Recurrence Quantification Analysis of Glottal Signal as non Linear Tool for Pathological Voice Assessment and Classification

Mohamed Dahmani and Mhania Guerti
Laboratoire Signal et Communications, Ecole Nationale Polytechnique, Algiers, Algeria

Abstract: *Automatic detection and assessment of Vocal Folds pathologies using signal processing techniques knows an extensively challenge use in the voice or speech research community. This paper contributes the application of the Recurrence Quantification Analysis (RQA) to a glottal signal waveform in order to evaluate the dynamic process of Vocal Folds (VFs) for diagnosis and classify the voice disorders. The proposed solution starts by extracting the glottal signal waveform from the voice signal through an inverse filtering algorithm. In the next step, the parameters of RQA are determined via the Recurrent Plot (RP) structure of the glottal signal where the normal voice is considered as a reference. Finally, these parameters are used as input features set of a hybrid Particle Swarm Optimization-Support Vector Machines (PSO-SVM) algorithms to segregate between normal and pathological voices. For the test validation, we have adopted the collection of Saarbrucken Voice Database (SVD) where we have selected the long vowel /a:/ of 133 normal samples and 260 pathological samples uttered by four groups of subjects : persons having suffered from vocal folds paralysis, persons having vocal folds polyps, persons having spasmodic dysphonia and normal voices. The obtained results show the effectiveness of RQA applied to the glottal signal as a features extraction technique. Indeed, the PSO-SVM as a classification method presented an effective tool for assessment and diagnosis of pathological voices with an accuracy of 97.41%.*

Keywords: *Glottal Signal, Recurrence Quantification Analysis, Saarbrucken Voice Database, PSO-SVM, Pathological Voice Detection.*

Received December 2, 2018; accepted March 23, 2020
https://doi.org/10.34028/iajit/17/6/4

1. Introduction

Vocal Folds (VFs) present an important part in the human phonatory apparatus, allowing the production of voiced sounds in speech languages. Thus, the presence of pathologies causes significant changes in the phonation process. In the field pathological voice detection practitioners already have a set of measuring medical devices, more or less sophisticated such as electroglottography, Laryngoscopy and stroboscopy. 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 visual examination during phonation process [11]. The nonlinear analysis methods of dynamics can analyze the irregularities behaviors of voice production and may be important to differentiate the voice' impairments, classify voices in different degrees of modification [15].

To investigate the vibration patterns of VFs under pathologies, several methods and algorithms called features extraction techniques have been developed. Mel Frequency Cepstral Coefficients (MFCC) combined with fundamental frequency (f0), jitter and shimmer as features set for VFs pathologies diagnosis using naive bays networks [15]. Modulation spectrum metrics was performed in different frequency regions in order to find the best parameterization for disordered voice detection. The obtained results showed the effectiveness of these metrics in this context. However, its use in the automatic classification is advisable and should provide good results [22].

In a study proposed by AL-Nasheri *et al.* [5], a robust features extraction technique for assessment and classifying pathological voices was investigated. Entropy and maximum peak and their corresponding lag values are extracted by autocorrelation in voiced signal frames. These parameters are evaluated in different frequency regions in order to determine the most bands contributing in classification processes. SVM was used as classification method. The bests accuracies achieved were obtained using combined features. Parameters extracted from Glottal Signal (GS) provide various information, not only for VFs dysfunction diagnosis, but it conducts also to identify vocal aging and synthesizing [20]. These parameters (describe the articulator aspect of vocal fold obtained by an inverse filtering algorithm just after VFs and before vocal tract) are used as input feature vector for classification of VFs pathologies using Artificial Neural Network (ANN), Support Vector Machine (SVM), K-Nearest Neighbours (KNN) and Hidden Markov Model (HMM). The result proved that GS

parameters are more relevant than MFCC for separate healthy from non-healthy voices.

Recurrence Plot (RP) is an advanced graphical technique for nonlinear data analysis. Conceived firstly to analyse the non-linearity properties of dynamical systems, permit not only the visualization of recurrence behaviour of signal trajectories but even for quantifying data sets of signals with short length [16]. It could successfully attract more attention to be applied in diverse other research area, such as in economy to study financial data time [10], Analysis of power dynamic systems [26]. In the field of biomedical engineering, RP knew a wide use to analyze the behaviour of biomedical signals [2, 3] especially for of Electro Encephalo Gram (EEG) signal analysis. ECG signal for heart beat describing [7, 21]. Emotion recognition through voice signal analysis [13].

Recurrence Quantification Analysis (RQA) is a non-linear signal analysis tool able to quantify the activity of dynamical systems performed directly to RP structure without the need for system identification but only the form of the signal. Hence, RQA proved to be ideally applied in several numerous real-world systems encompassing a multitude of disciplines potentially in biomedical to study and analyze the physiology of signals [25]. To discriminate between healthy and non-healthy patients RQA of heart beat signal has been successfully applied for heart diseases diagnosis [27]. In the field of pathological voice detection, which is our interest, Washington et al., presented an analysis of speech signal to detect laryngeal pathologies using quantification measures of RP structure, Linear Discriminate Analysis (LDA) and Quadratic Discriminate Analysis (QDA) are used as classification methods [14]. The results show an important discriminative potential. Nevertheless, the application of RQA to a specific signal describing a physiological phenomenon such as glottal signal, should give better results. RQA as non linear powerful tool have been used to detect drowsiness state by using different classifier: SVM, KNN and naive bayes networks. The determinism was suggested as the most reasonable features to occur the drowsiness state. An accuracy of 92.9 % have been achieved [24].

This paper aims to develop a new computational methodology to detect and classify vocal folds pathologies. The proposed method based on RQA indicators of the glottal signals. These are first extracted from the voice signal by an inverse filtering algorithm. Therefore, RP technique is applied on the obtained glottal signal waveform to provide a qualitative overview of the glottal signal. From the RP structures, RQA is performed to quantify the variation and shift of the glottal signal. Eight RQA indicators are studied and their mean values are used as input features vectors to hybrid algorithm classifiers using SVM and Particle Swarm Optimization (PSO). The obtained results show that RQA seems to be a very effective method to diagnose and analyse vocal folds pathologies.

