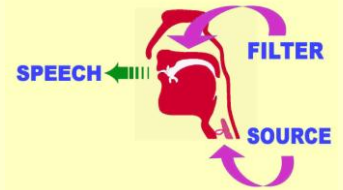


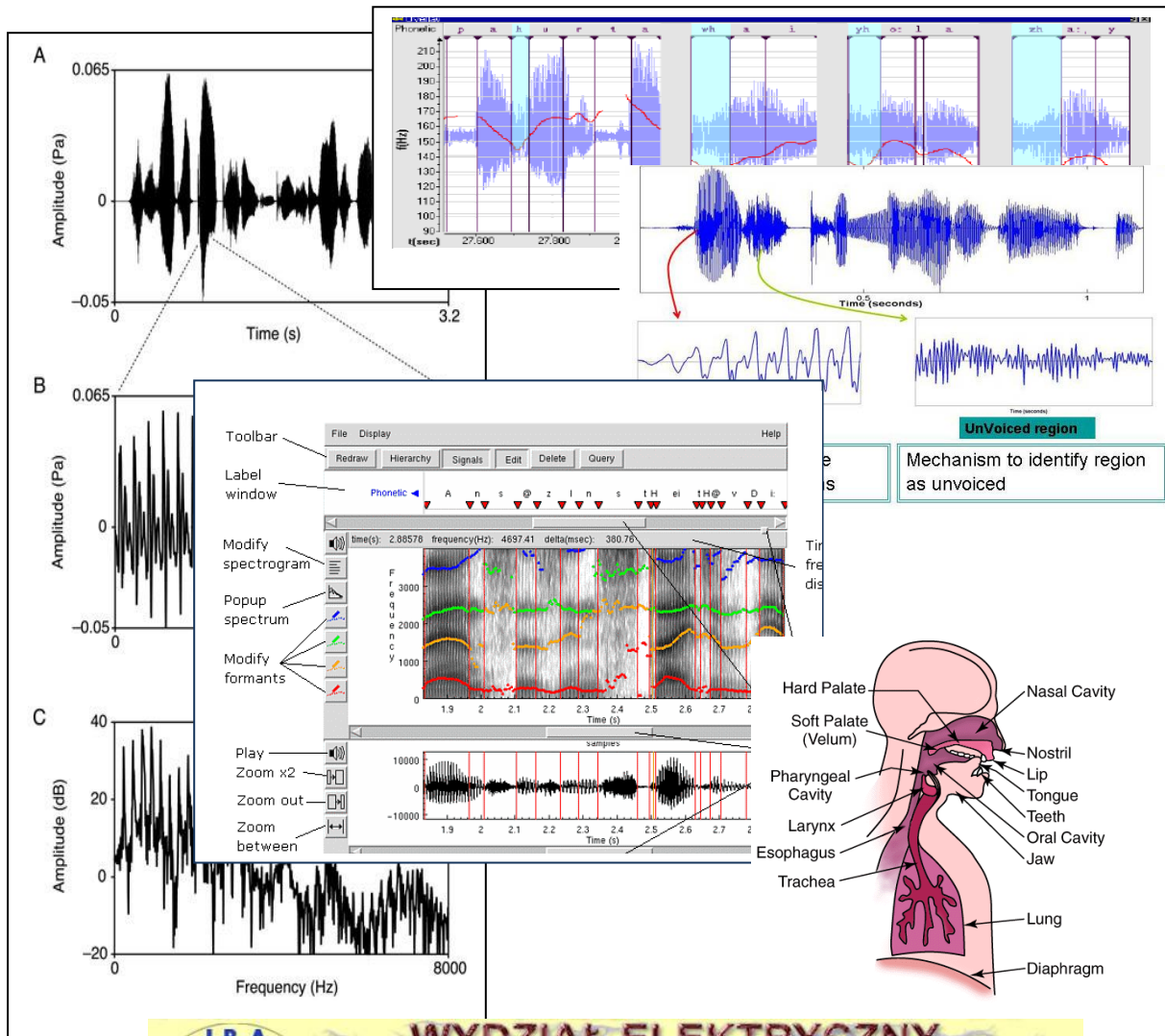
# SYGNAŁ MOWY

## LABORATORIUM



LABORATORIUM ANALIZY I PRZETWARZANIA SYGNAŁU MOWY – SALA C221

### HARMONOGRAM ORAZ INSTRUKCJE ĆWICZEŃ



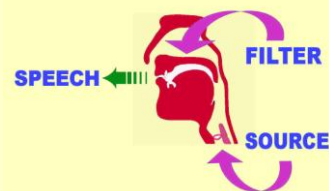
**IBA**

**WYDZIAŁ ELEKTRYCZNY**  
**POLITECHNIKI SZCZECIŃSKIEJ**  
**ZAKŁAD CYBERNETYKI I ELEKTRONIKI**  
**KIERUNEK : ELEKTRONIKA I TELEKOMUNIKACJA**  
**SPECJALNOŚĆ :**  
**Inżynieria Biomedyczna i Akustyczna**

Opracowanie: dr inż. Jerzy Sawicki  
 Szczecin 2008

# SYGNAŁ MOWY

## LABORATORIUM



## LABORATORIUM ANALIZY I PRZETWARZANIA SYGNAŁU MOWY – SALA C221

### WYKAZ ĆWICZEŃ LABORATORYJNYCH ORAZ HARMONOGRAM ZAJĘĆ

TEMATY ĆWICZEŃ LABORATORYJNYCH (w nawiasach symboliczne skróty)

1. Zapoznanie ze stanowiskami laboratoryjnymi, strukturą sprzętu i oprogramowania (BAZA).
2. Badanie wpływu częstotliwości próbkowania i rozdzielczości bitowej na zrozumiałość i subiektywnie odczuwaną jakość sygnału mowy (KOD).
3. Pomiar częstotliwości podstawowej głosu (F0).
4. Wygładzanie widma sygnału mowy (LPC).
5. Pomiar częstotliwości formantowych (FORMANT).
6. Rozpoznawanie dźwięków mowy na podstawie mapy formantów F1-F2 (ARM).
7. Synteza sygnału mowy (SYNTEZA).

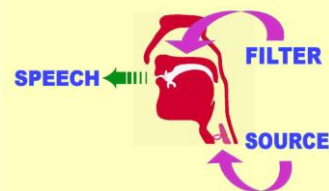
### HARMONOGRAM ĆWICZEŃ LABORATORYJNYCH

	Zajęcia 1	Zajęcia 2	Zajęcia 3	Zajęcia 4	Zajęcia 5	Zajęcia 6	Zajęcia 7
Zespół 1	BAZA	KOD	F0	LPC	FORMANT	ARM	SYNTEZA
Zespół 2							
