Journal of Automation, Electronics and Electrical Engineering

PARAMETRIC ANALYSIS OF PILOT VOICE SIGNALS IN PARKINSON'S DISEASE DIAGNOSTICS

Ewelina MAJDA-ZDANCEWICZ[1], **Anna POTULSKA-CHROMIK**[2], **Monika NOJSZEWSKA**[2] and **Anna KOSTERA-PRUSZCZYK**[2]

[1]Military University of Technology, Faculty of Electronics, Kaliskiego 2, 00-908 Warsaw Poland, ewelina.majda@wat.edu.pl
[2]Medical University of Warsaw, Company, Department of Neurology, Warsaw, Poland, apotulska@wum.edu.pl
[3]Medical University of Warsaw, Company, Department of Neurology, Warsaw, Poland, mnojszewska@wum.edu.pl
[4]Medical University of Warsaw, Company, Department of Neurology, Warsaw ,anna.kostera-pruszczyk@wum.edu.pl

DOI: https://doi.org/10.24136/jaeee.2022.003

Abstract – Parkinson's disease (PD) is a neurodegenerative disease of the central nervous system (CNS) characterized by the progressive loss of dopaminergic neurons in the substantia nigra. The article describes an analysis of pilot voice signal analysis in Parkinson's disease diagnostics. Frequency domain signal analysis was mainly used to assess the state of a patient's voice apparatus in order to support PD diagnostics. The recordings covered uttering the "a" sound at least twice with extended phonation. The research utilized real recordings acquired in the Department of Neurology at the Medical University of Warsaw, Poland. Spectral speech signal coefficients may be determined based on different defined frequency scales. The authors used four frequency scales: linear, Mel, Bark and ERB . Spectral descriptors have been defined for each scales which are widely used in machine and deep learning applications, and perceptual analysis. The usefulness of extracted features was assessed taking into account various methods. The discriminatory ability of individual coefficients was evaluated using the Fisher coefficient and LDA technique.. The results of numerical experiments have shown different efficiencies of the proposed descriptors using different frequencies scales.

Key words – features extraction, speech signal analysis, Parkinson's disease,

INTRODUCTION

Neurodegenerative diseases are associated with progressive damage to the cells forming the nervous system structures. These diseases are alarming primarily since mature cells of the nervous system (with very few exceptions) do not tend to regenerate. Thus, researchers' attention is focused on neurodegenerative diseases due to the fact that medicine still does not offer any satisfactory methods of their treatment. Neurodegenerative diseases are characterized by a progressive course, leading over time to major limitations in the patients' everyday life. Therefore, new methods are sought aimed at early detection and precise diagnosis of these disorders.

Parkinson's disease – after Alzheimer's disease – is the second most common neurodegenerative disease. A diagnosis is based on the patient's clinical evaluation, and t[1]he typical symptoms of the disease include movement disorders, such as *bradykinesia*, *resting tremor*, *muscle stiffness* or *postural function disorders*. However, Parkinson's disease (PD) is not only motor symptoms, but also a very wide spectrum of *extra-movement symptoms*. According to modern pathogenetic concepts, it is believed that the disease starts even outside the central nervous system. Pre-motor disease symptoms may precede the appearance of a motor syndrome even by several years, proving a slowly developing neurodegenerative process. *Speech disorders* are the examples of such symptoms. There is no known method for slowing down the disease – it is of progressing nature. Nevertheless, numerous drugs are known to alleviate its symptoms [1, 2].

I. SPEECH DISORDERS

Speech disorders, as one of PD symptoms were already approached by James Parkinson in his paper from 1817 [3]. PD speech disorder patomechanisms was first described by Darley et al. in 1969 [1, 3]. They distinguished characteristic speech features, such as *speech volume reduction*, *monotony*, *speech pitch and loudness disorders*, *accent and breathing reduction*, *hoarseness, inaccurate articulation, speech pace acceleration*. Research conducted by doctors suggests that speech disorders in patients with PD are

[1] **Article financed from the Polish Ministry of National Defence for the implementation of basic research within the research grant No. GBMON/13-996/2018/WAT "Basic research in the field of sensor technology using innovative data processing methods".**

diverse and do not share a common etiology. Their speech changes throughout the day, and it is also dependent on numerous coefficients, such as their *emotional state, taken dose of medication or disease phase*. Digital voice signal processing method based on its parametrization are used to objectively evaluate speech changes in people with Parkinson's disease. This provides a possibility to indicate potential sound features that may constitute differentiating information and a foundation for an engineering support of Parkin-son's disease diagnostics.