2. Database

Saarbruecken Voice Database (SVD) [9] was used to validate our Algorithm, it collects several recordings of normal, pathological voices and EGG signals of more than 2200 people with more than 70 types of pathologies. All samples are recorded in .wav format organized as follow:

- Vowels /a:/, /i:/, /u:/ pronounced at low, normal and high pitch.
- Sentence "Guten Morgen, wie geht es Ihnen?" (Hello, how are you?
- Electro glotto gram signal (EGG).

In our study, three types of pathological voice are investigated due it widespread. Also, their medical diagnosis knew a bit tricky to date. Spasmodic Dysphonia (SD) characterized by occasional breaks in voice caused by sudden involuntary movements of VFs, which perturb its normal vibration. VFs Paralysis (VFP) is a neurological disease that results by an immobility of one or both vocal folds causing a hoarse and diplophonic voice. Polyps is an organic disease affects often one Vocal fold in different sizes and forms producing vocal rough and fatigue. We have used normal sustained vowels /a:/ samples of around 1 and 3 seconds and sampled at 50 kHz with 16 bits of resolution. The number of samples and gender distribution are given in (Table 1).

Table 1. Details of the database used in this study.

	Subjects		**Range of age**		**Mean of age**	
	Male	Female	Male	Female	Male	Female
Normal	71	62	25-45	22-50	33.5	36
Polyps	45	30	26-50	25-45	38	35
VFP	57	40	25-40	20-38	32.5	29
SD	44	44	23-43	24-44	33	34

3. Proposed Method

The proposed approach consists mainly of the features extraction and classification method. To this end, the glottal signal was first extracted from the recorded samples of the vocal signal of the long vowel /a:/ using an Iterative Adaptive Inverse Filtering (IAIF) algorithm. RQA of RP was then applied to the glottal signal and RQA indicators were used as a feature extraction technique for the classification methods. The later was carried out using the Particle Swarm Optimization-Support Vector Machines (PSO-SVM) algorithm. In this combined algorithm, the SVM parameters were optimized using the PSO algorithm, providing a simple implementation and an easy to understand technique. SVM and KNN classification methods was employed to show the effectiveness of the proposed method. Unlike other techniques (KNN,

GA ...etc.,), this method, using test data set and optimized parameters (C; λ) , provided better accuracy of the pathological vocal folds recognition rate and investigated the prediction ability of the trained PSO-SVM model [1]. Figure 1 shows the flowchart of the proposed approach.

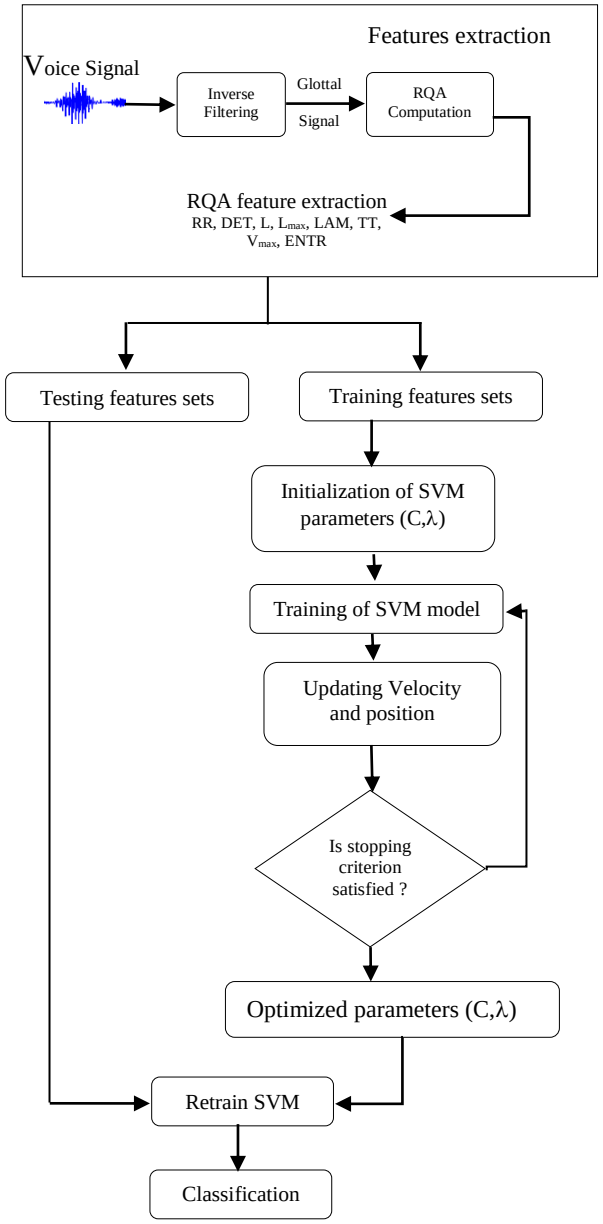


Figure 1. Flowchart of Proposed method based on RQA of Glottal signal.

3.1. Glottal Signal Extraction

Several algorithms and techniques exist for glottal signal estimation. In our study, we used the IAIF algorithm proposed firstly by Alku [4]. This method provides a robust estimation for the voice signal decompositions, which have been used for determination of any dysfunction in vibration patterns caused by pathological masses on the vocal folds [18, 19, 23]. In the IAIF algorithm, Voice signal is firstly filtered by a finite impulse response high pass filter. This permit to remove low frequency fluctuations. Linear Predictive Coding (LPC) eliminate vocal tract filter, which is cancelled out by an inverse filtering. By using the IAIF algorithm, the effect of vocal tract and lips is eliminated because in our case we concentrate on the glottal signal, which indicates the irregular vibration of vocal folds affected by pathologies, not the characteristics of vocal tract signal.

