BBE 359 1–16
ARTICLE IN PRESS
BIOCYBERNETICS AND BIOMEDICAL ENGINEERING XXX (2018) XXX–XXX

Available online at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/bbe

Original Research Article

Parkinson disease prediction using intrinsic mode function based features from speech signal

Q1 Biswajit Karan, Sitanshu Sekhar Sahu *, Kartik Mahto

Department of Electronics and Communication Engineering, Birla Institute of Technology, Mesra, Ranchi, India

ARTICLE INFO

Article history:
Received 7 December 2018
Received in revised form
3 May 2019
Accepted 8 May 2019
Available online xxx

Keywords:
Parkinson
Q2 Parkinson disease
Empirical mode decomposition
Intrinsic mode function
IMFCC
MFCC
Support vector machine
Receiver operating characteristics

ABSTRACT

Parkinson's disease (PD) is a progressive neurological disorder prevalent in old age. Past studies have shown that speech can be used as an early marker for identification of PD. It affects a number of speech components such as phonation, speech intensity, articulation, and respiration, which alters the speech intelligibility. Speech feature extraction and classification always have been challenging tasks due to the existence of non-stationary and discontinuity in the speech signal. In this study, empirical mode decomposition (EMD) based features are demonstrated to capture the speech characteristics. A new feature, intrinsic mode function cepstral coefficient (IMFCC) is proposed to efficiently represent the characteristics of Parkinson speech. The performances of proposed features are assessed with two different datasets: dataset-1 and dataset-2 each having 20 normal and 25 Parkinson affected peoples. From the results, it is demonstrated that the proposed intrinsic mode function cepstral coefficient feature provides superior classification accuracy in both datasets. There is a significant increase of 10–20% in accuracy compared to the standard acoustic and Mel-frequency cepstral coefficient (MFCC) features.

1. Introduction

Q3 Q4 Parkinson's disease (PD) is a degenerative disorder of the central nervous system and it affects the lifestyle of millions of people worldwide. The symptoms are due to the loss of neurons that produce a chemical messenger in the brain called dopamine which when decreases, causes abnormal brain activity, leading to signs of PD [1]. Parkinson leads to several motor and non-motor symptoms which include tremor, rigidity, loss of muscle control, along with thinking and behavioral problems [1,2]. PD can be identified through speech signal. Many speech impairments are found in Parkinson affected people. These speech impairments included reduction speech intensity, fluctuation in fundamental frequency, hoarseness in voice, irregularity in speech articulation, known as hypokinetic dysarthria [1–3]. Parkinson affected people show signs of increased aperiodicity of vocal fold vibration. PD appears to affect men and women differently with respect to their vocal performance. Average fundamental frequency

* *Corresponding author at*: Department of Electronics and Communication Engineering, Birla Institute of Technology, Mesra, Ranchi, India.
E-mail addresses: bkaranetc@gmail.com (B. Karan), sssahu@bitmesra.ac.in, sitanshusekhar@gmail.com (S.S. Sahu), kartikmahto@bitmesra.ac.in (K. Mahto).

https://doi.org/10.1016/j.bbe.2019.05.005

increases in male and decreases in female patients with Parkinson compared to healthy persons [3]. In the last few decades, a lot of researches have been conducted in the diagnosis of PD using a speech signal. Majority of the studies are based on acoustic, spectral, and cepstral measures. Acoustic measures include fundamental frequency, jitter and shimmer variation, harmonic to noise ratio (HNR), noise to harmonic ratio (NHR), glottis quotient (GQ), glottal to noise excitation (GNE), vocal fold excitation ratio (VFER), recurrence period density entropy (an RPDE), detrended fluctuation analysis (DFA), pitch period entropy (PPE), intensity parameter, and formant frequency [4–19].

Spectral and cepstral feature include Mel-frequency cepstral coefficient (MFCC), linear predictive coefficient (LPC), linear predictive cepstral coefficient (LPCC), perceptual linear predictive coefficient (PLP), relative spectra coefficient in the cepstral domain (RASTA-CEPS), and relative spectra coefficient in spectral domain (RASTA-SPEC) [21,22,28]. Little et al. [4] have reported many characteristics of the speech signal based on 17 acoustic features. Tsanas et al. [5] proposed 132 dysphonia measures for classification of PD and normal people along with four feature selection algorithm. They reported 99% overall accuracy using the random forest and support vector machine. Sakar et al. [6] have collected speech samples consisting of sustained vowels, words and sentences and tested the dataset using 26 acoustic features. Pérez et al. [7] have used 27 acoustic measures based on the Hoehn and Yahr scale and found accuracy upto 85%. The optimized acoustic features based on a genetical algorithm is proposed by Shahbakhi et al. [8]. Gupta et al. [9,10] have used a crow search and cuttlefish algorithm based on acoustic features for classification of PD and normal people. They have tested the optimized acoustic features using different classifiers. Belalcazar-Bolaños et al. [11] predicted PD using four sets of noise measures based on the acoustic features. Nilashi et al. [12] proposed the progression of PD using machine learning technique. They developed a new method for prediction of total unified Parkinson disease rating scale (UPDRS) and motor UPDRS based on 16 acoustic features. Arias-Londoño et al. [13] used 11 nonlinear time series features and found the best accuracy for classification of pathological voice. Vaiciukynas et al. [14] have analyzed 18 sets of speech features of sustained vowels and short sentences in Lithuanian language for diagnosis of PD. Rusz et al. [15] performed a study of vocal impairment present in PD affected patients. Different acoustic features are utilized for classification of PD and normal. Bocklet et al. [16] have employed acoustic features, prosodic features and features related to two mass models of the vocal fold. Jointly all features have contributed 80% accuracy in PD detection. Novotný et al. [17] have analyzed 13 features related to 6 articulation aspect of speech. They reported 88% successful rate to separate PD from normal. Benba et al. [18] have used acoustic features for detection of PD. Hlavnicka et al. [19] used an acoustic feature of connected speech for prediction of PD. The connected speech proved good biomarker for PD analysis. Brückl et al. [20] have presented six measure of speech signal related to the autocorrelation of fundamental frequency contour and amplitude contour.

Some studies have been performed for PD diagnosis based on spectral and cepstral features. Orozco-Arroyave et al. [21] used several spectral and cepstral feature for PD diagnosis. The authors reported an accuracyof 79% for 5-Spanish vowels and 74–87% for words. For detection of PD using running speech in three different languages is demonstrated by Orozco-Arroyave et al. [22]. They have extracted the energy content of unvoiced sounds using the Mel-frequency cepstral coefficient (MFCC) and used 25 frequency bands according to the Bark scale. The cross-lingual experiment is also performed with good accuracy. Khan et al. [23] have introduced cepstral separation disturbs (CSD) and MFCC coefficient to monitor speech symptoms in PD. The classification accuracy using a support vector machine is reported to be 85% in 3-levels of UPDRS scale and 92% in 2-levels of UPDRS scale. In a recent study to detect PD, done by Afonso et al. [24] proposed a recurrence plot-based approach, Gómez-Vilda et al. [25] proposed speech articulation dynamics for PD detection. Karan et al. [26] proposed variational mode decomposition (VMD) for detection of PD. A speaker model based on GMM-UBM and i-vector for PD progression is demonstrated by Arias-Vergara et al. [27]. Moro-Velázquez et al. [28] have proposed a framework based on the speech recognition model to detect PD. They used MFCC, RASTA-PLP and LPC feature. Torres et al. [29] have extracted Burg's spectral features from intrinsic mode functions. They have reported that spectral features of only three intrinsic mode functions of the speech signal are sufficient to classify healthy and pathological voices. The classification accuracy of healthy and pathological voice has been reported 99.00% while in the case of real voices the percentage of correct classification is found to be 93.40%. Pereira et al. [30] presented handwritten dataset for PD classification using deep learning approach. In another study, Pereira et al. [31] have presented a review on various techniques for PD prediction.

The various types of features extracted from the speech signal in literature are presented in Table 1.

The literature indicates that there are two important aspects that have not explored in speech features. First, the features used are either related to vocal fold or vocal tract information of the speaker. There is no feature proposed which combinely characterized vocal fold and vocal tract information. Second one is that the nonlinear nature of the speech signal has not taken into consideration for feature extraction.

In this paper, a new approach is proposed based on empirical mode decomposition (EMD) for detection of PD. It employs the dyadic filter banks which effectively capture the dynamics of the speech signal. EMD decomposes a signal into intrinsic mode functions (IMFs). First four IMFs give information about vocal tract information of speaker whereas higher order IMFs gives information about vocal fold vibration. This method gives the following advantages over conventional feature extraction techniques mentioned below.

(a) In the proposed EMD method, an effort is made to extract the most appropriate feature which characterized both vocal fold and vocal tract of the speaker.
(b) It helps to capture nonlinear characteristics of the speech signal using data-adaptive filter bank nature of empirical mode decomposition.

The IMFs of the speech signal is used to extract the relevant features related to PD. A new feature intrinsic mode function

Table 1 – Various features from the speech signal used in different research work for PD prediction.

