European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 7, Issue 11, 2020

Parkinson's Disease Identification using Glottal flow analysis

Harisudha Kuresan[1], Dhanalakshmi Samiappan[2*]
[1,2] SRM Institute of Science & Technology, Kattankulathur, Tamilnadu, India

Abstract:
Background:
Parkinson's is a neurological disorder not as rare as projected by maestros in this field, most commonly as it is considered a taboo in general societal norms or rebuffed as an old age syndrome. The complexity in detection resulting from the lack of reliable tests at the early stages of PD makes it hard to deal with chronic stages.
Methods:
This paper focuses on using extracted features from speech signals to detect an abnormality and deducing whether the patient is suffering from PD. Glottal Closure Instants (GCI) detection algorithms such as Speech Event Detection using Residual Excitation and the Mean Based Signal (SEDREAMS) is compared with electroglottographic (EGG) recordings as reference ground. The speech samples are taken from the UCI Machine Learning repository, which contains multiple speakers' recordings.
Results:
The early onset of PD can be predicted at benign stages of the disorder and help in treatment at later stages of Parkinson's disease. SEDREAMS showed an accuracy of 88.7%.
Conclusion:
SEDREAMS is used to diagnose Parkinson's disease early with precise detection of Glottal closing and opening instants.

Keywords: Electroglotography (EGG). Parkinson's disease (PD). Glottal closing instant (GCI). Glottal opening instant (GOI). Speech Event Detection using Residual Excitation and the Mean Based Signal (SEDREAMS)

Parkinson's disease is a rare disorder with neurological motor symptoms such as tremors in hands, arms, legs, face, body swelling, muscle rigidity, and movement problems. PD also has nonmotor symptoms such as slowing thoughts, dementia, focused attention, and speech difficulties. Dr. James Parkinson in 1817 described Parkinson's disease as "shaking palsy" [1].
Early detection or diagnosis and effectiveness of the treatment by correct and constant evaluation can improve PD patients' life expectancy. The assessment of the above two main things is usually performed using the Unified Parkinson's Rating Scale (UPDRS). [2]. Approximately 70%–90% of patients with PD show some form of vocal impairment, as reported by studies conducted, and these studies also suggest that this deficiency may also be one of the earliest indicators of the disease. [3] Specific symptoms can be alleviated by the use of medical treatment, including neuro-pharmacological and neurosurgical methods and early diagnosis of the disease plays a vital role in

improving patients' lives. [2][3]. To obtain higher classification accuracy by removing irrelevant, noisy, and redundant data according to specific criteria, we use a feature selection process. A suitable feature selection method improves the processing rate, predictive accuracy and avoids incomprehensibility. [4]
Using aperiodic vibrations in the voice, voice disorders can be measured by simple acoustic tools. To further improve the clinical usefulness of the voice disorder diagnosis , a complex nonlinear aperiodicity and turbulent, aeroacoustic, non-Gaussian randomness of the sound can be used [1]. During the production of voiced sounds, the speech signal exhibits quasi-periodicity. It happens due to the presence of periodic vibrations of the vocal folds. Thus, the fundamental frequency pitch is defined as the frequency of these vocal folds' vibrations. Also, the inverse of the interval between two successive epochs is termed as the pitch for the laryngeal cycle.[8] The Dynamic Plosion Index (DPI) is an extension of the Plosion Index, which is applied to the pre-processed signal to detect the epochal candidates. The use of DPI is suggestive as these prove to be effective even for low-level voiced segments. Another advantage of DPI is that it does not require preliminary pitch information, useful with a broad range of pitch (speech, music, etc.). [9]. Zero frequency filtering is a method that characterizes the glottal activity from the input speech signals. The FIR filter reduces the computational requirements for zero frequency filtering, including - 1) the use of single-precision floating-point and 2) stability of the filter. [10]

SEDREAM Algorithm:

SEDREAMS algorithm determines the location of GCIs from the speech waveform. GOIs are detected upon modification of the algorithm. This algorithm has two procedures: *i)* determining the existence of intervals in which GCIs appear utilizing a Mean-based signal and *ii)* refining the GCI locations using the residual excitation.