The main factor for Parkinson' s Disease is *UPDRS* (*The Unified Parkinson's Disease Rating Scale*). The UPDRS was originally developed in the 1980s and was updated by the Movement Disorder Society. Now it is formed by 4 parts: [4, 5]:

- Part I (Non-motor experiences of daily living),
- Part II (Motor experiences of daily living),
- Part III (Motor examination),
- Part IV (Motor complications).

II. Material

The research presented within the studies are a part of a project on collecting multimodal data set. Data recording set included infrared camera, a camera operating within the visible range, a microphone with a preamplifier, a graphic tablet and a portable computer, which acts as the controller and data integrator [6].

Material included in this research covered only speech signal. The recordings uttering the "a" sound at least twice with extended phonation are used. Such a configuration, regardless of the used computer, enables recording a pure, natural sound without any interference coming from the computer system. All the recordings were registered with a sampling rate of 44.1 kSa/s and a 16-bit resolution.

The recordings for people with PD are conducted at the *Department of Neurology at the Medical University of Warsaw*. The study involved both women and men. The material comprised of 12 patients with diagnosed Parkinson's disease.

The recording for healthy people are conducted at the *Institute of Electronic Systems at the Faculty of Electronics of the Military University of Technology in Warsaw. A comparative group comprised 12 persons without diagnosed PD.*

III. Research method

Acoustic voice analysis in PD patients can be a valuable and objective tool supporting the diagnosing diseases of neurodegenerative nature. Frequency domain signal analysis was mainly used to assess the state of the patient's voice apparatus, in order to support PD diagnostics. The usefulness of spectral signal representation – in relation to the speech signal in particular – results from its generation process. The articulation process mainly shapes the amplitude and frequency signal envelope, while the perception process prior to signal analysis in the brain neural networks extracts components of individual frequencies, using specialized inner ear structures. Based on input data obtained using *Matlab software*, the authors generated amplitude spectra using FFT (*Fast Fourier Transform*). The selection of features describing a given signal is the most important stage from the perspective of engineering various recognition systems. The conducted experiments primarily utilized the *audioFeatureExtractor* function, which enables determining a number of signal parameters in the form of easy-to-distinguish structures. The operation of the *audioFeatureExtractor* function within the Matlab environment is graphically depicted in Fig. 1.

The following descriptors were defined in the case of spectral analysis aimed at evaluating the voice apparatus in people with diagnosed Parkinson's disease using this technique [7, 8]:

- *Spectral Centroid, ,*
- *Spectral Crest,*
- *Spectral Decrease,*
- *Spectral Entropy,*
- *Spectral Flatness,*
- *Spectral Flux,*
- *Spectral Kurtosis,*
- *Spectral RolloffPoint,*
- *Spectral Skewness,*
- *Spectral Slope,*
- *Spectral Spread,*
- *Harmonic ratio,*
- *Pitch*

