

Ocena funkcji fonacyjnej krtani z zastosowaniem komputerowej analizy obrazów laryngowideostroboskopowych – badania pilotażowe

Assessment of vocal folds phonation by means of computer analysis of laryngovideostroboscopic images – a pilot study

BARTOSZ KOPCZYŃSKI^{1/}, PAWEŁ STRUMIĘŁO^{1/}, EWA NIEBUDEK-BOGUSZ^{2/}

^{1/} Instytut Elektroniki, Politechnika Łódzka

^{2/} Klinika Audiologii i Foniatrii, Instytut Medycyny Pracy im. prof. J. Nofera w Łodzi

Wprowadzenie. Komputerowe techniki analizy obrazów umożliwiają wprowadzenie nowych metod obrazów głośni podczas fonacji oraz wyznaczenie obiektywnych parametrów oceny drgań fałdów głosowych, wspomagających lekarza laryngologa/foniatrę w bardziej precyzyjnej diagnostyce narządu głosu.

Cel pracy. Zastosowanie algorytmów analizy obrazów do jakościowego i ilościowego opisu drgań fonacyjnych fałdów głosowych.

Materiał i metody. Badania wideostroboskopowe głośni przeprowadzono u 15 osób: 5 pacjentów ze stwierdzonymi guzkami głosowymi, 5 pacjentów z niedomykalnością głośni oraz 5 osób z głosem prawidłowym. Zastosowano algorytmy cyfrowego przetwarzania oraz segmentacji obrazów. Wyznaczono sygnały pola światła głośni dla kolejnych cykli fonacji oraz zbudowano glottovibrogramy stanowiące przestrzenno-czasowe zobrazowanie drgań fałdów głosowych.

Wyniki. Wyznaczono parametry geometryczne światła głośni dla każdego obrazu sekwencji wideostroboskopowej. Obliczono uśrednione profile szerokości światła głośni w fazie zamknięcia cyklu fonacyjnego dla poszczególnych grup badanych pacjentów.

Wnioski. W pilotażowych badaniach pacjentów potwierdzono przydatność opracowanych metod analizy obrazów w precyzyjnym obrazowaniu i ocenie ilościowej drgań fonacyjnych fałdów głosowych na podstawie filmów wideostroboskopowych.

Słowa kluczowe: zaburzenia głosu, diagnostyka obrazowa głośni, komputerowa analiza obrazów medycznych, wideostroboskopia

Introduction. Medical imaging techniques enable determination of novel visualisation modalities of the vocal folds during phonation and definition of parameters that can aid the otolaryngologist/phoniatrician in a more precise diagnosis of voice disorders.

Aim. Application of computer vision algorithms for qualitative and quantitative analysis of vocal-folds phonation vibrations.

Materials and methods. Videostroboscopic examinations of the glottis were carried out for 15 individuals divided into 3 groups including five subjects each: with diagnosed nodules, with glottal insufficiency, and with no voice disorders. Image pre-processing and image segmentation algorithms were applied. Signals of the glottis area for consecutive phonation cycles were derived. Glottovibrograms were also built which facilitate spatio-temporal visualisation of the vibrating vocal folds.

Results. The geometric parameters of the glottis area for each image in the stroboscopic video have been determined. The average width profiles of the glottis area for the closure phase of the glottal cycle have been computed for each group of the examined patients.

Conclusions. The conducted pilot study has confirmed that computer aided imaging methods could be applied in the qualitative and quantitative analysis of the videostroboscopic images showing the phonatory motions of the vocal folds.

Key words: voice disorders, medical imaging of glottis, medical image computing, videostroboscopy

Wykaz skrótów:ELS – *European Laryngological Society*

LWSS – laryngowideostroboskopia

ROI – *Region of Interest*OpenCV – *Open Source Computer Vision***WSTĘP**

Wraz ze wzrostem znaczenia procesu komunikacji wczesna diagnostyka zawodowych zaburzeń głosu staje się jednym z priorytetów zdrowia publicznego zarówno w Polsce, jak i w innych krajach Unii Europejskiej. Obecnie standardy europejskie podkreślają konieczność kompleksowej oceny zaburzeń głosu z uwzględnieniem oceny funkcjonalnej krtani, dokonywanej podczas badania laryngologicznego lub foniatrycznego [1-4]. Zgodnie z protokołem opracowanym przez Komitet Foniatrii Europejskiego Towarzystwa Laryngologicznego (ELS) w 2000 roku wśród 5 podstawowych narzędzi badawczych, stosowanych w diagnostyce zaburzeń głosu, znalazły się następujące metody: 1) ocena percepcyjna głosu, 2) samoocena głosu, 3) wideolaryngostroboskopia, 4) analiza akustyczna głosu, 5) ocena aerodynamiczna [5].

