

Grażyna Demenko

ANALIZA MATEMATYCZNA
CECH OSOBNICZYCH GŁOSU
W ZAKRESIE PARAMETRU F_0

24/1984

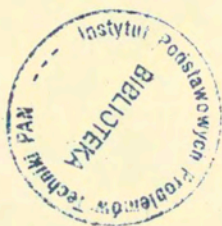
P. 269



WARSZAWA 1984

ISSN 0208-5658

Praca wpłynęła do Redakcji dnia 23 listopada 1983 r.



56972



N a p r a w a c h r ę k o p i s u

Instytut Podstawowych Problemów Techniki PAN

Nakład 140 egz. Ark.wyd. 2,4, Ark.druk. 3,25

Oddano do drukarni w czerwcu 1984 r.

Nr zamówienia 413/84.

Warszawska Drukarnia Naukowa, Warszawa,
ul. Śniadeckich 8

Grażyna Demenko
Pracownia Fonetyki Akustycznej
IPPT PAN

ANALIZA MATEMATYCZNA CECH OSOBNICZYCH GŁOSU W ZAKRESIE
PARAMETRU F_0 .^{1/}

Streszczenie

Przeprowadzono próbę analizy cech osobniczych w zakresie częstotliwości podstawowej. W przyjętym modelu wykorzystano czasowe zmiany parametru F_0 . Dla ustalenia różnicowań międzyosobniczych przeprowadzono analizę jednego zdania powtarzanego pięciokrotnie przez 10 osób. W celu określenia przydatności zastosowanej metody, dokonano szczegółowej analizy przykładowo przyjętych 16 przebiegów. Dokonano transformacji 6-wymiarowych wektorów, reprezentujących poszczególne fragmenty przebiegu częstotliwości podstawowej w badanym zdaniu, do układu współrzędnych utworzonych przez wektory własne macierzy kowariancji. Mimo zastosowania bardzo uproszczonego modelu otrzymane wyniki wskazują na jego wysoką użyteczność. Uzyskano w 95 % podobieństwo między wektorem pierwotnym 6-ciowymiarowym a wektorem dwuwymiarowym otrzymanym w wyniku transformacji. Stwierdzono bardzo duże osobnicze różnicowania w przebiegach parametru F_0 w niektórych fragmentach analizowanego zdania.

Uzyskane wyniki wskazują na przydatność przedstawionej metody do analizy parametru F_0 pod kątem praktycznych zastosowań np. w rehabilitacji, rozpoznawaniu głosów lub syntezie.

1. Wstęp.

Jednym z ważniejszych zagadnień analizy akustycznych para-

1/

Praca wykonana w ramach problemu węzłowego 06.9.01

metrów jest utworzenie odpowiedniego opisu czasowej zmienności częstotliwości podstawowej. Zasadniczą trudnością jest specyficzny charakter badanego parametru. Toteż niewystarczające jest przyjmowanie tradycyjnych metod stosowanych zwykle w analizie sygnałów. Ilościowy opis zmienności parametru F_0 jest utrudniony z powodu braku odpowiedniego modelu matematycznego. Zagadnieniom tym poświęcono szereg prac np. ([8], [4], [14]) są one jednak albo nieprecyzyjne i przez to uniemożliwiają właściwy opis albo są zbyt skomplikowane dla praktycznego zastosowania. Przedstawiona praca jest kontynuacją poprzednich badań ([3], [4]), których celem było stworzenie modelu matematycznego umożliwiającego opis charakteru zmian parametru F_0 . Przyjęta metoda powinna być uniwersalna tzn. musi umożliwiać badanie wszystkich cech analizowanego sygnału oraz być w miarę prosta. Wiele uwagi należy poświęcić zagadnieniu interpretacji zmienności parametru F_0 . Właściwe bowiem sformułowanie problemu w sposób decydujący wpływa na rezultaty pracy. Na zmienność przebiegów częstotliwości podstawowej (między innymi) istotny wpływ mają cechy osobnicze głosu. Temu zagadnieniu poświęcono szczególnie dużo uwagi w przedstawianej pracy. Za jej cel przyjęto opracowanie modelu matematycznego umożliwiającego ilościowy opis różnych cech sygnału w zakresie parametru F_0 . Założono wstępnie, iż należy dokonać matematycznego opisu cech międzypersonalnych głosu.

2. Cechy indywidualne w przebiegach częstotliwości podstawowej.

Ważnym czynnikiem wpływającym na zmienność przebiegów F_0 jest indywidualny charakter tonu krtańowego związany z charakterystycznymi własnościami głosu danego mówcy. Indywidualne cechy głosu związane ze zmianami częstotliwości podstawowej przedstawiono między innymi w pracach ([3], [12], [23]). W niektórych pracach (por. [1], [2]) sugeruje się częstotliwość podstawową jako dodatkowy parametr użyteczny przy rozpoznawaniu głosów. Najczęściej wykorzystuje się statystyczne cechy tego parametru: jego wartość średnią i cechy rozkładu (por. [3], [12], [23]). Informację o indywidualnych cechach głosu zawierają jednak nie tylko statystyczne cechy rozkładów parametru F_0 , lecz również czasowe zmiany tego parametru - kontur przebiegu.

Fakt ten może być między innymi istotny w problemie rozpoznawania głosów, ponieważ o ile jest możliwe naśladowanie takich parametrów jak wartość średnia przebiegu, to kłopotliwe staje się naśladowanie zmian częstotliwości podstawowej w czasie. Porównując wizualnie dostatecznie dużą liczbę tonogramów różnych osób można zauważyć pewne predyspozycje w stosowaniu określonych przebiegów parametrów F_0 . Uwzględnienie całej informacji a więc zarówno cech statystycznych jak i cech wynikających ze zmian częstotliwości podstawowej w czasie może stanowić dostateczną podstawę do wykorzystania tego parametru w rehabilitacji, rozpoznawaniu głosów lub syntezie. Jednym z zagadnień rehabilitacji osób z wadami wymowy lub osób głuchych jest problem związany z nauką intonacji ([17], [24]). W pracy [25] sugeruje się naukę intonacji poprzez analizę wyświetlanego obrazu - przebiegu F_0 na ekranie monitora. Szereg eksperymentów potwierdziło wysoką przydatność takiej metody. Konieczne staje się więc określenie typowych (dla głosów pełnosprawnych) cech częstotliwości podstawowej oraz ustalenie, jakie cechy wynikają z osobniczych zróżnicowań.

W zakresie badań statystycznych istotnych informacji dostarcza analiza wartości średniej oraz parametrów rozkładu (odchylenia od normy uwidaczniają się w wartości średniej jak również w charakterze rozkładu). W przypadku analizy zmian częstotliwości podstawowej głosów niepełnosprawnych można zauważyć następujące zjawiska: nienaturalną monotonię przebiegu, nieuzasadnione spadki częstotliwości, zaburzenia w periodyczności przebiegu wywołujące odpowiednie zmiany w przebiegu parametru F_0 . Pierwszoplanowym zagadnieniem staje się więc określenie cech osobniczych dla głosów pełnosprawnych, gdyż na tej podstawie można badać ewentualne odchylenia od normy (por. [11]).

W celu zbadania cech osobniczych w zakresie parametru F_0 wykonano następujący eksperyment. Na taśmie magnetofonowej zapisano 3 wypowiedzi (traktowane później jako wzorcowe): "pójdziemy do kina?", "pójdziemy do kina!", "pójdziemy do kina." Do przeprowadzenia eksperymentu wybrano losowo 10 głosów (5 żeńskich i 5 męskich). Każdej z tych osób postawiono zadanie odtworzenia usłyszanego wzorca. Odtworzenie to miało polegać na powtórzeniu

wypowiedzi (z zachowaniem naturalnych właściwości głosowych). Doświadczenie powtórzono czterokrotnie, w godzinę po pierwszym eksperymencie, w dniu następnym, po tygodniu oraz po miesiącu. W ten sposób uzyskano dla każdej z osób po pięć powtórzeń tego samego wzorca.

Z przyjętego materiału dokonano ekstrakcji parametru F_0 wyprawdzając wyniki pomiaru na ekran telewizora. Do szczegółowej analizy przyjęto zdanie - "pójdziemy do kina?". W pozostałych zdaniach zauważono podobne (jak w wybranym zdaniu) cechy osobnicze. Na rycinach 1,2 przedstawiono przykładowo wyniki ekstrakcji parametru F_0 dla czterech osób. Skala częstotliwości jest liniowa. Znaczniki umieszczono co 20 Hz, poczynając od najmniejszej mierzonej wartości równej 70 Hz. Znaczniki na skali czasu umieszczono co 8 danych (tzn. co 80 ms). Obszar czarno-biały odpowiada jednej wypowiedzi natomiast linia czarno-biała drugiej. Taki sposób wyświetlania obrazu umożliwi natychmiastowe porównanie dwóch przebiegów intonacyjnych.

Przedstawione zdjęcia ukazują bezpośrednio wynik ekstrakcji parametru F_0 przed jakimikolwiek matematycznymi procedurami (wygładzaniem, normalizacją). Porównując wypowiedzi tego samego zdania dla poszczególnych osób można zauważyć (bez jakiegokolwiek normalizacji), że bardzo podobne do siebie są przebiegi w wypowiedziach tej samej osoby, natomiast dla różnych osób kontury przebiegów znacznie się różnią. Można więc założyć, że zróżnicowania wewnątrzosobnicze są mniejsze niż międzyosobnicze.

3. Model matematyczny.

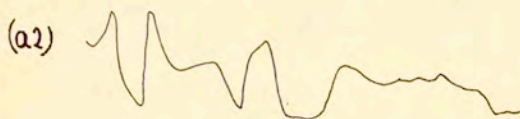
3.1. Przygotowanie danych otrzymanych z pomiaru częstotliwości podstawowej do analizy matematycznej.

3.1.1. Ekstrakcja parametru F_0 oraz redukcja danych.

Właściwy sposób pomiaru częstotliwości podstawowej ma zasadnicze znaczenie we wszystkich pracach związanych z analizą tego parametru. Należy dokładnie określić jakich błędów pomiarowych można się spodziewać i przyjęc taką metodę, która umożliwi poprawny opis zmian przebiegu częstotliwości podstawowej w czasie. Nie jest mankamentem, jak się nieraz sądzi, dostarczenie przez analizę instrumentalną zbyt dużej liczby dokładnych



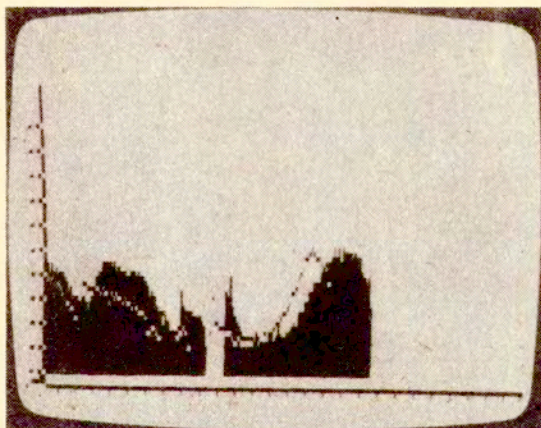
(a)



(a1) tonogram

(a2) oscylogram

(b)

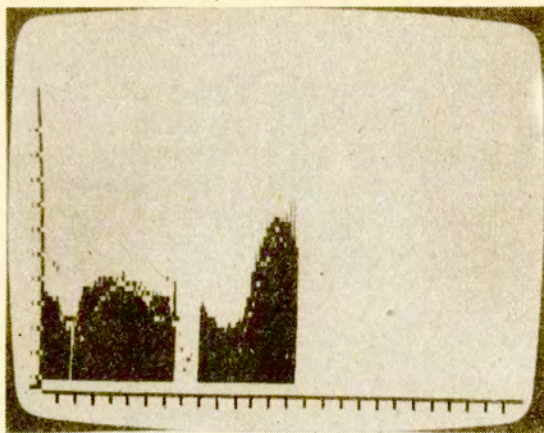


Ryc.1. Przebiegi częstotliwości podstawowej w zdaniu "pójdziemy do kina?" wymawianym przez osobę HK.

