10th Brazilian Congress on Computational Intelligence (CBIC'2011), November 8 to 11, 2011, Fortaleza, Ceará Brazil

Classification of Voice Pathologies Using Glottal Signal Parameters

Manoela Kohler[1], Leonardo Alfredo Forero Mendoza[1], Juan G. Lazo [1], Marley Vellasco[1] Edson Cataldo[2]

[1]Electrical Engineering Department - Pontifical Catholic University of Rio de Janeiro (PUC-Rio)
CEP 22.451-900 Rua Marquês de São Vicente, 225 Gávea - Rio de Janeiro, Brasil
{manoela, mendonza, juan, marley}@ele.puc-rio.br

[2]Applied Mathematics Department - Telecommunications Engineering - Universidade Federal Fluminense (UFF)
CEP 24210-240 Escola de Engenharia - Bloco D - Sala 502-B R. Passo da Pátria, 156, São Domingos Niterói, RJ Brasil
ecataldo@im.uff.br

Abstract – The classification of voice diseases has many applications in health, disease treatment, and the projection of new medical equipments for diagnosing these pathologies. This work uses the parameters of the glottal signal that are more likely to identify two types of voice disorders: vocal cord nodule and unilateral paralysis of vocal cords. The parameters of the glottal signal are obtained through the known inverse filtering method. The parameters of the glottal signal serve as input to a neural network that classifies into three different groups of speakers: speakers with pathology nodule on one's vocal cords; with unilateral vocal cord paralysis; and finally speakers with normal voices. The database is composed of 248 voice recordings containing samples of the three groups mentioned. In this study we have used a larger database for the classification compared with similar studies, and its classification rate is superior to other studies, reaching 95.83%.

Keywords – classification of diseases, parameters of the glottal signal, neural network.

1 Introduction

The most common pathologies of the vocal cords are vocal cord nodules and unilateral paralysis of vocal cords. The most common method for extracting voice features is directly from the voice signal [1]. However, many researchers have looked for some characteristics extracted from the so called glottal signal, which is the signal obtained just after the vocal folds and before the vocal tract. As most voice disorders are due to some discomfort on the vocal cords, it is best to work with parameters extracted from the glottal signal, since the signal is produced by the vocal cords. Nowadays, obtaining this signal is easier due to the development of algorithms that can perform an inverse filtering from the voice signal, eliminating the influence of the vocal tract. Before, it used to be necessary to use equipment coupled with micro cameras to record sounds just after air passed from the vocal folds. This was an invasive technique and very difficult to be performed.

In previous works, different methods have been used to classify diseases, such as Bayes, Hidden Mar*ko*v Models (HMM) [2], Gaussian Mixture Models (GMM) and neural networks [3], using as inputs- Mel-frequency cepstral Coefficients (MFCC) and the parameters such as *jitter* and *shimmer*. In [4] [5], 37 input parameters have been used to classify a database composed of 12 recordings for men and women, resulting in a maximum performance of 80% accuracy [4]. MFCC have also been proved to be effective in speaker recognition problems [6]. However, their performance is not as effective in the classification of voice pathologies. In [5], several models for the classification of voice pathologies are compared. The best performance has been provided by a neural network based model, differing from speaker recognition applications where best results are usually obtained with GMM and HMM. This is probably because classification of voice pathologies does not fully depend on temporal features of the voice, and the pathology causes change in the voice signal [4]. Therefore, the main objective of this work was to evaluate the performance of a voice pathologies classification model based only on parameters extracted from the glottal signal. Additionally, a new database was created, with a larger number of voice recordings, which allows a better evaluation of the influence of each parameter in the classification performance.

This paper is organized as follows: Section 2 explains how the glottal signal is obtained and the extracted signal features used in this work. Section 3 presents the neural network model used to classify voice pathologies. Section 4 presents the database details and Section 5 present results, analysis and discusses the conclusions of this work.

2 Feature Extraction from the Glottal Signal

2.1 The Glottal Signal

The voice signal, particularly the one related to voiced sounds, e.g. vowels, starts with the contraction-expansion of the lungs, generating a pressure difference between the air in the lungs and the air near the mouth. The airflow created passes through the vocal folds, which oscillate in a frequency called the fundamental frequency of the voice. This oscillation modifies the airflow coming from the lungs, changing it into air pulses. The pressure signal formed by the air pulses is quasi-periodic and it is called the glottal signal [13].

