

2.23 — rozpoznawanie mowy

Henryk Kubzdela

BADANIA NAD UDOSKONALENIEM
SPEKTROGRAMÓW BINARNYCH

24/1983

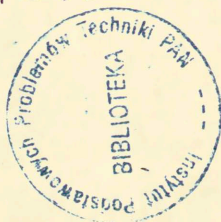
P. 269



WARSZAWA 1983

Praca wpłynęła do Redakcji dnia 22 listopada 1982 r.

57013



Na prawach rękopisu

Instytut Podstawowych Problemów Techniki PAN
Nakład 130 egz. Ark. wyd. 1 . Ark. druk. 1,75
Oddano do drukarni w maju 1983 r.
Nr zamówienia 706/83 M-13 .

Warszawska Drukarnia Naukowa, Warszawa,
ul. Sniadeckich 8

Henryk Kubzdela
Pracownia Fonetyki Akustycznej
IPPT PAN

BADANIA NAD UDOSKONALENIEM SPEKTROGRAMÓW BINARNYCH^{1/}

Streszczenie

Praca dotyczy zezegnienia spektrometrii binarnej. Zbadano trzy rodzaje przekształceń widmowych poprzedzających tworzenie widma binarnego oraz dwa warianty metody wartościowania jego parametrów. Przeprowadzono ocenę jakości spektrogramów binarnych przy zastosowaniu kombinacji poszczególnych przekształceń widmowych z każdym z wariantów metody wyznaczania spektrogramów binarnych. Podstawę oceny stanowiły kryteria jakości i podobieństwa spektrogramów binarnych wyrazów wymawianych wielokrotnie przez różne głosy. W wyniku przeprowadzonych badań wyłoniono sposób otrzymywania spektrogramów binarnych lepszy od dotychczas stosowanego w procesie automatycznego rozpoznawania wyrazów.

1. Wstęp

W Pracowni Fonetyki Akustycznej IPPT PAN w Poznaniu prowadzone są od kilku lat badania nad automatycznym rozpoznawaniem wyrazów na podstawie spektrogramów binarnych. Tok i wyniki tych badań relacjonowały prace : [3], [4], [5]. W kolejnych etapach zaopatrzone kryteria rozpoznawania w odniesieniu do rozmiaru słownika oraz liczby głosów. Dotychczasowe badania dotyczyły jednak wyłącznie rozpoznawania w oparciu o indywidualne zbiory wzorców dla poszczególnych głosów. Nie podejmowano natomiast prób rozpoznawania wyrazów w oparciu o zbiór wzorców wspólny dla pewnego ograniczonego grona głosów. Uznano bowiem za ko-

^{1/} Praca wykonana w ramach problemu międzyresortowego MR.I-24.

nieczne poprzedzenie takich prób zbadaniem rodzaju i zakresu różnic zachodzących pomiędzy obrazami binarnymi tego samego słowa wymawianego przez różne głosy. Praca niniejsza dotyczy po części tego zagadnienia. Przeglądając liczne spektrogramy binarne wykonane metodą dotychczas stosowaną zauważono następujący fakt. Kolejne parametry na spektrogramie binarnym dotyczącym głosu wyższego nie przyjmują zgodnie wartości "1" w określonych zakresach, tak jak należałoby tego oczekiwać i jak to ma miejsce w przypadku spektrogramu binarnego odnoszącego się do głosu niższego. Występujące w tych zakresach zróżnicowanie wartości kolejnych parametrów wykazuje pewną regularność, która ma swoje źródło w periodycznym charakterze dźwięku mowy. Niniejszą pracę poświęcono więc głównie zagadnieniu wyeliminowania wpływu wysokości głosu na jakość spektrogramu binarnego. Oprócz tego zbadano nową zasadę ustalania progu, względem którego dokonuje się oceny wartości poszczególnych parametrów widmowych podczas tworzenia widma binarnego. Ta nowa zasada wynika częściowo z określonych przesłanek percepcyjnych, podczas gdy zasada dotychczas stosowana nawiązywała raczej do pewnych przesłanek artykulacyjnych. Przeprowadzone w tym zakresie próby zmierzały również do poprawy jakości spektrogramów binarnych i stanowiły przyczynek do rozwiązania problemu rozpoznawania wyrazów na podstawie spektrogramów binarnych w oparciu o wspólny zbiór wzorców dla pewnego grona głosów.

2. Przekształcenia wygładzające widmo

Na rysunkach 1a i 1b przedstawiono 2 spektrogramy binarne słowa [spuwdzelna] odnoszące się odpowiednio do wypowiedzi głosem męskim M1 i głosem żeńskim Ż2. Oba spektrogramy wykonane zostały według zasad podanych w pracach [1], [2]. Parametry przypadające w zakresach formantowych powinny na spektrogramach binarnych przyjmować wartość "1". Zachodzi to konsekwentnie na rys. 1a. Na spektrogramie z rys. 1b wymaganie to nie zostaje w całości spełnione. System analogowo-cyfrowy, jakim posługiwano się w badaniach, umożliwia obecnie pokaz lub wydruk dowolnego z widm wpisanych do minikomputera z wielokanałowego analizatora analogowego i składających się na spektrogram wypowiedzi, jak również różnych form przejściowych któregośkolwiek z tych widm.

Na rys. 2a i 2b oraz 3a i 3b zamieszczono wydruki dwóch widm pochodzących z analogicznych fragmentów wypowiedzi wyrazu [spuwdzelpa] głosem M1 i głosem Z2 w takiej wersji, w jakiej wpisane one zostały do minikomputera oraz w wersji uśrednionej. Z widma w postaci uśrednionej wyznaczone zostaje bezpośrednio widmo binarne. Poszczególne parametry \bar{p}_i widma uśrednionego pozostają w następującej zależności od parametrów p_i widma pierwotnego wpisane z analizatora do minikomputera :

$$(1) \quad \bar{p}_i = \frac{1}{4} \sum_{j=0}^3 p_{i+j-2}$$

Takie przekształcenie widmowe stosowano dotychczas. Było ono wystarczające, gdy w grę wchodziło automatyczne rozpoznawanie spektrogramów binarnych wypowiedzi tego samego głosu przy użyciu jego indywidualnego zbioru wzorców.