Zespół 3							
Zespół 4							

**Zajęcia odbywają się co dwa tygodnie po dwie godziny.**

# SYGNAŁ MOWY

## LABORATORIUM



## LABORATORIUM ANALIZY I PRZETWARZANIA SYGNAŁU MOWY – SALA C221

## ĆWICZENIE NR 1 (BAZA)

### ZAPOZNANIE ZE STANOWISKAMI LABORATORYJNYMI, STRUKTURĄ SPRZĘTU I OPROGRAMOWANIA

WYMAGANA jest podstawowa wiedza z zakresu reprezentacji sygnału mowy w dziedzinie czasu i częstotliwości.

## PRZEBIEG ZAJĘĆ

Zapoznanie się z następującymi stanowiskami:

1. System edycji i analizy dźwięków *SonoLab* firmy Young Digital Poland.
2. Program do edycji i analizy sygnału mowy *MultiSpeech* firmy KAY-Pentax.
3. Program do analizy sygnału mowy *Praat* z Institute of Phonetic Sciences, Amsterdam.

Nauka podstawowych czynności na każdym ze stanowisk:

1. Ustawianie parametrów rejestracji sygnału z mikrofonu (częstotliwość próbkowania, ilość kanałów, preemfaza, czas rejestracji, wizualizacja przebiegu).
2. Zapisywanie i odczytywanie plików dźwiękowych.
3. Dodawanie uzupełniających informacji fonetycznych do pliku dźwiękowego.
4. Edycja przebiegu czasowego sygnału mowy.
5. Podstawowe operacje w dziedzinie czasu i częstotliwości (przycinanie pliku, filtracja).
6. Obserwacja przebiegów czasowych, widm i spektrogramów dla izolowanych samogłosek, spółgłosek szumowych i innych dźwięków. Porównanie poziomów sygnałów i szerokości pasma różnych dźwięków.
7. Obserwacja wpływu rodzaju okna czasowego na analizy widmowe sygnału mowy.

