ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/332754244

Normalized Mutual Information of phonetic sound to distinguish the speech of Parkinson's disease

Conference Paper · April 2019

DOI: 10.1109/EMBC.2019.8857112

CITATIONS
0

READS
83

6 authors, including:

Dinesh Kumar
RMIT University
368 PUBLICATIONS **2,964** CITATIONS
SEE PROFILE

Sridhar P Arjunan
SRM Institute of Science and Technology
126 PUBLICATIONS **819** CITATIONS
SEE PROFILE

Rekha Viswanathan
RMIT University
5 PUBLICATIONS **1** CITATION
SEE PROFILE

Beth Jelfs
RMIT University
44 PUBLICATIONS **367** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Estimating Muscle Conduction Velocity from High Density Surface Electromyography View project

Human movement View project

All content following this page was uploaded by Rekha Viswanathan on 30 April 2019.

The user has requested enhancement of the downloaded file.

Normalized Mutual Information of phonetic sound to distinguish the speech of Parkinson's disease

R. Viswanathan, A. Bingham, S. Raghav, Sridhar P. Arjunan, B. Jelfs, P. Kempster, Dinesh K. Kumar

***Abstract*— This study has investigated the use of inter-personnel mutual information computed from the phonetic sound recordings to differentiate between Parkinson's disease (PD) and control subjects. The normalized mutual information (NMI) denotes the amount of information shared between the voice recordings of people within the same group: PD and Control. The hypothesis of this study was that within group NMI will be significantly different when compared with inter- group NMI. For each phonetic sound, the NMI was computed for every pairing of recordings for both the PD and control groups. Pearson correlation coefficient analysis was used to determine the association of NMI with clinical parameters including Unified Parkinson's Disease Rating Scale (UPDRS), Montreal cognitive assessment (MoCA) and disease duration. ANOVA test for the three phonetic sounds of control and PD subjects showed that there is significant difference between the intra-group mean NMI for the two groups ($p < 0.003$) and also showed significant association with the UPDRS motor examination score, MoCA and disease duration.**

Keywords: Parkinson's Disease, Normalized mutual information, Sustained phonemes, Speech

I. Introduction

Parkinson's Disease (PD) is a progressive degenerative neurological disorder with fluctuating clinical symptoms [1]. Typical motor symptoms include tremors, bradykinesia, and postural instability and hypokinetic dysarthria. PD also extends to non-motor symptoms including alterations in mood, behaviour, and [2]. The current methods used for diagnosing a patient with PD are subjective and can easily result in a misdiagnosis [1, 3].

Speech impairment is seen in more than 85% of the PD population and is largely due to the limitation in movements of the speech musculature [4]. Studies have also reported that speech could be one of the earliest indicators of PD [3, 5]. Hence, speech has and still is being investigated for it's potential to objectively diagnose PD.

The assessment of speech impairment in PD is currently performed using two methods. The first being the perception evaluation method which measures the voice quality, pitch variation, nasality, articulatory precision, and speech rhythm [6]. One of the major disadvantages of the perception method is that even a well-trained speech analyst could easily miss speech markers associated with PD [7]. The second method to evaluate PD related speech impairment is based on the computerized evaluation of speech features related to dysarthria. Conventionally acoustic features including jitter, shimmer, harmonics to noise ratio, and variability in fundamental frequency were used to evaluate dysarthria [8]. In recent years, non-linear time series analysis is gaining popularity in the evaluation of PD speech impairment [9, 10]. Non-linear speech feature analysis facilitates the study of the aperiodicity in the speech. The nonlinear speech features included in the recent studies are correlation dimension (D_2), recurrence period entropy (RPDE), pitch period entropy (PPE) [8, 11]. Non-linear features based on information theory have also been proposed to study the aperiodic nature of the speech [12].

The objective of this study is to evaluate a phonetic sound feature based on the mutual information. The mutual information has previously been applied to speech signals as a measure of speech intelligibility and therefore could provide a valuable indicator of the presence and/or progression of PD [13, 14]. To distinguish the changes in the speech due to the Parkinson's Disease, normalized mutual information (NMI) is computed from three sustained phonemes based on the amount of information shared between the sound recordings within the group (PD or control).

