International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-3, January 2020

Automatic Detection and Classification of Voice Pathology

Vikas Mittal, R. K. Sharma

Abstract: ***The result of rough vocal use is commonly voice pathology. Poor vocal practice can result in worse exceptional of voice, vocal fatigue, and vocal stress. This research utilizes glottal signal (signal produced by vocal folds) parameters to help out in identify voice disorders linked to vocal folds pathologies. For each recorded speech, the respective glottal signal is acquired. We select the most relevant as far as pathological / normal discrimination is concerned from the enormous set of parameters obtained. In this paper a new glottal signal parameter Maximum Opening Quotient (MOQ) is calculated to find Pathological / Normal speech discrimination. Using distinct options, the outcomes are compared. Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) algorithms are used for classifications. Result shows that the average efficiency rise 2.1% using the newly studied glottal parameter Maximum Opening Quotient (MOQ), which is a major contribution of this research.***

Keywords: ***Pathologic voice, Glottal signal parameters, Support Vector Machine (SVM) and K-Nearest Neighbors (KNN).***

I. INTRODUCTION

The pathological voice diagnostic science has drawn particular attentions from academic speech processing society over the past centuries. Voice disorders can be divided into three primary classifications: organic, functional and two-category combination. Organic speech disorders split into two groups: structural disorders and neurogenic disorders [1, 2].

We are proposing a non-invasive technique to assist speech therapists in early identification of vocal fold disease that can enhance evaluation precision. Automatic voice disorder classification is a method based in two main steps. First, the speech utterance extracts a number of parameters. Secondary, the pattern recognition method uses these parameters to classify the voice of the disease and normal one [3,4]. The glottal signal is used to identify vocal fold-related pathologies [5]. In London and Llorente studies [5] and Pedro et al. [6], discussed the MFCCs were not as efficient in classifying voice pathologies. Hence, the primary goal of this work is classification of voice pathology based on parameters obtained from the glottal signal. In addition, a novel glottal parameter maximum opening quotient (MOQ) is suggested in relation to the maximum opening of the vocal fold, which enables better classification efficiency.

Revised Manuscript Received on January 30, 2020.

***** Correspondence Author

Vikas Mittal*, Ph.D. Scholar, Department of ECE, School of VLSI Design and Embedded Systems, NIT, Kurukshetra, (Haryana) India. E-mail: vikasmittal2k7@gmail.com

R. K. Sharma, Professor, Department of ECE, NIT kurukshetra, (Haryana) India. E-mail: mail2drrks@gmail.com

The paper is structures as follows. In section II provides a complete methodology used to classify speech pathologies. Results are provided in section III. Discussions are presented in chapter IV on the outcomes. Section V concludes the paper

II. METHODOLOGY

The suggested methodology is shown in Figure 1.

(i) Database collection.

(ii) Extraction of existing glottal parameters and new parameter.

(iii) Classify the voice / normal pathology

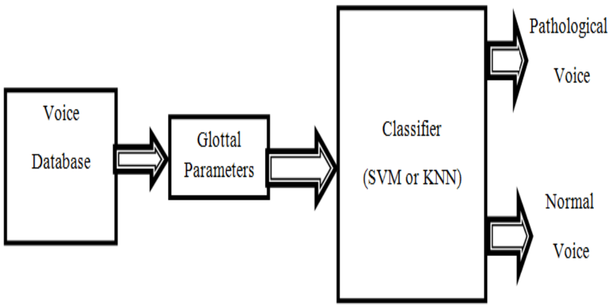

Figure 1 Methodology used to classify pathological voice

A. Database

The analyzed voices were obtained from 40 speakers divided into 20 dysphonic and 20 normal speakers. To analyze the proposed work, we utilized the Saarbrucken Voice Database (SVD) [7] and database developed from MMIMSR, Mullana, and hospital with the help of Dr. Shantanu. The voices recorded in a special sound proof room, using software called "Dr. Speech Software". Table 1 show the total number of voice types was 7 (Six pathologies and one normal group).

Table 1: Pathologies Found In The Analyzed Sample

Laryngeal Diseases	No. of cases
Normal	20
Carcinoma	2
Cyst	4
Nodule	6
Unilateral Paralysis	2
Polyp	4
Edema	2

B. Time-domain parameters of the glottal signal:

The glottal signal parameters are obtained with the help of a tool box Aparat [8]. The parameters are:

A glottal pulse is shown in figure 2 to define parameters.