For example *Pitch* is the elementary frequency of the generated sound, therefore the frequency of the fundamental tone generated during the vibration of the vocal folds of the larynx. *SpectralCentroid* is defined as the centre of gravity of the power density spectrum, i.e. the weighted average frequency of the spectral coefficients. It is a descriptor indicating the frequency at which the sound is focused. This parameter is also related to the so-called *sensation of feeling the sound*. The spectral centroid is identified with perceptual properties such as *brightness* and *sharpness* of sound. *SpectralCrest* is a descriptor that defines the so-called spectral ridge signal. In other words, it reveals how extreme the peaks are in relation to the effective value of the signal [7,8]. A crest factor of 1 means there are no peaks. Higher values of this parameter indicate peaks. For instance, sound waves tend to have high crest factor values. *SpectralFlatness* it is the ratio of the geometric mean to the arithmetic mean of the spectrum. The task of this descriptor is to capture the presence of a large number of peaks in the spectrum. It will assume larger values for sounds showing harmonics or consisting of many distinct individual tones. *SpectralKurtosis* is a descriptor that defines the measure of spectrum distribution flatness around its mean value [7,8,9]. Statistically, it is fourth-order moment. *SpectralRolloffPoint* is a descriptor that determines the frequency below which 95% of the signal energy is contained. *SpectralSkewness* gives a measure of the spectrum distribution asymmetry around its mean value. *Spectral slope* has been used in speech analysis, especially in modeling speaker stress [9,10]. The slope is directly related to the resonant characteristics of the vocal folds and has also been applied to speaker identification.

IV. Experiment Results

The analysis involved 5-second long recordings of the vowel "a". In order to minimize individual coefficients, this part of the experiments was based on the recordings of two women (healthy and sick) with similar fundamental frequency values, so that this parameter did not directly determine the diagnosed morbidity. The waveforms of changes over time in the values of selected parameters over time are shown in the fig. 2. The waveforms on the left are for a healthy person, while the right-side wave-forms are for a person diagnosed with Parkinson's disease.

The first stage of the experiments covered observation of changes to the values of defined features, based on the presented feature extraction method. The obtained results constitute a foundation for assessing the usefulness of extracted descriptors in terms of engineering support for Parkinson's disease diagnostics".

Very large differences between the amplitude of recorded signals are clearly noticeable. This proves reduced volume of a sound articulated by a person diagnosed with Parkinson's disease. The next one parameter is pitch. Analysis shows that the contour of the this descriptors for HC sample is more stable than the contour obtained from the PD patient.

It is worth focusing on the waveform of changes in the *Spectral Centroid* parameters, which is defined as the spectrum's centre of gravity. This parameter is associated with the so-called sound perception sensation. The spectral centroid is a parameter that defines the frequency at which sound is concentrated and is associated with such perceptual properties as sound sharpness and clarity. When looking at the waveforms in fig. 2, the first significant difference between the two recordings is associated with the frequency range. In the case of a sick person, we can see spectral centroid fluctuations. In addition, along with the increase in the "a" sound articulation sound, the analysed voice loses stability, and the spectrum's centre of gravity is reduced by an average of 50 Hz with each successive second of the sound articulated on the recording.

The Spectral Skewness descriptor, which is a measure of asymmetry in the spectral distribution around its mean value is also worth w more comprehensive commentary. Primarily, the values of this coefficient, that are practically twice as high in the case of analysing a sick person's sound, should be stressed. The mean value for a healthy person was 2.03, while in a sick person it was 4.71. In addition, the different nature of the change waveform and the growth in the value of this coefficient along with the time of the recording are also noteworthy. The most likely cause behind such changes is the potential fatigue of the speech apparatus, impacted by the neurodegenerative disease in question.

The changes in the *SpectralCrest* descriptor as waveforms of the sequence of crest coefficients determined for each fundamental tone interval are also worth highlighting. High fluctuations of changes in the values and the growing nature of these changes in the case of a person diagnosed with Parkinson's disease can also be observed when analysing the changes of this parameter. A similar nature of changes has been observed in the case of the feature, which defines the flatness of spectrum distribution around its mean value – *SpectralKurtosis*.