Jak wynika z rys. 3b, uśrednianie widma według wzoru (1) nie wystarcza, aby w widmie binarnym wysokiego głosu zatrzeć harmoniczny charakter dźwięku mowy a zachować jedynie cechy formantowe mające dla spektrogramu binarnego wartość dystynktywną. W związku z tym postanowiono przeanalizować inne sposoby przekształceń widmowych, które zapewniałyby pożądaną eliminację wpływu wysokości głosu na jakość spektrogramu binarnego. Założono, iż utrzymane zostaną dotychczasowe zasady porównywania spektrogramów binarnych w trakcie automatycznego rozpoznawania wyrazów.

Jedno z przekształceń widmowych rozpatrzonych w niniejszej pracy stanowi rozwinięcie dotychczas stosowanego uśredniania widma. Zasadą tego przekształcenia wyraża wzór (2) :

$$(2) \quad \bar{p}_i = \frac{1}{k} \sum_{j=1}^k w_j \cdot p_{\left(\frac{k-1}{2} + j\right)}$$

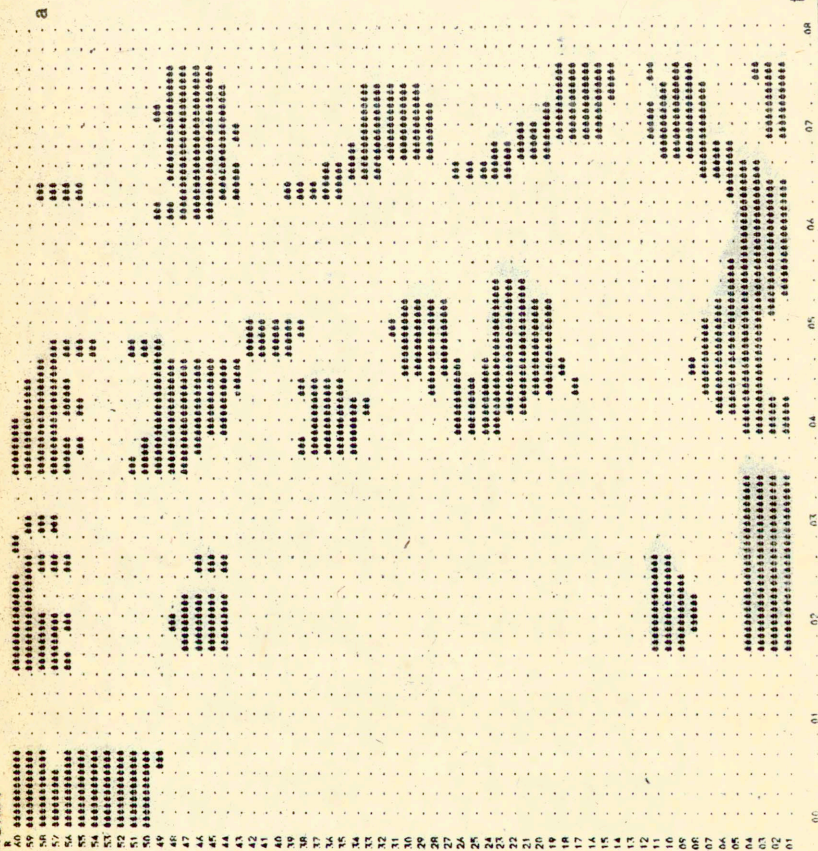
Kształt okna w_j określono w oparciu o kryterium, według którego przekształcenie zapewniłoby idealne wygładzenie widma prążkowego, w którym poszczególne prążki oddalone są od sąsiednich o stałą liczbę pasm podziału zakresu częstotliwości i są jednakowej wysokości. Wspomniane pasma utożsamiano z kanałami analizatora widma używanego w pracy. Kształt okna ma

Charakter dyskretny i jak wyliczono zależy od jego szerokości, to znaczy ilości pasm podziału zakresu częstotliwości jakie obejmuje, zaś szerokość ta wynika z maksymalnej odległości kolejnych prążków widma, wyrażonej także liczbą kanałów. Parametry okna wyliczono z warunków, jakie zilustrowano na rys. 4 przedstawiającym różne wzajemne pozycje okna i widma dla trzech wariantów odległości pomiędzy kolejnymi prążkami. Okno w przykładzie na rys. 4 jest siedmioelementowe, co oznacza, że obejmuje 7 kanałów. Kolejne rzędne tego okna mają następujące wartości : $0, \frac{1}{6}, \frac{1}{2}, \frac{5}{6}, 1, \frac{5}{6}, \frac{1}{2}, \frac{1}{6}, 0$. Szerokość okna decyduje o zakresie uśredniania i powinna być tak dobrana, aby nie dochodziło np. do zagubienia słabego formantu usytuowanego blisko sąsiadującego formantu silnego. Uwzględniając to wymaganie przyjęto okno o 2 elementy węższe. Jego parametry wyliczone na zasadzie wyżej przedstawionej wynoszą :