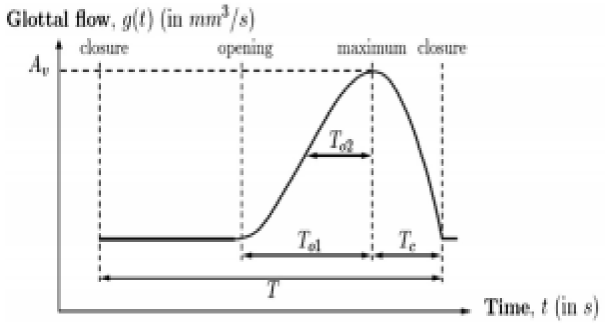

Figure 2 Representation of one glottal pulse[9]

(a) Opening quotient (OQ): the ratio of the complete moment of the opening vocal folds to the complete moment of the glottal pulse (T). It can be considered in two types OQ1 and OQ2 [10].
(b) Closing quotient (CIQ): the ratio of the closing phase parameter to a glottal pulse (T) complete length [10].
c) Amplitude quotient (AQ): The proportion of the glottal signal amplitude (Av) to the glottal signal derivative's minimum value [10].
(d) Normalized amplitude quotient (NAQ): is calculated by the proportion of the AQ to the glottal pulse (T) complete time length [10].
(e) Speed quotient (SQ): the ratio between the length of the opening stage and the length of the closing phase. It is also defined in two terms SQ1 and SQ2 [10].
(f) Maximum opening quotient (MOQ): is calculated by the proportion of time interval between the instant when the vocal folds start to oscillate to reach their highest opening point, which is represented by T_{o1} to the total length of glottal cycle or period (T). In case of pathological voice the value of MOQ is increased, which indicates vibration speed of vocal folds slows down. So, in normal (healthy) cases, the value of this parameter will need to be lower. This parameter is important because response time of this parameter is fast as compared to other. The MOQ is computed as:

$$MOQ=\frac{T_{o1}}{T} \qquad (1)$$

C. Frequency-domain parameters of glottal signal:

(a) Harmonic Difference (DH12): Also known as H1 H2, this is the difference between the first and second harmonic values of the glottal signal [10].
(b) Harmonics richness factor (HRF): refers to the first harmonic (H1) with the energy amount of the other harmonics (Hk) [10]

(D. Parameters related to the fundamental frequency:

(a) Jitter: fundamental frequency variations between consecutive cycles of vibration.
(b) Shimmer: glottal flow amplitude variations between consecutive vibrational cycles.

III. RESULT

(i) Analysis of the parameters for classification

The variation of glottal parameters, described in the following subsections and in Figures 2-13.

(a) Time-domain parameters of the glottal signal

Figures 3-10 demonstrate the respective box-plots obtained from the glottal signal for the so-called time-domain parameters, where some interesting observations can be obtained. OQ1, OQ2, CIQ, AQ and NAQ (Figures 3-7) parameters indicate that when compared to pathologies, normal voices have more intensity and better speech quality. In pathology, the values for the parameters SQ1 and SQ2 are smaller in pathological voices, which indicate a shortening in the structure of vocal folds (Figure 8 and 9).The value of a new investigated parameter MOQ as defined by equation 1 is lower for normal voices, which indicate fast and smooth opening of vocal folds during phonation as compared to one has diseases the vocal folds have slow opening or it acquired more time to reach maximum opening point from the start of oscillation.

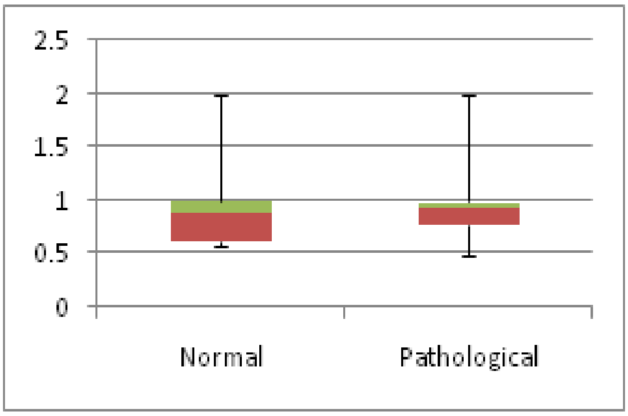

Figure 3 Open Quotient (OQ1)

