



Cross-Recurrence Plots and Quantification of Glottal Signal for Pathological Voice Assessment

Mohamed Dahmani*, Mhania Guerti

Signal and Communication Research Laboratory, Ecole Nationale Polytechnique, Algiers 16200, Algeria

Corresponding Author Email: mohamed.dahmani@g.enp.edu.dz

<https://doi.org/10.18280/ts.370217>

ABSTRACT

Received: 31 November 2019

Accepted: 15 January 2020

Keywords:

assessment, cross recurrence quantification analysis, glottal signal, vocal folds

Non-linear signal processing techniques know an extensively challenge use in the field of pathological voice diagnosis and assessment. This paper contributes the application of an important tool, Cross Recurrence Quantification Analysis (CRQA) for showing how it can be adapted to compare, assess and quantify the articulation changes of Vocal Folds (VFs) after a voice therapy. To achieve this aim, four types of pathological voice: VFs paralysis, VFs nodules, VFs polyps and spasmodic dysphonia have been considered in which the normal voice was considered as a reference. The glottal signal was firstly extracted using an inverse filtering algorithm, in the next step; the CRQA was performed on the Cross Recurrent Plot (CRP) structure of the glottal signal. Eight (08) CRQA measures were extracted to evaluate the improvement degree of voice quality after a voice therapy operation. By comparing these parameters variation before and after therapy to the normophonic voices, pathological voice stages were easily discriminated. For the test validation, a local collection database has been recorded in three different Algerian Ear-Nose-Throat services, where the long vowel [a:] was selected. A number of 60 pathological samples was achieved of the two genders adults male and female uttered in different stages before and after voice therapy. The obtained results shown the effectiveness of CRQA applied to the glottal signal. Indeed, it presented an effective tool to assess the improvement of voice quality after therapy.

1. INTRODUCTION

The assessment of pathological voice can be divided into two majors types: subjective or objective. In the subjective case, the Ear-Nose-Throat doctors and speech therapists already bank on hearing or by using a set of measuring medical devices, more or less sophisticated such as: Laryngoscopy, electroglottography and stroboscopy. The subjective assessment necessarily depends on the experience of the practitioners, so to attribute such decision on the pathology, several specialists evaluate the same case to reach a consensus [1]. However, despite its importance in the morphological observation of the phonatory apparatus, it is poorly adapted to the study of laryngeal functioning due to the hardly intubation access for the visual examination during the phonation process. Hence, instabilities, phonation onset, irregular vibrations and investigation of the singing voice transition cannot be investigated [2]. In addition, their prohibitive price and maintenance limit their popularization use in clinical practice.

The objective assessment, which is our interest, is conducted by a computer program designed to qualify vocal dysfunction, performed directly to the patient during the vocal production. This assessment is basically based on sophisticated algorithms allowing the analysis of voice signals. These methods can be considered as a primary screening tool to evaluate the functional quality of voice production and decision making through results analysis [3].

With the wide proliferation of signal processing techniques

and filtering algorithms, the use of voice and speech signals contributes the development of expert aided systems for the detection of voice impairments. These systems facilitate the understanding of the vibration of vocal folds and pathologies formation. They constitute also a useful tool as a medical document helping in the trailing of persons whose suffering from several voice disorders [4]. To investigate the mechanism of voice production under pathologies, several methods and algorithms were carried out known as features extraction techniques, parametric analysis, linear and non-linear signal analysis and statistical methods have been developed.

In the literature, the so-called features extraction techniques have been also developed to investigate the vibration patterns of VFs, these techniques mainly include various parametric signal analysis and, describing the signal variation for several applications. In order to demonstrate the effectiveness of voice therapy in patients with vocal cord disorders, the acoustic parameters change before and after treatment have been illustrated in the study [5]. The results showed clearly the importance of voice therapy, especially for longer treatment periods. As in the research [6], the (i) Mel Frequency Cepstral Coefficients (MFCC) combined with fundamental frequency, (ii) jitter and (iii) shimmer as features set for VFs pathologies diagnosis using Nave Bays Networks. The obtained results construed the ability of perceptual analysis and acoustics parameters for discriminating disordered from normophonic. The advantage of Acoustic parameters is its nonintrusive provided a quantitative analysis of the voice signal. So also, Al-Nasheri et al. [7] proposed a robust parameters extraction

technique for pathological voices assessment. Whereas, the entropy and maximum peak and their corresponding values are extracted by autocorrelation applied to the voiced signal frames. These parameters were evaluated in different frequency regions to determine the most contributed bands affected by the pathology. Experiments were accomplished in three different databases to study the reliability of the tuning process in an inter-database scenario, this made the results more reliable and preferment. However, these results should be more representative in a context of database recorded in several sessions of voice therapy.

The glottal signal articulation parameters (describe the articulator aspect of VF; obtained by an inverse filtering algorithm just after VFs and before vocal tract) were used to discriminate some VFs pathologies using artificial neural network, SVM, KNN and hidden Markov model [8, 9]. In spite of their good results, these parameters are more relevant to the study of voice quality before and after voice therapy, providing a suitable state discrimination of disordered voices.

Recurrence Quantification Analysis (RQA) is a non-linear signal analysis tool able to quantify the activity of dynamical systems performed directly to the Recurrence Plot structure without the need for system identification but only the form of the signal [10]. RQA attracted the attention of many researchers in diverse areas especially, biomedical engineering to discriminate between healthy and non-healthy persons. RQA has been successfully applied for heart diseases diagnosis based on the heart beat signal to distinguish between healthy and unhealthy patients [11]. It can be stated that RQA has not studied deeply in the field of pathological voice detection, de A. Costa et al. [12] analyzed speech signals to detect laryngeal pathologies using quantification measures of RP structure. In their investigation, Linear Discriminate Analysis (LDA) and Quadratic Discriminate Analysis (QDA) were used as a detection method. The results show an important discriminative potential to distinguish healthy voices from pathological ones. However, the application of RQA to a specific signal describing a physiological phenomenon such as glottal signal, should give better results.

CRQA is used in many research fields such as: underwater acoustic [13] to estimate the time-difference of arrival underwater acoustic signals, emotions recognition to evaluate the similarity of multivariate patterns [14], Characterizing of Electroencephalogram's dynamics signal in different states during the anesthesia [15] ... etc.

This paper aims to develop a new computational methodology technique based on the Cross Recurrence Quantification Analysis. A method that is more suited for assessing the similarities and compare the repeated articulatory over the time. The main advantage of this method in our field is that is able to quantify the variability that characterizes the articulatory aspect and vibration of vocal folds during voice production. Also, it provides a means to compare, assess and distinguish normal voices from pathological ones. This paper mainly focuses on the analysis of improvements degree of pathological voice quality before and after voice therapy to achieve the normal voice. In this context the CRQA is applied on the CRP maps of the two glottal signals waveforms, eight (08) CRQA indicators are studied, where, their mean values are used as indicators sings to detect the most pertinent parameter that practitioners should take it into consideration during the re-education process. The application has been successfully carried out using an Algerian pathological voice database recorded in the Ear-Nose-Thorax

service of three different hospitals. The obtained results show that the CRQA seem to be a very efficient method to diagnose and assess vocal folds pathologies.