$$0, \frac{1}{4}, \frac{3}{4}, 1, \frac{3}{4}, \frac{1}{4}, 0.$$

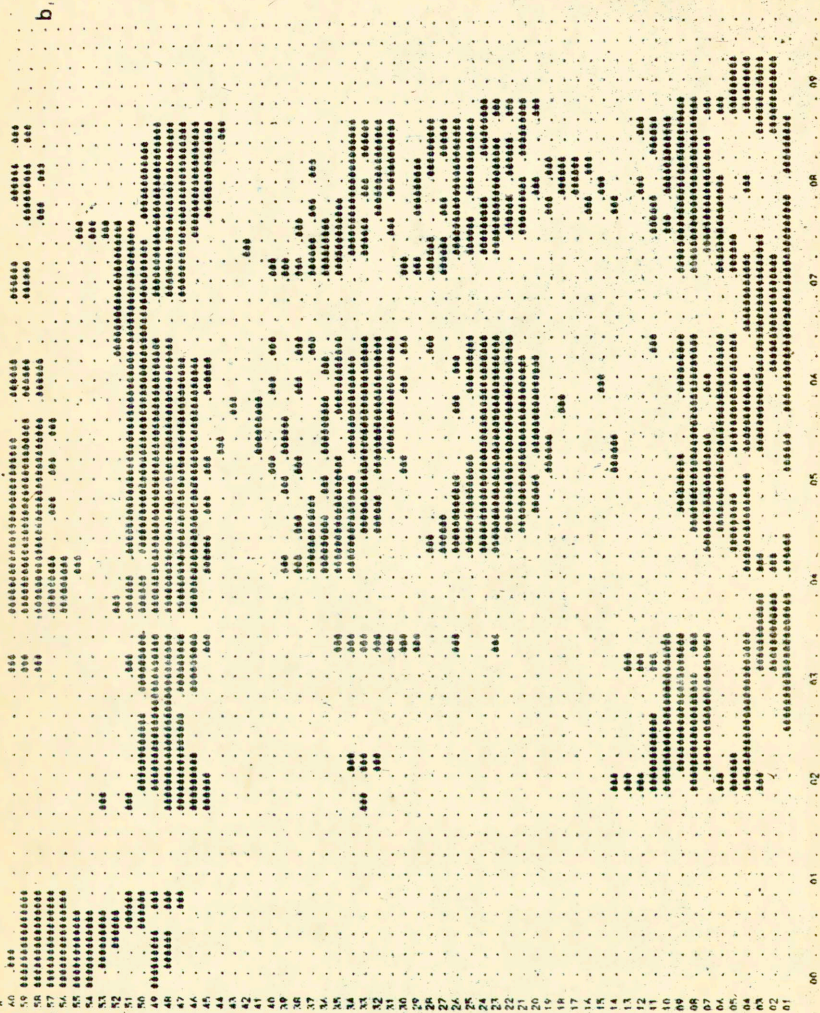
Można wykazać, że stosując takie okno nie uzyskuje się pełnego wygładzenia widma, gdy wartości kolejnych niezerowych parametrów widmowych różnią się o stały przyrost. Wyliczono, że uśrednianie gwarantuje wówczas wygładzenie widma z odchyleniami w granicach 10 %. W rzeczywistych widmach wysokiego głosu, uzyskanych przy pomocy wielokanałowego analizatora z podziałem zakresu częstotliwości na pasma o szerokości 80 Hz, oprócz składowych reprezentujących w przybliżeniu poszczególne harmoniczne występują też liczne składowe nieobecne w sygnale mowy, będące natomiast tworem analizy. Pojawiają się one z powodu zbyt rozległego zasięgu zbczoty filtrów analizujących. Ta okoliczność jest niekorzystna z punktu widzenia uśredniania widma, gdyż z jej powodu uśrednianie może wywoływać błędne deformacje obwiedni widma zwłaszcza w przypadkach, gdy w pewnym zakresie widmowym następuje znaczna i gwałtowna zmiana poziomu. Szczególnie niekorzystny rezultat uśredniania może mieć miejsce wówczas, gdy harmoniczna przypada na granicy pasm przepustowych sąsiadujących z sobą filtrów.

PASMA



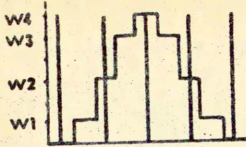
Rysunek 1a. Spektrogramy binarne słowa [apawdzelpa] wyznaczone dotychozasową metodą dla wypowiedzi głosem męskim M1

PASMA

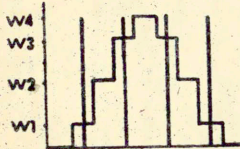


Rysunek 1b. Spektrogramy oznacze słowa [spawczelna] wyznaczone dotychczasową metodą dla wypowiedzi głosem żeńskim 22

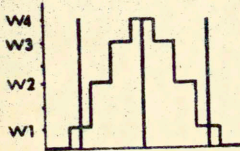
$$\Delta f = 2$$



$$2W_2 + W_4 = K$$

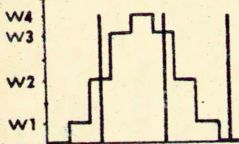


$$2W_1 + 2W_3 = K$$

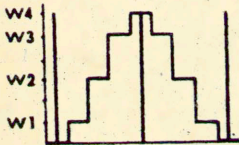


$$\Delta f = 3$$

$$2W_1 + W_4 = m$$

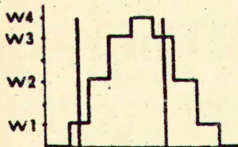


$$W_2 + W_3 = m$$

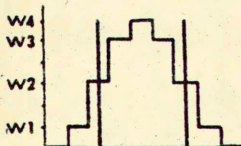


$$\Delta f = 4$$

$$W_4 = n$$



$$W_1 + W_3 = n$$



$$2W_2 = n$$

Rysunek 4. Wzajemne położenie okna wagowego i prążków widmowych dla trzech wariantów odległości pomiędzy kolejnymi prążkami oraz warunki pełnego wygładzenia przytoczonych widm.

Inną podjętą w tej pracy próbą wyeliminowania wpływu wysokości głosu na jakość spektrogramu binarnego stanowiło zastosowanie przekształcenia widmowego polegającego na tym, że poszczególnym parametrom p_i widma przekształconego nadaje się wartość największego elementu zbioru :

$$P_i = \{p_{i-k}, \dots, p_i, \dots, p_{i+k}\}$$

złożonego z $(2k+1)$ kolejnych parametrów widma pierwotnego. Wartość k zależy od spodziewanej wysokości głosu i od podziału zakresu częstotliwości na pasma analizy. Ustalono $k=1$ wychodząc z założenia, że dla wysokiego głosu kolejne harmoniczne przypadają najrzadziej w co trzecim paśmie analizatora.

Na rys. 5a, 5b, 5c przedstawiono przykładowe widmo wyznaczone za pomocą wielokanałowego analizatora oraz wyniki trzech przekształceń tego widma. Widmo oznaczone przez W_{UBW} jest rezultatem dotychczas stosowanego przekształcenia, tzn. uśredniania bez ważenia. Widma uzyskane przy zastosowaniu przekształceń omówionych wyżej oznaczono umownie przez W_{UZW} /uśrednione z ważeniem / i W_{EML} /przekształcone z zastosowaniem zasady ekspansji maksimum lokalnego/.

