2016

Analysis of glottal source parameters in Parkinsonian speech

Jane Hanratty
Technological University Dublin, jane.hanratty@tudublin.ie

Catherine Deegan
Technological University Dublin, catherine.deegan@tudublin.ie

Mary Walsh
St Mary's Hospital, Dublin

See next page for additional authors

Recommended Citation

Hanratty J, Deegan C, Walsh M, Kirkpatrick B. Analysis of glottal source parameters in Parkinsonian speech. Annu Int Conf IEEE Eng Med Biol Soc. 2016 Aug;2016:3666-3669. doi: 10.1109/EMBC.2016.7591523. PMID: 28269089.

Analysis of Glottal Source Parameters in Parkinsonian Speech

Jane Hanratty[1], Catherine Deegan[1], Mary Walsh[2] and Barry Kirkpatrick[1]

Abstract — **Diagnosis and monitoring of Parkinson's disease has a number of challenges as there is no definitive biomarker despite the broad range of symptoms. Research is ongoing to produce objective measures that can either diagnose Parkinson's or act as an objective decision support tool. Recent research on speech based measures have demonstrated promising results. This study aims to investigate the characteristics of the glottal source signal in Parkinsonian speech. An experiment is conducted in which a selection of glottal parameters are tested for their ability to discriminate between healthy and Parkinsonian speech. Results for each glottal parameter are presented for a database of 50 healthy speakers and a database of 16 speakers with Parkinsonian speech symptoms. Receiver operating characteristic (ROC) curves were employed to analyse the results and the area under the ROC curve (AUC) values were used to quantify the performance of each glottal parameter. The results indicate that glottal parameters can be used to discriminate between healthy and Parkinsonian speech, although results varied for each parameter tested. For the task of separating healthy and Parkinsonian speech, 2 out of the 7 glottal parameters tested produced AUC values of over 0.9.**

I. Introduction

Parkinson's disease (PD) is a neurodegenerative disease estimated to have between seven to ten million cases worldwide [1]. It is the second most common neurodegenerative disease after Alzheimer's and age is the most significant risk factor for PD onset. Incidence rates of Parkinson's are predicted to increase consistently with increases in life expectancy. Currently no cure exists, but with early diagnosis and intervention, quality of life can be improved in most cases [2]. Timely diagnosis and ongoing monitoring of symptoms in PD is critical, but presents a number of challenges.

It is known that a reduction in dopamine producing cells in the basal ganglia of the brain causes Parkinson's, although the underlying cause of the loss of these cells is unknown. The reduction in these cells has an impact on the function of neural circuitry in the basal ganglia resulting in PD symptoms. Typical symptoms of PD include muscular rigidity, bradykinesia (slow movement), resting tremor, postural instability and cognitive impairment [3]. In certain cases Parkinson's like symptoms can be the result of exposure to neurotoxins or drugs and is referred to as Parkinsonism. It is estimated that approximately 90% of PD patient's exhibit speech related symptoms [4]. Speech related symptoms reported include under-articulation, mono-pitch, reduced volume, harsh or breathy speech and vocal tremor.

Diagnosis of PD is a complex process and relies on subjective evaluations by experts in a clinical setting. A commonly adopted approach is to employ the Unified Parkinson's Disease Rating Scale (UPDRS) [5]. The UPDRS consists of 44 sections under 3 categories (1) mentation, behaviour and mood, (2) activities of daily living and (3) motor control. A score is assigned for each section, with the overall score used as an indicator of disease progression. This procedure is subjective and time consuming, requiring the patient to attend clinics and be assessed by a medical expert, to monitor disease progression. Despite the range of symptoms in PD, no definitive biomarkers exist [3] and there is a lack of objective measures to facilitate diagnosis. Recent research results have made progress in this area, most notably in speech related measures [6]. A need exists for the development of objective measures that can support early diagnosis and monitor disease progression. Ideally such a measure would be non-invasive and could be acquired outside the clinical environment without the need for expert assistance.

The aim of this study is to investigate the behaviour of the glottal waveform, estimated from the speech signal in recordings of Parkinsonian speech, and identify what parameters behave differently in Parkinsonian speech in contrast with healthy speech. This research was motivated by studies that have identified the potential of speech based measures to diagnose PD [6-11] and also from studies describing glottal behaviour in PD from laryngoscope examination [12].

