

Ewa Niebudek-Bogusz¹Jacek Grygiel²Paweł Strumiłło²Mariola Śliwińska-Kowalska¹

NIELINIOWA ANALIZA AKUSTYCZNA W OCENIE ZAWODOWYCH ZABURZEŃ GŁOSU

NONLINEAR ACOUSTIC ANALYSIS IN THE EVALUATION OF OCCUPATIONAL VOICE DISORDERS

¹ Instytut Medycyny Pracy im. prof. J. Nofera / Nofer Institute of Occupational Medicine, Łódź, Poland
Klinika Audiologii i Foniatrii / Department of Audiology and Phoniatrics² Politechnika Łódzka / Technical University of Lodz, Łódź, Poland
Instytut Elektroniki / Institute of Electronics

STRESZCZENIE

Wprowadzenie: W ostatnich latach podkreśla się, że tworzenie głosu podlega również procesom nieliniowym, powodującym aperiodyczne drgania fałdów głosowych. Drgania te nie zawsze można scharakteryzować za pomocą konwencjonalnych parametrów akustycznych, takich jak wskaźniki perturbacji częstotliwości i amplitudy. Z tego powodu wzrasta zainteresowanie nieliniowymi metodami analizy akustycznej. Celem pracy była ocena możliwości zastosowania nieliniowej analizy cepstralnej z wyznaczaniem współczynników cepstralnych (mel frequency cepstral coefficients – MFCC) w diagnozowaniu zawodowych zaburzeń głosu. **Materiał i metody:** Badaniem objęto 275 próbek głosów patologicznych (głoski „a” w przedłużonej fonacji oraz 4 standaryzowanych zdań) rejestrowanych u nauczycielek z zawodowo uwarunkowanymi łagodnymi zmianami przerostowymi fałdów głosowych (tj. guzkami głosowymi, polipami) oraz 200 próbek głosów prawidłowych kobiet z grupy kontrolnej. Średnia wieku grupy badanej wynosiła 45 lat, a grupy kontrolnej – 43 lata. W badaniach wykonano analizę próbek głosowych za pomocą nieliniowej analizy akustycznej, w której zastosowano parametry MFCC. **Wyniki:** Po klasyfikacji MFCC przy pomocy odwzorowania Sammona oraz tzw. Maszyny Wektorów Nośnych uzyskano dużą trafność badanej metody. W testach wykonanych dla 475 zarejestrowanych próbek głosu zaburzenia głosu zostały wykryte z 91-procentową czułością i 83-procentową specyficznością dla głoski „a”, a dla badanych zdań z czułością i specyficznością w granicach 87–100%. **Wnioski:** Nieliniowa analiza akustyczna z zastosowaniem współczynników MFCC może być przydatnym, obiektywnym narzędziem potwierdzającym zmiany patologiczne głosi o podłożu zawodowym. Prowadzone są dalsze badania w tym kierunku. Med. Pr. 2013;64(1):29–35

Słowa kluczowe: nieliniowa analiza akustyczna, zawodowe zaburzenia głosu, łagodne zmiany przerostowe fałdów głosowych, ŁZFG

ABSTRACT

Background: Over recent years numerous papers have stressed that production of voice is subjected to the nonlinear processes, which cause aperiodic vibrations of vocal folds. These vibrations cannot always be characterized by means of conventional acoustic parameters, such as measurements of frequency and amplitude perturbations. Thus, special attention has recently been paid to nonlinear acoustic methods. The aim of this study was to assess the applicability of nonlinear cepstral analysis, including the evaluation of mel cepstral coefficients (MFCC), in diagnosing occupational voice disorders. **Material and methods:** The study involved 275 voice samples of pathologic voice (sustained vowel “a” and four standardized sentences) registered in female teachers with the occupation-related benign vocal fold masses (BVF), such as vocal nodules, polyps, and 200 voice samples of normal voices from the control group of females. The mean age of patients and controls was similar (45 vs. 43 years). Voice samples from both groups were analyzed, including MFCC evaluation. **Results:** MFCC classification using the Sammon Mapping and Support Vector Machines yielded a considerable accuracy of the test. Voice pathologies were detected in 475 registered voice samples: for vowel “a” with 86% sensitivity and 90% specificity, and for the examined sentences the corresponding values varied between 87% and 100%, respectively. **Conclusions:** Nonlinear voice analysis with application of mel cepstral coefficients could be a useful and objective tool for confirming occupational-related lesions of the glottis. Further studies addressing this problem are being carried out. Med Pr 2013;64(1):29–35