Przykład widma na rys. 5 pochodzi z samogłoski [e] w wyrazie [spuwdzelnja] wymówionym przez głos żeński. Podobny przykład, lecz odnoszący się do wypowiedzi głosem męskim, zamieszczono na rys. 6a, 6b, 6c. Jak należało oczekiwać, nowe przekształcenia widmowe przynoszą pożądany rezultat w przypadku głosu wysokiego, gdyż eliminują dyskretyzację widma wywołaną harmonicznym charakterem dźwięku nie uszczuplając przy tym informacji o formantach. Z tego względu są one korzystniejsze od przekształcenia stosowanego dotychczas i gwarantują uzyskiwanie doskonalszych spektrogramów binarnych.

3. Próba zmocyfikowania zasady wyznaczania widme binarnego

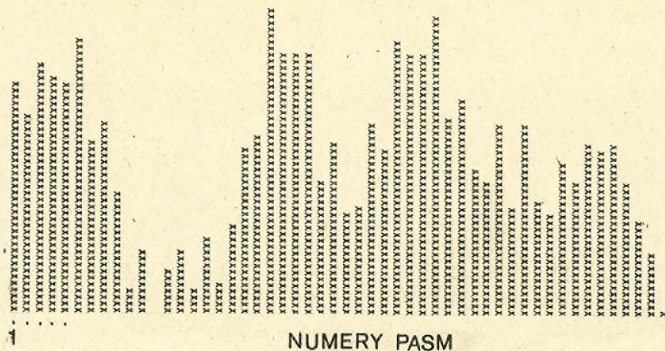
Według stosowanej dotychczas zasady wyznaczania widma binarnego poszczególne parametry $(p_{WB})_i$ widma binarnego otrzymują wartość "1.", jeżeli spełniony jest warunek :

$$(4) \quad (P_{WP})_i \geq \frac{(P_{WP})_{i-4} + (P_{WP})_{i+4}}{2} ,$$

POZIOM
-0-

-10-

-20-



WUBW

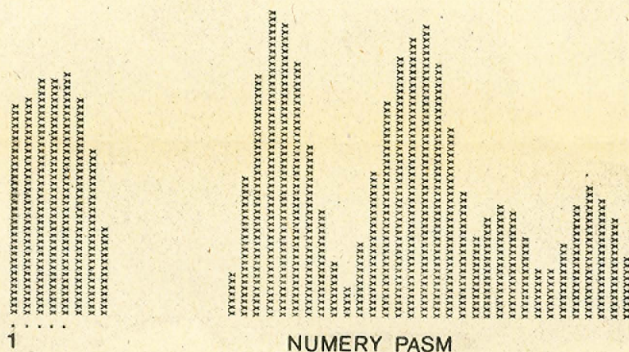
a

63

POZIOM
-0-

-10-

-20-



WUWZ

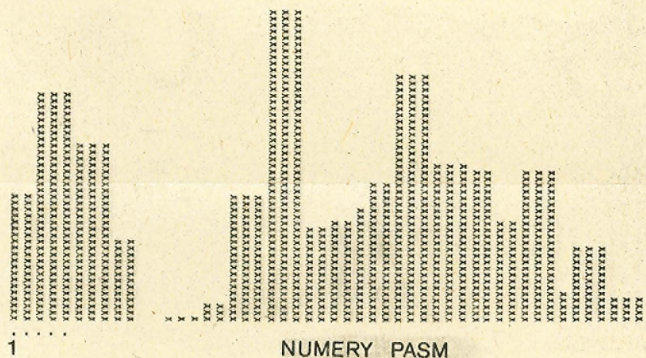
b

63

POZIOM
-0-

-10-

-20-



WEML

c

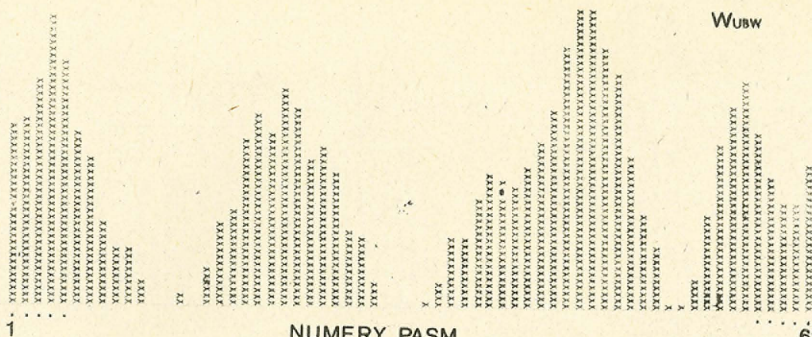
63

Rysunek 5. Widma zamglonek [e] w wyrazie [spudzelni] wypowiedzianym glosom zoladzim, uśrednione drogi trzech różnych przekształceń.