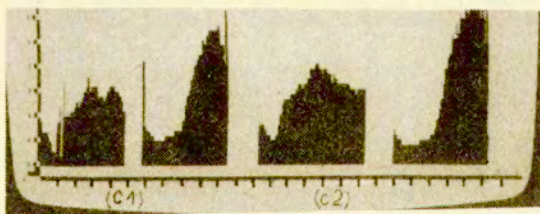
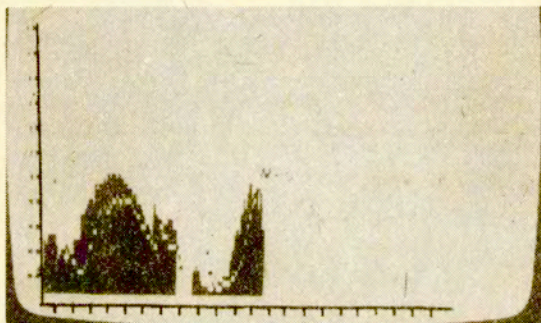
(a) ekstrakcja analogowa

(b) ekstrakcja cyfrowa. Przebieg odpowiadający obszarowi czarnemu na fotografii odpowiada przebiegowi z ryc. 4a. Przebieg oznaczony linią czarno-białą odpowiada powtórzeniu wypowiedzi przez osobę HK.

(a)



(b)



Ryc. 2. Przebiegi częstotliwości podstawowej w zdaniu "pójdziemy do kina ? " wymawianym dwukrotnie przez osoby :

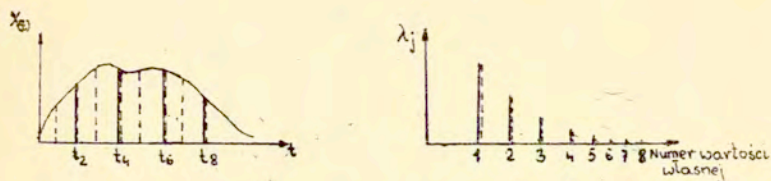
(a) JI

(b) WJ

(c) PJ (ryc. c1 odpowiada jednemu zdaniu, ryc. c2 drugiemu).

danych, lecz niewłaściwy sposób ich interpretacji. Na błąd cyfrowej ekstrakcji parametru F_0 ma wpływ szereg różnorodnych czynników. Każdy ze składników układu pomiarowego wprowadza swój stały błąd związany z charakterystyką przenoszenia sygnału oraz błędy losowe. Przyjęcie stałej częstotliwości próbkowania ułatwia numeryczną analizę danych lecz wprowadza błędy trudne do określenia. Przyjmując częstotliwość zbyt niską traci się informację o zmianach częstotliwości między momentami pomiaru (co zwłaszcza w wysokich żeńskich głosach może mieć istotne znaczenie), zaś ustalając większą częstotliwość próbkowania należy się liczyć z możliwością dwu- lub kilkukrotnego pomiaru tego samego okresu. Najpoprawniejszą metodą jest dokonywanie pomiaru każdego okresu. Zaletą tego rozwiązania jest otrzymanie dokładniejszych danych reprezentujących przebieg F_0 , wadą dodatkowe utrudnienia obliczeniowe. Najwłaściwsze wydaje się zastosowanie metody pozwalającej na dokładny pomiar przy możliwie niewielkich nakładach obliczeniowych. Metodę ekstrakcji parametru F_0 oraz przetwarzania wyniku pomiaru na postać cyfrową przedstawiono w pracach ([15], [19]). Dane otrzymane z pomiaru każdego okresu wprowadzono do pamięci minikomputera. Następnie za pomocą prostego algorytmu dokonano powtórnego odczytu w liniowej skali czasu, przyjmując kolejne dane jako średnie wartości częstotliwości w stałych przedziałach czasowych. Ważnym zagadnieniem jest ustalenie właściwej liczby danych reprezentujących analizowany przebieg. Przyjęcie zbyt dużej liczby danych nie tylko komplikuje algorytmy obliczeniowe ale również nie uwidacznia w sposób natychmiastowy istotnych cech badanego przebiegu. Ustalenie zaś zbyt małej liczby danych reprezentujących przebieg jest oczywistym błędem. Powyższy problem można zilustrować prostym przykładem.

Załóżmy, że przebieg x_t jest próbkowany w 4 momentach czasowych oznaczonych t_2, t_4, t_6, t_8



Ryc. 3. Przebieg funkcji x_t i jego wartości własne.

Macierz autokorelacji zdefiniowana następująco

$$R = \begin{bmatrix} R(t_1, t_1) & \dots & R(t_1, t_n) \\ \vdots & & \vdots \\ R(t_n, t_1) & \dots & R(t_n, t_n) \end{bmatrix}$$

posiada w tym przypadku 4 wartości własne przedstawione na ryc. 3. Zakładając, że podwoimy liczbę próbek, otrzymamy przebieg próbkowany w 8 momentach czasowych oznaczonych $t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8$ (punkty te zaznaczono linią przerywaną). Macierz R ma teraz 8 wartości własnych (oznaczonych na ryc. 3 linią przerywaną). Jeśli tak jak w przedstawionym przykładzie dodatkowe wartości własne (powstałe z przyjęcia podwójnej liczby punktów próbkowania) są małe to można przyjąć, że wystarczająca jest reprezentacja przebiegu taką liczbą danych, jaka odpowiada największym wartościom własnym. W powyższym przykładzie widać, iż próbkowanie w 4 momentach czasowych wystarczająco dobrze odzwierciedla charakter przebiegu. Dwukrotne powiększenie częstotliwości próbkowania nie spowodowało dodania istotnych cech do badanego przebiegu. Kryterium pozwalające dobrać właściwą liczbę danych reprezentujących dany przebieg sformułowane jest następująco (por. [7]).

$$I_n = \frac{1}{2} \frac{\sum_{i=1}^n \{ R(t_{2i}, t_{2i}) - R(t_{2i}, t_{2i-1}) - R(t_{2i-1}, t_{2i}) + R(t_{2i-1}, t_{2i-1}) \}}{\sum_{i=1}^n \{ R(t_{2i}, t_{2i}) + R(t_{2i-1}, t_{2i-1}) \}}$$

Jeżeli wartość $J_n \ll 1$ to należy uznać, że n próbek wystarczy do reprezentowania całego przebiegu. Taka redukcja danych ma duże znaczenie ponieważ umożliwia pozbycie się już na początku analizy nadmiaru danych i upraszcza ją w istotny sposób. W przypadku dysponowania skromnym sprzętem komputerowym ma to decydujące znaczenie.

3.1.2. Opisy przebiegów częstotliwości podstawowej.

Przy tworzeniu jakiegokolwiek zbioru sygnałów (przebiegów F_0) należy zwrócić uwagę na charakterystyczne własności elementów zawartych w tym zbiorze oraz na źródła zmienności sygnału. Na charakter przebiegów częstotliwości podstawowej ma wpływ szereg (na ogół znanych) różnorodnych przyczyn.

W rezultacie duże zróżnicowanie analizowanych przebiegów powoduje kłopoty z ich interpretacją. Z punktu widzenia analizy sygnałów interesujące jest nie tyle źródło (lub źródła) zmienności sygnału co skutki jakie one wywołują. Należy więc zwrócić uwagę na to, czym różnią się badane sygnały. W odniesieniu do przebiegów częstotliwości podstawowej można stwierdzić, iż różnią się one następującymi cechami :

- (a) liczbą odcinków, w których przebieg można uważać za ciągłą funkcję czasu,
- (b) charakterem i układem przerw między tymi odcinkami,
- (c) położeniem przebiegów F_0 w skali częstotliwości i w skali czasu,
- (d) szybkością zmian częstotliwości,
- (e) sposobem zmian częstotliwości.

Aby możliwa była analiza tak złożonego sygnału, konieczne jest przyjęcie pewnych założeń pozwalających na opracowanie w miarę prostych algorytmów. Założenia te zwykle dotyczą zagadnień związanych z normalizacją czasową oraz częstotliwościową parametru F_0 . Proponowane sposoby normalizacji częstotliwościowej dotyczą w większości następujących przekształceń :

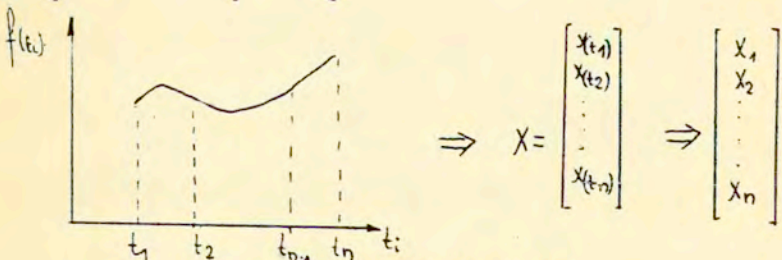
1. Przyjęcie logarytmicznej skali częstotliwości. Skala taka umożliwia bezpośrednie porównywanie głosów o różnych wysokościach; nie eliminuje różnic spowodowanych położeniem głosów w różnych zakresach skali częstotliwości.

2. Podzielenie wartości częstotliwości przez wartość średnią badanego zbioru przebiegów lub przez odchylenie standardowe. Normalizacja taka w znacznej mierze usuwa różnicowania spowodowane zmiennymi wysokościami głosu, nie likwidując ich całkowicie.
3. Odejmowanie wartości średniej zbioru.
Metoda ta w przypadku przebiegów F_0 nie ma większego sensu z uwagi na to, że centralizowanie przebiegów znajdujących się w różnych zakresach skali częstotliwości bez zmiany skali nie daje porównywalnych wyników.
4. Sprowadzenie wszystkich analizowanych przebiegów do przebiegów o jednakowej średniej.
5. Dzielenie wartości częstotliwości przez wartość maksymalną lub minimalną przebiegu.

Ponieważ w niniejszej pracy przyjęto za cel zbadanie różnic międzysobniczych wynikających jedynie ze sposobu oraz szybkości zmian częstotliwości podstawowej, przyjęto taką metodę która pozwoliła na pozbycie się pozostałych różnic w przebiegach parametru F_0 .

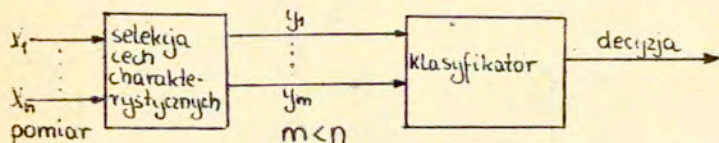
3.2. Selekcja charakterystycznych własności przebiegów F_0 .

Większość fizycznych sygnałów jest przedstawiana jako funkcje jednej lub dwóch zmiennych. W przypadku analizy częstotliwości podstawowej otrzymujemy zbiór ciągłych funkcji zależnych od czasu. Najprostszym sposobem pomiaru charakterystycznych cech sygnału jest próbkowanie go w określonych momentach czasowych t_1, \dots, t_n . W ten sposób każdy przebieg jest określony przez wektor w przestrzeni n-wymiarowej.



Ryc. 4. Pomiar wartości sygnału.

Wektor taki zawiera całą pomierzoną informację o danym obiekcie. Jest to bardzo dogodny sposób przedstawienia informacji ze względu na numeryczną analizę. W przypadku jednak dużej liczby n (punktów w których dokonywany jest pomiar) proces klasyfikacji i rozpoznawania komplikuje się. Dogodniejsze jest wówczas oparcie klasyfikacji na mniejszej liczbie danych - na cechach charakterystycznych rozpoznawanego przebiegu (ryc. 5).



Ryc. 5. Układ analizy sygnału.

Podstawowym zagadnieniem w każdym procesie rozpoznawania (przez człowieka czy też maszynę) jest selekcja cech charakterystycznych. Oczywiście jest, że liczba właściwości potrzebna do poprawnego rozwiązania problemu zależy od siły dyskryminacyjnej wybranych właściwości. Zagadnienie to jest skomplikowane przez fakt, że większość cech charakterystycznych nie jest łatwo i bezpośrednio mierzalna. Istotny jest więc wybór tych cech oraz ich ekstrakcja. W rozpoznawaniu przez maszynę zwykle używa się cech matematycznych, takich jak cechy statystyczne, współczynniki korelacji, wektory, wartości własne macierzy kowariancji itp. W odniesieniu do opisów przebiegów częstotliwości podstawowej analizowano różne cechy sygnału za pomocą bardzo dokładnych matematycznych modeli (por. [8], [23]) oraz nieprecyzyjnych metod opisowych.

Wszystkie matematyczne metody analizy parametru F_0 opierają się na sztywno przyjętym układzie współrzędnych (istnieje jedynie możliwość wyboru logarytmicznej lub liniowej skali częstotliwości oraz liniowej lub odcinkowo liniowej skali czasu). W przypadku modelowania złożonych przebiegów parametru F_0 opis komplikuje się do tego stopnia, że nie może być przydatny w procesie rozpoznawania lub klasyfikacji. Przynajmniej korzystne

byłoby przedstawienie przebiegów parametru F_0 nie w tradycyjnym układzie współrzędnych ale w takim układzie, który by uwidaczniał cechy charakterystyczne przebiegu. Należy przyjąć, że nie powinien to być układ skomplikowany a więc zawierać dwie do trzech współrzędnych. Pożądane jest wprowadzenie takiej transformacji, która umożliwi skoncentrowanie informacji na kilku pierwszych składowych i wyeliminuje te współrzędne w których wariancja jest mała. W takim kontekście transformacja polega na rozwinięciu badanego obiektu w zbiór wektorów bazowych i odrzuceniu części rozwinięcia w sposób zapewniający mały błąd.