1) Determining the existence of intervals in which GCIs appears utilizing a Mean-based signal:

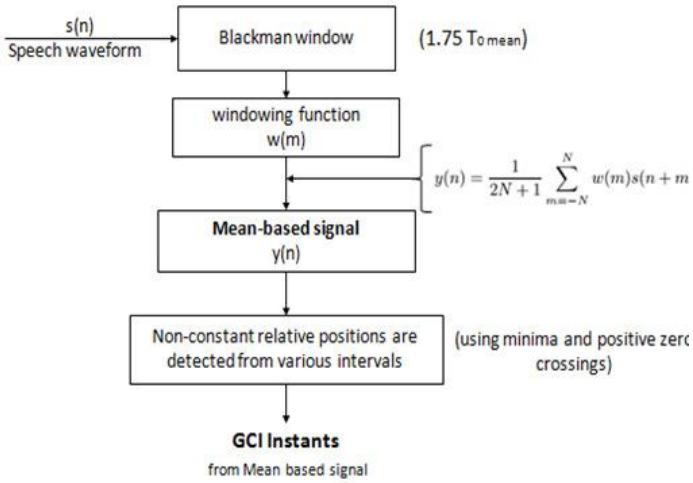

Fig 1. Flowchart for GCIs from the mean based signal

The idea used in this step is that when a stable system is introduced to a stimulus in an impulse, there will be a discontinuity in its output. In the voiced speech, the rapid closing of the vocal cords produces an impulse like excitation. It always causes a discontinuity in the vocal tract's production. Since the zero frequency is effectively secured from time-varying resonances of the vocal tract, it is selected to detect discontinuities. Hence, the mean-based signal is used in this step, which is procured by computing the mean of the sliding window method's speech segments.
A Blackman window of length $1.75T_{0,\text{mean}}$ is in the mean of $T_{0,}$ which is the average pitch period. The windowing function is obtained, which is then applied to the formula for a mean based signal. The mean based signal oscillates around the local pitch period. The GCI's appear at non-constant relative positions. The maxima and minima of the mean-based signal and positive and negative zero crossings are used to detect the G CIs and GOIs. But the output obtained isn't sufficient to detect the GCIs and GOIs accurately.

2) Refining the GCI and GOI locations using the residual signal.

The intervals in which the GCIs and GOIs occur in the above step consist of short indistinct regions where the instants would occur. This step refines these intervals. The LP residual signal is used to check for large discontinuities in the signal within a given interval, containing the GCI and GOI locations. The LP residual signal is obtained from the speech signal in how the LP residual signal is obtained in the DYPSA algorithm.

Combining the intervals obtained from the mean based signal and the residual signal and subjecting it to a peak picking algorithm gives the GCIs and GOIs accurate and unambiguous.

Methodology:

1. PRAAT

Praat is speech analysis software mainly used for identification and analyzing the speech signal in phonetics. This software helps in converting the dual-channel or multi-channel speech signal to a mono channel file.

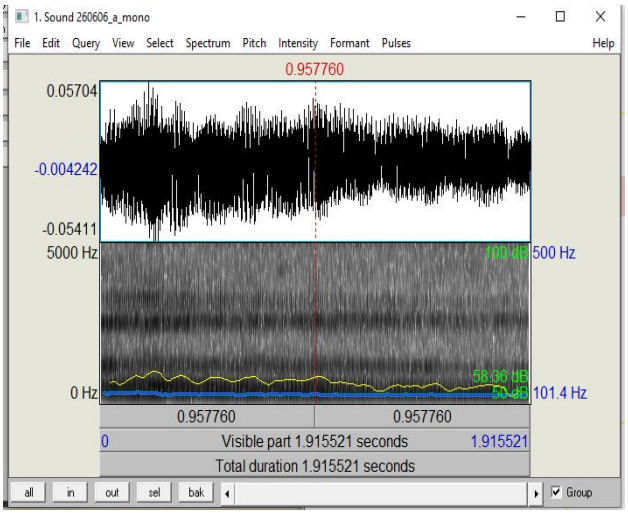

Fig 2: Praat waveform detection

2. SOUND FORGE:

Sound forge is a sound editing application used to truncate the speech files to a uniform 100ms interval. This application was also used to create the text files necessary for speech signal processing in the MATLAB program.

