35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Adaptive Neuro-Fuzzy Inference System for Acoustic Analysis of 4-Channel Phonocardiograms Using Empirical Mode Decomposition

Miguel A. Becerra[1], Diana A. Orrego[2] and Edilson Delgado-Trejos[3]

***Abstract*— The heart's mechanical activity can be appraised by auscultation recordings, taken from the 4-Standard Auscultation Areas (4-SAA), one for each cardiac valve, as there are invisible murmurs when a single area is examined. This paper presents an effective approach for cardiac murmur detection based on adaptive neuro-fuzzy inference systems (ANFIS) over acoustic representations derived from Empirical Mode Decomposition (EMD) and Hilbert-Huang Transform (HHT) of 4-channel phonocardiograms (4-PCG). The 4-PCG database belongs to the National University of Colombia. Mel-Frequency Cepstral Coefficients (MFCC) and statistical moments of HHT were estimated on the combination of different intrinsic mode functions (IMFs). A fuzzy-rough feature selection (FRFS) was applied in order to reduce complexity. An ANFIS network was implemented on the feature space, randomly initialized, adjusted using heuristic rules and trained using a hybrid learning algorithm made up by least squares and gradient descent. Global classification for 4-SAA was around 98.9% with satisfactory sensitivity and specificity, using a 50-fold cross-validation procedure (70/30 split). The representation capability of the EMD technique applied to 4-PCG and the neuro-fuzzy inference of acoustic features offered a high performance to detect cardiac murmurs.**

I. INTRODUCTION

Cardiac auscultation is a clinical procedure where the state of the cardiac valves are evaluated by analyzing the heartbeat sounds. A graphic representation of these sounds is known as phonocardiogram or phonocardiographic (PCG) signal, which is inexpensive and non-invasive [1]. Cardiac murmurs are generated when the blood flow becomes turbulent near damaged valves [2]. In order to develop a better diagnosis, the heart sounds should be analyzed from 4-Standard Auscultation Areas (4-SAA), one for each cardiac valve, as cardiac murmurs do not generally appear in every single area of auscultation [3].

The morphological changes in the PCG signal when a murmur appears have been studied using energy and temporal measurements [4], but cardiac murmurs have a nonstationary nature and exhibit sudden frequency changes and transients [1]. The analysis of fractal features and the optimization of the embedding parameters have been investigated in order to improve the training and classification stages [5], [6], although the increment in processing time becomes a big problem for real-time applications. Wavelet approaches have also been proposed because of the time-frequency disturbances caused by cardiac murmurs [7]. However, other decomposition methods, such as Empirical Mode Decomposition (EMD) and Hilbert-Huang Transform (HHT), express the signal in a better way as an expansion of signal-dependent basis functions, via an iterative procedure called sifting [8]. For example, the foetal heart sounds could be extracted from a recorded single channel abdominal PCG, using an EMD approach proposed in [9]. In another way, the acoustic analysis by Mel-Frequency Cepstral Coefficients (MFCC) [10] has been proposed to analyze the acoustical disturbances caused by heart murmurs, but these procedures are very sensitive to artifacts or noises frequently involved in the acquisition stage [1]. For this reason, the combination between MFCC and statistical moments of HHT on appropriate combination of different intrinsic mode functions (IMFs) would be suitable. Additionally, the learning capability of Adaptive Neuro-Fuzzy Inference Systems (ANFIS) can improve the classification, as the fuzzy logic and neural networks are combined into a single technique in learning nonlinearities [11]. The use of Linear Discriminant Analysis (LDA) and ANFIS to detect heart valve disorders was studied in [12] with promising results. In [13], an ANFIS model for evaluation of foetal health status using PCG signals was implemented effectively and provided high accuracy for antepartum antenatal care. Likewise, a biomedical-based decision support system was developed in [14] for the heart sound signal classification, where the reduced features of three types of heartbeat sounds were used as input patterns of an ANFIS classifier. However, all these studies have been developed using a single auscultation signal, and fail when a murmur is missing or attenuated in the standard single derivation. In a previous study [15], the representation capability of EMD applied to 4-PCG and stochastic analysis performed by an ergodic HMM of acoustic features offered a high performance, however this stochastic classifier was demonstrated to be highly dependent on the signal representation and parameter initialization for the model optimization.

