33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Complexity-based analysis for the detection of heart murmurs

J.A. Gómez-García*, J.D Martínez-Vargas*, G. Castellanos-Dominguez*

***Abstract*— While a healthy human heart produce a rhythmic pattern of sounds, some heart disorder induce deviations perceived as abnormal sounds called murmurs. Despite many murmurs can be considered harmless, other constitute the first basis of a heart disorder. In this sense, a correct diagnosis remains essential; however, due to the subjectivity on using human ear to make diagnosis, automatic detection systems appear as useful tools for helping medical specialists on improving diagnosis accuracy. Complexity analysis has become one important tool for the study of physiological signals, because tracking sudden alteration on the inherent complexity on biological processes might be useful for detecting pathologies. The present paper presents a complexity-based analysis methodology, which uses regularity features for the detection of heart murmurs, including Approximate Entropy, Sample Entropy, Gaussian Kernel Approximate Entropy, and Fuzzy Entropy. The results show the high discriminative power, up to 90%, of the Gaussian Kernel Approximate Entropy and Fuzzy Entropy for the proposed labour.**

I. INTRODUCTION

Functioning of human heart produces perceptible sounds that can be sensed by tools as stethoscopes on a process called *auscultation*. Perceived sounds come in pairs, S_1 and S_2, each one being generated by a different physiological phenomena; thus S_1, is produced by the closing of mitral and tricuspid valve, while S_2, is produced by the closing of aortic and pulmonar valve. Time intervals between sounds are also defined: systole for the one between S_1 and S_2, and diastole for the one between S_2 and S_1. It is expected that in a healthy heart both systole and diastole remain silent; the contrary, some perceived turbulence called *murmur*, might be an indicator of an abnormal condition. Despite some murmurs are harmless, others may be related to a serious cardiac disease for which an accurate diagnosis remains as an essential step for medical treatment. However, due to overlapping of murmurs with the cardiac beat, those can not be easily separated by human ear [1]. In this respect, automatic murmur detection system, which utilize signal processing techniques, might be a valuable tool for specialist for making more accurate diagnosis.

A characteristic of physiologic systems is their complexity, which arises from the interaction of a vast number of structural units and regulatory feedback loops enabling the organism to adapt to the stresses of everyday life [2]. Furthermore, it has been hypothesized that dynamics of a healthy physiological system produce an apparently irregular and highly complex type of variability, whereas disease or aging is often associated with more regularity and less complexity [3]. Thus, by quantifying the complexity of physiologic signals in health and disease potentially important applications have arose, with respect to evaluating both dynamical models of biologic control systems and bedside diagnostics [4]. As a way to measure complexity of a system, nonlinear dynamic analysis (NDA) tools have been used. In this respect, complexity features such as dimensions of the reconstructed attractor of a system, the rate of divergence of trajectories traced by reconstructed states, among others have been typically utilized. However, NDA features require the signal dynamics to be deterministic; an assumption that is not entirely valid due to stochastic components produced by effects such as noise. As a result, entropy-based features which do not need to assume determinism or stochasticity for its calculation, are important on pathology detection labours. In this category, complexity estimators which quantify the regularity of a time series, have provided successful results on characterizing heart dynamics [5], [6].

The present work presents a methodology for the detection of heart murmurs using complexity analysis. Regularity features as Approximate Entropy (ApEn), Sample Entropy (SampEn), Gaussian Kernel Approximate Entropy (GapEn) and Fuzzy Entropy (FuzzyEn) are used for characterization of phonocardiographic signals (PCG); a k-nn classifier, and a Gaussian Kernel Support Vector Machine (SVM) are used for classification; and ROC curves are used for results assessment. Results superior to 90% demonstrate the discriminative capability of FuzzyEn and GapEn for automatic murmur detection labours.

This paper is organized as follows: Section II describes the methodology used on this work. Section III presents the experimental setup and results. Finally, Section IV presents the conclusions of this work.

II. MATERIALS AND METHODS

The automatic murmur detection system proposed in this work is shown on Fig 1, while the most important sections are explained in the next subsections.