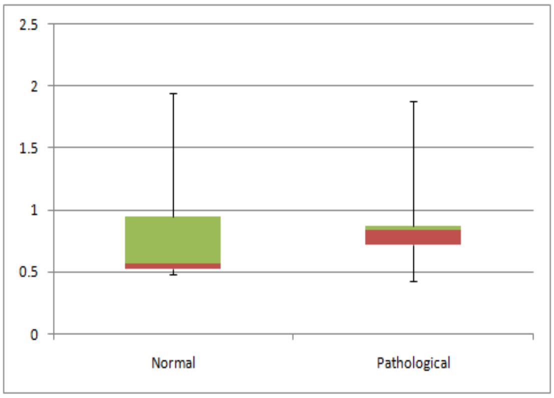

Figure 4 Open Quotient (OQ2)

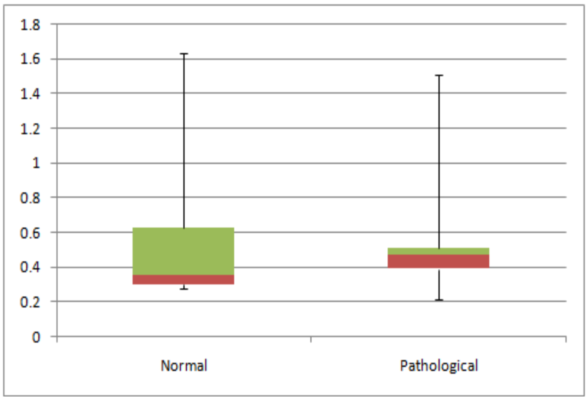

Figure 5 Closed Quotient (CIQ)

Retrieval Number: B7643129219/2020©BEIESP
DOI: 10.35940/ijitee.B7643.019320
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

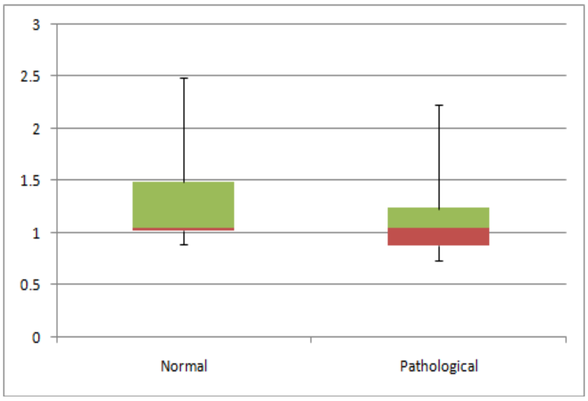

Figure 6 Amplitude Quotient (AQ)

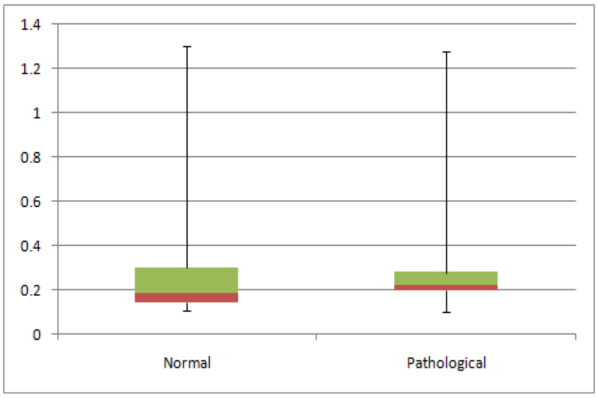

Figure 7 Normalized Amplitude Quotient (NAQ)