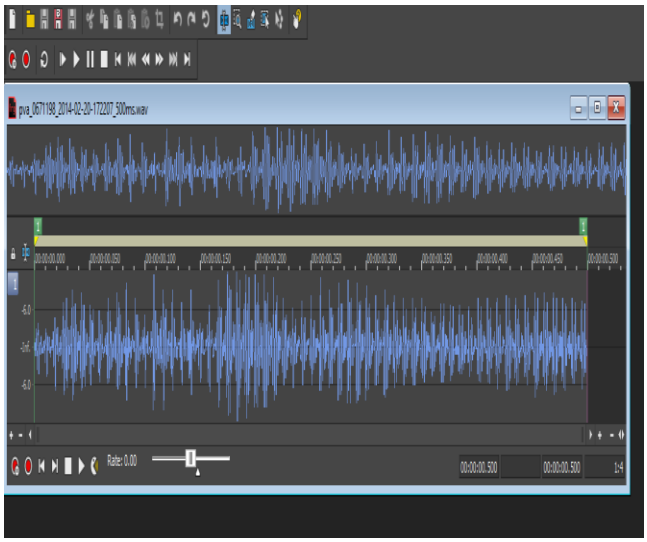

Fig 3: Sound forge editing application

3. SEDREAMS:

Speech Event Detection using the Residual Excitation and a Mean-based Signal (SEDREAMS) algorithm is a method that provides reliable results for locating both GCIs and GOIs from the speech waveform. Since we are only focusing on GCIs, the GOI location determination by the SEDREAMS algorithm is omitted. The two steps involved in this technique are i) short intervals detection where GCIs are expected to occur and ii) the GCI location refinement within these intervals. [5]

1) The intervals of presence are determined using a mean-based signal. Reflection of a discontinuity in the excitation is observed over the whole spectral band, including the zero frequency. This observation inspires the analysis to focus on a mean-based signal. The speech waveform is denoted as s(n), the mean-based signal y(n) is defined as:

$$y(n) = \frac{1}{2N+1} \sum_{m=-N}^{N} w(m)s(n+m) \qquad (1)$$

Where windowing function w(m) is of length 2N + 1.

2) GCI location refining is done using the residual excitation. The previous step indicates the intervals of presence that give short fuzzy regions where a GCI should happen. The next step in the process has the primary purpose of refining these intervals to determine the precise GCI

locations inside it. Therefore, assuming that the signal's most considerable discontinuity within the provided interval corresponds to the GCI location, the LP residual is inspected.

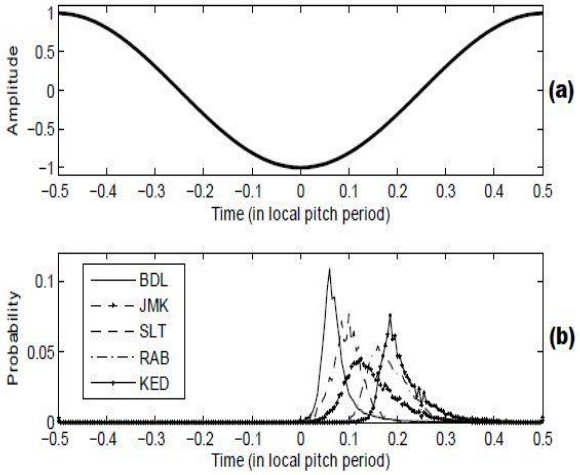

Fig 4. Their actual representation shows GCI locations.

Simulation Results:

The speech signal of the voice samples is pitch tracked, and polarities are detected. The signal is plotted with Electroglottography (EGG) signal as a reference signal. The EGG signal of a patient gives an accurate number of glottal closure instants (GCI) and is used here to reference the algorithm. The dataset was referred from the UCI Machine learning repository.