3.2. Recurrent Plots

Recurrent Plots constitutes a square matrix graphical tool for visualization of recurrence behaviour of states of dynamical systems. Especially useful for investigate the non-stationary property of high dimensional systems. Proposed firstly for visualizing patterns of genetic nucleotides by Maizel and Lenk [19]. It is represented by a symmetrical matrices N×N, in which elements $R_{i,j}$ are defined as below:

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

Where: ||.||: Euclidian distance and X is the studied time series, in our case it represents the glottal signal's time series. ε: neighborhood radius at the point X_i.

The Equation (1) means that if the distance between X_i and X_j is less than ε, $R_{i,j}$ =1: a dot is placed at (i, j) point, else $R_{i,j}$ =0 a blank is placed at (i, j) position. To analyze the glottal signal's form, the Table 2 summarizes the most common explanations of RP topologies.

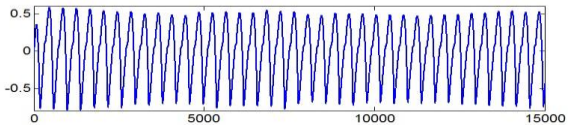
The figures below shows the structure of 15000 points calculated from a typical glottal flow waveform during production of vowel /a:/ of normal voice and patients presented Polyps, VFs polyps and spasmodic dysphonia diseases. The RP images composed of white space segments and concentrated recurrent points. The difference is visually distinguishable, so it may help doctors to take a decision, either the voice is normal of pathological. We use Table 2 to interpret RP.

Table 2. Interpretation of recurrent plots structures.

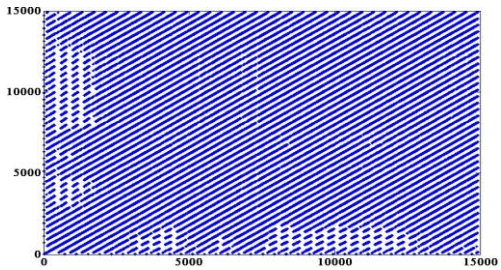
RP structure	Interpretation
Diagonal Parallel Lines	Deterministic Process, Periodic aspect
Dispersion Isolated Points	Anti-correlation process, weighty fluctuations
Periodic/quasi-periodic patterns	Long diagonal lines
White bands	Non-stationary, brusque transitions
Vertical and horizontal line	Laminar states, unchangeable states or slowly changements
Long bowed line	Dynamic of the system could change

- In the case of normal voice, recurrent plot structure shows the presence of parallel diagonal lines and homogeneity of black and white spots density, which indicate the periodicity and low variation in amplitude of glottal signal (Figure 2). However, in

(Figure 3) we note a loss of periodicity and decreasing of black spots is noticed. The white bands indicate a breathy or absent of voice or vocal loss produced by an excessive glottis closure, which characterize spasmodic dysphonia disease.

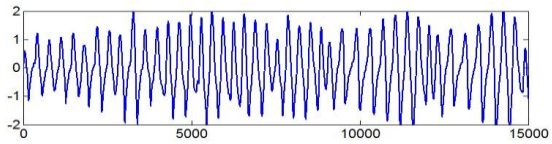


a) Glottal signal of a normal voice.

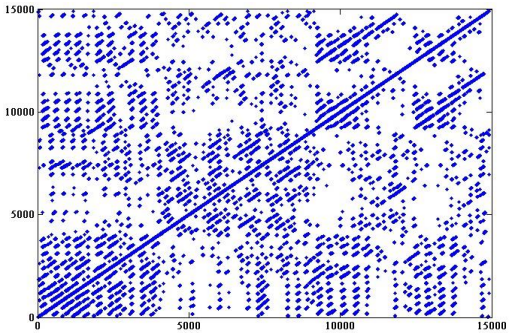


b) RP structure of a normal glottal signal.

Figure 2. Example of a glottal signal ant it RP structure for normal voice.

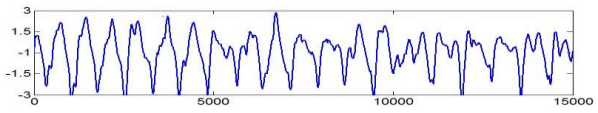


a) Glottal signal of a spasmodic dysphonia disease.

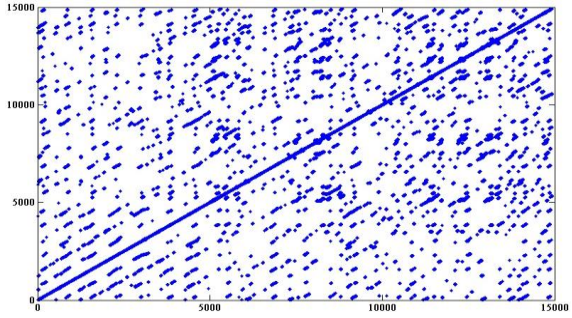


b) RP structure of a spasmodic dysphonia disease.

Figure 3. Example of a glottal signal ant it RP structure for SD disease.



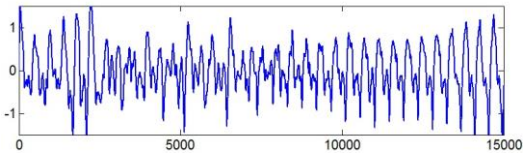
a) Glottal signal of a VFs Paralysis disease.



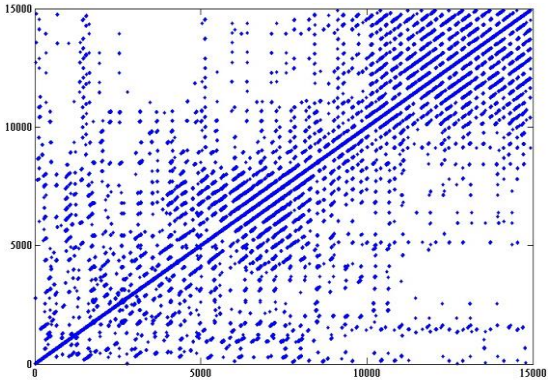
b) RP structure of a VFs paralysis disease.

