Estimation of Parkinson's disease severity from voice features of vowels and consonant

Rekha Viswanathan, *Member, IEEE*, Sridhar P Arjunan, *Senior Member, IEEE*, Peter Kempster, Sanjay Raghav, Dinesh Kumar, *Senior Member, IEEE*

***Abstract*— This study has investigated the efficiency of voice features in estimating the motor Unified Parkinson's Disease Rating Scale (UPDRS) score in Parkinson's disease (PD) patients. A total of 26 PD patients (mean age = 72) and 22 control subjects (mean age = 66.91) were recruited for the study. The sustained phonation /a/, /u/ and /m/ were collected in both off-state and on-state of Levodopa medication. The average motor UPDRS for PD off-state patients was 27.31, on-state was 20.42 and that of controls was 2.63. Voice features were extracted from the phonation tasks and were reduced to the most relevant 6 features for each phonation task using the Least Absolute Shrinkage and Selection Operator (LASSO) feature ranking method. The correlation between the reduced features and motor UPDRS was tested using the Spearman correlation coefficient test. AdaBoost regression learner was trained and used for automatically estimating the motor UPDRS score using the voice features. The results show that the vocal features for /m/ performed best by estimating the motor UPDRS score for PD off-state with the mean absolute error (MAE) of 3.52 and 5.90 for PD on-state. This study shows that assessment of voice can be used for day to day remote monitoring of PD patients.**

I. Introduction

Parkinson's Disease (PD) is a progressive neurological condition with loss of motor and non-motor functions of the body [1]. The observable motor and non-motor symptoms can be based on the progression and duration of the disease. The study conducted by [1] reported that in the early stage, the observable symptoms from patients' perspective are slowness, tremor, stiffness, pain, and loss of smell or taste whereas in the later stage the symptoms are mood changes, drooling, sleep problems, tremor and dyskinesia. The most widely used tool for the assessment and severity measurement of PD is the unified Parkinson's disease rating scale (UPDRS) [2]. The UPDRS assessment has four parts: 1. Nonmotor aspects of experiences of daily living, 2. Motor aspects of experiences of daily living, 3. Motor examination and 4. Motor complications [3]. Periodic UPDRS assessment of PD patients facilitates the evaluation of disease progression[4]. More importantly, evaluation of motor fluctuations over time aids in the proper adjustment of medication and clinical decision making [5]. Some of the pitfalls in the assessment of UPDRS are a) intra and inter-rater variability of the score, b) patients should make regular visits to the clinic which is often not possible due to various reasons c) in patients having rigorous motor fluctuations, UPDRS assessments made would sometimes poorly represent the clinical status [6]. By automating the assessment of UPDRS, makes it objective, would facilitate daily observation of motor fluctuations and facilitates better and objective clinical decision making.

Speech disturbances are sometimes considered as one of the first signs of PD [7] and often called Hypokinetic dysarthria (HD). HD is characterized by imprecise articulation, weak voice, cracked speech [8]. HD gets developed in nearly 70-80% of people with PD [9]. HD results directly from the neuropathological effects on the speech production system [10] hence leading to the limitation in movements and coordination of the speech production system [11]. Hence, speech has been used as a non-invasive tool for the diagnosis of PD [11]. Studies have also evaluated how acoustic features changes during the progression of PD [12, 13]. Speech features from speech tasks such as sustained phonation of vowels, reading sentences, diadochokinetic rate (DDK), spontaneous speech, etc. are commonly employed for the detection of PD. The acoustic features extracted from the speech tasks have been reported to correlate with the characteristics like voice quality, breathiness, pitch from the speech perception perspective. The non-conventional speech features are based on the fact that as PD progress the voice becomes more aperiodic and chaotic. Fractal dimension (FD), correlation dimension (D_2), pitch period entropy (PPE), sample entropy (SE) are some of the features in the non-conventional group [9].