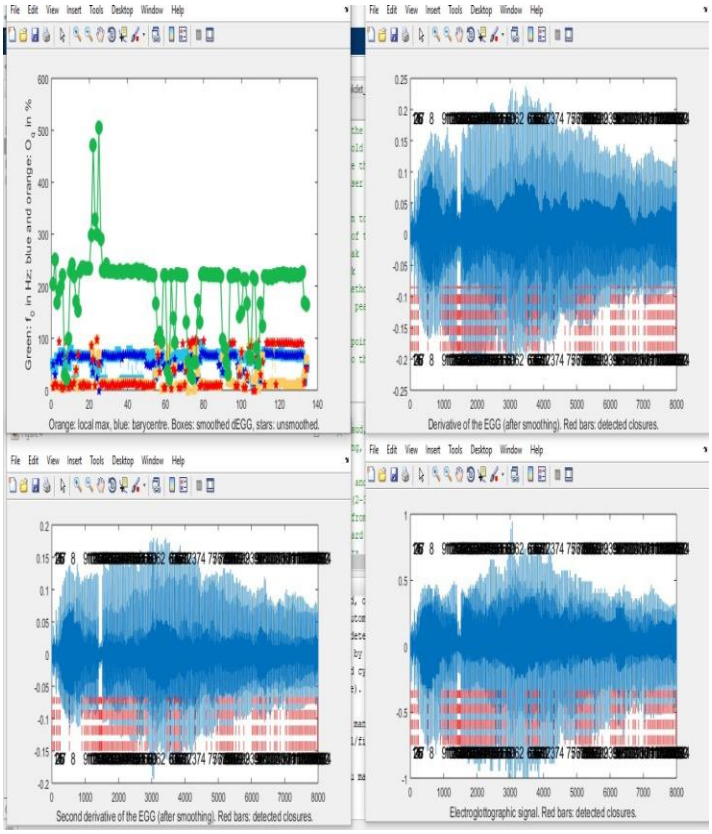

Fig 5: 1st and 2nd derivative graphs of EGG

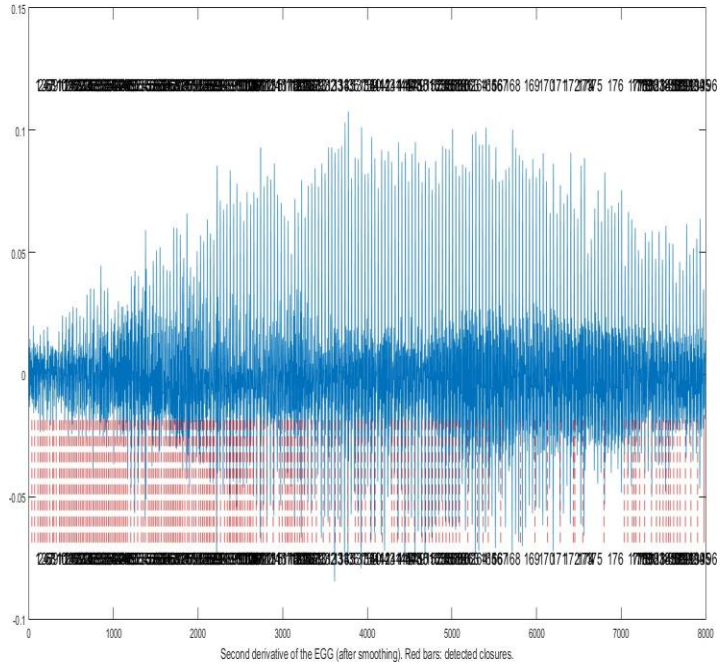

Fig 6: Pre-processed EGG Signal

SEDREAMS is the state of art algorithm used for detecting the correct number of GCI in a speech signal.