Section II of this paper describes the necessary background on glottal source analysis. Section III describes the experiment conducted and provides details of the speech database used in the study and how it was analyzed. In section IV the results are presented and section V presents the conclusions.

II. Background on Glottal Source Analysis

In order to characterise the behaviour of the glottal signal two key steps were required; estimation of the glottal signal from the speech signal and parameterization of the glottal signal. Estimation of the glottal signal is typically achieved by an inverse filtering algorithm [13]. Estimation of glottal parameters typically requires computing the optimum best fit parameters for a selected model of the glottal waveform.

Estimation of the glottal waveform is a well-known problem in speech processing and continues to be an active area of research, in particular for more challenging scenarios such as pathological speech [14]. Studies comparing methods to estimate the glottal waveform have indicated that the Iterative Adaptive Inverse Filter (IAIF) is a consistent and robust method [15]. The IAIF algorithm was adopted as the algorithm to estimate the glottal waveform in this study.

[1]Jane Hanratty, Barry Kirkpatrick and Catherine Deegan are with the Dept. of Engineering in the Institute of Technology, Blanchardstown, Dublin, Ireland (e-mail: jane.hanratty@itb.ie).

[2]Mary Walsh is with the Speech and Language Dept. of St Mary's Hospital, Phoenix Park, Dublin, Ireland.

A number of models of the glottal flow waveform exist and the most widely adopted is the Liljencrants-Fant (LF) model [16]. In this study the LF model was employed for glottal parameters estimation. The LF model and glottal parameters considered in this study are discussed in the remainder of this section.

A. Liljencrants-Fant model of the glottal source

The LF model is a five parameter model of the differentiated glottal flow signal. The model can be fully specified by the timing parameters t_e, t_p and t_a along with the pitch period T and the single amplitude parameter E_e. The model has two components representing the open and closed phases of the glottal cycle. The glottal closure instant (GCI) occurs at time t_e with amplitude E_e. The GCI is the primary landmark within the glottal signal. Although the model has further parameters they are implicitly constrained by the five primary parameters and model properties. The glottal flow derivative, $g(t)$, is defined in the LF model as in (1).

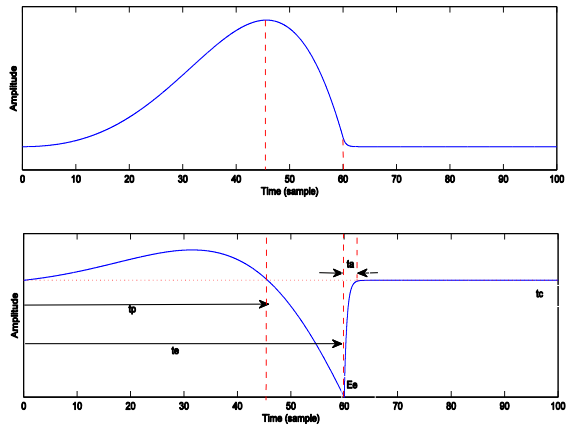
$$g(t) = \begin{cases} E_0 e^{\alpha t} \sin(\omega_g t), & 0 \le t \le t_e \\ \frac{E_e}{\epsilon t_a}\left[e^{-\epsilon(t-t_e)} - e^{\epsilon(t_c-t_e)}\right], & t_e < t \le T \end{cases} \quad (1)$$

This model is illustrated in Fig. 1. Following estimation of the glottal signal, optimisation algorithms are typically employed to estimate the parameters of the model [17]. In this study the LF model parameters E_e, t_e and t_p were considered.

B. NAQ and QOQ parameters

Two popular parameterisations of the glottal signal are the quasi-open quotient (QOQ) [18] and the normalized amplitude quotient (NAQ) [19]. The QOQ is measured from the instant the glottal pulse reaches 50% of its maximum value at time t_{q0} until the instant the pulse amplitude falls below this threshold at time t_{q1}, this duration is normalized with respect to the pitch period T, as in (2).