2.2 Parameters Obtained from the Glottal Signal

As mentioned before, to obtain the glottal signal it used to be necessary to perform an invasive process, by installing micro cameras to record sounds just after the air passed through the vocal folds. Nowadays, it is possible to obtain the glottal signal using noninvasive methods, by performing an inverse filtering on the voice signal, which consists of eliminating the influence of the vocal tract and the voice radiation caused by the mouth, preserving the glottal signal characteristics [7]. Algorithms that obtain the glottal signal can be classified into two categories: semi-automatic and manual. In this paper, the inverse filtering algorithm used is of the semi-automatic category, called PSIAIF (Pitch Synchronous Iterative Adaptive Inverse Filtering) [8] [9]. The PSIAIF Algorithm was chosen due to its high performance and ease development. There is a toolbox implementation in Matlab, called Aparat [10], which was constructed especially based on the PSIAIF method to obtain the glottal signal and to extract its main features or parameters. The parameters which can be extracted from the glottal signal can be divided into three groups: time domain, frequency domain, and the ones that represent the variations of the fundamental frequency [8].

2.2.1 Time Domain Parameters

The time domain parameters that can be extracted from the glottal signal are described below [8], [9].

(i) Closing phase (*Ko*): it describes the interval between the instant of the maximum opening of the vocal folds and the instant where they close [8], *Ko* is shown in Figure 1.

(ii) Opening phase (*Ka*): it describes the interval between the instant where the vocal folds start the oscillation up to their maximum opening [8], *Ka* is shown in Figure 1.

(iii) Opening quotient (*OQ*): The ratio between the total time of the vocal folds opening and the total time of a cycle (or period) of the glottal signal (T). It is inversely proportional to the intensity of the voice. The smaller it is, the higher the voice intensity [1][8].

(iv) Closing quotient (*CIQ*): The ratio between the closing phase parameter (*Ko*) and the total length of a glottal pulse (T) [8]. It is inversely proportional to the voice intensity. The smaller it is, the higher the voice intensity.

(v) Amplitude quotient (*AQ*): The ratio between the glottal signal amplitude (Av) and the minimum value of the glottal signal derivative (dmin) [11]. It is related to the speaker phonation [9].

(vi) Normalized amplitude quotient (*NAQ*): It is calculated by the ratio between the amplitude quotient (*AQ*) and the total time length of the glottal pulse (T) [12].

(vii) Opening quotient defined by the Liljencrants-Fant model (*OQ*a): This is another opening quotient but calculated by the Liljencrants-Fant model for inverse filtering. Details about this model can be found in [13].

(viii) Quasi opening quotient (QoQ): It is the relationship between the glottal signal opening at the exact instant of the oscillation and the closing time [9]. It has been used in some works to classify emotions [14].

(ix) Speed quotient (SQ): defined as the ratio of the opening phase length to the closing phase length [8].

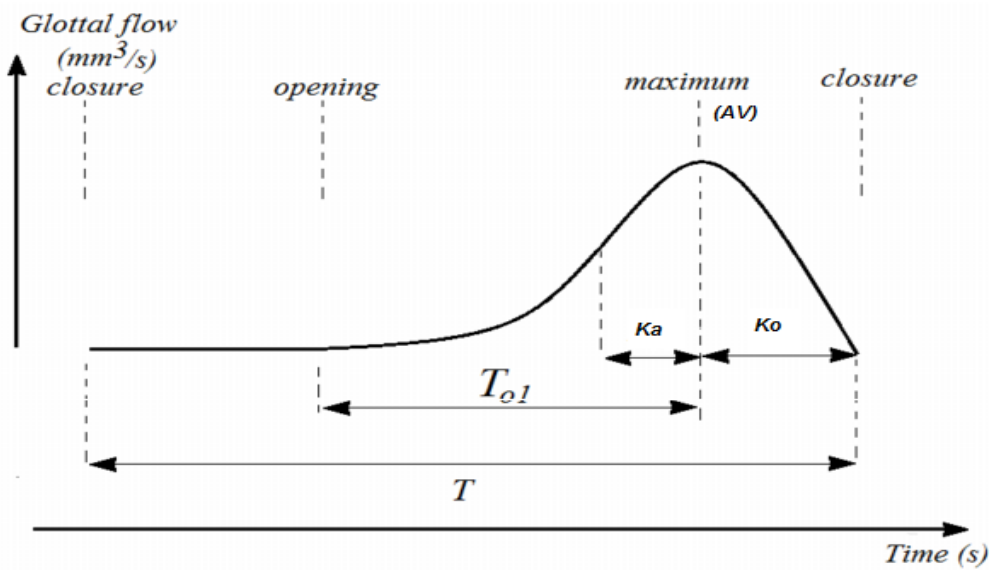


Figure 1 – Parameters *Ko* e *Ka* obtained from the glottal signal.

