34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Complex Networks: Application to Pathology Detection in Voice Signals

Juan Sebastián Hurtado-Jaramillo[1] *Student Member, IEEE*, Diego L. Guarín*[2] *Student Member, IEEE*, and Álvaro Orozco[1]

Abstract— **The method of complex networks has been proposed as a novel approach to analyze time series from a new perspective. However, only few studies have applied this methodology to certain types of pseudo-periodic signals. In this article, the network-based technique is applied on voice signals, a kind of pseudo-periodic signals which has not been analyzed using complex networks, to differentiate between a healthy subject and subjects with pathological disorders. The results obtained demonstrated that through a set of statistic computed from the complex networks is possible to differentiate between healthy and non-healthy subjects, contrary to what was observed using well known non-linear statistics, such as Lempel-Ziv complexity and sample entropy. We conclude that by seeing voice signals as complex networks new information can be extracted from the time series that may help in the diagnosis of pathologies.**

I. INTRODUCTION

Voice signals are graphical representations of sound waves produced by the human vocal apparatus. They are mainly used as support in the diagnosis of physiological conditions of the vocal tract of a patient [1]. Also, voice signals have been used in automatic speech recognition, to build machines capable of understanding what a person says, and to make machines able to express themselves clearly and concisely. It is essential to note that voice signals are dynamic, sometimes show transient behavior, and generally have non-stationary properties, i.e., their statistical moments and probability distribution are time-dependent [1]. In addition, evidence suggests that voice signals represent a non-linear process [2].

There has been widespread interest in the feature extraction of voice signals to contribute to the development of its application; several analytical techniques in the time, frequency [3], and time-frequency [4] domains have been used to fulfill this purpose. In particular, Mel frequency Cepstral coefficients are one of the most employed [5]. Yet, these previously mentioned methods require linearity [6], and it has been proved that nonlinear techniques can extract more information in voice signals than linear techniques [7].

Knowing that voice signals represent a non-linear process [2] and that non-linear methods work better on extracting information from these types of signals [7], there is an overriding interest in developing new techniques to analyze this kind time series. One of those new techniques is the transformation of signals into complex networks, an approach that holds a different view from the past and classical practices [8], [9]. This complex network technique creates a new way to visualize time series and therefore provides a new insight into a wide range of concepts that can be derived from this transformation [10], [11]. For example, few studies have applied the network-based technique to audio signals [12] or to pseudo-periodic signals [8], [13]. In [12] Yang et al. complex networks were constructed from audio signals in order to distinguish different types of music. Although, this aim was not achieved. In [13] Zangjie et al. the complex network technique was applied to electrocardiogram (ECG) signals; one ECG from a healthy patient and another ECG from a patient with heart arrhythmia. In this study, they obtained better results when differentiating between both ECG signals with the complex network approach than with classical techniques. This new practice was also able to detect small differences among pseudo-periodic signals which were not possible when working with more commonly used methods [13].

In this article, the concept of complex networks is applied to voice signals to distinguish between a voice signal from a healthy subject and a voice signal from a pathological subject. It is also intended to make a comparison with well known non-linear features such as the Lempel-Ziv complexity and the sample entropy so as to provide a factual example of the power of this new approach to the analysis of time series.

This document is organized as follows: in section II the voice signal database is described, as well as the basic idea behind the pseudo-periodic time series transformation into complex networks, and the set of statistics applied on the complex networks. In section III the procedure to construct complex networks from voice signals is explained; the methodology to apply the battery of statistics and the results from these statistics are also depicted in this section. Finally, conclusions are drawn and future work is proposed.

II. MATERIALS AND METHODS

A. Database

The voice signals employed in this study correspond to the KayPENTAX database [14]. This database was developed by the Massachusetts Eye and Ear Infirmary Voice and Speech Laboratory; it consists of more than 1400 voice samples of approximately 700 healthy subjects as well as those with a pathological disorder. The recordings consist of the sustained speech phoneme */ah/*. All voice signals were recorded in a

*D. Guarín was supported by the program "Jóvenes investigadores e innovadores 2011" of COLCIENCIAS.

