INTERSPEECH 2017
August 20–24, 2017, Stockholm, Sweden

Zero Frequency Filter Based Analysis of Voice Disorders

Nagaraj Adiga[1], Vikram C M[1], Keerthi Pullela[2], and S R M Prasanna[1]

[1]Department of Electronics and Electrical Engineering
Indian Institute of Technology Guwahati, India
[2] VIT University, Chennai, India

{nagaraj, cmvikram, prasanna}@iitg.ernet.in, keerthipullela70@gmail.com

Abstract

Pitch period and amplitude perturbations are widely used parameters to discriminate normal and voice disorder speech. Instantaneous pitch period and amplitude of glottal vibrations directly from the speech waveform may not give an accurate estimation of jitter and shimmer. In this paper, the significance of epochs (glottal closure instants) and strength of excitation (SoE) derived from the zero-frequency filter (ZFF) are exploited to discriminate the voice disorder and normal speech. Pitch epoch derived from ZFF is used to compute the jitter, and SoE derived around each epoch is used compute the shimmer. The derived epoch-based features are analyzed on the some of the voice disorders like Parkinson's disease, vocal fold paralysis, cyst, and gastroesophageal reflux disease. The significance of proposed epoch-based features for discriminating normal and pathological voices is analyzed and compared with the state-of-the-art methods using a support vector machine classifier. The results show that epoch-based features performed significantly better than other methods both in clean and noisy conditions.

Index Terms: Zero-frequency filter, epoch-based features, jitter, and shimmer

1. Introduction

There are millions of people suffering from a significant disability due to the voicing disorders. Assessment of voice disorders is a challenging task and there are no known biomarkers that can be used for diagnosis. Voice disorders are vocal conditions involving mainly abnormalities in the pitch, the strength of excitation (SoE), and thereby affecting the perceptual quality of produced speech. Voice disorders are broadly divided into two categories namely neurological and organic. Neurological voice disorders include voice problems caused by abnormal control, coordination, or strength of voice box muscles due to an underlying neurological disease such as stroke, Parkinson's disease, multiple sclerosis, myasthenia gravis, and amyotrophic lateral sclerosis. Organic disorders include vocal cord nodules, polyps, vocal cord paralysis, gastroesophageal reflux disease (GERD), cyst, etc. Acoustic analysis based automatic detection of voice disorders become an interest of research, because of non-invasive and objective quantification of pathological information. Researchers have been focused on the development of acoustic features, which can efficiently represent the pathological condition of the speech production system.

1.1. Prior art

Different signal processing based methods like, temporal, spectral, and non-linear dynamic based features are proposed in the literature for voice disorder analysis [1]. A detailed focus is given to the issues involved in the development of pathological voice detection system [1]. Further, machine learning techniques are incorporated to develop an automatic voice disorder detection system [2–4]. Among the different features, the voice source based attributes like a cycle-to-cycle change in the glottal vibration and noise parameters like harmonic-to-noise ratio become popular [5], [6]. The quasi-periodic glottal vibration may be affected due to the organic mass formation as in cyst, polyp, or due to the neurological reasons like Parkinson's disease, stroke. Hence, perturbation in the glottal vibration cycles in terms of the pitch period and amplitude are widely used features for the analysis of pathological speech [5]. Jitter (perturbation measure for the period) and shimmer (perturbation measure for the amplitude of glottal vibration) are considered as the effective features for the pathological speech analysis. Both features have been largely used to detect voice pathologies [7], [8]. In [9, 10], a mathematical model is incorporated for jitter estimation and then its significance is shown for pathological voices. Jitter and shimmer are commonly measured from the long sustained vowels. The values above a certain threshold are considered being related to pathological voices, which are usually perceived by humans as breathy, rough or hoarse voices. Vocal pathologies introduce the local aperiodicity and irregularities in the speech waveform, where the rate of vibration of the vocal folds may change from one glottal cycle to the next cycle. Hence, the methods based on the assumption of periodicity or block based pitch estimation may not give the cycle-to-cycle variations in the glottal vibrations accurately.