Podstawowym badaniem specjalistycznym pozwalającym precyzyjnie ocenić stan narządu głosu, uznawanym jako złoty standard, jest laryngowideostroboskopia (LWSS), w której dzięki zastosowanemu efektowi stroboskopowemu można ocenić drgania fonacyjne fałdów głosowych [6]. Laryngowideostroboskopia składa się z laryngoskopii lupowej i stroboskopii. Zjawisko stroboskopii polega na optycznym złudzeniu wynikającym z bezwładności wzroku, tj. istniejącego opóźnienia pomiędzy chwilą podania bodźca świetlnego a wytworzeniem wrażenia wzrokowego. Każdy bodziec świetlny docierający do siatkówki pozostawia wrażenie wzrokowe trwające około 200 milisekund (ms), czyli oko ludzkie jest w stanie rozróżnić do 5 drgań w ciągu 1 sekundy (prawo Talbota). Pojedyncze bodźce wzrokowe występujące z przerwami krótszymi niż 200 ms nie mogą być rozróżniane jako oddzielne wrażenia. Zatem drgania fałdów głosowych są za szybkie, aby mogły być postrzegane jako osobne bodźce: fałdy głosowe podczas fonacji drgają przeciętnie u kobiet z częstotliwością 200-250 Hz, a u mężczyzn z częstotliwością 100-150 Hz [6]. Z tego powodu nie można drgań fonacyjnych zaobserwować podczas rutynowej laryngoskopii. Jeżeli jednak oświetlimy krtani światłem stroboskopowym (migającym) o częstotliwości drgań różnej od częstości drgań fałdów głosowych, to zaobserwujemy ruchy fałdów głosowych w zwolnionym tempie, tzw. zwolniony

ruch fonacyjny głośni (slow motion). Za najbardziej skuteczną uznaje się sekwencję obrazów stroboskopowych drgań fonacyjnych uzyskiwaną w świetle stroboskopowym, gdy różnica między częstotliwością migotania światła a drganiem fałdów głosowych wynosi ok. 1 Hz. Koniecznym warunkiem poprawności badania jest m.in. niezakłócona praca mikrofonu rejestrującego dźwięk fonacyjny pacjenta. Rejestrowany dźwięk jest poddawany analizie w czasie rzeczywistym, umożliwiając dostosowanie częstotliwości migotania światła stroboskopowego synchronicznie z częstotliwością podstawową drgań fonacyjnych fałdów. Synchronizacja częstotliwości migotania światła stroboskopu umożliwia prawidłową rejestrację obrazów wideo poruszających się fałdów głosowych w czasie fonacji. Sekwencje obrazów laryngowideostroboskopowych są rejestrowane na komputerowych nośnikach danych. Odtwarzanie zapisu wideo czynności fonacyjnej krtani umożliwia precyzyjną diagnostykę zaburzeń głosu, gdyż prawidłowe vibracje fałdów głosowych odgrywają kluczową rolę w procesie fonacji. Podstawową wadą tej powszechnie stosowanej w praktyce klinicznej metody instrumentalnej jest brak obiektywizacji jej wyników, gdyż ocena dokonywana jest jakościowo, zależy od kompetencji i doświadczenia osoby przeprowadzającej badanie.

Ostatnio podkreśla się konieczność ilościowego opisu rejestrowanych obrazów drgań fonacyjnych fałdów głosowych [7-11], co może służyć obiektywizacji tej metody – istotnej w diagnostyce zmian patologicznych krtani. Obserwuje się zwiększoną liczbę doniesień naukowych, w których stosuje się zaawansowane metody cyfrowego przetwarzania i analizy obrazów krtani [12], do ich ilościowej analizy [13] oraz wizualizacji z wykorzystaniem wykresów przestrzenno-czasowych w postaci obrazów kimograficznych [14] i glottowibrogramów [12].

W niniejszej pracy podjęto badania pilotażowe, ukierunkowane na opracowanie i testy algorytmów komputerowej segmentacji i analizy obrazów wideostroboskopowych do jakościowego i ilościowego opisu drgań fonacyjnych fałdów głosowych rejestrowanych u osób z prawidłowym głosem oraz ze stwierdzoną patologią głośni.

MATERIAŁ I METODY

Filmy obrazujące funkcję fonacyjną krtani nagrano z zastosowaniem techniki wideostroboskopowej w Klinice Audiologii i Foniatrii Instytutu Medycyny Pracy w Łodzi. Badanie wideostroboskopowe wykonywano za pomocą wideostroboskopu firmy Xion EndoStrob DX. Do analizy komputerowej wybrano sekwencje wideostroboskopowe zareje-

strowane u 15 osób: 5 – ze stwierdzonymi guzkami głosowymi (średni wiek – 40 lat), 5 – z niedomykalnością fonacyjną głośni (średni wiek – 49 lat) oraz 5 – z prawidłowym głosem i prawidłową funkcją fonacyjną głośni (średni wiek – 25 lat). Rejestracje wideostroboskopowe były wykonywane w czasie przedłużonej fonacji głoski „a” i zapisywane na dysku komputera w zestawie firmy Xion. Przykładowy obraz głośni zarejestrowany w badaniu wideostroboskopowym pokazano na rycinie 1.