[1]J. S. Hurtado-Jaramillo and Á. Orozco are with Department of Electrical Engineering, Universidad Tecnológica de Pereira, Pereira, COLOMBIA. jshurtado at utp.edu.co

[2]D. L. Guarín is with the Department of Biomedical Engineering, McGill University, Montréal, QC H3A 2T5, CANADA. diego.guarinlopez at mail.mcgill.ca

controlled environment, as follows: low ambient noise, fixed length distance between the speaker and microphone, direct digital sampling of $16-bit$, robust signal conditioning, and sampling frequency of $25\ kHz$ or $50\ kHz$. Fig. 1 shows three cycles of a voice signal from a healthy subject and three cycles from a subject with a pathological disorder; both signals were sampled at $25\ kHz$.

B. Complex Networks

To transform a pseudo-periodic time series into a complex network, the procedure described in [8] and [13] is applied into the pseudo-periodic voice signals. A brief description of this process is as follows: start with a pseudo-periodic time series (a voice signal in our case) $\{x_i\}_1^n$ of n observations. Divide the time signal into M suitable cycles, either from maximum or minimum points of a cycle. Embed the M cycled pseudo-periodic time signal using one of the standard techniques, creating an attractor of $\{C_1, C_2, \ldots, C_M\}$ cycles embedded in m dimensions with an optimal delay τ (for a more detailed explanation on embedding see [15]). Each point in the m-dimensional space is considered as a point in the complex network. Define K as the number of nodes to be connected with each node in the complex network; keeping in mind that for nodes to be connected they have to be close to each other, but nodes from the same cycle C_i cannot be considered as connected. Finally, create the complex network with K-node connections.

C. Network Statistics on Complex Networks

1) Global Clustering Coefficient: The local clustering coefficient of a node v depicts the density of connections in the direct neighborhood of node v in terms of the density of connections between all vertices that are incident with v [16]:

$$\mathcal{C}_v = \frac{2}{k_v\left(k_v-1\right)} N_v^{\Delta}, \tag{1}$$

where N_v^{Δ} is the total number of closed triangles including node v, which is limited by the maximum value of $k_v\left(k_v-1\right)/2$ (k refers to the degree of centrality, i.e. the number of neighborhoods directly connected with the node v). For isolated or treelike nodes (i.e., $k_v = 0$ or 1) the clustering coefficient is marked out as $\mathcal{C}_v = 0$ by definition. If a complex network has a high clustering coefficient, it reveals a specific kind of structure in a network. This high clustering coefficient has to do with the tendency to be associated with only a selected group of nodes.

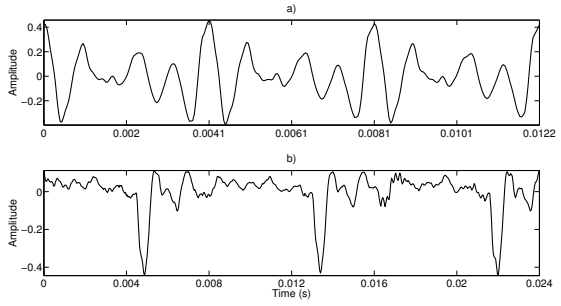

Fig. 1. a) Voice signal from healthy subject. b) Voice signal from subject with pathological disorder.

The global clustering coefficient represents the average value of all the local clustering coefficients (N) of the complex network:

$$\mathcal{C} = \frac{1}{N}\sum_{v=1}^{N}\mathcal{C}_v. \tag{2}$$

2) Assortativity: Assortativity refers to the tendency of a node v with degree k to be connected with other nodes of similar degree k. If nodes of degree k tend to be connected with other nodes of degree k, then the complex network is said to be assortative. Conversely, disassortative refers to the tendency of nodes of degree k to be connected with nodes of degree different than k [16].