3.3. Zastosowanie dyskretnej transformacji liniowej do selekcji cech.

3.3.1. Teoretyczne podstawy metody.

Optymalną metodą kompresji danych spośród wszystkich transformacji liniowych jest transformacja Karhunen-Loevego (K-L), ponieważ błąd całkowity (kwadrat różnicy między wektorem pierwotnym i wektorem o wyzerowanych składowych) dla układu K-L osiąga minimum. W praktycznych zastosowaniach używa się dyskretnej transformacji K-L. Wersję tą oraz wersję uogólnioną dla sygnałów ciągłych przedstawiono w pracach [15], [7], [20], [22]. Transformacja K-L polega na wyrażeniu funkcji $x_i(t)$ przez liniową kombinację funkcji $\phi_j(t)$ i współczynników c_{ij} . Zakłada się, że funkcje $\phi_j(t)$ tworzą zbiór ortonormalny w przedziale T_1, T_2 .

$$X_i(t) = \sum_{j=1}^n c_{ij} \phi_j(t) \quad (2)$$

Jeśli funkcje $x_i(t)$ są próbkowane równomiernie w przedziale T_1, T_2 to mogą one być przedstawione w formie wektorów.

$$\underline{X}_i = \begin{bmatrix} X_i(t_1) \\ X_i(t_2) \\ \vdots \\ X_i(t_n) \end{bmatrix} \quad (3)$$

gdzie n - liczba próbek w przedziale T_1, T_2
 $T_1 < t < T_2$

Równanie (2) sprowadza się wówczas do następującej postaci

$$\underline{X}_1 = \sum_{j=1}^n c_{1j} \phi_j \quad (4)$$

gdzie

$$\phi_j = \begin{bmatrix} \phi_j(t_1) \\ \phi_j(t_2) \\ \vdots \\ \phi_j(t_n) \end{bmatrix} \quad \text{oraz} \quad E \{c_{1j}\} = 0$$

Jeśli współczynniki c_{1j} przedstawione są w formie wektora

$$\underline{c}_{1j} = \begin{bmatrix} c_{11} \\ c_{12} \\ \vdots \\ c_{1n} \end{bmatrix}$$

to równanie (4) może być przedstawione w formie macierzowej

$$\underline{x}_1 = \phi \underline{c}_1 \quad \text{gdzie } \phi \text{ jest macierzą.} \quad (5)$$

Funkcja autokorelacji określona w zbiorze składającym się z M klas jest zdefiniowana następująco

$$R_{(t,s)} = \sum_{i=1}^M p(\omega_i) E \{x_{i(t)} x_{i(s)}\} \quad (6)$$

gdzie $p(\omega_i)$ jest prawdopodobieństwem pojawienia się obiektu z klasy ω_i .

$E \{x_{i(t)} x_{i(s)}\}$ jest wartością oczekiwaną z całego zbioru obserwacji.

Dyskretny analog funkcji autokorelacji określonej równaniem (6) jest zdefiniowany następująco

$$\underline{R} = \sum_{i=1}^M p(\omega_i) E (\underline{x}_i \underline{x}_i^T) \quad (7)$$

Po podstawieniu równania (5) do (7) otrzymujemy

$$\underline{R} = \sum_{i=1}^M p(\omega_i) E \{ \phi \underline{c}_i \underline{c}_i^T \phi^T \} = \phi \left\{ \sum_{i=1}^M p(\omega_i) E (\underline{c}_i \underline{c}_i^T) \right\} \phi^T \quad (8)$$

Jeśli założymy, że

$$\sum_{i=1}^m p(\omega_i) E(\underline{c}_i \underline{c}_i^T) = \underline{D}_\lambda \quad (9)$$

gdzie \underline{D}_λ jest macierzą diagonalną

$$\underline{D}_\lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \quad (10)$$

to równanie (8) redukuje się do postaci

$$\underline{R} = \underline{\Phi} \underline{D}_\lambda \underline{\Phi}^T \quad (11)$$

Przemnażając równanie (11) przez macierz $\underline{\Phi}$ otrzymujemy

$$\underline{R} \underline{\Phi} = \underline{\Phi} \underline{D}_\lambda \underline{\Phi}^T \underline{\Phi} \quad (12)$$

Kolumny macierzy $\underline{\Phi}$ tworzą układ ortonormalny tzn.

$$\underline{\Phi}_i^T \underline{\Phi}_j = \begin{cases} 1 & \text{jeśli } i = j \\ 0 & \text{jeśli } i \neq j \end{cases} \quad (13)$$

Równanie (12) upraszcza się do postaci

$$\underline{R} \underline{\Phi} = \underline{\Phi} \underline{D}_\lambda \quad (14)$$

Z równania (14) oraz z definicji wektorów i wartości własnych widać, że j -ty wektor własny użyty w rozwinięciu (4) jest wektorem własnym macierzy korelacji odpowiadającym j -tej wartości własnej.

Wektory własne są ortonormalne, stąd współczynniki rozwinięcia c_i mogą być obliczone następująco

$$\underline{\Phi} \underline{c}_i = \underline{x}_i \quad (15)$$

$$\underline{\Phi}^T \underline{\Phi} \underline{c}_i = \underline{\Phi}^T \underline{x}_i \quad (16)$$

$$\underline{c}_i = \underline{\Phi}^T \underline{x}_i \quad (17)$$

Optymalne rezultaty uzyskuje się, jeśli jest spełniony warunek $E\{a\} = 0$ tzn. jeśli wszystkie badane obiekty są scentralizowane. Założmy, iż wybieramy tylko $m (< n)$ składników c_i i chcemy dokonać estymacji wektora x . Wprowadzony wówczas do rozwinięcia

błąd będzie równy sumie n-m wartości własnych.

$$\underline{e} = \underline{x}_i - \sum_{j=1}^m c_{ij} \phi_j \quad (18)$$

Jeśli w rozwinięciu zastosuje się wszystkie wektory własne, to błąd będzie równy zeru. Informatywność każdej cechy jest określona przez odpowiednią wartość własną. Każda wartość własna równa jest wariancji zgodnie z zależnością

$$G_i^2 = \phi_i^T R \phi_i = \lambda_i \quad (19)$$

Jeśli więc wartości własne uporządkowane są w następujący sposób

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0,$$

to cechy odpowiadające tym wartościom własnym są uporządkowane w ten sam sposób. W transformacji K-L największe wariancje (wartości własne są związane z rzutami wariancji na osie współrzędnych) odpowiadają określonym kolumnom macierzy Φ (kolumnom przedstawiającym kierunki największej zmienności danych). Układ taki pozwala jednocześnie (po uporządkowaniu wartości własnych) stwierdzić, o ile poprzednia cecha jest lepsza - bardziej informatywna od poprzedniej. Drugą istotną cechą transformacji jest fakt, że cechy są nieskorelowane (macierz D_λ jest diagonalna). Stosując transformację do układu składowych głównych decydujemy, czy wybrać wektor własny odpowiadający danej wartości własnej. Oczywiście absolutna wielkość wartości własnej nie daje adekwatnej informacji. Dlatego używa się stosunku danej wartości własnej do sumy wszystkich wartości własnych (równej śladowi analizowanej macierzy)

$$J = \frac{\lambda_i}{\sum_{i=1}^n \lambda_i} = \frac{\lambda_i}{tr R} \quad (20)$$

Stosunek ten wyrażony w procentach określa błąd średniokwadratowy wprowadzony przez eliminację któregoś z wektorów własnych badanej macierzy. Wyróżnia się kilka typów metod selekcji informacji w których korzysta się z rozwinięcia K-L. Są to metody wykorzystujące informację zarówno o średnich jak i o wariancjach obserwacji z poszczególnych klas, bądź metody wykorzystujące jedynie wariancję (dokonuje się centralizacji przez wyzerowanie składowych wektorów średnich w klasach).

Różnice w poszczególnych metodach wynikają z różnych sposobów konstruowania macierzy kowariancji. Wybór określonego wariantu musi być uzasadniony fizycznym aspektem problemu. Ponieważ w pracy przyjęto założenie, że należy zbadać wyłącznie różnice w konturach przebiegów parametru F_0 , przyjęto metodę wykorzystującą informację jedynie o wariancji przebiegów. Wszystkie przebiegi zostały unormowane. Macierz kowariancji obliczono na podstawie przebiegów scentralizowanych ze względu na wszystkie klasy. W celu znalezienia cech przebiegów w optymalnym układzie współrzędnych wykonano następującą procedurę :

1. Normowanie przebiegów tak aby spełniona była zależność

$$\underline{X} \underline{X}^T = 1$$

2. Obliczenie wektora średniego

$$\underline{\mu} = E \underline{X} \quad (\text{uśrednianie odbywa się po całym zbiorze } X_N)$$

3. Obliczenie macierzy kowariancji

$$\underline{S} = E \left[(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})^T \right] \quad (21)$$

Obliczając macierz kowariancji dokonano zmiany skali polegającej na pomnożeniu różnic między poszczególnymi wektorami a wektorem średnim w celu uzyskania postaci macierzy dogodniejszej do przeprowadzania obliczeń. Przemnożenie takie nie zmienia układu wartości własnych a tylko sprowadza się do przemnożenia wszystkich wartości własnych przez tę wartość.

4. Obliczenie wartości i wektorów własnych.

5. Uporządkowanie wartości własnych według malejących wartości

$$\lambda_1 \geq \lambda_2 \geq \dots \geq 0$$

6. Obliczenie informatywności poszczególnych cech według wzoru(20).

7. Transformacja przebiegów do układu współrzędnych wektorów własnych odpowiadających największym wartościom własnym według zależności (17).

Centralizowania zbioru dokonano przez odjęcie od każdego wektora reprezentującego daną klasę wektora średniego, gdzie uśrednianie odbywa się po całym zbiorze uczącym. Mankamentem przekształcenia K-L, tak jak wszystkich metod opartych na bazie wektorów własnych macierzy są wysokie wymagania obliczeniowe. Zagadnieniu obliczania wektorów i wartości własnych poświęca się odrębne prace.