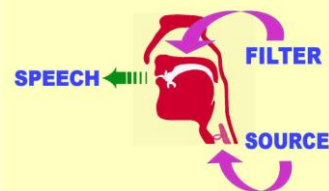
SPRAWOZDANIE Z ĆWICZENIA nie jest wymagane.

## LITERATURA

1. System edycji i analizy dźwięku *SonoLab* – instrukcja obsługi (dostępna do wglądu w sali C221).
2. *MultiSpeech* – instrukcja obsługi (dostępna do wglądu w sali C221).
3. *Praat* – program oraz instrukcja obsługi dostępne w Internecie: [www.fon.hum.uva.nl](http://www.fon.hum.uva.nl)

# SYGNAŁ MOWY

## LABORATORIUM



LABORATORIUM ANALIZY I PRZETWARZANIA SYGNAŁU MOWY – SALA C221

### ĆWICZENIE NR 2 (KOD)

#### BADANIE WPŁYWU CZĘSTOTLIWOŚCI PRÓBKOWANIA I ROZDZIELCZOŚCI BITOWEJ NA ZROZUMIAŁOŚĆ I SUBIEKTYWNIIE ODCZUWANĄ JAKOŚĆ SYGNAŁU MOWY

WYMAGANA jest znajomość zasad poprawnego próbkowania i kwantowania sygnałów.

#### PRZEBIEG ZAJĘĆ

1. Rejestracja za pomocą mikrofonu krótkiego komunikatu głosowego (np. data dzienna, godzina itp.) z maksymalną częstotliwością próbkowania na jaką pozwala dany system wprowadzania danych i zapisanie go w pliku.
2. Wykonywanie kolejnych filtracji i przepróbowań sygnału oraz odsłuchowa ocena utraty jakości i pogorszenia zrozumiałości w porównaniu z redukcją objętości informacyjnej sygnału w bitach/sekundę.
3. Obserwacja zjawiska aliasingu – włączanie i wyłączanie filtra antyaliasingowego.
4. Rejestracja za pomocą mikrofonu krótkiego komunikatu głosowego z częstotliwością próbkowania 32 kHz i zapisanie go w pliku. Należy zwrócić uwagę, by sygnał wykorzystywał większość dostępnych poziomów kwantowania.
5. Zmniejszanie rozdzielczości bitowej wszystkich próbek zarejestrowanego sygnału, przywracanie oryginalnego poziomu i odsłuchowa ocena utraty jakości i pogorszenia zrozumiałości w porównaniu z redukcją objętości informacyjnej sygnału w bitach/sekundę.
6. Rejestracja za pomocą mikrofonu krótkiego komunikatu głosowego z częstotliwością próbkowania 16 kHz, dodanie dwóch jego powtórzeń i zapisanie go w pliku. Należy znaleźć taką częstotliwość graniczną  $f_g$ , która podzieli zarejestrowany sygnał na dwie subiektywnie równoważne części: dolnoprzepustową (poniżej  $f_g$  – powtórzenie pierwsze) i górnoprzepustową (powyżej  $f_g$  – powtórzenie drugie). Należy zacząć od  $f_g = 4$  kHz, a następnie korygować wartość stosownie do subiektywnego wyniku. W przypadku zmiany poziomu sygnału po filtracji dokonać odpowiedniej korekcji wzmocnienia.

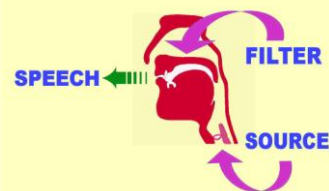
SPRAWOZDANIE Z ĆWICZENIA ma zawierać zwały opis przeprowadzonych badań oraz wnioski. W celu ilustracji można wykorzystać zapisane podczas badań kopie obrazu komputerowego.