Assortativity can be calculated using the Pearson correlation coefficient of the node degrees on both ends of all edges:

$$\mathcal{R} = \frac{\frac{1}{L}\sum_{j>i} k_i k_j A_{i,j} - \left[\frac{1}{L}\sum_{j>i}\frac{1}{2}\left(k_i+k_j\right)A_{i,j}\right]^2}{\frac{1}{L}\sum_{j>i}\frac{1}{2}\left(k_i^2+k_j^2\right)A_{i,j} - \left[\frac{1}{L}\sum_{j>i}\frac{1}{2}\left(k_i+k_j\right)A_{i,j}\right]^2}, \tag{3}$$

where L is the total number of edges in the recurrence network and $A_{i,j}$ is the number of nodes $j \neq i$ that are directly connected with i.

3) Path Length: The following assumption has to be made in order to calculate the path length of a complex network: complex networks are undirected and unweighted, i.e., all of the edges of the network are assumed to be of unit length in terms of geodesic distance. Now, the shortest path length $l_{i,j}$ is defined as the length of the distance between any two vertices, i and j, of the complex network. The nomenclature $l_{i,j}$ in the network indicates the minimum number of edges that have to be passed on a graph from a node i to a node j [16].

Conversely, the average path length $\mathcal{L}$ is the mean value of the shortest path lengths for all the pairs of vertices of the complex network:

$$\mathcal{L} = \langle l_{i,j}\rangle = \frac{2}{N\left(N-1\right)}\sum_{i<j} l_{i,j}. \tag{4}$$

It is worth noting that $l_{i,j}$, the shortest path length, is set to zero by definition for a disconnected pair of nodes; nonetheless, this situation does not affect in a great manner the statistic.

D. Other Non-linear Statistics

The following non-linear statistics have been selected in this study because two of the authors proved that these non-linear statistics can find differences among pseudo-periodic signals [17]. Besides, a comparison is intended to be made between these non-linear statistics and network statistics.

1) Lempel-Ziv Complexity: (LZC) is an important non-linear statistic capable of differentiating between several states of a dynamical system. The complexity is a measurement of the regularity of a symbolic sequence. Before estimating the complexity from a time series, it is necessary to convert it into a sequence of elements from a finite and small set of symbols. The complexity is defined as the number of sequences one observes in a symbolic sequence as a fraction of the maximum number of sequences. The maximum number of sequences that would be observed are those observed for a random sequence of symbols [17].

2) Sample Entropy: (SampEn) is a measure based on comparing patters within a time series to approximately calculate its complexity through estimation of entropy rates. Usually, the SampEn is calculated with the parameters suggested for short data segments, i.e., $L = 2$, and $\epsilon = 20\%$ of the standard deviation of the observed time series. Higher values of SampEn are associated with higher irregularity and lower values of SampEn to a high degree of regularity [17].

III. RESULTS AND DISCUSSION

A. Cycle Selection

As described in II-A, the database consists of voice recordings from healthy subjects as well as subjects with a pathological disorder. To follow the procedure mentioned in [8] and [13], voice signals have to be limited to an integer number of pseudo-periodic cycles. To do this, the "Applied Non-linear Time Series Analysis (ANTA)" Toolkit of Matlab (The Mathworks Inc.) was used to find the local maximum for each cycle [15]. Another parameter taken into account was that the difference between the first and the last datum (end-point mismatch) of the signal should be less than 1% of the maximum value [15].

Ten signals from healthy subjects and ten signals from subjects with pathological disorders were selected for this preliminary study. The application of this method is straightforward for voice signals from healthy subjects, since the pseudo-periods are well defined. However, for some voice signals from subjects with pathological disorders this is not the case. Therefore, the ten pathological signals were selected given a visual identification of the pseudo-periods. Fig. 1 shows two signal with three noticeable cycles. Note that if the cycles in the signal are not detectable then this method cannot be applied.

B. Graphical Representation of a Complex Network

To create a graphical representation of a complex network, one must first obtain the time delay embedding and the embedding dimension of each voice signal. Then, a matrix containing the node connections is created using the preceding parameters as well as the defined threshold K (60 neighbors in the attractor). It is important to keep in mind that for embedding dimensions greater than 3, the visualized complex network is a projection in only 3 dimensions. Due to computational and visual limitations, complex networks were created using three cycles for each voice signal. Fig. 2 displays four complex networks created using the software Pajek [18], the two complex networks in the upper panel correspond to healthy subjects while the ones in the lower panel to subjects with pathological disorder. At a simple glance one can tell a visual distinction between the complex networks in the upper panel from the complex networks in the lower panel. Also, if by comparison a complex network seems to be more distributed than other, as in Fig. 2, then it can be said that the underlying signal is more complex than the other [11].

