previously collected and processed by Behbehani, et al [4]. Twelve (12) normal subjects' volunteers were recruited for the study. None of them had any known history of SDB. Fourteen (14) patients, previously diagnosed with OSA at an accredited sleep lab were also recruited. Appropriate approval from the Institutional Review Board for human subject testing was obtained for post-processing of existing data. During the mentioned study, a certified sleep expert, blind to the study, scored the sleep stages and SDB events for all subjects, and apnea/hypopnea events were qualitatively described using the apnea/hypopnea index (AHI). Table I summarizes the subject population demographics.

Data Source: Different physiological markers were reordered during the NPSG study, including ECG, EEG, SaO_2, and airflow, at a 1000 sample/sec rate. Lead ECG-I was used to extract the HRV time series. The data was clipped to 15-minutes (900 seconds) clips based on the recommendations of Hilton [5]. There were 198 total usable 900-second clips; 92 normal clips (NOR) and 106 event clips (OSA).

The HRV was interpolated using cubic spline, and resampled at 10 Hz rate. This resulted in a 9000-point series.

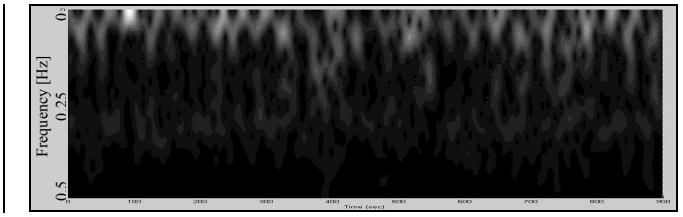
TABLE I
SUBJECT DEMOGRAPHICS OF THE NORMAL AND OSA GROUPS INCLUDING THE APNEA/HYPOPNEA INDEX

Subject group (N)	Number of males/females	Age (mean ± std)	BMI (mean ± std)	AHI (mean ± std)
NOR (12)	7/5	46 ± 9.38	25.34 ± 3.86	3.75 ± 3.11
OSA (14)	8/6	50.28 ± 9.60	31.33 ± 6.29	31.21 ± 23.89

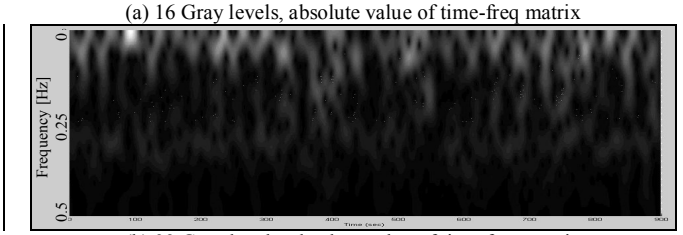
Plot Construction: STDFT is performed on the HRV equally-spaced discrete time series. A Hanning window is used with N=300, resulting in 9000 point in time and 1250 in frequency (0-5 Hz). Since the frequency range that was shown most promising is 0-0.5Hz [4], only the lower 125 rows were included, and the columns were reduced by one

fourth to decrease the processing time, without greatly affecting the resolution.

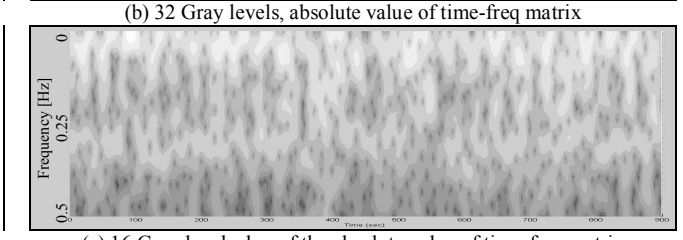
The resultant 125x2250 point complex matrix was then translated to 4 characteristically different gray-level images. Figure 1 shows a graphical illustration of these images representing the same 900-second clip. Figure 1(a) shows Image-1, which was generated by finding the magnitude of the complex STDFT matrix, and quantizing it to 16 equally spaced gray levels (darker shades represents lower power intensities and brighter shades represent higher power intensities). Image-2 in Figure 1(b) was generated using the same method for Image-1, but with 32 gray levels. Figure 1(c) shows Image-3, which was found by finding the log of the magnitude of the images before quantizing it to 16 gray levels. And Image-4 in Figure 4(d) was found by finding the histogram of the magnitude values, then allocating them to un-equal size quantization bins, so that each gray shade would be equally represented in the image.



