

Identification of Patients with Congestive Heart Failure by Recognition of Sub-Bands Spectral Patterns

Abdulnasir Hossen, and Bader Al-Ghunaimi

Abstract—A new simple technique based on recognition of power spectral densities patterns of decomposed sub-bands of R-R interval (RRI) data for identification of patients with Congestive Heart Failure (CHF), is investigated. This method uses a soft-decision algorithm for estimating the PSD of the decomposed sub-bands by a probability measure of energy contents in those sub-bands. Both trial and test data used in this work are drawn from MIT databases. Two standard patterns of the base-2 logarithmic values of the reciprocal of the probability measure of the approximated PSD of CHF patients and normal subjects are derived by averaging all corresponding values of all sub-bands of 12 CHF data and 12 normal subjects in the trial set. The computed pattern of each data under test is then compared band-by-band with both standard patterns of CHF and normal subjects to find the closest pattern by finding the minimum sum of squares of distances from the standard pattern. The new technique results in an overall identification accuracy of 90% on both data sets.

Keywords—Congestive Heart Failure, Pattern Recognition, Sub-Band Decomposition, Soft-Decision, Power Spectral Density.

I. INTRODUCTION

HEART failure is a common condition that usually develops slowly as the heart muscle weakens and needs to work harder to keep blood flowing through the body. Heart failure develops following injury to the heart such as the damage caused by heart attack, long-term high blood pressure, or an abnormality of one of the heart valves [1]. Heart failure is often not recognized until a more advanced stage of heart failure, commonly referred to as congestive heart failure (CHF), in which fluid may leak into the lungs, feet, and in some cases the liver or abdominal cavity. Physicians often assess the stage of heart failure according to the New York Heart Association (NYHA) functional classification system. This system relates symptoms to everyday activities and the patient's quality of life. According to this system, heart failure is classified into 4 classes [1]:

1. Class I (Mild), Symptoms with more than ordinary activity.
2. Class II (Mild), Symptoms with ordinary activity.
3. Class III (Moderate), Symptoms with minimal activity.
4. Class IV (Severe), Symptoms at rest

Physicians often order a number of tests when exploring a possible diagnosis of heart failure. The most important of these tests is an echocardiogram. Echocardiography is a noninvasive, entirely safe test that uses ultrasound to image the heart as it is beating [2]. Cardiac ultrasounds provide information about:

- Accurate indications of valve function.
- The amount of blood flow through the heart's chambers.
- The location of the failure, whether has occurred on the left side, the right side, or both.

Physicians use information from the echocardiography for calculating the ejection fraction (EF), which is the percent of the blood pumped out during each heartbeat. EF is important for determining the severity of heart failure. People with a healthy heart usually have an EF of 50 percent or greater. Most people with heart failure, but not all, have an EF of 40 percent or less. The echocardiogram is the most accurate diagnostic test but an expensive one. So there is a need to a noninvasive simple test that helps in determining patients who most likely do not need an echocardiogram test.

II. HEART RATE VARIABILITY

Heart rate variability (HRV) is referred to as the beat-to-beat variation in heart rate. Instantaneous heart rate is measured as the time in seconds between peaks of two consecutive R waves of the ECG signal. This time is referred to as the RRI. The beat-to-beat fluctuations in the rhythm of the heart provide us with an indirect measure of the heart health, as defined by the degree of balance in sympathetic and vagus nerve activity [3]. HRV provides a window to observe the hearts ability to respond to normal regulatory impulses that affect its rhythm. HRV analysis serves as a marker for cardiovascular disease because cardiac dysfunction is often manifested by systematic changes in the variability of the R-R interval sequence relative to that of normal controls. The variation of heart rate accompanies the variation of several physiological activities such as breathing, thermoregulation and blood pressure changes [4].

Several HRV abnormalities have been described in patients with congestive heart failure (CHF). It has been shown that patients with HF have decreased HRV, particularly those who are at risk of cardiac death [5].

Analysis of the HRV using many signal processing techniques can help to detect abnormalities, by manipulating RRI data either in the time-domain or in the frequency domain [6].

The frequency spectrum of the RRI data is divided into three main bands:

- The very low-frequency band (VLF): $f \in (0.0033 - 0.04) \text{ Hz}$.
- The low-frequency band (LF): $f \in (0.04 - 0.15) \text{ Hz}$.
- The high-frequency band (HF): $f \in (0.15 - 0.4) \text{ Hz}$.