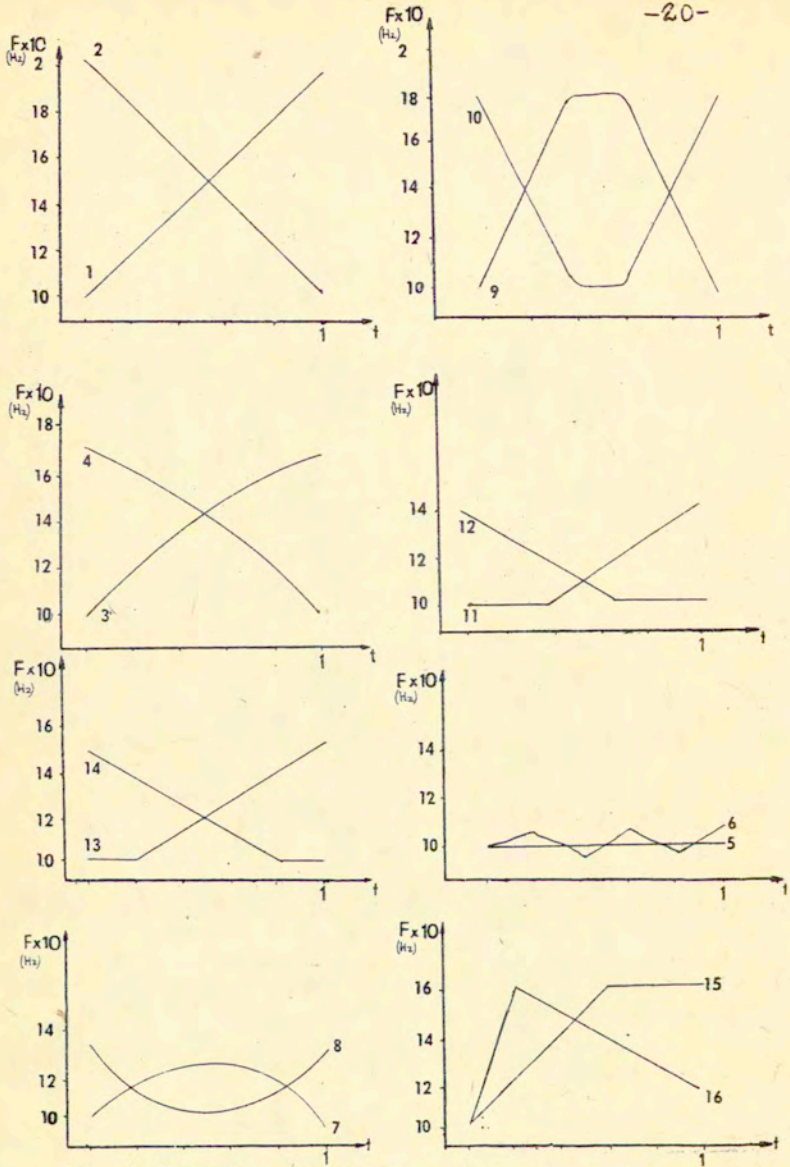
W przypadku dużej wartości próbek (dużej częstotliwości próbkowania) liczba obliczeń staje się bardzo duża i wyklucza użycie tej transformacji. W przypadku analizy częstotliwości podstawowej nie jest wymagana aż tak duża liczba próbek, co stwarza możliwości dla wykorzystania tej metody. Ałal w pracy [1] przedstawił metodę opartą na przebiegach czasowych częstotliwości podstawowej dokonując kompresji 500 danych reprezentujących każdy kontur F_0 do 10 danych w układzie współrzędnych Karhunen-Loevego. Spodziewać się należy, że zastosowanie tej metody umożliwi bardzo znaczną redukcję danych z minimalną utratą informacji o przebiegu. W celu sprawdzenia stopnia przydatności tej metody do analizy przebiegów częstotliwości podstawowej zbadano wstępnie przykład ilustrujący możliwości transformacji K-L w przypadku analizy zbioru dość zróżnicowanych przebiegów. Z uwagi na posiadany bardzo skromny sprzęt komputerowy przyjęto analizę przebiegów reprezentowanych przez 6 danych. Zastosowane algorytmy można zaadoptować do analizy większej liczby wymiarów. Metodę obliczania wektorów oraz wartości własnych przedstawiono w załączniku.

3.3.2. Transformacja przyjętych przykładowo przebiegów.

Przyjęto 16 fikcyjnych przebiegów dość często spotykanych w analizie częstotliwości podstawowej. Przebiegi te przedstawiono na ryc. 6. Dane reprezentujące poszczególne przebiegi zamieszczono w tabeli 1.

Tabela 1.

Numer przebiegu															
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
100	200	100	170	100	100	100	130	100	180	100	140	100	150	100	100
120	180	120	160	100	105	120	120	140	140	100	130	100	140	120	160
140	160	140	150	100	100	130	100	180	100	100	120	120	130	140	150
160	140	150	140	100	105	130	100	180	100	120	100	130	120	160	140
180	120	160	120	100	100	120	120	140	140	130	100	140	100	160	130
200	100	170	100	100	105	100	130	100	180	140	100	150	100	160	120



rys. 6. Przebiegi przyjęte do analizy matematycznej. Numery umieszczone na rycinach odpowiadają poszczególnym przebiegom.

Do układu K-L przetransformowano więc różnice między poszczególnymi wektorami a wektorem średnim. Różnice te dla poszczególnych przebiegów przedstawiono w tabeli 2.

Tabela 2.

Numer przebiegu							
1	2	3	4	5	6	7	8
-12.376	14.1543	-10.1657	9.9543	1.909	0.9043	-4.1257	6.3143
-8.669	7.261	-6.009	5.491	0.316	1.299	1.241	1.241
-3.359	1.959	-0.251	2.619	0.324	-0.681	4.729	-5.711
1.644	-3.646	2.3244	-0.5456	0.0294	1.012	4.434	-6.006
7.235	-8.695	5.466	-0.035	0.29	-0.715	1.215	1.215
12.573	-13.957	8.373	-11.747	0.328	1.311	-5.707	4.733

Tabela 2 c.d.

Numer przebiegu							
9	10	11	12	13	14	15	16
-10.3143	12.2043	-3.758	10.304	-6.196	10.162	-10.146	-8.6257
-0.749	-0.749	-5.351	5.197	-7.789	5.3	-5.983	7.951
10.615	-12.103	-5.343	1.688	-1.236	2.036	-0.221	4.929
10.3204	-12.3976	1.393	-5.6376	1.741	-1.531	5.234	1.6094
-0.778	-0.778	5.17	-5.377	5.274	-7.815	5.495	-1.155
-12.099	10.623	8.723	-5.339	8.583	-7.777	5.533	-4.147

Wektor średni obliczony z 16 analizowanych przebiegów ma postać

$$\bar{x} = \begin{bmatrix} 0.389157 \\ 0.40509 \\ 0.4052 \\ 0.407956 \\ 0.40536 \\ 0.40497 \end{bmatrix}$$

Obliczanie macierzy kowariancji według wzoru (21) zaprogramowano w języku BASIC na minikomputerze Sinclair. Dla przedstawionych różnic otrzymano następującą macierz kowariancji

$$S = \begin{bmatrix} 82.20601 & 31.0588 & -11.966 & -35.798 & -34.982 & -30.2158 \\ & 27.721 & 9.79 & -8.357 & -23.0414 & -35.753 \\ & & 24.885 & 17.7536 & -7.75 & -32.375 \\ & & & 25.484 & 9.01 & -7.509 \\ & & & & 23.44 & 33.95 \\ & & & & & 72.73 \end{bmatrix}$$

Ponieważ macierz ta jest symetryczna podano wartości znajdujące się po jednej stronie przekątnej. W celu obliczenia wartości oraz wektorów własnych powyższą macierz przekształcono do postaci trójdzielnej

82.20601	67.276043	38.220025				
67.276043	77.315021	90.367871	-9.0639066			
	38.220025		4.6005319	-0.21427006		
			-0.21427006	1.1263437	0.12492907	
				0.12492907	0.85023254	

Po powyższym przekształceniu suma elementów znajdujących się na głównej przekątnej nie ulega zmianie

$$\text{tr } S = 256.46607$$

Suma wartości własnych podanych macierzy powinna być równa

$$\sum_{i=1}^6 \lambda_i = 256.46607$$

Poniżej podano obliczone kolejno wartości własne

Wartości własne
158.380297707
89.081065449
6.3900111
1.20255
0.81385
0.59831
TR = 256.46607
I = 96.48893

Ostatecznie $\sum_{i=1}^6 \lambda_i = 256.46607$

W celu sprawdzenia czy wystarczające jest przyjęcie układu współrzędnych składających się z dwóch wektorów własnych odpowiednich dla całego zbioru przebieg obliczono stosunek

$$J = \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6} \cdot 100\% = 96.4889$$

Można przyjąć więc, że pominięcie wektorów własnych odpowiadających czterem pozostałym wartościom własnym nie wprowadza

istotnego błędu.

Dwie pierwsze wartości własne zawierają 96.5 % informacji o przebiegach. Wektory własne odpowiadające wartościom własnym λ_1 i λ_2 przedstawia poniższy wydruk

$$\text{dla } \lambda_1 = 158,38 \quad \text{dla } \lambda_2 = 89,08$$

Wektory własne

V1	V2
-0.60918278	0.49765729
-0.3804716	-0.082241435
-0.10599526	-0.50166607
0.16840348	-0.46847752
0.37835178	0.028729929
0.54899265	0.52338618

Zgodnie z wzorem (17) dokonano następującego przekształcenia

$$C_i = \Phi^T X_i$$

W wyniku otrzymano zbiór 16 dwuwymiarowych wektorów reprezentujących wektory z tabeli 2. Na ryc. 7 przedstawiono dwuwymiarowe wektory w nowym układzie współrzędnych V1, V2. W tabeli 3 podano składowe tych wektorów.

Tabela 3.

Numer przebiegu						
1	2	3	4	5	6	7
21.11	-23.159	15.56185	-17.255	-1.0227	-0.3533	-0.3868
2.2399	-0.368	-1.001	-2.8666	0.92836	0.87895	-9.55443

Tabela 3 c.d.

Numer przebiegu						
8	9	10	11	12	13	14
-1.66674	-0.36859	-2.417	11.87107	-14.34811	13.869	-15.4786
11.23356	-21.6891	23.5509	5.30102	3.55628	1.989	0.0324

Tabela 3 c.d.

Numer przebiegu	
15	16
14.4786	-0.7357
-3.857	-10.361

Wektory własne odpowiadające wartościom własnym λ_i i λ_j są ortogonalne

$$V_i^T V_j = 0$$

$$[-0.60918278 \quad -0.3804716 \quad -0.10599526 \quad 0.16840348 \quad 0.37835178 \quad 0.54899265]^*$$

$$x \begin{bmatrix} 0.49765729 \\ -0.082241435 \\ -0.50166607 \\ -0.46847752 \\ 0.028729929 \\ 0.52338618 \end{bmatrix} = 0$$

Każda wartość własna równa jest wariancji

$$\sigma_i^2 = \lambda_i \quad \begin{array}{l} \lambda_1 = 158.379 \\ \lambda_2 = 89.079 \end{array}$$

Dla pierwszej transformowanej zmiennej otrzymujemy

$$\begin{bmatrix} -0.609 & -0.3805 & -0.106 & 0.1684 & 0.3783 & 0.549 \end{bmatrix} x \begin{bmatrix} 82.206 & 31.0588 & -11.966 & -35.798 & -34.982 & -30.2158 \\ 31.059 & 27.72 & 9.79 & -8.357 & -23.041 & -35.753 \\ -11.966 & 9.79 & 24.885 & 17.754 & -7.75 & -32.375 \\ -35.798 & -8.357 & 17.754 & 25.484 & 9.01 & -7.509 \\ -34.982 & -23.04 & -7.75 & 9.01 & 23.44 & 33.95 \\ -30.215 & -35.75 & -32.37 & -7.509 & 33.95 & 72.73 \end{bmatrix} x \begin{bmatrix} -0.609 \\ -0.3805 \\ -0.106 \\ 0.1684 \\ 0.3783 \\ 0.549 \end{bmatrix} = 158.379$$

Analogicznie dla drugiej zmiennej

$$\begin{bmatrix} 0.4976 & -0.0802 & -0.50166 & -0.468 & 0.0287 & 0.5234 \end{bmatrix} x \begin{bmatrix} 82.206 & 31.0588 & -11.966 & -35.798 & -34.982 & -30.2158 \\ 31.059 & 27.72 & 9.79 & -8.357 & -23.041 & -35.753 \\ -11.966 & 9.79 & 24.885 & 17.754 & -7.75 & -32.375 \\ -35.798 & -8.357 & 17.754 & 25.484 & 9.01 & -7.509 \\ -34.982 & -23.04 & -7.75 & 9.01 & 23.44 & 33.95 \\ -30.215 & -35.75 & -32.37 & -7.509 & 33.95 & 72.73 \end{bmatrix} x \begin{bmatrix} 0.4976 \\ -0.0802 \\ -0.50166 \\ -0.4684 \\ 0.0287 \\ 0.5234 \end{bmatrix} = 89.079$$

Wariancję tę możemy obliczyć sumując wariancje w poszczególnych zmiennych w przetransformowanych przebiegach.

