IDENTIFICATION OF PARKINSON'S DISEASE USING ACOUSTIC ANALYSIS OF POEM RECITATION

Ján Mucha
Doctoral Degree Programme (1.), FEEC BUT
E-mail: muchajano@phd.feec.vutbr.cz

Supervised by: Zdeněk Smékal
E-mail: smekal@feec.vubtr.cz

Abstract: Parkinson's disease (PD) is the second most frequent neurodegenerative disorder. It is estimated that 60–90 % of PD patients suffer from speech disorder called hypokinetic dysarthria (HD). The goal of this work is to reveal influence of poem recitation on acoustic analysis of speech and propose concept of Parkinson's disease identification based on this analysis. Classification methods used in this work are Random Forests and Support Vector Machine. The best achieved accuracy of disease identification is 70.66 % with 59.25 % sensitivity for Random Forests classifier fed mainly with articulation features. These results demonstrate a high potential of research in this area.

Keywords: poem recitation, acoustic analysis, binary classification, Parkinson's disease, hypokinetic dysarthria

1 INTRODUCTION

Parkinson's disease (PD) is the second most frequent neurodegenerative disorder with a chronic progressive course in the world [1]. PD affects approximately 1.5 % of population aged over 65, whereas the risk increases with age [2]. It is estimated that 60–90 % of PD patients suffer from a distinctive motor speech disorder called hypokinetic dysarthria (HD). HD affects especially subsystems involved in the formation of speech such as phonation, articulation, prosody and speech fluency [3]. Clinical symptoms such as speech rate abnormalities, increased voice nasality, reduced voice intensity, increased acoustic noise, reduced variability of pitch and loudness, harsh breathy voice quality, rapid repetition of words or syllables, imprecise articulation, unintentional introduction of pauses and sudden deceleration or acceleration in speech rate have been observed in PD patients [4, 5].

According to previous studies, significant reduction of F_0 variability was observed in PD patients compared with healthy controls [6, 7]. Moreover, Skodda et al. [8] have analysed dysprosody in patients with the same gender and they revealed increased rate of pauses during standardised reading task. For quantification of different aspects of HD in PD a wide range of speech tasks was used in previous studies, e. g. sustained vowel phonation [9, 10], diadochokinetic tasks [11, 12], several types of reading tasks [5, 8], and running speech [13, 8]. Subsequently, disorders of HD have been analysed using conventional, clinically interpretable speech parameters describing articulation [11, 14], prosody [13, 4, 5, 8], speech fluency [13, 5, 8, 14], and quality of speech/voice [10, 9].

The aims of this work are: 1) to find the best features for analysis of poem recitation, where the prosodic and articulation speech features are expected to be the most significant; 2) to determine which classification method has higher HD identification accuracy; 3) to optimize the acoustic analysis of poem recitation in PD patients in order to get the best results of sensitivity and specificity.

2 DATASET AND METHODOLOGY

2.1 DATASET

For the purpose if this study, we used a database of poem recitation recorded in PD patients and healthy controls. The healthy participants had no history or presence of speech disorders or brain diseases, including neurological and psychiatric illnesses. All patients were examined on their regular dopaminergic medication approximately 1 hour after the L–dopa dose. All participants were enrolled at the First Department of Neurology, St. Anne's University Hospital in Brno, Czech Republic. All participants signed an informed consent form that has been approved by the local ethics committee. Database consist of 152 Czech native speakers, where 99 are PD patients (59 men with: age (66.4 ± 8.7) years, UPDRS III – Unified Parkinson's disease rating scale, part III: evaluation of motor function 26.8 ± 10.5 and LED – L–dopa equivalent daily dose (1110.4 ± 562.5) mg; 40 women with: age (68.9 ± 7.7) years, UPDRS III 22.1 ± 13.6 and LED (870.2 ± 479.3) mg) and 53 are healthy control (26 men with age (65.6 ± 8.9) years and 27 women (67.3 ± 9.4) years). Speakers first read and tried to recite a poem for themselves. Consequently, they recited the poem into a microphone. The poem recitation task is the following one (2 rhymes): Czech original – *Chcete vidět velký lov? Budu lovit v džungli slov. Osedlám si Pegasa, chytím báseň do lasa!*; English translation – *Would you like to see a big hunt? I will be hunting in a jungle of words. I will saddle the Pegasus, I will catch a poem into a lasso.*