The above has shown that it is possible to identify the differences between the voice of PD and controls. The next step is to estimate the severity of the disease from the voice which could facilitate remote monitoring of PD and for observing the efficacy of the medication [14]. Phonatory tasks and sentence reading tasks have been employed to extract both conventional and non-conventional features for the automatic assessment of UPDRS scores[15, 16]. Parkinson's telemonitoring dataset developed by Tsanas and Little [17] employed different regression learners like Random forest (RF), linear Support vector machines (SVM), ridge regression, neural network regression for the automatic estimation of UPDRS scores by most of the studies [18, 19].

In our previous work, the efficiency of voice features from phonation task /m/ in differentiating PD and control subjects was demonstrated [20]. In this study, we investigate the efficiency of voice features extracted from three phonation /a/, /u/ and /m/ in automatically estimating the motor UPDRS score of the patient. While most of the studies have used a common database for the automatic estimation of the UPDRS score, we have tried to evaluate the efficiency of speech

R Viswanathan, S Raghav, DK Kumar are with School of Engineering, RMIT University, Melbourne, Australia. P Kempster is with Monash Health.

SP Arjunan is with SRM Institute of Science and Technology, Chennai, India (corresponding author: sridhar_arjunan@ieee.org).

978-1-7281-1990-8/20/$31.00 ©2020 IEEE

features in the same context by considering both the off-state and on-state PD patients.

II. MATERIALS AND METHODS

A. *Participants*

The study ethics was approved by Monash Health, Australia (LNR/16/MonH/319) and ratified by RMIT University, Australia (BSEHAPP2215KUMAR). Study participants provided written consent before the experiment. Speech task included producing three sustained phonation /a/, /u/ and /m/. The subjects were informed to perform the speech task in their comfortable pitch. Each voice recording contained one sustained phonation. The speech task was recorded using Apple iPhone 6s plus with an omnidirectional head-worn microphone connected to it. The microphone was placed approximately 4 cm from the subject's mouth. The speech task was recorded at a sampling rate of 48 kHz and 16-bit resolution in WAV format. The waveforms for sustained phoneme /a/, /u/ and /m/ is shown in Fig 1 for a single control subject, PD patient in both on-state and off-state of Levodopa medication.

24 patients (14 males and 12 females) with PD were recruited for the study and they were diagnosed based on Square Brain Bank criteria for idiopathic PD. 22 healthy age-matched individuals (8 females and 14 males) were also recruited as control subjects. PD patients with other significant neurological and psychiatric conditions were excluded from the study. Control subjects were recruited in the absence of the following conditions: a) no neurological disorders, b) no limb surgery in the last three months, c) pregnancy, d) no other active medical conditions. All participants were tested by an expert clinical nurse who also performed the motor UPDRS assessment and Montreal Cognitive Assessment (MoCA). Speech task was carried out in both off-state and on-state of Levodopa medication for PD patients. Participant information is presented in Table I.

TABLE I. PARTICIPANT INFORMATION

	PD patients (Mean (SD))	**Control subjects (Mean (SD))**
Total subjects	26 (12 females, 14 males)	22 (8 females, 14 males)
Age, years	72(7.47)	66.91(6.22)
UPDRS off-state	27.31(2.59)	2.64(1.37)
UPDRS on-state	20.42(10.08)	-
Duration of disease, years	5.50(2.96)	-
Total Levodopa mg/day	490.39(313.22)	-
MoCA	27.31(2.59)	28.46(1.37)

B. *Automatic estimation of motor UPDRS score from the voice features*

Conversations preceding and following the phonetic speech tasks were removed manually from the recordings. Filtering of the recorded speech tasks was carried out using a 2nd order bandpass filter of 80 Hz – 24kHz to remove high-frequency noise and low-frequency artifacts. This was followed by feature extraction.