II. Methodology

A. Participants

The study included 24 participants with PD and 22 age matched controls. All the participants recruited for the study were between the ages of 53-85 years. The participants were evaluated by a clinical nurse, prior to the speech recording, using the Unified Parkinson's Disease Rating Scale (UPDRS) motor examination and the Montreal Cognitive Assessment (MoCA), which is used to identify mild cognitive dysfunction. PD subjects who were confined to wheelchair, had a MoCA score < 20, or were unable to read and write English were excluded from the study.

Table I shows the clinical information for both the PD and the control subjects. The UPDRS motor examination for the control subjects shows that the population exhibits some age-related motor slowness.

TABLE I. PARTICIPANTS CLINICAL INFORMATION

	PD subjects	**Control subjects**
Number of subjects	24	22
Age	71.83 ± 7.67	66.91±6.22
UPDRS Motor assessment	27.58 ± 2.58	2.64±3.65
MoCA	27.58 ± 2.48	28.45±1.37
Duration of disease	5.63 ± 3.00	-
Range of speech score in UPDRS	0-1	0

R. Viswanathan, A. Bingham, S. Raghav, B Jelfs are with Biosignal lab, School of Engineering, RMIT University, Melbourne, AUSTRALIA (e-mail: s3427376@student.rmit.edu.au); S.P.Arjunan is with SRM Institute of Science and Technology, Chennai, India (email: sridhar_arjunan@ieee.org); P. Kempster is with Department of Neurosciences, Monash Medical Centre, Clayton, VIC, Australia

B. Speech data

The speech data used for the study contained three sustained phonetic sounds /a/ (as in car), /u/ (as in wool) and /m/ (as in mum). The phonetic sounds were recorded in a noise restricted room during morning (9.00 AM and 10.AM) for all the subjects. The sounds were recorded using an Apple iPhone 6S with wired head worn omni-directional microphone at a sampling rate of 48 kHz with a 16-bit resolution.

C. Experimental protocol

The study protocol was approved by the Human research ethics committee of Monash Health and RMIT University, Melbourne, Australia and conducted in accordance with Helsinki Declaration (revised 2004). The recordings were collected after obtaining signed informed consent from the participants.

Each phonetic sound recording consisted of only one sustained utterance of the respective sound. The participants were seated comfortably on a chair for the experiment. The procedure to produce the sustained sound was demonstrated to the participants by the researcher. Before each sustained sound production, the participants were required to take a deep breath and then exhale by uttering the sound with a sustain of minimum 10 seconds. Between each phonetic sound recording a minimum of 60 seconds gap was provided to regain the normal breathing.

All PD participants were in the off-state of Levodopa medication during the phonetic sound recording. PD off-state is defined in literature as a minimum of 12 hours of medication withdrawal [15]. The PD participants were showing normal to slight loss of expression, diction, and volume in the speech evaluation section under UPDRS motor examination conducted by the clinical nurse.

D. Data analysis

In the proposed method, two features the intra-group mean and SD of the normalized mutual information (NMI) for all PD and control sound recordings for three phonemes were evaluated. The mean and SD of NMI is evaluated for each phonetic signal by determining the NMI shared by the signal with all other signals in the same group and thus we term the two features as intra-group NMI features. The sound recordings were pre-processed by performing high pass filtering with a low cut-off frequency of 70 Hz and then trimming them to a duration of 2 seconds.

NMI was computed based on the probability distribution of the signals and histogram method was used to determine the probability distribution. The phonetic signals were first windowed using a window size of 480 samples. The probability distribution of each window was obtained from the histogram. The histogram had equal bin partition and the number of bins was determined using Rice Rule [16]. The probability distribution and the joint probability distribution of the signals were then estimated for each window. From there the MI for each window was evaluated using the following equation:

$$I(X;Y) = \sum_{x\in X}\sum_{y\in Y} P_{XY}(x,y)\log_n \frac{P_{XY}(x,y)}{P_X(x)P_y(y)}, \quad (1)$$

where $I(X; Y)$ is the mutual information between a single window of signal X and signal Y, $P_x(x)$ and $P_y(y)$ are the probability distribution functions of X and Y respectively, and $P_{xy}(x,y)$ is the joint probability distribution of X and Y. The value of n in (1) determines the unit used for $I(X;Y)$, in our study $n = 2$ resulting in the units being bits.