V. Parameter Evaluation

Spectral speech signal coefficients may be determined based on different defined frequency scales. Based on the

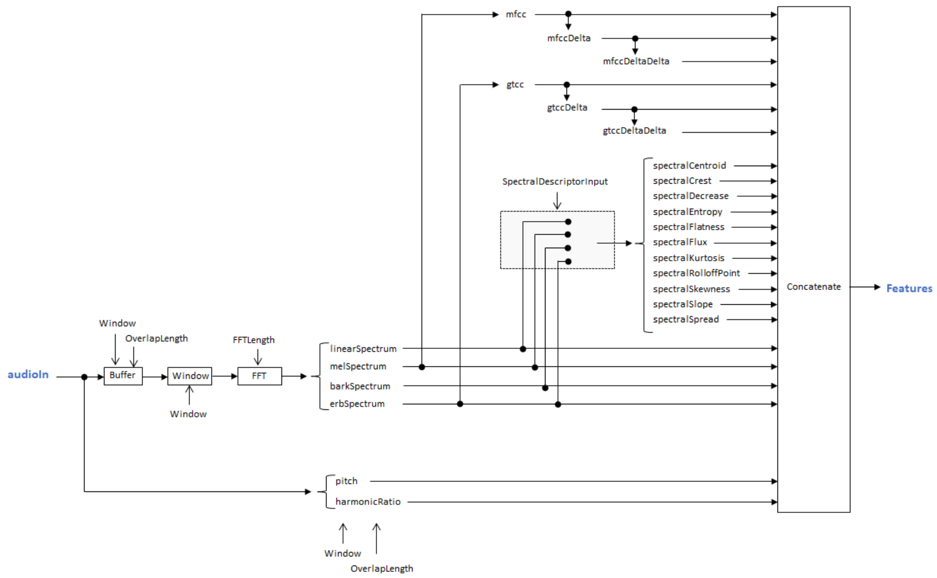

Fig. 1. Graphical depiction of the *audioFeatureExtractor* function within the Matlab environment [5]

algorithm schematically shown in Fig. 1, it is possible to use four frequency scales: *linear, Mel, Bark and ERB* [10].

According to Weber-Fechner's law, the response of a biological system is proportional to the logarithm of the stimulus that excites it. Hence, it can be concluded that the subjective impression of a human does not simply depend on objectively measurable arousal. This means that the human ear does not respond linearly to increasing frequency. And this is the reason why, based on the measurements of the intensity of hearing sounds at different frequencies, perceptual scales were developed, the most important of which are the scales that are expressed in mels and barks. The definitions of these scales are provided below [11,12.].

As far as the *Bark scale* is concerned, the frequency ranges that the ear perceives as a single tone are evenly distributed, and therefore correspond to the critical bandwidth (a concept supported by Corti's organ anatomy, hearing theories and experiments). The scale is defined in the range from 1 to 24, and corresponds to the critical bands of human hearing. The transformation (according to Zwicker) from the Hz to Bark scale is described by the following formula [12]

$$f_{[\mathrm{Bark}]} = 13\ \arctan(0.00076f) + 3.5\arctan((f/7500)^2) \quad (1)$$

ERB scale (*Equivalent Rectangular Bandwidth*) is an auditory scale with 42 critical bands, and is approximately of logarithmic character. ERB is based on an analytically created frequency scale. The centre frequencies of the 42 critical bands are evenly spaced on a scale [12,13]. In terms of numbers, it is equal to the bandwidth of an ideal rectangular filter with a transmittance value equal to the maximum transmittance of an auditory filter, while the power of the noise passing through this filter is equal to the power of the noise passing through the auditory filter. The course of the dependence of an auditory filter equivalent bandwidth on

Fig. 2 Sample waveforms of selected spectral speech signal features.
The waveforms on the left are for a healthy person, while the right-side waveforms are for a person diagnosed with Parkinson's disease

the frequency is described by the following equation: ERB scale (*Eqivalent Recangulat Badnwidth*) is defined by the formula:

$$f_{[\mathrm{Erb}]} = 21.4\, log_{10}(0.00437f + 1) \qquad (2)$$

The *Mel scale* was constructed based on the subjective perception of tones of various pitch, searching for such pitch for a given harmonic that would be audible as half the pitch. The tests were carried out in the entire range of audibility, and the obtained scale corresponds to the subjective impression of pitch [14]. This name originates from the English word *melody*. The formula below shows the transformation of the scale in Hz to the Mel scale. The Mel scale proposed by Stevens, Volkmann and Newman in 1937

$$f[Mel] = 2595\, log_{10}(1 + \frac{f_{[Hz]}}{700}) \qquad (3)$$

Fig. 3 presents various frequency scales used in non-linear voice processing techniques.