1.2. Contributions

The current work is motivated to exploit the instantaneous fundamental frequency and SoE for the computation of jitter and shimmer. The vocal folds vibration results in a sequence of glottal pulses with major excitation taking place around the glottal closure instants (GCIs) or epoch. Epoch-based pitch detectors estimate the pitch period by locating the instants at which glottis closes and then measuring the time interval between two such events [11, 12]. Pitch period computed from epoch-based methods gives pitch period corresponding to each glottal cycle, whose inversion is considered as instantaneous fundamental frequency [11]. Hence, it is more appropriate to use the instantaneous fundamental period computed from every glottal cycle for jitter and shimmer estimation. In our work, we used zero-frequency filtering (ZFF) based method for extracting instantaneous fundamental frequency and SoE [13]. This method is considered to be robust against noise as it extracts the epochs by filtering of speech signal around zero Hz, which is least affected by vocal tract response and external noise. Extraction of shimmer directly from a speech signal by measuring the peak-to-peak amplitude at each glottal cycle may not give the accurate information about the cycle-to-cycle amplitude pertur-

  http://dx.doi.org/10.21437/Interspeech.2017-589

bation. The SoE measured from ZFF gives the approximately proportional to source amplitude which is used to characterize the glottal activity [14, 15]. Jitter and shimmer computed from the proposed method is used for the discrimination of normal and voice disorder speech. Further, the robustness of proposed epoch-based features under white and babble noise conditions is tested for the analysis of voice disorders. The paper is organized as follows: Section 2 describes the database, extraction of ZFF based features is explained in Section 3. Analysis of voice disorders using epoch-based features is explained in Section 4. Classification results are given in Section 5, and Section 6 provides conclusion and direction for the future work.

2. Database

Speech samples are collected from controlled normal (CN) speakers and individuals with voice disorders (VD). Table 1 describes the details about the database used in the current work. The database consists of samples from Parkinson's disease, vocal chord cyst, Gastroesophageal reflux disease (GERD), and vocal chord paralysis. Samples are recorded at 16 kHz frequency using a unidirectional microphone and PRAAT software [16] at All Indian Institute of Speech and Hearing, Mysore, India [17]. It had been experimentally showed earlier that compared to continuous speech sustained phonation contains more information about vocal pathology [18]. Therefore, in this work, the sustained vowel $/aa/$ of duration around 2 s is recorded. From each person, single or multiple time recordings are considered.

Table 1: *DATABASE*

Class	No. of subjects	No. of samples
Controlled normal (CN)	44	64
Parkinson's	25	55
Cyst	04	12
GERD	07	21
Vocal fold Paralysis	05	15

3. Zero-frequency filter based features

Quasi-periodic puffs of airflow due to the adduction/abduction process of vocal fold excite the vocal tract system to produce the voiced speech. The resulting speech signal consists of highest energy around the epoch [13]. Hence, in the speech production mechanism, it is assumed that around the epoch location, impulse-like excitation signal excites the vocal tract system. Such impulse-like information is present everywhere in the speech spectrum including zero frequency. ZFF compute these epochs by extracting impulse like information around the zero frequency. Epoch extraction from ZFF involves the following steps.

- Pass the speech signal twice through a cascade of ideal zero Hz resonator whose transfer function ($H(z)$) is given by

$$H(z) = \frac{1}{(1 - z^{-1})^2} = \frac{1}{1 - 2z^{-1} + z^{-2}}. \quad (1)$$

- Compute the average pitch period using a frame size of 30 ms with a shift of 10 ms by autocorrelation method.
- Remove the trend present in the resonator output using a moving average filter with a window, whose length equal to average pitch period [13].
- The trend removed signal is called as zero-frequency filtered signal (ZFFS).

Positive or negative zero-crossings of ZFFS (depending on the polarity) gives the epoch locations [19]. Slope measured at the zero crossing corresponding to epoch is considered as a measure of the strength of glottal vibration and called as the SoE [14]. Pitch period derived from ZFFS by the successive difference of the epoch locations is referred as the instantaneous pitch period (T_0). Figure 1 shows the speech signal, ZFFS, instantaneous pitch period contour, first order absolute difference computed for the pitch period contour, SoE, and first order absolute difference computed for the SoE contour. First order difference of the pitch period and SoE derived from ZFF shows the significant difference between the normal and pathological speech. Cycle-to-cycle changes in terms of amplitude and strength information are effectively captured from ZFF. Hence, ZFF can be used to derive the time and amplitude perturbation features to discriminate the normal and pathological voices.