Feature measure	Name	Description	References
Acoustic measure	Jitter variants	It captures the cycle to cycle change in fundamental frequency.	[4–19]
	Shimmer variants	It quantifies the cycle to cycle change in amplitude.	
	Fundamental frequency	It is defined as the frequency of vocal fold vibration.	
	Harmonicity	It captured the noise component due to incomplete vocal fold closure.	
	Recurrence period density entropy (RPDE)	It described the amount of deviation of stable vocal fold vibration from the fundamental frequency.	
	Detrended fluctuation analysis (DFA)	It is the measure of stochastic self-similarity of the turbulent noise.	
	Pitch period entropy (PPE)	It accesses the abnormal variation of the fundamental frequency on a logarithmic scale.	
	Intensity parameter	It quantifies the power of speech signal in dB.	
	Formant frequency	It is defined as the frequencies produced by the vocal tract.	
	Glottis quotient (GQ)	It gives the opening and closing duration of the glottis.	
	Glottal to noise excitation (GNE)	It captures the amount of turbulent noise due to incomplete closure of vocal fold.	
	Vocal fold excitation ratio (VFER)	VEFR gives the amount of noise in terms of nonlinear energy and entropy value produced due to pathological vocal fold oscillation.	
Spectral and cepstral features	MFCCs	It captures the PD effects in the vocal tract.	[21,22,28]
	LPC	Linear predictive coefficient performs articulation analysis of speech signal. It gives a model of the vocal tract as a filter	
	LPCC	It is represented by Linear predictive coefficient in the cepstral domain	
	PLP	Perceptual linear prediction coefficients	
	RASTA-CEPS	Relative spectra coefficients in cepestral domain.	
	RASTA-SPEC	Relative spectra coefficients in spectral domain	
	Cepstral separation disturbs (CSD)	It is defined as the pressure wave disturbance of glottal closure	[23]
	Burg's spectral coefficient	Spectral properties of intrinsic mode function of the speech signal.	[29]
Recurrence plot		It provides the visualization of periodic nature of trajectory through a phase space	[24]
Speech articulation dynamics		It is defined as the probability distribution of the absolute kinematic velocity of the jaw–tongue system.	[25]

cepstral coefficient is proposed in this study which outperformed standard acoustic and MFCC features.

2. Database

In the present work, two different datasets are used to assess the reliability of the proposed method. During the recording of both dataset, proper acoustic environment, and standard setting of recording instrument is maintained [49].

Dataset-1: Datasets consist of 20 healthy and 25 PD subjects. Each subject has recorded 6 samples (sustained phonation of /a/ and /o/ three times).

- The PD subject's data is collected from the UCI machine learning respiratory provided by Sakar et al. [6] (https://archive.ics.uci.edu/ml/machine-learning databases/00301/). The dataset was collected in the Department of Neurology in Cerrahapasa, Faculty of Medicine, Istanbul University. The recording of the patient is done by using Trust – MC-1500 microphone. The recording instrument is set at 44.1 kHz sampling frequency and 30 dB quantization level and is placed at 10 cm distance from the subject. The PD patients' age group is between 39 and 79 years. During the collection of this dataset, 25 PD (15 males and 10 females) patients are asked to say only the sustained vowels 'a' and 'o' three times respectively which makes a total of 150 recordings.
- The recordings of 20 healthy people (10 males and 10 females) have been collected at Birla Institute of Technology, Ranchi, India. Studying the physiological history of the subjects, they are assumed to be healthy. Speech data is collected from the age group of 50–70 years. During recording, similar condition and recording instrument is followed as done by Sakar et al. [6]. The voice recording is collected with the help of Trust – MC-1500 microphone with a sampling frequency of 44.1 kHz and 30 dB quantization levels. The microphone is placed at 10 cm from the subject. The average duration of the recording is 10 s. All are asked to say sustained vowels 'a' and 'o' three times. Thus a total of 120 recordings are collected from healthy people. There is a good balance of age and gender in the database.

Dataset-2: This dataset consists of a speech recording of 20 healthy (10 males and 10 females) and 25 PD (12 males and 13 females) subject randomly chosen from PC-GITA database [32]. All the speakers belong to the Spanish language. The average

Please cite this article in press as: Karan B, et al. Parkinson disease prediction using intrinsic mode function based features from speech signal. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.05.005

age of subject with PD is from 33 to 77 years and healthy control (HC) is from 31 to 86. The database is well balanced in terms of gender and age. All the PD patients were diagnosed under the supervision of a neurological expert. The UPDRS-III scales of patients are 36.7 ± 18.7. The voice recordings are collected in ON-state. The recording is done in the noise controlled condition in a sound proof room with the help of a professional microphone (M-Audio, Fast Track C400). Speech signals are recorded at a sampling frequency of 44.1 kHz. The statistics of phonation duration both databases are summarized in Table 2.

3. Methodology

The complete process of the PD identification task is illustrated in Fig. 1. The proposed method includes a recording of the speech signal, pre-processing of recorded data, signal decomposition using EMD, feature extraction from intrinsic mode function, and classification using support vector machine (SVM) and random forest (RF). The signal was recorded at 44.1 kHz sampling rate. Hence pre-processing is done by downsampling the speech signal to 16 kHz as most of the latent features are within 8 kHz bandwidth [48].

3.1. *Feature extraction*

In various literature [1–20], it is shown that speech of Parkinson patient is affected even at the early stage of the disease. Different speech features are applied to the diagnosis of PD. Acoustic, Mel-frequency cepstral coefficients (MFCCs) and the proposed empirical mode decomposition based features are demonstrated in this study. We used Voice analysis toolbox [5] for extracting 27 acoustic features and MIR toolbox [33] for extracting 13-coefficient MFCC features. All the simulation is carried out using MATLAB-2017b environment.

3.1.1. *Acoustic feature*

This feature subset consists of many variations of speech signal such as jitter, shimmer, harmonicity (harmonic to noise ratio (HNR), noise to harmonic ratio (NHR), recurrence period density entropy (RPDE), detrended fluctuation analysis (DFA), pitch period entropy (PPE), and formant frequency. Total 27 number of features are used in this study.

3.1.2. *Mel-frequency cepstral coefficients (MFCC)*

MFCC has been used as a popular feature extraction technique for speech recognition [34] and speaker identification. It is similar to cepstral analysis and closely related to the human auditory system response because of its Mel filter bank. The computation of MFCC is shown in Fig. 2. First, Fourier transform is applied to the windowed signal to get frequency domain information. Then FFT coefficients are filtered by triangular bandpass filter bank also known as Mel scale filter [35]. The conversion from linear frequency to Mel frequency is given by Eq. (1):

$$\mathrm{Mel}(f) = 2595 \log_{10}\left(1 + \frac{f}{700}\right) \tag{1}$$

The output of Mel filter bank is converted to log energy and then passed through discrete cosine transform block to get the MFCC coefficients.

Table 2 – Statistics of phonation of duration of databases.

	PD (mean and standard deviation) in seconds	HC (mean and standard deviation) in seconds	*p*-Value using Mann Whitney *U*-test
Database-1	6.64 (5.48)	9.23 (1.36)	0.00378
Database-2	2.79 (1.311)	2.34 (1.63)	0.0122

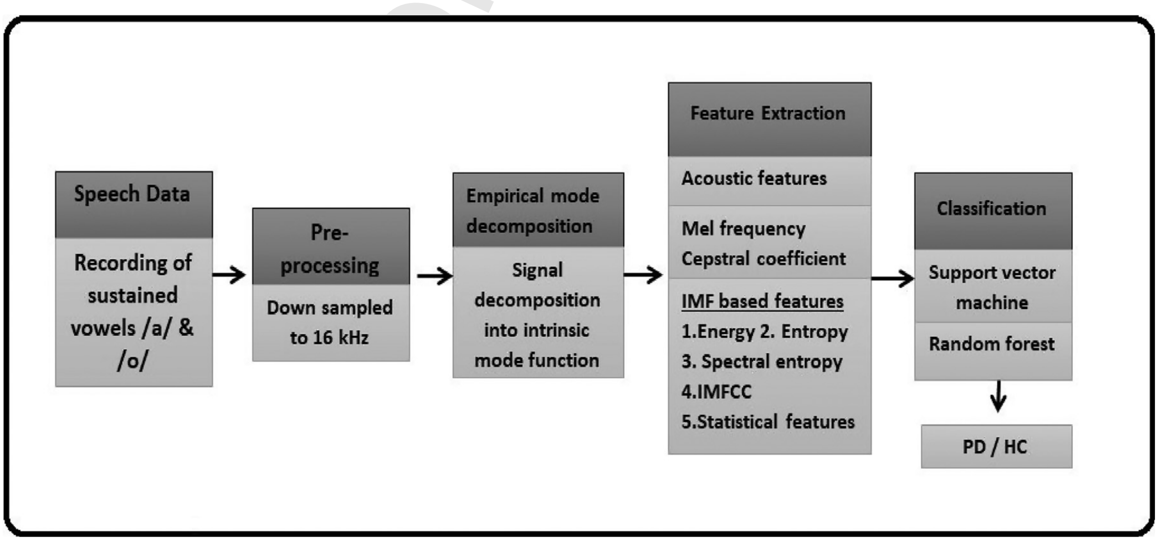

Fig. 1 – Overview of the proposed PD identification system.

Please cite this article in press as: Karan B, et al. Parkinson disease prediction using intrinsic mode function based features from speech signal. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.05.005

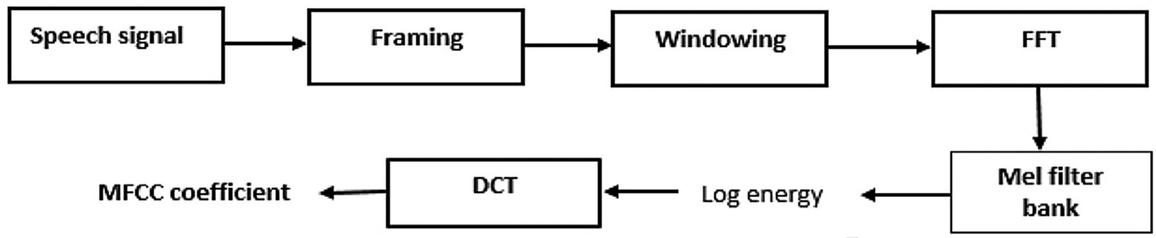

Fig. 2 – Block diagram of MFCC feature extraction

The MFCC features have some limitation as follows:

i) The short time processing of speech signal leads to loss of information at the edge.
ii) MFCC has fixed filter bank nature.
iii) The inability of capturing the dynamics of the speech signal.