In this study, a classification approach based on a MFCC-HHT-ANFIS hybrid technique applied on the combination of different IMFs of 4-PCG joined with a relevance analysis to reduce the number of features is presented, in order to provide an objective and accurate mechanism for more reliable heart murmur diagnosis.

This work was supported by the Research Center of the Instituto Tecnologico Metropolitano –ITM under grant PM11117, and by the research project number 0403 of the Institucion Universitaria Salazar Herrera in Medellin Colombia.

[1]Miguel A. Becerra is with GEA Research Group of the Institucion Universitaria Salazar Herrera, Carrera 70 No. 52-49, Medellin, Colombia. migb2b@gmail.com

[2]Diana A. Orrego is with SINERGIA Research Group of the Instituto Tecnologico Metropolitano ITM, Calle 73 No. 76A-354, Medellin, Colombia. dianaorrego@itm.edu.co

[3]Edilson Delgado-Trejos is with MIRP Lab of the Research Center of the Instituto Tecnologico Metropolitano ITM, Calle 73 No. 76A-354, Medellin, Colombia. edilsondt@gmail.com

II. Materials and methods

A. Database

The database is made up of 143 de-identified adult subjects, who gave their formal consent, and underwent a medical examination with the approval of the ethical committee. The valve lesion severity was evaluated by cardiologists according to a clinical routine. 55 patients were labeled as normal, while 88 had evidence of cardiac murmurs (aortic stenosis, mitral regurgitation, etc). From each patient, 8 recordings were recorded according to the four standard auscultation areas (4-SAA), i.e., mitral, tricuspid, aortic and pulmonic areas, in the phase of post-expiratory and post-inspiratory apnea. Each recording lasts 8 s and was obtained with the patient standing in dorsal decubitus position. The signals were acquired at 44.1 kHz with 16-bits per sample with an electronic stethoscope (WelchAllyn® Meditron model). Finally, 400 individual beats were chosen, 180 normal and 180 with evidence of cardiac murmur. The individual beats picked out were the best from each cardiac sound signal, according to a visual and audible inspection by a cardiologists.

B. Theoretical background

1) Hilbert-Huang Transform (HHT) and Empirical Mode Decomposition (EMD): Instantaneous frequency and its magnitude of heart sound signals can be extracted by HHT, which is used to adaptively decompose non-stationary and nonlinear signals and extract the instantaneous frequency. In general, HHT consists of two steps: Empirical Mode Decomposition (EMD) and Hilbert transform. EMD is used to adaptively decompose the signal into a series of intrinsic mode functions (IMFs). Hilbert transform is then carried out to acquire the instantaneous frequency and amplitude of each IMF and constitute the time-frequency-energy distribution in the Hilbert-Huang spectrum of the signal [16]. The EMD method, reported in [15] and [16], adaptively decomposes a multi-component signal $x(t)$ into a number L of Intrinsic Mode Functions (IMFs), $h^{(i)}(t)$, $1 \leq i \leq L$,

$$x(t) = \sum_{i=1}^{L} h^{(i)}(t) + d(t) \tag{1}$$

where $d(t)$ is a remainder which is a non zero-mean slowly varying function with only few extrema. Each one of the IMFs, say the ith one $h^{(i)}(t)$, is estimated with the aid of an iterative process, called sifting, applied to the residual multi-component signal [8].