POZIOM
-0-

-10-

-20-



WuBw

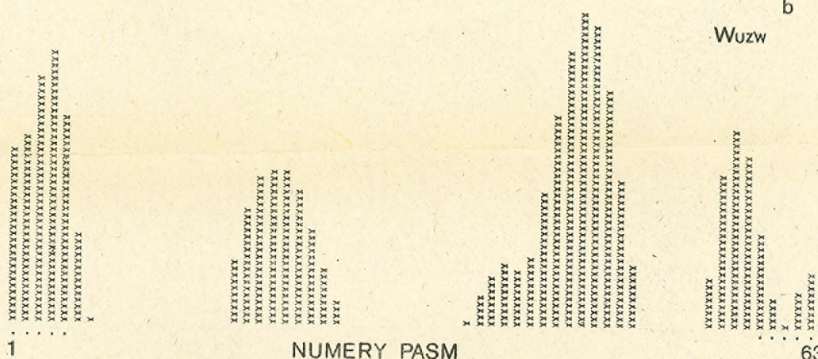
a

63

POZIOM
-0-

-10-

-20-



Wuzw

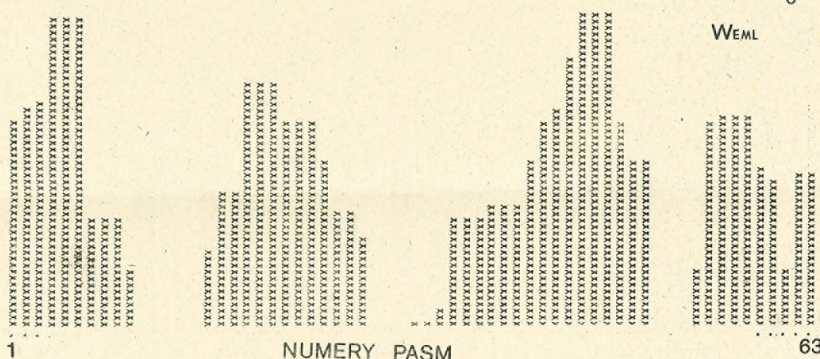
b

63

POZIOM
-0-

-10-

-20-



WeML

c

63

Рисунок 6. Видна сонгогтооки [e] в wyрyzie [pandzelju] wyрoвe-
дзiтoкy гiлoзeм мoдeлi, урoвнiе дoгoд трeч
рiзнyх прeкcaтoкeй.

gdzie $(p_{WP})_i$ oznacza parametry widma w formie przekształconej. Warunek (4) jest uproszczoną formą zapisu :

$$(5) \quad (P_{WP})_i \geq \frac{1}{3} \left[(P_{WP})_{i-4} + (P_{WP})_i + (P_{WP})_{i+4} \right]$$

Prawa strona nierówności (5) wyraża średni poziom sygnału w poszczególnych pasmach obejmujących zakresy częstotliwości poszczególnych grup dwunastu kolejnych kanałów analizatora, natomiast wyrażenie po lewej stronie stanowi zapis ogólny parametrów widma, uśrednionego zgodnie ze wzorem (1). Wyrażenie prawostronne należy interpretować jako próg będący funkcją lokalnej konfiguracji widma. W zakresach, w których widmo uśrednione przewyższa ten próg, widmo binarne otrzymuje wartość "1". Próba modyfikacji zasady wyznaczenia widma binarnego polegała na wprowadzeniu innej definicji progu. Według nowej definicji progami dla i-tego parametru widmowego p_i jest pomniejszona o stałą różnicę Δ najwyższa z wartości parametrów :

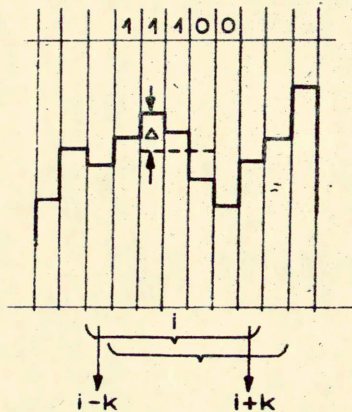
$$\{ p_{i-k}, \dots, p_i, \dots, p_{i+k} \}$$

Ilustrację tej definicji stanowi rys. 7. Metoda wyznaczania widma binarnego oparta o powyższą definicję progu będzie w dalszej części pracy oznaczana umownie literą M, natomiast metoda dotychczas stosowana literą S.

Zarówno dotychczasową jak i nową metodę wyznaczania widma binarnego wyposażono obecnie w kryterium stałego progu, który ustala się na poziomie o stałą różnicę Δ niższym od najwyższego poziomu, jaki osiąga dane widmo w którymkolwiek paśmie. Według tego kryterium parametry widma binarnego przyjmują w pierwszej kolejności wartość 0, jeśli odnośne parametry p_i widma cyfrowego, dla którego wyznaczane jest widmo binarne, spełniają jeden z następujących dwóch warunków :

- I. $p_i \leq p_{\max} - \Delta$, gdy $p_{\max} - \Delta \geq 0$
- II. $p_i \leq p_0$, gdy $p_{\max} - \Delta < 0$,

gdzie p_0 oznacza pewien umowny poziom najniższy, wyższy niż 0. Dotychczas stosowano jedynie warunek II.



Rysunek 7. Ilustracja zmodyfikowanej zasady wyznaczania widma binarnego.

4. Technika badań

System analogowo-cyfrowy, stosowany dotychczas w pracach nad automatycznym rozpoznawaniem wyrazów na podstawie spektrogramów binarnych, poszerzony został ostatnio o dodatkowe urządzenie peryferyjne o nazwie MEMOSKOP oraz o zwykły odbiornik telewizji czarno-białej.

Memoskop skonstruowany przez P.Janickiego ma być przedstawiony w osobnej publikacji. Urządzenie to współpracując z minikomputerem MERA 303 oraz poprzez wejście antenowe z odbiornikiem telewizyjnym, umożliwia wyświetlanie obrazów binarnych na ekranie TV. Obrazy są najpierw programowo formowane z informacji zawartych w pamięci minikomputera, a następnie zostają zapisane w postaci ciągu liczb binarnych w pamięci typu RAM w memoskopie. Stąd mogą być cyklicznie odczytywane, kodowane według standardu telewizyjnego i wyświetlane na ekranie odbiornika TV. Zestaw MEMOSKOP-ODBIORNIK TV stwarza w badaniach dogodną możliwość szybkiej kontroli zależności obrazu spektrogramu binarnego od różnych kryteriów formalnych, występujących w procedurze

wyznaczania spektrogramów binarnych. Zestaw ten pozwolił skrócić znacznie czas wykonywania części doświadczalnej, oszczędził drukarkę mozaikową oraz co najmniej kilkadziesiąt arkuszy papieru do niej.

W ramach niniejszej pracy zbudowano w języku wewnętrznym minikomputera MERA 303 szereg nowych programów dla następujących operacji :

- a/ uśrednianie widma z ważeniem,
- b/ przekształcenie widmowe eskalujące wartość silniejszych parametrów na najbliższe parametry okoliczne,
- c/ wyznaczanie widm binarnych metodą M,
- d/ kodowanie spektrogramu binarnego w spektrogramie cyfrowym,
- e/ wpis jednego lub dwóch spektrogramów binarnych do memoskopu,
- f/ wpis jednego lub dwóch widm cyfrowych lub binarno-cyfrowych do memoskopu,
- g/ wydruk w postaci graficznej widm cyfrowych lub binarno-cyfrowych.