W prosty sposób można dokonać przekształcenia odwrotnego zgodnie z zależnością

$$\underline{X}_i = \underline{\Phi} \underline{C}_i$$

Przykładowo dokonano transformacji odwrotnej dla pierwszego przebiegu reprezentowanego w przestrzeni dwuwymiarowej następująco :

$$\underline{C}_i = \begin{bmatrix} 21,1103 \\ 2,2399 \end{bmatrix}$$

Zgodnie z równaniem (5) otrzymujemy

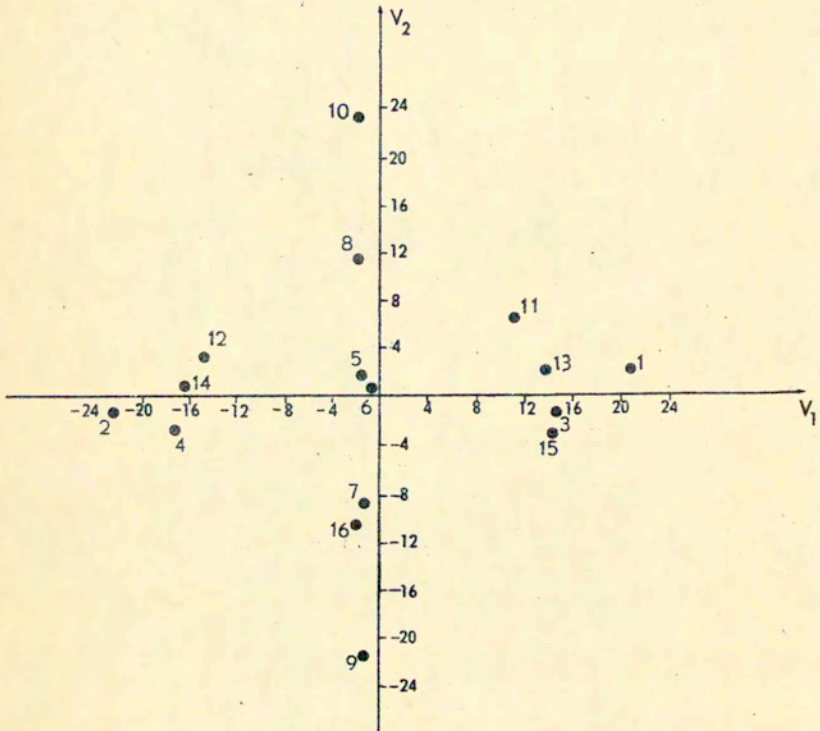
$$\begin{bmatrix} -0,609 & 0,4977 \\ -0,3805 & -0,0822 \\ -0,106 & -0,50166 \\ 0,168 & -0,4685 \\ 0,3783 & 0,0287 \\ 0,549 & 0,5234 \end{bmatrix} \times \begin{bmatrix} 21,1103 \\ 2,2399 \end{bmatrix} = \begin{bmatrix} -11,745 \\ -8,216 \\ -3,362 \\ 2,48 \\ 8,05 \\ 12,762 \end{bmatrix}$$

Porównując otrzymany wektor z wektorem pierwotnym \underline{X} równym

$$\underline{X} = \begin{bmatrix} -12,376 \\ -8,669 \\ -3,359 \\ 1,644 \\ 7,235 \\ 12,573 \end{bmatrix}$$

można zauważyć, iż wektor pierwotny jest reprezentowany (zgodnie z założeniem) z pewnym niewielkim błędem. Jeśli w transformacji odwrotnej wykorzystamy macierz \underline{S} złożoną z wszystkich sześciu wektorów własnych to otrzymamy dokładne wartości wektora pierwotnego, a błąd transformacji będzie równy zero. W omawianym przykładzie błąd spowodowany odrzuceniem 4 składowych jest niewielki i wynosi 3,6 %. Analogicznych przekształceń można dokonać na pozostałych wektorach.

Na ryc. 7 przedstawiono przetransformowane przebiegi w układzie współrzędnych złożonym z dwóch wektorów własnych V_1, V_2 z zachowaniem 96,4 % podobieństwa między nimi.



Ryc. 7. Przebiegi z ryc. 6 umieszczone w układzie współrzędnych składowych głównych. Numery umieszczone przy poszczególnych punktach oznaczają odpowiednie przetransformowane przebiegi.

Z ryciny widać, iż w bardzo prosty sposób badany zbiór został podzielony na 5 podzbiorów

- podzbiór przebiegów "jednokierunkowych" rosnących
- podzbiór przebiegów "jednokierunkowych" opadających
- podzbiór przebiegów dwukierunkowych rosnąco-opadających
- podzbiór przebiegów dwukierunkowych opadająco-rosnących
- podzbiór przebiegów równych.

Najbliższej osi V_1 (odpowiadającej 1 wektorowi własnemu) znajdują się przebiegi "jednokierunkowe", natomiast przebiegi "dwukierunkowe" znajdują się w pobliżu osi V_2 (odpowiadającej drugiemu wektorowi własnemu). Po przeciwnych stronach wykresu najbliższej osi V_1 znajdują się przebiegi rosnące (11,13,1) oraz opadające (12,14,4,2). Przebieg nr 3 znajduje się bliżej środka układu (ma mniejszy kąt nachylenia) niż przebieg 1. Po przeciwnych stronach wykresu najbliższej osi V_2 znajdują się przebiegi opadająco-rosnące (8,9) oraz rosnąco-opadające (7,16,9). Ponieważ dane były uprzednio scentralizowane więc miarą podobieństwa do wektora średniego między poszczególnymi przebiegami jest odległość od środka układu. Z ryciny widać, że przebiegi 10 i 9 oraz 1 i 2 najbardziej się różnią od przebiegu średniego, a przebiegi 5 i 6 są najbardziej podobne do przebiegu średniego. Zasadniczy wpływ na interpretację wektorów własnych ma sposób konstruowania macierzy kowariancji. Dla zilustrowania tego faktu przedstawiono poniżej prosty przykład.

Załóżmy, że chcemy dokonać rozwinięcia K-L następujących wektorów

$$X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad X_3 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad X_4 = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$$

Macierz kowariancji ma w tym przypadku postać

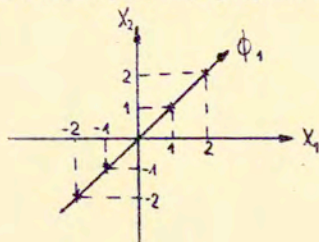
$$\underline{S} = \begin{bmatrix} \frac{10}{4} & \frac{10}{4} \\ \frac{10}{4} & \frac{10}{4} \end{bmatrix}$$

Jeśli wyeliminuje się wektor odpowiadający λ_2 , to błąd transformacji będzie równy zeru.

Otrzymujemy dwa wektory własne

$$\text{dla } \lambda_1 \quad \phi_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \text{dla } \lambda_2 \quad \phi_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Na ryc. 8 zilustrowano powyższy przykład.



Ryc. 8. Przykład transformacji K-L.

Dokonyjąc transformacji zgodnie z zależnością (17) otrzymujemy dla poszczególnych wektorów następujące wartości

$$c_1 = \begin{bmatrix} 2 \\ \sqrt{2} \end{bmatrix}, \quad c_2 = \begin{bmatrix} 4 \\ \sqrt{2} \end{bmatrix}, \quad c_3 = \begin{bmatrix} -2 \\ \sqrt{2} \end{bmatrix}, \quad c_4 = \begin{bmatrix} -4 \\ \sqrt{2} \end{bmatrix}$$

W celu sprawdzenia, czy błąd transformacji równa się zeru można dokonać przekształcenia odwrotnego zgodnie z równaniem (5)

$$x_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} 2 \\ \sqrt{2} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} 4 \\ \sqrt{2} \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Podobnie dla punktu x_3 oraz x_4 otrzymujemy błąd równy zeru.

Wynika z tego, iż analizowane wektory można bez błędów przedstawić za pomocą jednego wektora własnego. Należy zwrócić uwagę, że wzdłuż wektora własnego następuje teraz wyłącznie zmiana wartości średniej. Wybór określonego sposobu macierzy kowariancji wpływa na interpretację wyników.

3.3.3. Percepcyjne skalowanie przebiegów częstotliwości podstawowej.

Bardzo ciekawe jest porównanie wyników niniejszego przykładu (przedstawienia 16 przebiegów w postaci dwuwymiarowych wektorów)

z rezultatami pracy [9] opartej na badaniu percepcyjnych wymiarów tonu podstawowego. W pracy tej za pomocą skalowania wielowymiarowego opracowano doświadczenie polegające na badaniu percepcji tonu przez słuchacza. W eksperymencie wzięło udział 250 słuchaczy z różnych grup językowych. Materiał zawierał zbiór 19 przebiegów F_0 występujących w większości języków w świecie. Przebiegi te nałożone były na syntetyczne monosylaby. Były to 3 przebiegi rosnąco-opadające, 3 przebiegi opadająco-rosnące, 4 przebiegi rosnące, 4 opadające, 5 przebiegów równych, różniące się między sobą (w poszczególnych grupach) wartością średnią. Przebiegi były przedstawione w pięciopunktowej skali częstotliwości. Zadaniem pracy było zbadanie natury psychologicznego procesu stanowiącego podstawę percepcji tonu oraz określenie liczby i charakteru percepcyjnych wymiarów tonu. Poprzednio wykonane prace przez Gandoura sugerują stosowanie 5 percepcyjnych wymiarów :

1. wartość średnia
2. kierunek przebiegu
3. długość przebiegu
4. punkty ekstremalne
5. nachylenie przebiegu

Hombert (por. [10]) zaś w analizie wielowymiarowej sugeruje 2 wymiary - pierwszy odniesiony do kierunku przesuwania się zmian F_0 , drugi do wielkości spadku F_0 . W przytoczonej pracy [9] autor potwierdza statystyczną stabilność dwóch wymiarów oraz dochodzi do wniosku, iż dodatkowe wymiary nie wnoszą istotnych zmian i trudne są do interpretacji. Pierwszy wymiar autor zinterpretował jako oś, wzdłuż której następuje zmiana wartości średniej przebiegu F_0 , drugi wymiar jako oś, wzdłuż której następuje zmiana kierunku przebiegu. Porównując rezultaty percepcyjnego eksperymentu z rezultatami matematycznej analizy przyjętych tutaj 16 przebiegów można stwierdzić dużą zbieżność wyników. W obu pracach uznano za wystarczające przyjęcie 2 wymiarów (w pracy [9] wymiary wyznaczono na drodze percepcyjnego eksperymentu, w niniejszej pracy wymiary są wyprowadzone matematycznie). W obu pracach jeden wymiar zinterpretowano jako wymiar zmiany kierunku przebiegu. Drugi wymiar w pracy [9] został

zinterpretowany jako wymiar zmiany wartości średniej. W niniejszej pracy składową V_2 zinterpretowano jako oś przedstawiającą zmiany kąta nachylenia przebiegów "jednokierunkowych". Różnica ta głównie wypływa z faktu innego sformułowania problemu w przyjętym modelu matematycznym. Wynika ona głównie z dwóch powodów: z normalizacji przebiegów przed analizą (przebiegi różniące się jedynie wartością średnią uznane zostały za jednakowe - posiadają one taki sam kontur) oraz z centralizowania zbioru przez odjęcie wektora średniego. Na ryc. 7 przedstawiono odległości punktów reprezentujących poszczególne przebiegi od przebiegu średniego, który jest w punkcie zerowym układu. Po uwzględnieniu w matematycznym modelu nie tylko informacji o konturze, ale również o wartościach średnich można spodziewać się wyników analogicznych jak w pracy [9]. Ponieważ założono, iż w przyjętym tutaj modelu badane będą jedynie różnice w konturach przebiegu F_0 , wyniki transformacji przedstawione na ryc. 7 należy interpretować pod tym właśnie kątem.

4. Analiza przebiegów parametru F_0 .

4.1. Przygotowanie materiału do analizy.

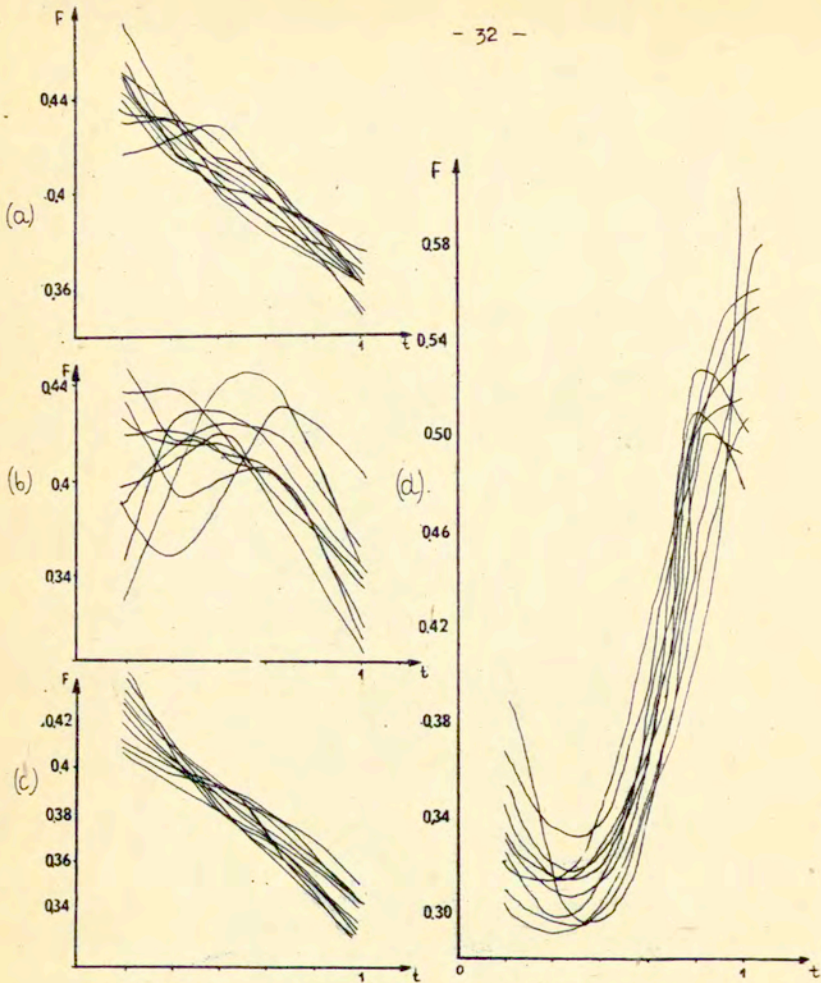
Porównując wizualnie tonogramy można zauważyć, że w pewnych przypadkach bardzo łatwo jest dokonać podziału przebiegów częstotliwości podstawowej w zdaniu na poszczególne odrębne jednostki intonacyjne, w innych przypadkach podział jest problematyczny. Najwygodniejsze byłoby wprowadzenie danych reprezentujących przebieg parametru F_0 w całym zdaniu do pamięci komputera i opracowanie algorytmów umożliwiających analizę tego parametru. Wiązałoby się to jednak z jednoczesną analizą dużej liczby danych, co w przypadku maszyny cyfrowej o małej pamięci jest kłopotliwe. Konieczne staje się podzielenie przebiegów na pewne fragmenty. Powstaje wówczas problem segmentacji przebiegów częstotliwości podstawowej. W niniejszej pracy przyjęto następujące zasady podziału: najmniejszą jednostką intonacyjną jest sylaba (por. [13]), zaś w przypadku prostszych przebiegów można przyjąć do analizy dwie lub więcej sylab, z tym że poszczególne jednostki powinny być ciągłymi funkcjami czasu (dość częste są nagłe zmiany częstotliwości przy przejściach od przebiegów samogłoskowych do spółgłoskowych następujące w szerokich granicach i w