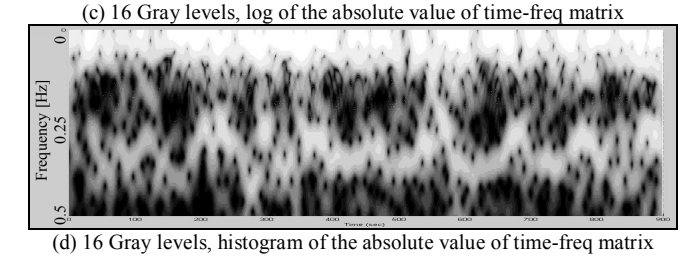
(a) 16 Gray levels, absolute value of time-freq matrix



(b) 32 Gray levels, absolute value of time-freq matrix



(c) 16 Gray levels, log of the absolute value of time-freq matrix



(d) 16 Gray levels, histogram of the absolute value of time-freq matrix

Figure 1: Four characteristically different gray level images describing the time-frequency plots of a 15-minute HRV time sequence of an OSA patient. The horizontal axis represents the time scale from 0 to 900 seconds.

Co-occurrence Matrices and Textural Features: Normalized gray level co-occurrence matrices (NGLCM) [6] were used to quantitatively analyze the resultant images. Each image from each clip was cut into 18 125×125 segments, and an NGLCM was calculated for each of these segments. Textural features were found from these NGLCMs, and were averaged out for each clip. Two NGLCMs were calculated for 64×125 size segments. This was an effort to create images that account for the lower frequency range only.

Ten NGLCMs were extracted from these four images. From Image-1, four NGLCMs were found:

a. GLCM-1 (125×125, $d = 5$, $\theta = 90º$)
b. GLCM-2 (125×125, $d = 1$, $\theta = 90º$)
c. GLCM-3 (64×125, $d = 5$, $\theta = 90º$)
d. GLCM-4 (125×125, $d = 5$, $\theta = 0º$)

From Image-2, three NGLCMs were found:

a. NGLCM-5 (125×125, $d = 5$, $\theta = 90º$)
b. NGLCM-6 (125×125, $d = 3$, $\theta = 90º$)
c. NGLCM-7 (125×125, $d = 1$, $\theta = 90º$)

From Image-3, two NGLCMs were found:

a. NGLCM-8 (125×125, $d = 5$, $\theta = 90º$)
b. NGLCM-9 (64×125, $d = 5$, $\theta = 90º$)

From Image-4, one NGLCM was found:

NGLCM-10 (125×125, $d = 5$, $\theta = 90º$)

Nine textural features [7], [8] are chosen and defined herein. These Textural features are used as inputs to the detector. For all the subsequent measures, the following definitions are applicable:

$M(i,j)$ is the i, j element of the gray level co-occurrence matrix.. Since M is symmetric, $M(i,j) = M(j,i)$.

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

1. Entropy (ENT): $ENT = -\sum_{i}^{N_g}\sum_{j}^{N_g} M(i,j) \cdot \log(M(i,j))$

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

3. Contrast: $CON = \sum_{i}^{N_g}\sum_{j}^{N_g} M(i,j) \cdot (i-j)^2$

4. Correlation (COR): $COR = \sum_{i}^{N_g}\sum_{j}^{N_g} \frac{M(i,j) \cdot (i-\mu) \cdot (j-\mu)}{\sigma^2}$

5. Dissimilarity (DIS): $DIS = \sum_{i}^{N_g}\sum_{j}^{N_g} M(i,j) \cdot |i-j|$

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

7. Inverse Difference Moment (IDM): $IDM = \sum_{i}^{N_g}\sum_{j}^{N_g} \frac{M(i,j)}{1+(i-j)^2}$

8. Variance (VAR): $VAR = \sum_{i}^{N_g}\sum_{j}^{N_g} M(i,j)\cdot(i-\mu)^2$

9. Inverse Recursivity (INR)

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

Feature Selection: 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 suited 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 deviation-to-mean ratio.