C. Application of Complex Network Statistics

To calculate the already-mentioned statistics in II-C, the procedure differs from III-B in the number of cycles for the creation of the correspondence matrix. In this case, 50 cycles were chosen for each signal so as to generate correspondence matrices. These signals vary in size from 5000 to 10000 data points.

Fig. 3 shows the calculated values for the clustering coefficient, assortativity and path length for healthy subjects (+) as well as for subjects with pathological disorders (o).

As observed, the clustering coefficient can differentiate between subjects and patients. Proving that the node connections of the complex network derived from the healthy subject are denser than those of a complex network coming from a patient.

The difference seen in the clustering coefficient cannot be noticed when referring to the assortativity. But, this result was expected because all the networks were created with a constant number of connections ($K = 60$). Furthermore, the path length behaves in a similar way as the assortativity; there is no difference between the two types of complex network. But again, since both types of complex networks were constructed using the same parameters, the path length of all the complex networks is expected to be similar. This could possibly mean that it is unimportant if the subject is in a healthy or pathological state; these statistics depend on the number of connected nodes.

With the intention of demonstrating how powerful the theory of complex networks is with respect to traditional approaches, Fig. 4 shows the values of the Lempel-Ziv complexity and the sample entropy for both types of voice

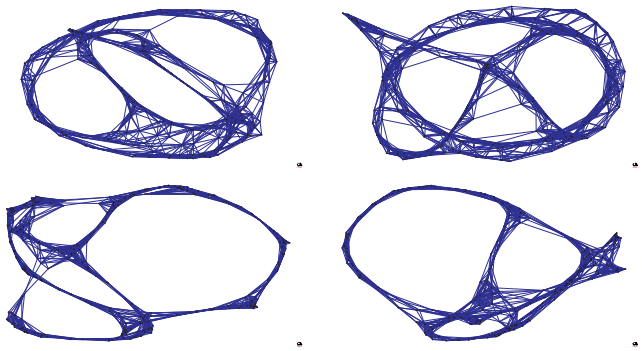

Fig. 2. Upper panel: complex networks for healthy subjects. Lower panel: complex networks for subjects with pathological disorder.

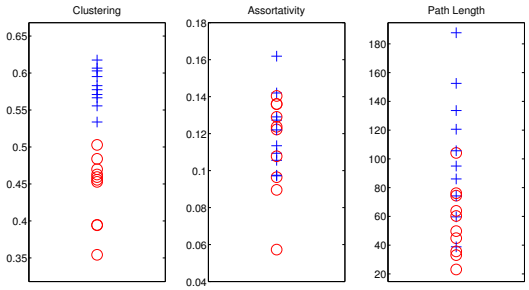

Fig. 3. Clustering coefficient, assortativity and path length for the generated complex networks. The plus sign (+) represents the calculated value for the complex networks from healthy subjects, while the circle (o) from subjects with pathological disorder.

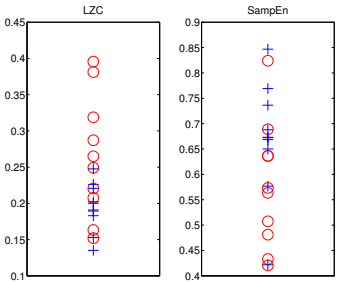

Fig. 4. Lempel-Ziv complexity and Sample Entropy for the selected voice signals. The plus sign (+) represents the calculated value for voice signal from healthy subjects, while the circle (o) from subjects with pathological disorder.

signals. It can be seen that they are not able to distinguish a healthy subject from a patient. One possible reason is that the regularity (or irregularity) of both kind of signals is considerable similar; because of the fact that voice signals from subjects with pathological disorder were selected to have a recognizable pseudo-periodic wave form. Another probable cause is that both the Lempel-Ziv complexity and the sample entropy were extracted with commonly used parameters, as specified in II-D, which may not be the adequate values for these particular types of voice signals.