3.1.3. *Empirical mode decomposition (EMD)*

To mitigate the limitation of MFCC, many efforts have been proposed by several researchers. Cosi [36] have introduced the measure of delta and double delta coefficient of MFCC to minimize the effect of short time processing. Jung et al. [37] have proposed an algorithm based on variable frame length and rate to capture the time-varying properties of the speech signal. It is not enough to describe the characteristics of the speech signal. However, wavelet decomposition somehow able to catch the nonlinearity of speech signals. Again, because of fixed filter bank nature, it produces the redundant components. In this direction, some efforts have been made by Deshpande et al. [38]. They have proposed an alternative method MDA (multiband demodulation analysis) to break the fixed filter bank structure. This effect is not adaptive to every speech signal. There is a need for data-adaptive technique which can split the speech signal into meaningful time domain components. In this context, empirical mode decomposition is found as an effective speech analysis technique.

It is a data-adaptive technique which decomposes a signal into intrinsic mode function (IMF) or modes [39]. The greatest advantage of EMD is to decompose the signal adaptably [40] because of its data dependent and dyadic filter bank nature [42]. IMFs satisfy the following properties: a number of local maxima and zero crossing are at most one and mean value of the upper and lower envelope is equal to zero. For a speech signal, the decomposed IMFs are represented by

$$s(n) = r_k(n) + \sum_{i=1}^{k} c_i(n) \tag{2}$$

where $r_k(n)$ is residue and $c_i(n)$ is the intrinsic mode function of ith mode.

The EMD algorithm is summarized in Fig. 3. It recursively decomposes the signal into several intrinsic modes which carry the latent information of speech signal.

3.1.3.1. *Analysis of intrinsic mode function*. Extensive studies have been done on IMFs to find the relationship between IMFs and the original Parkinson speech. From past studies [40,41], it is shown that IMFs carry information of vocal tract and vocal fold. Vocal tract information related to formant frequency whereas vocal fold related to pitch or fundamental frequency. To examine the fact, a synthesized speech utterance of fundamental frequency 100 Hz and formant frequency 700 Hz, 1400 Hz, 2800 Hz, and 3500 Hz is generated. This utterance is decomposed into IMFs. The first four IMFs are analyzed to see the spectrum using linear prediction (LP) analysis. Thirteen order LP analysis of speech signal is shown in Fig. 4(a). The LP analysis of the first five IMFs is shown in Fig. 4(b). It is observed that IMF-1 carried the information of all formants frequencies, IMF-2 carries the information about 1st and 2nd formant. IMFs from 3rd to 5th does not carry any information regarding formant frequency. These IMFs have resonant peaks at low frequency but these are not formant frequency. These peaks may associate with fundamental frequency or pitch.

This is also a close agreement with the work of Schlotthauer et al. [43] which reported that the pitch tracking is carried in higher IMFs like 5th, 6th, and 7th IMFs. This indicates that the IMFs capture the important signature of the speech signal which may effectively distinguish the Parkinson and normal speech.

3.1.3.2. *Feature extraction from IMFs*. As the IMFs represent peaker-specific information, the speech signal is decomposed into several IMFs. It characterizes speaker-specific information into successive IMF having both vocal tract and vocal fold information. Thus, IMFs based feature can be useful to discriminate between Parkinson and healthy peoples. From IMFs following features are extracted.

1) *Statistical features*: The Parkinson speech has a different amplitude and time variation than healthy speech. These variations can be captured by using the mean, variance, kurtosis, and skewness of IMFs obtained from the signal.
2) *Energy*: In this method energy of first 8 IMFs is calculated. Energy is normalized by dividing it with a number of samples to remove the dependency on the sample length. It is defined as

$$E_{\mathrm{imfi}} = \frac{1}{N}\sum_{i=1}^{N}(imf_i)^2 \tag{3}$$

N: number of samples of each IMFs.

Please cite this article in press as: Karan B, et al. Parkinson disease prediction using intrinsic mode function based features from speech signal. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.05.005

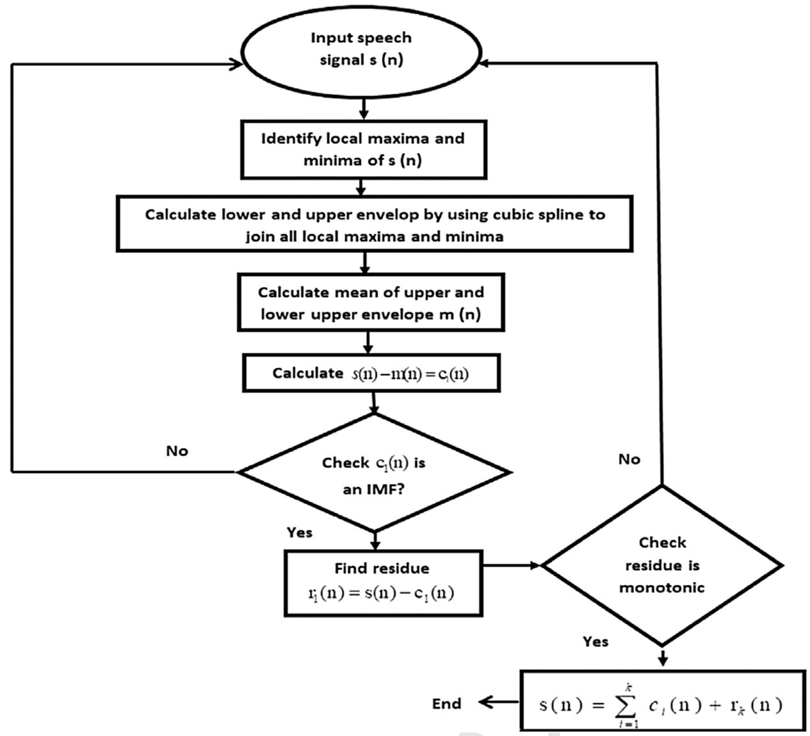

Fig. 3 – Block diagram of the empirical mode decomposition algorithm.

3) *Intrinsic mode functioncepstral coefficient (IMFCC)*: This feature is similar to the MFCC feature extraction process. First, the speech signal is decomposed into several intrinsic mode functions using EMD. Then the log energy of each IMF is computed and passed through discrete cosine transformation (DCT). The output of DCT is the intrinsic mode function cepstral coefficients (IMFCC). The entire process is shown in Fig. 5.

$$\mathrm{IMFCC}_{\mathrm{imf}} = \mathrm{DCT}(\log(E_{\mathrm{imf}})) \tag{4}$$

4) *Entropy*: It represents the information content of IMFs. The short-term entropy of energy can be interpreted as a

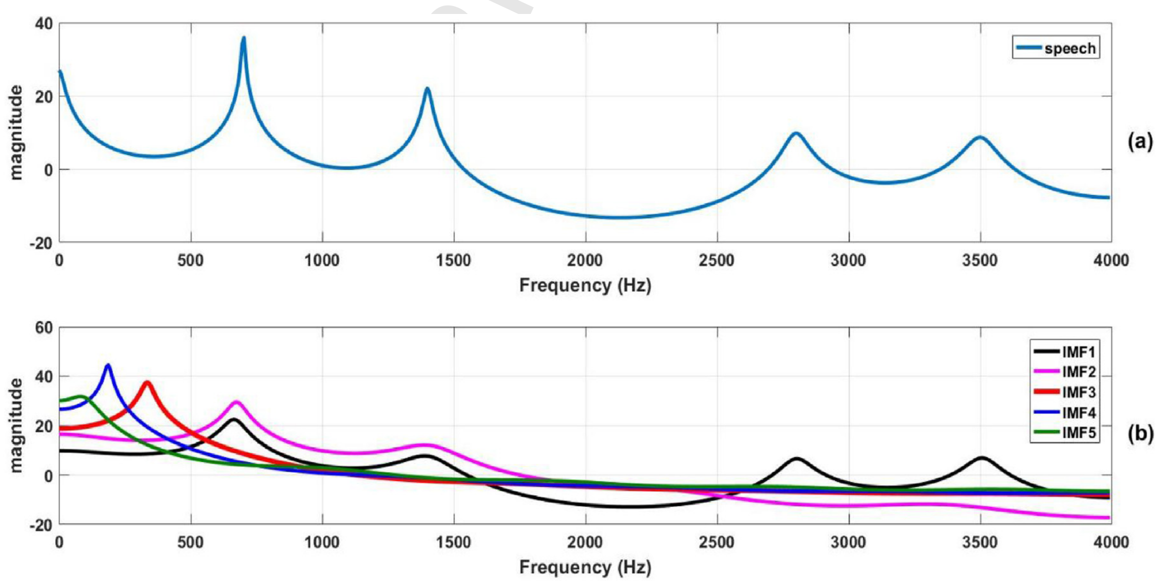

Fig. 4 – Magnitude spectrum of LP filter of (a) speech signal and (b) first five IMFs.

Please cite this article in press as: Karan B, et al. Parkinson disease prediction using intrinsic mode function based features from speech signal. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.05.005

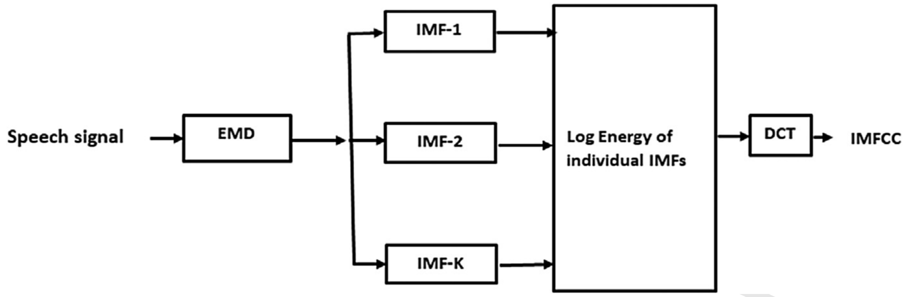

Fig. 5 – Computation of intrinsic mode function cepstral coefficient (IMFCC) feature from IMFs.

measure of abrupt changes in the energy level of a speech signal [44]. It implicitly represents the information content of the EMD filter bank resulting in the IMFs. The feature extraction process is shown in Fig. 6.It is defined by

$$H(i) = -\sum_{j=1}^{k} e_j \log_2(e_j), \quad j = 1, 2, \ldots, k \text{ and } i = \text{No. of IMFs} \tag{5}$$

where $e_j = E_{\text{subframe}_j}/E_{\text{shortframe}_i}$ and $E_{\text{shortframe}_i} = \sum_{k=1}^{K} E_{\text{subframe}_k}$.