Figure 4. Example of a glottal signal ant it RP structure for VFs paralysis.

- Figure 4 shows an example of glottal signal and it RP structure of VFs Paralysis disease. It is obvious a discontinuity of recurrent points in diagonal lines, which explain an extent of fluctuation in the amplitude of the signal. The highness in amplitude indicate that the patient exercise more force for vowel production. In the other hand in the case of vocal fold polyps (Figure 5), a predominance of isolated points is clear and loss of parallel diagonal lines resulting perturbation in fundamental frequency and disagreement in closing and opening instances of vocal folds.



a) Glottal signal of a VFs Polyps disease.



b) RP structure of a VFs polyps disease.

Figure 5. Example of a glottal signal ant it RP structure for VFs polyps.

3.3. Recurrence Quantification Analysis (RQA)

RQA is a non-linear data analysis permit to examine the recurrence patterns of a dynamic system, it was proposed firstly by Webber and Zbilut [26] to measure and visualize the recurrence plots points (R_{ij}) behaviors, also to examine the shifting of a given signal [8]. In our case, we quantify the characteristics RP structure of glottal signal of different types of voices signal normal and pathological.

3.3.1. Recurrent Plots Indicators

- Recurrence Rate (RR) represents the ration between all recurrent states of RP structure and all possible states in other word it is the probability of recurrence Equation (2).

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

- Determinism (DET) is represented by the proportion between number of recurrence points, forming diagonal structures and all R_{ij} points of *RP* form. High determinism values indicate long diagonals. The mathematical formula of

determinism is given by:

$$DET = \frac{\sum_{l=l_{min}}^{N} l \times P(l)}{\sum_{l=1}^{N} l \times P(l)} \quad (3)$$

Where: $P(l)$ represents the frequency distribution of the lengths of the diagonal structures in the RP. l_{min} is the minimum number of points to form a diagonal structure in the RP.

- Maximum diagonal line length (L_{max}) is corresponding to the longest diagonal line in RP Structure it's calculated by (Equation (4)):

$$L_{\max} = \max(\ l_i, i = 1,..., N_l) \quad (4)$$

Where: N_l correspond to the number of all diagonal lines present in RP structure.

- Average diagonal line length (MEANLINE); which represents the average time that two segments of the RP structure are close to each other. It can be interpreted as the mean prediction time of dynamical systems.

$$L = \frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{l=l_{min}}^{N} P(l)} \quad (5)$$

- Laminarity (LAM) quantifies the percentage of recurrent points of vertical lines over the entire RP. It represents the frequency of presence of laminar states in the system. Low Laminarity values indicate the presence of isolated recurrent points.

$$LAM = \frac{\sum_{v=v_{min}}^{N} vP(v)}{\sum_{v=1}^{N} vP(v)} \quad (6)$$

Where: $P(v)$ is the distribution of the vertical lines of length v. N is the length of the signal.

- Entropy (ENTR): Entropy quantifies the frequency distribution of the diagonal line length, it associated to the complexity of RP structure. The recurrence plots of a sine-wave shows a close distribution of diagonal lines, which result low values of ENTR.

$$ENTR = -\sum_{l=l_{min}}^{N} P(l) \ln P(l) \quad (7)$$

- Longest vertical line (V_{max}): as (L_{max}), (V_{max}) measures the longest vertical line in the RP structure. It is related to singular states where the system is stuck in a wait pattern, describing rectangles in the RP.

$$V_{\max} = \max\left(\{v_l\}_{l=1}^{N_v}\right) \quad (8)$$

- Trapping Time (TT) provides information about the amount and length of vertical structures, indicating the average time that the system will stand a specific state. It's given by the formula:

$$TT = \frac{\sum_{v=v_{min}}^{N} vP(v)}{\sum_{v=v_{min}}^{N} P(v)} \quad (9)$$

3.3.2. Discrimination of Measurements Analysis

Figure 6 shows the boxplots of the different RP measures for normal and pathological voices, it's a graphical tool on MATLAB allowing a visual analysis and comparison of features set data which make easier and effective decision making. The median value of feature data set is represented by red horizontal line in the boxplot, 25 and 75% percentiles are illustrated by bleu horizontal line, which represent the first and third quartile of the boxplot. The whiskers represent the minimum and maximum value in the serial of features. Red crosses mark represent the residual current data set and outliers.

- The RR, which quantifies recurrence points, shows a clear difference between the medians of Normal and pathological voices. Normal voice presents the high median value around 0.043 while the lower value is 0.012 in the case of spasmodic dysphonia disease, also Polyps and Vocal folds paralysis diseases show less values than normal voice, 0.015 and 0.018 respectively. This indicate that the tendency of recurrence point in RP structure decreases by the presence of pathologies. (Figure 6-a).
- The Determinism shows an approximate median values but we note that the whiskers maximum and minimum values present a remarkable difference so it can be used as discriminative indicators. (Figure 6-b).
- Most of the maximal length of the vertical lines (V_{max}) means values for participants across pathological and normal voices within a similar range. The median values are between 20 and 30, the uppermost values ranging from 80 and 70 in the case of normal voice and spasmodic dysphonia respectively. Figure 6-c).