Key words: nonlinear acoustic analysis, occupational voice disorders, benign vocal fold masses, BVF

Adres 1. autorki: Klinika Audiologii i Foniatrii, Instytut Medycyny Pracy im. prof. J. Nofera w Łodzi,
ul św. Teresy, 91-348 Łódź, e-mail: ebogusz@imp.lodz.pl
Nadesłano: 21 września 2012, zatwierdzono: 5 grudnia 2012

WSTĘP

Głos jako barometr zdrowia fizycznego i psychicznego wymaga kompleksowej oceny. Spośród metod diagnostycznych, zalecanych przez Komitet ds. Foniatrii Europejskiego Towarzystwa Laryngologicznego (Committee on Phoniatrics of the European Laryngological Society) (1) analiza akustyczna wyróżnia się jako metoda obiektywna i nieinwazyjna.

Ostatnio obserwuje się rozwój zaawansowanych metod cyfrowego przetwarzania dźwięków w celu ich analizy. Klasyczna Fourierowska analiza akustyczna głosu jest coraz częściej uzupełniana o metody nieliniowe (2,3). Nieliniowa analiza homomorficzna z zastosowaniem współczynników cepstralnych umożliwia wyznaczenie precyzyjnych parametrów sygnału pobudzającego (tj. dźwięków wytwarzanych bezpośrednio przez głośnie) na podstawie rejestrowanych próbek mowy (4). Analiza taka daje pełniejszy obraz parametrów charakteryzujących drgania fałdów głosowych niż analiza Fourierowska, w wyniku której wyznaczane są konwencjonalne parametry akustyczne, takie jak parametry perturbacji częstotliwości (grupa jittera) i amplitudy (grupa shimmera). Parametry te przeznaczone są do oceny głosów prawie periodycznych ze stałą dominantą, która warunkuje jednakową częstotliwość podstawową w przebiegu fali akustycznej (5). Z kolei w literaturze przedmiotu podkreśla się ostatnio, że tworzenie głosu obejmuje nie tylko procesy linearne, ale również nielinearne, które powodują aperiodyczne drgania fali akustycznej (6,7).

Z tego powodu do oceny klinicznej zaburzeń głosu coraz częściej próbuje się stosować nieliniowe metody akustyczne (8). Pozwalają one na bardziej precyzyjną akustyczną ocenę zaburzeń głosu, w tym także małych zmian organicznych głośni, w których etiopatogeneza ma często podłoże zawodowe. Wymienione zmiany patologiczne krtani, określane od niedawna w piśmiennictwie polskim jako łagodne zmiany fałdów głosowych (ŁZFG), są synonimem benign vocal fold masses (BVFM) – nazwy od wielu lat stosowanej w literaturze angloamerykańskiej. W etiologii tych zmian bardzo istotną rolę odgrywa nadmierny wysiłek głosowy. Z tego powodu ŁZFG pojawiają się u osób obciążających głos zawodowo, często w efekcie wadliwej emisji głosu lub braku jego higieny (9).

W wyniku działania sił wibracyjnych podczas nieprawidłowej fonacji dochodzi do uszkodzenia błaski właściwej błony podstawnej fałdu głosowego i powstania zgrubień zlokalizowanych na wolnych brzegach

fałdów głosowych (10). Na skutek tych przewlekłych mikrourazów głośni (fonotraumy) tworzą się ŁZFG w postaci guzków głosowych, polipów, zmian obrzkowo-przerostowych fałdów głosowych czy torbieli fałdów głosowych, które powodują nielinearne procesy w trakcie przepływu powietrza przez głośnie w fazie fonacyjnej. Łagodne zmiany fałdów głosowych zaburzają jakość głosu i wydolność aparatu głosowego w różnym stopniu (11). Z kolei dla zaburzeń głosu o podłożu zawodowym istotna jest obiektywna ocena dysfunkcji narządu głosu, także z uwzględnieniem ich nielinearnej dynamiki.