5) *Spectral entropy*: Its computation is similar to the entropy. It is frequency domain feature. Here we divide the spectrum into L sub-bands (bins) and compute its entropy as defined by

$$H = -\sum_{F=0}^{L-1} n_f \log_2(n_f) \tag{6}$$

where

$$n_f = \frac{E_f}{\sum_{f=0}^{L-1} E_f}, \quad f = 0, 1 \ldots, L-1 \tag{7}$$

and E_f: energy of fth sub-band.

4. Classification

This section briefly explains the widely used machine learning methods such as support vector machine (SVM) and random forest (RF). The input to the classifier is a set of feature vector extracted from the speech signals. SVM and RF are good for binary classification and successfully used in many applications like pathological speech classification [45] and face recognition [46].

4.1. *Support vector machine (SVM)*

SVM is a type of machine learning technique which optimally separates two classes using a hyperplane. The training dataset is defined by $x_m \in R^N$ and $y_m \in (1, -1)$. The data vector is mapped into a feature space using the function $\varphi : x_m \rightarrow \varphi(x_m)$. The separation between the two classes is $2/\|d\|$ in feature space. The objective is to maximize the distance between two groups of classes and minimize the training error ξ_m. According to Cortes and Vapnik [47], it is done by

$$\text{Minimize}\left(\frac{d^T d}{2} + C\sum_{m=1}^{M} \xi_m\right) \tag{8}$$

Subject to:

$$y_m(d^T \varphi(x_m) + b) \geq 1 - \xi_m$$

$$\xi_m \geq 0$$

where C: trade-off between training error and margin.

These optimization problems can be solved as a dual optimization problem defined as: Maximize

$$w(\alpha) = \sum_{m=1}^{M} \alpha_i - \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \alpha_i \alpha_j y_i y_j K(x_i x_j) \tag{9}$$

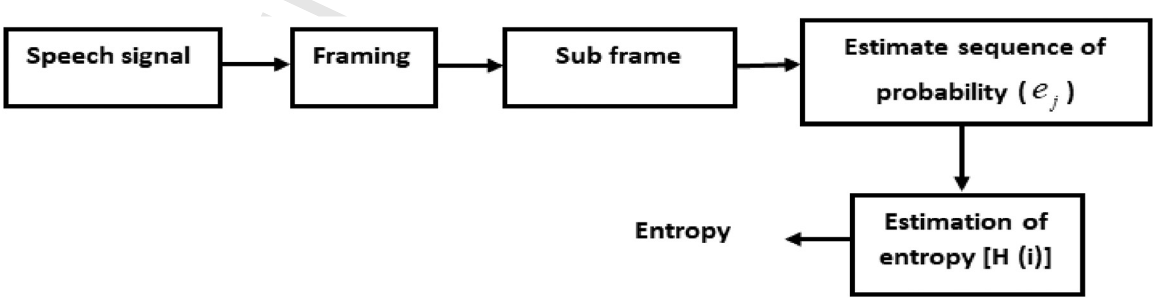

Fig. 6 – Steps for calculation of energy entropy

Please cite this article in press as: Karan B, et al. Parkinson disease prediction using intrinsic mode function based features from speech signal. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.05.005

Subject to:

$$\alpha_i \geq 0 \text{ and } \sum_{i=1}^{m} \alpha_i y_i = 0 \tag{10}$$

where α_i is the Lagrange multiplier and correspond to (x_i, y_i) and $K(x_i x_j)$ (known as kernel function).

In this work, the radial basis function (RBF) is used and the kernel function

$$K(x_i x_j) = \exp\left(-\frac{\|x_i - x_j\|}{2\sigma^2}\right) \tag{11}$$

where σ^2 is the variance of RBF function.

4.2. Random forest (RF)

Random forest is a type of machine learning algorithm which consists of an ensemble of multiple decision trees. The way it works is that many decision trees are used for predictions. Then the final prediction is the mean of all the predictions made by individual decision trees. The steps involved in this process are:

- Pick random 'm' data points from the training set with replacement.
- Use multiple decision trees to model the data with the 'm' data points.
- The new predictions are made by aggregating the results of every decision tree by assigning that category to the data point which most of the trees have predicted.

To evaluate the performance of classifier, following parameters are used:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{12}$$

$$SENS = \frac{TP}{TP + FN} \tag{13}$$

$$SPEC = \frac{TN}{TN + FP} \tag{14}$$

$$EER = FPR - FNR = \frac{FP.TP - FN.TN}{(FP + TN)(TP + FN)} \tag{15}$$

where TP: true positive; TN: true negative; FP: false positive; FN: false negative; ACC: accuracy; SENS: sensitivity; SPEC: specificity; EER: equal error rate; FPR: false positive rate; FNR: false negative rate.

5. Results and discussion

The performances of the proposed features are assessed using SVM and RF. All classifier models are evaluated using10-fold cross-validation, leave one out cross-validation (Loocv) and leave one subject out (Loso) test. The entire database is divided into 80% training data and 20% testing data. The optimized model parameter is computed from the 80% training data. Then the model is tested using the optimized parameters. The optimized parameters are obtained by the grid search algorithm by taking C values (10^{-1} to 10^3) and γ values (0.1–1) for SVM classifier. For random forest classifier, the optimized parameters are computed using grid search algorithm by taking two parameters, *n*_estimator (number of trees in the forest) and *n*_features (max number of features considered for splitting a node). The values of *n*_estimator are 200, 300, 400, and 500 and the range of *n*_features is in between 1 and a maximum number of features. The remaining 20% of the data is used as an independent test set for testing the model. The classification task is performed in python-3.6 environment. The time consumption to build the model is around 40–50 s.

5.1. Comparative study of the feature pattern

A comparative study among the extracted IMF based features is carried out to visualize their distinguishing behavior of the PD and healthy subject.

Figs. 7–10 show the IMFs and its extracted feature variation of healthy and Parkinson peoples. Feature patterns are obtained using 300 speech samples of Parkinson and healthy people. Almost all speech signals followed the same pattern. From Fig. 7 we can clearly distinguish the variation in the pattern of healthy and Parkinson affected speech. It is noticed that the average energy value of normal speech IMFs is more than Parkinson affected speech. The same kind of variation is observed in the case of IMFCC. The entropy of each IMF for healthy people showing constant variation whereas for PD people it shows a noticeable variation. From the pattern of spectral entropy shown in Fig. 10 noticeable variation is observed between healthy and Parkinson affected speech signal. Thus proposed IMF based features energy, entropy, spectral entropy, IMFCC, and the statistical feature can be a good marker for prediction of PD.

5.2. Statistical analysis

Various statistical tests have been performed to test the effectiveness of feature for classification. The Spearman correlation result of proposed feature IMFCC and phonation duration is summarized in Table 3. The larger the magnitude of correlation coefficient and the lesser the value *p*, the stronger is the statistical relationship between the variables. From Table 3, it is noted that the proposed feature has more contribution toward classification.

5.2.1. Features importance

To measure the importance of each feature in classification, the *p*-value of each feature is evaluated using the *t*-test. 1/*p*-value gives the relative importance of each feature. The *p*-value of EMD based features, MFCC, and acoustic features are mentioned in Table 4. The lower *p*-value reveals that the two groups PD and HC are more statistically different. Fig. 11 gives the plot of relative heights of the importance of each feature. The IMF based features have more significance than MFCC and acoustic features. Among the IMF based features, the IMFCC shows more significance for classification.

Please cite this article in press as: Karan B, et al. Parkinson disease prediction using intrinsic mode function based features from speech signal. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.05.005

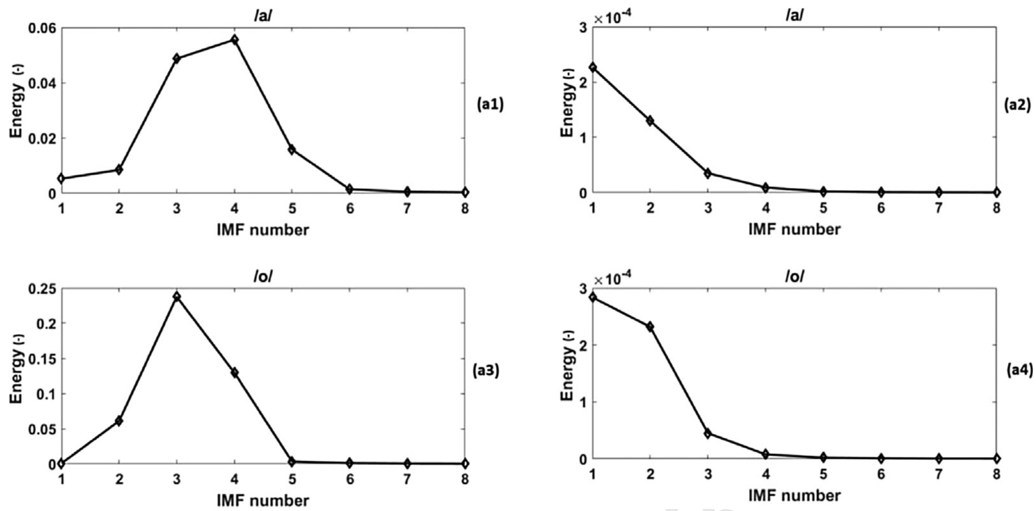

Fig. 7 – Energy variation of IMFs: (a1) normal speech of vowel /a/; (a2) Parkinson speech of vowel /a/; (a3) Normal speech of vowel /o/; (a4) Parkinson speech of vowel /o/.