In [7-8], a method of screening of obstructive sleep apnea is implemented using the estimated PSD by a soft-decision algorithm on decomposed sub-bands. The same idea is applied in [9] for estimating the spectral analysis using a wavelet analysis.

In this work a pattern recognition technique based on PSD estimation of decomposed sub-bands of RRI data is implemented for the purpose of identification of patients having CHF from normal controls.

III. RRI DATA GROUPS

Two different groups of data are used for trial and test. Both groups were drawn from MIT databases [10].

A. Trial Data

The CHF records and Normal Sinus Rhythm (NSR) records were drawn from MIT database [9], which includes 12 records from normal subjects (age 29-64 years) and 12 records from patients with CHF (age 22-71 years).

B. Test Data

This group contains 17 CHF and 53 NSR recordings that are used to test the performance of the classification algorithm. The CHF subjects are selected from a larger set that contains 29 long-term recordings. The selected 17 records have CHF with NYHA class 3 and 4, and the remaining 12 records have NYHA class 1 and 2, and therefore have been excluded from the study. The subjects for the selected records are 8 men, aged 39 to 68, and 2 women aged 38 and 59; gender is unknown for the 7 remaining records, but aged between 35 and 64 years. The NSR data of this group contains 53 long-term (about 24 hours) RRI records. The subjects are 30 men, aged 28.5 to 76, and 24 women aged 58 to 73.

The RRI data, for both groups, are generated from the annotation file, which is associated with each record, using WFDB software [11]. In the beginning, all records are truncated to a length of 75821 RR intervals, since the shortest record has such a length.

IV. SOFT DECISION TECHNIQUE

A. Sub-band-DFT

In the well-known radix-2 decimation-in-time FFT algorithm, the time sequence $x(n)$ is decomposed into even and odd subsequences, in order to arrange the computation into smaller transforms. As a result of this arrangement, the DFT complexity is reduced by calculating smaller size DFTs. The approach to the SB-DFT is decomposition, too. However,

the decomposition in SB-DFT is done into low- and high-frequency bands rather than into even and odd samples, giving a physical interpretation to the decomposition process. Fig.1 shows a two-band decomposition for calculating the DFT of N -points [12-13].

In Fig. 1 $a(n)$ and $b(n)$ are the low-pass and high-pass filtered versions of $x(n)$, with $g(n)$ and $h(n)$ their factor-2 down-sampled versions, respectively:

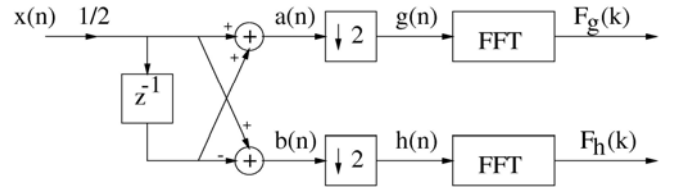


Fig. 1 DFT calculation using a two-band decomposition

$$\begin{aligned} g(n) &= \frac{1}{2}[x(2n) + x(2n+1)] \\ h(n) &= \frac{1}{2}[x(2n) - x(2n+1)] \end{aligned} \quad (1)$$

The full-band size- N DFT $X(k)$ can be obtained by multiplying $F_g(k)$ and $F_h(k)$ (the size- $N/2$ DFTs of $g(n)$ and $h(n)$ respectively) with the suitable correction factors as:

$$X(k) = (1 + W_N^k)F_g(k) + (1 - W_N^k)F_h(k) \quad (2)$$

Eq.(2) is approximated for calculating only the low-pass band:

$$X(k) \approx (1 + W_N^k)F_g(k), \quad k \in (0, 1, \dots, N/4 - 1) \quad (3)$$

B. Adaptive Sub-band-DFT

The decomposition process in Fig. 1 can be repeated m times to get $M = 2^m$ subbands. If there is no information about the energy distribution of the input sequence, a band-selection [14] algorithm can be used. This method depends on the energy comparison between the low- and high-frequency subsequences after the down sampling in Fig. 1:

$$B = \sum_{n=0}^{\frac{N}{2}-1} (g(n))^2 - (h(n))^2 \quad (4)$$

According to the sign of B , the decision is taken: If B is positive, the low-frequency band is considered, and if B is negative, the high-frequency band is considered. Since only the sign of B is important, Eq.4 could be simplified to

$$\text{sgn}(B) = \text{sgn} \sum_{n=0}^{\frac{N}{2}-1} |g(n)| - |h(n)| \quad (5)$$

C. Estimation of Power Spectral Density

The decomposition stages in Fig. 1 are computed with all branches up to a certain stage m to obtain 2^m subbands.