Z tego powodu celem pracy była ocena możliwości zastosowania nieliniowej metody analizy akustycznej w oparciu o wyznaczanie współczynników cepstralnych (mel cepstral coefficients – MFCC) do badania zawodowych zaburzeń głosu w przebiegu łagodnych zmian fałdów głosowych.

MATERIAŁY I METODY

Zarejestrowano 275 próbek tzw. głosów patologicznych u 55 nauczycielek, stanowiących grupę badaną, oraz 200 próbek głosów prawidłowych u 40 kobiet z grupy kontrolnej, u których dokonano percepcyjnej oceny głosu. Średnia wieku grupy badanej wynosiła 45 lat, a grupy kontrolnej – 43 lata. Do grupy badanej pacjentki z łagodnymi zmianami przerostowymi fałdów głosowych kwalifikowano na podstawie badania foniatrycznego z laryngowideostroboskopią.

Do badań akustycznych nagrywano 5 próbek głosu: głoskę „a” w przedłużonej fonacji oraz 4 wystandaryzowane dla języka polskiego zdania: 1) Ten dzielny żołnierz był z nim razem, 2) Czy jestem zdrowy?, 3) Tak, jestem zdrowy, i 4) Już jestem zdrowy! Wykonano analizę próbek głosowych uzyskanych od obu grup. Wszystkie próbki były randomizowane.

Przeprowadzono nieliniową homomorficzną analizę akustyczną (12), w której oceniano parametry MFCC oraz wartości 3 pierwszych formantów. W przetwarzaniu sygnału głosu w celu jego parametryzacji najpierw przeprowadzono proces wydzielania cech głosu za pomocą obliczania MFCC, a następnie dokonano klasyfikacji tych współczynników. W opisywanej analizie nieliniowej zastosowano odwrotne przekształcenie Fouriera, otrzymując tzw. cepstra, które dla uproszczenia można zapisać w postaci równania:

$$C_s(t_q) = C_x(t_q) + C_h(t_q) \quad [1]$$

gdzie:

$C_s(t_q)$ – suma cepstrów definiowanych w tzw. dziedzinie que-frency mającej wymiar czasu (w odróżnieniu od frequency – oznaczającego częstotliwość),

$C_x(t_q)$ – cepstrum opisującym własności dźwięków wytwarzanych przez fałdy głosowe (m.in. ton krtaniowy F_0),

$C_h(t_q)$ – cepstrum związanym z „wysokotonową” charakterystyką traktu głosowego.

Cepstra (nazywane też współczynnikami cepstrów dla dyskretnej skali que-frency) występujące po prawej stronie ww. równania można łatwo rozdzielić za pomocą odpowiedniej filtracji liniowej.

Taka analiza umożliwia uzyskanie tzw. współczynników cepstralnych, które można łatwo filtrować za pomocą odpowiedniego banku filtrów. W proponowanej metodzie nieliniowej analizy akustycznej zastosowano bank 12 filtrów definiowanych w tzw. melowej skali częstotliwości. Skala melowa jest jedną z perceptualnych skal częstotliwości. Człowiek dokładniej rozróżnia małe częstotliwości od częstotliwości dużych. Aby odwzorować te właściwości systemu słuchowego człowieka, przekształca się skalę częstotliwości wyrażoną w hertzech (Hz) do skali częstotliwości wyrażonej w melach. Wartości współczynników MFCC są wynikiem filtracji cepstrów za pomocą kolejnych filtrów z banku filtrów melowych.

Rejestrowany sygnał głosu badanych pacjentów analizowano dla kolejnych okien czasowych o czasie trwania 20 ms. Dla każdego okna wyznaczano następujące grupy parametrów:

- 12 współczynników mel-cepstralnych MFCC,
- 12 pochodnych 1. rzędu współczynników mel-cepstralnych Δ MFCC,
- 12 pochodnych 2. rzędu współczynników mel-cepstralnych $\Delta\Delta$ MFCC,
- energię sygnału w oknach czasowych E oraz pochodne tej energii, odpowiednio: pierwszego ΔE i drugiego rzędu $\Delta\Delta E$,
- wartości trzech pierwszych formantów.