2.2 SPEECH FEATURES

Speech features used in this study can be divided into four categories describing various aspects of speech: prosodic, articulatory, phonatory and features describing speech quality. Considering the selected speech task, it is expected that the prosodic and articulatory speech features will have the highest discrimination power. Prosodic speech features used in this study are: fundamental frequency (F_0), short-time energy (STE), Teager-Kaiser energy operator (TKEO), total speech time (TST), total pause time (TPT), and net speech time (NST). Articulatory features used in this study are: first three formant frequencies and their bandwidths, and spectral flux (SF). Phonatory features used in this study are: jitter, shimmer, and pitch period entropy (PPE). Features describing speech quality used in this study are: harmonic-to-noise ratio (HNR), noise-to-harmonic ratio (NHR), zero-crossing rate (ZCR), and median of power spectral density (MPSD).

2.3 STATISTICAL ANALYSIS AND CLASSIFICATION

Some features are extracted from segments of speech signal, so the result is a vector. For the consequent classification it's necessary to express the statistical properties of these features. To describe the statistical properties of the selected speech features the following statistics (or their estimates) were used: mean, median, standard deviation (std), range, relative standard deviation (relative std), relative range, mean absolute deviation (mad), 1st quartile, 3rd quartile, interquartile range (IQR).

Two classification methods were used: Random Forests (RF) and Support Vector Machine (SVM) with Gaussian radial basis function kernel. We employed a 10–fold cross-validation with 5 repetitions. Univariate and multivariate classification was performed. In the case of multivariate classification, we used the mRMR (minimal redundancy and maximal relevance) feature selection algorithm to identify relevant and non-redundant speech parameters [15]. The classification performance was evaluated using accuracy (ACC), sensitivity (SEN), specificity (SPE) and Matthew's correlation coefficient (MCC).

3 RESULTS

3.1 UNIVARIATE CLASSIFICATION

Results of the univariate classification are summarized in Table 1–2. Features are ordered according to MCC. The 5 best features are chosen for both classification methods. Values of ACC, SEN, SPE, MCC, Spearman correlation coefficients (SK) and related p values (pSK), and Pearson correlation coefficients (PK) and related p values (pPK) for RF are listed in Table 1 and in Table 2 for SVM.

Table 1: Overview of the best features identified in the univariate classification using **RF**.

Feature	**ACC** [%]	**SEN** [%]	**SPE** [%]	**MCC**	**PK**	***p*PK**	**SK**	***p*SK**
relative std 1st formant	66.58±0.9	47.17±2.8	76.97±0.6	0.2953	0.05	0.51	0.05	0.54
1st quartile STE	64.87±0.7	51.32±2.8	72.12±1.4	0.2586	0.15	0.07	0.11	0.17
std SF	64.74±1.8	46.42±2.9	74.55±1.8	0.2436	0.06	0.44	0.08	0.35
relative std 3rd formant	64.08±2.6	40.00±6.3	76.97±2.7	0.2175	0.01	0.94	0.01	0.90
shimmer	64.34±1.8	43.40±1.9	75.56±3.3	0.2050	0.18	0.03	0.14	0.08

Table 2: Overview of the best features identified in the univariate classification using **SVM**.