Based on the varying effect of Levodopa medication reported [21, 22], we consider different sets of features for off state and on state PD patients. Perturbation features (jitter and shimmer variants), noise features (harmonics to noise ratio) were extracted using PRAAT[23], glottal and glottal based spectral features were extracted using VoiceSauce [24] and TKK Voice Source Analysis and Parameterization (APARAT) toolbox[25]. The features mentioned in [26, 27] were extracted using algorithms developed in MATLAB. In this study, we have used the Least Absolute Shrinkage and Selection Operator (LASSO) as the feature selection algorithm as this method reduces the prediction error by providing the most relevant predictors. The most relevant 6 features from LASSO ranking were used in the automatic estimation of the motor UPDRS score. The feature selection was performed for both off-state and on-state features. The automatic estimation of the motor UPDRS score was achieved using the AdaBoost (Adaptive Boosting) regression learner. AdaBoost is an ensemble method formulated by Freund and Schapire [28] which combines the weak learners or base decision trees to a strong learner. The performance of the regression learners was measured using leave one out validation method. Apart from that, we have used mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE) and R squared (R^2) values to measure the learner performance.

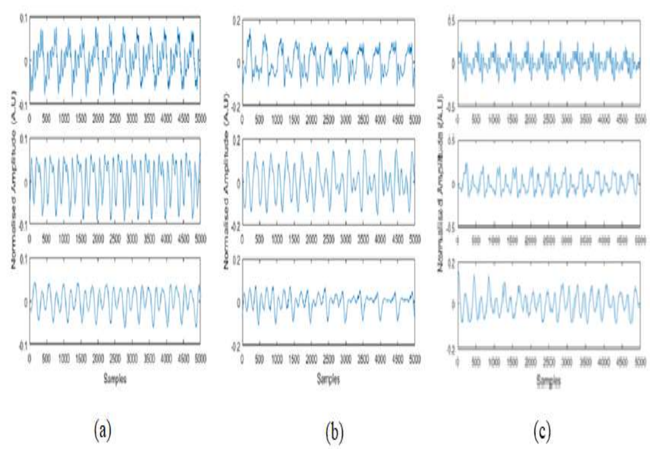

Fig. 1. Illustrative example of the three phonetic signals for (a) PD off-state subject, (b) Control subject and (c) PD on-state subject

C. *Statistical analysis*

The strength of association of the vocal features extracted for the phonation tasks with the motor UPDRS score was evaluated using the Spearman correlation coefficient test. The normality of the features was tested using the Anderson Darling test and then the correlation analysis was performed.

III. RESULTS

Literature reports that Levodopa alters some aspects of voice (features) in PD patients [22, 29] and in the light of this, we decided to have two separate feature sets for finding the association between motor UPDRS and voice features of off-state and on-state PD patients respectively. We have also considered control subjects in the automatic estimation of motor UPDRS.

The feature extraction step was followed by feature reduction using LASSO learner and 6 features from off-state and on-state for each phonetic task were selected for the automatic estimation of motor UPDRS score. Before implementing the regression model for the automatic estimation, the Spearman correlation test was carried out to determine the statistical correlation of these features with the motor UPDRS score. The results are as shown in Table II and III for off state and on state respectively.

TABLE II. SPEARMAN CORRELATION OF **UPDRS** OFF STATE VOICE FEATURES WITH MOTOR UPDRS

Voice features /a/	r	Voice features /u/	r	Voice features /m/	r
F_2	+0.670	A_2	+0.192	F_2	+0.704
Fractal dimension	-0.341	Jitter (local, absolute)	+0.285	H_1-A_3	+0.612
MFCC C_2	+0.089	F_1/F_2	-0.070	F_2 - F_1	+0.658
MFCC C_4	-0.119	Fractal dimension	-0.405	H_1	+0.477
SD F_0	-0.278	SD F_0	+0.276	MFCC C_4	+0.128
OQ_a	-0.348	H_2	-0.269	F_{0min}	+0.214

AdaBoost regression was implemented for the automatic estimation of motor UPDRS. For each phonation task, the most relevant 6 voice features from LASSO feature reduction were provided to the AdaBoost regression learner to estimate the motor UPDRS score of the PD patients in off-state and on-state. The performance measures (assessed by MSE, RMSE, MAE, R^2) of the model are reported in Table IV.