2.2.2 Frequency Domain Parameters

(i) Difference between harmonics (*DH*12): Also known as H1-H2 and it is the difference between the values of the first and second harmonics of the glottal signal [15][16]. This parameter has been used to measure vocal quality.

(ii) Harmonics relation factor (*HRF*): It relates the first harmonic (H1) with the sum of the energy of the other harmonics (Hk) [17]. It has also been used to measure vocal quality.

2.2.3 Parameters that Represent Variations and Perturbations in the Fundamental Frequency

(i) *Jitter*: variations in fundamental frequency between successive vibratory cycles [18] [19]. Changes in *jitter* may be indicative of neurological or psychological difficulties [1].

(ii) *Shimmer*: variations in amplitude of the glottal flow between successive vibratory cycles [18] [19]. Changing the *shimmer* is found mainly in the presence of mass lesions in the vocal folds, such as polyps, edema, or carcinomas [1].

3 Model Used for Voice Pathologies Classification

The model has two stages, the first stage is to obtain the abovementioned parameters from the glottal signal, and the second stage is the classification of pathologies of the voice by using a multilayer perceptron type neural network. The steps are shown in Figure 2.

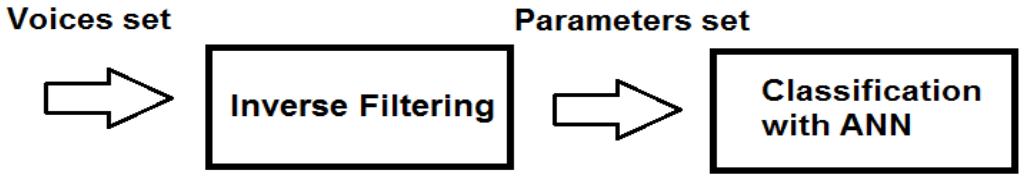


Figure 2 – Two stages of model

3.1 Inverse Filtering

For each vocal register, the corresponding glottal signal was obtained by inverse filtering PSIAIF and the parameters were extracted using the Aparat [10] and Praat [21] software. The following parameters were obtained: fundamental frequency (fo), *jitter*, *shimmer*, *Ko*, *Ka*, *NAQ*, *AQ*, *CIQ*, *OQ*1, *OQ*2, *Oq*a, Qoq, SQ1, SQ2, *DH*12, and *HRF*. The parameters were separated according to the groups to which they belonged. *OQ* was divided into *OQ*1 and *OQ*2, the open quotients calculated from the so-called primary and secondary openings of the glottal flow. The difference between *OQ*1 and *OQ*2 is that *OQ*1 is calculated from the closure of the glottal flow until the closure of the next glottal flow, and *OQ*2 is calculated from de opening until the closure of the glottal flow; SQ, as well, was divided into speed quotients calculated from the primary and secondary openings of glottal signal.

3.2 Classification of pathologies of the voice Using Neural Networks

With the parameters of the glottal signal, pathologies are classified with a neural network type (MLP) [18]. The neural network classification model is developed to classify the speaker into the three groups: speakers with nodule on the vocal cords; speakers with vocal cords paralysis; and speakers with normal voices. These groups will be the outputs of the neural network in which 70% of the database was used for training, 20% as the validation set (to avoid the network overtraining and to choose the number of processors and the hidden layers), and 10% for testing. For this neural network model was chosen early stopping to avoid over training the neural network, with this purpose we used the validation set.

4 Database

Most of the works on disease classification just classify speakers into two groups: speakers with disease (all kinds of disease) and speakers with normal voices [2], [3], [4], and [20]. The major difference on this work is the fact that the type of disease is classified, letting the patient knows if he/she has nodule or paralysis on the vocal cords, or neither one.

The speakers that belong to the pathology groups (nodule and paralysis) have different types of the disease, as show in Tables 1, 2.

The developed database is composed of 249 records consisting of voices of both women and men, with different ages, and it is divided into three groups: 12 speakers with nodule on the vocal cords; 8 speakers with vocal cords paralysis; and 11 speakers with normal voices. Eight voice records were taken from each speaker. This database was obtained randomly in Rio de Janeiro among people in treatment. The following tables describe the speakers.

Table 1 – Speakers with Nodule on the Vocal Cords (F – Female, M – Male).