Figure 1. Illustration of the LF model, indicating the glottal pulse waveform (above) and glottal flow derivative (below).



$$QOQ = \frac{t_{q0} - t_{q1}}{T} \quad (2)$$

The NAQ is computed from the amplitude parameter of the glottal flow derivative E_e and the maximum amplitude of the glottal pulse A and is normalized with respect to the pitch period, as in (3).

$$NAQ = \frac{A/E_e}{T} \quad (3)$$

The QOQ and NAQ are often selected as glottal parameters as they are robust to measurement noise and do not require the difficult task of estimating the instant of glottal opening. Both the QOQ and NAQ were tested in this study.

C. R parameters - R_g, R_k and R_d

A transformed set of parameters, the R parameters, can be computed from the LF model parameters and characterise the shape of the LF model pulse [20].

$$R_g = \frac{T}{2t_p} \quad (4)$$

$$R_k = \frac{t_e - t_p}{t_p} \quad (5)$$

$$R_d = 1000\left(\frac{A}{E_e}\frac{f0}{110}\right) \quad (6)$$

The parameter R_g defined in (4) represents the glottal formant normalized to the pitch frequency. The parameter R_k defined in (5) is a measure of asymmetry in the glottal cycle and R_d defined in (6), captures the co-variation of the LF parameters [20] and is dependent on the pitch frequency $f0$. The parameter R_d is proportional to the NAQ and as such only the NAQ parameter was tested in this study. Both R_g and R_k were tested.

III. The Experiment

The objective of this study was to identify if the parameters of the glottal signal have distinct characteristics in PD speech as compared to the glottal parameters of healthy speech. To evaluate this each candidate glottal parameter was tested in a binary classification task to discriminate between PD and healthy speech. Receiver operating characteristic (ROC) curves and the area under the ROC curve (AUC) [21] were employed to quantify the performance of each glottal parameter in discriminating between healthy and Parkinsonian speech.

A. Parkinsonian speech database

Patients were recruited for this study in St Mary's Hospital in Dublin, Ireland. Ethical approval was granted by both collaborating institutions involved in the study, the Institute of Technology Blanchardstown and St Mary's Hospital. Patients presenting to St. Mary's hospital for assessment, respite or rehabilitation with a diagnosis of PD were asked to participate in the study. All participants provided witnessed verbal consent, and written informed consent within capacity. No distinction was made between Parkinson's and Parkinsonism for the purposes of this study. The only information retained on each participant was the speech recording, no further data was retained.

A total of 22 recordings were made of patients attempting to make a sustained 'ah...' sound. The recordings were made using a ZOOM H2n portable recorder in a quiet environment. Following data collection, 6 recordings were discarded as they did not contain a sustained phonation of at least 500 ms.

B. *Healthy speech database*

The database used for healthy speech was taken from [22]. Healthy recordings of sustained 'ah… ' sounds provided in this database were used, which was 52 in total. Note that 2 recordings were discarded as reliable glottal estimates could not be achieved. The remaining 50 files were used in the study.

C. *Glottal feature extraction from database*

For each speaker in both healthy and PD databases 500 ms of voiced speech was extracted for analysis. The analysis was performed using Aparat software [23]. For each speaker the extracted speech had the IAIF algorithm applied to estimate the glottal waveform followed by estimating the LF model and estimation of each of the glottal parameters under consideration in this study. For each speaker an average was computed for each glottal parameter, from all instances of that parameter computed over the 500 ms analysis window. Note that LF parameters t_e and t_p were normalized with respect to the pitch period T to ensure that variations in the pitch between speakers do not impact the results. The LF amplitude parameter E_e was normalized using min-max scaling [24] to ensure that relative differences in amplitude due to different recording conditions do not impact the results.

IV. Results

The results for each candidate glottal parameter are presented in Table 1. This table provides the mean parameter value μ and standard deviation σ computed for all speakers for both PD speech and healthy speech for each glottal parameter. Notable differences can be observed in the values for PD and healthy speech.