All estimator results up to stage m are stored, and a probability measure is assigned to each path (i.e., frequency band) to bear the primary information.

If $J(L)$ is the assigned probability of the input signal being primarily low-pass, the number $J(H) = 1 - J(L)$ is the probability that the signal is primarily high-pass. One simple way to make the probability assignments is to use the ratio of the number of positive comparisons between $|g(n)|$ and $|h(n)|$ in Eq.(5) to the total number of comparisons for a given stage.

At the following stage, the resulting estimate can be interpreted as the conditional probability of the new input sequence containing primarily low (high) frequency components, given that the previous branch was predominantly low (high)-pass. Using this reasoning and laws of probability, the assignments for the probability measure of the resulting sub-bands should be made equal to the product of the previous branch probability and the conditional probability estimated at a given stage. Fig. 2 shows this step of probability assignment of 8 sub-bands.

The probabilities $P(B_i)$ derived from the estimator outputs, where i is the index of the band, may be interpreted themselves as a coarse measurement of the PSD: The higher the probability value of any band, the higher is its power-spectral content

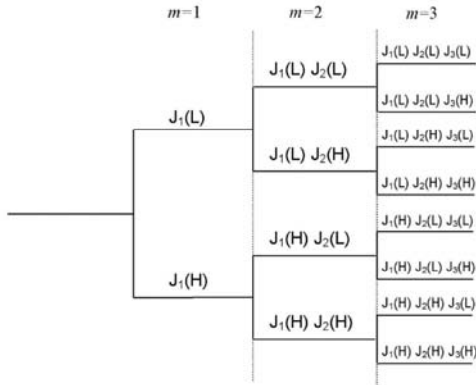


Fig. 2 Probability Assignments of 8 Sub-bands

V. PATTERN RECOGNITION ANALYSIS

The new pattern recognition technique is implemented using the steps:

1. The probability measures are computed for each data of the trial set up to stage m to obtain 2^m probabilities $P(B_i)$, where i is the index of the band. Each band covers $(0.5/2^m)$ Hz of the RRI spectrum.

2. The logarithmic based-2 values of the reciprocal of the probabilities can be obtained as:

$$I(B_i) = \log_2 \left(\frac{1}{P(B_i)} \right) \quad (6)$$

A staircase approximation of the values of $I(B_i)$ can be plotted for all bands.

3. The average plot can be also found by averaging all values of corresponding bands for all 12 CHF data and normal data in the trial set to obtain two standard plots for CHF and normal as in Fig. 3 for 8 sub-bands.

4. The $I(B_i)$ values for each data under test are found for 8 sub-bands. A classification factor CF is then determined as:

$$CF = \sum_{i=1}^8 (I(B_i) - I_H(B_i))^2 - \sum_{i=1}^8 (I(B_i) - I_N(B_i))^2 \quad (7)$$

Where $I_H(B_i)$ and $I_N(B_i)$ are the standard patterns for CHF and normal respectively.

Depending on the sign of CF, the algorithm can decide whether the data is CHF or normal. The data with a negative CF is considered as normal and the data with a positive CF is a CHF data.

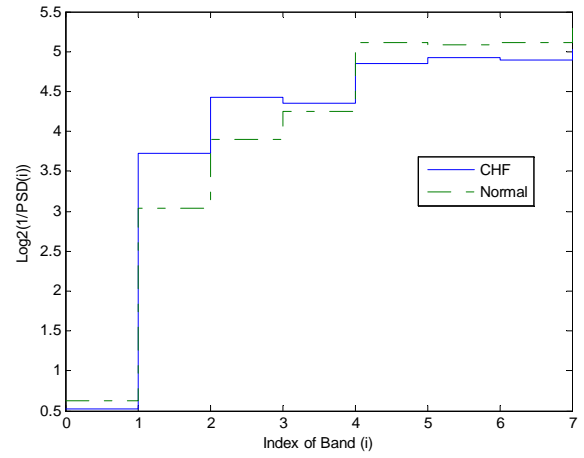


Fig. 3 Standard Patterns of CHF and Normal with 8 Sub-bands

VI. CLASSIFICATION PERFORMANCE

The performance of a classifier is evaluated by three main metrics: sensitivity, specificity, and accuracy as defined below [15]:

TABLE I
CONFUSION MATRIX

	Normal	CHF
Normal	TN	FP
CHF	FN	TP

$$\text{Sensitivity (\%)} = \frac{TP * 100}{TP + FN} \quad (8)$$

$$\text{Specificity (\%)} = \frac{TN * 100}{TN + FP} \quad (9)$$

$$\text{Accuracy (\%)} = \frac{(TN + TP) * 100}{T} \quad (10)$$

Where TP, TN, FN, FP are defined as in Table I, and T is the total number of data under test.

Sensitivity represents the ability of a classifier to detect the positive cases, e.g. CHF. Specificity indicates the ability of a classifier to detect negative cases, e.g. normal subjects. Accuracy represents the overall performance of a classifier. It indicates the percentage of correctly classified positives and negative cases from the total cases [15].

VII. RESULTS

Fig. 4 and Fig. 5 show the values of CF for all CHF and normal data using 8 Sub-bands for trial and test data respectively. The algorithm succeeds to classify CHF and normal trial data with a specificity of 100% (12/12) and a sensitivity of 91.63% (11/12) and an accuracy of 95.83% (23/24), as shown in (Fig. 4). The new simple technique results in a sensitivity of 76.5% (13/17) and a specificity of 90.5% (48/53) with an overall identification accuracy of 87.14% (61/70) by implementation on test data set as shown in (Fig. 5). The total classification efficiency on both data sets is about 90%.

VIII. CONCLUSION

A soft decision algorithm of PSD estimation of decomposed sub-bands is implemented to obtain two standard plots of CHF and normal subjects for the purpose of classification between them in a technique based on pattern recognition.

The number of sub-bands used is 8 and the decomposition is done using simple Hadamard filters. An accuracy of 90% is obtained on both data sets.