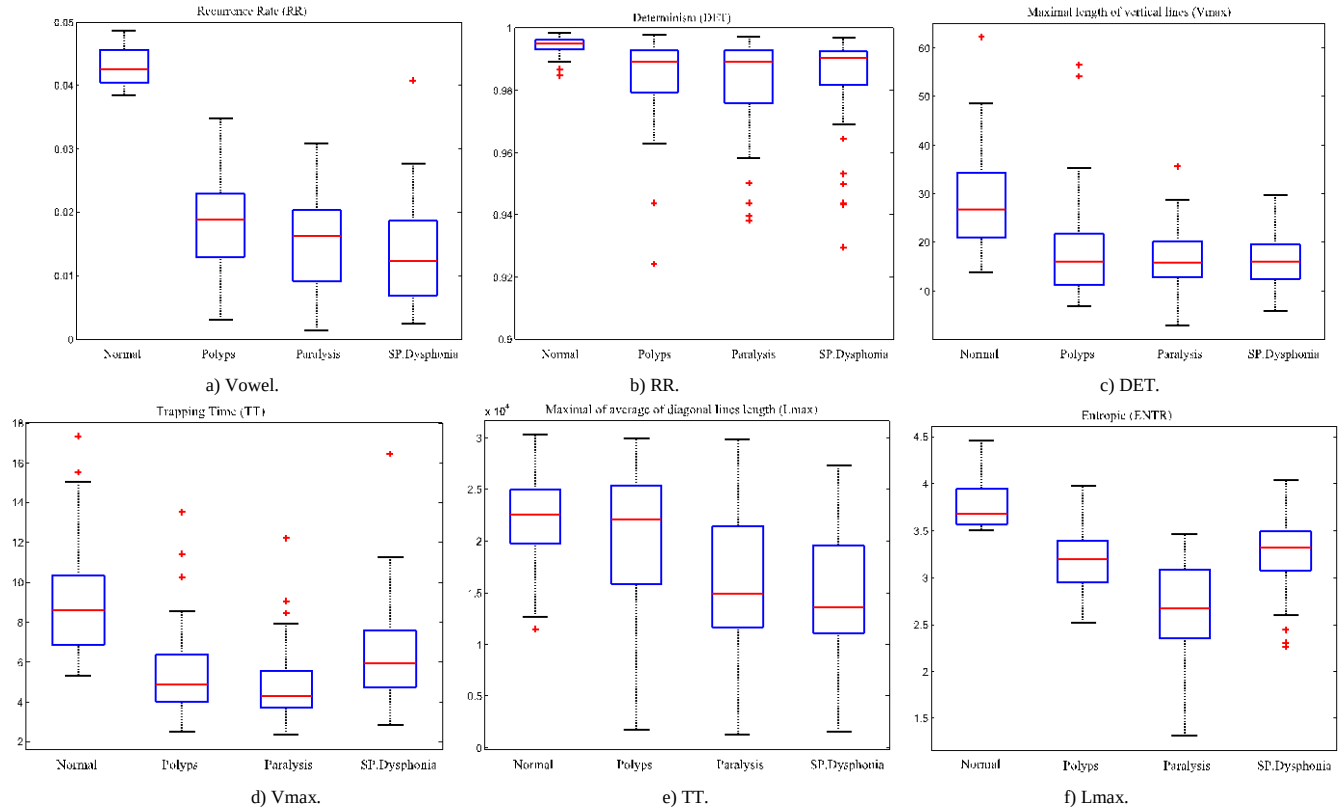


Figure 6. Boxplot data distribution of each kind of voice (Normal, Polyps, Vocal folds Paralysis, Spasmodic dysphonia) for each measurement extracted from Recurrence plot structure of glottal signal [a]: ENTR).

- Trapping Time (TT) have the smallest data distribution (min=2, max=8) and the lowest median value is 4 in the case VFs Paralysis disease. Normal voice present the larger range data distribution (min =5, max=15) and the highest median value nearly to 9, High values of TT indicate the absence or slight changes of glottal signal in time. Spasmodic dysphonia and polyps present second and third largest data range of TT parameter with 6 and 5 median value respectively. (Figure 6-d).

The L_{max}, average time that two segments of the trajectory on the RP are close to each other. Show a large variation in the case of Polyps and VFs paralysis, it varies between (0.2 and 3)*104 compared to those Normal and Polyps. The medians values are approximately the same for Normal and Polyps 2.25*104 and 1.5*104 for Paralysis and Spasmodic Dysphonia. Figure 6-e).

- Entropy (ENTR) is an important indicator parameter used for measuring the complexity of a given signal and periodicity behavior. As showing in (Figure 6-f) normal voice has the maximum data distribution (3.5 and 4.5) an mean value 3.75 this indicate the periodicity form of glottal signal wave form, however the mean values of others pathologies are less than Normal one, varying between 2.75 and 3.25 Which indicate the loss of periodicity by presence of pathologies. The lowest median value and minimum values are in the case of vocal folds paralysis 2.25 and 1.15 respectively. Figure 6-f).

After observing (Figure 6), the main conclusion that we can draw is: the medians of Normal and pathological voices differ with significance levels in all cases. This leads us to an application of an Optimized classification algorithm in order to discriminate between healthy and pathological voices.

3.4. Classification Method

A hybrid model PSO and SVM, PSO-SVM is proposed as an automated learning tool to optimize SVM parameters (kernel parameter "C" and regularization parameter "λ") and improve the accuracy of classification. The PSO is a popular Optimization method, inspired from the social behaviours of swarm birds or fish searching the optimal position to find food (BAI, 2010) [8]. The application of the PSO algorithm, as any others evolutionary algorithms, is influenced by some factors such as the stopping criterion or iteration number, the particle structure, and the objective function which is intended to minimize the error of generalization. This algorithm begins with a random initialization of the particles in the search space.

At each iteration, the PSO algorithm chose the best particle from the entire population. Each particle represents a candidate solution in the search space.

The algorithm of PSO-SVM algorithm is given as

follow:

- Initialize a Particles population with random positions and velocities. Initialize P_i and P_{ibest} respectively by starting position and starting fitness.
- For each particle, evaluate the fitness of the position Pi which means for each particle, the SVM classifier evaluates its performance using the cross-validation method. The fitness of the position P_i is the error of generalization calculated by cross validation method.
- Compare the fitness of the particles with the P_{ibest} .If the current value is better, copy it to P_{ibest} and replace P_i with the current position.
- Identify the most successful particle in the neighborhood and change its index to determine the global best position g_{ibest}.
- Make a loop in step 2, until the termination criterion is satisfied and the optimized parameters are obtained.
- Change the velocity and position of the particle according to the Equation:

$$V_i^d(t+1) = W V_i^d(t) + c_1 r_1(t)(pbest^d(t) - p_i^d(t)) - c_2 r_2(t)(gbest^d(t) - p_i^d(t)) \quad (10)$$

$$p_i^d(t+1) = p_i^d(t) + V_i^d(t+1) \quad (11)$$

Where: c_1, c_2 are positive constants, named acceleration coefficients $r_1(t)$ and $r_2(t)$ are random variables varied between 0 and 1.

