Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bbe

Original Research Article

Classification of speech intelligibility in Parkinson's disease

Taha Khan [a,b,*], *Jerker Westin* [b], *Mark Dougherty* [b]

[a] *School of Innovation, Design and Technology, Malardalen University, Vasteras, Sweden*
[b] *School of Technology and Business Studies, Computer Engineering, Dalarna University, Falun, Sweden*

ARTICLE INFO

Article history:
Received 7 August 2013
Received in revised form
20 September 2013
Accepted 4 October 2013
Available online 16 October 2013

Keywords:
Parkinson's disease
Speech processing
Dysarthria
Support vector machine
Tele-monitoring

ABSTRACT

A problem in the clinical assessment of running speech in Parkinson's disease (PD) is to track underlying deficits in a number of speech components including respiration, phonation, articulation and prosody, each of which disturbs the speech intelligibility. A set of 13 features, including the cepstral separation difference and Mel-frequency cepstral coefficients were computed to represent deficits in each individual speech component. These features were then used in training a support vector machine (SVM) using *n*-fold cross validation. The dataset used for method development and evaluation consisted of 240 running speech samples recorded from 60 PD patients and 20 healthy controls. These speech samples were clinically rated using the Unified Parkinson's Disease Rating Scale Motor Examination of Speech (UPDRS-S). The classification accuracy of SVM was 85% in 3 levels of UPDRS-S scale and 92% in 2 levels with the average area under the ROC (receiver operating characteristic) curves of around 91%. The strong classification ability of selected features and the SVM model supports suitability of this scheme to monitor speech symptoms in PD.

1. Introduction

Parkinson's disease (PD) is caused by the progressive deterioration of dopamine producing nerve cells in the mid-brain [1]. The dopamine serves as a messenger that allows communication between the mid-brain and other parts of the brain that are responsible for producing smooth and controlled body movements. A lack of dopamine causes a number of motor symptoms including reduced muscular movement, tremor and speech dysfunctions. These symptoms advance with the disease progression and degrade the quality of life of people with PD. Medication and surgical intervention can alleviate some of these symptoms but there is no cure available. PD treatments are optimized by following up the patients at regular intervals; this is problematic given the physical restrictions of patients and the established assessment procedures. Tele-monitoring of symptoms through internet or mobile devices have potential to complement traditional clinical practices and may relieve the workload of clinicians as well as reduce treatment cost [2]. In this aspect objective assessment algorithms are developed that record biometric signals associated with PD symptom severity and quantified on standard clinical scale such as the UPDRS (Unified

* *Corresponding author at*: School of Technology and Business Studies, Computer Engineering, Dalarna University, 79188 Falun, Sweden.
E-mail addresses: tkh@du.se (T. Khan), jwe@du.se (J. Westin), mdo@du.se (M. Dougherty).

0208-5216/$ – see front matter
http://dx.doi.org/10.1016/j.bbe.2013.10.003

Parkinson's Disease Rating Scale) [3]. Speech is particularly suitable in this regard as it is convenient to self-record without supervision and expensive equipment.

Speech disturbance is an early indicator of PD and previous investigations revealed that speech degradation and general PD symptom severity are strongly interlinked [4]. Several methods on PD speech classification are reported to have analyzed speech signals to discriminate between PD patients and healthy controls [5–8]. Traditional investigations involved voice signal analysis to estimate dysphonic symptoms manifested in sustained-vowel phonation. In the recent methods [6–8], the running speech is analyzed to demonstrate deficits in motor speech, suggesting that PD can affect all different subsystems of speech including respiration, phonation, articulation and prosody.

The speech item utility in motor UPDRS was previously examined by Zraick et al. [9]. According to them, a standard speech protocol to identify symptom severity should include reading of an unfamiliar passage containing different linguistic structures and a description to assess the reading ability. A strong inter-rater reliability coefficient was produced between symptom severity ratings, performed separately by the medical (neurologists) and non-medical (speech pathologists) experts, when a standard speech protocol was utilized in the motor examination. It was inferred that the running speech with standard formulation has potential to exploit capacious symptoms in PD speech, providing a broader perspective of evaluation.