związku z tym przebiegu czasowego nie można aproksymować krzywą ciągłą). W przyjętym materiale możliwe było przyjęcie prostego naturalnego podziału (por. ryc. 1 i 2). Przebieg częstotliwości podstawowej w zdaniu został podzielony na 4 części, na przebiegi na sylabie "pój", na sylabach "dziemy", na sylabie "do" i na sylabach "kina". Wybrane fragmenty analizowanego przebiegu w różnych powtórzeniach tego samego zdania miały niejednakową długość, natomiast procent czasu jaki zajmowały one w poszczególnych zdaniach był w przybliżeniu stały. Przykładowo dla jednej z osób przedstawiono wyniki pomiaru czasu poszczególnych fragmentów przebiegu w tabeli 4.

Tabela 4.

	Długość poszczególnych sylab (ms)			
	"pój"	"dziemy"	"do"	"kina"
1	110	258	149	200
2	160	275	148	179
3	110	240	125	240
4	112	288	138	194

W celu ustalenia liczby danych reprezentujących wybrane fragmenty przebiegu F_0 obliczono dla najbardziej zróżnicowanego przebiegu na sylabach "kina" wartość funkcji J (por. str.10) zgodnie z zależnością (1). Dla 6 danych $J = 0.019$, zaś dla 12 danych $J = 0.0048$. Przyjęto, że wystarczający będzie opis przebiegów za pomocą sześciu danych reprezentujących przebiegi na poszczególnych odcinkach. Rozwiązanie takie uniemożliwia wzajemne porównywanie fragmentów przebiegu w zdaniu, lecz pozwala na analizę przebiegu na tej samej sylabie (lub sylabach) w badanych zdaniach. Otrzymane dane dla poszczególnych fragmentów przebiegu częstotliwości podstawowej przedstawiono w Załączniku 2 w tabeli 9. Następnie zgodnie z procedurą podaną w punkcie 3.3.2. dokonano normowania wektorów oraz obliczono dla poszczególnych osób wektory średnie reprezentujące cztery przyjęte fragmenty przebiegu. Uśrednione przebiegi przedstawiono na ryc. 9. Z ryc. 9 widać, że przebiegi na sylabach "pój" i



Ryc. 9. Uformowane przebiegi częstotliwości podstawowej. Skala czasu względna.

- (a) przebiegi na sylabie "pój"
- (b) przebiegi na sylabach "dziemy"
- (c) przebiegi na sylabie "do"
- (d) przebiegi na sylabach "kina"

Na rycinie przedstawiono uśrednione przebiegi dla poszczególnych osób.

Obliczono wartości własne tej macierzy oraz wektory własne

Wartości własne

35.1317

14.90554

2.651

1.98305

1.1365

0.0065

TR = 55.81429

I = 89.649515

Wektory własne

V1

V2

-0.799

0.31182

0.06

-0.2967

0.187

-0.5772

-0.102

-0.215

0.291

0.1696

0.477

0.6375

Dla poszczególnych osób wektory przetransformowane do układu dwóch wektorów własnych przedstawiono w tabeli 5.

Tabela 5.

Transformacja przebiegów częstotliwości podstawowej.

1	2	3	4	5	6	7	8	9	10
-1.7	0.62	-3.3	1.14	1.32	-1.5	-0.88	0.77	0.72	-3.6
-0.5	2.14	-1.55	2.3	-1.18	-0.7	-0.54	-0.05	0.02	-0.08

Na ryc. 10a przedstawiono wynik transformacji.

4.2.2. Analiza sylab "dziemy".

Macierz kowariancji dla danych reprezentujących przebiegi F_0 na sylabach "dziemy" ma postać :

106.5609	18.2572	-14.693	-30.4305	-40.242	-26.86438
	37.4148	10.9568	-16.59	-21.77	-28,578
		12.951	1.891	-2.764	-11.899
			15.33	14.54	12.267
				23.698	23.524
					33.643

Powyższą macierz sprowadzono do postaci trójdzielnej

106.5609	61.776306						
61.776306	68.195655	35.30822					
	35.30822	41.303172	-4.438667				
		-4.438667	10.350645	1.04208			
			1.04208	1.4422886	-0.08969		
				-0.08969	1.742939		

Wartości i wektory własne macierzy mają postać

Wartości własne

156.10178

59.153

10.9405

1.764425

1.381

0.251

TR = 229.59171

I = 93.755479

Wektory własne

V1

V2

-0.766

0.4987

-0.302

-0.5769

0.0241

-0.4229

0.2689

0.0321

0.37084

0.1441

0.3342

0.4667

Przetransformowane wektory przedstawiono w tabeli 6.

Tabela 6
Transformacja przebiegów F_0 .

PJ	JI	WJ	GD	HK	ESz	PD	AM	DG	ES
1.48	5.95	6.42	2.24	-5.35	1.03	-1.99	-2.13	-5.37	-2.294
-1.3	5.07	-3.01	-3.1	-2.12	0.15	-0.007	0.57	1.18	2.6

Na ryc. 10c przedstawiono przetransformowane przebiegi.

4.2.3. Analiza sylaby "do"

Macierz kowariancji dla danych przedstawiających przebiegi na sylabie "do" ma postać

41.2728	-3.727	-2.4505	-8.7841	-20.8202	-14.8098
	4.4378	2.2998	0.4535	-0.1107	-3.8947
		3.791	-0.073	-2.343	-1.8537
			2.246	5.03366	3.30758
				14.456	9.4388
					13.221

Dla powyższej macierzy otrzymano następującą macierz trójdziagonalną

41.2728	-27.38369				
	22.92938	-6.546584			
	-6.546584	9.614627	-1.703722		
		-1.703722	4.7662176	-0.061825	
			-0.061825	0.563718	-0.03943
				-0.03943	0.277526

Wartości i wektory własne podanych macierzy mają postać

Wartości własne

61.26598

12.66829

4.64869

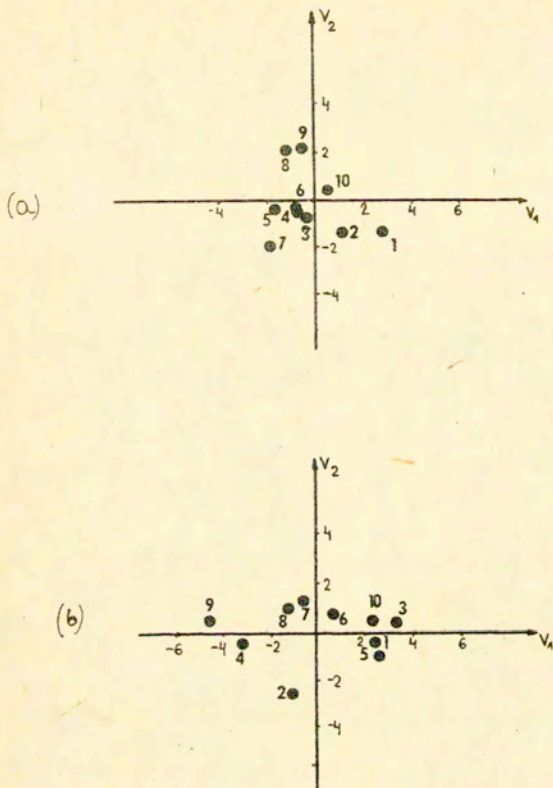
0.792

0.04

0.0065

TR = 79.42146

I = 93.091858



Ryc. 10. Przetransformowane przebiegi częstotliwości podstawowej

(a) sylaba "pój"

(b) sylaba "do"

Numerы umieszczone przy poszczególnych punktach oznaczają przetransformowane przebiegi dla każdej osoby

Dla JI oznaczono przebieg numerem				
"	WJ	"	"	1
"	GD	"	"	2
"	PJ	"	"	3
"	ESz	"	"	4
"	BS	"	"	5
"	DG	"	"	6
"	AM	"	"	7
"	PD	"	"	8
"	HK	"	"	9
				10

Wektory własne

V1	V2
-0.8054	0.32569
0.029	-0.5586
0.0059	-0.4933
0.1776	-0.025
0.4466	0.1769
0.3456	0.6073

Przetransformowane wektory przedstawiono w tabeli 7.

Tabela 7
Transformacja przebiegów F_0 .

PJ	JI	WJ	GD	HK	ESz	PD	AM	DG	ES
-3.232	2.318	-1.37	3.12	2.438	2.487	-4.365	-1.345	-0.6763	0.6296
-0.1612	-0.15	-2.9	0.23	0.098	-0.81	0.36	0.99	1.4	0.87

Na ryc. 10b przedstawiono przetransformowane przebiegi.

4.2.4. Analiza sylab "kina"

W tym przypadku otrzymano następującą postać macierzy kowariancji

72.256396	27.7151	8.3553	-18.59618	-15.3473	-43.16646
	21.5774	15.543	0.71538	-14.2999	-25.77805
		18.23547	13.1197	-9.5909	-23.688
			51.9836	38.758	-66.181
				78.8	-78.6546
					116.5105

Macierz tę sprowadzono do postaci macierzy trójdzielnej

72.256396	57.291356				
57.291356	88.58876	79.7003			
	79.7003	20.35795	45.91899		
		45.91899	50.022	7.5682	
			7.5682	3.9700142	-0.151
				-0.151	0.9462

Poniżej podano wartości i wektory własne tych macierzy

Wartości własne

256.22662

117.12848

43.88

5.639

0.967

0.421

TR = 424.2621

I = 88.001049

Wektory własne

V1

V2

-0.142 0.742

-0.089 0.346

-0.095 0.1584

-0.339 -0.2924

-0.4136 -0.4431

0.823 -0.1518

Przetransformowane wektory przedstawiono w tabeli 8.

Tabela 8

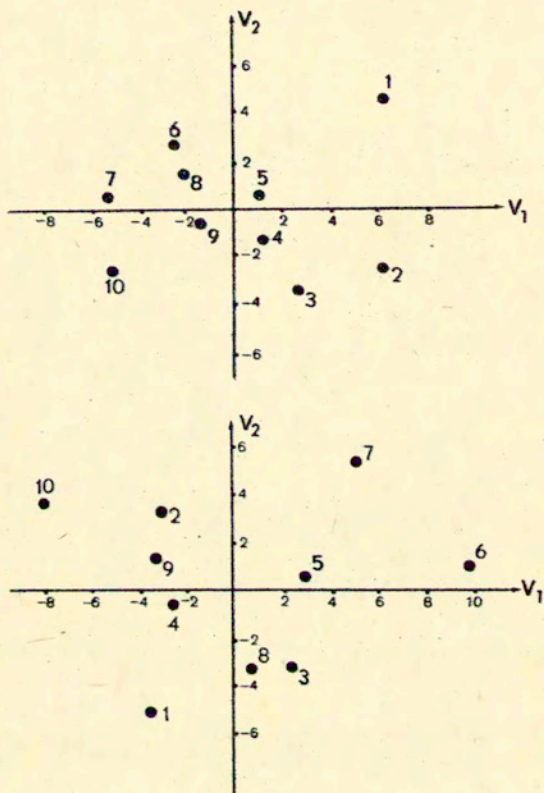
Transformacja przebiegów F_0

HK	ESz	PD	AM	PJ	JI	DG	WJ	BS	GD
-8.05	2.96	-3.87	0.78	-2.42	-3.83	4.98	-3.25	10.26	2.41
3.33	0.313	1.38	-3.3	-0.5	-5.25	5.42	3.32	0.9	-2.9

Na ryc. 10d przedstawiono przetransformowane przebiegi.