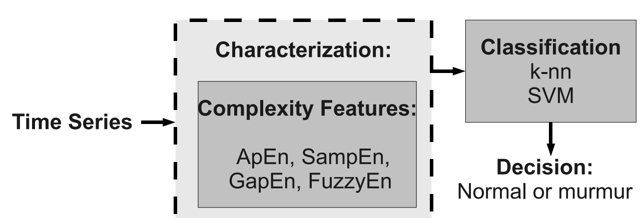

Fig. 1. Automatic murmur detection system

* Grupo de Control y Procesamiento Digital de Señales. Universidad Nacional de Colombia, sede Manizales. Km 7. Vía al Magdalena. Manizales, Colombia. (e-mail: {jorgomezg; jmartinezv; cgcastellanosd }@unal.edu.co

A. Characterization

NDA makes use of a process, called *embedding*, which maps a given time series $s = \{x_1, x_2, ...x_n\}$ into an m-dimensional space called *phase space*. The reconstructed state vectors of the system are obtained by using (1).

$$\mathbf{x}(t) = [s(t), s(t-\tau), s(t-2\tau), ..., s(t-(m-1)\tau)] \quad (1)$$

where $\mathbf{x}(.)$ is called *state*; m is the embedding dimension, related to the least number of needed coordinates to embed space vectors into state space; and τ is the time lag, related to the spread of the reconstructed states as they traces out trajectories on the state space. In order to calculate m and τ the false nearest neighbours algorithm, and the mutual information of the time series are used respectively [7].

The collection of states as they evolve, originate a geometric object called *attractor* where typically NDA features are extracted.

1) Regularity features: Are quantifiers of the complexity of a system, based on entropy estimation.

For a single random variable, entropy quantifies its uncertainty, being measured by means of the Shannon entropy $H(X)$. For a discrete random variable X, whose set of possible values is given by $\Theta = \{x_1, x_2, .., x_n\}$, and its probability mass function is given by $p(x_i) = Pr\{X = x_i\}$; $H(X)$ will be defined as:

$$H(X) = -\sum_{x_i \in \Theta} p(x_i) \log p(x_i)$$

Extending the latter concept to phase space reconstructions, an information production rate, called Kolmogorov-Sinai entropy (H_{KS}), can be found.

Let the m-dimensional phase space be divided on N hypercubes of radius r and volume r^m, and let n measurements be done, spaced τ between them. Let also be $PC = p(k_1, k_2, ..., k_N)$, the joint probability of the system of being on hypercube k_1 at time $t = \tau$, hypercube k_2 at time $t = 2\tau$ and hypercube at $t = N\tau$. H_{KS} can then be defined as in (2) [8]:

$$H_{KS} = -\lim_{\substack{\tau \to 0 \\ r \to 0 \\ n \to \infty}} \frac{1}{n\tau} \sum_{k_1, ..., k_N} PC \log PC \qquad (2)$$

Because of the limits on which H_{KS} is established, it is desirable to find another way of quantifying the entropy of a time series without heavy computational load, the need of large amount of information, while being robust to the presence of low amplitude noise [9].

In this sense, regularity features have been developed to somehow estimate entropy of a system while making front to the implicit problems of H_{KS}. One of first regularity features, ApEn, was proposed by Pincus [10]; examining similar epochs onto a time series, and quantifying the average negative logarithm of the conditional probability that two sequences that are similar for m points remain similar (within a tolerance r), at the next point [11].

ApEn is defined as in (3), where r is a tolerance measure, and $C_i^m(r)$ as in (4).

$$\text{ApEn} = \phi^m(r) - \phi^{m+1}(r) \qquad (3)$$

$$\phi^m(r) = \frac{1}{n-m+1} \sum_{i=1}^{n-m+1} \log C_i^m(r)$$

$$C(r) = \lim_{n \to \infty} \frac{1}{n^2} \sum_{i,j=1}^{n} \Theta(r - \|\mathbf{x}(i) - \mathbf{x}(j)\|) \qquad (4)$$

Despite ApEn has been successfully used on biosignal characterization, it suffers from a phenomena called *self-matching* which makes it a biased estimator. This happens because when comparing embedding vectors, looking for similar epochs, self-comparisons are also made. To overcome that bias, Richman [11] proposed SampEn, which is defined as in (5).

$$\text{SampEn} = -\log\left(\frac{A^m(r)}{A^{m+1}(r)}\right) \qquad (5)$$

with A as in (4) but without self-matching.