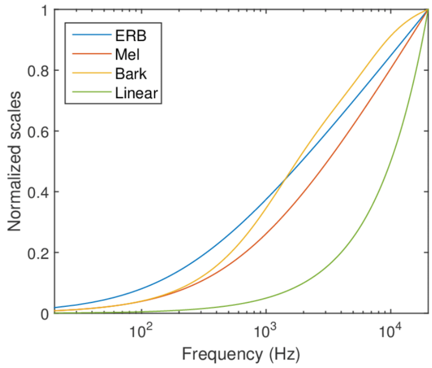

Fig. 3. Different frequency scales using in spectral technique

The response will be a set of 11 features defined in 4 different scales. that remain parameters defined within the frequency domain. 44 features in four different scales, which provided a total of 4 feature subgroups with 11 parameters each, were de-fined based on the suggested algorithm. The *pitch* and the *harmonicRatio* is defined only in linear scale.

The list of all extracted parameters is shown in the Table 1 and table 2.

Extracted parameters (features) can be evaluated in many different ways. The authors of the tests decided to evaluate descriptor quality using the Fisher's method, which is based on calculating the so-called *Fisher's significance coefficient*. The tests were conducted in four separate sets. Each group consisted of 11 spectral parameters extracted using four different frequency scales.

Table 1. Spectral descriptor groups defined within feature extraction stage

Number of descriptors	Spectral descriptor
c_1-c_{13}	Spectral features based on linear frequency scale
c_{14}-$_{24}$	Spectral features based on MEL frequency scale
c_{25}-c_{35}	Spectral features based on BARK frequency scale
c_{35}-c_{45}	Spectral features based on ERB frequency scale

Table 2. List of spectral descriptor defined within feature extraction stage

Number of descriptors	Spectral descriptor
$c_1, c_{14}, c_{25}, c_{36}$	"spectralCentroid"
$c_2, c_{15}, c_{26}, c_{37}$	"spectralKurtosis"
$c_2, c_{16}, c_{27}, c_{38}$	"spectralCrest"
$c_4, c_{17}, c_{28}, c_{39}$	"spectralDecrease"
$c_5, c_{18}, c_{29}, c_{40}$	"spectralEntropy"
$c_6, c_{19}, c_{30}, c_{41}$	"spectralFlatness"
$c_7, c_{20}, c_{31}, c_{42}$	"spectralRolloffPoint"
$c_8, c_{21}, c_{32}, c_{43}$	"spectralSkewness"
$c_9, c_{22}, c_{33} c_{44}$	"spectralSlope"
$c_{10}, c_{23}, c_{34}, c_{45}$	"spectralSpread"
$c_{11}, c_{24}, c_{35}, c_{46}$	"spectralFlux"
$c_{11},$	"pitch"
$c_{12},$	"harmonicRatio"

The Fisher score is a factor which describes the Fisher method. This measure defined the significance of the feature for recognizing the samples belonging to two classes: [15,16]

$$S_{ij}(f) = \frac{|c_i - c_j|}{\sigma_i + \sigma_j} \qquad (4)$$

where c_1, c_2, represent the mean values and σ_1, σ_2 represent the standard deviations of feature in the first and second classes, respectively. If the value of factor is higher than the features is more significant in recognition between classes 1 and 2. Each group consisted of set of spectral parameters extracted using four different frequency scales. Fig 4 ,5, 6 and 7. collectively shows the total Fisher coefficients of each group of descriptors. Table 3 shows results for all the groups.