2) Mel-Frequency Cepstral Coefficients (MFCC): Psychophysical studies have shown that human perception of the frequency content of audio sounds does not follow a linear scale but as a Mel-warped frequency, which spaces linearly for low-frequency contents and logarithmical at high frequencies [1]. So, MFCC are a family of parameters that are estimated as [17]:

$$c[p] = \sum_{m=0}^{M-1} X_F[m] \cos\left(\pi p(m-0.5)/M\right),\ 0 \leq p \leq M \tag{2}$$

where $X_F[m] = \ln\left(\sum_{i=0}^{N-1} |X[i]|^2 H_m[i]\right)$. Here, $X[i]$ is the Fourier transform of an input random sequence $x[n]$ and $H_m[i]$ is a triangular band-pass filter with central frequency in $f[m]$. Thus, in order to simplify the signal spectrum without any significant loss of data, a set of M triangular band-pass filters must be used, which are nonuniform in the original spectrum and uniformly distributed at the Mel-warped spectrum. Each filter is multiplied by the spectrum so that only a single value of magnitude is returned per filter.

3) Adaptive Neuro-Fuzzy Inference System (ANFIS): It is a simple data learning technique that uses Fuzzy Logic to transform given inputs into a desired output through highly interconnected Neural Network processing elements and information connections, which are weighted to map the numerical inputs into an output [11]. For simplicity, it is assumed in the ANFIS architecture, two fuzzy IF-THEN rules based on a first order Sugeno model [18]:

$\mathrm{Rule}_{(1)}$: IF x is A_1 AND y is B_1, THEN $f_1 = p_1x + q_1y + r_1$

$\mathrm{Rule}_{(2)}$: IF x is A_2 AND y is B_2, THEN $f_2 = p_2x + q_2y + r_2$

Where, x and y are inputs, A_i and B_i are fuzzy sets, f_i are outputs within the fuzzy region specified by the fuzzy rule, and p_i, q_i, r_i are design parameters which are adjusted during the training process. In Fig. 1, these two rules are

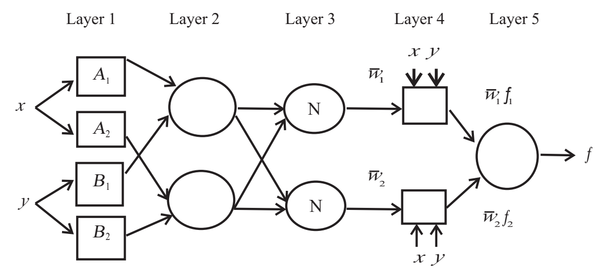

Fig. 1. ANFIS architecture

implemented by means of a five-layer ANFIS architecture, where Π is an AND operator to fuzzify the inputs, the N-nodes indicate a normalization to the firing strengths from the previous layer. In the 4th layer, the nodes are adaptive and the output of each node is the product of the normalized firing strength with a first order polynomial (for a first order Sugeno model). The overall output f of the model is given by one single fixed Σ-node, i.e., $f = \sum_i \bar{w} f_i$.

4) Fuzzy-Rough Feature Selection (FRFS): Fuzzy-rough sets (FRS) encapsulate the related but distinct concepts of vagueness (for fuzzy sets) and indiscernibility (for rough sets), both of which occur as a result of uncertainty in knowledge. FRFS provides a means by which discrete or real-valued noisy data (or a mixture of both) can be effectively reduced without the need for user-supplied information, and as such can be applied to regression as well as classification datasets [19]. The fuzzy lower and upper approximations can be defined using a Υ-transitive fuzzy similarity relation to approximate a fuzzy equivalence class X [20]:

$$\mu_{\underline{R_P}X}(x) = \inf_{y} \Psi\left(\mu_{R_P}(x,y), \mu_X(y)\right) \tag{3}$$

$$\mu_{\overline{R_P}X}(x) = \sup_{y} \Upsilon(\mu_{R_P}(x,y), \mu_X(y)) \tag{4}$$

where Ψ is a fuzzy implication and Υ a t-norm. The membership of an object $x \in \mathbb{U}$, belonging to the fuzzy positive region can be defined by [21]

$$\mu_{POS_P(\mathbb{D})}(x) = \sup_{X \in \mathbb{U}/\mathbb{D}} \mu_{\underline{R_P}X}(x) \tag{5}$$