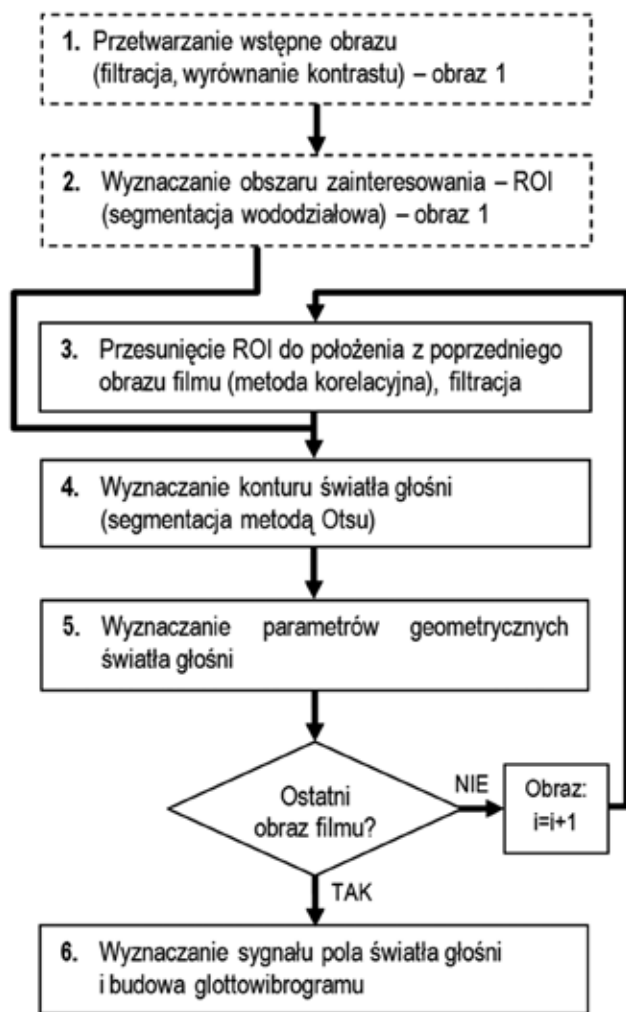
Ryc. 1. Obraz głośni w badaniu wideostroboskopowym – widoczny brak zwarcia fonacyjnego na całej długości fałdów głosowych

Opracowano komputerowe algorytmy przetwarzania, segmentacji i analizy obrazów głośni, które umożliwią automatyczną parametryzację światła głośni w czasie fonacji. Należy podkreślić, że wideostroboskopowe obrazy głośni są trudnym materiałem badawczym do analizy ilościowej. Głośnia w trakcie badania podlega ciągłemu przemieszczaniu się względem kamery wideostroboskopu. Powoduje to nieustanną zmianę położenia głośni w kadrze obrazu, zmianę kontrastu kolejnych obrazów, zmianę wzajemnej odległości kamery od głośni oraz kierunku jej obserwacji przez kamerę. Takie warunki rejestracji wideostroboskopowej znacząco utrudniają automatyczną, ilościową analizę obrazów głośni w czasie fonacji. W zastosowanym podczas badań pilotażowych zestawie algorytmów wzięto pod uwagę ww. niedogodności i starano się zniwelować ich wpływ na końcowe wyniki.

W badaniach zaimplementowano algorytmy w postaci programów napisanych w językach Python, C++ z wykorzystaniem biblioteki funkcji analizy obrazów OpenCV [15]. Ogólny schemat blokowy opracowanych algorytmów pokazano na rycinie 2. Algorytmy zaznaczone w blokach wykreślonych linią przerywaną (algorytmy 1 i 2) są wykonywane tylko jednokrotnie dla pierwszego obrazu

z sekwencji wideostroboskopowej. Celem tych dwóch pierwszych kroków przetwarzania obrazu jest poprawa jego jakości. Zastosowano w pierwszej kolejności filtrację cyfrową obrazu w celu redukcji zakłóceń zawartych w obrazie, a następnie tzw. modelowanie histogramu obrazu do poprawy kontrastu obrazu [16]. Te kroki przetwarzania wstępnego są konieczne do skutecznego przeprowadzenia tzw. segmentacji obrazu, tj. działania polegającego na automatycznym wydzieleniu w obrazie obszaru tzw. obszaru zainteresowania (*Region of Interest, ROI*) zawierającego obraz głośni. Zastosowano, podobnie jak w pracy [17], transformację wododziałową obrazu uważaną za jedną z najbardziej skutecznych metod segmentacji obrazu. Działanie tej metody polega na podziale obrazu gradientowego głośni wg kryterium jasności, która jest interpretowana jako wysokość hipotetycznej powierzchni odwzorowującej jasność obrazu. Obszary o zadanych jasnościach (hipotetycznych wysokościach) są przydzielane do tego samego wododziału i traktowane jako spójny obszar. Zastosowanie tego algorytmu do obrazu wideostroboskopowego pozwala wyznaczyć zgrubnie światło głośni. Wynik działania tej metody segmentacji pokazano na rycinie 3 (obszar ROI zaznaczony jednolitą jasnością). Wadą tej metody jest jednak mała dokładność wyznaczania brzegów obszarów oraz wykryte inne obszary niestanowiące światła głośni. Konieczne jest zatem zastosowanie kolejnych metod segmentacji obrazu pokazanych na schemacie blokowym z ryciny 2. Metody te są poprzedzone procedurą przemieszczenia cyfrowego obszarów głośni w każdym kolejnym obrazie sekwencji wideo do położenia światła głośni wykrytego w pierwszym obrazie (algorytm 3, szerzej opisany w pracy [18]). Procedura ta umożliwiła zbudowanie filmu obrazującego fonację fałdów, w którym obszar głośni pozostaje w stałym położeniu w kadrze obrazu. Dzięki takiemu zabiegowi można precyzyjniej śledzić cykliczny ruch fałdów głosowych w sekwencji wideostroboskopowej. Przykładowy film, po zastosowaniu wyżej opisanej procedury, udostępniono pod łączem internetowych www.eletel.p.lodz.pl/obrazowanie_faldow. W algorytmie 3 schematu blokowego każdy obraz wideostroboskopu jest również poddawany przetwarzaniu wstępnemu jak w kroku 1 (ryc. 2).