Fuzzy Logic System: Fuzzy logic systems have showed promising results in previous studies as classifier system with overlapping input values [9]. This prominent fuzzy logic property comes from the use of membership functions (MFs) that assign a membership degree between 0 and 1 to the range of possible inputs. This makes the fuzzy logic more promising for certain application. Sigmoid MFs was chosen for the present study. The sigmoidal function can be written as $\mu_x(x) = 1/1-e^{-a(c-x)}$

The sigmoid functions have the following advantages: 1) simple function with 2 parameters *a* and *c*, 2) ease of training due to its continuity, and 3) ability to use for two classes better than triangular or trapezoidal MFs.

A singleton fuzzy system is made up of input fuzzifier, rule-fed inference engine, and a defuzzifier. The input is made up from the selected feature values [9]. Each value is given two membership degrees (MDs); NOR and OSA. The MDs are allocated using the sigmoidal MFs discussed above. The choice of initial c_L and c_H values are the means of lower and higher NOR and OSA means, respectively. a_L and a_H values are initialized to be 1. These values are then trained to find best fitting MFs that give optimum classification results, as described in [9] algorithm.

To formulate fuzzy rules, suppose for a given NGLCM-X, the selected features had the following means:

	COR	IND	IDM
NOR	0.835	0.923	0.997
OSA	0.891	0.947	0.993

Then there are 2^3 = 8 fuzzy rules. Two easy to formulate rules are in the following form:

IF *COR* is *low* AND *IND* is *low* AND *IDM* is *high* THEN *clip is NOR*

IF *COR* is *high* AND *IND* is *high* AND *IDM* is *low* THEN *clip is OSA*

But, it is cumbersome to formulate the other six rules, since there is no clear subjective knowledge to describe the relationship between the input statistics and the desired output that can be extracted from the given data. However, since all the selected features are exclusively disjoint, in which each feature independently describes the image, rule reduction can be applied here, and these rules can be broken into simpler, easier to formulate rules [10]. The following are the 2×3=6 rules used in the inference engine of the system that classifies the NGLCM-X:

IF *COR* is *low* THEN *clip is NOR*
IF *COR* is *high* THEN *clip is OSA*
IF *IND* is *low* THEN *clip is NOR*
IF *IND* is *high* THEN *clip is OSA*
IF *IDM* is *high* THEN *clip is NOR*
IF *IDM* is *low* THEN *clip is OSA*

For the deffuzification part, height defuzzification is used [9]. A value of "1" was assigned to OSA clip diagnosis and "-1" to NOR clip diagnosis. A crisp decision boundary was set at "0".

Training: Each of the FLS was recursively trained, as described in [9].

Clip Classification and Weighed Majority Rule: With the ten FLS results optimized, the results were combined using weighed majority rule to find a classification for each clip. Table II shows an example of the results for 6 randomly chosen clips, their classification, and final decision.

Sensitivity, Specificity, and Accuracy: The performance of the system is described by.

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

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

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

where OSA_c is the number of correctly detected OSA clips and NOR_c is the number of correctly detected NOR clips [11].

III. Results

Time-Frequency Plots: Figure 1 shows the time-frequency plots we obtained from one clip from one OSA subject (900 second, 0-0.5 Hz).