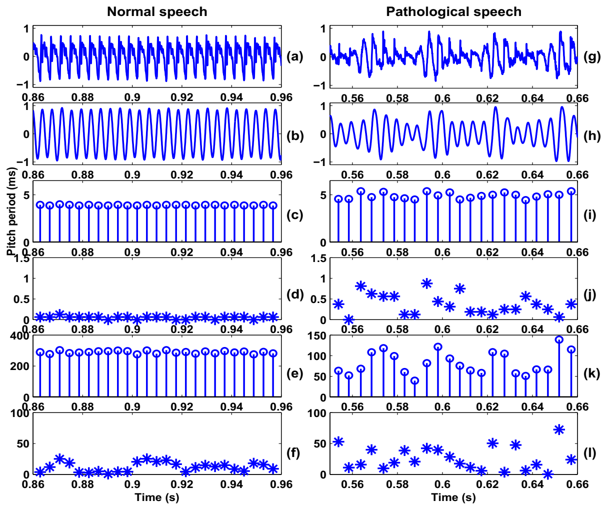

Figure 1: *Extraction of instantaneous pitch period and SoE from ZFF: (a)-(f) and (g)-(l) represent the speech signal, ZFFS, instantaneous pitch period (T_0) contour, first order absolute difference of T_0, SoE contour, first order absolute difference of SoE, for normal and pathological speech signals, respectively.*

3.1. Computation of Epoch-based perturbation features

The features derived from ZFF are considered for the discrimination between normal and pathological speech. Further, in this work, both pitch period and amplitude perturbation features are calculated in different ways. Pitch period perturbation features used here are jitter (abs), jitter (local), relative pitch perturbation (rap) – pitch period perturbation computed over 3 successive pitch periods, and pitch perturbation quotient (ppq) – perturbation computed over 5 successive pitch periods. These features are computed to represent the cycle-to-cycle change in the instantaneous fundamental period. Computation method along with the equations are given in Table 2. Amplitude perturbation features are generally computed using the peak-to-peak amplitude of speech measured at each glottal cycle. In this work, SoE (γ) derived from ZFF is used to compute the amplitude perturbation features. Shimmer (abs), shimmer (local, dB), amplitude perturbation quotient-3 (apq3) i.e. shimmer computed over three successive cycles, and amplitude perturbation quotient-5 (apq5) i.e. shimmer computed over five successive cycles are used to represent the cycle-to-cycle-change in the SoE. Computation of these features with relevant expressions are given in Table 3.

Table 2: *Pitch period perturbation features*

Sl. No	Features	Calculation method
1	Jitter (local, abs)	$\frac{1}{N-1}\sum_{i=1}^{N-1}\lvert T_{i+1}-T_i\rvert$ where, N : Total number of epochs
2	Jitter (local, %)	$\frac{\frac{1}{N-1}\sum_{i=1}^{N-1}\lvert T_{i+1}-T_i\rvert}{T_{mean}}\times 100$ where, T_{mean} : Mean pitch period $T_{mean}=\frac{1}{N}\sum_{i=1}^{N}T_i$
3	Jitter (3 point) Relative average perturbation : smoothed over 3 pitch periods (rap)	$\frac{\frac{1}{N-2}\sum_{i=2}^{N-1}\lvert\frac{T_{i+1}+T_i+T_{i-1}}{3}-T_i\rvert}{T_{mean}}\times 100$
4	Jitter (5 point) Pitch perturbation quotient: smoothed over 5 pitch periods (ppq5)	$\frac{\frac{1}{N-4}\sum_{i=3}^{N-2}\lvert\frac{\sum_{k=i-2}^{i+2}T_k}{5}-T_i\rvert}{T_{mean}}\times 100$