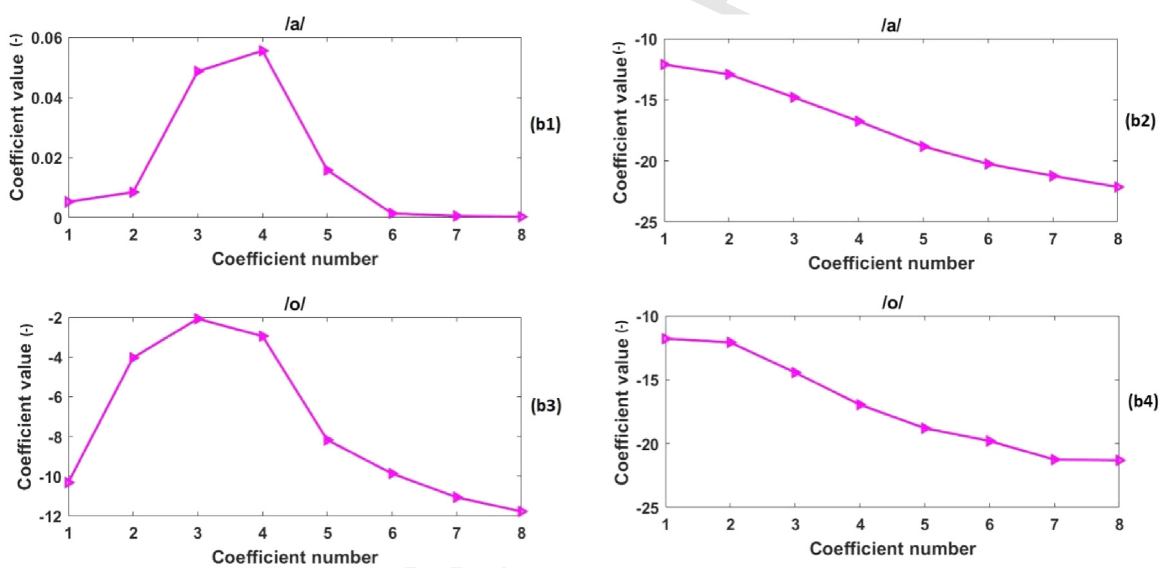

Fig. 8 – IMFCC variation of (b1) normal speech of vowel /a/; (b2) Parkinson speech of vowel /a/; (b3) normal speech of vowel /o/; (b4) Parkinson speech of vowel /o/.

From Table 4, it is shown that the *p*-value of EMD based features is less than that of MFCC and acoustic features. The same is shown in Fig. 11 that proposed EMD based features has more relative feature importance compare to traditional feature MFCC. From this analysis, we can say that EMD based features could provide more contribution to classification accuracy.

5.3. *Experimental result*

Experiment 1: Various IMFs are selected to compute the proposed features energy, energy entropy, spectral entropy, IMFCC and statistical features. Table 5 shows the classification accuracy obtained from a different number of IMFs used.

In Table 5, it is highlighted that set of IMFs (1–4, 1–5, and 1–6) gives the highest accuracy. This reveals that lower IMFs have more contribution to relevant information than higher index IMFs.

From Fig. 12, it is noticed that classification accuracy is better between the IMFs 4–6. It is concluded that speaker-specific information is well distributed among IMFs 4–6. So in further successive experiments, the first six IMFs is considered (Figs. 13–14).

Experiment 2: The performance of the extracted features of dataset-1 is evaluated with sustained vowels /a/ and /o/. Table 6 shows the classification accuracy using dataset-1. It is found that among all EMD based features, IMFCC feature gives the highest classification accuracy for both the sustained

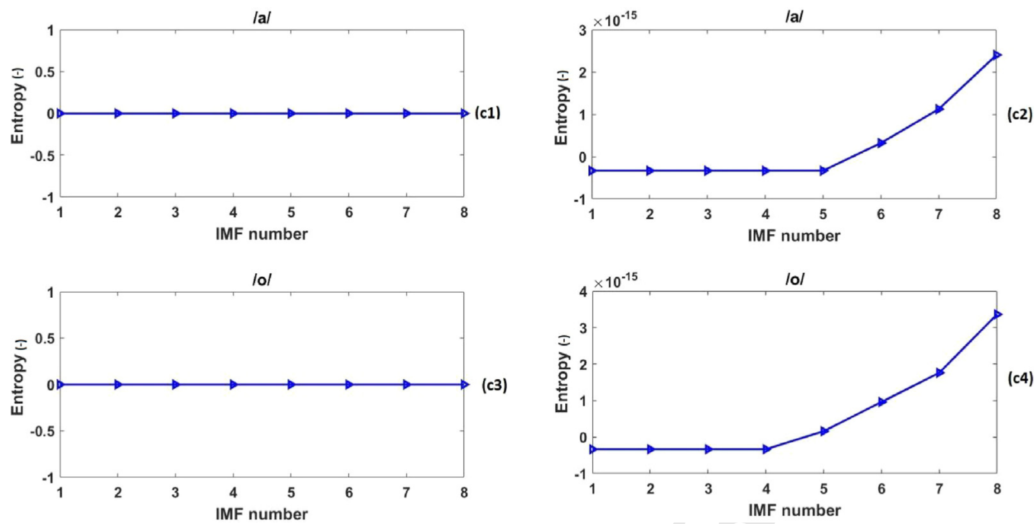

Fig. 9 – Entropy variation of IMFs (c1) normal speech of vowel /a/; (c2) Parkinson speech of vowel /a/; (c3) normal speech of vowel /o/; (c4) Parkinson speech of vowel /o/.

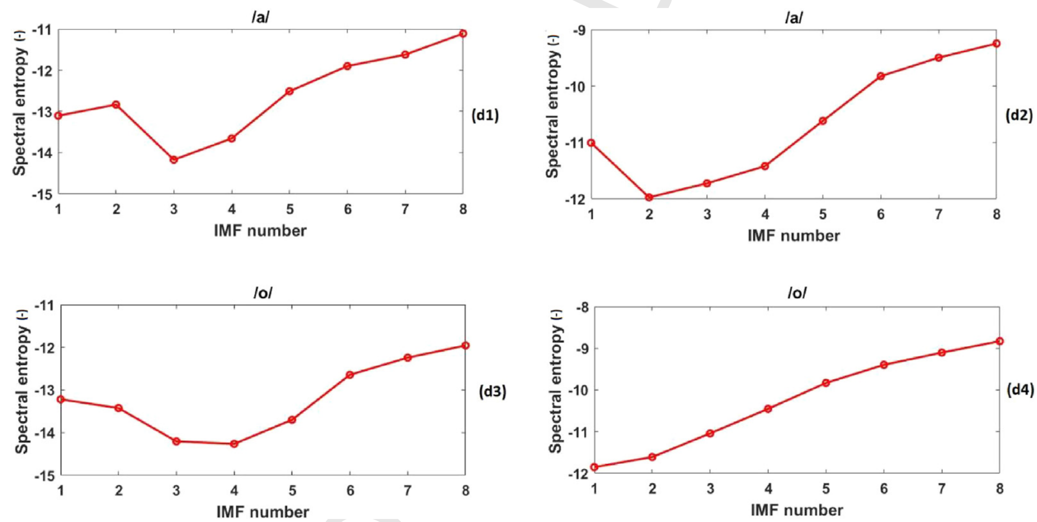

Fig. 10 – Spectral entropy variation of IMFs (d1) normal speech of vowel /a/; (d2) Parkinson speech of vowel /a/; (d3) Normal speech of vowel /o/; (d4) Parkinson speech of vowel /o/.

Table 3 – Spearman correlation analysis of proposed features and phonation duration.

Proposed features	Database-1		Database-2	
	Correlation coefficient	*p*-Value	Correlation coefficient	*p*-Value
IMFCC 1st coefficient	0.588504981	7.60×10^{-1}	0.617605379	1.47×10^{-7}
IMFCC 2nd coefficient	0.568102251	2.2×10^{-6}	0.621637768	1.15×10^{-7}
IMFCC 3rd coefficient	0.55900414	3.4×10^{-6}	0.629098589	7.31×10^{-8}
IMFCC 4th coefficient	0.539109172	8.82×10^{-6}	0.61138975	2.12×10^{-7}
IMFCC 5th coefficient	0.550799539	5.12×10^{-6}	0.609250177	2.39×10^{-7}
IMFCC 6th coefficient	0.563957931	2.7×10^{-6}	0.60834156	2.35×10^{-7}
Duration of phonation	0.14287302	0.27615	0.380141174	0.002736

Table 4 – *p*-Value of EMD based features, MFCC and acoustic feature of the signal using *t*-test.