Reasumując, dla każdego analizowanego okna czasowego wyznaczono wektor o liczbie 42 parametrów (39 współczynników mel-cepstralnych MFCC oraz wartości 3 pierwszych formantów) opisujących chwilowe własności sygnału głosu. Parametry obliczane z kolejnych okien uśredniono dla wyznaczenia jednego 42-elementowego wektora charakteryzującego właściwości sygnału głosu badanego pacjenta. Najpierw rzutowano 42-wymiarową przestrzeń do przestrzeni o mniejszej liczbie wymiarów z zachowaniem istotnych parametrów

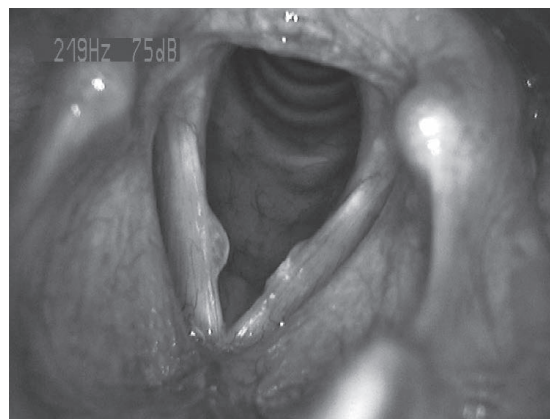
za pomocą odwzorowania Sammona. W dalszej analizie do klasyfikacji uzyskanych parametrów zastosowano Klasyfikator Minimalnoodległościowy (Minimum Distance Classifier – MDC), a następnie Maszynę Wektorów Nośnych (Support Vector Machine – SVM) (13).

Skuteczność rozpoznawania głosów nieprawidłowych za pomocą współczynników MFCC określano, obliczając dokładność rozpoznania za pomocą parametrów stosowanych do wyznaczania charakterystyki ROC (receiver operating characteristic – charakterystyka jakości klasyfikatora), tj. czułość, swoistość, przewidywaną częstość u osób z wynikiem nieprawidłowym – dodatnim (wartość predykcyjna dodatnia), oraz przewidywaną częstość u osób z wynikiem prawidłowym – ujemnym (wartość predykcyjna ujemna). Parametry te obliczono w odniesieniu do wyników laryngowideostroboskopii – metody diagnostycznej, która jest uznawana jako złoty standard w diagnozowaniu patologii narządu głosu.

Na prowadzenie opisanych badań akustycznych otrzymano zgodę Komisji Bioetycznej Instytutu Medycyny Pracy im. J. Nofera w Łodzi (decyzja nr 6/2012).

WYNIKI

W grupie badanej obejmującej 55 kobiet z łagodnymi zmianami przerostowymi fałdów głosowych o podłożu zawodowym znalazły się 33 osoby z guzkami głosowymi, 18 z polipami fałdów głosowych oraz 14 osób z małymi zmianami organicznymi głośni, które trudno zaklasyfikować jednoznacznie. Przykładem tych ostatnich zmian patologicznych krtani jest obraz wideostroboskopowy głośni nauczycielki przedstawiony na rycinie 1.



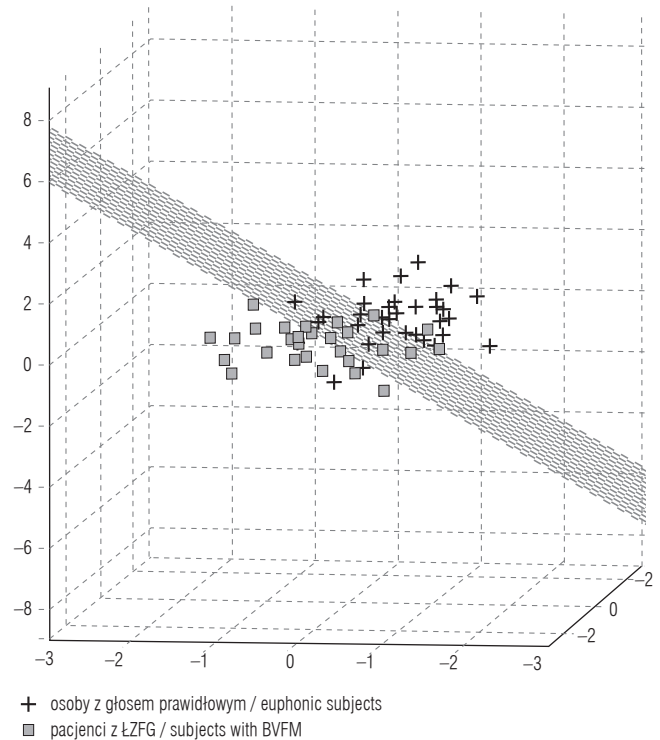
Ryc. 1. Łagodne zmiany przerostowe fałdów głosowych (ŁZFG) u nauczycielki (opis w tekście)