4.3. Ocena otrzymanych wyników.

We wszystkich przypadkach analizy fragmentów przebiegu parametru F_0 otrzymano zadowalające przedstawienie wektorów 6-wymiarowych w układzie dwuwymiarowym. Analizując przebiegi na sylabach "pój" oraz "do" można stwierdzić, iż zróżnicowania międzysobnicze są tam bardzo małe. Przebiegi na tych sylabach są opadające. Ponieważ sylaby te są stosunkowo krótkie (w porównaniu z innymi fragmentami przebiegu) można spodziewać się, że do ich opisu wystarczyłaby mniejsza liczba danych (między

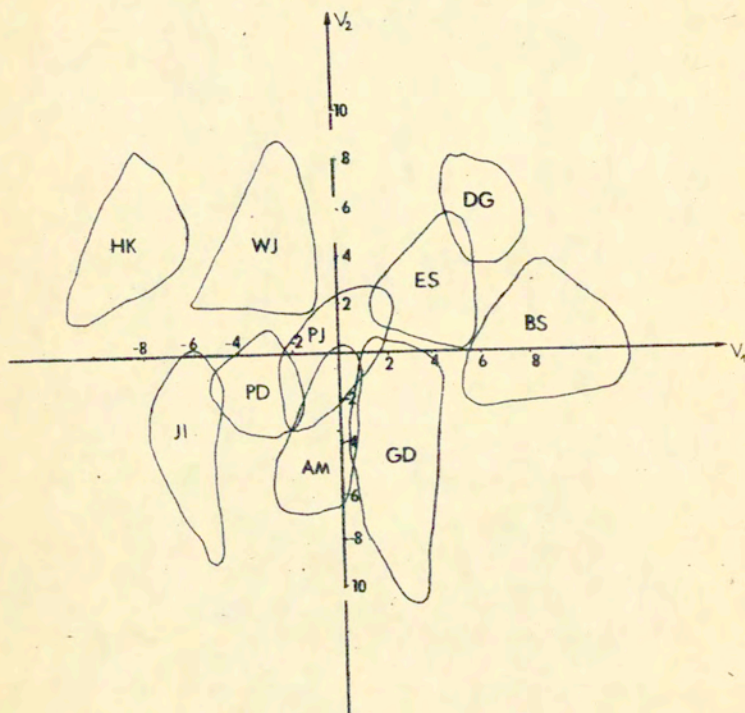


Ryc. 10. Przetransformowane przebiegi częstotliwości podstawowej
(c) na sylabach "dziemy"
(d) na sylabach "kina".

danymi istnieje silna korelacja). Otrzymane wartości własne będące wariancjami równe są $\lambda_1 = 35$ oraz $\lambda_2 = 61$. Analizując sylaby "dziemy" oraz "kina" można zauważyć duże osobnicze zróżnicowania przebiegów ($\lambda_1 = 156$, $\lambda_2 = 256$). Charakterystyczne jest to, że w obu przypadkach istnieją przebiegi częstotliwości pewnej grupy osób mało różniące się od przebiegu średniego (są to głosy AM, PD, ESz, PJ) oraz przebiegi grupy głosów wyraźnie zróżnicowanych. Macierz kowariancji była skonstruowana na podstawie wektorów średnich w poszczególnych fragmentach. Jeśli jednak dokonać transformacji nie tylko dla wektorów średnich lecz dla wszystkich wektorów reprezentujących dany fragment, otrzymamy zbiory punktów (5-ciu punktów dla każdej osoby) reprezentujące poszczególne głosy. Na ryc. 11 przedstawiono wyniki takiej transformacji dla sylab "kina". Można zauważyć, że niektóre obszary pokrywają się częściowo, inne są wyraźnie odległe od pozostałych. Ocena różnic między- i wewnątrzosobniczych zależy od sposobu konstrukcji macierzy kowariancji. Macierz kowariancji należy skonstruować tak, aby stosunek średniej odległości między klasami do średniego rozproszenia każdej klasy był jak największy (tzn. stosunek wariancji międzyklasowej do wewnątrzklasowej). Zastosowany sposób konstrukcji macierzy kowariancji przyjęto pod kątem optymalnego opisu danych, a nie optymalnej klasyfikacji. Dlatego też nie przeprowadzono statystycznych testów badających istotność zróżnicowań między głosami. Z ryc. 10 zróżnicowania międzyosobnicze są na tyle wyraźne, iż można ustalić zakres zmian przebiegów parametru F_0 dla poszczególnych osób.

5. Wnioski.

W odróżnieniu od innych opracowań przedstawiona metoda jest z jednej strony uniwersalna, z drugiej nie wymaga rozbudowanego systemu obliczeniowego, co umożliwia jej szeroką dostępność. Z przyjętych dość różnorodnych przykładów wynika, iż w sposób matematycznie uzasadniony możliwa jest znaczna redukcja danych w przebiegach parametru F_0 . Przytoczony przykład (por. [9]) ukazuje możliwość opisu przebiegów częstotliwości podstawowej w innym niż powszechnie stosowany układzie współrzędnych. W przypadku analizy podobnego zbioru przebiegów w



Ryc. 11. Transformacja przebiegów parametru F_0 na sylabach "kina".

niniejszej pracy oraz w pracy [9] ustalono układ 2 współrzędnych wystarczających do opisu przyjętych przebiegów. Zastosowany matematyczny model został maksymalnie uproszczony (co nie pozwoliło na optymalne rozwiązanie wszystkich problemów). Mimo to możliwa była dość szczegółowa analiza przebiegów. Możliwe było również przeprowadzenie analizy różnicowań międzypersonalnych. Można stwierdzić, iż nawet fragmentaryczne badanie konturu częstotliwości podstawowej dostarcza istotnych informacji w zakresie różnicowań międzypersonalnych. Możliwe jest zbadanie zakresu zmienności konturów częstotliwości podstawowej dla poszczególnych osób, co może mieć istotne znaczenie w badaniach klinicznych głosu i nauce intonacji osób głuchych. Należy również uznać przydatność takiej analizy w syntezie mowy i rozpoznawaniu głosów.

Załącznik I.

Obliczanie wartości własnych i wektorów własnych macierzy.

Sposób numerycznego wyznaczania wektorów i wartości własnych macierzy zależy od jej cech szczególnych. W ogólnym przypadku jest to zadanie dość skomplikowane. Istotne jest również, czy chcemy wyznaczać wszystkie wartości własne czy też tylko kilka z nich. W przypadku macierzy symetrycznej dodatnio określonej, posiadającej pierwiastki rzeczywiste dodatnie, można stosować dość dogodne algorytmy. Ponieważ w początkowej fazie pracy możliwe było opracowanie algorytmów jedynie na minikomputer MERA 303, konieczne było minimalizowanie wymiarów macierzy. Zastosowany algorytm można zaadoptować do obliczania wartości i wektorów własnych macierzy wyższego stopnia. Opracowane algorytmy służą do wyznaczania macierzy kowariancji 6 stopnia oraz odpowiadających jej wartości wektorów własnych. Algorytmy zaprogramowano w języku Basic na minikomputer Sinclair. Obliczenia przeprowadzono z dokładnością 4 cyfr po przecinku. Dokładność obliczania wartości własnych odgrywa dużą rolę w niektórych metodach obliczania wektorów własnych. W celu sprawdzenia poprawności obliczeń wykonano powtórnie obliczenia z dokładnością ośmiu cyfr po przecinku. Nie stwierdzono rozbieżności między wynikami obliczeń, a więc w przyjętej metodzie nie jest wymagane bardzo dokładne wyznaczanie wartości własnych. Wygodną metodą obliczania wartości własnych macierzy niezbyt wysokiego stopnia jest przekształcenie macierzy do postaci dogodniejszej (bez zmieniania jej wartości własnych) a następnie skorzystanie ze szczególnych cech tej macierzy. Jest to metoda przekształceń ortogonalnych. Dla dowolnej macierzy każde rozwiązanie równania

$$AX = \lambda X$$

spełnia

$$M AX = \lambda MX$$

oznaczając $MX = Y$

otrzymujemy

$$MAM^{-1}Y = \lambda Y$$

Jeśli więc λ i X są odpowiednio wartością własną i wektorem własnym macierzy A to λ i MX są odpowiednio wartością własną i wektorem własnym macierzy MAM^{-1} .

Należy dobrać M tak, aby uprościć iloczyn MAM^{-1} , konstruując macierz trójdagonalną (metoda Givensa).

Daną macierz przekształcamy za pomocą mnożenia przez macierz ortogonalną

$$\Omega_{ij} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & c & & s & \\ & & & \ddots & & \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} \begin{array}{l} \text{i-ty wiersz} \\ \text{j-ty wiersz} \end{array}$$

gdzie $c^2 + s^2 = 1$ (warunek ten zostanie spełniony jeśli przyjmiemy

$$c = \cos \Theta, \quad s = \sin \Theta)$$

Przekształcając macierz Ω_{ij} należy tak dobrać c i s aby uzyskać w rezultacie macierz trójdagonalną

$$V = \begin{bmatrix} \alpha_1 & \beta_1 & & & & \\ \beta_1 & \alpha_2 & \beta_2 & & & \\ & \beta_2 & \alpha_3 & \beta_3 & & \\ & & & \ddots & \ddots & \\ & & & & \beta_{n-2} & \alpha_{n-1} & \beta_{n-1} \\ & & & & & \beta_{n-1} & \alpha_n \end{bmatrix}$$

Łatwo jest wyznaczyć następujące równości dla elementów przekształcanej macierzy

$$a_{ip} = ca_{ip} + sa_{jp}$$

$$a_{jp} = -sa_{jp} + ca_{jp}$$

$$a_{pi} = ca_{pi} + sa_{pj}$$

$$a_{pj} = -sa_{pi} + ca_{pj}$$

Dla elementów znajdujących się na przecięciu wierszy i kolumn o indeksach i, j

$$a_{11} = c^2 a_{11} + sc a_{1j} + a_{j1} + s^2 a_{jj}$$

$$a_{jj} = c^2 a_{jj} - sc a_{1j} + a_{j1} + s^2 a_{11}$$

$$a_{1j} = sc a_{jj} - a_{11} + c^2 a_{1j} - s^2 a_{j1}$$

$$a_{j1} = sc a_{jj} - a_{11} + c^2 a_{j1} - s^2 a_{1j}$$

Przyjmując dla przykładu $i=2, j=3$ otrzymujemy dla elementu położonego w pierwszym wierszu i trzeciej kolumnie macierzy zależność

$$a_{13} = -s a_{12} + c a_{13}$$

gdzie

$$c = \cos \vartheta, \quad \text{tg } \vartheta = \frac{a_{13}}{a_{12}}$$
$$s = \sin \vartheta$$

Analogicznie obliczamy pozostałe elementy. Po wyznaczeniu macierzy trójdzielnej w prosty sposób można rozwiązać jej równanie charakterystyczne. W przypadku macierzy szóstego stopnia zadanie sprowadziło się do rozwiązania prostego układu równań

$$Q = \alpha_1 - \lambda_1$$

$$T = (\alpha_2 - \lambda_1) \cdot Q - \beta_1^2$$

$$S = (\alpha_3 - \lambda_1) \cdot T - \beta_2^2 \cdot Q$$

$$R = (\alpha_4 - \lambda_1) \cdot S - \beta_3^2 \cdot T$$

$$P = (\alpha_5 - \lambda_1) \cdot R - \beta_4^2 \cdot S$$

$$O = (\alpha_6 - \lambda_1) \cdot P - \beta_5^2 \cdot R$$

Ponieważ wiadomo, iż wszystkie wartości własne są dodatnie, a ich suma równa się śladowi macierzy to najprostszym rozwiązaniem jest podstawianie pod niewiadomą λ_1 kolejno wartości od zera do sumy wartości własnych i badanie przejść przez zero funkcji. Mając wyznaczone wartości własne macierzy trójdzielnej po przemnożeniu przez macierz transformacji można obliczyć

wektory własne macierzy Ω_{1j} . Obliczanie wektorów własnych macierzy trójdiagonalnej jest bardzo proste i sprowadza się do rozwiązania następujących równań

$$\begin{aligned}(\alpha_1 - \lambda_1) + \beta_1 y_2 &= 0 \\ \beta_1 + (\alpha_2 - \lambda_1) y_2 + \beta_2 y_3 &= 0 \\ \dots & \dots \\ \beta_{n-2} y_{n-2} + (\alpha_{n-1} - \lambda_1) y_{n-1} + \beta_{n-1} y_n &= 0\end{aligned}$$

Ostatnie równanie może posłużyć jako sprawdzenie :

$$\beta_{n-1} y_{n-1} + (\alpha_n - \lambda_1) y_n = 0$$

Bezpośrednio można otrzymać wektory własne przez obliczanie podwyznaczników danej macierzy. Należy wykonać wówczas następującą procedurę

- 1/ dla każdej wartości własnej należy obliczyć macierz $A - \lambda_1 I$ (odejmując λ_1 od elementów diagonalnych macierzy A)
- 2/ obliczenie podwyznaczników macierzy $A - \lambda_1 I$. Kolumny macierzy $A - \lambda_1 I$ są wzajemnie proporcjonalne.
- 3/ podzielenie elementów każdego wektora własnego przez pierwiastek z sumy kwadratów tych elementów.