TABLE III. SPEARMAN CORRELATION OF ON STATE VOICE FEATURES WITH MOTOR UPDRS

Voice features /a/	r	Voice features /u/	r	Voice features /m/	r
F_2 - F_1	+0.682	F_2	+0.701	H_1	+0.547
F_2	+0.729	Shimmer apq11	+0.469	F_2	+0.726
F_{0min}	+0.360	MFCC C_{10}	-0.237	Shimmer apq11	+0.265
Fractal dimension	-0.377	MFCC C_1	+0.240	QOQ	+0.226
A_2	+0.144	MFCC C_7	+0.375	MFCC C_2	-0.155
MFCC C_{13}	-0.168	MFCC C_6	+0.354	MFCC C_3	+0.498

TABLE IV. RESULT OF AUTOMATIC ESTIMATION OF MOTOR UPDRS FROM PHONETIC TASKS USING AdaBoost FOR OFF STATE AND ON STATE PD PATIENTS

	MSE	**RMSE**	**MAE**	**R^2**
Significance	*MSE lower, better the prediction*	*RMSE lower, lower the spread of data from the fitted line*	*MAE lower, lower the error in prediction*	*R^2 higher, better the fit*
Off-state /a/	75.90	8.71	5.56	0.53
Off-state /u/	138.50	11.77	7.88	0.14
Off-state /m/	31.52	5.61	3.52	0.80
On-state /a/	74.15	8.61	6.23	0.46
On-state /u/	96.96	9.85	6.96	0.29
On-state /m/	63.31	7.96	5.90	0.54

From Table IV it is observed that the voice features from the task /m/ had the lowest estimation error of 3.52 with 80% fit. It was also noted that the voice features from the tasks /a/ and /m/ yielded higher MAE in the automatic estimation of motor UPDRS. The results also revealed that under the effect of Levodopa medication, voice features from the task /m/ were able to represent the motor UPDRS score with the lowest MAE of 5.90 whereas, the highest MAE was achieved by the voice features of task /u/.

IV. DISCUSSION AND CONCLUSION

In this study, we have performed the automatic estimation of the motor UPDRS score using voice features extracted from three phonatory tasks in off-state and on-state of PD patients. We aimed at finding the best set of voice features from the phonatory tasks which could aid in estimating the motor UPDRS score with the lowest estimation error both in off-state and on-state of medication which could have remote monitoring and decision-making applications for PD. The features extracted from the phonation tasks corresponding off state and on states were reduced to the most relevant 6 features using the LASSO feature selection algorithm. In most of the studies from literature dysphonia features extracted from phonation task /a/was for the automatic assessment of total and motor UPDRS scores [15, 18]. Very few studies had used other tasks such as diadochokinetic rate (DDK), phonation /e/, /i/, /o/, /u/, reading sentences for the automatic estimation of UPDRS score [16]. We investigated phonation task /a/ together with phonation task /u/ and /m/ independently for the evaluation of motor UPDRS.

We implemented the AdaBoost regression for the estimation of the motor UPDRS score. The best estimate of motor UPDRS was provided in the off state by the AdaBoost learner with MAE 3.52 (80% fit) using the /m/ phonation task. One limitation of the phonation task /a/ the literature mentions is its inability to provide more information about the articulators [30]. Moreover, there is evidence from the literature that the production of consonants, especially the voiced nasal consonants requires the co-ordination of the following: timing of breathing, larynx, articulators and contact force between lips, tongue, and palate [31]. Thus, it can be assumed that the nasal consonant /m/ could convey more information on the vocal mechanism compared to the phonation /a/ and this could be one of the possible reasons of obtaining lowest estimation error (MAE = 3.52) in off state by using features from the phonation task /m/ than /a/ which obtained a MAE of 5.56. Estimating the motor UPDRS score for On-state PD patients will facilitate the regular and remote monitoring of the progression of the disease and the patient's responsiveness to the medication. It is observed from Table IV that for the estimation of motor UPDRS score in on state using the voice features extracted from task /m/ were more efficient than that from task /a/ by yielding the lowest MAE.