Fig. 1. Benign vocal fold masses in the female teacher (description in the text)

Na prawym fałdzie głosowym na jego wolnym brzegu widoczna jest niewielka torbiel, a na lewym fałdzie głosowym symetrycznie występuje „odcisk” – zgrubienie błony podstawnej wolnego brzegu jako efekt przewlekłej fonotraumaty. W takich przypadkach najbardziej adekwatną nazwą patologii krtani jest ŁZFG.

Po wyznaczeniu współczynników mel-cepstalnych MFCC dokonano ich klasyfikacji opisanej w metodach. Po zastosowaniu odwzorowania Sammona w celu zredukowania wektora danych do 3 wymiarów otrzymano płaszczyznę oddzielającą współczynniki MFCC charakteryzujące głosy prawidłowe od współczynników MFCC wyznaczonych dla głosów patologicznych (ryc. 2). Błąd funkcji dla charakterystyki Sammona wynosił ok. 1,26 (10–14). Po zastosowaniu MDC zaburzenia głosu u osób z łagodnymi zmianami fałdów głosowych zostały wykryte z dobrą czułością i specyficznością.

Tabela 1. przedstawia wartości parametrów krzywej ROC w zależności od rodzaju próbek akustycznych. Dla badanego testu – analizy mel-cepstalnej – uzyskano dobrą wykrywalność głosów patologicznych osób z ŁZFG, o czym świadczą prezentowane w tej tabeli wyniki czułości, swoistości, wartości predyktywnej dodatniej i ujemnej. Jeszcze wyższą rzetelność dla różnicowania głosów patologicznych i eufonicznych dla badanych grup uzyskano po zastosowaniu klasyfikatora SVM (tab. 2). Głosy nieprawidłowe osób z ŁZFG zostały wykryte z czułością 86% dla głoski, a dla poszczególnych nagrywanych zdań z czułością w granicach 87–100%. Specyficzność dla głoski wynosiła 91%, a dla poszczególnych zdań osiągała wartości w granicach 86–100%. Otrzymano też wysoką wartość predyk-



Ryc. 2. Rozkład współczynników cepstralnych (MFCC) charakteryzujących głosy nauczycielek z łagodnymi zmianami przerostowymi fałdów głosowych (ŁZFG) oraz prawidłowe głosy kobiet z grupy kontrolnej

Fig. 2. Visualisation of mel cepstral frequency coefficients (MFCC), distinguishing between voices of female teachers suffering from benign vocal fold masses (BVFM) and normal voices of females from the control group

tywną ujemną – dla wszystkich próbek głosowych wynosiła ona średnio 88%, co ma duże znaczenie z punktu widzenia klinicznego.

Tabela 1. Parametry rzetelności testu – nieliniowej analizy mel-cepstalnej w ocenie głosu nauczycielek ze zmianami przerostowymi fałdów głosowych (ŁZFG) – po zastosowaniu MDC

Table 1. The accuracy parameters of the test – non-linear mel-cepstral analysis in the voice evaluation of female teachers with benign vocal fold masses (BVFM) – application of the MDC

MDC MDC	Czułość Sensitivity [%]	Swoistość Specificity [%]	Wartość predyktywna dodatnia Positive predictive value [%]	Wartość predyktywna ujemna Negative predictive value [%]
Głoska „a” / Vowel „a”	71	73	73	61
Zdanie 1 / Sentence 1	73	86	85	74
Zdanie 2 / Sentence 2	78	81	89	65
Zdanie 3 / Sentence 3	70	76	85	57
Zdanie 4 / Sentence 4	80	76	86	67

MDC – Klasyfikator Minimalnoodległościowy / Minimum Distance Classifier.