Features	Energy	Entropy	Spectral entropy	IMFCC	MFCC-1st coefficient	MFCC-2nd coefficient	MFCC-4th coefficient
p-Value	0.0005	0.00047	0.00043	0.00028	0.0006	0.00058	0.00076

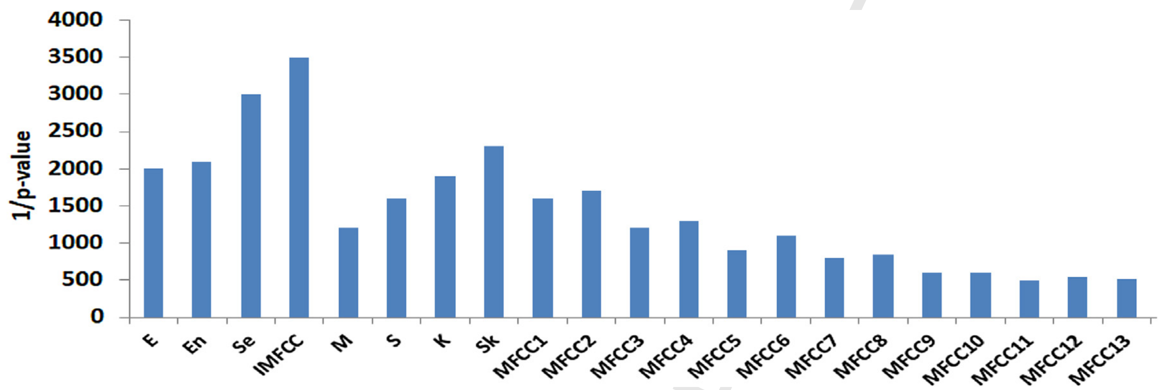

Fig. 11 – Comparison of features importance of proposed features. E: energy; En: entropy; Se: spectral entropy; IMFCC: intrinsic mode function based cepstral coefficient; M: mean; S: standard deviation; K: kurtosis; Sk: skewness; MFCC1-13: Mel frequency cepstral coefficient; F0: fundamental frequency; F1 and F2: 1st and 2nd formant frequency.

Q6 Table 5 – Variation of classification accuracy of proposed features with no. of IMFs used for feature extraction.

Number of IMFs	Energy	Entropy	Spectral entropy	IMFCC	Statistical features
1–2	88	87	92	91	81
1–3	89	82	91	88	82
1–4	**92**	**93**	**96**	**94**	**86**
1–5	**96**	**96**	**94**	**95**	**85**
1–6	**96**	**93**	**96**	**98**	**84**
1–7	93	92	94	90	85
1–8	92	92	93	89	80

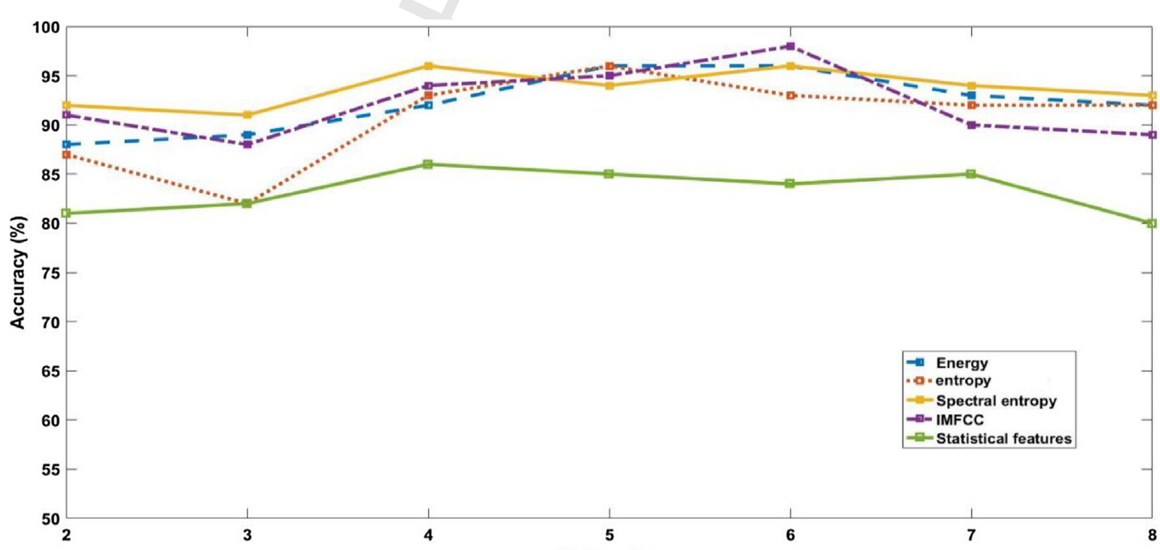

Fig. 12 – Variation of accuracy with respect to no. of IMFs used as a feature

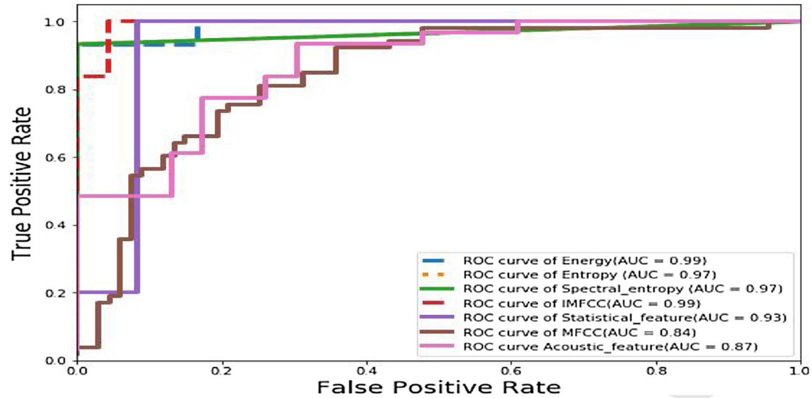

Fig. 13 – Receiver operating characteristics curve for the database-1.

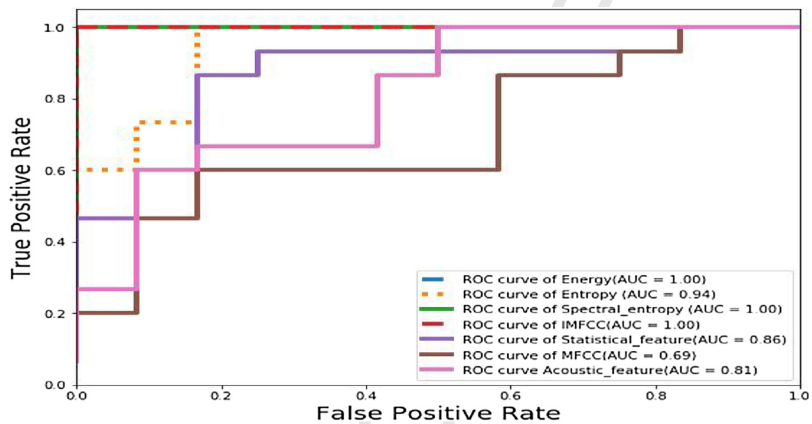

Fig. 14 – Receiver operating characteristics curve for the database-2.

Table 6 – Performance of proposed IMF based features, MFCC and acoustic features using SVM of dataset-1.

Feature		/a/					/o/				
		ACC (%)	SPEC (%)	SEN (%)	AUC	EER	ACC (%)	SPEC (%)	SEN (%)	AUC	EER
Energy	10-fold	93	93	92	0.93	0.083	85	75	93	0.98	0.06
	Loocv	92	93	92	0.93	0.06	89	93	83	0.97	0.08
	Loso	90	88	93	0.99	0.08	84	89	0.78	0.85	0.87
Entropy	10-fold	96	93	100	0.97	0.08	93	87	100	0.93	0
	Loocv	96	100	93	0.97	0.06	92	100	87	0.93	0.12
	Loso	93	93	92	0.99	0.06	89	100	80	0.99	0.08
Spectral entropy	10-fold	93	93	92	0.97	0.08	89	87	91	0.99	0.13
	Loocv	89	87	92	0.99	0.08	96	100	92	1	0
	Loso	90	88	93	0.99	0.06	89	100	80	1	0
IMFCC	10-fold	**98**	**92**	**100**	**0.99**	**0**	**97**	**96**	**98**	**0.99**	**0.02**
	Loocv	**96**	**100**	**92**	**1**	**0**	**93**	**93**	**92**	**0.98**	**0.08**
	Loso	**95**	**92**	**100**	**0.99**	**0.06**	**90**	**92**	**88**	**0.98**	**0.08**
Statistical feature	10-fold	93	100	83	0.93	0.11	85	87	83	0.94	0.12
	Loocv	93	93	92	0.89	0.08	85	75	92	0.96	0.08
	Loso	93	92	93	0.99	0.06	81	75	87	0.93	0.19
Acoustic feature	10-fold	82	92	72	0.91	0.27	81	86	75	0.91	0.19
	Loocv	81	100	67	0.92	0.08	78	58	93	0.97	0.06
	Loso	79	80	78	0.84	0.25	78	58	93	0.97	0.06
MFCC	10-fold	85	86	85	0.84	0.32	82	86	83	0.80	0.32
	Loocv	85	100	67	0.92	0.08	77	60	93	0.97	0.06
	Loso	80	76	84	0.85	0.23	78	86	68	0.83	0.22

Table 7 – Performance of proposed IMF based features, MFCC and Acoustic features using SVM of dataset-2.

Features		/a/					/o/				
		ACC (%)	SPEC (%)	SEN (%)	AUC	EER	ACC (%)	SPEC (%)	SEN (%)	AUC	EER
Energy	10-fold	85	80	92	0.95	0.08	93	93	92	1	0.06
	Loocv	88	100	83	0.94	0.16	96	100	92	1	0.08
	Loso	78	67	87	0.90	0.16	91	100	78	0.92	0.08
Entropy	10-fold	93	100	83	0.94	0.16	93	93	93	0.98	0.16
	Loocv	92	100	92	0.96	0.08	92	92	93	0.97	0.06
	Loso	90	88	94	0.98	0.06	87	80	100	0.93	0.08
Spectral entropy	10-fold	93	100	83	0.99	0.17	95	100	92	1	0.08
	Loocv	89	100	75	0.99	0.06	96	100	92	1	0.08
	Loso	87	97	73	0.87	0.26	93	98	87	0.87	0.13
IMFCC	10-fold	**93**	**92**	**92**	**0.99**	**0.08**	**96**	**100**	**92**	**1**	**0.08**
	Loocv	**85**	**83**	**87**	**0.98**	**0.13**	**96**	**100**	**92**	**1**	**0.08**
	Loso	**82**	**75**	**87**	**0.91**	**0.13**	**93**	**83**	**100**	**1**	**0**
Statistical feature	10-fold	82	75	83	0.91	0.17	85	67	83	0.86	0.17
	Loocv	81	80	83	0.81	0.17	85	93	75	0.87	0.13
	Loso	81	80	83	0.81	0.17	82	83	80	0.9	0.19
Acoustic features	10-fold	70	86	50	0.81	0.33	55	65	40	0.64	0.41
	Loocv	70	86	50	0.81	0.33	60	64	50	0.68	0.39
	Loso	66	50	76	0.68	0.28	54	64	45	0.39	0.6
MFCC	10-fold	70	75	80	0.92	0.26	70	65	86	0.91	0.19
	Loocv	68	73	78	0.69	0.22	65	72	80	0.7	0.20
	Loso	67	63	70	0.71	0.35	60	56	62	0.68	0.38

vowels /a/ and /o/. Table 7 shows the prediction accuracy of dataset-2. In this database, again the IMFCC features show superior performance to other EMD based features.