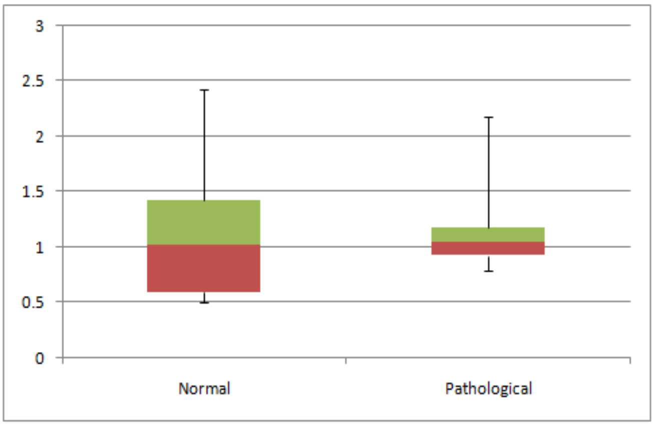

Figure 8 Speed Quotient 1(SQ1)

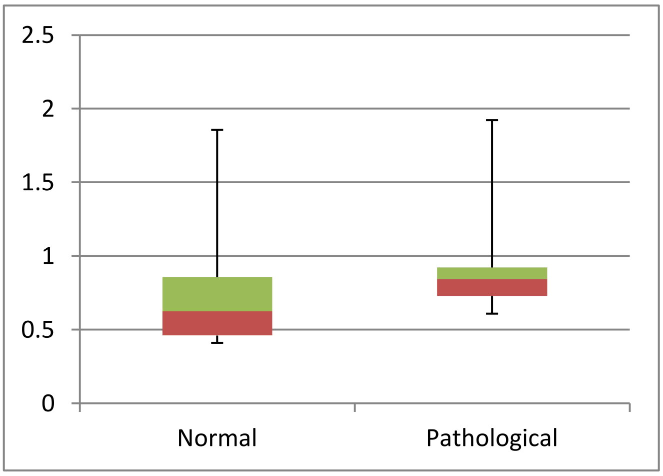

Figure 9 Speed Quotient 2(SQ2)

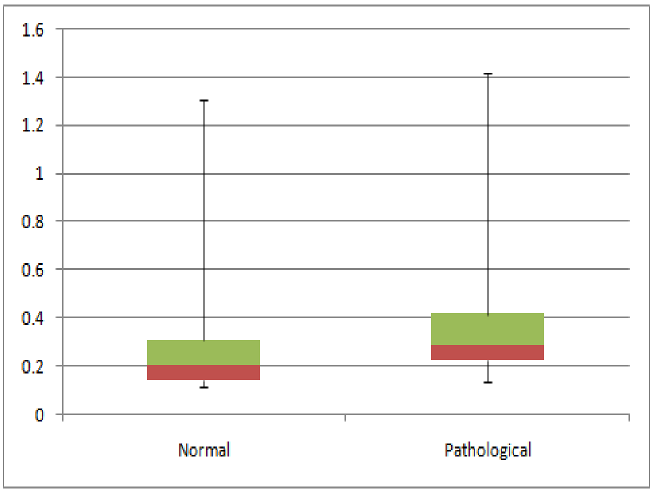

Figure 10 Maximum Opening Quotient (MOQ)

(b) Frequency-domain parameters

Figure 11 and 12 shows the resultant box-plots to frequency-domain parameters. The pathological voices have more variations in frequency as compared to normal voices.

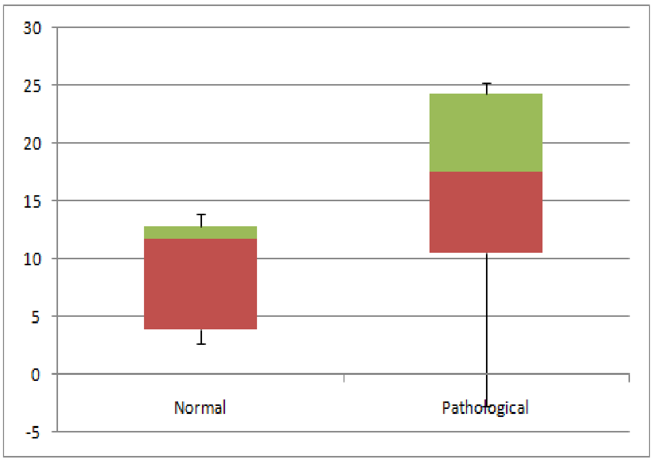

Figure 11 Difference between harmonics (DH12).