The remaining of this paper is organized as follow, the section II provides the methodology of the proposed approach for human voice disorders assessment including the description of CRQA measurements. The section III concerns the experimental setups containing the details of the adopted database. The parametric evaluation and results are discussed in section IV. The overall summary and conclusions are provided in the last section.

Cross Recurrence Quantification Analysis is a tool able identifies the degree to which two time-series trajectories are in a similar state. As well as it, permits to examine coupling between the glottal signal changes caused by the presence of pathologies and evaluate the improvement of voice quality after a voice therapy.

2. DATABASE

In order to validate the proposed methodology, an Algerian disordered voice database has been adopted. The data collects more than 100 samples of normal and pathological voices recorded in 3 different hospitals between October 2016 and December 2018. In the present paper, four types of pathological voice have been investigated. These pathologies have been selected according to their widespread as well as their medical diagnosis knew a bit tricky to date which deal a miss-diagnosis.

The spasmodic dysphonia is characterized by an occasional break in voice caused by sudden involuntary movements of VFs which perturb its normal vibration. VFs paralysis is a neurological disease that results by an immobility of one or both vocal folds causing a hoarse and diplophonic voice. The polyps are considered as an organic disease affects often one Vocal fold in different sizes and forms producing vocal rough and fatigue. VFs nodule caused usually by a vocal hyperfunction, resulting a turbulent passage of air through glottis. Normal sustained vowels [a:] samples (of around 1 and 3 seconds and sampled at 50 kHz with 16 bits of resolution) have been used.

The Table 1 illustrates the number of samples and gender distribution. The control group consisted in this study comported of 53 healthy adult males and females with a range of age 25-50 years who did not have any historical voice disorders.

Sixty patients completed the speech therapy program, which lasted 6 weeks, a one-hour session a week. They were programmed as indicating the table above with 53 subjects considered as a control group. The therapy sessions include: muscle relaxation, vocal hygiene, carryover and respiration.

- **Vocal hygiene:** it is a very important step in vocal therapy, which included an education on how the normal voice is produced by avoiding vocal abuse, it consists of system and local hydration.
- **Muscle relaxation:** The goal is to avoid a false gesture that causes vocal forcing and teach you the right musculoskeletal gesture by applying easy gymnastics, and massaging of the shoulders, neck, and laryngeal muscles. It comports also yawn-sigh movement, which help and reduce vocal hyperfunction.

- **Respiration:** It is important to note that voice therapy is essentially based on abdominal breathing, so stretching and stretching the back regularly allows the rib cage to open properly by releasing the passage of air to the vocal tract.
- **Carryover:** it consists of applying attempts by transferring newly learned vocal behaviors to situations of daily life does not fall within the therapeutic framework, for example recitation of Quran, communication, give oral lectures.

Table 1. Details of the data used in this investigation

	Subjects		Range of age		Mean of age		Standard deviation of age	
	Male	Female	Male	Female	Male	Female	Male	Female
Normal	27	21	25-45	20-50	35	35	6.06	7.34
Polyps	07	09	26-50	25-45	38	35	7.61	5.48
Nodule	08	10	40-57	38-50	42.5	44	5.61	4.48
VFP	06	07	25-40	20-38	32.5	29	4.80	4.97
SD	05	08	23-43	24-44	33	34	5.69	6.08

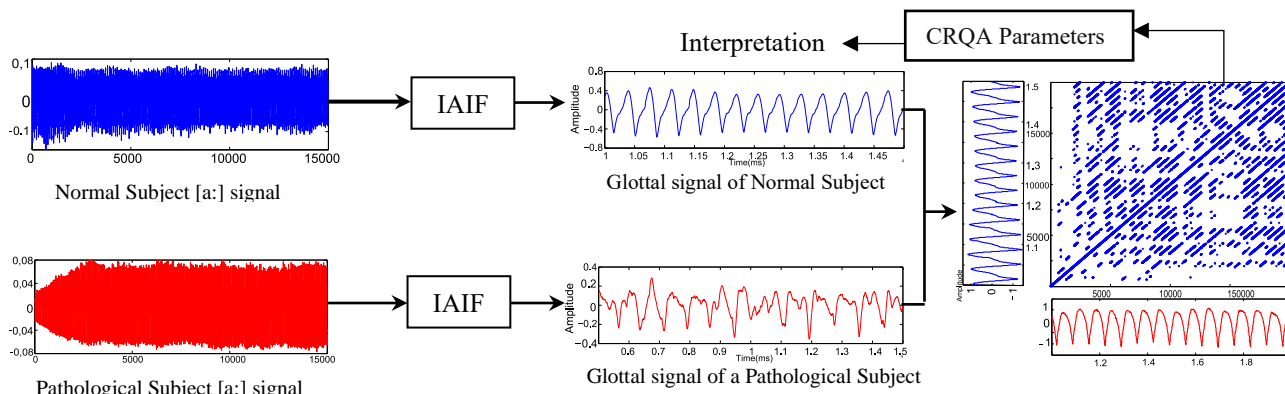


Figure 1. Block diagram of the proposed approach

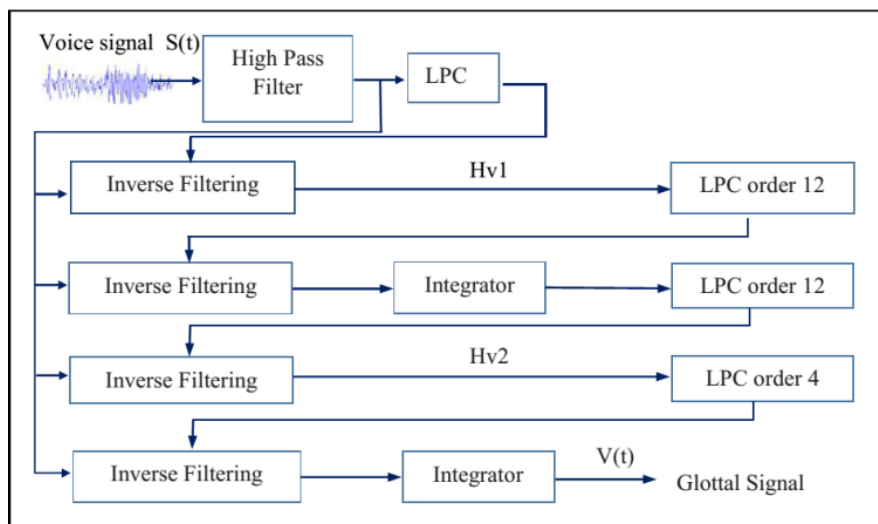


Figure 2. Block diagram of the IAIF Algorithm

3. PROPOSED APPROACH

This section concerns the methodology of the proposed approach used for the evaluation and assessment of human vocal folds disorders. The proposed solution consists mainly on the generation of a parametric model based on the cross recurrence quantification analysis of voice quality of patients with different voices pathologies (VFs paralysis, VFs nodules, VFs polyps, spasmodic dysphonia and the normophonics considered as references). For this purpose, by the main of an Iterative Adaptive Inverse Filtering (IAIF) algorithm, the glottal signal is firstly extracted from the recorded samples of the long vowel [a:]. The Cross-RQA parameters have been

evaluated by using the mean values of each parameters before and after the voice therapy. This permits not only to evaluate the performance of parametric Cross-RQA analysis, but also to detect the most pertinent parameters that the clinician or speech therapist should take attention to it in order to reestablish the voice to the patients. Eight parameters are extracted (Cross Recurrent Rate, determinism, Maximal diagonal line length, Average diagonal line length, Entropy, Trapping time, Linearity and the longest diagonal line) were used as indicators measurements for the comparison between the control group and dysphonic patients. The study was examined on a corpus included 120 samples. The control group consisted of 53 healthy peoples; the rest of the whole

dataset constitutes the dysphonic patients. Four types of vocal folds pathologies are selected: VFs: Nodules, polyps and Paralysis and Spasmodic dysphonia. In summarizing, Figure 1 shows the block diagram of the proposed approach.