Where $\mathbb{D}$ is a set of decision features. Thus, the fuzzy-rough degree of dependency of $\mathbb{D}$ on the attribute subset P can be defined

$$\gamma'_P(\mathbb{D}) = \frac{\sum_x \mu_{POS_P(\mathbb{D})}(x)}{|\mathbb{U}|} \tag{6}$$

C. Proposed procedure

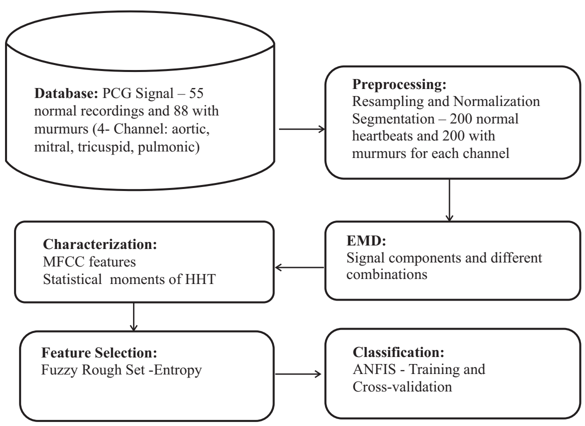

Fig. 2. General proposed procedure

According to Fig. 2, the 4-SAA PCG recordings were resampled to 4410 Hz by applying a FIR low-pass anti-aliasing filter. Next, the signals were normalized in [-1, 1] and 200 normal heartbeats and 200 with murmurs were segmented for each channel. Features derived from the acoustic and time-frequency analysis were estimated from different signals obtained from the combination of different IMFs, which were estimated using the sifting algorithm, with the following parameters: resolution 40 dB, residual energy 40 dB and the gradient step size 0.01. Particularly, a Mel-scaled filter bank was used to calculate the Mel-warped spectrum, so the first 8 and 12 MFCC were estimated using 24 Hamming shaped filters and a sliding hamming window (50% overlap) on the combination of different IMFs derived from the whole beats. Additionally, the first 4 statistical moments in function of the instant frequency and instant amplitude obtained by HHT were also considered, using the same combinations of EMD components. The representation space was normalized in order to improve the classification performance. With the aim of obtaining a minimal feature subset, a FRFS algorithm with entropy estimation was implemented (following the procedure presented in [19]), where the neighbor distance tolerance and the inclusion rate were adjusted in 0.05 and 0.5 respectively, taking into account an evaluation measure based on the entropy estimation. Finally, an ANFIS network was implemented for the classification of heartbeat sounds, i.e., normal or murmur, where the reduced feature set was used as the input vector and the training parameters are described in Table I. This classification stage was carried out by a 50-fold cross-validation procedure using a 70/30 split, where consistency and representation capability of the feature space were analyzed.

TABLE I
ANFIS TRAINING PARAMETERS

Number of layers	5
FIS type	Takagi-Sugeno
Number of system input	10
Partitioning type	Subtractive clustering
Radius of influence sub-clustering	0.5
Learning algorithm type	Least squares & gradient descent
Number of fuzzy rules	70
Training error	0.01

III. RESULTS AND DISCUSSION

Table II shows the classification accuracy of a cardiac murmur detection system for 4-PCG signals based on FRFS-ANFIS, where sets of 8 and 12 MFCC were tested over two sets of constructions based on IMFs (EMD components), which were considered after making several different combinations: IMF-C1= $\{3,5,7\}$ and IMF-C2= $\{1,3,5,7\}$. Additionally, the first 4 statistical moments calculated over the Hilbert spectrum were included in these feature sets. These results show that MFCC 9, 10, 11 and 12 of the IMF 1 contain relevant acoustical information related to heart valve damages. The FRFS algorithm reduced the features from 160 to 12. Table III presents statistical measures of the ANFIS performance for MFCC-HHT features over EMD components IMF-C2, considering each auscultation area. Finally, this classification approach is compared with other ANFIS-based classifiers (see Table IV), where a greater performance is evidenced, although this was on other databases.