Do badań wybrano dwa wyrazy : [spuwdzɛlna] i [vambiez#tɕe]. Wyrazy te zawierają przykłady wszystkich typów polskich fonemów. Dwa głosy męskie i dwa żeńskie wypowiedziały w warunkach studyjnych trzykrotnie każdy z tych wyrazów. Wypowiedzi nagrano wpięrcw na taśmie magnetofonowej. Następnie przy użyciu wielokanałowego analizatora widma oraz sprzężonego z nim poprzez konwerter A/C minikomputera MERA 303 wykonano dla nich spektrogramy cyfrowe, które zapisano na dyskach elastycznych. Stworzono 5 zestawów programowych umożliwiających pobieranie z dysku poszczególnych spektrogramów cyfrowych i wyznaczanie z nich spektrogramów binarnych przy zastosowaniu różnych przekształceń widmowych i w oparciu o różne definicje progu, wobec którego następuje wartościowanie binarne poszczególnych parametrów widma. Łącznie wyznaczono 120 spektrogramów binarnych, które także zgromadzono na dysku elastycznym, skąd następnie mogły być pojedynczo w dowolnej kolejności pobierane i wyświetlane na ekranie telewizyjnym, po dwa jednocześnie. Niezależnie od tego, w sporadycznych przypadkach wyświetlano w formie graficznej widma cyfrowe, z których powstały wątpliwej treści widma

binarne. W celu udokumentowania wybranych widm cyfrowych lub spektrogramów binarnych można było sporządzić ich wydruki.

5. Ocena zbadanych sposobów wyznaczania spektrogramów binarnych i wnioski

Każdy z badanych sposobów wyznaczania widma binarnego zeterminowany był przez typ przekształcenia widmowego i zastosowaną definicję progę. W grę wchodziły :

- 1/ trzy wyżej omówione rodzaje przekształceń widmowych, które odąd nazywane będą w skrócie symbolami :
U /uśrednianie bez ważenia/, UW /uśrednianie z ważeniem/, EM /przekształcenie na zasadzie eskalacji wartości najsilniejszego parametru na najbliższe parametry okoliczne/,
2/ dwa warianty metody wyznaczania spektrogramów binarnych, które w rozdziale trzecim oznaczono umownie symbolami S i M.

W badaniach wariant S stosowano łącznie z przekształceniami U, UW i EM+U, natomiast wariant M był stosowany z przekształceniami UW i EM. Ocenie poddano zatem pięć sposobów tworzenia widma binarnego. Sposoby te dalej nazywane będą umownie metodami i oznaczane odpowiednio symbolami : S/U, S/UW, S/EM+U oraz M/UW, M/EM.

Na ocenę poszczególnych metod składała się jakość spektrogramów binarnych. O niej decydowały głównie dwa kryteria oraz stopień podobieństwa spektrogramów binarnych tego samego wyrazu wypowiedzianego kilkakrotnie różnymi głosami. Wspomniane dwa kryteria brzmiały następująco-:

- 1/ w zakresach formentów dystynktywnych parametry widm binarnych powinny konsekwentnie mieć wartość "1",
2/ jedynek spektrogramu binarnego nie powinny układać się w chaotyczną mozaikę, lecz muszą tworzyć w kierunku czasowym regularne przebiegi.

Ocenę jakości spektrogramów binarnych przeprowadzono drogą obserwacji ich obrazów na ekranie telewizyjnym. Przypomina się, że memoskop umożliwia jednoczesne oglądanie dwóch spektrogramów binarnych.

Porównując, w oparciu o wyżej sformułowane kryteria, spektro-

gramy binarne wyznaczone z tego samego spektrogramu cyfrowego różnymi metodami, stwierdzono wstępnie, że najlepszą wydaje się być metoda S/UW. Spektrogramy binarne wyznaczone tą metodą przyjęto więc jako odniesienie i porównywano z nimi spektrogramy binarne wyznaczone z tego samego spektrogramu cyfrowego innymi metodami. W trakcie takiego porównywania punktowano zawsze spektrogram lepszy lub oba równocześnie, jeśli różnice w jakości były nieznaczne. Tą drogą poszczególne metody otrzymały ocenę wyrażoną w punktach, przy czym maksymalna liczba punktów wynosić mogła 24. Wyniki tej oceny zamieszczono w tabelicy 1.

Tablica 1
Wyniki oceny badanych metod wyznaczania
spektrogramów binarnych

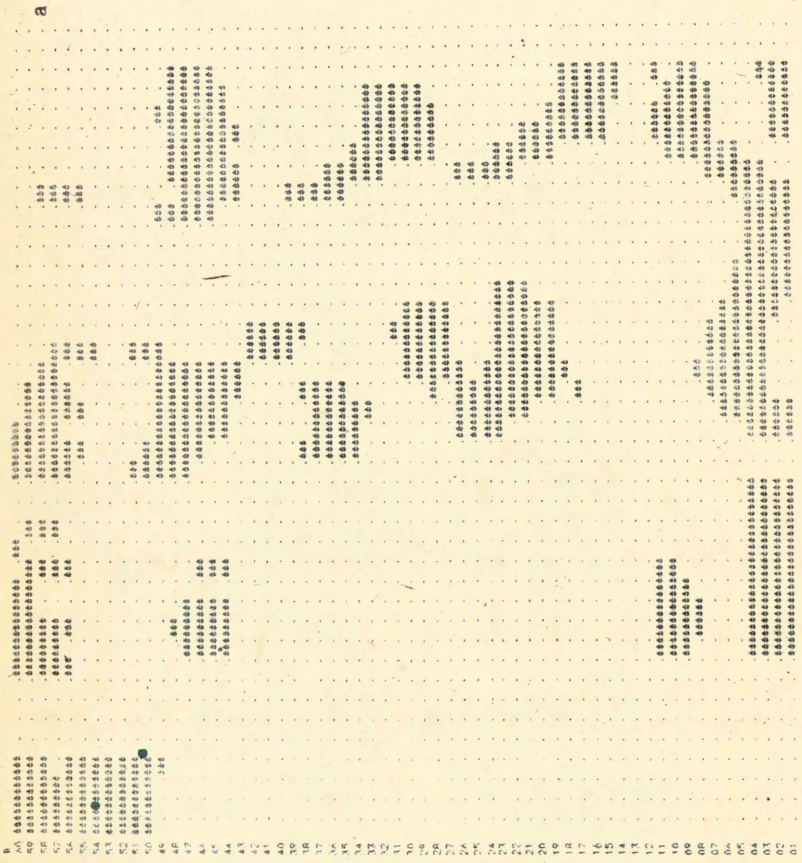
L.p.	Metoda	Ocena w punktach
1.	S/U	13
2.	S/UW	17
3.	S/EM+U	16
4.	M/UW	10
5.	M/EM	9