IV. CONCLUSIONS

Through the technique of complex networks, a new representation of voice signals was presented. This network-based transformation provided alternative ways to analyze pseudo-periodic signals. We demonstrated that using this novel technique is possible to differentiate voice signals from healthy subjects and from patients; this may help in the diagnosis of pathologies. Additionally to the extraction of quantitative features through some statistics, complex networks present a graphical interpretation that does intuitively depict a visual distinction between healthy and non-healthy subjects.

From the battery of statistics applied to the complex networks generated from voice signals, the clustering coefficient was able to entirely differentiate between the two types of voice signals under analysis. Moreover, classical statistics such as Lempel-Ziv complexity and sample entropy, could not accomplish the same results. This discrepancy gives the complex network approach more relevance among the techniques to study time series, particularly pseudo-periodic signals such as recordings of the human voice.

For future work, the complex network methodology could be applied to a larger number of subjects. Also, subjects could be divided into specific pathological disorders in order to assess the capability of the network-based methodology in making a distinction among different pathologies.

REFERENCES

[1] R. M. Rangayyan, *Biomedical Signal Analysis: A Case-Study Approach*. New York: IEEE Press, 2002.

[2] M. Small, C. K. Tse, and T. Ikeguchi, in *Proceedings of the European Conference on Circuit Theory and Design*, Cork, Ireland, 2005, pp. 169–172.

[3] J. R. Deller Jr., J. H. Hansen, and J. G. Proakis, *Discrete-Time Processing of Speech Signals*. Piscataway: IEEE Press, 2000.

[4] R. Chiodi and D. Massicotte, in *First International Conference on Advances in Satellite and Space Communications*, 2009, pp. 54–57.

[5] L. S. Chee, O. C. Ai, M. Hariharan, and S. Yaacob, in *IEEE Student Conference on Research and Development*, 2009, pp. 146–149.

[6] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N.-C. Yen, C. Tung, and H. Lui, *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 903–995, 1998.

[7] J. D. Arias-Londoño, J. I. Godino-Llorente, N. Sáenz-Lechón, V. Osma-Ruiz, and C. G. Castellanos-Domínguez, *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 370–379, 2011.

[8] J. Zhang and M. Small, *Physical Review Letters*, vol. 96, pp. 238 701–1–238 701–4, 2006.

[9] M. Small, J. Zhang, and X. Xu, in *Complex Sciences*, J. Zhou, Ed. Springer, 2009, vol. 2, pp. 2078–2089.

[10] X. Xu, J. Zhang, and M. Small, *Proceedings of the National Academy of Sciences*, vol. 105, no. 50, pp. 19 601–19 605, 2008.

[11] R. V. Donner, M. Small, J. F. Donges, N. Marwan, Y. Zou, R. Xiang, and J. Kurths, *International Journal of Bifurcation and Chaos*, vol. 21, no. 4, pp. 1019–1046, 2011.

[12] Q. Yang, Q. Gao, and R. Fan, in *2nd International Conference on Intelligent Control and Information Processing*, Harbin, 2011, pp. 929–930.

[13] J. Zhang, J. Sun, X. Luo, K. Zhang, T. Nakamura, and M. Small, *Physica D*, vol. 237, pp. 2856–2865, 2008.

[14] KayPENTAX, "Disordered voice database and program, model 4337," Web: http://www.kayelemetrics.com/.

[15] M. Small, *Applied nonlinear time series analysis : Applications in physics, physiology and finance*. World Scientific Publishing, 2005.

[16] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, and J. Kurths, *New Journal of Physics*, vol. 12, pp. 1–40, 2010.

[17] D. L. Guarín, , and A. A. Orozco, in *The 11^{th} International Conference on Information Science, Signal Processing and their Applications, ISSPA*, Montreal, QC, Canada, 2012, (To appear).

[18] V. Batagelj, A. Mrvar, and M. Zaveršnik, "Pajek," [Software]. Available from http://pajek.imfm.si.