The healthy and Parkinsonian distributions of each parameter were analyzed using ROC curves. AUC values were computed to quantify the level of separation between the distributions of PD and healthy speech for each glottal parameter. An AUC value of 0.5 indicates no discriminating information for a given parameter and a value of 0 or 1 indicates full separation of PD and healthy distributions. The AUC value can be interpreted as the probability of making a correct classification, with an AUC value of 0.5 representing the level of pure chance. The AUC values for each parameter are presented in Table 2. The standard error for each AUC value is also presented and was computed according to [21]. To facilitate the comparison of AUC values that are above and below 0.5, each AUC value is also presented as a performance metric. To compute the performance, AUC values below 0.5 are subtracted from 1 and converted to a percentage and AUC values greater than 0.5 are converted directly to a percentage. The ROC curves for all parameters tested are presented in Fig. 2.

The results for the glottal parameters NAQ and QOQ both exceed 90%, indicating that these parameters were different in PD and healthy speech in the test databases. The parameter E_e was found to have the lowest performance with a value of 55.7%. The remaining parameters had intermediate performance ranging from 64.5% to 79%.

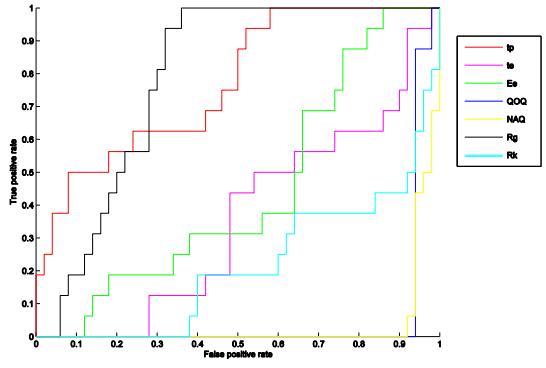
TABLE I. Results - Mean and Standard Deviations of Glottal Parameters for PD and Healthy Speech

Glottal Parameter	PD Speech		Healthy Speech	
	μ	σ	μ	σ
NAQ	0.045	0.013	0.119	0.300
QOQ	0.179	0.059	0.656	0.230
E_e	0.587	0.082	0.601	0.118
t_e	0.902	0.043	0.922	0.047
t_p	0.701	0.100	0.541	0.167
R_g	2.087	0.663	1.502	1.230
R_k	0.366	0.275	1.177	1.227

TABLE II. Results - AUC values for each glottal parameter tested

Glottal Parameter	Results		
	AUC	*Standard Error*	*Performance %*
NAQ	0.03	±0.020	96.60%
QOQ	0.06	±0.027	94.50%
E_e	0.44	±0.081	55.70%
t_e	0.36	±0.075	64.50%
t_p	0.77	±0.075	77.10%
R_g	0.79	±0.072	79.00%
R_k	0.21	±0.062	78.60%

Figure 2. ROC curves for each glottal parameter tested for the task of discriminating between PD and healthy speech.



V. Conclusion

The results presented in this study indicate that speech related symptoms of PD are evident in the glottal flow signal. Particularly notable results were recorded for the NAQ and QOQ glottal parameters, with performance of over 90% recorded for the task of discriminating between healthy and PD speech in the test databases.

It should be noted that estimating the parameters of the glottal waveform is a challenging task and that certain parameters can be more robustly estimated. In this study the parameters that can be robustly estimated provided the best results, namely the NAQ and QOQ. Most timing parameters, with the exception of the QOQ, are sensitive to noise. The timing parameters t_e and t_p produced AUC values above the level of chance but were significantly outperformed by the more reliable QOQ timing parameter. This indicates that glottal timing information is important in PD but the specific representation employed can influence results. The least reliable estimate in this study was the amplitude based parameter E_e. The amplitude of recordings was not sufficiently controlled to use E_e values without normalization. The normalization process employed is likely to have removed information that could have aided diagnosis. The NAQ is an amplitude parameter that depends on E_e but is independent of signal scaling and was found to produce a significantly higher AUC value then E_e. This indicates that glottal amplitude information is important in PD but the specific representation employed can influence results.

The results presented are positive indicators that both timing and amplitude based measures derived from the glottal signal show significant potential as objective measures of PD.

Acknowledgment

The authors would like to acknowledge the contribution of the participants of the study and Aideen Lawlor for facilitating the arrangements that made the data collection possible. The authors would also like to acknowledge the Institute of Technology Blanchardstown for funding this research project.

References

[1] Parkinson's Disease Foundation, "Statistics on Parkinson's," 2015. [online]. Available: http://www.pdf.org/ . [Accessed 10/12/15].