Kolejnym krokiem przetwarzania obrazu jest algorytm 4 służący do dokładnego wyznaczenia konturu światła głośni. Do tego zadania segmentacji obrazu zastosowano tzw. metodę Otsu [19] opublikowaną w roku 1979 i do dzisiaj uważaną za jedną z najbardziej skutecznych metod segmentacji obrazów. W metodzie tej jako kryterium segmentacji jest wykorzystana informacja o rozkładzie staty-



Ryc. 2. Schemat blokowy pokazujący kolejność zastosowanych algorytmów przetwarzania, segmentacji i analizy obrazów wideostroboskopowych (bloki oznaczone linią przerywaną są wykonywane tylko dla pierwszego obrazu sekwencji wideostroboskopowej)

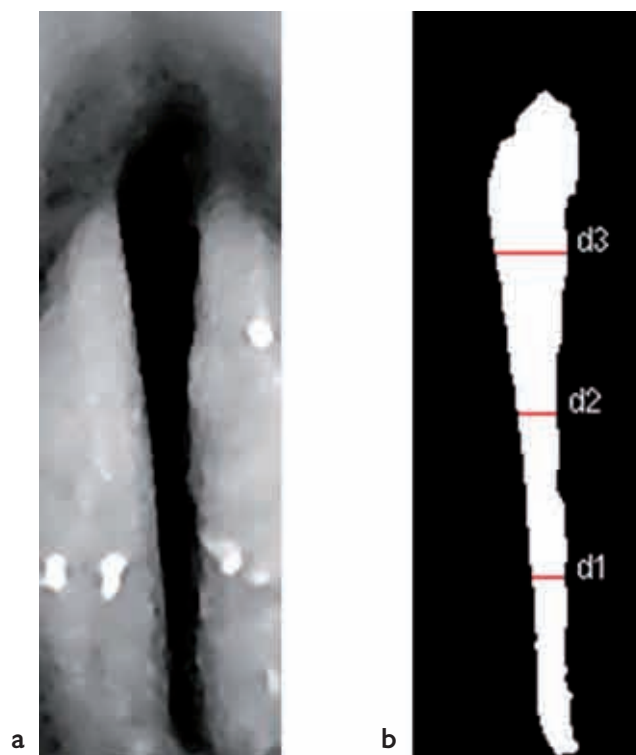


Ryc. 3. Wynik segmentacji wododziałowej obrazu krtani z ryc. 1; widoczny jest obszar głośni wyznaczony w sposób przybliżony oraz drugi obszar stanowiący artefakt (widoczny w prawym górnym rogu obrazu)

stycznym jasności punktów obrazu w obszarze zainteresowania (ROI). W metodzie tej minimalizowane są wewnątrzklasowe wariancje jasności punktów obrazu w segmentowanych obszarach. Wynik zastosowania tej metody segmentacji do wyznaczenia pola światła głośni pokazano na rycinie 4a.

Podstawowe parametry geometryczne jakie obliczono dla wyznaczonego pola światła głośni to: długość liczona jako najdłuższa przekątna światła głośni i przekroje wzdłuż linii prostopadłych do tej przekątnej (algorytm 5 z ryc. 2). Na rycinie 4b wyróżniono przekroje dla odpowiednio d1 – 25%, d2 – 50% i d3 – 75% długości głośni wyznaczone dla obrazów wideostroboskopowych rejestrowanych u osób z głosem prawidłowym oraz pacjentów ze stwierdzonymi schorzeniami narządu głosu (niedomykalność głośni oraz guzki głosowe).

Algorytmy 3-5 z ryciny 2 są stosowane do wszystkich kolejnych obrazów sekwencji wideostroboskopowych. Jeden cykl fonacyjny jest obrazowany w ok. 20 do 30 obrazów. Analizowanych jest średnio ok. 4 cykli, zatem analizowane filmy wideostroboskopowe zawierają średnio ok. 100 obrazów (najkrótszy składa się z 80 obrazów a najdłuższy z 200). Filmy są nagrywane z częstotliwością próbkowania 25 obrazów na sekundę, zatem średni czas trwania filmu wynosi ok. 4 sekund.

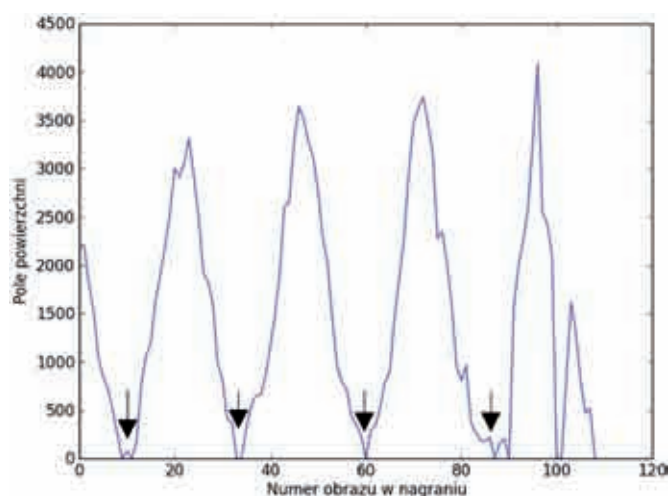


Ryc. 4. Obraz głośni pokazany w powiększeniu po korekcji kontrastu i filtracji (a) oraz pole światła krtani wyznaczone metodą segmentacji Otsu (b); symbole d1, d2 d3 oznaczają przekroje światła głośni dla odpowiednio 25%, 50% i 75% jej długości