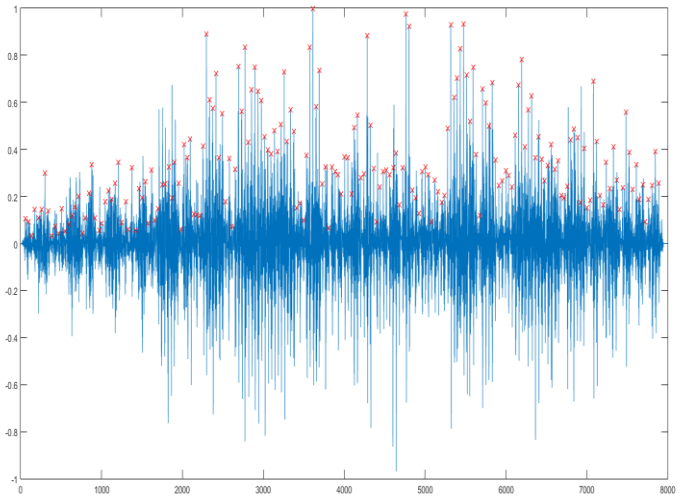

Fig 7. The figure shows the SEDREAMS signal of a PD patient where the x and y-axis are time vectors, and normalized speech signal and x represent GCI.

The algorithm pitch tracks the input signal, finds the threshold, and analyses the signal to detect polarities and recognize the positive polarities as GCIs. When we use these GCI points and

compare them, it is evident that we can differentiate between a normal and a Parkinson's patient. After analyzing the graphs, the threshold value is fixed to 200, which depicts that if the GCI value is equal or more than the threshold (i.e.,) 200, then the person is a healthy individual. Less than 200 is at high risk of being a Parkinson affected individual. Also, in the case of GCI, far much less than the threshold value, it can be curtained that the particular individual is in the advanced stage of Parkinson's.

Comparison of the EGG signal and SEDREAMS signal:

The speech signal has been given as input to both the programs to find a person's GCI. The programs counted the GCIs of 54 people and the standard deviation of the algorithm from the reference signal was found to be 0.89% and an accuracy of 99.11%.

The "n" number of people (patient and non-patients) was observed, and it was observed that for 49 PD patients, the threshold of GCI count was found to be 220. Only healthy people had a count above 220 GCI count. PD patients showed a pattern of newer and younger patients of lesser age or under medication or only in the early stages of PD had GCI count closer to 220 or just falling short of 220. In contrast, people of middle ages or in intermediate PD stages had GCI count in the range of 200-100. Patients in later stages or greater age had GCI count even below 100, as shown in the figure.

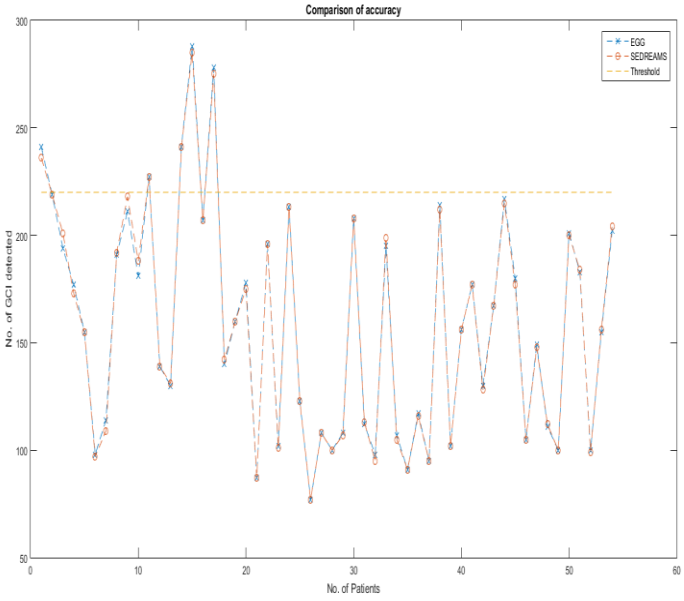

Fig.8. Comparison of accuracy:

The GCI counts of each patient's EGG and SEDREAMS is plotted. The x and y-axis are the numbers of patients, and the number of GCI detected.

Performance Measures:

The system reliability is found by the algorithm's accuracy and standard deviation w.r.t to the reference signal.

The accuracy of the algorithms:

• Accuracy to ±0.25 ms (A25): accuracy is defined by the percentage of detections for which the identification error must be $t \leq 0.25$ ms.