[2] Singh, N., Pillay, V., Choonara, Y.E. 2007 Advances in the treatment of Parkinson's disease, Progress in Neurobiology, 81, 29-44.

[3] A.E. Lang, A.M. Lozano, "Parkinson's disease", The New England Journal of Medicine, 339 (1998), pp. 1044–1053.

[4] A. K. Ho, R. Iansek, C. Marigliani, J. L. Bradshaw, and S. Gates, "Speech impairment in a large sample of patients with Parkinson's disease," Behav. Neurol., vol. 11, pp. 131–137, 1998.

[5] Ramaker, C., Marinus, J., Stiggelbout A.M., and van Hilten, B.J. 2002 Systematic evaluation of rating scales for impairment and disability in Parkinson's disease, Movement Disorders, 17, 867-876.

[6] Little, M.A., McSharry, P.E., Hunter, E.J., Spielman, J., Ramig, L.O. 2009 Suitability of dysphonia measurements for telemonitoring of Parkinson's disease, IEEE Transactions Biomedical Engineering, 56, 1015-1022.

[7] A. Tsanas, M.A. Little, P.E. McSharry, J. Spielman, L.O. Ramig. "Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease", IEEE Transactions on Biomedical Engineering, 59(5):1264-1271, 2012.

[8] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," IEEE Trans. Biomed. Eng., vol. 56, no. 4, pp. 1015–1022, Apr. 2009.

[9] A. Tsanas, M.A. Little, P.E. McSharry, L.O. Ramig, "Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity", Journal of the Royal Society Interface, **8**(59):842-855, 2010.

[10] B. Harel, M. Cannizzaro, and P. J. Snyder, "Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease: A longitudinal case study," Brain Cognition, vol. 56, pp. 24–29, 2004.

[11] I. Gath, E. Yair, "Analysis of vocal tract parameters", Journal of the Acoustical Society of America, 84, 1628, 1988.

[12] Hanson DG, Gerratt BR, Ward PH. "Cinegraphic observations of laryngeal function in Parkinson's disease." Mar;94(3):348-53, 1984.

[13] T. F. Quatieri, *Discrete-Time Speech Signal Processing: Principles and Practice,* New Jersey: Prentice Hall PTR, 2002, p. 55-67.

[14] Thomas Dubuisson, PhD Thesis, "Glottal Source Estimation and Automatic Detection of Dysphonic Speakers". 2010/11??

[15] T. Drugman, B. Bozkurt, and T. Dutoit, "Comparative study of glottal source estimation techniques," Comput. Speech Lang., vol. 26, no. 1, pp. 20–34, Jan. 2012.

[16] G. Fant, J. Liljencrants, and Q. Lin, "A four parameter model of glottal flow", STL-QPSR, vol. 4, no. 1985, pp. 1–13, 1985.

[17] H. Strik and L. Boves, "Automatic estimation of voice source parameters", in Proceedings International Conference on Spoken Language Processing (ICSLP)"94, pp. 155–158, 1994.

[18] Hacki, T., (1989) "Klassifizierung von glottisdysfunktionen mit hilfe der elektroglottographie" Folia Phoniatrica, pp. 43-48.

[19] Alku, A. and Backstrom, T., "Normalized amplitude quotient ¨ for parameterization of the glottal flow" J. Acoust. Soc. Am., 112(2):701–710, 2002.

[20] Fant, G., Liljencrants, J., and Lin, Q. The LF-model revisited. transformations and frequency domain analysis. KTH, Speech Transmission Laboratory, Quarterly Report, 1995.

[21] Hanley, J. A. and McNeil, B. J. The meaning and use of the area under the receiver operating characteristic (ROC) curve. Radiology, 143:29–36, 1982.

[22] Childers,D.G. (2000) Speech Processing and Synthesis Toolboxes. New York: John Wiley & Sons inc.

[23] [Online], 2008. http://aparat.sourceforge.net/index.php/main page. TKK Aparat Main Page.

[24] Jiawei Han (Author), Micheline Kamber (Author), Jian Pei, Data Mining: Concepts and Techniques, Third Edition (The Morgan Kaufmann Series in Data Management Systems) 3rd Edition, 2012.