Table 3: *Amplitude perturbation features*

Sl. No	Features	Calculation method
2	Shimmer (local, %)	$\frac{\frac{1}{N-1}\sum_{i=1}^{N-1}\lvert\gamma_{i+1}-\gamma_i\rvert}{\gamma_{mean}}\times 100$ where, γ_{mean} : Mean SoE $\gamma_{mean}=\frac{1}{N}\sum_{i=1}^{N}\gamma_i$
2	shimmer (local, dB)	$\frac{1}{N-1}\sum_{i=1}^{N-1}20\times log\frac{\gamma_{i+1}}{\gamma_i}$
3	shimmer(3 point) Amplitude perturbation quotient : smoothed over 3 pitch periods (apq3)	$\frac{\frac{1}{N-2}\sum_{i=2}^{N-1}\lvert\frac{\gamma_{i+1}+\gamma_i+\gamma_{i-1}}{3}-\gamma_i\rvert}{\gamma_{mean}}\times 100$
4	Shimmer (5 point) Amplitude perturbation quotient: smoothed over 5 pitch periods (apq5)	$\frac{\frac{1}{N-4}\sum_{i=3}^{N-2}\lvert\frac{\sum_{k=i-2}^{i+2}\gamma_k}{5}-\gamma_i\rvert}{\gamma_{mean}}\times 100$

4. Analysis of voice disorders

The proposed features derived from the ZFF are evaluated on the collected pathological speech database mentioned in Table 1. Combination of pitch period and amplitude perturbation features, a total 8 features are used for evaluation. Box-plots for the pitch period perturbation or jitter measures for the proposed and PRAAT-based features are as shown in Figures 2 and 3, respectively. In the plot, results from all the samples of both controlled normal (CN) speaker and voice disorder (VD) speaker is shown. It can be seen that all 4 proposed parameters having higher feature values for VD speaker compared to CN speaker, whereas, this discrimination is less in case of PRAAT-based features.

For evaluating the significance of proposed epoch-based features, F-statistics analysis is made. F-statistics is a statistical significance test and compares the two distributions under the null hypothesis [20]. The input to the distributions is proposed features values computed from the normal speech and voice disorder speech. Higher values of F-statistics indicates that the two distributions which are under test are significantly discriminated and indicates that two speech signals can be classified with higher accuracy. For each pathology, F-statistics is analyzed for each feature individually. The comparison is done with existing PRAAT software [16]. PRAAT-based voice report is widely used by phoneticians and clinicians for the analysis of pathological speech [21]. From the PRAAT software, jitter and shimmer features are extracted. PRAAT computes perturbation parameters by directly locating the pitch epochs from the speech signal and peak-to-peak amplitude at each pitch epoch. F-statistics for the pair of normal and pathological voices for the proposed features and PRAAT-based features are depicted in Table 4. As shown in the table, F-statisticss for the proposed method is higher than that of PRAAT-based features. Further, it can be seen that jitter parameter (jitter local % and jitter local, abs) calculated for two cycles is performed better for all 4 voice disorder conditions, compared to 3 and 5 pitch periods. Whereas, for PRAAT-based feature jitter calculated for 3 and 5 glottal cycles also performed better for GERD and vocal fold paralysis case. The results indicate that computation of pitch period perturbation features using epochal information derived from ZFF are better able discriminate between normal and voice disordered classes, than that of PRAAT-based features.

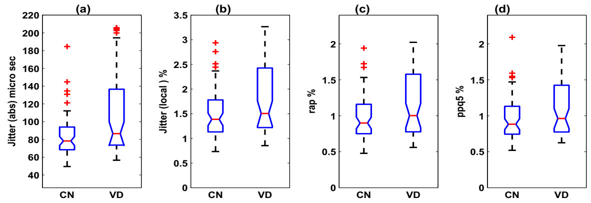

Figure 2: *Pitch period perturbation features using proposed method. (a) Jitter (abs) micro sec, (b) Jitter (local) %, (c) relative average perturbation (rap) %, and (d) pitch period perturbation-5 point (ppq5) %*