3.5. Experimental Setup

Each sustained /a:/ vowel, which was taken from the SVD database correspond to normal and the three common considered pathologies in this study is used as input signal to an Iterative Adaptive Inverse Filtering (IAIF) to generate the corresponding glottal flow signal. This procedure has three fundamental parts: analysis, inverse filtering and integration. RP topological structure is obtained by plotting Equation (1) allowing detecting the non-stationary in time series. RQA offers multiple measures that help to quantify the recurrence pattern of RP diagram. Eight parameters are extracted from each sample, which represent the input features vector of the classifier. These vectors are organized as training and test sets. The training sets take 70% of the samples randomly from the dataset. For the test data set, we used other data remaining in the data set, which contained all the attributes that the model is supposed to predict.

In the classification phase, PSO-SVM is used as classifier. To obtain the optimal parameters values of SVM kernel parameter "C" and regularization parameter "λ". The program was running 20 times and the final accuracy was obtained by averaging accuracies of 20x10 instances (20 runs, and 10 iterations per run). For classification, the LIBSVM library [12], one-versus-one approach was utilized.

4. Classification Results

In order to assess the separability aspect of recurrence quantification features sets and classes distribution, 3 dimensions scatters plots are used, as showing in figures below. According to scatters figures we notice the presence of many overlap among recurrences parameters. Furthermore, in each scatter plot RQA mean values corresponding to training and test phases, features are round up in four groups related to normal and pathological voices distinguished. Such representation suggests that RQA of glottal signal parameters are directly related to vocal folds stats normal or pathological. These demonstrate the capability of these tools to differentiate normal vocal folds from pathological ones. Figures 7 and 8.

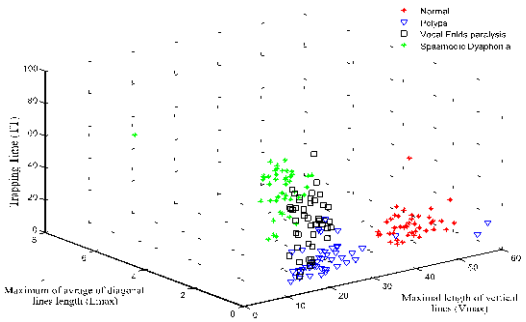


Figure 7. 3-D scatter plot to show the discriminative capability of the three features. TT, Lmax, Vmax.

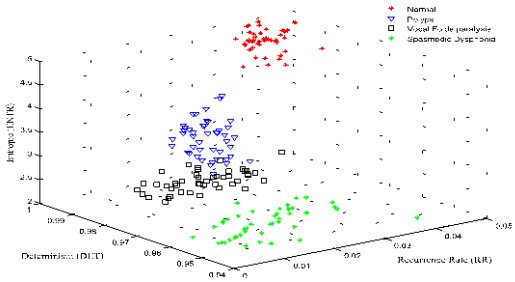


Figure 8. 3-D scatter plot to show discriminative capability of three Features ENTR, DET and RR.

4.1. Classification Results Using SVM and KNN

SVM and KNN are employed to evaluate and gauge the robustness of our prediction method. The reason of these choices goes back to its simplicity of implementation; it gives an excellent generalization and low computational cost. As it mention above SVM has been employed in different such problem and offered always bests accuracies. KNN requires no training and attributes the most common class corresponding to the maximum k nearest neighbors. Four types of distances are employed: Euclidian, city block, cosine and correlation, which allow to determine the nearest neighbors given the best accuracy.

The results of our performed described experiments for vocal folds classification are expressed in terms of

accuracy rate, which represent the ratio between correctly detected samples and the total number of samples. Table 3 resume the obtained accuracies for both methods used KNN and SVM. According to (Figure 9), the KNN Classifier used in our study demonstrate to be an effective tool and compromised for classification of vocal fold disorders selected. Cityblock distance gave the best accuracy, with neighbors equal to 106. The classification rate was around 95%.

Table 3. Accuracy and parameters value of the proposed method.

Method	KNN	SVM	SVM-PSO
Accuracy Rate (%)	95	96.21	**97.41**
Neighbors, Distance(KNN)	106/Cityblock	/	/
Optimized parameters (C,λ)	/	/	**(78.085416,8.333069)**

The (Table 3) depict the accuracy obtained by SVM. Correct decision and classification rate of 96.21 % have been achieved by taking 70 % of data for training phase and the rest 30 % for recognition phase using polynomial kernel function.

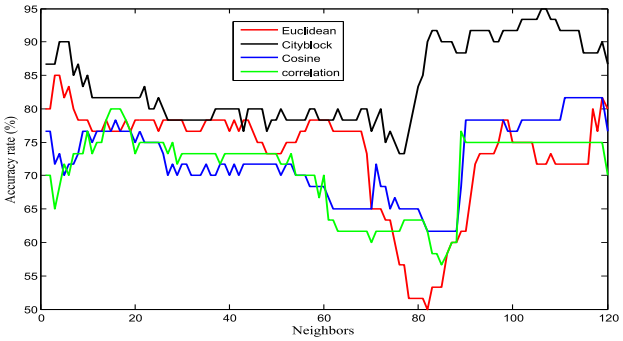


Figure 9. Classification accuracy of KNN for different types of distances.

4.2. Classification Method Using the Proposed Method

In fact, to achieve better classification results in term of accuracy it is necessary to conceive an algorithm in which the characteristics corresponded appropriately to the choice of feature extraction technique, this may uncover a critical research area by applying other diagnosis techniques to others types of pathologies. The Figure 10 show that our proposed method outperforms the KNN and SVM classifier in detection accuracy based on RQA features. The classification rate achieved the 97.41 %, which correspond to C and λ parameters 78.085416 and 8.333069 respectively. Consequently, the combination of RQA of glottal signal as discriminative features extraction technique and PSO-SVM algorithm using SVD database presented good performance.