According to [12], SampEn and ApEn had problems on validity and precision due to their formulation on the non continuous Heaviside function. A first attempt to overcome that problem was proposed by [12] with GapEn, which replaces the Heaviside function by a Gaussian Kernel function. GapEn is defined as in (3), but replacing (4) by (6).

$$C_i^m(r) = \frac{\sum_{j=1, j\neq i}^{n-m+1} \exp\left(-\frac{(\|\mathbf{x}(i), \mathbf{x}(j)\|)^2}{10r^2}\right)}{n-m} \qquad (6)$$

A second attempt was proposed by [13] with FuzzyEn, which replaces Heaviside function by a Fuzzy membership function. FuzzyEn is defined as in (7).

$$\text{FuzzyEn} = \ln \phi^m(ne, r) - \ln \phi^{m+1}(ne, r) \qquad (7)$$

$$\phi^m(ne, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m-1} \sum_{j=1, j\neq i}^{N-m} D_{ij}^m\right)$$

where D_{ij}^m is a fuzzy membership fuzzy as in (8), ne is a parameter determining shape of the fuzzy membership function [14], d_{ij}^m is a distance function, and $u_0(i)$ removes baseline of space state vectors.

$$D_{ij}^m = \exp(-(d_{ij}^m)^{ne}/r) \qquad (8)$$

$$d_{ij}^m = \max_{k \in (0, m-1)} |\mathbf{x}(i+k) - u_0(i) - (\mathbf{x}(j+k) - u_0(j))|$$

$$u_0(i) = m^{-1} \sum_{j=0}^{m-1} \mathbf{x}(i+j) \qquad (9)$$

B. Classification

In order to make decisions, a 9 neighbours k-nn classifier and a Gaussian kernel SVM were used. The number of neighbours on the k-nn classifier were chosen as they produced the highest classification rates among tested neighbours, from 1-11. SVM Gaussian Kernel classifier was chosen for its generalization capability, while its parameters were tuned in a cross-validation scheme for obtaining highest accuracy rates.

III. EXPERIMENTAL SETUP

A. Subjects

The database used in this study is made up of 148 de-identified adult subjects. An electronic stethoscope was used to acquire the heart sounds simultaneously with a standard 3-lead ECG. Signals were digitized at 44.1 kHz with 16 bits per sample. A diagnosis was carried out for recordings of each patient and the severity of the valve lesion was evaluated by cardiologists according to clinical routine. A set of 50 patients were labeled as normal, while 98 were labeled as exhibiting cardiac murmurs, caused by valve disorders (aortic stenosis, mitral regurgitation, etc.). Furthermore, for training and validation of the algorithms, PCG signals labeled as normal and those labeled as murmur were separated, then, 360 individual beats were extracted, 180 for each class. The individual beats were picked out as the best from each cardiac sound signal, after a visual and audible inspection by a cardiologists; this was done to select beats without artifacts and other types of noise that can impair the performance of the algorithms [1].

B. Experiments

Individual beats were used as input to the presented methodology, where every recording was firstly zero-one normalized on amplitude, such that the dynamic range of signals remained constant. In order to use regularity features, the tolerance parameter r should be calculated, to not to depend on the absolute amplitude of the signal [11]. That parameter is typically given by $r = r_c std(.)$, where $std(.)$ is the standard deviation of the signal and r_c is a value varying between 0 and 1. Experimentally the highest accuracy rate, using all regularity features and SVM classifier, was achieved by utilizing $r_c = 0.15$ as shown in Fig 2, therefore fixing it as parameter.

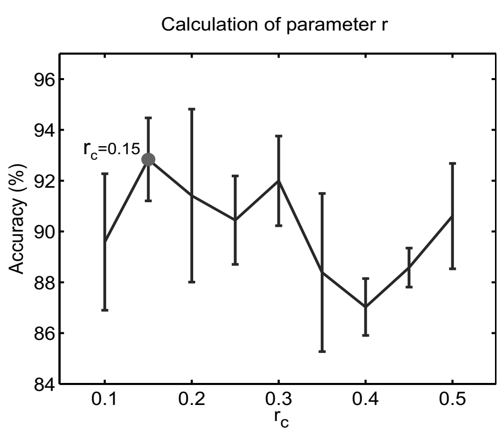

Fig. 2. Accuracy for different r_c values. Highest achieved accuracy is marked with a circle.

To evaluate performance, a 11 fold cross-validation was used. For assessment ROC curves were also utilized.

C. Results

A series of tests were performed with both k-nn and SVM classifiers, and different feature sets: features working individually, combination of GapEn and FuzzyEn (the two individual features which produced the best performance), and all regularity features working together. The obtained results are shown on table I.