O dodatniej ocenie metod S/UW i S/EM+U nad metodą S/U zadecydowały głównie spektrogramy głosów wysokich. Stąd wniosek, że gdyby w badaniach wzięło udział więcej głosów żeńskich, przewaga pozytywnej oceny metody S/UW byłaby jeszcze znaczniejsza. Niskie oceny metod M/UW i M/EM dowodzą, że dotychczasowa zasada ustalania progu względem którego wartościowane są poszczególne parametry widma binarnego, jest korzystniejsza od zasady jaką zastosowano po raz pierwszy w tej pracy. Uśrednienie z ważeniem oraz przekształcenie kombinowane EM+U okazały się korzystniejsze od dotychczas stosowanego przekształcenia SU.

Ocena poszczególnych metod w oparciu o kryterium podobieństwa spektrogramów binarnych tego samego wyrazu wymówionego wielokrotnie przez różne głosy, wypadła także najkorzystniej dla metody S/UW a najgorzej dla metod M/UW i M/EM.

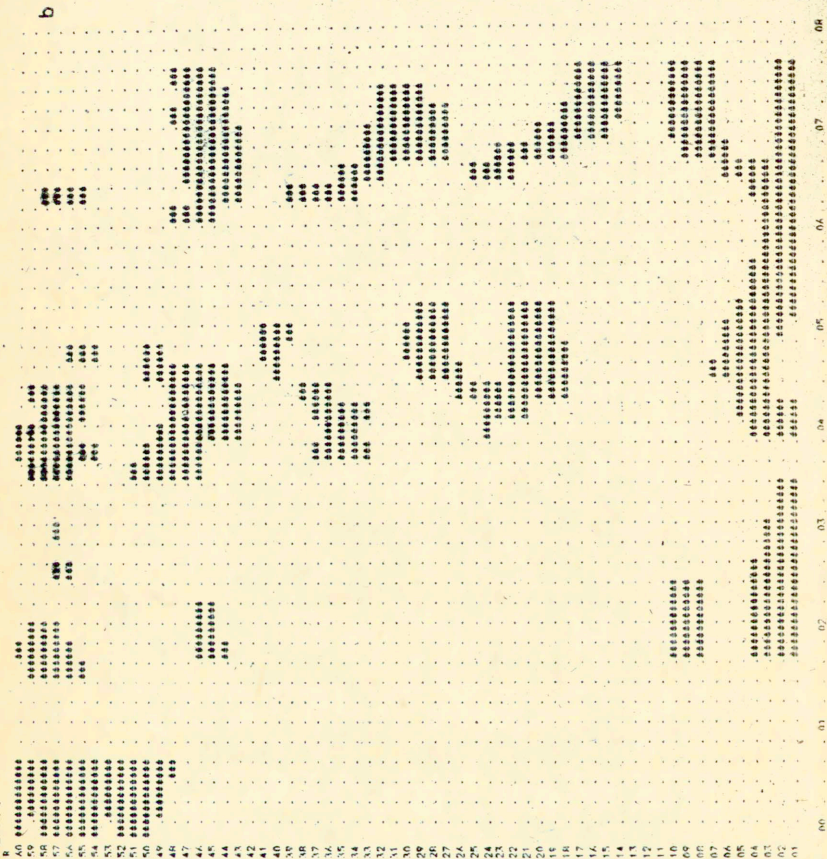
Ograniczone ramy tej pracy nie pozwalają na udokumentowanie otrzymanych rezultatów odpowiednimi ilustracjami spektrogramów binarnych. Poniżej przytoczono jedynie dla przykładu dwa zestawy spektrogramów binarnych wyrazu [spuwdzelna] wymówionego głosem męskim i żeńskim. Spektrogramy dotyczące głosu męskiego zamieszczono na rys. 8 a głosu żeńskiego na rys. 9.

Na zakończenie wspomnieć należy, iż przeprowadzono także szereg innych prób udoskonalenia spektrometrii binarnej, lecz nie przyniosły one zadowalających rezultatów i stąd nie omówiono ich w niniejszej pracy.