Feature Selection: Using the statistical significant difference between NOR and OSA extracted textural features as a feature selection method, it was found that for all 10

NGLCMs, COR, IND, and IDM showed the most significant difference and lowest spread to mean ratio.

TABLE II
SAMPLE RESULT OF THE OUTPUT OF THE FLSS USED FOR SIX TEST CLIPS. THE NOR CLIPS ARE MARKED AS (-1) AND OSA MARKED AS (1)

Clip #	1	2	3	4	5	6
FLS-1	-1	1	1	1	-1	1
FLS-2	-1	1	1	1	1	1
FLS-3	-1	1	1	1	1	-1
FLS-4	1	1	1	1	1	1
FLS-5	-1	-1	-1	1	1	-1
FLS-6	-1	1	1	1	1	-1
FLS-7	-1	1	1	1	1	1
FLS-8	-1	1	-1	1	-1	-1
FLS-9	-1	1	1	1	1	-1
FLS-10	-1	1	1	1	1	-1
Majority Rule	-1	1	1	1	1	-1
Clip Diagnosis	-1	-1	1	1	1	-1

These features were the inputs to the FLS layer. Each FLS was trained with the corresponding 3 feature from each NGLCM. So, FLS-1 was trained using features COR_1, IND_1, and IDM_1, and so on through FLS-10 being trained using COR_{10}, IND_{10}, and IDM_{10}.

Training and Testing Results: All 92 NOR clips and 106 OSA clips were randomly divided into training set and testing set: 132 clips for training and 66 for testing. The means of the training set was used as initial values for c_L and c_H. The test set was classified and the output of the FLSs was then acquired and combined using weighed majority rule, as shown in Table II. This was repeated 100 times, following the randomized clip selection, to obtain the average training and testing detection results.

Detection Results: Table III shows summary of the training and testing detection results.

TABLE III
SENSITIVITY, SPECIFICITY, AND ACCURACY OF THE COMBINED FUZZY LOGIC CLASSIFIERS

	Training	Testing
Sensitivity	86.87%	83.22%
Specificity	71.72%	68.54%
Accuracy	79.29%	75.88%

IV. DISCUSSION

By studying the images in Figure 1, it is noticed that the power distribution within the clip is not uniform, and is focused in the low frequency range (0-0.1Hz). The band of 0.2-0.3Hz shows some power variations, as previously described by [4], but not as high as the previous band.

The four different gray-level encoding schemes used in this research have allowed for different distribution of the gray-scale shading that represent each point in the plot. This was translated into differences in the extracted textural features.

The features selection method used is a 1st degree statistical analysis and may not result in optimum feature selection. However, these features are sufficient to produce acceptable results.

Future investigation for improvement of the detection rates will include the exploration of alternative techniques such as neural networks.

V. CONCLUSIONS

The combination of the HRV series time and frequency domains in one plot and quantitatively analyze the resultant images has been shown to be a suitable classification scheme. Image processing and pattern recognition schemes and classifiers show significant promise in improving the detection rate as an alternative to conventional temporal or spectral methods.

REFERENCES

[1]. 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.M.

[2]. 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.

[3]. 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.

[4]. S. Vijendra, K. Behbehani, E.A. Lucas, J.R. Burk, D.N. Burli, and D.H. Dao, "The Use of R-wave morphology in the Detection of Sleep-Disordered Breathing using the Electrocardiogram – A Comparison between Leads," *Conference Proceedings, 26th Annual International Conference of the IEEE EMBS*, 2, pp.3881 - 3884 2004.

[5]. 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.

[6]. 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.

[7]. 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.

[8]. 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.

[9]. J.M. Mendel, "Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions", Prentice Hall, NJ, 2001.

[10]. Setnes and R. Babuska, "Rule Base Reduction: Some Comments on the Use of Orthogonal Transforms", *IEEE Transactions on Systems, Man & Cybernetics, Part C: Applications & Reviews*, 31, pp. 199-206, 2001.

[11]. S.R. Suhas, "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.