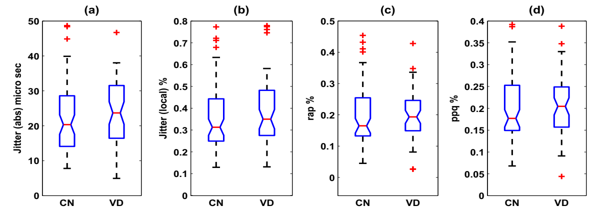

Figure 3: *Pitch period perturbation features PRAAT. (a) Jitter (abs) micro sec, (b) Jitter (local) %, (c) relative average perturbation (rap) %, and (d) pitch perturbation quotient-5 point (ppq5) %*

Table 4: *F-statistics Analysis of pitch period perturbation features and comparison with PRAAT*

Disorder	Proposed method				PRAAT			
	Jitter (local, abs)	Jtter (local)	rap	ppq	Jitter (local, abs)	Jitter (local)	rap	ppq
Parkinson's disease	5.95	5.59	5.88	5.39	1.79	1.82	1.05	0.87
Cyst	3.97	6.19	3.60	4.3	4.71	2.62	1.87	1.96
GERD	9.07	9.63	9.23	8.37	3.86	3.23	2.18	3.88
Vocal fold paralysis	5.9	6.10	6.02	5.75	5.86	5.09	5.91	2.4

Box-plots for the amplitude perturbation or shimmer measures for the proposed and PRAAT-based features are as shown in Figures 4 and 5, respectively. The amplitude perturbation calculated from proposed parameters also follows a similar trend as that of pitch period perturbation and gives higher features values for VD speaker when compared to CN speakers. F-statistics for the pair of normal and pathological voices for the proposed features and PRAAT-based features are as shown in Table 5. Similar to pitch period perturbation measures, amplitude-based features computed from SoE are able to discriminate between the normal and pathological voice better than that of PRAAT-based features.

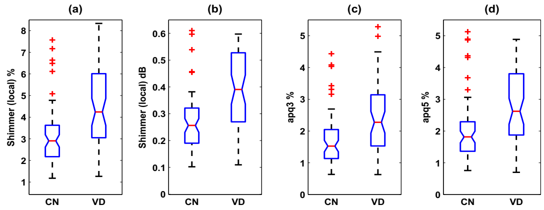

Figure 4: *Amplitude perturbation measures using proposed method. (a) Shimmer (local)% , (b) Shimmer (local) dB, (c) amplitude perturbation quotient 3 point (apq 3) %, and (d) amplitude perturbation quotient 5 point (apq 5) %*

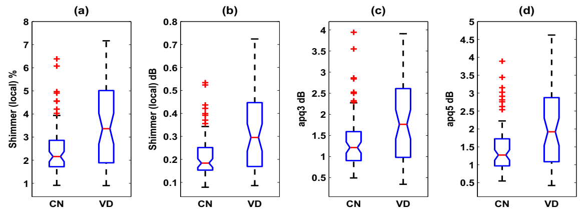

Figure 5: *Amplitude perturbation measures using PRAAT. (a) Shimmer (local)% , (b) Shimmer (local) dB, (c) amplitude perturbation quotient 3 point (apq 3) %, and (d) amplitude perturbation quotient 5point (apq 5) %*

Table 5: *F-statistics analysis of amplitude perturbation features and comparison with PRAAT*

	Proposed method				PRAAT			
Disorder	Shimmer (local)	Shimmer (local dB)	apq3	apq5	Shimmer (local)	Shimmer (local dB)	apq3	apq5
Parkinson's disease	5.86	9.11	8.85	8.41	4.74	5.37	3.41	4.22
Cyst	8.10	6.10	7.40	9.09	9.50	9.70	9.10	9.40
GERD	8.57	10.58	11.93	9.39	5.81	2.77	3.27	8.73
Vocal fold paralysis	9.92	8.49	15.99	10.99	5.90	3.35	7.90	6.71

5. Classification results

In order to get better separability between CN and VD speakers, also to avoid the threshold for classification, proposed features are tested using support vector machine (SVM) classifier. The proposed 8 dimensional epoch-based features are used as input vectors for SVM classifier and used to classify the normal and pathological voices. Here, the polynomial kernel of order 3 is used. 70% normal and pathological voice samples are used for training and 30% for testing. For the given train and test set, SVM classifier results in 92.87% and 66.66% classification for the proposed features and PRAAT-based features, respectively. The discriminating capability of proposed features for the normal and pathological voices is higher than that of PRAAT. The results also indicate that cycle-to-cycle variations in terms of the pitch period and amplitude of glottal vibration estimated from ZFF are better than that of direct measurement from speech. Here, only jitter and shimmer based measures are used for analysis and the inclusion of noise parameters like HNR improve the performance of both proposed and PRAAT methods.