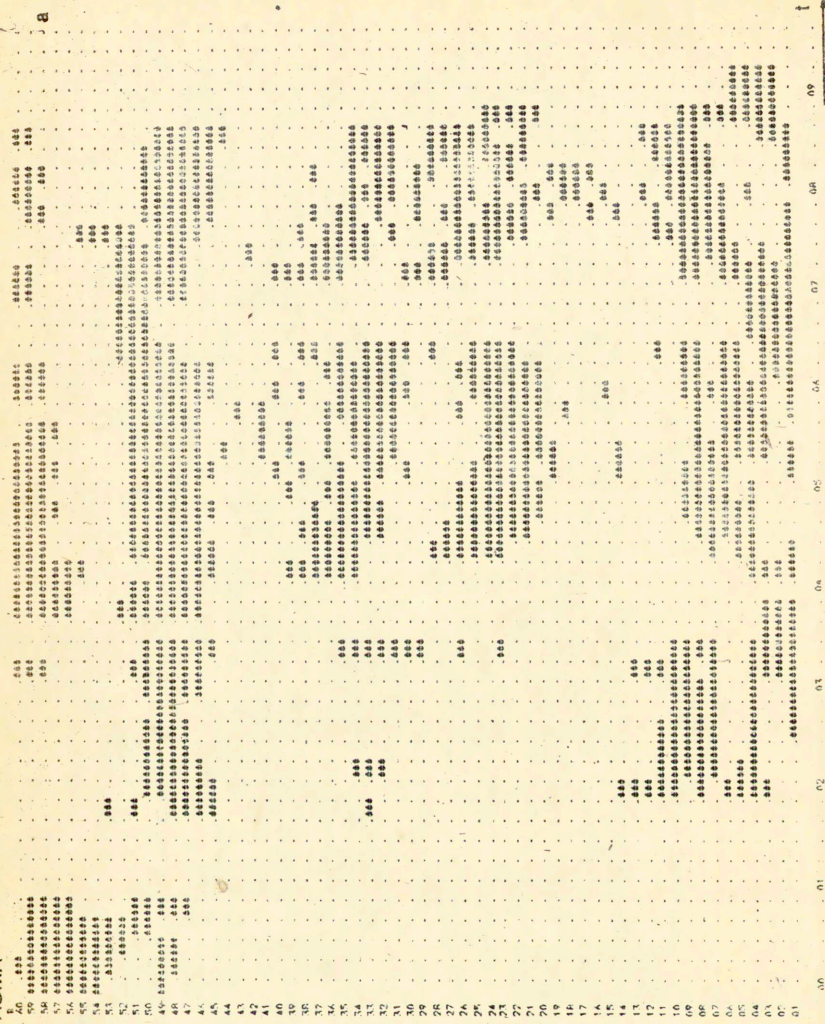
Rysunek 8a. Spektrogramy wypowiedzi głosem męskim wyrazu [spwazdzelpa] wyznaczone metodą S/U

PASMA



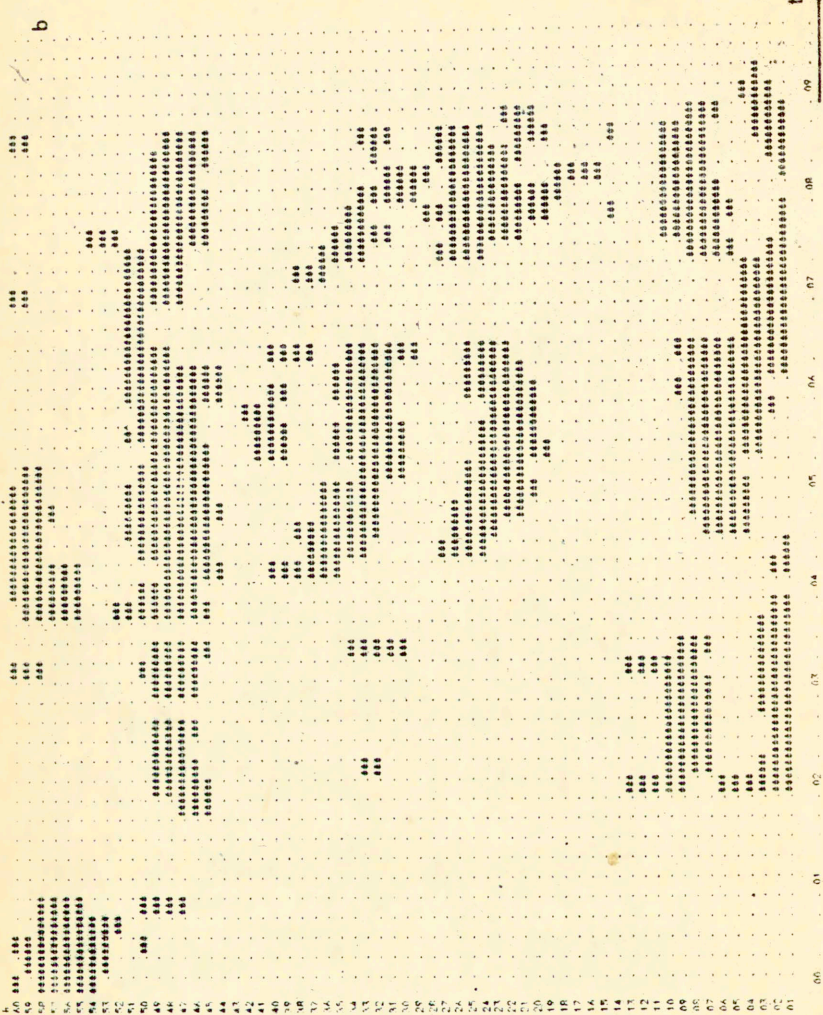
Rysunek 8b. Spektrogramy wypowiedzi głosem męskim wyrazu [apawdzelpa] wyznaczone metodą S/UW

PASMA



Rysunek 3a. Spektrogramy wypowiedzi głosem żeńskim wyrazu [spuwdzjelna] wyznaczone metodą 3/U

PASMA



Rysunek 9b. Spektrogramy wypowiedzi głosem żeńskim wyrazu [spudzelka] wyznaczone metodą S/UW

LITERATURA

- [1] KUBZDELA, H. : Wizualizacja sygnału mowy w formie spektrogramów binarnych, Materiały XXVII Otwartego Seminarium z Akustyki, s. 167-171, Warszawa-Puławy 1980.
- [2] KUBZDELA, H. : Metoda automatycznego rozpoznawania wyrazów w oparciu o spektrogramy binarne, Prace IPPT, 14, 1981.
- [3] KUBZDELA, H. : Automatyczne rozpoznawanie wyrazów na podstawie spektrogramów binarnych, Prace IPPT, 15, 1981.
- [4] KUBZDELA, H. : Automatyczne rozpoznawanie wyrazów na podstawie spektrogramów binarnych, Materiały XXVIII Otwartego Seminarium z Akustyki, s. 178-181, Gliwice 1981.
- [5] KUBZDELA, H. : Weryfikacja i optymalizacja metody rozpoznawania wyrazów w skończonych zbiorach hasłowych w oparciu o spektrogramy binarne, Prace IPPT, 10, 1982.