3.1 Glottal signal extraction

The representation of all relevant characteristics of the glottal source excitation requires the separation of the vocal tract component. Iterative Adaptive Inverse Filtering (IAIF) algorithm is used to extract the glottal signal [16]. According to [8, 17], this method provides a robust estimate of the voice signal decompositions, which were used to determine any dysfunction in vibration patterns caused by pathological masses on VFs. In this study, the speech signal was pre-emphasized prior to a Linear Predicting Coding (LPC) inverse filtering analysis to obtain a better approximation to the vocal tract filter and the glottal source signal $v(t)$ from the speech signal $s(t)$. Although, this model processes some limitations in terms of representing some sounds, such as the nasals sounds, because these sounds are independent to the vocal tract [18]. Figure 2 illustrates the procedures required to eliminate ambient noise, which would affect the results of the LPC analysis.

In the Iterative Adaptive Inverse Filtering (IAIF) method, the voice signal is high-pass filtered by a linear phase finite impulse response (FIR) which is the most widely used method for glottal signal estimation in Figure 2. The IAIF method permits to decompose automatically the voiced sounds into vocal tract transfer function and the glottal source signal.

The filtered speech signal is then analyzed by a first-order LPC providing $Hg_1(t)$ preliminary estimate of the glottal flow and lip radiation effects. Using this obtained first-order LPC filter, the high-pass filtered speech signal is inverse filtered in the third step, cancelling the effects estimated by $Hg_1(t)$. The output of this is analyzed by a p-order LPC in the fourth step and the resulting estimation of the filtering effects, indicated by $Hvt_1(z)$, is used to reverse the filtering of the filtered speech signal again. High pass filter is used to reduce the effects of the vocal tract in the fifth step. The final estimate of vocal tract effects is calculated by another p-order LPC analysis to provide $Hvt_2(z)$. This is used to inverse filter the high pass filtered acoustic waveform one last time at step eleven by integrating the result to give the final estimate of the shape of the glottal signal $v(t)$ with cancellation of the vocal tract effects.

3.2 Recurrence plots maps

Recurrence Plot (RP) technique constitutes a graphical tool to visualize the recurrence behaviour of the various states of dynamical systems. Especially useful for investigate the non-stationary property of high dimensional systems. RP was proposed for visualizing patterns of genetic nucleotides [19]. It is represented by a symmetrical matrices $N \times N$, in which elements R_{ij} are defined as below:

$$R_{i,j} = \begin{cases} 1, & \|X_i - X_j\| \leq \varepsilon \\ 0, & \|X_i - X_j\| > \varepsilon \end{cases} \quad i, j = 1, \dots, N \quad (1)$$

in which, $\|\cdot\|$ is the Euclidian distance and X is the studied time series; it represents the glottal signal's time series. The coefficient " ε " represents the neighborhood radius at X_i point.

The expression (1) means that if the distance between X_i and X_j is less than " ε ", $R_{ij} = 1$. In this case, a dot is placed at $(i; j)$ point in the RP map. Otherwise, $R_{ij} = 0$ and a blank mark is placed at $(i; j)$ position.

3.3 Cross Recurrence Plots

Cross-Recurrence Plots (CRP) is an extension variant of recurrence plots employed to investigate the dependencies and similarities evolution between two different systems by comparing their states. It based on the behavioral recurrence of the phase trajectory of the two signals [20]. To make the concept clearer, considering two measured signals $x(t)$ and $y(t)$ with lengths N_x and N_y respectively and t as a sampling interval. In the case of our contribution the two signal $x(t)$ and $y(t)$ represents the time series of glottal signal of normal voice subject and pathological. The time series of N samples is given as follow:

$$NS(n) = [x(1), x(2), \dots, x(N_x)] \quad (2)$$

$$PS(n) = [y(1), y(2), \dots, y(N_y)] \quad (3)$$

where, NS and PS are the representation of glottal signal: Normal subject and Pathological subject. By phase space X , Y of N dimensions and time-delay embedding method, the glottal signals can be expressed as follow:

$$X_{NS}(t_i) = [x(t_i), x(t_i - \tau), \dots, x(t_i - (m-1)\tau)] \quad (4)$$

$$Y_{PS}(t_i) = [y(t_i), y(t_i - \tau), \dots, y(t_i - (m-1)\tau)] \quad (5)$$

in which, m is the embedding dimension of the phase space, is a delay and $i = [1; 2 \dots; N]$ is the sample index. Comparing the degree of similarity between these two vectors leads to the similarity matrix defined as:

$$d(i, j) = Sim(X_{NS}(i), Y_{PS}(j)) \quad (6)$$

CRPA is obtained then through the comparison of each coefficient of the similarity matrix to a threshold defined as:

$$CRP_{i,j}^{X_{NS}, Y_{PS}} = \Theta(\varepsilon - \|X_i - Y_j\|) \quad i, j = 1, 2, \dots, N \quad (7)$$

where, " Θ " represents the Heaviside function and " ε " is an appropriately chosen threshold.

3.4 Cross-Recurrence Quantification Analysis

CRQA is an extension of the recurrence quantification analysis, a tool based on the statistical description of the parallel line and points distribution among the CRP. Measures of complexity defined using the recurrence point density and diagonal line structures in the CRP maps provide a qualitative description of the dynamics of the studied glottal signal time series [21]. In this study eight parameters have been suggested considered as the most pertinent parameters in several classification or assessment contexts:

- **Cross Recurrence Rate (CRR):** It represents the probability of appearance of similar states in the two

signals with a certain delay t . A high density of recurrence points in a diagonal implies high RR values. In this case the trajectories of the two systems often visit the same regions of phase space.

$$CRR = \frac{1}{N^2} \sum_{i,j=1}^N R_{i,j(i \neq j)}(\varepsilon) \quad (8)$$

where, N is the number of recurrence points on the phase space trajectory.