Speaker	Gender	Age	Description of the Disease
Speaker 1	*F*	*42*	*Bilateral nodules and small slit*
Speaker 2	*F*	*38*	*Bilateral nodules with mid-posterior slit*

Speaker 3	*F*	*24*	*Vocal nodules with moderate and severe cleft significant anterior and posterior*
Speaker 4	*F*	*53*	*Vocal nodules slotted hourglass*
Speaker 5	*F*	*53*	*Vocal nodules with fissure*
Speaker 6	*F*	*38*	*Bilateral nodules with mid-posterior slit*
Speaker 7	*F*	*34*	*Bilateral nodules with mid-posterior slit*
Speaker 8	*F*	*32*	*Fibrous nodules - mid-posterior slit - great vocal effort*
Speaker 9	*F*	*29*	*Bilateral nodules with mid-posterior slit*
Speaker 10	*F*	*33*	*Vocal nodules with fissure*
Speaker 11	*F*	*28*	*Vocal nodules with slight fissure*
Speaker 12	*F*	*28*	*Bilateral nodules with mid-posterior slit*

Table 2 – Speakers with Vocal Cords Paralysis (F – Female, M – Male).

Speaker	**Gender**	**Age**	**Description of the Disease**
Speaker 13	*M*	*50*	*Paralysis of right vocal fold with scar retraction in the middle 1/3 - cleft anterior fusiform sequel laryngeal trauma*
Speaker 14	*M*	*50*	*Idiopathic palsy of the right hemilarynx with mild bowing of the free edge*
Speaker 15	*M*	*24*	*Vocal cord paralysis with right fusiform slit*
Speaker 16	*F*	*69*	*Right vocal cord paralysis in paramedian position with a slight bend and a slight cleft spindle - paralytic falsetto*
Speaker 17	*F*	*45*	*Vocal cord paralysis in the left median position and para-median*
Speaker 18	*F*	*43*	*Palsy and idiopathic right hemilarynx positioned at the median*
Speaker 19	*M*	*66*	*Pearrencial palsy with moderate bend to the left of the free edge (trauma of intubation)*
Speaker 20	*M*	*53*	*Right vocal cord paralysis in paramedian position - left vocal fold stiffness*

Table 3 – Speakers with No Disease (F – Female, M – Male).

Speaker	**Gender**	**Age**
Speaker 21	*F*	*56*
Speaker 22	*M*	*30*
Speaker 23	*F*	*41*
Speaker 24	*M*	*46*
Speaker 25	*F*	*61*

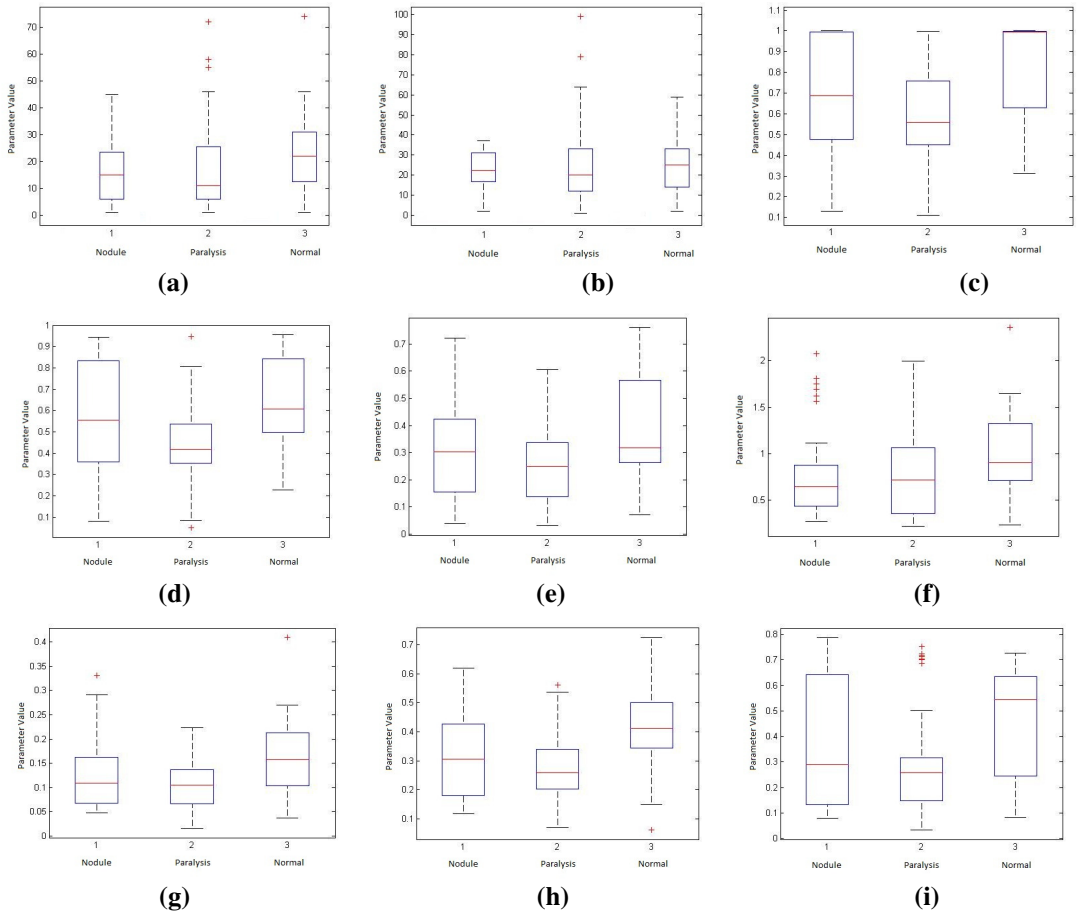
Speaker 26	*M*	*35*
Speaker 27	*M*	*63*
Speaker 28	*M*	*48*
Speaker 29	*M*	*26*
Speaker 30	*F*	*56*
Speaker 31	*F*	*56*