• Identification Accuracy (IDA): The standard deviation of the identification error ζ or how much on an average deviation was observed from the reference signal; it was found to be 0.89%

The modification of voice characteristics is facilitated by speech technology platforms such as statistical speech synthesis and voice transformation systems. The results suggest that improvement in the GCI detection rate occurs due to a reduction in false positives. GCI identification accuracy is further improved when there is a lack of prominent LP residual peaks for a breathy and harsh voice. [11] The parameters can be calculated using the formula

- Miss Alarm Rate = FN/(TP+FN)
- False Alarm Rate = FP/(TP+FP)
- Sensitivity = TP/(TP+FN)
- Accuracy = (TP+TN)/(TP+TN+FP+FN)

Where TP-True positive, TN-True negative, FP-False positive, and FN-False negative.

Continuous and sustained analyses are done for calculating the accuracy of SEDREAMS and EGG. For determining the accuracy, we calculate the deviation of GCI locations obtained from the EGG signal after processing with that of the SEDREAMS.

Performance Parameters	Result
FAR	10.4%
MAR	10.98%
Sensitivity	88.5%
Identification Rate	88.7%

Table 1: Performance metrics of SEDREAMS

Conclusion:

SEDREAMS program worked efficiently to determine the number of GCI locations, and the accuracy rate was as high as 88%. The threshold value was set at 220. Individuals with a number of GCI points less the 220 were acknowledged to be suffering from PD. Ensemble models perform better in precise detection of Glottal Closure Instants (GCI's) necessary for accurate detection of Parkinson's disease. Here the SEDREAMS was found out to be efficient in the precise detection of the number of GCI locations, thereby identifying the normal and abnormal patient. Hence speech signals can be used for the determination of Parkinson's disease in an individual. Further comparison of other state of the art algorithms can be done, and optimization can be done for GCI detection of Parkinson disease.

References:

1. Betul Erdogdu Sakar, M. Erdem Isenkul, C. Okan Sakar, Ahmet Sertbas, Fikret Gurgen, Sakir Delil, Hulya Apaydin, Olcay Kursun.(2013) Collection and Analysis of a Parkinson's Speech Dataset With Multiple Types of Sound Recordings. *IEEE Journal of Biomedical and Health Informatics.* Vol 17, No. 4.

2. Alex Frid, Edmond J. Safra, Hananel Hazan, Lorrey L. Lokey, Dan Hilu, Larry Manevitz, Lorraine O. Ramig, Shimon Sapir. (2014). Computational Diagnosis of Parkinson's Disease Directly from Natural Speech using Machine Learning Techniques. *2014 IEEE International Conference on Software Science, Technology, and Engineering*.

3. J. Rusz, R. Cmejla, H. Ruzickova, and E. Ruzicka. (2011). Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *J. Acoust. Soc. Am.* 129(1), 350-367.

4. Ismail El Moudden, Mounir Ouzir and Souad ElBernoussi,. (2017). Feature selection and extraction for class prediction in dysphonia measures analysis: A case study on Parkinson's disease speech rehabilitation. *Technology and Health Care* 25 693–708 DOI 10.3233/THC-17082.

5. Thomas Drugman, Mark Thomas, Jon Gudnason, Patrick Naylor, Thierry Dutoit,(2012). Detection of Glottal Closure Instants from Speech Signals: a Quantitative Review. *Article in IEEE Transactions on Audio Speech and Language Processing.* DOI: 10.1109/TASL.2011.2170835

6. Jindrich Matousek, Daniel Tihelka.(2017) Classification-Based Detection of Glottal Closure Instants from Speech Signal. *INTERSPEEH.* August 20–24, 2017, Stockholm, Sweden

7. Jindrich Matousek, Daniel Tihelka(2018). Glottal Closure Instant Detection from Speech Signal Using Voting Classifier and Recursive Feature Elimination. *INTERSPEECH* 2018,2-6 September 2018, Hyderabad

8. P.Sujith, A.P.Prathosh, R.A.G.,and P.K.Ghosh(2015). An error correction scheme for GCI detection algorithms using pitch smoothness criterion. *INTERSPEECH*, Dresden, Germany, 2015, pp. 3284–3288.

9. A. P. Prathosh, T. V. Ananthapadmanabha, and A. G. Ramakrishna.(2013). Epoch extraction based on integrated linear prediction residual using plosion index. *IEEE Transactions on Audio, Speech and Language Processing.* Vol. 21, no. 12, pp. 2471–2480, 2013.