- **Determinism (DET):** represented by the proportion between the number of recurrence points, forming the diagonal structures and all the R_{ij} points of CRP map. A high determinism values indicates a long diagonals lines, i.e. the both systems possess similar phases behaviour. The mathematical formula of the determinism is given by:

$$DET = \frac{\sum_{l=l_{min}}^N l \times P(l)}{\sum_{l=1}^N l \times P(l)} \quad (9)$$

where, $P(l)$ stands for the probability distribution of a diagonal line of length l . For fully deterministic signals, the DET is close to 1 because most recurrent points form diagonal lines. For a white noise signal, the DET is close to 0 because most recurrent points in the RP structure are single isolated points and form a few diagonal lines. $l_{min} = 2$ is the minimum length of the diagonal line [21]. $P(l)$ can be expressed as

$$p(l) = (N_l) / \left(\sum_{a=l_{min}}^{l_{max}} N_a \right) \quad (10)$$

where, l_{max} denotes the maximum length of diagonal line. N_a and N_l denotes the number of diagonal lines with number a and l respectively.

- **The Maximum diagonal line length (L_{max})** on a cross recurrence plot maps (L_{max}) represents the longest uninterrupted period in which the two systems remain attunement, this indicator expresses the stability of coordination. Example: noise sensitivity and external disturbances create unstable sequences, therefore a diagonal longer and shorter. The inverse of (L_{max}) is related to the exponential divergence of the phase-space trajectory $DIV = 1/L_{max}$.

$$L_{max} = \max(l_i, i = 1, \dots, N_l) \quad (11)$$

where, $N_l = \sum_{l \geq l_{min}} P(l)$ correspond to the total number of diagonal lines present in RP structure.

- **Average diagonal line length (MEANLINE):** Reveals to the duration of such similarity in the signal dynamics. A high coincidence of the two systems resulting in an increase in the length of these diagonals. In other way these lines are related to the divergence of the trajectory segments. It can be interpreted also as the mean

prediction time of dynamical systems [22]. The Eq. (12) expresses the formula of average diagonal line length.

$$L = \frac{\sum_{l=l_{min}}^N l P(l)}{\sum_{l=l_{min}}^N P(l)} \quad (12)$$

- **Laminarity (LAM):** quantifies the percentage of density of recurring points forming the structure of vertical lines in the cross recurrence map. It delineates the time intervals during which the state of the two signals representing systems dynamics is relatively constant compared to brutal activity intervals. A low laminarity value indicates that the RP map consists more single recurrence points than vertical lines [23].

$$LAM = \frac{\sum_{v=v_{min}}^N v P(v)}{\sum_{v=1}^N v P(v)} \quad (13)$$

where, $P(v)$ is the frequency distribution of the vertical lines of length v , which have at least a length of v_{min} in order to minimize the tangential motion effect. N is the length of the signal.

- **Entropy (ENTR)** Provides information about the complexity of CRP the attunement between the two signals, if the diagonal lines tend to have the same length so the attunement is very regular which result low values of $ENTR$, otherwise it is more complex (high $ENTR$ value). This indicator is referred to the Shannon entropy of the probability $p(l)$ to find a diagonal line with an exactly length l in the CRP structure. It is computed according to the formula:

$$ENTR = - \sum_{l=l_{min}}^N P(l) \ln[P(l)] \quad (14)$$

where, l_{min} represents the minimal diagonal line length which is fixed to $l_{min} = 2$ here, because larger value of l_{min} necessary only for very smooth, continuous data. Small values of $ENTR$ indicate a low complexity of CRP structure [24].

- **Longest vertical line (V_{max}):** analogously to (L_{max}), (V_{max}) measures the longest vertical line in the CRP structure. It is related to the singular states where the two systems remain in a laminar state [25].

$$V_{max} = \max \left(\{ v_l \}_{l=1}^{N_v} \right) \quad (15)$$

where, N_v is the absolute number of vertical lines. Hence, this measure allows the investigation of the intermittency of non-stationary data series.

- **Trapping time (TT):** Its calculation requires also as laminarity, the consideration of a minimal value of V_{min} . This measure provides information about the amount and length of vertical structures, it describes the average time in which the two trajectories stay in the same region stand a specific state. It is given by the formula:

$$TT = \frac{\sum_{v=v_{\min}}^N vP(v)}{\sum_{v=v_{\min}}^N P(v)} \quad (16)$$

4. EXPERIMENT AND RESULTS

4.1 Cross recurrence plots maps analysis

The Figure 3 shows the cross-recurrence plot obtained using the two-time series: the glottal signal of a normal voice vertical signal and the glottal signal corresponded to a subject presenting VFs polypes horizontal one. A visual inspection of

the CRP structure map denoted the independence of recurrence patterns between the two signals. The absence of parallels diagonal lines expresses the loss of periodicity caused by the presence of polyps in the vocal folds. The presence of vertical lines means that the recurrence behaviour of the signal is not stable in different steps over the time series. The homogeneity of point distribution deals to the stability of the two-time series. In the case of VFs paralysis, it is obvious that CRP structure corresponded to the normal voice follow a gross change in glottal signals (disappearance of diagonals lines) which is referenced to the fundamental frequency perturbation. In this situation, some vertical segments of signals appear in same phase space region in some time series. The white bands indicate a breathy, absent of voice or vocal loss produced by an excessive glottis closure in Figure 5.