Finally, it can be concluded that the task /m/can reflect the nuances in vocal mechanism better than /a/ and /u/ as its production is much more complex than the other phonemes reported here. Though not directly comparable with the existing literature on the automatic estimation of motor UPDRS score due to the limitation of small study size, this study produces comparatively lower MAE for the evaluation of motor UPDRS using voice features of the phonation task /m/ for both off state and on state PD groups. Another advantage of this study is that it has attempted to estimate the on state UPDRS score with a different feature set from off state by considering both the positive and negative impact of the medication on the vocal mechanism.

ACKNOWLEDGEMENT

We acknowledge the exchange fellowship for Rekha Viswanathan supported by SPARC Project Scheme, MHRD, India.

We acknowledge Jennifer K Nagao, Kit Wong from Monash Health for their support in this study.

REFERENCES

[1] M. Politis, K. Wu, S. Molloy, P. G. Bain, K. R. Chaudhuri, and P. Piccini, "Parkinson's disease symptoms: the patient's perspective," *Mov Disord,* vol. 25, no. 11, pp. 1646-1651, 2010.

[2] C. Ramaker, J. Marinus, A. M. Stiggelbout, and B. J. Van Hilten, "Systematic evaluation of rating scales for impairment and disability in Parkinson's disease," *Mov Disord,* vol. 17, no. 5, pp. 867-876, 2002.

[3] C. G. Goetz *et al.*, "Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results," *Mov Disord,* vol. 23, no. 15, pp. 2129-2170, 2008.

[4] B. Kostek, K. Kaszuba, P. Zwan, P. Robowski, and J. Slawek, "Automatic assessment of the motor state of the Parkinson's disease patient--a case study," *Diagnostic Pathology,* vol. 7, no. 1, p. 18, 2012.

[5] N. Piro, L. Piro, J. Kassubek, and R. Blechschmidt-Trapp, "Analysis and visualization of 3D motion data for UPDRS rating of patients with Parkinson's disease," *Sensors,* vol. 16, no. 6, p. 930, 2016.

[6] M. Giuberti *et al.*, "Automatic UPDRS evaluation in the sit-to-stand task of Parkinsonians: Kinematic analysis and comparative outlook on the leg agility task," *IEEE J. Biomed. Heal Informatics* vol. 19, no. 3, pp. 803-814, 2015.

[7] J. R. Duffy, *Motor speech disorders: Substrates, differential diagnosis, and management*. Elsevier Health Sciences, 2013.

[8] F. L. Darley, A. E. Aronson, and J. R. Brown, *Motor speech disorders*. Saunders, 1975.

[9] L. Brabenec, J. Mekyska, Z. Galaz, and I. Rektorova, "Speech disorders in Parkinson's disease: early diagnostics and effects of medication and brain stimulation," *Journal of Neural Transmission,* vol. 124, no. 3, pp. 303-334, 2017.

[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 and cognition,* vol. 56, no. 1, pp. 24-29, 2004.

[11] L. A. Ramig, R. C. Scherer, I. R. Titze, and S. P. Ringel, "Acoustic analysis of voices of patients with neurologic disease: rationale and preliminary data," *Ann Otol Rhinol Laryngol,* vol. 97, no. 2 Pt 1, pp. 164-72, Mar-Apr 1988.

[12] R. J. Holmes, J. M. Oates, D. J. Phyland, and A. J. Hughes, "Voice characteristics in the progression of Parkinson's disease," *International Journal of Language & Communication Disorders,* vol. 35, no. 3, pp. 407-418, 2000.