To illustrate the classification results, we present in Table 4 the classification of only 12 samples chosen randomly from the test set. The classification using our PSO-SVM method was successfully done for these samples. However, even if KNN classifier provides three misclassifications and two in SVM, out of 12, it's seemed to be effective selected methods to evaluate the proposed classifier performance. The RR and ENTR appear excellent indicators parameters, as it is seen from the table these two parameters presented the greatest contribution for detection and classification.

In the present study, the application was performed on an opened access database and showed encouraging results. These motivates us to extend it big data framework for voice pathology assessment, seeing that this efficiency in terms of accuracy and time requirement in different studies [6, 17].

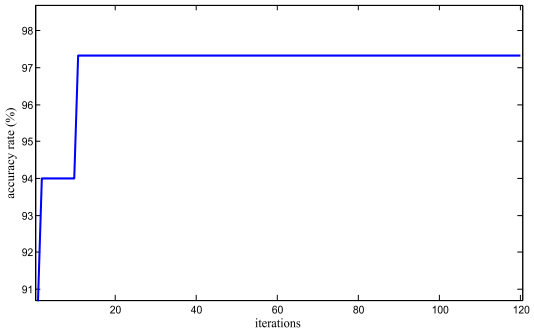


Figure 10. Classification accuracy using the proposed method PSO-SVM.

5. Conclusions

This paper presents an approach for diagnosing and classifying the vocal folds pathologies by RQA of glottal signal and PSO-SVM. The RQA is a more efficient tool to quantify the recurrence behaviour of dynamic system, with eight parameters that are extracted and analyzed in three types of vocal folds pathologies, where normal voice is considered as reference. Three dimensional scatter plots pointed out the compatibility between features extraction technique and glottal signal under vocal folds pathologies. Indeed, the four groups are clearly corresponding. Therefore, these indicators are proposed as an input feature set to an optimized classification problem, which is solved by means of PSO-SVM. The Results of the proposed approach reaches 97.41%, whereas 96.21% in the SVM and 95 % in the case of KNN classifier. Hence, the results demonstrate well that our developed approach is more suitable for vocal folds pathologies.

Table 4. Classification results.

Samples	RR	DET	L	L_{max}	LAM	TT	V_{max}	ENTR	Assigned class	Obtained class		
										KNN	SVM	PSO-SVM
1	0.0418	0.9962	0.9983	22383	8.7505	39.5522	28	4.7770	Normal	VFP*	Normal	Normal
2	0.0077	0.9500	0.9755	13367	5.5239	16.9547	21	2.9094	SD	SD	SD	SD
3	0.0091	0.9707	0.9245	25319	2.9518	12.9291	12	2.7167	VFP	VFP	VFP	VFP
4	0.0126	0.9906	0.9827	24889	6.4516	22.7472	24	3.3405	Polyps	Polyps	Polyps	Polyps
5	0.0187	0.9697	0.9002	15732	4.0673	14.2030	8	2.7194	VFP	VFP	VFP	VFP
6	0.0486	0.9976	0.9971	24897	12.3168	48.7418	36	4.4125	Normal	Normal	Normal	Normal
7	0.0095	0.9515	0.9622	13491	5.7926	16.1749	26	2.8147	SD	VFP*	Normal*	SD
8	0.0091	0.9707	0.7705	25319	2.9518	27.9291	19	3.7167	Polyps	Polyps	Polyps	Polyps
9	0.0439	0.9926	0.9982	17383	6.6825	41.2136	26	4.5951	Normal	Normal	Normal	Normal
10	0.0165	0.9716	0.8965	10020	8.6874	18.3325	10	2.7054	VFP	VFP	VFP	VFP
11	0.0100	0.9476	0.9900	16674	5.0394	14.4865	28	3.1647	SD	Normal*	Normal*	SD
12	0.0168	0.9901	0.9048	18288	4.8305	26.8312	15	3.3926	Polyps	Polyps	Polyps	Polyps

We can express our findings as follow:

- RQA of glottal signal measures present an effective tool as features set and appear very promising for representing the dynamics and articulation of vocal folds.
- Combining RQA measures as features set and PSO-SVM algorithm present good performances for vocal folds pathologies diagnosis and classification.

Currently, our methodology is applied only to SVD database with specific diseases due to its high incidence. However, it could be extended also to other diseases. For this reason, we plan to perform in the next studies, new experiments on a larger number of patients in order to generalize the obtained results. To achieve this aim, we are involved in collaboration with the medical staff of Ear-Nose-Throat service in Algerians hospitals to collect a large number of subjects of different age and diseases. This, permit not only to create a new voice disorders database, but also enforce our proposed method in the use for examination by Ear-Nose-Throat clinicians and speech therapists.

References

[1] Abbas N. and Kabudian J., "Speech Scrambling based on Independent Component Analysis and Particle Swarm Optimization," *The International Arab Journal of Information Technology*, vol. 14, no. 4, pp. 521-527, 2017.

[2] Acharya U., Fujita H., Sudarshan V., Bhat S., and Koh J., "Application of Entropies For Automated Diagnosis of Epilepsy Using EEG Signals: A Review," *Knowledge-Based Systems*, vol. 88, pp. 85-96, 2015.

[3] Acharya U., Sree S., Chattopadhyay S., Yu W., and Ang A., "Application of Recurrence Quantification Analysis for The Automated Identification of Epileptic EEG Signals," *International Journal of Neural Systems*, vol. 21, no. 03, pp. 199-211, 2011.

[4] Alku P., "Glottal Wave Analysis with Pitch Synchronous Iterative Adaptive Inverse Filtering," *Speech Communication*, vol. 11, no. 2-3, pp. 109-118, 1992.

[5] Al-Nasheri A., Muhammad G., Alsulaiman M., Ali Z., Malki K., Mesallam T., and Ibrahim M., **"**Voice Pathology Detection and Classification Using Auto-Correlation and Entropy Features Indifferent Frequency Regions," *IEEE Access*, vol. 6, pp. 6961-6974, 2018.