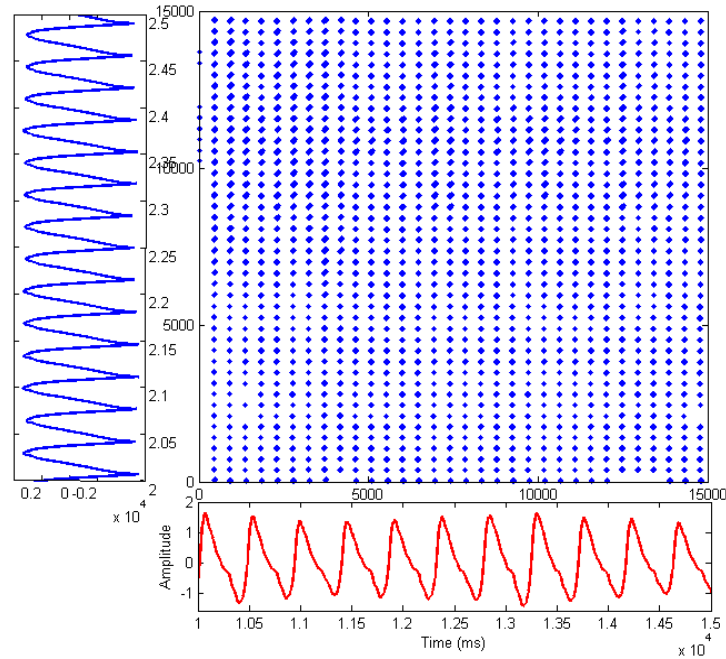


Figure 3. CRP maps Normal Vs VFs Polypes

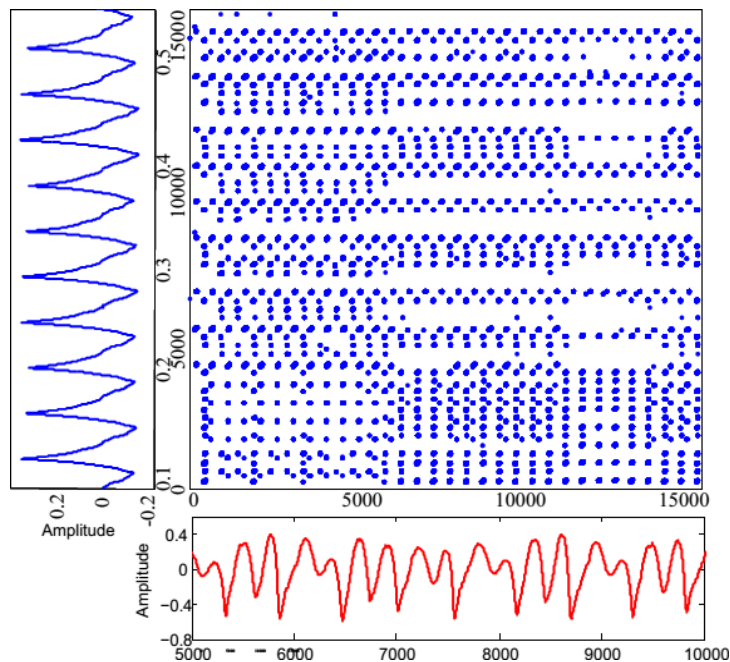


Figure 4. CRP maps Normal Vs Spasmodic dysphonia

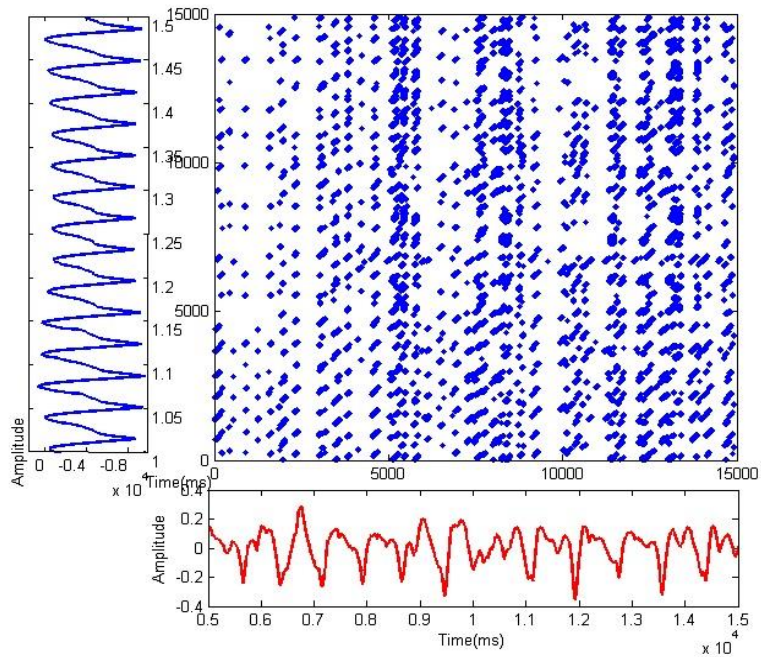


Figure 5. CRP maps Normal Vs VFs Paralysis

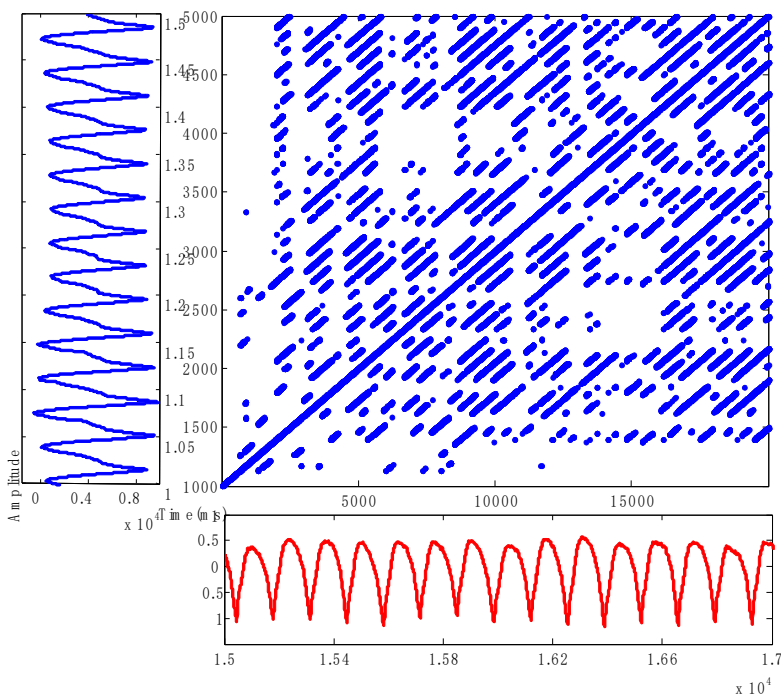


Figure 6. CRP maps Normal Vs VFs Nodules

In Figure 4, the CRP map of glottal signal between a Normal subject and that of a patient presented a spasmodic dysphonia shows a disappearance of longer diagonal lines, which indicate that both signals have any similar phase space dynamics. The distributions of recurrence points forming the lines along the vertical lines are spaced by a white band, this provides an information on voice production process. Its present a breathy vocal loss caused by the presence of pathology. In Figure 6, which illustrates the cross recurrence plot between the normal glottal signal and that of a patient presented a VFs Nodule, we have used a sample of vowel in an advanced state of vocal therapy in order to observe the changements in glottal signal form comparing to control voice (the normophonics). It's clear

through the figure, the occurrence of main diagonal line, which indicates the beginning of similarities between the two signals. We notice also discontinuous secondary diagonal lines. This one means the adjustment of fundamental frequency and intensity perturbation to normal voice.

4.2 Cross-RQA parameters analysis before and after voice therapy

The major purpose of voice therapy is to recover the normal automatisms of voice production behavior by decreasing the vocal effort and stress in order to establish an appropriate hearing feedback. Now days, ENT doctors, laryngologists and

speech therapists approve the importance of voice therapy and apply it to several voice disorders. The interest of this study is to take a care of dysphonic patients before any surgical intervention, by helping specialists in their diagnosis and decisions making. This study was enrolled on the vocal evolution of sixty (60) dysphonic adults patients (Males and Females) with dysfunctional symptomatology (VFs paralysis: n=13, Spasmodic dysphonia: n=13) or organics dysphonia (VFs Nodules: n=18, VFs polyps: n=16). n represents the number of patients (males and females). By applying the Cross Recurrence Quantification Analysis on the glottal signal waveform for the different groups of patients before and after rehabilitation and comparing these parameters to the control group (normal voice), the effect of voice therapy is clearly observed through a parametric analysis.