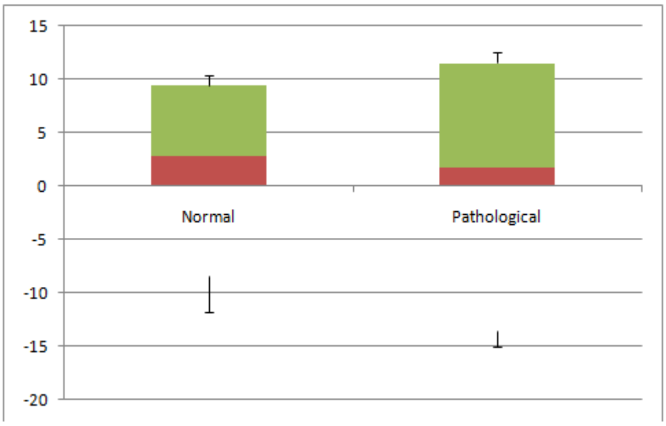

Figure 12 Harmonic richness factor (HRF)

(c) Parameters related to fundamental frequency

The shimmer and HNR values in pathological voices are very large, as shown in figures (13-14).

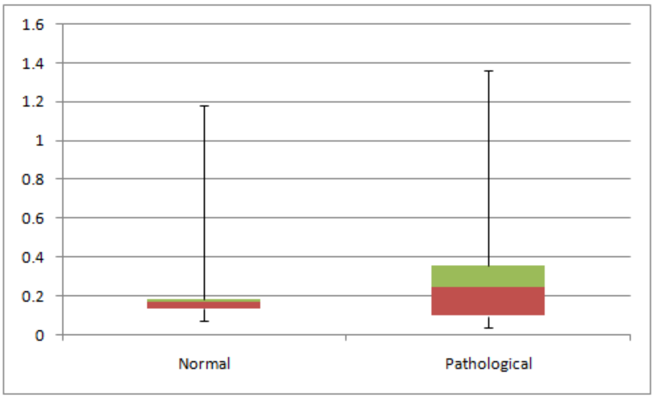

Figure 13 Shimmer

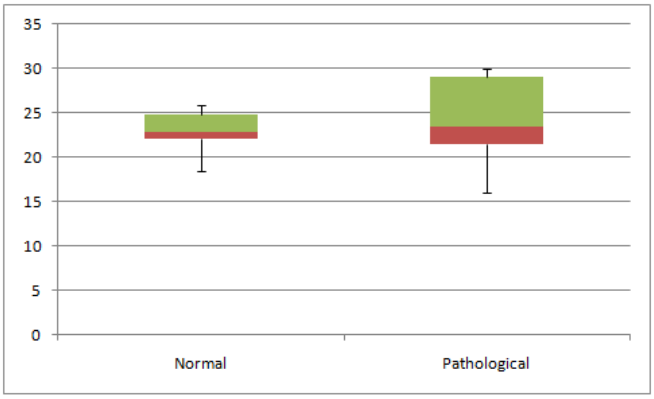

Figure 14 Harmonic to Noise Ratio (HNR)

(ii) Analysis of Results

Pathology classification was carried out using various classifiers: SVM and KNN. Three instances were discussed for the input parameters for each classifier: (1) only the extracted glottal parameters, (2) glottal parameters including new investigated parameter Maximum Opening Quotient (MOQ)

(1) Classification Results with only glottal parameters

The table 1 shows that classification accuracy considers SVM as classifier is obtained 77.5% and with KNN classifier is 67.5%.

Table 1: Confusion matrix of only the glottal parameters

SVM Classifier			
True Class	Normal	18	2
	Pathological	7	13
		Pathological	Normal
KNN Classifier			
True Class	Normal	17	3
	Pathological	10	10
		Normal	Pathological

(2) Classification Results with glottal parameters including new investigated parameter (MOQ)

The input of the classifiers is composed of 16 glottal parameters and plus one new investigated Maximum Opening Quotient MOQ. The confusion matrix is given below in table 2.

The table 2 shows that classification accuracy considers SVM as classifier is obtained 95% and with KNN classifier is 90%.