Classification accuracy, up to 90% with both k-nn and SVM classifier, demonstrate the discriminative capability of FuzzyEn, and GapEn working separately. Combination of both did not significantly improved accuracy. On the other hand ApEn produced the worst classification results, being even outperformed by SampEn despite the later was largely surpassed in performance by GapEn and FuzzyEn. Combination of all features, increased accuracy to nearly 94% using the SVM classifier.

Features	k-nn	SVM
FuzzyEn	90,91± 0,90	91,51±1,29
GapEn	89,29± 2,52	90,62±1,59
SampEn	79,55± 1,92	75,21±4,63
ApEn	63,96± 3,30	71,98±3,05
GapEn+Fuzzy	91,06± 0,91	91,23±1,20
All	91,23± 0,11	93,66±1,58

TABLE I
ACCURACY OBTAINED BY USING K-NN AND SVM CLASSIFIERS

A boxplot of individual features distributed per class using the SVM classifier is shown on Fig. 3. For assessment, ROC curves are shown on Fig. 4. Moreover, area under ROC curves (AUC) are also computed because their significance on performance evaluation: the AUC of a classifier is equivalent to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance [15]; thus, the higher the AUC, the better the classifier.

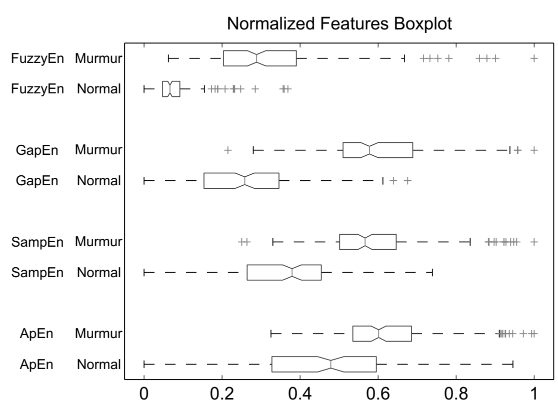

Fig. 3. Boxplot of individual features distributed per class, by using SVM classifier

As shown on Fig. 3, the complexity analysis hypothesis seems to be fulfilled with all tested features: Complexity in normal class (perceived as less regular signals) is higher than in the pathological class, suggesting the possibility on using complexity analysis for accurately detecting pathology by measuring changes on the natural complexity of biosignals.

It is also noticeable the small between-class overlapping between GapEn and FuzzyEn boxes, which might explain

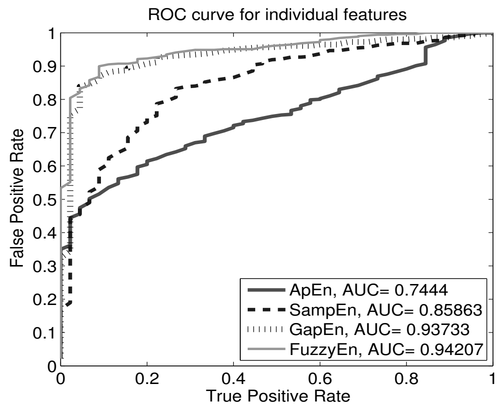

Fig. 4. ROC curves for individual features by using SVM classifier

the high accuracy rates obtained. On the other hand, besides ApEn has a small overlap between boxes, presents a large dispersion of values between whiskers which might explain its low discriminative capability. The good performance given by GapEn and FuzzyEn is revalidated as shown by the ROC curves of Fig. 4 and the high AUC obtained, up to 0.93 and 0.94 respectively. The other extrema is given by the ApEn and SampEn curves and its 0.74 and 0.85 AUC respectively; a much lower performance compared to the given by GapEn and Fuzzy.

The obtained results were comparable with some in the state of the art. In [5] for example, a dog model was used as an attempt to investigate prognostic value of heart murmur in man. In this study, regularity features were used for characterization of dogs with murmurs having aorta stenosis and innocent murmurs. Results showed that dogs with the disorder presented higher values of SampEn compared to dogs without aorta stenosis. In [6] evidence was found of SampEn as a suitable tool for quantification of cardiovascular murmurs on humans, however finding no evidence of either nonlinear or chaotic behaviour in recorded signals. Results also verified the lower regularity and hence higher complexity of healthy heart recordings than on those with murmurs.