The obtained results are grouped in the Table 2. The table shows that all the values have been improved compared to those measured before the speech therapy operation according to the degree of disturbance of the voice. Before the voice therapy, we note that the lowest mean values of the parameters CRR, DET and L are recorded in the case of the pathologies: Nodule, Spasmodic Dysphonia and VFs Paralysis respectively. The lowest value of cross Laminarity, Lmax and TT are observed in the case of VFs paralysis and spasmodic

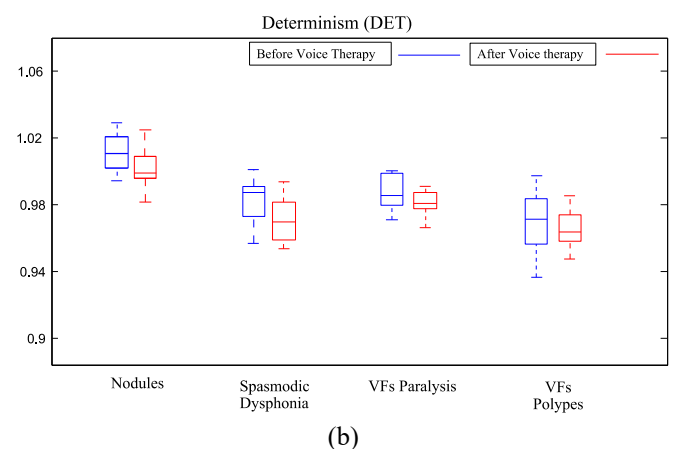
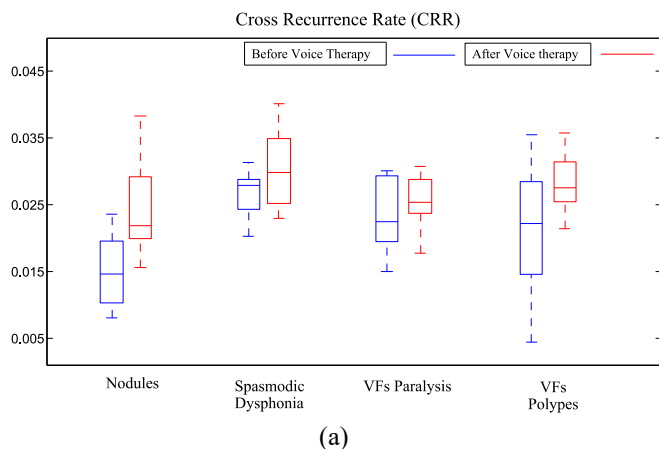
dysphonia. For ENTR we notice significant changes before and after voice therapy in all pathologies cases. After the operation voice therapy, the improvement of the voice is expressed by either an increase or a reduction of the value of the parameter according to the normative value specific to the control group (normophonic subjects). The CRR and cross ENTR were improved in the four groups of patients and a statically difference was mostly noticed in VFs polyps and VFs nodules.

4.3 Discussions

The Figure 7 shows the boxplots of the six (06) CRQA indicators for normal and pathological voices. This representation is a graphical tool available on MATLAB, allowing a visual analysis and comparison of data features sets to make facilitate decision-making. The median value of feature data set is represented by a central horizontal line in the boxplot. The 25 and 75% percentiles are illustrated by blue or red horizontal line, representing the first and third quartile of the boxplot. The whiskers represent the minimum and maximum value in the serial of features. Red Cross marks represent the current residual data set and outliers.

Table 2. Cross recurrence quantification parameters between patients with different voice disorders, before and after voice therapy (mean ± standard deviation)

Group		CRR	DET	L	LAM
Vocal Folds Polypes	Before voice therapy	0.0126 ± 0.0092	0.9906±0.0095	0.9727±0.0254	6.4511±3.7755
	After Voice therapy	0.0339 ± 0.0127	0.9985±0.0025	1.225±0.2101	9.3104±2.2512
Vocal folds Paralysis	Before voice therapy	0.0165 ± 0.0064	0.9716±0.0125	0.9245±0.1125	4.0673±1.2265
	After Voice therapy	0.0318 ± 0.0154	0.9926±0.0254	0.9993±0.0214	11.2366±3.1522
Sapsmodic dysphonia	Before voice therapy	0.0183 ± 0.0055	0.9476±0.0135	0.9622±0.1540	5.5239±2.7418
	After Voice therapy	0.0330 ± 0.0122	0.9976±0.0022	1.4332±0.2015	9.2254±2.1421
Nodules	Before voice therapy	0.0100±0.0052	0.9912±0.0192	0.9694±0.0120	7.1654±0.1655
	After Voice therapy	0.0303±0.0021	0.9965±0.0092	1.0226±0.1445	8.2600±1.2205
All patients	Before voice therapy	0.0740±0.0065	0.9752±0.0136	0.9572±0.1003	5.8019±1.9773
	After Voice therapy	0.0322±0.0106	0.9963±0.0098	1.1700±0.1925	9.5081±2.1915
		Lmax	TT	Vmax	ENTR
Vocal Folds Polypes	Before voice therapy	24889±5225	25.7472±6.0221	24±06	3.3926±1.5488
	After Voice therapy	25055±10225	27.4155±9.4877	30±05	4.0215±0.7658
Vocal folds Paralysis	Before voice therapy	15732±2665	18.3325±5.3325	18±02	2.7144±0.9894
	After Voice therapy	17544±2548	23.1447±11.2254	21±08	3.9755±1.2247
Sapsmodic dysphonia	Before voice therapy	13141±2900	21.2548±8.2554	21±04	3.1636±1.0788
	After Voice therapy	18362±0955	23.0995±16.6658	28±05	4.1250±1.9220
Nodules	Before voice therapy	19224±1540	27.2258±2.6598	22±15	3.6602±1.2450
	After Voice therapy	19100±0145	29.3695±7.1654	27±10	3.9544±1.2356
All patients	Before voice therapy	18246±3082	21.1400±5.5674	21±06	3.9827±1.2155
	After Voice therapy	18665±3468	27.2573±11.1360	26±07	4.0191±1.2870