[13] S. Skodda, H. Rinsche, and U. Schlegel, "Progression of dysprosody in Parkinson's disease over time—a longitudinal study," *Mov Disord,* vol. 24, no. 5, pp. 716-722, 2009.

[14] F. Majdinasab, S. Karkheiran, M. Soltani, N. Moradi, and G. Shahidi, "Relationship between voice and motor disabilities of Parkinson's disease," *Journal of Voice,* vol. 30, no. 6, pp. 768. e17-768. e22, 2016.

[15] A. Tsanas, M. A. Little, P. E. McSharry, and 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* vol. 8, no. 59, pp. 842-855, 2011.

[16] A. Bayestehtashk, M. Asgari, I. Shafran, and J. McNames, "Fully Automated Assessment of the Severity of Parkinson's Disease from Speech," *Comput Speech Lang,* vol. 29, no. 1, pp. 172-185, Jan 2015.

[17] A. Tsanas, "Accurate telemonitoring of Parkinson's disease symptom severity using nonlinear speech signal processing and statistical machine learning," DPhil, University of Oxford, 2012.

[18] T. Peterek, P. Dohnálek, P. Gajdoš, and M. Šmondrk, "Performance evaluation of Random Forest regression model in tracking Parkinson's disease progress," in *13th International Conference on Hybrid Intelligent Systems (HIS 2013)*, 2013, pp. 83-87.

[19] Ö. Eskidere, F. Ertaş, and C. Hanilçi, "A comparison of regression methods for remote tracking of Parkinson's disease progression," *Expert Systems with Applications,* vol. 39, no. 5, pp. 5523-5528, 2012.

[20] R. Viswanathan *et al.*, "Efficiency of Voice Features Based on Consonant for Detection of Parkinson's Disease," in *2018 IEEE Life Sciences Conference (LSC)*, 2018, pp. 49-52: IEEE.

[21] H. Im, S. Adams, A. Abeyesekera, M. Pieterman, G. Gilmore, and M. Jog, "Effect of Levodopa on Speech Dysfluency in Parkinson's Disease," *Movement disorders clinical practice,* vol. 6, no. 2, pp. 150-154, 2019.

[22] A. K. Ho, J. L. Bradshaw, and R. Iansek, "For better or worse: The effect of levodopa on speech in Parkinson's disease," *Mov Disord,* vol. 23, no. 4, pp. 574-580, 2008.

[23] D. W. Paul Boersma (22 january). *Praat: doing phonetics by computer*. Available: http://www.praat.org./

[24] *VoiceSauce*. Available: http://www.phonetics.ucla.edu/voicesauce/documentation/contents.html

[25] M. Airas, "TKK Aparat: An environment for voice inverse filtering and parameterization," *Logopedics Phoniatrics Vocology,* vol. 33, no. 1, pp. 49-64, 2008.

[26] D. Michaelis, T. Gramss, and H. W. Strube, "Glottal-to-noise excitation ratio–a new measure for describing pathological voices," *Acustica / acta acustica,* vol. 83, no. 4, pp. 700-706, 1997.

[27] M. A. Little, "Biomechanically informed nonlinear speech signal processing," DPhil, University of Oxford, 2007.

[28] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.,* vol. 55, no. 1, pp. 119-139, 1997.

[29] C. Mawdsley and C. Gamsu, "Periodicity of speech in Parkinsonism," *Nature,* vol. 231, no. 5301, p. 315, 1971.

[30] M. Nespor, M. Peña, and J. Mehler, "On the different roles of vowels and consonants in speech processing and language acquisition," *Lingue e linguaggio,* vol. 2, no. 2, pp. 203-230, 2003.

[31] K. M. Kurowski, S. E. Blumstein, C. L. Palumbo, R. S. Waldstein, and M. W. Burton, "Nasal consonant production in Broca's and Wernicke's aphasics: speech deficits and neuroanatomical correlates," *Brain Lang,* vol. 100, no. 3, pp. 262-75, Mar 2007.