The estimated MI is normalized by the smaller of the entropies of X and Y. The NMI evaluation performed in our study is based on [17]. The NMI is thus given by

$$NMI(X;Y) = \frac{I(X;Y)}{\min(H(X),H(Y))}, \quad (2)$$

where H(X) and H(Y) represents the entropies of X and Y which are given below.

$$H(X) = -\sum P_X(x)\, log_2 P(x), \quad (3)$$

$$H(Y) = -\sum P_Y(y)\, log_2 P(y), \quad (4)$$

The term min (H(X), H(Y)) determines the minimum entropy between H(X) and H(Y).

The proposed methodology to determine the intra-group NMI features can be summarized as follows.

1. The phonetic signals (e.g. */a/*) for the PD subjects are considered to determine the intra-group NMI features.
2. The phonetic signal of PD subject (S1) is paired with every other signal in the same group (S2, S3…, S24) to evaluate the NMI features. The signals S2, S3..,S24 represent the phonetic signals of subject 2 to subject 24 respectively.
3. The next step in determining the features is to window the signals and the windows are given by w_1, w_2, w_3…., w_m where w_m represents the m^{th} window.
4. The marginal and joint probability distributions are estimated for each window between S1 and S2. The marginal entropies for each window is then evaluated using these probability distributions.
5. The MI is evaluated for each window of S1 and S2 using Eqn (1).
6. The NMI for each window is then evaluated using Eqn (2).
7. All NMI values of S1 and S2 are then averaged resulting in a single NMI value.
8. The above steps 4-7 are repeated to find the average NMI shared between S1 and S2, S3…, S24 in the PD group.
9. The mean and SD of the NMI between S1 and S2, S3…, S24 was computed.
10. The same procedure is repeated for all recordings in the PD group to evaluate the mean and SD of the NMI between each signal and every other signal in the same group.
11. The same procedure is followed to determine the mean and SD of the NMI within the control group.

E. Statistical analysis

The difference in the intra-group mean and SD of NMI between the control group and the PD group is demonstrated through descriptive statistics. One-way ANOVA test was performed to show the statistical difference of the intra-group NMI features between PD and control subjects. To determine the correlation between the values of the intra-group mean and SD of NMI and the clinical parameters like UPDRS, MoCA, disease duration, first a normality test was performed using the Anderson-Darling test and the Pearson correlation coefficient was then computed to determine the correlation of NMI features with the clinical parameters.

III. RESULTS

The 95 % confidence interval (CI) plots of the intra-group mean and SD of NMI for three different phonetic sounds for PD and control subjects are shown in Fig. 1 and Fig. 2. It can be observed from the interval plots that there is a significant difference in intra-group mean NMI between control and PD sound recordings for all three phonemes. The intra-group mean NMI values are low for the /a/ phonetic sound recordings in both PD and control subjects. There is comparitively high intra-group mean NMI for the phonemes /u/ and /m/ in both controls and PD. It can be clearly seen from the 95% CI interval plot that there is no overlap between the intra-group mean NMI values of the control and PD subjects for all three phonemes. This indicates that there exists a significant difference in the intra-group NMI between the two groups.

For the intra-group SD of NMI feature, the /a/ and /u/ sound recordings of the PD group showes a slight reduction compared to the control group. However, intra-group SD of NMI of the /m/ sound for the PD group shows an increase compared to the controls.

The range of intra-group mean and SD of NMI for PD and control subjects for the three phonemes are shown in Table II. It could be seen that there is minor overlap between the intra-group mean SD of NMI values for PD and control subjects.

TABLE II. MEAN AND SD NMI RANGE FOR PD AND CONTROL SUBJECTS

Phonetic sound	PD subjects	Control subjects
Mean NMI (bits)		
/a/	0.212 – 0.304	0.192 – 0.276
/u/	0.293 – 0.361	0.284 – 0.358
/m/	0.291 – 0.414	0.289 – 0.384
SD NMI		
/a/	0.016 – 0.040	0.014 – 0.037
/u/	0.017 – 0.042	0.019 – 0.040
/m/	0.015 – 0.040	0.018 – 0.042