In [1], a review of linear and nonlinear methodologies for heart murmur detection was presented. In the nonlinear stage, three nonlinear dynamic features: Correlation dimension, Largest lyapunov Exponent and Hurst Exponent were used for characterization, in conjunction with a 9 neighbours k-nn for classification. Results superior to 97% of classification accuracy and 0.99 AUC, reflect the capability of the NDA-based schema on the detection of heart murmurs. Despite results in the latter work were superior to those found in the present paper, it should be pointed out that the discriminative power of sole GapEn or FuzzyEn was remarkable, suggesting that a methodology including other NDA features in conjunction with regularity features might be useful for increasing accuracy rates.

IV. Conclusions

The results verify the complexity analysis hypothesis on all tested features: normal signals are more complex than pathological ones, and hence less regular. Also, the results evidence the good performance given by GapEn and FuzzyEn while suggesting its utility on automatic murmur detection systems for clinical applications. Moreover, the possibility of varying tolerance parameter allows system to deal with noise induced into recordings, as respiratory noise.

Despite SVM classifier produced the highest accuracy rates among all tests, k-nn accuracy was nearly as good in most of them and at a lower computational cost.

As future work, combination of regularity features with other NDA features is proposed in order to more accurately characterize dynamical complexity of biosignals, and thus seeking for better classification performance.

V. Acknowledgement

This research is carried out under the grants: *Servicio de monitoreo remoto de actividad cardíaca para el tamizaje clínico en la red de telemedicina del departamento de Caldas*, funded by Universidad Nacional de Colombia and Universidad de Caldas; and *Convocatoria de fortalecimiento de Programas de Maestrías, Doctorados y Especialización* funded by Universidad Nacional de Colombia.

References

[1] E. Delgado-Trejos, A. Quiceno-Manrique, J. Godino-Llorente, M. Blanco-Velasco, and G. Castellanos-Dominguez, "Digital auscultation analysis for heart murmur detection," *Annals of Biomedical Engineering*, vol. 37, pp. 337–353, 2009.

[2] A. Goldberger, C. Peng, and L. Lipsitz, "What is physiologic complexity and how does it change with aging and disease?" *Neurobiology of aging*, vol. 23, no. 1, pp. 23–26, 2002.

[3] D. Kaplan, M. Furman, S. Pincus, S. Ryan, L. Lipsitz, and A. Goldberger, "Aging and the complexity of cardiovascular dynamics," *Biophysical journal*, vol. 59, no. 4, pp. 945–949, 1991.

[4] M. Costa, A. Goldberger, and C. Peng, "Multiscale entropy analysis of complex physiologic time series," *Physical Review Letters*, vol. 89, no. 6, 2002.

[5] C. Ahlstrom, K. Hoglund, P. Hult, J. Haggstrom, C. Kvart, and P. Ask, "Assessing aortic stenosis using sample entropy of the phonocardiographic signal in dogs," *Biomedical Engineering, IEEE Transactions on*, vol. 55, no. 8, pp. 2107–2109, 2008.

[6] S. Schmidt, M. Graebe, E. Toft, and J. Struijk, "No evidence of nonlinear or chaotic behavior of cardiovascular murmurs," *Biomedical Signal Processing and Control*, 2010.

[7] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. Cambridge University Press, 1 2004.

[8] M. Costa, A. Goldberger, and C. Peng, "Multiscale entropy analysis of biological signals," *Physical Review E*, vol. 71, no. 2, 2005.

[9] S. M. Pincus, I. Gladstone, and R. Ehrenkranz, "A regularity statistic for medical data analysis," *J Clin Monit 1991;7:335-345*, 1991.

[10] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proc. Nati. Acad. Sci. USA Vol. 88, pp. 2297-2301*, 1991.

[11] J.-S. Richman and J.-R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Am J Physiol Heart Circ Physiol 278: H2039–H2049*, 2000.

[12] L. Xu, K. Wang, and L. Wang, "Gaussian kernel approximate entropy algorithm for analyzing irregularity of time-series," *Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 18-21*, 2005.

[13] W. Chen, Z. Wang, H. Xie, and W. Yu, "Characterization of surface emg signal based on fuzzy entropy," *IEEE Transactions on neural systems and rehabilitation engineering, vol. 15, No. 2*, 2007.

[14] W. Chen, J. Zhuang, W. Yu, and Z. Wang, "Measuring complexity using fuzzyen, apen, and sampen," *Medical Engineering & Physics 31 (2009) 61-68*, 2009.

[15] T. Fawcett, "Roc graphs: Notes and practical considerations for researchers," *Machine Learning*, vol. 31, no. HPL-2003-4, pp. 1–38, 2004.