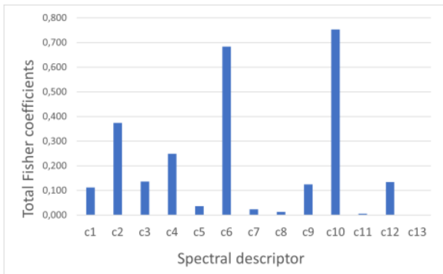

Fig 4. Total Fisher coefficients of spectral descriptor groups defined within the experiment

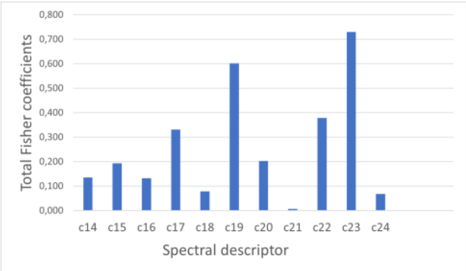

Fig 5. Total Fisher coefficients of spectral descriptor groups defined within the experiment

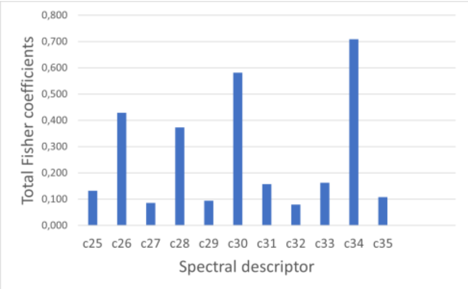

Fig 6. Total Fisher coefficients of spectral descriptor groups defined within the experiment

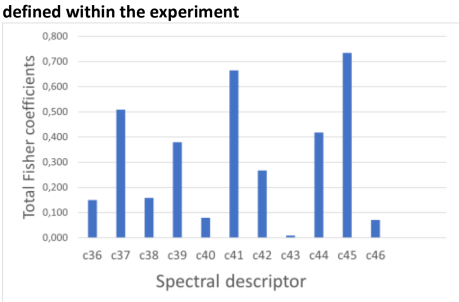

Fig 7. Total Fisher coefficients of spectral descriptor groups defined within the experiment

Table 3. Total Fisher coefficients of spectral descriptor groups defined within the experiment

Descriptors	linear scale	MEL scale	BARK scale	ERB scale
"spectralCentroid"	0,112	0,135	0,150	0,132
"spectralKurtosis"	0,374	0,194	0,509	0,429
"spectralCrest"	0,136	0,132	0,158	0,086
"spectralDecrease"	0,249	0,331	0,379	0,373
"spectralEntropy"	0,036	0,078	0,080	0,094
"spectralFlatness"	**0,683**	**0,602**	**0,665**	**0,582**
"spectralRolloffPoint"	0,023	0,202	0,268	0,157
"spectralSkewness"	0,013	0,006	0,009	0,080
"spectralSlope"	0,125	0,378	0,418	0,163
"spectralSpread"	**0,753**	**0,730**	**0,735**	**0,709**
"spectralFlux"	0,006	0,068	0,071	0,108
"spectralCentroid"	0,112	0,135	0,150	0,132
"Pitch"	0,133	-	-	
"Harmonicratio"	0,001	-	-	-

The following descriptors have reached the highest values for each of the defined features: "*spectralSpread*" and "*spectralFlatness*". Moreover, in the case of coefficients extracted using the Mel scale, good values are achieved by "*spectralDecrease*" and "*spectralSlope*", "*spectralKurtosis*" and "*spectralDecrease*" for ERB, and "*spectralKurtosis*", "*spectralDecrease*" and "*spectralSlope*" for the Bark scale. Therefore, it can be seen that different descriptors achieve various values when using varying frequency definition scales.

A reliable spectral descriptor evaluation requires an expanded selection or applying an evaluation methodology using a classifier. Based on the Fisher's measure values obtained in the subgroups, the authors extracted these descriptors that seem to be a good prognostic for creating a resultant feature vector differentiating between a voice of a healthy and sick person. The fusion of the features from the four subgroups assumed defining a feature vector based on these descriptors, for which the Fisher's measure value exceeded the 0.3 threshold adopted by the authors. It gives 16-elementts sets of features. An LDA classifier result example for two features vectors are shown in Fig. 8.