The structural analysis of running speech is complex due to linguistic confounds and annotations at different levels of processing e.g. separating syllables, phonemes and prosodic units. Instead of processing individual speech units for symptom analysis, acoustic features such as variation in fundamental frequency, sound pressure level, speech rate, pause intervals and signal-to-noise ratio have been relied upon to identify PD speech impairment [6–8,10]. In a recent method for evaluating spastic dysarthria, the Mel-frequency cepstral coefficients (MFCC), glottal-to-noise energy and harmonic-to-noise ratio were evaluated in running speech samples [11] and indicated high correlation. Llorente et al. [12] proposed a scheme for detection of voice impairment from text-dependent running speech. They parameterized MFCCs from 140 recorded running speech samples. These MFCCs were then used for classification between117 dysarthric and 23 normal speech samples with an accuracy of 96%.

For an accurate monitoring of speech symptom status in PD, statistical mapping between the computed features and clinical ratings of speech symptom severity is an important step. A difficulty in the clinical assessment of running speech is to track underlying deficits in individual speech components which as a whole disturb the speech intelligibility. The aim of this work is to extract signal features from running-speech samples computing deficits in individual speech components, and to utilize these features for classification between speech symptom severity levels in accordance with the UPDRS-S using support vector machines (SVM) [13]. A recently introduced speech measure, cepstral separation difference (CSD) [14] has been explored in pursuit to categorize the level of speech impairment.

2. Patients and data

The data were obtained from a feasibility study of an at-home testing device [2] conducted at the University of California, San Francisco (UCSF) in collaboration with Parkinson's Institute. A total of 80 subjects (48 males and 32 females) with an average age of 63.8 years, participated in this study over a course of a year (i.e. from June 2009 to June 2010). 60 participants (40 males and 20 females) had a mean PD duration of 75.4 weeks and 20 other participants were normal controls. Speech samples were recorded during examinations of speech by a clinician. The recording equipment consisted of a microphone connected to a computer-based test-battery called QMAT. Subjects were asked to recite static paragraphs displayed on the QMAT screen in 3 standard running speech tests (RST). The paragraphs [15], ''The North Wind and the Sun'', ''The Rainbow Passage'' and ''The Grandfather Passage'' were recited by the subjects in RST type 1, 2 and 3 respectively. These paragraphs were devised in a way such that the level of textual difficulty increases from RST 1 to 3, demanding a greater stress in reading [14,16].

Each subject was rated by a clinician based on his/her reading performance in each RST using the UPDRS examination of speech. The speech examination is item 18 in UPDRS part III and is abbreviated as UPDRS-S [3]. The UPDRS-S is ranged from 0 to 4 where '0' represents normal speech, '1' represents mildly impaired speech, '2' represents moderately impaired speech, '3' represents severely impaired speech and '4' represents unintelligible speech. Out of the 80 subjects, 24 subjects were rated '0', 25 subjects were rated '1', 28 subjects were rated '2' and 3 subjects were rated '3'. The speech signals were sampled at 48 kHz with 16 bit resolution. In total, 240 speech samples (80 subjects $\times$ 3 RST types) have been utilized for classification between the symptom severities.

3. Methods

The intelligibility of speech can be disturbed by a number of PD symptoms. Pinto et al. [4] identified the relation between PD symptoms and anatomical substrates of speech components. According to them, vocal impairment in PD is associated with pathological changes to mainly three components of speech: respiration, phonation and articulation, attributed to the dysfunction of musculatures at subglottis (lungs, trachea, windpipes etc.), glottis (larynx) and supraglottis (jaw, lips, tongue, velum, pharynx etc.) respectively. The collective dysfunction in these components gives rise to the dysfunction in the fourth speech component called prosody.

In this work, several acoustic features were extracted from running speech signals to estimate symptoms in each speech component. For the sake of description, these features were organized into groups as: (1) measures relating to the phonatory symptoms, (2) measures relating to the articulatory symptoms, and (3) measures relating to the prosodic symptoms. The respiratory symptoms (e.g. reduced loudness) are manifested in speech prosody. The phonatory measures represent symptoms which emerge due to the in-coordination between phonation and respiration and cause harshness and