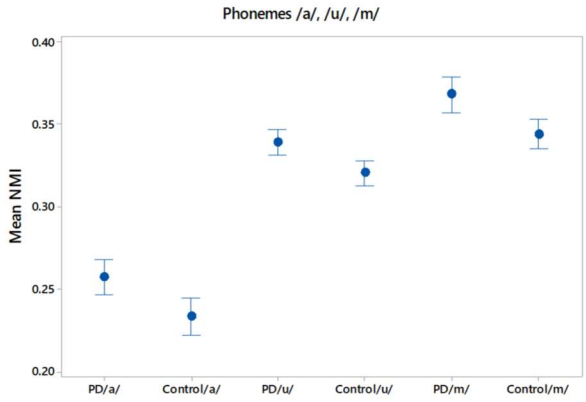

Fig. 1. 95% CI plot showing Mean NMI for /a/, /u/, /m/ sound recordings for PD and control subjects

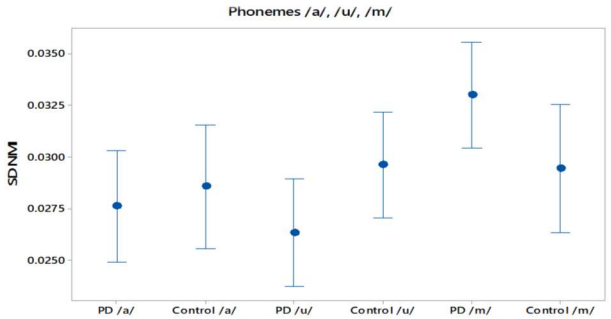

Fig. 2. 95% CI plot showing SD NMI for /a/, /u/, /m/ sound recordings for PD and control subjects

TABLE III. ANOVA RESULTS SHOWING THE DIFFERENCE IN FEATURES BETWEEN PD AND CONTROL GROUPS FOR THREE PHONETIC SOUNDS

Feature	F-value	p value
Mean NMI /a/	10.18	0.003
Mean NMI /u/	12.39	0.001
Mean NMI /m/	12.38	0.001

The statistical difference of the NMI between the PD and control phonetic sounds were evaluated using a one-way ANOVA with the results as shown in Table III. It can be seen from Table III that only the mean NMI feature shows significant difference between the PD and control subjects for all phonemes. The NMI SD feature did not show any difference between the PD and control subjects.

The correlation of the mean and SD of intra-group NMI with the clinical parameters were tested using the Pearson correlation coefficient test and are shown in Table IV. NMI SD did not show correlation with any clinical parameters. Table IV indicates that there is significant correlation between the mean NMI values of all phonetic recordings and UPDRS motor examination scores. The only feature exhibiting significant correlation with MoCA is the sound /u/. It can also be seen from the correlation analysis that the disease duration also exhibits significant correlation with all the three phonetic sound recordings.

TABLE IV. CORRELATION OF MEAN NMI AND NMI SD WITH CLINICAL PARAMETERS

Feature	MoCA	UPDRS	Duration of disease
Mean NMI /a/	-	0.415 (0.004)	0.439 (0.002)
Mean NMI /u/	-0.482 (<0.001) [a]	0.454 (0.002)	0.485 (0.001)
Mean NMI /m/	-	0.488 (0.001)	0.417 (0.004)

[a] Values expressed as correlation coefficient (p value)

IV. DISCUSSION

This study has investigated the difference in the amount of information shared within control and PD participants for sustained vowel sounds. This method is based on the shared information between each sound recording in the group with all other recordings. Features including intra-group mean and SD of the shared information were evaluated. The shared information was evaluated within control and PD groups for three phonetic sounds. It was seen that the intra-group range of the NMI values for the phonetic sounds /u/ and /m/ were similar. The intra-group range of the phonetic sound /a/ presented lower values of NMI for both the groups.

The CI plots indicate that there is a clear separation between the intra-group NMI of PD and control participants for all three phonemes. This was also verified using the ANOVA test which showed significant differences between PD and controls for all the three phonetic sounds with $p < 0.003$. Our results indicate that the features based on intra-group NMI could be used to differentiate PD and control subjects with voice as the underlying factor.

Moreover, the correlation analysis also showed that the mean NMI feature for all three phonetic sounds were significantly correlated with the clinical parameters. The mean NMI of /u/ sound showed significant negative correlation with the MoCA score. All three phonemes showed significant positive correlation with the UPDRS motor examination score and with the duration of the disease with $p < 0.004$. This association indicates that the mean NMI can be used as a feature in differentiating PD and control voices.