#### LITERATURA

1. Tadeusiewicz R.: Sygnał mowy, WKiŁ, Warszawa 1988 (dostępna również w Internecie: <http://winntbg.bg.agh.edu.pl/skrypty/0004/>)

# SYGNAŁ MOWY

## LABORATORIUM



LABORATORIUM ANALIZY I PRZETWARZANIA SYGNAŁU MOWY – SALA C221

### ĆWICZENIE NR 3 (F0)

#### POMIARY CZĘSTOTLIWOŚCI PODSTAWOWEJ GŁOSU

WYMAGANA jest znajomość procesu wytwarzania mowy, modelu typu źródło-filtr, budowy akustycznej sygnału głosek dźwięcznych, relacji między przebiegiem czasowym a widmem sygnału okresowego.

#### PRZEBIEG ĆWICZENIA

1. Rejestracja za pomocą mikrofonu samogłosek i krótkich słów. Obserwacja struktury przebiegu czasowego. Znajdowanie granic kolejnych okresów częstotliwości podstawowej (tonu krtaniowego). Obserwacja widma jednego okresu tonu krtaniowego (zmiany poziomu harmonicznych, ilość harmonicznych).
2. Wykorzystanie sygnału błędu LPC do automatycznego wyznaczania granic okresu tonu krtaniowego w przebiegu czasowym.
3. Pomiar częstotliwości podstawowej głosów członków zespołu.
4. Zdolność maksymalnej zmiany częstotliwości podstawowej głosu – eksperyment z generowaniem bardzo niskich i bardzo wysokich częstotliwości podstawowych. Obliczanie przedziału zmian w oktawach.
5. Rejestracja za pomocą mikrofonu zdań pytających i ich zapisanie w pliku.
6. Automatyczne wyznaczanie dla zarejestrowanych zdań znaczników tonu krtaniowego oraz konturu intonacyjnego. Obserwacja pojawiających się błędów.

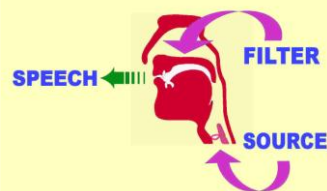
SPRAWOZDANIE Z ĆWICZENIA powinno zawierać wyniki pomiarów z oszacowaniem dokładności, porównanie otrzymanych wyników F0 z danymi statystycznymi, klasyfikację muzyczną wysokości głosu.

#### LITERATURA

1. Tadeusiewicz R.: Sygnał mowy, WKiŁ, Warszawa 1988 (dostępna również w Internecie: <http://winntbg.bg.agh.edu.pl/skrypty/0004/>)
2. Basztura Cz.: Źródła, sygnały i obrazy akustyczne, WKiŁ, Warszawa 1988.

# SYGNAŁ MOWY

## LABORATORIUM



LABORATORIUM ANALIZY I PRZETWARZANIA SYGNAŁU MOWY – SALA C221

### ĆWICZENIE NR 4 (LPC)

#### WYGŁADZANIE WIDMA SYGNAŁU MOWY

WYMAGANA jest znajomość podstaw teoretycznych techniki liniowej predykcji (LPC) oraz analizy kepstralnej.

#### PRZEBIEG ĆWICZENIA

1. Rejestracja za pomocą mikrofonu izolowanych samogłosek. Analiza LPC w różnych miejscach przebiegu czasowego samogłosek.
2. Badanie wpływu rzędu LPC na poprawność znajdowania formantów F1-F4.
3. Określenie wpływu preemfazy na wyniki analizy LPC.
4. Dla tych samych samogłosek wykonać analizę kepstralną, określić długość liftru i wygładzić widmo odrzucając szybkozmienną część kepstrum.
5. Porównać wyniki wygładzania widma za pomocą LPC i kepstrum.

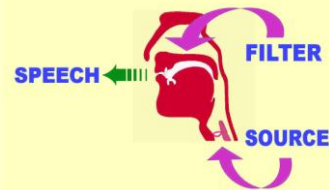
SPRAWOZDANIE Z ĆWICZENIA powinno zawierać wyniki zwarty opis przeprowadzonych badań z zestawieniem wyników częstotliwości formantowych otrzymanych w wyniku zastosowania obydwu rodzajów analiz.