Tabela 2. Parametry rzetelności testu – nieliniowej analizy mel-cepstralnej w ocenie głosu nauczycielek ze zmianami przerostowymi fałdów głosowych (ŁZFG) – po zastosowaniu klasyfikatora SVM

Table 2. The accuracy parameters of the test – non-linear mel-cepstral analysis in the voice evaluation of female teachers with benign vocal fold masses (BVFM) – application of SVM classifier

Klasyfikator SVM SVM Classifier	Czułość Sensitivity [%]	Swoistość Specificity [%]	Wartość predykcyjna dodatnia Positive predictive value [%]	Wartość predykcyjna ujemna Negative predictive value [%]
Głoska „a” / Vowel ”a”	86	91	92	82
Zdanie 1 / Sentence 1	100	86	85	100
Zdanie 2 / Sentence 2	89	88	89	95
Zdanie 3 / Sentence 3	87	88	93	78
Zdanie 4 / Sentence 4	93	100	100	86

SVM – Maszyna Wektorów Nośnych / Suport Vector Machine.

OMÓWIENIE

W artykule przedstawiono możliwość oceny zaburzeń głosu w przebiegu łagodnych zmian fałdów głosowych za pomocą nowej metody – nieliniowej analizy akustycznej – z zastosowaniem współczynników mel-cepstralnych. Opisywane zmiany patologiczne głosu występują w postaci guzków głosowych, małych polipów, torbieli czy zmian obrzękowych, zlokalizowanych na brzegach fałdów głosowych. Zmiany te, mimo że niewielkich rozmiarów, mogą istotnie wpływać na charakterystykę akustyczną głosu, powodując nieprawidłowy, czasami nieliniowy rozkład wszystkich składowych, tj. częstotliwości, natężenia i barwy.

Nieliniowy przepływ powietrza przez głośnie w ŁZFG powoduje zaburzenia przesunięcia brzeźnego, czego efektem może być aperiodyczność drgań, która nasila się podczas hyperfonacji. Hyperkinetyczny sposób tworzenia głosu jest charakterystyczny dla osób zawodowo obciążających głos – szczególnie tych, które nie opanowały jego prawidłowej emisji. Dysfonia o typie hyperfunkcyjnym może być przyczyną lub następstwem ŁZFG. W wyniku ŁZFG nie zawsze występuje nasilona chrypka, natomiast występują dyskretnie zmiany w pasmach energii akustycznej, które często wpływają negatywnie na jakość i wydolność głosu. Zmiany te są szczególnie trudne do detekcji w głosach aperiodycznych (5,14,15).

Przedstawiana metoda analizy cepstralnej z wykorzystaniem współczynników MFCC jest przeważnie stosowana do rozpoznawania głosów rozmówców (14). Z kolei prezentowane badania są pierwszą w Polsce

próbą jej aplikacji klinicznej w diagnozowaniu patologii krtani. Badaniami objęto 55 nauczycielek z łagodnymi zmianami przerostowymi głosi, potwierdzonymi podczas laryngowideostroboskopii, oraz 40 kobiet z grupy kontrolnej o prawidłowym głosie. W sumie testowano 275 próbek tzw. głosów patologicznych oraz 200 próbek głosów prawidłowych, ponieważ od każdej badanej osoby uzyskiwano po 5 próbek głosowych: głoskę „a” w przedłużonej fonacji oraz 4 wystandaryzowane zdania.

W testach otrzymano wysoką wykrywalność głosów patologicznych osób z ŁZFG. Z dużą trafnością wykrywano głosy nieprawidłowe szczególnie po zastosowaniu dla otrzymanych MFCC klasyfikatora SVM. Dla głoski „a” uzyskano czułość 86%, natomiast dla zdań czułość mieściła się w granicach 87–100%. Specyficzność dla głoski wynosiła 90%, a dla zdań przyjmowała wartości podobnie wysokie jak dla czułości. Należy podkreślić też otrzymaną wysoką wartość predykcyjną ujemną badanego testu – dla wszystkich próbek głosowych wynosiła ona średnio 88%. Jest to istotne dla oceny jakości głosu zarówno w praktyce foniatrycznej, jak i laryngologicznej. W klasyfikowaniu głosów prawidłowych i nieprawidłowych większym błędem z punktu widzenia klinicznego będzie zaklasyfikowanie głosu osoby z ŁZFG jako głosu eufonicznego niż wskazanie głosu prawidłowego jako patologicznego.