Feature	**ACC** [%]	**SEN** [%]	**SPE** [%]	**MCC**	**PK**	***p*PK**	**SK**	***p*SK**
relative std 1st formant	72.63±1.9	52.08±5.4	83.64±2.5	0.3945	0.05	0.51	0.05	0.54
std SF	72.37±0.9	37.74±2.3	90.91±1.6	0.3633	0.06	0.44	0.08	0.35
jitter	71.18±2.1	31.70±3.4	92.32±1.5	0.3459	0.09	0.26	0.05	0.57
1st quartile STE	68.42±1.5	33.21±4.5	87.23±1.8	0.2437	0.15	0.07	0.11	0.17
relative std 3rd formant	68.95±2.2	32.08±3.5	88.69±1.8	0.1978	0.01	0.94	0.01	0.90

3.2 MULTIVARIATE CLASSIFICATION

Results of the multivariate classification for both classifiers are summarized in Table 3. Features are ordered according to MCC. The table shows number of parameters in a feature set established using mRMR, and related values of ACC, SEN, SPE, and MCC.

Table 3: Multivariate classification results ordered according to MCC.

RF				
Number of features in group	**ACC** [%]	**SEN** [%]	**SPE** [%]	**MCC**
36	70.66±2.6	59.25±6.3	76.77±2.5	0.3484
30	70.00±3.5	59.25±2.9	75.76±4.7	0.3244
64	69.61±3.6	60.75±8.9	74.35±4.6	0.3141
43	68.95±2.2	56.60±6.1	75.56±4.4	0.3100
32	69.08±2.6	57.74±6.9	75.15±2.9	0.3099
SVM				
Number of features in group	**ACC** [%]	**SEN** [%]	**SPE** [%]	**MCC**
34	68.95±2.3	27.92±4.1	90.91±1.6	0.2808
36	68.95±1.8	27.92±4.7	90.91±1.0	0.2378
46	68.55±1.6	23.77±3.9	92.53±1.5	0.2256
31	68.95±2.3	29.06±2.9	90.30±2.7	0.2241
33	69.34±2.1	27.92±4.5	91.52±1.5	0.2183

4 DISCUSSION

4.1 UNIVARIATE CLASSIFICATION

The highest ACC in both classifiers provides relative std of 1st formant (see Table 1 and Table 2), which represents variability of the 1st formant. This feature describes reduced momentum of tongue caused by muscle rigidity, which is an ordinary symptom of PD. SF was identified as significant too. SF describes a rate of vocal tract position change, which is also associated with the rigidity. Values of ACC are by approximately 5 % higher in SVM, but values of SEN are approximately by 13 % lower than in RF. Therefore, RF is more successful classification method in terms of SEN. Articulatory features significantly occur, which confirms initial assumption. Absolute values of correlation coefficients are low and related *p* values point out that neither are statistically significant. These results highlight the importance of feature combination in order to improve the classification performance.

4.2 MULTIVARIATE CLASSIFICATION

In opposite to the univariate classification, the discrimination power of RF is better than the one of SVM (see Table 3). Values of SEN are in the case of SVM approximately by 30 % lower than in RF. Considering this, the best group of features was selected from groups classified by RF according to the highest ACC. The best group contains 36 features selected by mRMR, where 58 % of them are articulatory features (formants, SF) and 36 % are prosodic features (STE, TST, F_0, TKEO). This indicates their significance in analysis of poem recitation. It is obvious from these results that for the acoustic analysis of poem recitation the multivariate classification provides better discrimination power than the univariate one.

5 CONCLUSION

This work deals with the identification of PD based on the acoustic analysis of poem recitation. In the case of univariate analysis, the SVM classifier seems to be better than RF, but considering sensitivity, the RF is more suitable. Results in univariate classification point to the fact that it is convenient to combine features to improve accuracy of classification. A disadvantage of some features is their dependence on gender. This means that further division of participants (based on their gender) should improve the results.

Based on the results of multivariate classification, we can conclude that RF provides more adequate results than SVM. Accuracy of 70.66±2.6 % and sensitivity of 59.25±6.3 % was achieved by a group of features selected using mRMR. These results demonstrate a high potential of research in this area. This work has some limitations. We hypothesise that optimisation of classifiers' parameters (especially in SVM) would improve the results. Moreover, gender should be considered as a covariate. Finally, to generalize the results, a bigger and multilingual dataset must be analysed.

ACKNOWLEDGEMENT

This work was supported by the grant of the Czech Ministry of Health NV16-30805A (Effects of non-invasive brain stimulation on hypokinetic dysarthria, micrographia, and brain plasticity in patients with Parkinson's disease).