#### LITERATURA

1. Tadeusiewicz R.: Sygnał mowy, WKiŁ, Warszawa 1988 (dostępna również w Internecie: <http://winntbg.bg.agh.edu.pl/skrypty/0004/>)
2. Basztura Cz.: Źródła, sygnały i obrazy akustyczne, WKiŁ, Warszawa 1988.
3. Bolc L., Borodziewicz W., Cytowski J., Jaszczak K.: Komputerowe przetwarzanie sygnału mowy. Metody. Technologia. Zastosowania, Wydawnictwa UW, Warszawa 1990.
4. Cytowski J.: Metody i algorytmy sztucznej inteligencji w cyfrowym przetwarzaniu sygnałów, Akademicka Oficyna Wydawnicza PLJ, Warszawa 1999

# SYGNAŁ MOWY

## LABORATORIUM



LABORATORIUM ANALIZY I PRZETWARZANIA SYGNAŁU MOWY – SALA C221

### ĆWICZENIE NR 5 (FORMANTY)

#### POMIAR CZĘSTOTLIWOŚCI FORMANTOWYCH

WYMAGANA jest znajomość podstaw teoretycznych techniki liniowej predykcji oraz umiejętność jej praktycznego wykorzystania.

#### PRZEBIEG ĆWICZENIA

1. Rejestracja za pomocą mikrofonu ( $f_p = 8$  kHz) sześciu podstawowych wypowiedzianych w izolacji samogłosek [a, e, i, o, u, y] przez każdego uczestnika ćwiczenia.
2. Wykonanie preemfazy i zachowanie materiału dźwiękowego w pliku.
3. Określenie wartości częstotliwości formantowych F1, F2, F3 i F4 dla każdej samogłoski (należy kontrolować poprawność identyfikacji formantów przez porównywanie z przeciętnymi wynikami publikowanymi w literaturze).
4. Powtórzyć punkty od 1 do 3 dla samogłosek wypowiedzianych w kontekście (zaproponować i wykorzystać wyrazy zawierające dużą liczbę różnych samogłosek).

SPRAWOZDANIE Z ĆWICZENIA powinno zawierać wyniki zwarty opis przeprowadzonych badań z zestawieniem wyników częstotliwości formantowych dla wszystkich mówców oraz wykres położenia samogłosek na mapie formantów w układzie współrzędnych F1-F2 z zaznaczeniem obszarów wspólnych dla tego samego typu samogłoski..

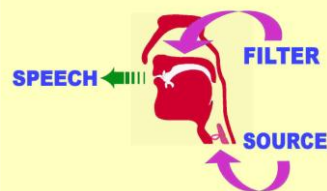
#### LITERATURA

1. Tadeusiewicz R.: Sygnał mowy, WKiŁ, Warszawa 1988 (dostępna również w Internecie: <http://winntbg.bg.agh.edu.pl/skrypty/0004/>)
2. Basztura Cz.: Źródła, sygnały i obrazy akustyczne, WKiŁ, Warszawa 1988.
3. Jassem W.: Podstawy fonetyki akustycznej, PWN, Warszawa 1973.
4. Jassem W.: Mowa a nauka o łączności, WKiŁ, Warszawa 1974.



# SYGNAŁ MOWY

## LABORATORIUM



LABORATORIUM ANALIZY I PRZETWARZANIA SYGNAŁU MOWY – SALA C221

### ĆWICZENIE NR 6 (ARM)

#### ROZPOZNAWANIE DŹWIĘKÓW MOWY NA PODSTAWIE MAPY FORMANTÓW F1-F2

WYMAGANA jest znajomość podstaw teoretycznych techniki liniowej predykcji oraz umiejętność jej praktycznego wykorzystania.