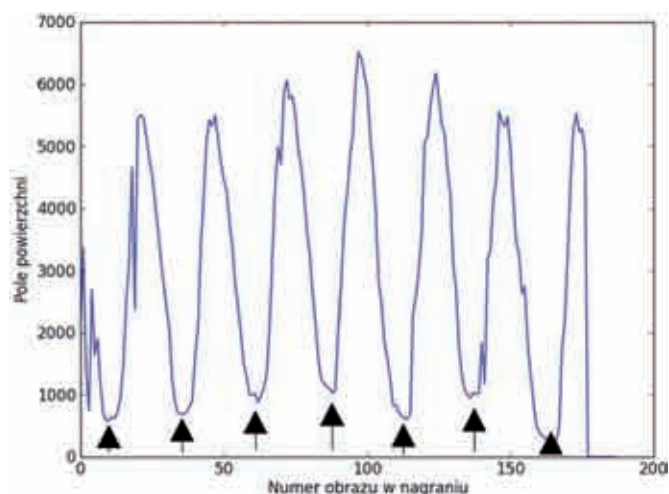
WYNIKI

Dla każdego obrazu filmu wyznaczono parametry geometryczne światła głośni (pole powierzchni oraz profile szerokości światła głośni na różnych jej poziomach), które wykorzystano do wykreślenia sygnałów pola światła głośni oraz wyznaczenia glottowibrogramów.

Sygnał pola światła głośni definiujemy jako przebieg zmian pola powierzchni światła głośni dla kolejnych obrazów wideostroboskopowych. Przykładowe wykresy takich sygnałów pokazano na rycinach 5 i 6. Sygnał pola światła głośni uzyskany dla osoby z głosem prawidłowym obrazuje rycina 5. Na podstawie tego wykresu można wnioskować



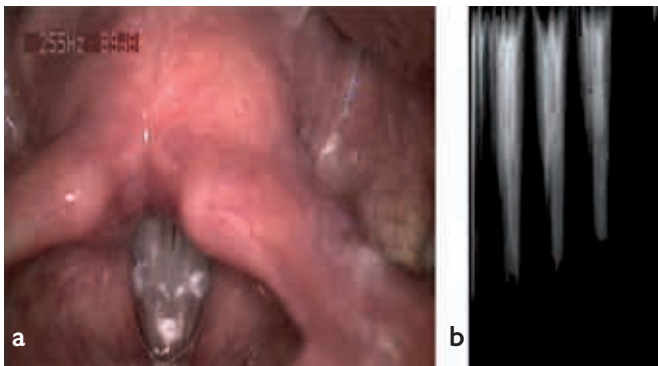
Ryc. 5. Sygnał pola światła głośni wyznaczony dla osoby z głosem prawidłowym (pole powierzchni wyrażono w liczbie pikseli obrazu); amplitudy sygnału osiągają wartości zerowe w fazie zamknięcia (wskazane strzałkami)



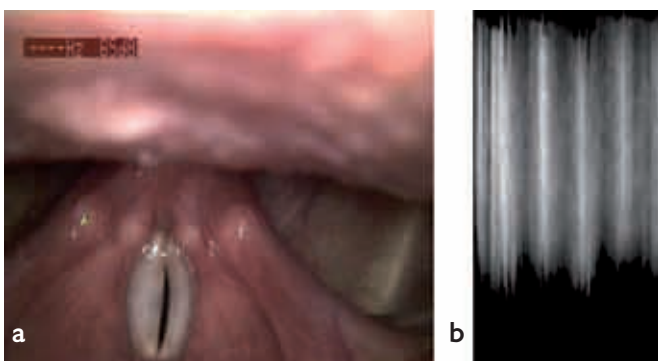
Ryc. 6. Sygnał pola światła głośni wyznaczony dla osoby ze stwierdzoną niedomykalnością głośni (pole powierzchni wyrażono w liczbie pikseli obrazu); minima amplitud sygnału w fazie zamknięcia przyjmują znaczne wartości świadczące o braku zwarcia fonacyjnego fałdów głosowych (wskazane strzałkami)

o pełnym zamknięciu fałdów głosowych (pole przyjmuje wartości zerowe). Na rycinie 6 wykreślono z kolei sygnał pola światła głośni uzyskany dla pacjenta ze stwierdzoną niedomykalnością fałdów głosowych. Na wykresach tych można zauważyć, że minima amplitud nie osiągają wartości zerowych. Wartości pól powierzchni światła głośni wyrażono w liczbie pikseli, tj. liczbie punktów obrazu.