REFERENCES

[1] M. C. de Rijk, L. J. Launer, K. Berger, M. M. Breteler, J. F. Dartigues, M. Baldereschi, L. Fratiglioni, J. Lobo, A. Martinez-Lage, C. Trenkwalder, and A. Hofman, "Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts," *Neurology*, vol. 54, pp. 21–23, 2000.

[2] S. Sapir, L. Ramig, and C. Fox, "Speech and swallowing disorders in Parkinson disease," *Curr. Opin. Otolaryngol. Head Neck Surg.*, vol. 16, no. 3, pp. 205–10, 2008.

[3] L. O. Ramig, C. Fox, and S. Sapir, "Speech treatment for Parkinson's disease," *Expert Rev. Neurother.*, vol. 8, no. 2, pp. 297–309, 2008.

[4] J. Mekyska, E. Janousova, P. Gomez-Vilda, Z. Smekal, I. Rektorova, I. Eliasova, M. Kostalova, M. Mrackova, J. B. Alonso-Hernandez, M. Faundez-Zanuy, and K. L. de Ipiña, "Robust and complex approach of pathological speech signal analysis," *Neurocomputing*, vol. 167, pp. 94–111, 2015.

[5] Z. Galaz, J. Mekyska, Z. Mzourek, Z. Smekal, I. Rektorova, I. Eliasova, M. Kostalova, M. Mrackova, and D. Berankova, "Prosodic analysis of neutral, stress-modified and rhymed speech in patients with parkinson's disease," *Comput. Methods. Programs. Biomed.*, vol. 127, pp. 301 – 317, 2016.

[6] A. M. Goberman, "Correlation between acoustic speech characteristics and non-speech motor performance in Parkinson's disease," *Med. Sci. Monit.*, vol. 11, no. 3, pp. CR109–116, 2005.

[7] A. J. Flint, S. E. Black, I. Campbell-Taylor, G. F. Gailey, and C. Levinton, "Acoustic analysis in the differentiation between Parkinson's disease and major depression," *J. Psycholinguist. Res.*, vol. 21, pp. 383–399, 1992, 10.1007/BF01067922.

[8] S. Skodda, W. Visser, and U. Schlegel, "Gender-related patterns of dysprosody in Parkinson's disease and correlation between speech variables and motor symptoms," *J. Voice*, vol. 25, no. 1, pp. 76–82, 2011.

[9] J. Mekyska, Z. Galaz, Z. Mzourek, Z. Smekal, I. Rektorova, I. Eliasova, M. Kostalova, M. Mrackova, D. Berankova, M. Faundez-Zanuy, K. L. de Ipina, and J. B. Alonso-Hernandez, "Assessing progress of Parkinson's disease using acoustic analysis of phonation," *2015 4th International Work Conference on Bioinspired Intelligence (IWOBI)*, pp. 111–118, June 2015.

[10] 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," *J. R. Soc. Interface*, vol. 8, no. 59, pp. 842–855, 2010.

[11] S. Skodda, W. Visser, and U. Schlegel, "Vowel articulation in Parkinson's disease," *J. Voice*, vol. 25, no. 4, pp. 467–472, 2011.

[12] J. Rusz, R. Cmejla, T. Tykalova, H. Ruzickova, J. Klempir, V. Majerova, J. Picmausova, J. Roth, and E. Ruzicka, "Imprecise vowel articulation as a potential early marker of Parkinson's disease: effect of speaking task," *J. Acoust. Soc. Am.*, vol. 134, no. 3, pp. 2171–2181, 2013.

[13] S. Skodda, H. Rinsche, and U. Schlegel, "Progression of dysprosody in Parkinson's disease over time–a longitudinal study," *Mov. Disord.*, vol. 24, pp. 716–722, 2009.

[14] 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.

[15] A. Tsanas, M. A. Little, and P. E. McSharry, *A Methodology for the Analysis of Medical Data, Chapter 7 in Handbook of Systems and Complexity in Health*. Springler, 2013.