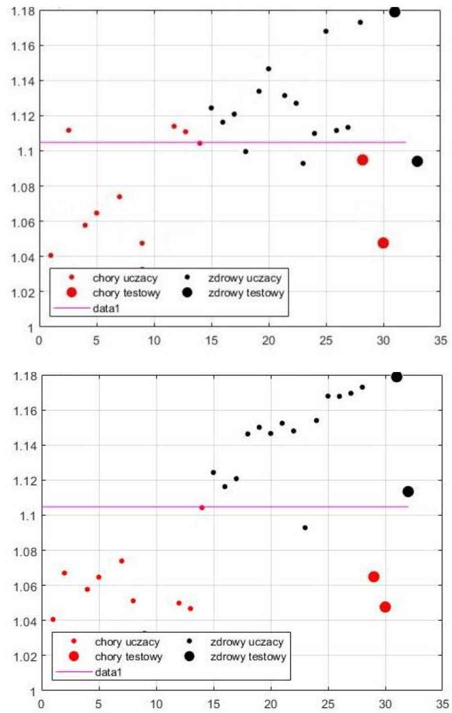

Fig. 8. Sample LDA classifier results for spectral features, the Fisher's measure value of which has reached >0.3 (chory – sick person, zdrowy – healthy person, uczący – learning, testowy – test)

VI. Conclusions

The objective of this article was to analyse the recording of pilot voice signals aimed at detecting pathological voice change in people with Parkinson's disease. The first stage targeted preliminary speech signal transformations in order to obtain a set of parameters, the value of which will constitute the foundation for a system supporting PD diagnostics. The author mainly focused on spectral speech signal features. The variability of extracted features was assessed using different methods!

The first one is to make a observation of changes to the values of spectral descriptor in lineal scale. Global analysis shows that the contour of the all descriptors for HC sample is more stable than the contour obtained from the PD patient The difference in experiment results is caused primarily by weaker operation of the vocal folds. It is determines as well as by the stiffness of the pharynx, oral cavity and larynx muscles in patients with Pakinsons Disease.

Additionally, the variability of the extracted features was assessed with the use of various features (Fischer coefficients and LDA) The main aim for this research is to determine the spectral coefficients which is the most usefulness of the extracted feature regardless of used frequency scale. Fisher's coefficients obtained similar high value for two descriptors in all subgroup. It is "*spectralSpread*" anf "*spectralFlatness".* Spectral flatness is an indication of the peakiness of the spectrum. A higher spectral flatness indicates noise, while a lower spectral flatness indicates tonality. The spectral spread represents the "instantaneous bandwidth" of the spectrum. It is used as an indication of the dominance of a tone. For example, the spread increases as the tones diverge and decreases as the tones converge

What's more we used another frequency scale and different descriptors achieved good values. If we want to determine unique spectral features sets we can combine different spectral features defined using different frequency scale. Moreover we defined threshold based on value of total Fisher coefficients of spectral descriptor groups defined within the experiment. In results we selected 16 features which give us the best results.

The results obtained in the PD research studies showed that the choice of feature extraction and learning algorithms directly influences the accuracy and reliability of the proposed system. Furtherome the choice of used frequency scale in spectral descriptors give us new usefull information to determine unique features sets.