#### PRZEBIEG ĆWICZENIA

1. Wyłączyć sprzętowo (w sposób uniemożliwiający przypadkowe włączenie funkcji odsłuchu) zestawu głośnikowe wchodzące w skład stanowiska.
2. Wczytać wskazany przez prowadzącego dźwiękowy plik testowy z dysku.
3. Dla każdego dźwięku z pliku określić wartość częstotliwości podstawowej  $F_0$  i na tej podstawie ustalić płeć osoby mówiącej oraz metodą LPC częstotliwości formantowe  $F_1$  i  $F_2$  i za pomocą mapy formantów sporządzonej w ćwiczeniu laboratoryjnym nr 5 określić rodzaj samogłoski. Uwaga plik testowy może zawierać spółgłoski i dźwięki nie należące do sygnałów mowy (należy to odnotować i nie prowadzić dla nich dalszych analiz).
4. Włączyć zestaw głośnikowe i skonfrontować otrzymane wyniki z rzeczywistymi dźwiękami. W przypadku błędnych rozpoznań znaleźć przyczynę błędów.
5. Powtórzyć punkty od 1 do 4 dla kolejnych wskazanych przez prowadzącego zajęcia plików testowych.

SPRAWOZDANIE Z ĆWICZENIA powinno zawierać wyniki zwarty opis przeprowadzonych badań ze wskazaniem odsetka poprawnie rozpoznanych dźwięków w kolejnych testach oraz omówienie najważniejszych przyczyn błędnych rozpoznań. Wskazać trudności w zalgorytmizowaniu czynności wykonywanych podczas ćwiczenia.

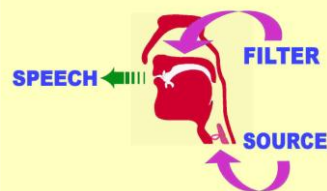
#### LITERATURA

1. Tadeusiewicz R.: Sygnał mowy, WKiŁ, Warszawa 1988 (dostępna również w Internecie: <http://winntbg.bg.agh.edu.pl/skrypty/0004/>)
2. Basztura Cz.: Źródła, sygnały i obrazy akustyczne, WKiŁ, Warszawa 1988.
3. Jassem W.: Podstawy fonetyki akustycznej, PWN, Warszawa 1973.
4. Jassem W.: Mowa a nauka o łączności, WKiŁ, Warszawa 1974.



# SYGNAŁ MOWY

## LABORATORIUM



LABORATORIUM ANALIZY I PRZETWARZANIA SYGNAŁU MOWY – SALA C221

### ĆWICZENIE NR 7 (SYNTEZA)

#### SYNTEZA SYGNAŁU MOWY

WYMAGANA jest znajomość podstaw teoretycznych syntezy mowy metodami kompilacyjnymi i formantowymi.

#### PRZEBIEG ĆWICZENIA

1. Zaplanować i przeprowadzić eksperymenty z konkatenacją wyodrębnionych fragmentów wyrazów (np. [MOWA]=[doMOfon]+[koWA]).
2. Zaplanować i przeprowadzić eksperymenty ze zmianą znaczenia zdań (np. [to jest nieprawda] → [to jest prawda]).
3. Przeprowadzić eksperyment z kompilacyjną syntezą liczebników typu *rdzeń + sufiks* (np. [jeden]+[naście], [dwa]+[naście]).
4. Przeprowadzić testy systemu syntezy mowy *EXPRESSIVO*. Znaleźć słabe punkty systemu syntezy (liczebniki, skróty, wtrącone wyrazy obcojęzyczne itd.).

SPRAWOZDANIE Z ĆWICZENIA powinno zawierać zwięzły opis przeprowadzonych eksperymentów oraz ocenę systemu syntezy *EXPRESSIVO*.

#### LITERATURA

1. Tadeusiewicz R.: Sygnał mowy, WKiŁ, Warszawa 1988 (dostępna również w Internecie: <http://winntbg.bg.agh.edu.pl/skrypty/0004/>)
2. Basztura Cz.: Źródła, sygnały i obrazy akustyczne, WKiŁ, Warszawa 1988.
3. Jassem W.: Podstawy fonetyki akustycznej, PWN, Warszawa 1973.
4. *EXPRESSIVO* – Twój osobisty lektor – strona internetowa [www.expressivo.com/pl/](http://www.expressivo.com/pl/)
5. Bolc L., Borodziewicz W., Cytowski J., Jaszczak K.: Komputerowe przetwarzanie sygnału mowy. Metody. Technologia. Zastosowania, Wydawnictwa UW, Warszawa 1990.
6. Cytowski J.: Metody i algorytmy sztucznej inteligencji w cyfrowym przetwarzaniu sygnałów, Akademicka Oficyna Wydawnicza PLJ, Warszawa 1999