5 Results

5.1 Analyses of Parameters for Classification

In order to evaluate the influence of each input parameter in the classification of voice diseases, Matlab boxplot function was used. The boxplot is shown for each of the parameters extracted from the glottal signal for this work, in order to see their behavior in each type of pathology in normal voices and to compare their behavior. Boxplot is a convenient way of graphically depicting groups of numerical data, and it was used in this work to analyze the influence of each parameter in correctly classifying each disease [22].

The fundamental frequency does not undergo major changes between the three groups compared. The parameter *Ko* that shows the closing phase of the vocal cords is higher in normal voices as shown in Figure 3 (a). *CIQ* parameter, *AQ*, *NAQ*, *OQ*1 and *OQ*2 show that normal voices have more intensity in the voices compared with pathologies as shown in Figure 3 (f), (g), and (h). The parameters SQ1 and SQ2 are lower in normal voices, which indicate a shortening in the structure of the vocal cords when one has these diseases, especially paralysis, as shown in Figure (l), (m). In such condition, the structure of the vocal cords is greatly compromised this is indicated for *jitter* and *shimmer*, as shown in Figure 3 (o) and (p). *Jitter* and *shimmer* parameters vary the most in the voice when paralysis occurs. *Jitter* and *Shimmer* are very high in voices with paralysis, proving to have affected the most the vocal cords.