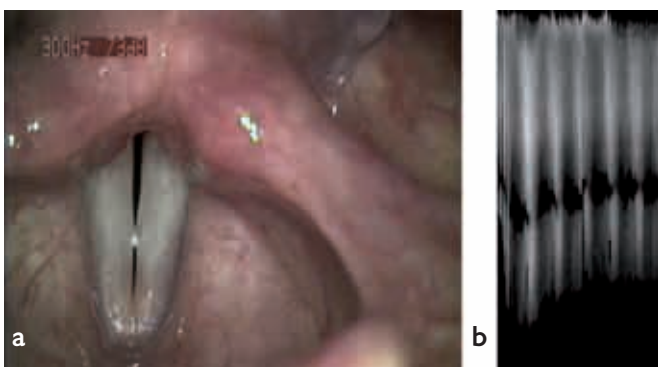
Glottowibrogramy są przestrzenno-czasowym zobrazowaniem ruchów fonacyjnych fałdów głosowych. Zawierają one informację o szerokości rozwarcia fałdów głosowych dla kolejnych przecięć poprzecznych fałdów (stanowiących wiersze glottowibrogramu) i dla kolejnych faz cyklu fonacji (kolumny obrazu glottowibrogramu). Jasność punktu glottowibrogramu należy interpretować jako stopień rozwarcia fałdów głosowych w cyklu fonacyjnym, tj. im jaśniejszy punkt tym większe otwarcie głośni. Zatem obszary ciemne glottowibrogramu charakteryzują fazę zamknięcia głośni, a obszary jasne fazę jej otwarcia. Na rycinach 7-9 pokazano obrazy wideostroboskopowe głośni dla fazy zamknięcia fałdów oraz glottowibrogramy fonacji fałdów odpowiednio dla osoby z głosem prawidłowym, pacjenta ze stwierdzoną niedomykalnością głośni oraz pacjenta z guzkami głosowymi. Należy podkreślić, że wzrokowa ocena glottowibrogramów pozwala tylko na jakościową ocenę funkcji fonacyjnej. W glottowibrogramie z ryciny 7b, wyznaczonym dla osoby z prawidłowym głosem, występują naprzemienne, regularne fazy otwarcia (obszary jasne) i zamknięcia głośni (obszary ciemne). Przy czym obszar otwarcia zwęża się ku dołowi wykresu, co świadczy o najpełniejszym zwarciu fonacyjnym głośni w części przedniej. Glottowibrogram z ryciny 8b to obraz fonacji dla przypadku niedomykalności głośni z powodu astenii mięśni wewnętrznych krtani, szczególnie w fałdzie głosowym lewym. Widać tylko niewielkie zmiany w jasności glottowibrogramu świadczące o istotnej przewadze fazy otwarcia w cyklu fonacyjnym, czyli praktycznie brak jest zwarcia fonacyjnego na całej długości głośni – konfiguracja jej niedomykalności o kształcie wrzeciona. Rycina 9b obrazuje glottowibrogram głośni pacjenta ze stwierdzonymi guzkami głosowymi. Wyraźnie widać zwężenie światła głośni na ok. 1/3 jej długości, czyli w miejscu występowania guzków, co wskazuje, że do pełnego zamknięcia głośni w cyklu fonacyjnym dochodzi tylko w miejscu guzków. Jest to charakterystyczna dla tej patologii niedomykalność głośni w kształcie klepsydry. Zmiana położenia zwężenia głośni w kolejnych cyklach fonacji wynika z przemieszczania się wideostroboskopu względem obrazowanego pola głośni (w związku z minimalnymi ruchami ręki badającego).



Ryc. 7. Obraz wideostroboskopowy w fazie zamknięcia głośnia (a) oraz glottowibrogram (b) obrazujący ruchy fonacyjne fałdów głosowych u osoby z prawidłowym głosem



Ryc. 8. Obraz wideostroboskopowy w fazie zamknięcia głośnia (a) oraz glottowibrogram (b) obrazujący ruchy fonacyjne fałdów głosowych u osoby z niedomykalnością głośnia



Ryc. 9. Obraz wideostroboskopowy w fazie zamknięcia głośnia (a) oraz glottowibrogram (b) ilustrujący ruchy fonacyjne fałdów głosowych u osoby z guzkami głosowymi: ciemny nieregularny obszar na ok. 1/3 wysokości glottowibrogramu wskazuje zamknięcie głośnia w miejscu występowania guzków

Na podstawie informacji zawartej w obrazach glottowibrogramów wyznaczono profile szerokości światła głośnia dla wszystkich wykrytych faz zamknięcia głośnia, a następnie uśredniono wynik otrzymując profil stopnia zamknięcia. W tabeli I zawarto syntetyczne zestawienie szerokości światła

głośnia wyznaczonej w przekrojach d1, d2, d3 czyli 25%, 50% i 75% długości głośnia (zob. też ryc. 4). Odpowiednie szerokości dla d1, d2 i d3 w fazie zamknięcia wyrażono jako wartości procentowe odniesione do długości głośnia badanego pacjenta. Taki wybór miary dla szerokości światła głośnia pozwala uniezależnić jej wartości od skali obrazów wideostroboskopowych krtani, która w czasie rejestracji podlega ciągłym zmianom podczas nawet minimalnych ruchów. Należy zaznaczyć, że takie rozwiązanie należy traktować jako pewne przybliżenie. Nie we wszystkich obrazach widoczna jest cała głośnia. Również jej długość podlega zmianom zależnie od rodzaju fonacji np. wysokości tonu.