Experiment 3: The EMD based features are assessed with the classifier random forest. The results are shown in Table 8. It is observed that the EMD based feature is better than MFCC and acoustic features.

Among all features, EMD based features performed well in both the database, whereas, the new features IMFCC is superior among EMD based features.

5.3.1. *Cross-database evaluation*

The cross-database evaluation is performed to analyze the model performance at different recording conditions (includ-

Table 8 – Performance of proposed EMD based features, MFCC and acoustic features using RF classifier of dataset-1 and dataset-2.

Features		Database-1						Database-2					
		/a/			/o/			/a/			/o/		
		ACC (%)	AUC	EER	ACC (%)	AUC	EER	ACC (%)	AUC	EER	ACC (%)	AUC	EER
Energy	10-fold	92	0.93	0.06	96	1	0	81	0.91	0.16	96	0.99	0.06
	Loocv	89	0.97	0.08	96	1	0	82	0.93	0.16	96	1	0.06
	Loso	90	0.99	0.04	94	0.98	0.09	81	0.83	0.16	94	0.99	0.08
Entropy	10-fold	96	0.97	0	93	0.93	0.11	89	0.96	0.08	96	0.99	0.06
	Loocv	96	0.97	0.06	93	0.93	0.11	89	0.96	0.08	96	0.99	0.06
	Loso	93	0.94	0.11	87	0.89	0.18	89	0.95	0.12	94	0.98	0.06
Spectral entropy	10-fold	96	1	0.06	96	1	0	82	0.89	0.16	96	0.99	0.08
	Loocv	93	1	0	96	1	0	82	0.89	0.16	96	0.99	0.08
	Loso	92	0.99	0.04	89	0.98	0.08	81	0.92	0.16	95	1	0.04
IMFCC	10-fold	**100**	**1**	**0**	**93**	**0.99**	**0.06**	**100**	**1**	**0.02**	**96**	**0.99**	**0.08**
	Loocv	**96**	**1**	**0**	**89**	**0.99**	**0.06**	**100**	**1**	**0.02**	**96**	**0.99**	**0.08**
	Loso	**93**	**0.99**	**0.09**	**95**	**1**	**0**	**93**	**0.98**	**0.14**	**92**	**1**	**0.03**
Statistical features	10-fold	86	0.94	0.16	96	0.99	0.08	82	0.94	0.16	96	1	0
	Loocv	85	0.94	0.16	96	0.99	0.08	82	0.94	0.16	96	1	0
	Loso	85	0.94	0.16	94	0.99	0.06	83	0.95	0.13	90	0.99	0.06
Acoustic features	10-fold	82	0.94	0.08	93	0.98	0.08	63	0.69	0.42	67	0.74	0.35
	Loocv	82	0.94	0.08	93	0.98	0.08	56	0.67	0.46	74	0.72	0.3
	Loso	80	0.95	0.1	90	0.98	0.08	54	0.58	0.42	72	0.83	0.25
MFCC	10-fold	78	0.85	0.23	76	0.85	0.23	77	0.83	0.23	70	0.81	0.29
	Loocv	76	0.87	0.23	73	0.8	0.25	68	0.73	0.29	67	0.80	0.31
	Loso	75	0.85	0.25	72	0.84	0.22	69	0.82	0.31	65	0.78	0.31

Please cite this article in press as: Karan B, et al. Parkinson disease prediction using intrinsic mode function based features from speech signal. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.05.005

Table 9 – Cross-database performance for classification of PD and healthy in two different databases using support vector machine.

Train database	Test database	Features	ACC (%) /a/	ACC (%) /o/
Database-1	Database-2	IMFCC	54.2	56.1
		Acoustic features	58.3	55.6
		MFCC	52.1	53.2
Database-2	Database-1	IMFCC	56.4	55.3
		Acoustic features	51.6	53.8
		MFCC	54.1	50.2

ing recording instrument used, position of recording device, different room acoustics) or languages. The results are listed in Table 9.

In Table 9, from the result, it is shown that all the features, i.e. acoustic, MFCC and proposed EMD based features failed to predict PD and normal subjects. The decrease in performance may be due to different recording conditions or language difference in two dataset.

5.4. *Limitation of the study*

There are some limitations of the present study. The present study based on a small number of the PD patients. For database -1, some clinical data such as speech severity assessment, patient medication status ON or OFF condition is not available. There is difference in the acoustic condition and language between PD patient and normal people, which may affect the classification result.

6. Conclusion

In this paper, an intrinsic mode function based feature, IMFCC is proposed to efficiently detect the PD. From the empirical study, it is observed that the IMFs provide better speaker-specific information for distinguishing between PD and normal speech. It captures the non-stationary dynamics of the speech. The performance of the proposed method is evaluated with two standard datasets. The results show that the proposed feature is superior, providing highest accuracy of 100% for dataset-1 and 96% for dataset-2. There is a significant improvement of 10–20% in accuracy compared to the standard acoustic and Mel-frequency cepstral coefficient (MFCC) features. From the exhaustive study, it is found that the proposed intrinsic mode function cepstral coefficient can be used to efficiently detect the PD.

Authors's contribution

SSS conceived and supervised the study and drafted the manuscript. BK carried out the design, computational analyses, and implementation of the algorithms. KM performed the result analysis and drafting of the manuscript. All authors read and approved the final manuscript.

Conflict of interest

All the authors confirm that there is no conflict of interest to anybody.

REFERENCES

[1] Dashtipour K, Tafreshi A, Lee J, Crawley B. Speech disorders in Parkinson's disease: pathophysiology, medical management and surgical approaches. Neurodegener Dis Manag 2018;337–48. http://dx.doi.org/10.2217/nmt-2018-0021
[2] Roberts A, Post D. Information content and efficiency in the spoken discourse of individuals with Parkinson's disease. J Speech Lang Hear Res 2018;61(9 (Sep 19)):2259–60. http://dx.doi.org/10.1044/2018_JSLHR-L-17-0338
[3] Mühlhaus J, Frieg H, Bilda K, Ritterfeld U. Game-based speech rehabilitation for people with Parkinson's disease. International Conference on Universal Access in Human–Computer Interaction. Cham: Springer; 2017 Jul 9. p. 76–85.
[4] Little MA, McSharry PE, Hunter EJ, Spielman J, Ramig LO. Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. IEEE Trans Biomed Eng 2009;56 (4):015–1022. http://dx.doi.org/10.1109/TBME.2008.2005954
[5] Tsanas A, Little MA, McSharry PE, Spielman J, Ramig LO. Novel speech signal processing algorithms for high accuracy classification of Parkinsons disease. IEEE Trans Biomed Eng 2012;59:1264–70. http://dx.doi.org/10.1109/TBME.2012.2183367
[6] Sakar BE, Isenkul ME, Sakar CO, Sertbas A, Gurgen F, Delil S, et al. Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. IEEE J Biomed Heal Inform 2013;17:828–34. http://dx.doi.org/10.1109/JBHI.2013.2245674
[7] Pérez CJ, Campos-Roca Y, Naranjo L, Martín J. Diagnosis and tracking of Parkinson's disease by using automatically extracted acoustic features. J Alzheimers Dis Parkinsonism 2016;6:260. http://dx.doi.org/10.4172/2161-0460.1000260
[8] Shahbakhi M, Taheri Far D, Tahami E. Speech analysis for diagnosis of Parkinson's disease using genetic algorithm and support vector machine. J Biomed Sci Eng 2014;147–56. http://dx.doi.org/10.4236/jbise.2014.7401
[9] Gupta D, Sundaram S, Khanna A, Hassanien AE, De Albuquerque VHC. Improved diagnosis of Parkinson's disease using optimized crow search algorithm. Comput Electr Eng 2018. http://dx.doi.org/10.1016/j.compeleceng.2018.04.014
[10] Gupta D, Julka A, Jain S, Aggarwal T, Khanna A, Arunkumar N, et al. Optimized cuttlefish algorithm for diagnosis of Parkinson's disease. Cogn Syst Res 2018. http://dx.doi.org/10.1016/j.cogsys.2018.06.006
[11] Belalcazar-Bolaños EA, Orozco-Arroyave JR, Arias-Londono JD, Vargas-Bonilla JF, Noth E. Automatic detection of Parkinson's disease using noise measures of speech. Symp Signals Images Artif Vis – 2013 STSIVA 2013. http://dx.doi.org/10.1109/STSIVA.2013.6644928
[12] Nilashi M, Ibrahim O, Ahmadi H, Shahmoradi L, Farahmand M. A hybrid intelligent system for the prediction of Parkinson's disease progression using machine learning techniques. Biocybern Biomed Eng 2018;38:1–15. http://dx.doi.org/10.1016/j.bbe.2017.09.002

Please cite this article in press as: Karan B, et al. Parkinson disease prediction using intrinsic mode function based features from speech signal. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.05.005

[13] Arias-Londoño JD, Godino-Llorente JI, Sáenz-Lechón N, Osma-Ruiz V, Castellanos-Domínguez G. Automatic detection of pathological voices using complexity measures, noise parameters, and mel-cepstral coefficients. IEEE Trans Biomed Eng 2011;58(2):370–9. http://dx.doi.org/10.1109/TBME.2010.2089052

[14] Vaiciukynas E, Verikas A, Gelzinis A, Bacauskiene M. Detecting Parkinson's disease from sustained phonation and speech signals. PloS One 2017. http://dx.doi.org/10.1371/journal.pone.0185613

[15] Rusz J, Cmejla R, Ruzickova H, Ruzicka E. Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's 2011. http://dx.doi.org/10.1121/1.3514381

[16] Bocklet T, Nöth E, Stemmer G, Ruzickova H, Rusz J. Detection of persons with Parkinson's disease by acoustic, vocal, and prosodic analysis. 2011 IEEE Workshop on Automatic Speech Recognition & Understanding. IEEE; 2011. p. 478–83. doi:0.1109/ASRU.2011.6163978.