[6] Andreu-Perez J., Poon C., Merrifield R., Wong S., and Yang G., "Big Data for Health," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 4, pp. 1193-1208, 2015.

[7] Bahari F. and Janghorbani A., "EEG-Based Emotion Recognition Using Recurrence Plot Analysis and K Nearest Neighbour Classifier," *in Proceedings of IEEE 20th Conference in: Biomedical Engineering (ICBME)*, Tehran, pp. 228-233, 2013.

[8] Bai Q., "Analysis of Particle Swarm Optimization Algorithm," *Computer and Information Science*, vol. 3, no. 1, pp. 180-184, 2010.

[9] Barry W. and Putzer M., "Saarbrucken Voice Database," Institute of Phonetics, University of Saarland, http://www.stimmdatenbank.coli.uni-saarland.de/, Last Visited, 2018.

[10] Belaire-Franch J., Bayarri D., and Tordera-Lledo L., "Assessing Nonlinear Structures in Real Exchange Rates Using Recurrence Plot Strategies," *Physica D: Nonlinear Phenomena*, vol. 171, no. 4, pp. 249-264, 2002.

[11] Catani G., Hamerschmidt R., Moreira A., Timi j., Wiemes G., Ido J., and Macedo E., "Subjective and Objective Analyses of Voice Improvement After Phonosurgery in

Professional Voice Users," *Med Probl Perform Art*, vol. 31, no. 1, pp. 18-24, 2016.
[12] Chang C. and Lin C., "Libsvm: a Library for Support Vector Machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 3-27, 2011.
[13] Chao L., Tao J., Yang M., Li Y., and Wen Z., "Long Short Term Memory Recurrent Neural Network Based Multimodal Dimensional Emotion Recognition," *in Proceedings of 5th International Workshop on Audio/Visual Emotion Challenge*, Brisbane, pp. 65-72, 2015.
[14] Costa W., Assis F., Neto B., Costa S., and Vieira V., "Pathological Voice Classification Based on Recurrence Quantification Measures," *in Proceedings of ISSNIP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living (BRC)*, Manaus, 2012.
[15] Dahmani M. and Guerti M., "Vocal Folds Pathologies Classification Using Naive Bayes Networks," *in Proceedings of IEEE 6th International Conference on Systems and Control*, Batna, pp. 426-432, 2017.
[16] Eckmann J., Kamphorst S., and Ruelle D., "Recurrence Plots of Dynamical Systems," *EPL (Europhysics Letters*), vol. 4, no. 9, pp. 973-1987, 1987.
[17] Hossain M. and Muhammad G., "Healthcare Big Data Voice Pathology Assessment Framework," *IEEE Access*, vol. 4, pp. 7806-7815, 2016.
[18] Kennedy J. and Eberhart R., "Particle Swarm Optimization," *in Proceedings of IEEE International Conference on Neural Networks*, Perth, pp. 1942-1948, 1995.
[19] Maizel J. and Lenk R., "Enhanced Graphic Matrix Analysis of Nucleic Acid and Protein Sequences," *Proceedings of the National Academy of Sciences*, vol. 78, no. 12, pp. 7665-7669, 1981.
[20] Mendoza L., Cataldo E., Vellasco M., Silva M., and Apolinàrio J., "Classification of Vocal Aging Using Parameters Extracted from the Glottal Signal," *Journal of Voice*, vol. 28, no. 5, pp. 532-537, 2014.
[21] Mewett D., Reynolds K., and Nazeran H., "Reducing Power Line Interference in Digitised Electromyogram Recordings by Spectrum Interpolation," *Medical and Biological Engineering and Computing*, vol. 42, no. 4, pp. 524-531, 2004.
[22] Moro-Velázquez L., Gómez-García J., and Godino-Llorente J., "Voice Pathology Detection Using Modulation Spectrum-Optimized Metrics," *Frontiers in Bioengineering and Biotechnology*, vol. 4, pp. 1-13, 2016.
[23] Muhammad G., Alsulaiman M., Ali Z., Mesallam T., Farahat M., Malki K., Al-Nasheri A., and Bencherif M., "Voice Pathology Detection using Interlaced Derivative Pattern on Glottal Source Excitation," *Biomedical Signal Processing and Control*, vol. 31, pp. 156-164, 2017.
[24] Rangaprakash D. and Pradhan N., "Study of Phase Synchronization in Multichannel Seizure EEG Using Nonlinear Recurrence Measure," *Biomedical Signal Processing and Control*, vol. 11, pp. 114-122, 2014.
[25] Shabani H., Mikaili M., and Noori S., "Assessment of Recurrence Quantification Analysis (RQA) of EEG for Development of a Novel Drowsiness Detection System," *The Korean Society of Medical and Biological Engineering and Springer*, vol. 2, pp. 196-204, 2016.
[26] Webber C. and Zbilut J., "Recurrence Quantification Analysis of Nonlinear Dynamical Systems," *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences*, vol. 12, pp. 26-94, 2005.
[27] Yang H., "Multiscale Recurrence Quantification Analysis of Spatial Cardiac Vector Cardiogram Signals," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 339-347, 2011.

Mohamed Dahmani received the degree of Engineer and Master's degree in electronics engineering from Ecole Nationale Polytechnique of Algiers in 2014. He is currently a PhD student at the Electronics Department, Signal and communication laboratory in the same School. His research areas include signal processing, automatic classification and assessment of pathological voice.

Mhania Guerti is currently a Professor in the Department of Electronics at National Polytechnic School of Algiers (Algeria). She received her MSc in 1984, from the ILP Algiers in collaboration with the CNET-Lannion (France). She got her PhD from ICP-INPG (Grenoble France), in 1993. She is specialized in Speech and Language Processing. He has supervised a lot of students in Master and PhD degree. Her current research interests include the areas of speech processing, acoustics and audiovisual systems.