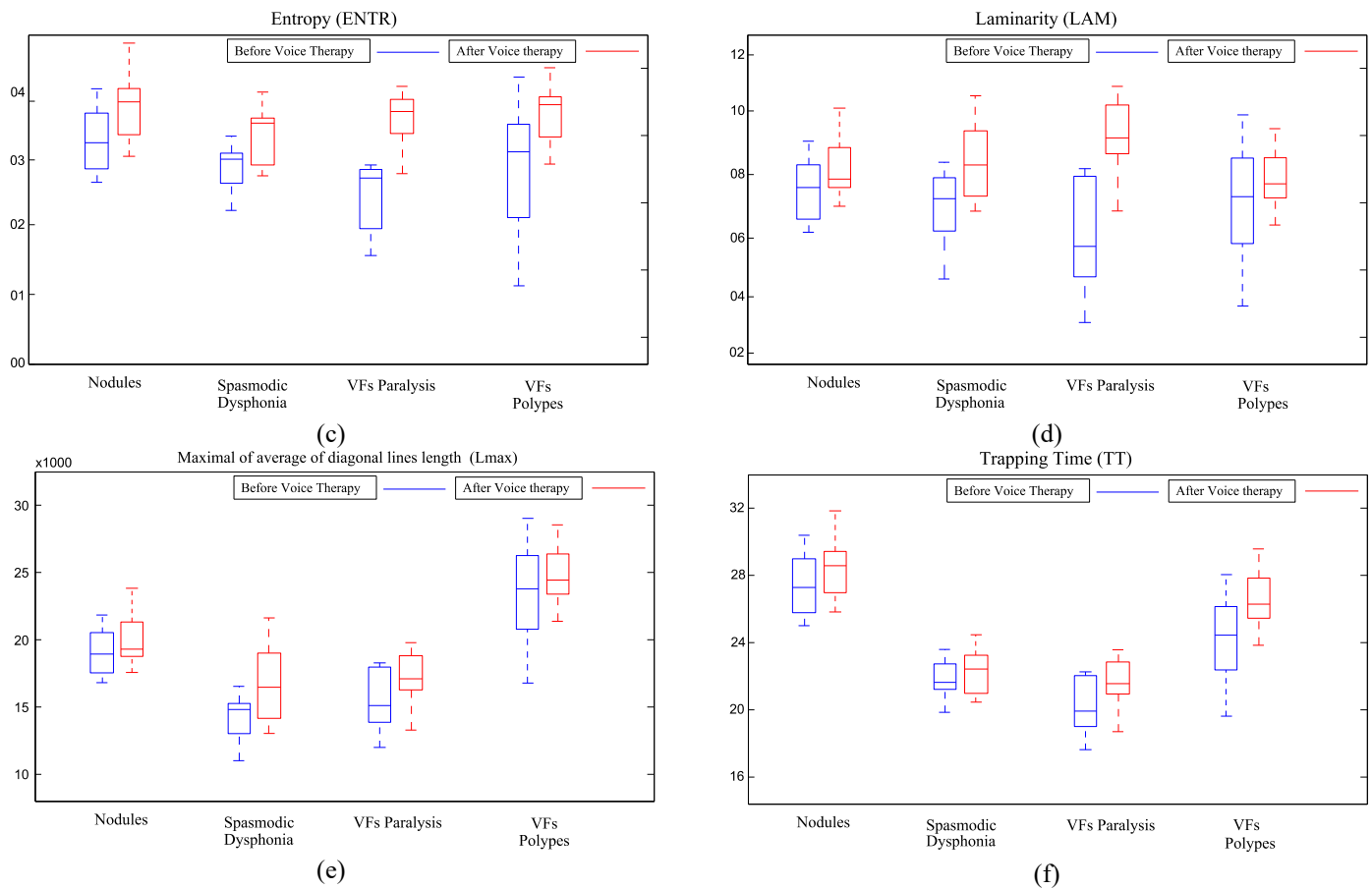


Figure 7. Boxplot data distribution of each kind of voice (Normal, Polyps, Vocal folds Paralysis, Spasmodic dysphonia) for each measurement extracted from the Cross Recurrence plot structure of the glottal signal [a:] vowel; (a): CRR, (b): DET, (c): ENTR, (d): Lmax, (e): LAM, (f): TT

The boxplots representation permits to interpret the results given in the Table 2, seeing that such representation allows visualizing the stage of parameters variation before and after voice therapy and detect the most pertinent measure.

In this study, there was significant differences in each parameter in all patient subgroups before and after voice therapy. This expresses the usefulness of the cross-RQA for such situation of VFs pathologies assessment through glottal the most pertinent parameter in the case of vocal folds nodules. The Determinism and longest diagonal line (L_{max}) are another two parameters used for description of voice quality improvement. generally, the determinism has a typical variation between 1 and 0.996 in normal voices. In the present study the percentage of improvement in the voice quality before and after voice therapy was 3.21% and 1.85% in the case of determinism and L_{max} respectively. Figure 7.b,e. signal. The lowest CRR value before voice therapy was in subjects presented VFs Nodules disease with a range of [0.07; 0.024]. However, after voice therapy, making patients relaxing and more confident speaking, a remarkable improvement in this parameter with a data distribution varied between [0.015, 0.038] in addition the mean value has been jumped from 0.015 to 0.027. We predict so that the CRR was the analysis of the data ranges provides that these two parameters are mostly discriminative in the case of vocal folds polyps, indeed the reduction of the ranges parameters achieve [24000 30000] after the therapy operation, which is closed to normal voice. In the others parameters a low variations was reported in the data sets, which related the small, sized of lesions nodules and polyps. Concerning the laminarity, in the Figure 7.d, before

the voice therapy the largest size data range was recorded in the case of vocal folds paralysis pathology with the lowest median values. However, the lowest data range was found in the case of VFs spasmodic dysphonia and vocal folds nodules diseases. After voice therapy, the box plot of patients with spasmodic dysphonia show a significant improvement (before: between: [0.45, 0.85] after: [7.00, 10.2]) this express the beneficial effect of voice therapy especially this disease which noticed no organic marks for states discrimination. Therefore, this study makes easier for voice monitoring during the treatments process. The cross entropy seems to be a good indicator for measuring voice quality improvement before and after therapy. Its pertinence is most significant in the two pathologies: VFs polyps and VFs paralysis. This parameter reported that normal voice possesses entropies values compromising between 3.50 and 4.50. for male adults and between 3.80 And 5.00 for females. In general, normal voices values are higher than pathological ones. Before voice therapy, the entropy shows a convergent mean values and different data distribution, the largest one is recorded in the case of vocal polyps Pathology with an interval between [1.23, 4.00]. After voice therapy, the values of entropies were improved in different voice disorders group particularly in patients presented VF polyps. We inferred that this indicator is very sensitive for VFs polypes pathology.

The performed study showed a promising result obtained by the application of the cross-RQA to a few numbers of pathologies. The main goal of this study was to see how these parameters could be adapted to the diagnosis of voice perturbation degree during the re-education process. A further

study with a large data has been considered. The data recording structure and parameters organization can be affected by the quality of instruments used, resulting in sampling bias so a precise set of instruments and protocol should be followed.

5. CONCLUSION

This paper presented a new approach for the assessment of the most common vocal fold(s) disorders namely vocal fold polyps, vocal fold paralysis, vocal fold nodules and spasmodic dysphonia. This approach used the Cross Recurrence Quantification Analysis technique applied the glottal signal to quantify and assess the articulation alterations of vocal folds caused by these pathologies. The Cross-RQA measurements offered an excellent tool to evaluate the voice impairments before and after voice therapy. The proposed solution allows also the discrimination the different stages during the protocol of reeducation. The results clearly demonstrate the effectiveness of the proposed approach for vocal folds pathologies assessment. Actually, our method has been developed for a few numbers of sample and it showed good results in the assessment of vocal folds pathologies. However, we have planned to extend our database to a large number of samples and others types of pathologies. For this, the authors are working in collaboration with the medical staff of Ear-Nose-Throat services in Algerian hospitals to collect a large number of subjects of different age and diseases. This not only creates a new database on voice disorders, but also allows the proposed method to be applied for examination by Ear-Nose-Throat clinicians and speech therapists.