10. S. S. Srinivas and K. Prahallad(2012). An 'FIR implementation of zero frequency filtering of speech signals,.*IEEE Trans. Audio, Speech, Lang.Process.*, vol. 20 no.9, pp. 2613–2617.

11. Kane and C. Gobl.. (2013). Evaluation of glottal closure instant detection in a range of voice qualities. *Speech Communication.* vol. 55, pp. 295–314.. [12]J. C. Catford,*(1977)* Fundamental Problems in Phonetics. Bloomington, IN: Indiana Univ. Press, 1977.

12. E. Moulines and F. Charpentier (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication.* vol. 9, no. 5–6, pp. 453–467,

13. N. D. Gaubitch and P. A. Naylor. (2007) Spatiotemporal averaging method for enhancement of reverberant speech. Proc. IEEE Int. Conf. DigitalSignal Process. (DSP), Cardiff, U.K.

14. D. Y. Wong, J. D. Markel, and J. A. H. Gray.(1979). Least squares glottal inverse filtering from the acoustic speech waveform. IEEE Trans.Acoust., Speech, Signal Process. vol. ASSP-27, no. 4, pp. 350–355,

15. Y. Stylianou (2001). Applying the harmonic plus noise model in concatenative speech synthesis. *IEEE Trans. Speech Audio Process.*. Vol. 9, no.1, pp. 21–29, Jan. 2001.

16. T. Drugman, G.Wilfart, and T. Dutoit, "A deterministic plus stochastic model of the residual signal for improved parametric speech synthesis," in *Proc. Interspeech Conf.*, 2009.

17. M. D. Plumpe, T. F. Quatieri, and D. A. Reynolds. (1999). Modeling of the glottal flow derivative waveform with application to speaker identification. IEEE Trans. Speech Audio Process..vol. 7, no. 5, pp. 569–576.

18. P. A. Naylor, A. Kounoudes, J. Gudnason, and M. Brookes. (2007). Estimation of glottal closure instants in voiced speech using the DYPSA algorithm,. *IEEE Trans. Speech Audio Process.* vol. 15, no. 1, pp. 34–43.

19. M. R. P. Thomas, J. Gudnason, and P. A. Naylor. (2012). Estimation of glottal opening and closing instants in voiced speech using the YAGA algorithm. *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 1, pp:82–91.

20. T. V. Ananthapadmanabha and B. Yegnanarayana. (1979). Epoch extraction from linear prediction residual for identification of closed glottis interval. *IEEE Trans. Acoust., Speech, Signal Process.* Vol. ASSP-27, no. 4, pp. 309–319.

21. C. Ma, Y. Kamp, and L. F. Willams. (1994). A Frobenius norm approach to glottal closure detection from the speech signal. *IEEE Trans. Speech Audio Process.* vol. 2, no. 2, pp. 258–265.

22. K. S. R. Murty and B. Yegnanarayana. (2008) Epoch extraction from speech signals. *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 8, pp.1602–1613.613.

23. S. K. Nataraj, F. Al-Turjman, A. H. Adom, R. Sitharthan, M. Rajesh and R. Kumar, "Intelligent Robotic Chair with Thought Control and Communication Aid Using Higher Order Spectra Band Features," in IEEE Sensors Journal, doi: 10.1109/JSEN.2020.3020971.

24. B. Natarajan, M. S. Obaidat, B. Sadoun, R. Manoharan, S. Ramachandran and N. Velusamy, "New Clustering-Based Semantic Service Selection and User Preferential Model," in IEEE Systems Journal, doi: 10.1109/JSYST.2020.3025407.

25. Ganesh Babu, R.; Obaidat, Mohammad S.; Amudha, V.; Manoharan, Rajesh; Sitharthan, R.: 'Comparative analysis of distributive linear and non-linear optimised spectrum sensing clustering techniques in cognitive radio network systems', IET Networks, 2020, DOI: 10.1049/iet-net.2020.0122

26. Rajalingam, B., Al-Turjman, F., Santhoshkumar, R. et al. Intelligent multimodal medical image fusion with deep guided filtering. Multimedia Systems (2020). https://doi.org/10.1007/s00530-020-00706-0