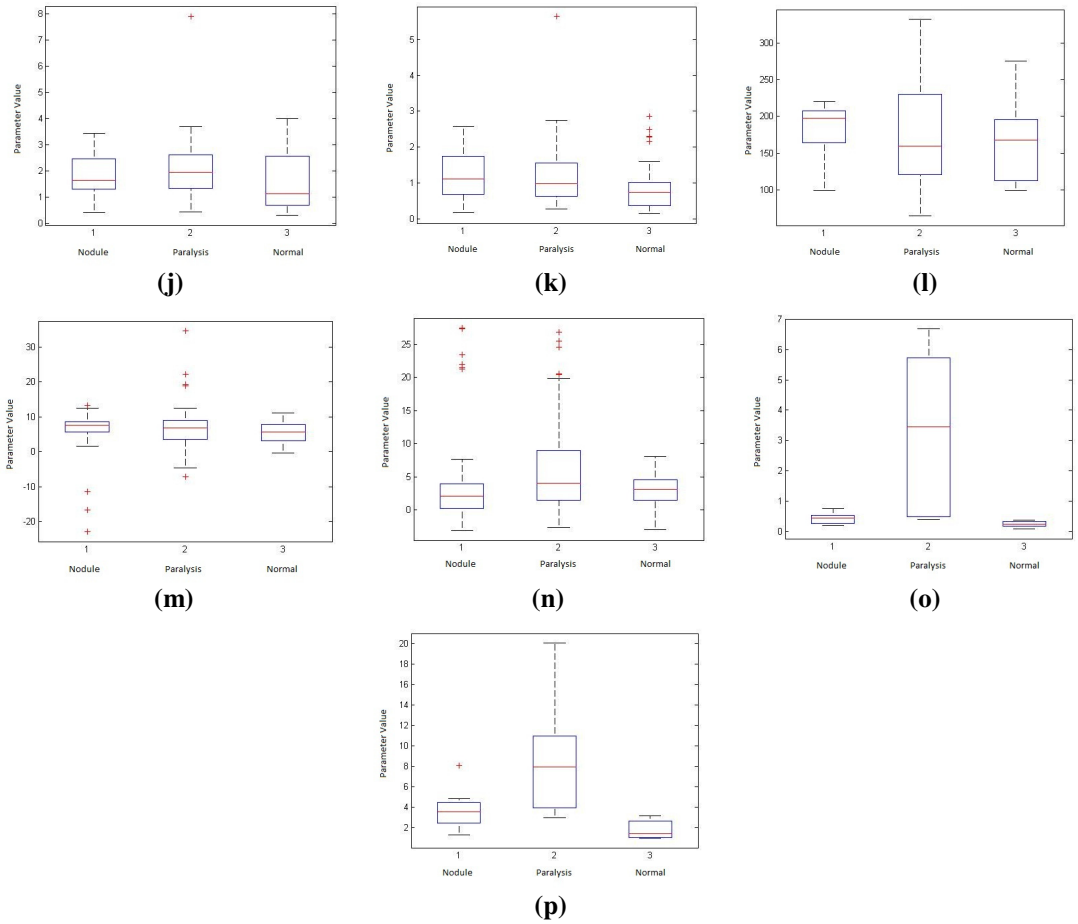


Figure 3 – Boxplot for each of the parameters extracted from the glottal signal**. K0 (a), *Ka* (b), *OQ*1 (c), *OQ*2 (d), ClQ (e), *AQ* (f), *NAQ* (g), *OQ*a (h), QoQ (i), SQ1 (j), SQ2 (k), F0 (l), *DH*12 (m), *HRF* (n), *Jitter* (o), *Shimmer* (p).**

5.2 Classification results with the parameters of the glottal signal

The original database was divided into training, validation and test sets, where 70% of the database was used for training, 20% as the validation set, and 10% for testing. After lots of tests varying the number of the processors, the best result came up with 8 processors in the hidden layer.

The Neural Network has three outputs: speakers with nodule on the vocal cords, containing 93 voice records, speaker with vocal cords paralysis, containing 67 records, and speaker with normal voice, containing 89 records. Correctly classified instances with 8 processors in the hidden layer achieved a result of 95.83% using the test set (10% of the original database), the confusion matrix is show in Table 4 (a).

5.3 Classification results with Mel-frequency cepstral coefficients (MFCCs)

Mel-frequency cepstral coefficients (**MFCCs**) are coefficients that collectively make up an MFC. They are derived from a type of cepstral representation of the audio clip (a nonlinear "spectrum-of-a-spectrum"). Are common in speaker recognition, which is the task of recognizing people from their voices [2]. To compare the results obtained in the classification of voice pathologies with the parameters of the glottal signal, it gets the 12 MFC coefficients of each of the recordings from the database through the MFCC technique, best known for that sort of speaker [2].

The original database was divided into training, validation and test sets, where 70% of the database was used for training, 20% as the validation set, and 10% for testing. After lots of tests varying the number of the processors, the best result came up with 6 processors in the hidden layer.

The Neural Network has three outputs: speakers with nodule on the vocal cords, containing 93 voice records, speaker with vocal cords paralysis, containing 67 records, and speaker with normal voice, containing 89 records. Correctly classified instances with 6 processors in the hidden layer achieved a result of 75% using the test set (10% of the original database), the confusion matrix is show in Table 4 (b).

The classification was successful with the parameters of the glottal signal, despite having fewer samples for voices with paralysis and factors such as gender and age difference between speakers, reaching the conclusion that the parameters of the glottal signal are good discriminators for classifying voice disorders.

Most studies of the classification of voice pathologies only use two classes, normal voices and voices patients [23], this study with three classes: two diseases and normal voices have superior results with the parameters of the glottal signal while the database is different.