Analogiczne wyniki parametrów charakterystyki ROC dla oceny głosów nieprawidłowych za pomocą współczynników MFCC otrzymali Godino-Llorente i wsp. (16,17). Autorzy korzystali z bazy komercyjnych próbek ludzkich głosów patologicznych i prawidłowych.

wych, opublikowanych na stronie MEEI (Massachusetts Eye and Ear Infirmary Voice and Speech Labs – Szpital Chorób Oczu i Uszu w Massachusetts, Laboratorium Głosu i Mowy) (18). Z kolei badania przedstawiane w niniejszym artykule prowadzone były w oparciu o próbki akustyczne głosu nagrywane u pacjentek ze ściśle określonymi zmianami patologicznymi głosi.

Zaletą parametrów MFCC istotną diagnostycznie, jak podkreślają Arias-Londono i wsp. (19), jest to, że w ich obliczaniu nie jest konieczna pierwotna estymacja częstotliwości podstawowej, co często nastręcza trudności w wyznaczaniu konwencjonalnych parametrów akustycznych, takich jak jitter lub shimmer. Problemy te pojawiają się w ocenie głosów dysfonicznych, cechujących się aperiodycznością drgań fonacyjnych, i mogą przyczyniać się do błędnego wyznaczania ww. konwencjonalnych parametrów akustycznych. Arias-Londono i wsp. stwierdzają, że zwiększona masa fałdów głosowych – spowodowana nawet małymi zmianami przerostowymi głosi – implikuje zaburzenia przesunięcia brzeżnego (mucosal wave), co ma odzwierciedlenie w niskich pasmach akustycznych, dobrze charakteryzowanych przez współczynniki MFCC.

Ponadto dzięki analizie cepstralnej można wydzielić dźwięki pochodzące od tonu krtaniowego z sygnału mowy, które obejmują dłuższe wypowiedzi, np. zdania lub dłuższe fragmenty, co wynika z równania przedstawionego w niniejszym artykule w opisie metod. Jest to możliwe również dla dźwięków aperiodycznych. W opisywanym modelu nieliniowej analizy akustycznej zastosowanie banku filtrów melowych skutkuje transformacją wyznaczanych ilościowych parametrów w sposób uwzględniający właściwości percepcji dźwięków przez system słuchowy człowieka. Ma to istotne znaczenie w ocenie jakości głosu. Dzięki takiemu podejściu można jeszcze dokładniej w rejestrowanych sygnałach mowy wydzielić współczynniki opisujące charakterystyki dźwięków tworzonych w trakcie głosowym człowieka, w tym także dźwięków nieperiodycznych. Wagę tego problemu podkreślają też inni badacze (20).

Reasumując, nieliniowa analiza akustyczna z zastosowaniem współczynników MFCC okazała się relatywnie rzetelną metodą, która pozwala na skuteczne różnicowanie głosów patologicznych i prawidłowych. Metoda ta cechuje się wysokim poziomem czułości i swoistości, przekraczającym 85% dla wszystkich rodzajów badanych próbek głosowych po zastosowaniu proponowanych klasyfikatorów. Z tego powodu metoda może być przydatnym obiektywnym narzędziem

charakteryzującym nawet dyskretne zmiany patologiczne głosi, w tym łagodne zmiany przerostowe fałdów głosowych o podłożu zawodowym. Prowadzone są dalsze badania w tym kierunku.