Tabela I. Średnie wartości szerokości światła głośnia wyznaczone oddzielnie dla poszczególnych grup badanych wyrażone jako wartości procentowe długości światła głośnia

Diagnoza	Szerokości światła głośnia		
	d1	d2	d3
Głos prawidłowy	0,1%	0,6%	1,2%
Niedomykalność głośnia	4,3%	4,6%	5,3%
Guzki głosowe	2,0%	0,2%	4,2%

DYSKUSJA

Obrazy wideostroboskopowe krtani umożliwiają dokładną obserwację funkcji głośnia w trakcie drgań fonacyjnych; zastosowanie coraz to bardziej zaawansowanych technik przetwarzania cyfrowego obrazu wideo daje możliwość zobiektywizowania wyników wideostroboskopowych. W niniejszej pracy opisano badania pilotażowe, których celem jest opracowanie komputerowych algorytmów przetwarzania, segmentacji i analizy obrazów głośnia. Przedstawiane wyniki badań są przykładem wykorzystania nowych technik analizy obrazów do jakościowego i ilościowego opisu drgań fonacyjnych fałdów głosowych na podstawie zapisów badania wideostroboskopowego. Za pomocą zastosowanych algorytmów cyfrowego przetwarzania oraz segmentacji obrazów dokonano automatycznej parametryzacji światła głośnia w czasie fonacji z wyznaczeniem parametrów geometrycznych światła głośnia dla każdego obrazu sekwencji wideostroboskopowej.

Należy podkreślić, że konfiguracja zamknięcia fonacyjnego głośnia jest jednym z parametrów oceny badania stroboskopowego obok amplitudy drgań fonacyjnych, ich regularności, różnicy faz oraz przesunięcia brzeźnego, zwanego też falą śluzówkową (mucosal wave), które jest efektem przemieszczania błony śluzowej względem mięśnia głosowego [1]. Podczas fonacji od strony okolicy podgłośniowej napiera na zwartą głośnień ciśnienie podgłośniowe, któremu przeciwstawiają się masa i napięcie fałdów

głosowych. Przewaga ciśnienia podgłośniowego doprowadza do cyklicznego rozwierania i zwierania krawędzi fałdów głosowych w trakcie drgań fonacyjnych.

Cykl drgań fonacyjnych obejmuje fazę otwierania, otwarcia, zamykania i zamknięcia głośni. W fazie otwarcia powietrze z okolicy podgłośniowej przeciska się przez głośnię i generuje powstawanie ww. fali śluzówkowej. Faza zamykania zaczyna się od stopniowego spadku ciśnienia podgłośniowego spowodowanego wciskaniem się powietrza do szpary głośni podczas fazy otwarcia. Natomiast naturalne siły elastyczne fałdów głosowych zaczynają zasysać fałdy głosowe do siebie, począwszy od krawędzi dolnych [4]. Przewaga efektu Bernoulliego oraz sił elastycznych powodują w tym etapie cyklu również zwieranie górnych krawędzi fałdów głosowych, co prowadzi do zamknięcia głośni. Powszechnie uważa się, że w głosach prawidłowych zwarcie fonacyjne powinno być pełne na całej długości głośni. Natomiast Woo wyróżnia następujące nieprawidłowe konfiguracje zamknięcia szpary głośni [6]:

1. Niedomykalność w kształcie klepsydry, charakterystyczna dla guzków głosowych – do zwarcia fonacyjnego dochodzi tylko w miejscu guzków (taka konfiguracja głośni prezentowana jest w pracy na rycinie 8 przedstawiającej obraz wideo-stroboskopowy i utworzony na jego podstawie glottowibrogram).
2. Niedomykalność w części międzybłoniastej głośni o kształcie wrzeciona: „łukowate”, wiotkie fałdy głosowe zwierają się tylko w przednim i tylnym spoidle (taki obraz głośni przedstawia ryc. 9).
3. Całkowity brak zwarcia fonacyjnego głośni na całej jej długości.
4. Niedomykalność w części przedniej.
5. Niedomykalność w części tylnej głośni.

Tę ostatnią foniatry skandynawscy: Södersten i wsp. [20] określają jako nie odbiegającą od normy u kobiet. Taki kształt głośni widoczny jest na ryc. 7 przedstawiającej obraz wideo-stroboskopowy i glottowibrogram badanej kobiety z głosem prawidłowym, u której widoczna niedomykalność w części tylnej głośni nie miała wpływu na jakość głosu.

W prowadzonych badaniach pilotażowych zanotowano dla osób z głosem prawidłowym najmniejsze uśrednione wartości (d) szerokości światła w fazie zamknięcia cyklu fonacyjnego (tab. I); z tym że najmniejsze wartości d występowały na poziomie

25% długości głośni: $d_1=0,2$, największe wartości na poziomie 75% $d_3=1,2$. Wyniki te potwierdzają obiektywnie wyżej opisaną obserwację autorów skandynawskich donoszącą, że u osób z głosem prawidłowym może występować niewielka niedomykalność w tylnej części głośni, co nie zaburza funkcji fonacyjnej krtani. Natomiast u osób z niedomykalnością głośni daje się zauważyć wzdłuż całej długości głośni brak jej zamknięcia w cyklu fonacyjnym: wskaźniki szerokości głośni dla fazy zamknięcia na poziomie d_1 , d_2 i d_3 przekraczają wartość 4% i znacząco różnią się od wartości wyznaczonych dla osób z głosem prawidłowym, ale także dla grupy osób z guzkami głosowymi. Z kolei w tej ostatniej grupie badanej można zauważyć charakterystyczny dla guzków kształt klepsydry profilu światła głośni w fazie zamknięcia, w którym pełne zamknięcie głośni występuje konsekwentnie dla przekroju 50% długości głośni – d_2 wynosi 0,2, co potwierdza obserwowane klinicznie najpełniejsze zwarcie fonacyjne w miejscu guzków.