Table 2: Confusion matrix of glottal parameters and plus one new investigated MOQ

SVM Classifier			
True Class	Normal	18	2
	Pathological	0	20
		Normal	Pathological
KNN Classifier			
True Class	Normal	18	2
	Pathological	2	18
		Normal	Pathological

IV. DISCUSSION

Table 3 summarizes the comparison of accuracy of both the classifiers. The classification was effective, as seen from the outcomes in table 3, concluding that glottal parameters are good discriminators of classifying voice disorders with new investigated glottal parameter Maximum Opening Quotient (MOQ) and classification performance improved.

Table 3: Comparison of Average Accuracy

Parameters	SVM	KNN	Average Accuracy
Only glottal parameters	77.5%	67.5%	72.5
Glottal parameters including MOQ parameter	95%	90%	92.5%

V. CONCLUSION

A technique has been developed to classify pathological voice using characteristics obtained from glottal and voice signals based on well-known SVM and KNN classifiers.

The main contribution of this work is the use of newly investigated glottal parameter (MOQ). The average accuracy of classification with both classifiers is 92.5 % achieved glottal with MOQ parameter, which is better as compared to other case. The combination of existing glottal parameters and with newly investigated glottal parameter MOQ offered average efficiency rise 2.1 % in classification the importance of the work or suggest applications and extensions.

ACKNOWLEDGMENT

Authors are thankful to Special Manpower Development Program for Chip to System Design (SMDP-C2SD) initiated by Govt. of India for providing hardware, software and other technical resources in School of VLSI Design and Embedded Systems, NIT, Kurukshetra.

REFERENCES

1. J. Cheolwoo,"Source Analysis of Pathological Voice," in Proc. International multiconference of engineers and computer scientists,2010.
2. Hariharan M.Paulraj MP, Yaacob S. identifiaction of Vocal Fold Pathology Based on Mel Frequrncy Band Energy Coefficient and singular Value Decomposition. Signal and Image Prcessing Applications(ICSIPA);2009:514-517.

3. Kohler M,Mendoza LF,Lazo J,Vellasco M,Cataldo E. Classification of voice pathologies using glottal signal parameters. Proc. Of Brazilian Congress on Computational Intelleligence(CBIC);2011.
4. Verdonck-de Leeuw IM, Mahieu HF. Vocal aging and the impact on daily life: a longitudinal study.J Voice.2004;18:193-202.
5. Londono J, Liorente J. An improved method of voice pathology detection by means of a HMM-based feature space transformation.Pattern Recogn. 2010;43:3100-3112.
6. Pedro Gomez-Vilda et al. Glottal source biometrical signature for voice pathology detection. Speech Communication 2009;51:759-781.
7. W.J.Barry and M.Putzer, " Saarbrucken Voice Database," Institute of Pjonetics,Univ. of Saarland http://www.stimmdatenbank.coli.uni-saarland.de/.
8. Software Aparat, Available at: http://aparat.sourceforge.net/index.php/main_page, Helsinki University of Technolgoly Laboratory of Acoustics and Audio Signal Processing, Accesessed June 15, 2015.
9. Leonardo A.Forero M. ,Manoela Kohler,marley M.B.R Vellasco, and Edson Cataldo, " Classification of Vocal Aging Using parameters extracted from glottal signal ",Journal of Voice,vol.28,pp. 532-537,2014.
10. Nelwamondo, Fulufhelo V., and Tshilidzi Marwala. "Faults Detection Using Gaussian Mixture Models, Mel-Frequency Cepstral Coefflcients and Kurtosis." *2006 IEEE International Conference on Systems, MAan, and Cybernetics October 8-11, 2006, Taipei, Taiwan*. 1-4244-0100-3/06. (2006): 290-295. Print.

AUTHORS PROFILE

Vikas Mittal, received his M.Tech in Electronics and Communication Engineering from Kurukshetra University Kurukshetra. Currently, he is pursuing Ph.D. at National Institute of Technology (NIT) in the School of VLSI Design and Embedded Systems. His research interests include Biomedical Signal Processing, VLSI Design and Embedded System.

R. K. Sharma, received his M.Tech in Electronics and Communication Engineering and PhD degree in electronics and communication from kurukshetra University, Kurukshetra (through National Institute of Technology Kurukshetra), India in 1993 and 2007, respectively. Currently he is Professor with the Department of Electronics and Communication Engineering, NIT Kurukshetra, India. His main research interests are in the field of low power VLSI design, Voice profiling, Microprocessor and FPG Abased systems.