[17] Novotný M, Rusz J, C?mejla R, Růžicka E. Automatic evaluation of articulatory disorders in Parkinson's disease. IEEE/ACM Trans Audio Speech Lang Process 2014. http://dx.doi.org/10.1109/TASLP.2014.2329734

[18] Benba A, Jilbab A, Hammouch A. Voice analysis for detecting patients with Parkinson's disease using the hybridization of the best acoustic features. Int J Electr Eng Inform 2016. http://dx.doi.org/10.15676/ijeei.8.1.20168

[19] Hlavnicka J, C?mejla R, Tykalová T, Šonka K, Růžicka E, Rusz J. Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. Sci Rep 2017;7:12. http://dx.doi.org/10.1038/s41598-2004017-00047. Neurology

[20] Brückl M, Ghio A, Viallet F. Measurement of tremor in the voices of speakers with Parkinson's disease. Procedia Comp Sci 2018;128:47–54. http://dx.doi.org/10.1016/j.procs.2018.03.007

[21] Orozco-Arroyave JR, Hönig F, Arias-Londoño JD, Vargas-Bonilla JF, Nöth E. Spectral and cepstral analyses for Parkinson's disease detection in Spanish vowels and words. Expert Syst 2015;32:688–97. http://dx.doi.org/10.1111/exsy.12106

[22] Orozco-Arroyave JR, Hönig F, Arias-Londoño JD, Vargas-Bonilla JF, Daqrouq K, Skodda S, et al. Automatic detection of Parkinson's disease in running speech spoken in three different languages. J Acoust Soc Am 2016 Jan;139(1):481–500. http://dx.doi.org/10.1121/1.4939739

[23] Khan T, Westin J, Dougherty M. Cepstral separation difference: a novel approach for speech impairment quantification in Parkinson's disease. Biocybern Biomed Eng 2014;34:25–34. http://dx.doi.org/10.1016/j.bbe.2013.06.001

[24] Afonso LC, Rosa GH, Pereira CR, Weber SA, Hook C, Albuquerque VH, et al. A recurrence plot-based approach for Parkinson's disease identification. Future Gener Comput Syst 2019;282–92. http://dx.doi.org/10.1016/j.future.2018.11.054

[25] Gómez-Vilda P, Mekyska J, Ferrández JM, Palacios-Alonso D, Gómez-Rodellar A, Rodellar-Biarge V, et al. Parkinson disease detection from speech articulation neuromechanics. Front Neuroinform 2017 Aug 25;11:56. http://dx.doi.org/10.3389/fninf.2017.00056

[26] Karan B, Mahto K, Sahu SS. Detection of Parkinson disease using variational mode decomposition of speech signal. 2018 International Conference on Communication and Signal Processing (ICCSP). IEEE; 2018. http://dx.doi.org/10.1109/ICCSP.2018.8524445

[27] Arias-Vergara T, Vásquez-Correa JC, Orozco-Arroyave JR, Nöth E. Speaker models for monitoring Parkinson's disease progression considering different communication channels and acoustic conditions. Speech Commun 2018 Jul 1;101:11–25. http://dx.doi.org/10.1016/j.specom.2018.05.007

[28] Moro-Velázquez L, Gómez-García JA, Godino-Llorente JI, Villalba J, Orozco-Arroyave JR, Dehak N. Analysis of speaker recognition methodologies and the influence of kinetic changes to automatically detect Parkinson's Disease. Appl Soft Comput 2018 Jan 1;62:649–66. doi:16/j.asoc.2017.11.001.

[29] Torres ME, Schlotthauer G, Rufiner HL, Jackson-Menaldi MC. Empirical mode decomposition. Spectral properties in normal and pathological voices. 4th European Conference of the International Federation for Medical and Biological Engineering. 2009. pp. 252–5.

[30] Pereira CR, Pereira DR, Rosa GH, Albuquerque VH, Weber SA, Hook C, et al. Handwritten dynamics assessment through convolutional neural networks: an application to Parkinson's disease identification. Artif Intell Med 2018 May 1;87:67–77. http://dx.doi.org/10.1016/j.artmed.2018.04.001

[31] Pereira CR, Pereira DR, Weber SA, Hook C, de Albuquerque VH, Papa JP. A survey on computer-assisted Parkinson's disease diagnosis. Artif Intell Med 2018 Sep 7. http://dx.doi.org/10.1016/j.artmed.2018.08.007

[32] Orozco-Arroyave JR, Arias-Londoño JD, Vargas-Bonilla JF, González-Rátiva MC, Nöth E. New Spanish speech corpus database for the analysis of people suffering from Parkinson's disease. Lr 2014. Proc Ninth Int Conf Lang Resour Eval 2014;342–7.

[33] Lartillot O, Toiviainen P. A Matlab toolbox for musical feature extraction from audio. International conference on digital audio effects; 2007.

[34] Murty KS, Yegnanarayana B. Combining evidence from residual phase and MFCC features for speaker recognition. IEEE Signal Process Lett 2006;13(1):52–5. http://dx.doi.org/10.1109/LSP.2005.860538

[35] Dhanalakshmi P, Palanivel S, Ramalingam V. Classification of audio signals using AANN and GMM. Appl Soft Comput J 2011;11:716–23. http://dx.doi.org/10.1016/j.asoc.2009.12.033

[36] Cosi P. Evidence against frame-based analysis techniques. 1998 Proceedings of NATO Advance Institute on Computational Hearing; 2009. p. 163–8. doi:10.1.1.39.4812.

[37] Jung C, Han KJ, Seo H, Narayanan SS, Kang H. A variable frame length and rate algorithm based on the spectral kurtosis measure for speaker verification. Perform Eval 2010;2754–7.

[38] Deshpande MS, Holambe RS. Speaker identification based on robust AM-FM features. 2009 2nd Int Conf Emerg Trends Eng Technol ICETET. 2009. pp. 880–4. http://dx.doi.org/10.1109/ICETET.2009.209

[39] Huang NE, Shen Z, Long SR, Wu M-LC, Shih HH, Zheng Q, et al. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond Ser A 1998;454:903–95.

[40] Sharma R, Prasanna SRM, Bhukya RK, Kumar Das R. Analysis of the intrinsic mode functions for speaker information. Speech Commun 2017;91:1–16. http://dx.doi.org/10.1016/j.specom.2017.04.006

[41] Sharma R, Prasanna SRM, Rufiner HL, Schlotthauer G. Detection of the glottal closure instants using empirical mode decomposition. Circuits Syst Signal Process 2018;37:3412–20. http://dx.doi.org/10.1007/s00034-017-0713-4

[42] Flandrin P, Rilling G, Goncalves P. Empirical mode decomposition as a filter bank. IEEE Signal Process Lett 2004;11(2):112–4. http://dx.doi.org/10.1109/LSP.2003.628216

Please cite this article in press as: Karan B, et al. Parkinson disease prediction using intrinsic mode function based features from speech signal. Biocybern Biomed Eng (2019), https://doi.org/10.1016/j.bbe.2019.05.005

[43] Schlotthauer G, Torres ME, Rufiner HL. Voice fundamental frequency extraction algorithm based on ensemble empirical mode decomposition and entropies. World Congress on Medical Physics and Biomedical Engineering, September 7–12, 2009, Munich, Germany; 2009. pp. 984–7. http://dx.doi.org/10.1007/978-3-642-2-2_2620388

[44] Giannakopoulos T, Pikrakis A. Introduction to audio analysis: a MATLAB approach; 2014. http://dx.doi.org/10.1016/C2012-0-4-70352

[45] Bayestehtashk A, Asgari M, Shafran I, McNames J. Fully automated assessment of the severity of Parkinson's disease from speech. Comput Speech Lang 2015;29:172–85. http://dx.doi.org/10.1080/01924036.1990.9289688

[46] Parveen P, Thuraisingham B. Face recognition using multiple classifiers. Proc – Int Conf Tools with Artif Intell ICTAI 2006;179–86. http://dx.doi.org/10.1109/ICTAI.200659

[47] Cortes C, Vapnik V. Cortes_Vapnik95Pdf, vol. 297. 1995;p. 273–97. http://dx.doi.org/10.1023/A:1102262741141

[48] Ai OC, Hariharan M, Yaacob S, Chee LS. Classification of speech dysfluencies with MFCC and LPCC features. Expert Syst Appl 2012 Feb 1;39(2):2157–60. http://dx.doi.org/10.1016/j.eswa.2011.07.065

[49] Rusz J, Novotný M, Hlavnicka J, Tykalová T, Růžicka E. High-accuracy voice-based classification between patients with Parkinson's disease and other neurological diseases may be an easy task with inappropriate experimental design. IEEE Trans Neural Syst Rehabil Eng 2017 Aug;25(8):1319–20. http://dx.doi.org/10.1109/TNSRE.2016.8852621