Table 4 – Confusion Matrix for both classifications

Nodule	Paralysis	Normal	
8	**0**	**0**	Nodule
0	**8**	**0**	Paralysis
0	**1**	**7**	Normal

Nodule	Paralysis	Normal	
6	**2**	**0**	Nodule
1	**5**	**0**	Paralysis
1	**1**	**7**	Normal

Table 5 – Measures statistics of the parameters for each class

	Nodule				Paralysis				Normal			
	Average	Stand.Desv.	Min	Max	Average	Stand.Desv.	Min	Max	Average	Stand.Desv.	Min	Max
K0	0,00171	0,00098	0,00020	0,00450	0,001797	0,001504	0,00010	0,00720	0,002589	0,001332	0,0004	0,0074
Ka	0,00243	0,00083	0,00040	0,00500	0,00279	0,00185	0,00030	0,00990	0,00307	0,00133	0,0012	0,006
OQ1	0,71213	0,25084	0,1309	1,001	0,60345	0,22928	0,11140	0,99860	0,85752	0,20443	0,3106	1,0013
OQ2	0,58514	0,23944	0,0809	0,9447	0,4523	0,17619	0,0514	0,9468	0,64491	0,19085	0,2295	0,9552
CIQ	0,311	0,1602	0,0376	0,7223	0,24801	0,12573	0,0321	0,6072	0,39724	0,18526	0,0697	0,7607
AQ	0,72093	0,40697	0,2635	2,0716	0,78916	0,50021	0,2179	1,995	0,99111	0,39366	0,2286	2,3553
NAQ	0,1254	0,06816	0,0472	0,3307	0,10521	0,05276	0,0162	0,2228	0,15842	0,06963	0,0374	0,4092
OQa	0,31158	0,14168	0,1188	0,6189	0,26952	0,11744	0,0716	0,561	0,41325	0,13794	0,0628	0,7244
QoQ	0,3732	0,25213	0,0789	0,7872	0,28758	0,18497	0,0342	0,7514	0,4509	0,21067	0,0826	0,7261
SQ1	1,79398	0,77605	0,4074	3,433	2,02462	1,07392	0,4432	7,9025	1,57107	1,0766	0,3146	4,0084
SQ2	1,21607	0,65665	0,1918	2,5774	1,17752	0,81481	0,2769	5,6347	0,83121	0,60137	0,1611	2,8615
Fo	189,5762	24,06232	98,6511	220,7276	183,9806	77,24535	64,8521	332,3147	165,4452	54,68968	99,3614	215,8327
DH12	6,19828	5,82417	-22,7849	13,2562	6,55555	6,69789	-7,2049	34,5436	5,32254	2,92909	-0,3915	11,0567
HRF	4,01594	6,87889	-3,0292	27,4274	6,19521	6,91366	-2,6794	26,7979	2,88172	2,44622	-2,9348	8,12
Jitter	0,43438	0,17807	0,753	3,29663	2,56625	0,394	6,676	0,24588	0,09279	0,09279	0,096	0,377
Shimmer	3,82713	1,94526	1,313	8,114	8,60825	5,34208	2,971	20,056	1,8183	0,83861	1,015	3,186

Table 5 shows that the statistical measurements of the parameters of each class to determine how each parameter behaves in each class. Ka and Ko their averages are nearly equal to the voices with nodule, and paralysis but its value is lower compared to normal voices, we can conclude that on average the patients voices glottal signal the opening and closing time is less than the shortening may denote vocal cords. For the group of parameters *OQ1, OQ2, CIQ, AQ, NAQ, OQA, QoQ*, which belong to the parameters in time domain and its standard deviation averages are higher than for normal voices to the voices with nodule, and paralysis. SQ1 and SQ2 to the values of standard deviation to the media and the voices are lower than for normal voices with nodule, and paralysis. In the parameters of the glottal signal in time domain the results of standard deviation and average for the voices with paralysis and nodule are very similar and follow the same trends.

For the parameter values DH12 media and standard deviation for the voices with paralysis and nodule are higher than normal voices. For the parameters of the glottal signal frequency-domain Jitter Shimmer HRF standard deviation and the average are too high for voices with the voices palsy compared with nodule and nodule voices are higher than in normal voices. The parameters in the frequency domain are better distinguish the voices with paralysis the voices with nodule.

6 Conclusions

This paper presented the proposal of a new voice pathologies classification model that is able to distinguish between nodule and paralysis on the vocal cords. This paper uses parameters from the glottal signal and two voice pathologies, which is a novelty, to develop a model for the classification made in this work, we created a database of speakers with these two diseases, paralysis or nodule on the vocal cords and speakers with normal voices. The speakers have different ages and gender. The results obtained were satisfactory and better than those achieved in other works. However, it is important to note that it is not conclusive that separately the parameters will correctly identify the voice pathology, but it is conclusive that when they are

combined the performance is improved. The method of the parameters of the glottal signal is compared with the traditional MFCC method has better performance.