An important outcome of this study is that all the PD subjects displayed normal to slight loss of expression, volume and diction in their speech based on the UPDRS motor examination. The overlapping range of intra-group NMI features between control and PD subjects might be due to the lack of significant speech impairment in the PD participants.

V. CONCLUSION

This study has investigated the difference in the amount of information shared within the voices of control and PD subjects with respect to three phonetic sounds. The intra-group features mean and SD of NMI represented the shared information in the study. It has been demonstrated that there is a significant difference in the information shared within-group control and PD groups for three phonetic sounds. The intra-group mean NMI showed significant correlation with clinical parameters associated to PD thus demonstrating the feature's ability to differentiate PD and control voices. This proposed methodology based on the NMI between the intra-group sound recordings can be extended to study the changes in PD voices with varying speech impairment and their association with disease progression.

ACKNOWLEDGEMENT

We would like to thank Dr Jennifer Nagao and Ms Kit Wong from Monash Health, Melbourne, Australia for their constant support during data collection and participant recruitment

REFERENCES

[1] A. Michell, S. Lewis, T. Foltynie, and R. Barker, "Biomarkers and Parkinson's disease," *Brain,* vol. 127, no. 8, pp. 1693-1705, 2004.

[2] J. Rusz, R. Cmejla, H. Ruzickova, and E. Ruzicka, "Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease," *The journal of the Acoustical Society of America,* vol. 129, no. 1, pp. 350-367, 2011.

[3] 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.

[4] 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.

[5] J. A. Logemann, H. B. Fisher, B. Boshes, and E. R. Blonsky, "Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients," *J Speech Hear Disord,* vol. 43, no. 1, pp. 47-57, Feb 1978.

[6] J. E. Sussman and K. Tjaden, "Perceptual measures of speech from individuals with Parkinson's disease and multiple sclerosis: intelligibility and beyond," *J Speech Lang Hear Res,* vol. 55, no. 4, pp. 1208-19, Aug 2012.

[7] R. D. Kent, "Research on speech motor control and its disorders: a review and prospective," *J Commun Disord,* vol. 33, no. 5, pp. 391-427; quiz 428, Sep-Oct 2000.

[8] 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.

[9] V. S. Lee, X. P. Zhou, D. A. Rahn, 3rd, E. Q. Wang, and J. J. Jiang, "Perturbation and nonlinear dynamic analysis of acoustic phonatory signal in Parkinsonian patients receiving deep brain stimulation," *J Commun Disord,* vol. 41, no. 6, pp. 485-500, Nov-Dec 2008.

[10] D. A. Rahn, 3rd, M. Chou, J. J. Jiang, and Y. Zhang, "Phonatory impairment in Parkinson's disease: evidence from nonlinear dynamic analysis and perturbation analysis," *J Voice,* vol. 21, no. 1, pp. 64-71, Jan 2007.

[11] A. Tsanas, M. A. Little, P. E. McSharry, J. Spielman, and L. O. Ramig, "Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease," *IEEE Transactions on biomedical engineering,* vol. 59, no. 5, pp. 1264-1271, 2012.

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

[13] J. Taghia, R. Martin, and R. C. Hendriks, "On mutual information as a measure of speech intelligibility," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, 2012, pp. 65-68: IEEE.

[14] J. Jensen and C. H. Taal, "Speech intelligibility prediction based on mutual information," *IEEE/ACM transactions on audio, speech, and language processing,* vol. 22, no. 2, pp. 430-440, 2014.

[15] R. Cilia *et al.*, "The modern pre-levodopa era of Parkinson's disease: insights into motor complications from sub-Saharan Africa," *Brain,* vol. 137, no. 10, pp. 2731-2742, 2014.

[16] A. S. Morris and R. Langari, *Measurement and instrumentation: theory and application*. Academic Press, 2012.

[17] A. Bingham, S. Arjunan, B. Jelfs, and D. Kumar, "Normalised Mutual Information of High-Density Surface Electromyography during Muscle Fatigue," *Entropy,* vol. 19, no. 12, p. 697, 2017.

View publication stats