PIŚMIENNICTWO

1. Dejonckere P.H., Crevier-Buchman L., Marie J.P., Moerman M., Remacle M., Woisard V.: Implementation of the European Laryngological Society (ELS) basic protocol for assessing voice treatment effect. *Rev. Laryngol. Otol. Rhinol.* 2003;124:279–283
2. Maciel C.D., Pereira J.S.: Identifying healthy and pathologically affected voice signals. *IEEE Signal Processing Mag.* 2010;27(1):120–123
3. Werth K., Voigt D., Dollinger M., Eysholdt U., Lohscheller J.: Clinical value of acoustic voice measures: a retrospective study. *Eur. Arch. Otorhinolaryngol.* 2010;267:1261–1271
4. Kumar B.R., Bhat J.S., Prasad N.: Cepstral analysis of voice in persons with vocal nodules. *J. Voice* 2010;24(6):651–653
5. Dejonckere P.H., Wieneke G.H.: Spectral, cepstral and aperiodicity characteristics of pathological voices before and after phonosurgical treatment. *Clin. Linguist. Phon.* 1994;8(2):161–169
6. Titze L.R., Baken R.J., Herzel H. Evidence of chaos in vocal fold vibration. W: Titze L.R. [red.]. In *Frontiers in basic science*. CA: Singular Publishing Group, San Diego 1993, ss. 143–188
7. Maryn Y., Roy N., De Bodt M., Van Cauwenberge P., Corthals P.: Acoustic measurement of overall voice quality: A meta-analysis. *J. Acoust. Soc. Am.* 2009;126(5):2619–2634
8. Maryn Y., Corthals P., Van Cauwenberge P., Roy N., De Bodt M.: Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *J. Voice* 2010;24(5):540–555
9. Niebudek-Bogusz E.: Zastosowanie analizy akustycznej głosu z próbą obciążeniową w dysfoniach zawodowych [rozprawa habilitacyjna]. Instytut Medycyny Pracy, Łódź 2010
10. Rubin J.S., Satalof R.T., Korovinn G.S. [red.]: *Diagnosis and treatment of voice disorders*. Thomson Delmar Learning, New York 2003
11. Niebudek-Bogusz E., Kotyło P., Politański P., Śliwińska-Kowalska M.: Acoustic analysis with vocal loading test in occupational voice disorders: outcomes before and after voice therapy. *Int. J. Occup. Med. Environ. Health* 2008;21(4):301–308. DOI: 10.2478/v10001-008-0033-9

12. Osowski S.: Sieci neuronowe w ujęciu algorytmicznym. WNT, Warszawa 1996
13. Theodoridis S., Koutroumbas K.: Pattern Recognition. Academic Press, London 2009
14. Awan S., Roy N., Jette M., Meltzner G., Hillman R.: Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: Comparisons with auditory-perceptual judgements from the CAPE-V. Clin. Linguist. Phon. 2010;24(9):742–758
15. Mehta D.D., Deliyiski D.D., Zeitels S.M., Quatieri T.F., Hillman E.R.: Voice production mechanisms following phonosurgical treatment of early glottic cancer. Ann. Otol. Rhinol. Laryngol. 2010;119(1):1–9
16. Godino-Llorente J.I., Fraile R., Saenz-Lechon N., Osma-Ruiz V., Gomez-Vilda P.: Automatic detection of voice impairments from text-dependent running speech. Biomed. Signal Processing Control 2009;4:176–182
17. Godino-Llorente J.I., Gomez-Vilda P., Cruz-Roldan F., Blanco-Velasco M., Fraile R.: Pathological likelihood index as a measurement of the degree of voice normality and perceived hoarseness. J. Voice 2009;24(6):667–677
18. Massachusetts Eye and Ear: The voice and speech lab [cytowany 9 września 2012]. Adres: <http://www.masseyeandear.org/specialties/otolaryngology/voice-laboratory>
19. Arias-Londofio J.D., Godino-Llorente J.I., Markaki M., Stylianou Y.: On combining information from modulation spectra and mel-frequency cepstral coefficients for automatic detection of pathological voices. Logoped. Phoniatr. Vocol. 2011;36(2):60–69
20. Saenz-Lechon N., Fraile R., Godino-Llorente J.I., Fernandez-Baillo R., Osma-Ruiz V., Arias-Londofio J.D.: Towards objective evaluation of perceived roughness and breathiness: An approach based on mel-frequency cepstral analysis. Logoped. Phoniatr. Vocol. 2011;36(2):52–59