Acknowledgments

The authors acknowledge the financial support from the Brazilian agencies: CNPq (Conselho Nacional de Pesquisa e Desenvolvimento), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and FAPERJ (Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro)

7 References

[1] Rosa I. S., Analise acústica da voz de indivíduos na terceira idade. **Tese de mestrado Universidade de São Carlos**, (2005) (in portuguese).

[2] Londoño J., Llorente J., An improved method for voice pathology detection by means of a HMM-based feature space transformation Pattern Recognition, Volume 43, Issue 9, September 2010.

[3] Wang X, Zhang J, Yan Y, Glottal Source biometrical signature for voice pathology detection, **Speech Communication**, 51 759–781 (2009).

[4] Hariharan, M. Paulraj, M.P. Yaacob, S., **Identification of vocal fold pathology based on Mel Frequency Band Energy Coefficients and singular value decomposition Signal and Image Processing Applications** (ICSIPA), volume 514 - 517 (2009).

[5] M. D. O. Rosa, J. C. Pereira, M. Grellet, "Adaptive Estimation of Residue Signal for Voice Pathology Diagnosis", **IEEE Trans. Biomedical Eng**., Vol. 47, No. 1, Jan. 2000.

[6] Campbell, Jr. J. P., "Speaker Recognition: A tutorial", **Proceedings of the IEEE**, vol. 85, no. 9, pp. 1437-1462, 1997.

[7] Alku P., Glottal wave analysis with Pitch Synchronous Adaptive Inverse Filtering", **Speech Communication**, vol. 11, pp. 109-118 (1992).

[8] Pulak*ka* H. Analysis of Human Voice Production Using Inverse Filtering, High-Speed Imaging, and Electroglottography. **University of Technology Helsinki** (2005).

[9] Juliano S. M., "Um estudo comparativo entre o sinal electroglotográfico e o sinal de voz", **Dissertação de mestrado em Engenharia de Telecomunicações**, UFF (2008).

[10] Software Aparat, http://aparat.sourceforge.net/index.php/Main_Page, **Helsinki University of Technology Laboratory of Acoustics and Audio Signal Processing**.

[11] Alku P. and Vilkman. E. Amplitude Domain Quotient for Characterization of the Glottal Volume Velocity Waveform Estimated by Inverse Filtering. **Speech Communication**, 18(2):131-138, (1996)

[12] Alku P., Bäckström T., and Vilkman E. Normalized Amplitude Quotient for Parameterization of the Glottal Flow, **Journal of the Acoustical Society of America**, 112(2):701-710, August (2002)

[13] Gobl, C., Chasaide. A., Amplitude-based source parameters for measuring voice quality, **In V*OQ*UAL**, 151-156 (2003)

[14] [14] A.-M. Lauk*ka*nen, E. Vilkman, P. Alku, Related to Stress and Emotional State: a Preliminary Study, **Journal of Phonetics**, 24(3):313-335, (1996).

[15] I. Titze and Sundberg J. Vocal intensity in speakers and singers, **Journal of the Acoustical Society of America**, 2936-2946, May (1992)

[16] Airas M., Methods and studies of laringeal voice quality analysis in speech production, **Dissertation for the degree of Doctor Helsinki University of Technology**, (2008)

[17] Childers D. G. and Lee. C. K Vocal quality factors: Analysis, synthesis, and perception, **Journal of the Acoustical Society of America**, 2394-2410 May (1990)

[18] Cataldo, E.; Rodrigues, Fabiana; BRANDÃO, A. ; LUCERO, Jorge . Usando Redes Neurais para Classificação de padrões de voz. In: **XXVIII CNMAC - Congresso Nacional de Matemática Aplicada e Computacional**, 2005, São Paulo. Anais do XXVIII CNMAC, (2005) (in portuguese)

[19] Vieira, M. N. Automated Measures of Dysphonias and the Phonatory Effects of Asymmetries in the Posterior Larynx, Ph.D Thesis, **University of Edinburgh, UK**, (2000).

[20] Behlau M, Pontes PP. Avaliação e tratamento das disfonias. São Paulo: **Lovise**, (1995)

[21] Software Praat, http://www.fon.hum.uva.nl/praat/, **University of Amsterdam.**

[22] Alagambigai, P., Thangavel, K.., Feature Selection for Visual Clustering. **ARTCom 2009 Kerala,** vol 1, pag 498, (2009)

[23] Lee J., Jeong S., Minsoo H. Classification of Pathological and Normal Voice Based on Linear Discriminant Analysis, **Lecture Notes in Computer Science,** vol 4432/2007, pag 382-390, (2007)