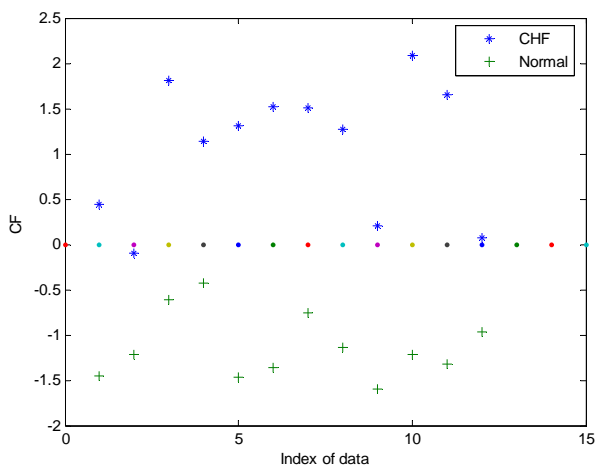


Fig. 4 Classification Factor Results of Trial Data with 8 Sub-bands

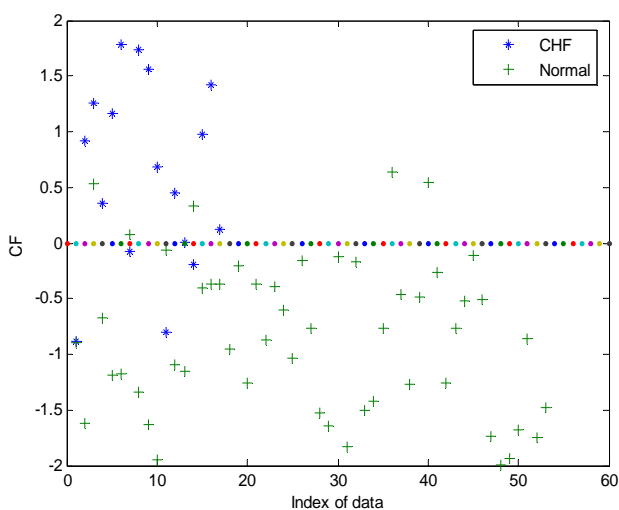


Fig. 5 Classification Factor Results of Test Data with 8 Sub-bands

REFERENCES

- [1] Heart Failure Society of America, available at: <http://www.abouthf.org>.
- [2] Congestive Heart Failure, University of Maryland Medicine, available at: <http://www.umm.edu>
- [3] Dr. Robert Nolan and Heart Rate Variability, available at: <http://www.bio-medical.com/news-display.cfm>.
- [4] Task Force of the European Society of Cardiology and the North American Society of pacing and Electrophysiology, Heart Rate Variability, standards of measurements, physiological interpretation, and clinical use, *Circulation* 93, pp. 1043-1065, 1996.
- [5] Ponikowski, P., Anker S.D., Chau T.P. et al; Depressed heart rate variability as an independent predictor of death in chronic congestive heart failure secondary to ischemic or idiopathic dilated cardiomyopathy. *AM J. Cardio.* 1997, 79, 1645-1650.
- [6] Teich M.C., Lowen S.B., Jost B.M., Vibe-Rheymer K., and Heneghan C., " Heart rate Variability: Measures and Models", *Nonlinear Biomedical Signal Processing, Vol. II, Dynamic Analysis and Modeling*, edited by M. Akay (IEEE Press, New York, 2001), Chapter 6, pp.159-213.
- [7] Hossen, A., Al-Ghunaimi, B., and Hassan, M. O., "A new simple algorithm for heart rate variability analysis in patients with obstructive sleep apnea and normal controls", *International Journal of Bioelectromagnetism*, 5, No. 1, pp. 238--239, 2003.
- [8] Hossen A., Al-Ghunaimi B., and Hassan M. O., "Subband Decomposition Soft Decision Algorithm for Heart rate Variability Analysis in Patients with OSA and Normal Controls. To be appeared in *Signal Processing Journal*.
- [9] Hossen, A.; "Power Spectral Density Estimation Via Wavelet Decomposition", *Electronics Letters*, Vol. 40, No. 17, 2004.
- [10] Physionet, an NIH--NCRR Research Resource, data files online, available at: <http://www.physionet.org/physiobank/database>.
- [11] Physionet, an NIH--NCRR Research Resource, WFDB Software package, available at: <http://www.physionet.org/physiotools/wfdb.shtml>.
- [12] Shentov; O. Mitra, S.; Heute, U.; Hossen, A.: "Sub-band DFT- Part I: Definition, Interpretation and Extensions", *Signal Processing*, Vol. 41, pp. 261-277, 1995.
- [13] Hossen, A.; Heute, U.; Shentov, O.; Mitra, S.: "Sub-band DFT- Part II: Accuracy, Complexity and Applications", *Signal Processing*, Vol. 41, pp. 261-277, 1995.
- [14] Hossen, A.; Heute, U.: "Fully Adaptive Evaluation of SB--DFT", *Proceedings of IEEE Int. Symp. on Circuits and Systems*, Chicago, Illinois, 1993.
- [15] R. M. Rangayyan, *Biomedical Signal Analysis: A Case-Study Approach*, IEEE Press, pp.466--472, 2001.

Abdulnasir Hossen is an Associate Professor at Sultan Qaboos University in Oman since 2005. He got his Ph.D. in Digital Signal Processing from the Ruhr-University at Bochum, Germany in 1994 and his Post. Doct. in Digital Signal Processing in 1997 from University of Kiel, Germany. Dr. Hossen is a member of IEEE and DGBMT. He has been awarded the DAAD research scholarship award at Kiel University in 2000 and 2003 and 2006. His interests are in Digital Signal Processing, Fast Algorithms, Sub-band and Wavelet Transforms, Biomedical Signal Processing, Image Compression, Speaker Recognition, Pattern Recognition and Neural Networks.

Bader Al-Ghunaimi received the B. Sc degree in computer engineering in 1995 from Sultan Qaboos University, Muscat, Oman. He received the M.Sc in electrical engineering (with concentration in biomedical digital signal processing) at the same university in 2003. His thesis subject was the screening of obstructive sleep apnea by using statistical signal characterization and subband decomposition. His principal interest includes digital signal processing, programming and control. He is working in Ministry of defense in Oman.