REFERENCES

- [1] Do Amaral Catani, G.S., Hamerschmidt, R., Moreira, A.T., Timi, J.R.R., Wiemes, G.R. M., Ido, J., Macedo, E. (2016). Subjective and objective analyses of voice improvement after phonosurgery in professional voice users. *Med Probl Perform Art*, 31(1): 18-24. <https://dx.doi.org/10.21091/mppa.2016.1004>
- [2] Dejonckere, P.H., Bradley, P., Clemente, P., Cornut, G., Crevier-Buchman, L., Friedrich, G., Van De Heyning, P., Remacle, M., Woisard, V. (2001). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. *European Archives of Oto-Rhino-Laryngology*, 258(2): 77-82. <https://dx.doi.org/10.1007/s004050000299>
- [3] Doellinger, M. (2009). The next step in voice assessment: High-speed digital endoscopy and objective evaluation. *Current Bioinformatics*, 4(2): 101-111. <https://dx.doi.org/10.2174/157489309788184774>
- [4] Henríquez, P., Alonso, J.B., Ferrer, M.A., Travieso, C.M., Godino-Llorente, J.I., Díaz-de-María, F. (2009). Characterization of healthy and pathological voice through measures based on nonlinear dynamics. *IEEE Transactions on Audio, Speech, and Language Processing*, 17(6): 1186-1195. <https://dx.doi.org/10.1109/TASL.2009.2016734>
- [5] Lu, D., Chen, F., Yang, H., Yu, R., Zhou, Q., Zhang, X., Ren, J., Wang, H. (2018). Changes after voice therapy in acoustic voice analysis of Chinese patients with voice disorders. *Journal of Voice*, 32(3): 386-e1. <https://dx.doi.org/10.1016/j.jvoice.2017.05.005>
- [6] Dahmani, M., Guerti, M. (2017, May). Vocal folds pathologies classification using Naïve Bayes Networks. In 2017 6th International Conference on Systems and Control (ICSC), Batna, pp. 426-432. <https://dx.doi.org/10.1109/ICoSC.2017.7958686>
- [7] Al-Nasheri, A., Muhammad, G., Alsulaiman, M., Ali, Z., Malki, K.H., Mesallam, T.A., Ibrahim, M.F. (2017). Voice pathology detection and classification using auto-correlation and entropy features in different frequency regions. *IEEE Access*, 6: 6961-6974. <https://dx.doi.org/10.1109/ACCESS.2017.2696056>
- [8] Mendoza, L.A.F., Cataldo, E., Vellasco, M.M., Silva, M.A., Apolinário Jr, J.A. (2014). Classification of vocal aging using parameters extracted from the glottal signal. *Journal of Voice*, 28(5): 532-537. <https://dx.doi.org/10.1016/j.jvoice.2014.02.001>
- [9] Kohler, M., Vellasco, M.M., Cataldo, E. (2016). Analysis and classification of voice pathologies using glottal signal parameters. *Journal of Voice*, 30(5): 549-556. <https://dx.doi.org/10.1016/j.jvoice.2015.06.010>
- [10] Rangaprakash, D., Pradhan, N. (2014). Study of phase synchronization in multichannel seizure EEG using nonlinear recurrence measure. *Biomedical Signal Processing and Control*, 11: 114-122. <https://dx.doi.org/10.1016/j.bspc.2014.02.012>
- [11] Yang, H. (2010). Multiscale recurrence quantification analysis of spatial cardiac vectorcardiogram signals. *IEEE Transactions on Biomedical Engineering*, 58(2): 339-347. <https://dx.doi.org/10.1109/TBME.2010.2063704>
- [12] de A. Costa, W.C., Assis, F.M., Neto, B.G.A., Costa, S.C., Vieira, V.J.D. (2012). Pathological voice assessment by recurrence quantification analysis. 2012 ISSNIP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living (BRC), Manaus, pp. 1-6. <https://dx.doi.org/10.1109/BRC.2012.6222177>
- [13] Le Bot, O., Mars, J.I., Gervaise, C., Simard, Y. (2015, December). Cross recurrence plot analysis based method for tdoa estimation of underwater acoustic signals. In 2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Cancun, pp. 1-4. <https://dx.doi.org/10.1109/CAMSAP.2015.7465293>
- [14] Lancia, L., Fuchs, S., Tiede, M. (2014). Application of concepts from cross-recurrence analysis in speech production: An overview and comparison with other nonlinear methods. *Journal of Speech, Language, and Hearing Research*, 57(3): 718-733. [https://dx.doi.org/10.1044/1092-4388\(2013\)12-0223](https://dx.doi.org/10.1044/1092-4388(2013)12-0223)
- [15] Shalhaf, R., Behnam, H., Sleight, J.W., Steyn-Ross, D.A., Steyn-Ross, M.L. (2014). Frontal-temporal synchronization of EEG signals quantified by order patterns cross recurrence analysis during propofol anesthesia. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 23(3): 468-474. <https://dx.doi.org/10.1109/TNSRE.2014.2350537>
- [16] Alku, P. (1992). Glottal wave analysis with pitch synchronous iterative adaptive inverse filtering. *Speech Communication*, 11(2-3): 109-118. [https://dx.doi.org/10.1016/0167-6393\(92\)90005-R](https://dx.doi.org/10.1016/0167-6393(92)90005-R)
- [17] Muhammad, G., Alsulaiman, M., Ali, Z., Mesallam, T.A.,

- Farahat, M., Malki, K.H., Al-nasheri, A., Bencherif, M.A. (2017). Voice pathology detection using interlaced derivative pattern on glottal source excitation. *Biomedical signal Processing and Control*, 31: 156-164. <https://dx.doi.org/10.1016/j.bspc.2016.08.002>
- [18] Cabral, J.P., Richmond, K., Yamagishi, J., Renals, S. (2014). Glottal spectral separation for speech synthesis. *IEEE Journal of Selected Topics in Signal Processing*, 8(2): 195-208. <https://dx.doi.org/10.1109/JSTSP.2014.2307274>
- [19] Maizel, J.V., Lenk, R.P. (1981). Enhanced graphic matrix analysis of nucleic acid and protein sequences. *Proceedings of the National Academy of Sciences*, 78(12): 7665-7669. <http://dx.doi.org/10.1073/pnas.78.12.7665>
- [20] Zbilut, J.P., Giuliani, A., Webber Jr, C.L. (1998). Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. *Physics Letters A*, 246(1-2): 122-128. [http://dx.doi.org/10.1016/S0375-9601\(98\)00457-5](http://dx.doi.org/10.1016/S0375-9601(98)00457-5)
- [21] Kanakambaran, S., Sarathi, R., Srinivasan, B. (2017). Identification and localization of partial discharge in transformer insulation adopting cross recurrence plot analysis of acoustic signals detected using fiber Bragg gratings. *IEEE Transactions on Dielectrics and Electrical Insulation*, 24(3): 1773-1780. <http://dx.doi.org/10.1109/TDEI.2017.006407>
- [22] Marwan, N., Romano, M.C., Thiel, M., Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, 438(5-6): 237-329. <http://dx.doi.org/10.1016/j.physrep.2006.11.001>
- [23] Afsar, O., Tirnakli, U., Marwan, N. (2018). Recurrence Quantification Analysis at work: Quasi-periodicity based interpretation of gait force profiles for patients with Parkinson disease. *Scientific Reports*, 8(1): 1-12. <http://dx.doi.org/10.1038/s41598-018-27369-2>
- [24] Elias, J., Namboothiri, V.N. (2014). Cross-recurrence plot quantification analysis of input and output signals for the detection of chatter in turning. *Nonlinear Dynamics*, 76(1): 255-261. <http://dx.doi.org/10.1007/s11071-013-1124-0>
- [25] Webber Jr, C.L., Zbilut, J.P. (1994). Dynamical assessment of physiological systems and states using recurrence plot strategies. *Journal of Applied Physiology*, 76(2): 965-973. <http://dx.doi.org/10.1152/jappl.1994.76.2.965>