ANALIZA PARAMETRYCZNA PILOTAŻOWYCH SYGNAŁÓW GŁOSU W DIAGNOSTYCE CHOROBY PARKINSONA

Choroba Parkinsona (PD) jest neurodegeneracyjną chorobą ośrodkowego układu nerwowego charakteryzującą się postępującą utratą neuronów dopaminergicznych w istocie czarnej. W artykule opisano analizę rejestracji pilotażowych sygnałów głosu w diagnostyce choroby Parkinsona. Rejestracji podlegało co najmniej dwukrotnie wypowiadanie głoski „a" o przedłużonej fonacji. Do badań wykorzystano nagrania zarejestrowane w Katedrze i Klinice Neurologii Warszawskiego Uniwersytetu Medycznego w Warszawie. Do oceny stanu aparatu głosu pacjenta celem wsparcia diagnostyki choroby Parkinsona wykorzystano w głównej mierze analizę sygnału w dziedzinie częstotliwości. Autorzy zastosowali cztery skale częstości: liniową, skalę typu Mel, skalę typu Bark oraz skalę typu ERB. Dla każdej z tych skali zdefiniowali deskryptory spektralne szeroko stosowane w aplikacjach uczenia maszynowego i głębokiego uczenia się oraz w analizie percepcyjnej. Ocena przydatności wyekstrahowanych cech została zrealizowana z uwzględnieniem różnych metod. Wykorzystano metodą oceny jakości cech przy użyciu współczynnika istotności Fischera oraz analizę LDA. Wyniki eksperymentów numerycznych wykazały różne wydajności proponowanych deskryptorów przy użyciu różnych skal częstości..

Słowa kluczowe: analiza sygnału mowy, choroba Parkinsona, ekstrakcja cech,

Bibliography

[1] Pawlukowska W., Honczarenko K., Gołąb-Janowska M (2013)., „Charakter zaburzeń mowy w chorobie Parkinsona", Neurologia i Neurochirurgia Polska, vol. 47, nr 3, pp. 263-269.

[2] AJ. Hughes, SE Daniel, L, Kilford, AJ. Lees (1992), "Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases" J Neurol Neurosurg Psychiatry, 55 (3),, spp181-184.

[3] Parkinson J., An Essay on the Shaking Palsy. 1817 (2002), "The Journal of Neuropsychiatry", vol. 4, issue 2, pp. 223-236.

[4] Braak H, Del Tredici K, Rüb U et al. (2003), "Staging of brain pathology related to sporadic Parkinson's disease". Neurobiol Aging, 24, pp.197-211

[5] *Movement Disorders Society- Updated Unified Parkinson's Disease Rating Scale, MDS-UPDRS* (Goetz et al., 2007, 2008a; Martinez–Martin et al., 2013)

[6] J. Chmielińska, K. Białek, A. Potulska-Chromik, J. Jakubowski, E. Majda-Zdancewicz, M. Nojszewska, A. Kostera-Pruszczyk and A. Dobrowolski (2019), "Multimodal data acquisition set for objective assessment of Parkinson's disease", Proc. SPIE 11442, Radioelectronic Systems Conference, 114420F, (2020).

[7] https://www.mathworks.com/help/audio/ug/spectral-descriptors.html

[8] Peeters G (2004)., "A large set of audio features for sound description (similarity and classification) in the CUIDADO project", Paris,

[9] Misra H, et al, (2014) "Spectral Entropy Based Feature for Robust ASR", June

[10] Alías F., Socoró J. C., Sevillano X. (2016), "A Review of Physical and Perceptual Feature Extraction Techniques for Speech", Music and Environmental Sounds, Appl. Sci., 6(5), 143;

[11] Valero X., Alias F. (2012), "Gammatone Cepstral Coefficients: Biologically InspiredFeatures for Non-Speech Audio Classification, "IEEE transactions on multimedia", vol. 14, no. 6,

[12] Francesc Alías, Joan Claudi Socoró, Xavier Sevillano (2016)," A Review of Physical and Perceptual Feature Extraction Techniques for Speech, Music and Environmental Sounds", Appl. Sci., 6(5), 143;

[13] Nikhil G.V., Keerthi A.M., and Premananda B.S. (2017)," Impact of ERB and Bark scales on Perceptual Distortion based Near-end Speech Enhancement", Department of Telecommunication Engineering,

[14] L. Jeancolas et al., (2017) "Automatic detection of early stages of Parkinson's disease through acoustic voice analysis with mel-frequency cepstral coefficients," 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP),

[15] F. Gil and S. Osowski (2021), "Feature selection methods in gene recognition problem", BULLETIN OF THE POLISH ACADEMY OF SCIENCES TECHNICAL SCIENCES, Vol. 69(3),
[16] I. Guyon and A. Elisseeff (2003)"An introduction to variable and feature selection", *J. Mach. Learn. Res.* 3, 1158–1182 .