5.1. Robustness to noise

Discrimination between normal and pathological voices under noisy conditions is a challenging task. Since pathological voices are added with noise due to turbulence generated at glottis due to irregular vibration of vocal folds. It is difficult to discriminate between external noise and noise due to pathology. ZFF method is shown to be robust under noisy conditions because it filters the signal around zero Hz, which is least affected by noise [13]. In this work, white and babble noise samples from NOISEX database [22] are added to normal and pathological samples for both train and test sets by varying the signal to noise ratio (SNR) at 0, 10, 20, 30 and 40 dBs. SVM is trained and tested for the noisy speech. Classification accuracy of SVM for white noise and babble noise are as shown in Figures 6.(a) and (b) respectively. These graphs show that proposed epoch-based features are robust for the addition of external noise above 10 dB, whereas the PRAAT-based features show severe degradation in the performance.

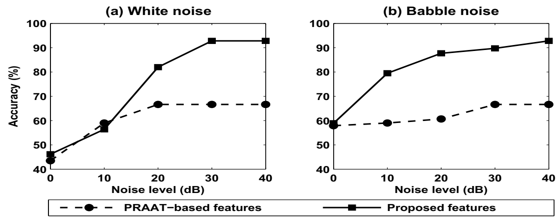

Figure 6: *Classification performance of SVM for the proposed and PRAAT-based features under noisy conditions*

6. Conclusion

We proposed the epoch-based features derived from ZFF for the discrimination of normal and pathological voices. The voice disorder speech consists of perturbation in pitch period and amplitude of each glottal cycles. Hence, the epoch-based jitter and shimmer are helpful in discriminating CN and VD speaker when compared to block processing based method. F-statistics analysis also showed that proposed features are better than widely used PRAAT features. Further, classification of CN and VD speakers are done using an SVM-based classifier, which shows better discrimination compared to PRAAT for the proposed pitch and amplitude perturbation features. Robustness of proposed features under white and babble noise conditions is evaluated, which shows that the proposed features are able to identify the normal and pathological voices under external noisy conditions also. The proposed features can be used to develop real-time pathological voice detection systems as they are robust to the addition of external noise. Future work includes an extension of epoch-based features to develop a multi-class classification of different pathological conditions.

7. Acknowledgement

This work is in part supported by the project grants, for the projects entitled "NASOSPEECH: Development of Diagnostic system for Severity Assessment of the Disordered Speech" funded by the Department of Biotechnology (DBT), Govt. of India and "ARTICULATE +: A system for automated assessment and rehabilitation of persons with articulation disorders" funded by the Ministry of Human Resource Development (MHRD), Govt. of India.

8. References

[1] N. Saenz-Lechon, J. I. Godino-Llorente, V. Osma-Ruiz, and P. Gomez-Vilda, “Methodological issues in the development of automatic systems for voice pathology detection,” *Biomedical Signal Processing and Control*, vol. 1, no. 2, pp. 120–128, 2006.

[2] M. K. Arjmandi, M. Pooyan, M. Mikaili, M. Vali, and A. Moqarehzadeh, “Identification of voice disorders using long-time features and support vector machine with different feature reduction methods,” *Journal of Voice*, vol. 25, no. 6, pp. e275–e289, 2011.

[3] J. Orozco-Arroyave, F. Hönig, J. Arias-Londoño, J. Vargas-Bonilla, K. Daqrouq, S. Skodda, J. Rusz, and E. Nöth, “Automatic detection of Parkinson’s disease in running speech spoken in three different languages,” *The Journal of the Acoustical Society of America*, vol. 139, no. 1, pp. 481–500, 2016.