TABLE II
FRFS-ANFIS USING HHT-EMD REPRESENTATIONS

	8-MFCC IMF-C1	8-MFCC IMF-C2	12-MFCC IMF-C1	12-MFCC IMF-C2
Accuracy (%)	76.5±3.2	95.1±3.5	84.7±2.9	**98.9**±**1.1**

TABLE III
STATISTICAL PERFORMANCE BY AUSCULTATION AREA

Channel	Accuracy (%)	Sensitivity (%)	Specificity (%)
Aortic	99.3±0.7	98.6±1.4	100±0.0
Mitral	98.9±1.1	99.1±0.9	98.8±1.2
Tricuspid	98.7±1.3	98.6±1.4	98.0±2.0
Pulmonic	98.8±1.2	98.3±1.7	99.0±1.0
Mean	98.9	98.7	99.0

IV. CONCLUSION

The interpretation of heart sounds depends on the physician's ability and experience. Such limitations can be reduced

TABLE IV

COMPARISON WITH OTHER APPROACHES

Approach	Accuracy (%)
Wavelets-LDA/ANFIS [12]	Se (95.9) - Sp (94)
Wavelets-PCA/ANFIS [18]	94.5
Wavelets/Entropy-ANFIS [14]	94.3
EMD/MFCC/HHT-HMM [15]	98.7
EMD/MFCC/HHT-FRFS/ANFIS (this work)	**98.9**

by developing biomedical-based decision support systems. In this study, an objective and accurate ANFIS mechanism of 4-PCG classification was obtained, for more reliable cardiac murmur detection, in terms of sensitivity and specificity. The EMD representation enhanced the acoustical content associated with cardiac murmurs and reduced the acoustical components related to normal heart sounds or noises included in the acquisition stage. The heart's mechanical activity could be well-represented by acoustic features derived from MFCC and statistical moments of HHT obtained from the combination of different IMFs of 4-SAA PCG signals. These features substantially outperformed the traditional MFCC applied directly on the signal. The relevance analysis based on FRFS allowed a reduced feature space to be found with low complexity and high representation capability, which is related to a high learning capability. According to Table II, it is noticeable that MFCC 9, 10, 11 and 12 of the IMF 1 contain relevant acoustical information in terms of the detection of heart murmurs. Regarding the ANFIS results, a high performance in terms of accuracy, sensitivity and specificity was achieved, and the fuzzy rules played an important role in the detection procedure. However, the operation parameters and the type of fuzzy functions were not optimized by a criterion function. In this sense, the automatic tuning of ANFIS parameters is proposed as future work.

ACKNOWLEDGEMENTS

We would like to thank the Research Center and the Biomedical Engineering Program of the Instituto Tecnologico Metropolitano ITM, as well as the Institucion Universitaria Salazar Herrera in Medellin Colombia.

REFERENCES

[1] E. Delgado-Trejos, A. F. Quiceno, J. I. Godino, M. Blanco, and G. Castellanos, "Digital auscultation analysis for heart murmur detection," *Annals of Biomedical Engineering*, vol. 37, no. 2, p. 337–353, 2009.

[2] A. G. Tilkian and M. B. Conover, *Understanding Heart Sounds and Murmurs: With An Introduction to Lung Sounds*, 4th ed. Philadelphia: W. B. Saunders, 2001.

[3] B. Karnath and W. Thornton, "Auscultation of the heart," *Hospital Physician*, pp. 39–43, September 2002.

[4] R. M. Rangayyan and R. J. Lehner, "Phonocardiogram signal analysis: a review," *Critical Reviews in Biomedical Engineering*, vol. 15, no. 3, pp. 211–236, 1987.