Teoretyczne podstawy obliczania wartości i wektorów własnych przedstawiono w pracach ([6], [18], [21]).

Wybór określonej metody zależy również od możliwości obliczeniowych. Omawianą metodę przyjęto z uwagi na jej prostotę i możliwość zaprogramowania na prostym kalkulatorze. Ponieważ w przyjętym modelu matematycznym wyznaczanie wartości i wektorów własnych odgrywa istotną rolę, zastosowany algorytm musi być krótki oraz nieskomplikowany.

Załącznik 2.

Tabela 9.

Wyniki ekstrakcji parametru F w zdaniu "pójdziemy do kina?" przedstawione w postaci sześciu danych reprezentujących poszczególne fragmenty przebiegu. Inicjalami oznaczono osoby biorące udział w eksperymencie.

	HK					ESz					PD					AM				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
pó	162	172	163	145	160	263	217	254	218	221	110	138	113	115	121	124	105	94	95	123
	160	168	170	140	154	205	216	214	212	207	118	112	105	113	118	104	104	92	98	105
	161	157	160	133	148	205	216	211	200	201	118	99	98	110	113	105	98	93	97	103
	155	147	153	127	138	195	214	209	198	192	114	99	97	110	108	98	95	94	90	98
	147	142	144	130	131	182	193	200	185	173	111	94	109	111	102	95	93	97	89	96
	151	143	145	115	128	178	181	183	170	165	104	95	103	108	100	97	92	96	100	97
dziemy	178	163	171	162	152	228	240	245	204	222	104	106	116	106	117	103	107	100	100	108
	178	171	170	161	149	227	242	245	226	235	106	108	110	106	122	100	105	98	98	106
	172	165	165	153	145	237	245	254	237	234	105	102	107	107	118	98	108	100	97	107
	169	155	160	148	140	240	248	260	248	206	100	100	110	107	119	96	105	98	96	101
	148	131	156	140	129	243	218	233	232	206	96	93	106	99	115	92	100	95	90	94
	124	116	139	127	115	222	208	223	212	188	87	89	97	92	101	87	93	88	87	86
do	137	142	138	143	130	241	240	228	219	203	136	113	126	130	131	100	114	110	102	111
	144	137	142	142	123	227	243	238	232	205	104	101	114	107	130	95	100	112	94	91
	140	132	141	133	118	221	236	227	226	197	100	99	112	107	120	91	100	96	93	90
	135	132	137	129	117	219	230	223	220	193	101	90	111	100	118	90	98	96	93	87
	136	126	135	125	117	211	225	215	213	187	86	90	109	95	107	88	96	90	90	81
	126	124	128	124	103	197	221	197	191	172	81	83	100	95	99	84	90	87	84	78
kina	149	145	122	126	126	197	198	231	192	175	111	99	131	106	122	88	95	94	82	86
	130	125	120	116	116	181	182	200	173	167	98	95	106	99	102	85	92	90	89	85
	129	116	124	120	110	177	188	195	169	176	108	97	106	103	101	89	97	95	94	92
	142	132	142	160	130	210	200	209	192	238	111	111	111	127	127	111	111	130	114	119
	197	182	189	190	167	287	275	279	267	322	164	166	149	155	165	158	136	165	153	163
	191	166	170	168	159	358	311	357	305	330	170	167	154	159	169	177	157	179	170	155

Tabela 9.

	PJ					JI					WJ					GD				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
py	162	151	131	140	145	97	115	112	110	110	107	103	103	104	113	245	243	265	226	209
	133	149	126	130	139	120	114	109	110	108	115	103	101	103	110	232	240	240	217	208
	131	147	121	126	138	125	113	106	106	121	114	105	97	100	105	224	231	232	210	201
	130	144	115	117	130	114	108	97	102	100	108	97	96	95	100	217	217	218	200	189
	127	133	110	110	126	108	104	95	100	108	102	92	93	94	97	204	206	200	197	179
	116	121	101	108	119	100	102	85	104	98	96	88	96	97	72	197	200	190	189	174
dzl emy	132	137	118	126	144	131	153	131	137	129	110	101	117	132	112	212	206	207	201	205
	146	154	122	128	143	130	130	128	135	125	136	129	137	132	131	250	246	254	231	217
	156	161	125	127	136	144	141	136	143	133	154	140	145	143	147	237	244	244	249	211
	155	160	126	125	141	161	156	147	155	145	164	142	140	142	146	236	239	231	238	210
	143	154	121	119	134	153	154	155	152	138	160	125	126	135	137	230	250	193	216	202
	127	139	103	104	124	148	139	146	145	134	139	109	115	110	120	210	211	210	192	185
do	146	149	136	120	145	161	163	150	144	142	145	118	121	114	120	207	225	221	206	194
	131	139	108	116	131	158	153	151	145	142	145	116	121	113	114	218	230	214	197	197
	126	149	107	113	130	164	149	148	140	139	138	108	112	110	107	217	213	209	194	192
	104	130	99	113	132	159	150	143	137	136	101	109	114	115	107	211	208	207	193	187
	96	121	88	112	124	150	146	140	131	132	101	103	99	112	105	210	205	211	193	184
	100	119	87	111	113	145	130	120	136	126	90	93	95	89	92	196	189	199	177	167
kina	133	139	103	107	109	111	94	108	104	119	130	122	102	93	99	215	170	230	217	212
	117	140	101	102	108	104	93	98	96	109	97	81	83	87	80	197	194	203	186	179
	112	125	106	103	108	107	101	105	101	103	83	79	81	80	81	197	196	195	192	183
	122	166	140	140	123	152	132	145	149	125	91	98	104	99	102	265	256	262	231	217
	175	209	180	184	160	167	170	192	205	175	150	158	141	137	148	387	374	330	341	305
	205	202	174	181	173	162	167	185	197	174	152	155	137	145	137	416	388	389	374	340

Tabela 9.

	DG					BS				
	1	2	3	4	5	1	2	3	4	5
pó	241	264	240	215	233	236	231	247	234	220
	230	240	235	195	217	240	239	225	220	213
	228	220	240	193	213	233	234	214	220	214
	215	216	244	188	209	228	217	203	210	216
	202	200	241	167	186	214	205	188	194	201
	200	179	220	160	180	213	214	190	192	197
dzieny	240	293	252	260	230	300	255	260	293	260
	206	244	240	245	242	239	240	254	270	240
	206	247	242	250	243	238	248	278	280	240
	208	244	220	243	236	250	250	293	290	220
	200	240	204	225	210	241	227	265	260	210
	184	220	170	194	184	207	221	230	271	200
do	280	239	214	194	220	254	260	269	260	262
	202	230	210	190	206	240	260	244	252	240
	200	210	201	198	192	240	236	234	242	230
	207	206	200	198	190	241	240	236	244	220
	213	202	192	194	180	240	230	217	243	222
	184	192	180	183	162	220	218	230	200	206
kina	212	208	180	192	184	220	216	215	214	210
	192	184	174	172	168	208	192	190	186	200
	192	182	174	164	170	212	204	180	180	190
	204	190	203	177	191	265	259	191	193	200
	241	249	265	225	240	316	286	259	260	261
	320	328	303	311	316	407	489	357	394	400

BIBLIOGRAFIA

- 1 ATAL, B., Automatic Speaker Recognition Based on Pitch Contours, *JASA*, vol. 52, No. 6, 1972.
- 2 ABBERTON, E., FOURCIN, J., Intonation and speaker identification, *Language and Speech*, vol. 21, Part 4, 1978.
- 3 DEMENKO, G., Statystyczne własności rozkładów chwilowych wartości parametru F_0 w mowie ciągłej, *Prace IPPT* 31/1980.
- 4 DEMENKO, G., Normalizacja przebiegów częstotliwości podstawowej, *Prace IPPT* 26/1981.
- 5 FRANKS, L.E., *Teoria sygnałów*, PWN, Warszawa, 1975.
- 6 FORTUNA, Z., MACUKOW, B., WASOWSKI, J., *Metody numeryczne*, WNT, Warszawa, 1982.
- 7 FUKUNAGA, K., *Introduction to statistical pattern recognition*, Academic Press, New York, 1972.
- 8 FUJISAKI, H., Dynamic characteristics of Voice Fundamental Frequency in Speech and Singing, *The Fourth F.A.S.E. Symposium Contributed papers*, 57-69, Wenecja, 1981.
- 9 GANDOUR, J., Tone perception in Far Eastern languages, *Journal of Phonetics*, vol. 11, No. 2, 1983.
- 10 HOMBERT, J., Perception of tones of bisyllabic nouns in Yoruba, *Studies in African Linguistics*, Supplement 6, 109-121, 1976.
- 11 GUBRYNOWICZ, R., MIKIEL, W., ŻARNECKI, P., Acoustical analysis for the evaluation of laryngeal dysfunction in the case of vocal cord paralysis, *Speech Analysis and Synthesis*, vol. 5, PWN, Warszawa.
- 12 JASSEM, W., Parametry statystyczne rozkładów średnich wartości F_0 jako cechy osobnicze głosu, *Prace IPPT* 58/1970.
- 13 JASSEM, W., *Akcent języka polskiego*, *Prace Językoznawcze*, PAN, Wrocław, 1962.
- 14 JASSEM, W., DEMENKO, G., Normalization and Mathematical Description of Linguistically Relevant Pitch Curves, *The Fourth F.A.S.E. Symposium, Contributed papers*, 145-148, Wenecja, 1981.
- 15 KUBZDELA, H., Automatyczna ekstrakcja częstotliwości tonu podstawowego oraz pierwszych trzech formantów sygnału mowy, *Prace IPPT* 51/1973.

- 16 KRZYSKO, M., Analiza dyskryminacyjna, Prace Instytutu Matematyki, UAM, Poznań, 1982.
- 17 LUCHSINGER, R., ARNOLD, G.E., Voice. Speech. Language. Clinical Communicology : Its Physiology and Pathology, Belmont, 1965.
- 18 LEGRAS, J., Praktyczne metody analizy numerycznej, WNT, Warszawa, 1974.
- 19 MYTKOWSKI, K., Kanał funkcji analogowych KF-01 do wprowadzania i wyprowadzania informacji w systemie "on-line" do/z pamięci minikomputera Momik 8B/100, Prace IPPT 39/1976.
- 20 MEISEL, W., Computer-oriented approaches to pattern recognition, Academic Press, New York, 1972.
- 21 RALSTON, A., Wstęp do analizy numerycznej, PWN, Warszawa, 1975.
- 22 SOBCZAK, W., MALINA, W., Metody selekcji informacji, WNT, Warszawa, 1978.
- 23 STEFFEN-BATOG, M., JASSEM, W., GRUSZKA-KOSCIELAK, H., Statistical distribution of short-term F₀ values as a personal voice characteristic, Speech Analysis and Synthesis, vol. 2, PWN, Warszawa, 1970.
- 24 STYCZEK, I., Logopedia, PWN, Warszawa, 1979.
- 25 WICHERN, P., BOVES, L., Visual feedback of F₀ curves as an aid in learning intonation-contours, Proceedings of the Institute of Phonetics of University Nijmegen, Nijmegen, 1980.