[4] J. I. Godino-Llorente, R. Fraile, N. Sáenz-Lechón, V. Osma-Ruiz, and P. Gomez-Vilda, “Automatic detection of voice impairments from text-dependent running speech,” *Biomedical Signal Processing and Control*, vol. 4, no. 3, pp. 176–182, 2009.

[5] P. Lieberman, “Some acoustic measures of the fundamental periodicity of normal and pathologic larynges,” *The Journal of the Acoustical Society of America*, vol. 35, no. 3, pp. 344–353, 1963.

[6] J. I. Godino-Llorente, V. Osma-Ruiz, N. Sáenz-Lechón, P. Gómez-Vilda, M. Blanco-Velasco, and F. Cruz-Roldán, “The effectiveness of the glottal to noise excitation ratio for the screening of voice disorders,” *Journal of Voice*, vol. 24, no. 1, pp. 47–56, 2010.

[7] A. Al-nasheri, G. Muhammad, M. Alsulaiman, Z. Ali, T. A. Mesallam, M. Farahat, K. H. Malki, and M. A. Bencherif, “An investigation of multidimensional voice program parameters in three different databases for voice pathology detection and classification,” *Journal of Voice*, 2016.

[8] R. J. Moran, R. B. Reilly, P. de Chazal, and P. D. Lacy, “Telephony-based voice pathology assessment using automated speech analysis,” *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 3, pp. 468–477, 2006.

[9] M. Vasilakis and Y. Stylianou, “Voice pathology detection based on short-term jitter estimations in running speech,” *Folia Phoniatrica et Logopaedica*, vol. 61, no. 3, pp. 153–170, 2009.

[10] ——, “Spectral jitter modeling and estimation,” *Biomedical Signal Processing and Control*, vol. 4, no. 3, pp. 183–193, 2009.

[11] B. Yegnanarayana and K. S. R. Murty, “Event-based instantaneous fundamental frequency estimation from speech signals,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 4, pp. 614–624, 2009.

[12] T. Drugman, M. Thomas, J. Gudnason, P. Naylor, and T. Dutoit, “Detection of glottal closure instants from speech signals: a quantitative review,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 3, pp. 994–1006, 2012.

[13] K. S. R. Murty and B. Yegnanarayana, “Epoch extraction from speech signals,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 8, pp. 1602–1613, 2008.

[14] K. S. R. Murty, B. Yegnanarayana, and M. A. Joseph, “Characterization of glottal activity from speech signals,” *IEEE signal processing letters*, vol. 16, no. 6, pp. 469–472, 2009.

[15] N. Adiga and S. Prasanna, “Detection of glottal activity using different attributes of source information,” *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 2107–2111, 2015.

[16] P. Boersma *et al.*, “PRAAT, a system for doing phonetics by computer,” *Glot international*, vol. 5, no. 9/10, pp. 341–345, 2002.

[17] “All India Institute of Speech and Hearing, Mysore, India.” [Online]. Available: http://www.aiishmysore.in/en/about_aiish.html

[18] V. Parsa and D. G. Jamieson, “Acoustic discrimination of pathological Voice: Sustained vowels versus continuous speech,” *Journal of Speech, Language, and Hearing Research*, vol. 44, no. 2, pp. 327–339, 2001.

[19] K. Deepak, K. Ramesh, N. Adiga, and S. Prasanna, “Speech and EGG polarity detection using Hilbert envelope,” in *Proc. TENCON*. IEEE, 2015, pp. 1–6.

[20] R. A. Fisher, “The use of multiple measurements in taxonomic problems,” *Annals of eugenics*, vol. 7, no. 2, pp. 179–188, 1936.

[21] J. Rohrer, S. Maturo, C. Hill, G. Bunting, C. Ballif, and C. Hartnick, “Pediatric voice analysis: Comparison of 2 computerized analysis systems,” *JAMA Otolaryngology–Head & Neck Surgery*, vol. 140, no. 8, pp. 742–745, 2014.

[22] Noisex-92. [Online]. Available: http://www.speech.cs.cmu.edu/comp.speech/Sectionl/Data/noisex.html