Reasumując wdrożenie opisanych technik analizy obrazów do praktyki medycznej może przyczynić się do obiektywizacji wyników badań laryngostroboskopowych; konieczna jest dalsza współpraca lekarzy foniatorów/laryngologów oraz programistów, w celu opracowania interfejsu graficznego dostosowanego do wymagań klinicznych.

WNIOSKI

1. Metody komputerowej analizy obrazów wideo-stroboskopowych umożliwiają automatyczne wyznaczanie nowych zobrazowań fałdów głosowych podczas fonacji.
2. Glottowibrogramy oraz sygnały pola światła głośni umożliwiają odwzorowanie kinematyki fałdów głosowych podczas fonacji (zakres ruchów fałdów, czas faz zamykania i otwierania głośni).
3. Profile światła głośni umożliwiają ilościową ocenę stopnia zwarcia fonacyjnego fałdów głosowych na całej długości głośni.

Oświadczamy, że w czasie badań były przestrzegane prawa pacjentów zgodnie z zasadami zawartymi w Konwencji Helsińskiej.

Praca przygotowana w ramach projektu statutowego IMP 18.12/2014 oraz w ramach działalności statutowej Instytutu Elektroniki Politechniki Łódzkiej.

Piśmiennictwo

1. Rubin JS, Satalof RT, Korovin GS. Diagnosis and treatment of voice disorders. Thomson Delmar Learning, New York 2003.
2. Schutte HK. Proceedings of 7th Pan European Voice Conference Aug 29-Sept 1 2007, Groningen, The Netherlands. Voice Research Lab, Groningen 2007.
3. De Jong F. An introduction to the teacher's voice in a biopsychosocial perspective. *Folia Phoniatri Logop.* 2010; 62: 5-8.
4. Obrębowski A. Narząd głosu i jego znaczenie w komunikacji społecznej. Poznań, 2008.
5. Dejonckere PH, Crevier-Buchman L, Marie JP, Moerman M, Remacle M, Woisard V. Implementation of the European Laryngological Society (ELS) basic protocol for assessing voice treatment effect. *Rev Laryngol Otol Rhinol (Bord)* 2003; 124(5): 279-83.
6. Woo P. Stroboscopy. Plural Publishing, United Kingdom 2010.
7. Kelley RT, Colton RH, Casper J, Paseman A, Brewer D. Evaluation of stroboscopic signs. *J Voice* 2011; 25(4): 490-5.
8. Svec JG, Sundberg J, Hertegård S. Three registers in an untrained female singer analyzed by videokymography, strobolaryngoscopy and sound spectrography. *J Acoust Soc Am* 2008, 123(1): 347-53.
9. Krausert ChR, Olszewski AE, Taylor LN, McMurray JS, Dailey SH, Jiang JJ. Mucosal wave measurement and visualization techniques. *J Voice* 2011; 25(4): 395-405.
10. Bonilha HS, Deliyski DD, Whiteside JP, Gerlach TT. Vocal fold phase asymmetries in patients with voice disorders: a study across visualization techniques. *Am J Speech Lang Pathol* 2012; 21(1): 3-15.
11. Gaber AGH, Liang FY, Yang JS, Wang YJ, Zheng YQ. Correlation among the Dysphonia Severity Index (DSI), the RBH voice perceptual evaluation, and minimum glottal area in female patients with vocal fold nodules. *J Voice* 2011; 28(1): 20-3.
12. Karakozoglou S-Z, Nathalie Henrich, d'Alessandro C, Stylianou Y. Automatic glottal segmentation using local-based active contours and application to glottovibrography. *Speech Communication* 2012; 54(5): 641-54.
13. Tao C, Zhang Y, Jiang JJ. Extracting physiologically relevant parameters of vocal folds from high-speed video image series. *IEEE Trans Biomed Eng* 2007; 54(5): 794-801.
14. Chodara AM, Krausert CR, Jiang JJ. Kymographic characterization of vibration in human vocal folds with nodules and polyps. *The Laryngoscope* 2012; 122(1): 58-65.
15. OpenCV. Open Source Computer Vision Library, www.opencv.org (strona dostępna 20.05.2014)
16. Materka A, Strumiłło P. Wstęp do komputerowej analizy obrazów, Politechnika Łódzka, 2009.
17. Osmá-Ruiz V, Godino-Llorente JI, Saenz-Lechon N, Fraile R. Segmentation of the glottal space from laryngeal images using the watershed transform. *Comput Med Imaging Graph.* 2008; 32(3): 193-201.
18. Kopczyński B. Analiza pracy strun głosowych na podstawie filmów i obrazów kymograficznych. Praca Inżynierska, Politechnika Łódzka, Instytut Elektroniki, 2013.
19. Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Sys Man Cyber* 1979; 9(1): 62-6.
20. Södersten M, Ternstom S, Bohman M. Loud speech in environmental noise: phonetogram data, perceptual voice quality, subjective ratings, and gender differences in healthy speakers. *J Voice* 2005; 19(1): 29-46.