[5] E. Delgado-Trejos, J. A. Jaramillo, A. F. Quiceno, and G. Castellanos, "Parameter tuning associated with nonlinear dynamics techniques for the detection of cardiac murmurs by using genetic algorithms," in *Computers in Cardiology*, vol. 34, 2007, pp. 403–406.

[6] D. Kumar, P. Carvalho, R. Couceiro, M. Antunes, R. P. Paiva, and J. Henriques, "Heart murmur classification using complexity signatures," in *20th International Conference on Pattern Recognition*, 2010, pp. 2564–2567.

[7] B. Ergen, Y. Tatar, and H. O. Gulcur, "Time-frequency analysis of phonocardiogram signals using wavelet transform: a comparative study," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 15, no. 4, pp. 371–381, 2012.

[8] Y. Kopsinis and S. McLaughlin, "Development of EMD-based denoising methods inspired by wavelet thresholding," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1351–1362, 2009.

[9] A. D. Warbhe, R. V. Dharaskar, and B. Kalambhe, "A single channel phonocardiograph processing using EMD, SVD, and EFICA," in *3rd International Conference on Emerging Trends in Engineering and Technology*, 2010, pp. 578–581.

[10] J. Vepa, "Classification of heart murmurs using cepstral features and support vector machines," in *31st Annual International Conference of the IEEE/EMBS (EMBC)*, 2009, pp. 2539–2542.

[11] J.-S. R. Jang, "ANFIS: Adaptive-network-based fuzzy inference system," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665–684, 1993.

[12] A. Sengur, "An expert system based on linear discriminant analysis and adaptive neuro-fuzzy inference system to diagnosis heart valve diseases," *Expert Systems with Applications*, vol. 35, no. 1-2, pp. 214–222, 2008.

[13] V. S. Chourasia, A. K. Tiwari, and R. Gangopadhyay, "Adaptive neuro-fuzzy inference system for antepartum antenatal care using phonocardiography," *International Journal of Biomedical Engineering and Technology*, vol. 8, no. 4, pp. 357–374, 2012.

[14] H. Uguz, "Adaptive neuro-fuzzy inference system for diagnosis of the heart valve diseases using wavelet transform with entropy," *Neural Computing and Applications*, vol. 21, no. 7, pp. 1617–1628, 2012.

[15] M. A. Becerra, D. A. Orrego, C. Mejia, and E. Delgado-Trejos, "Stochastic analysis and classification of 4-area cardiac auscultation signals using empirical mode decomposition and acoustic features," in *Computing in Cardiology*, vol. 39, September 2012, pp. 529–532.

[16] Y.-L. Tseng, P.-Y. Ko, and F.-S. Jaw, "Detection of the third and fourth heart sounds using Hilbert-Huang transform," *BioMedical Engineering OnLine*, vol. 11, p. 8, February 2012.

[17] A. Acero and H. W. Hon, *Spoken Language Processing: A Guide to Theory, Algorithm and System Development*. NJ, USA: Prentice Hall, 2001.

[18] E. Avci and I. Turkoglu, "An intelligent diagnosis system based on principle component analysis and ANFIS for the heart valve diseases," *Expert Systems with Applications*, vol. 36, no. 2 part 2, pp. 2873–2878, 2009.

[19] D. A. Orrego, M. A. Becerra, and E. Delgado-Trejos, "Dimensionality reduction based on fuzzy rough sets oriented to ischemia detection," in *34th Annual International Conference of the IEEE/EMBS (EMBC)*, 2012, pp. 5282–5285.

[20] A. M. Radzikowska and E. E. Kerre, "A comparative study of fuzzy rough sets," *Fuzzy Sets and Systems*, vol. 126, no. 2, pp. 137–155, 2002.

[21] N. Mac-Parthalin and R. Jensen, "Measures for unsupervised fuzzy-rough feature selection," in *9th International Conference on Intelligent Systems Design and Applications (ISDA)*, 2009, pp. 560–565.