

Wydział Elektryczny

Katedra Telekomunikacji i Aparatury Elektronicznej

Instrukcja do pracowni specjalistycznej

Temat ćwiczenia:

Badanie wybranych własności sygnału audio

Numer ćwiczenia: **1-2**

Pracownia Specjalistyczna z przedmiotu:

Techniki Multimedialne

Kod: **TS1C611258**

Opracował:
Dr hab. inż. Ewa Świercz

Białystok 2013

Pracownia Specjalistyczna 1-2 z przedmiotu Techniki Multimedialne

Temat: Badanie wybranych własności sygnału audio.

1. Cel pracowni: Badanie wpływu częstotliwości próbkowania na jakość sygnału mowy.
Badanie reprezentacji częstotliwościowej sygnału mowy –
spektrogramu (krótco-czasowa transformacja Fouriera – STFT).

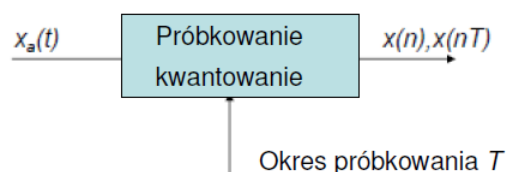
Cyfrowe przetwarzanie sygnałów w przetwarzaniu sygnału mowy obejmuje następujące zagadnienia:

- *filtracja* — kodowanie sygnału mowy, filtracja post, pre-filtracja, redukcja szumu
- *analiza spektralna* — projekty wokoderów, synteza mowy, rozpoznania mowy, rozpoznanie mówcy, wyostrenie sygnału mowy.
- *implementacja struktur* — systemy syntezy mowy, systemy analiza-synteza, struktury kodowania/dekodowania audio w MP3, AAC
- *konwersja próbkowania* — różne systemy audio wykorzystują różne częstotliwości próbkowania
 - Magnetofony DAT , telewizja HDTV — 48 kHz
 - Dyski CD Audio — 44.06 kHz
 - Mowa telekomunikacja — 6, 8, 11, 16 kHz
 - telefonia komórkowa — TDMA, GSM, CDMA
 - Dyski DVD Audio — 96,0 kHz

Najczęściej spotykane częstotliwości próbkowania: 44.1 kHz, 48 kHz, 96 kHz i 192 kHz. Sygnały o różnych częstotliwościach próbkowania często wymagają konwersji częstotliwości próbkowania.

Przetwarzanie wieloczęstotliwościowe (multirate processing) jest ważnym obszarem cyfrowego przetwarzania sygnałów dyskretnych, który nie ma odpowiednika w dziedzinie analogowej. Konwersja do różnych częstotliwości próbkowania ma na celu dopasowanie sygnału do określonego standardu, pojemności kanału transmisyjnego lub mocy obliczeniowej procesora DSP. Operacja zmniejszania częstotliwości próbkowania sygnału (o czynnik całkowity) nazywana jest decymacją. Jeżeli potrzebne jest zwiększenie częstotliwości przetwarzania, np. w celu dokładniejszej reprezentacji sygnału, stosuje się operację odwrotną, czyli interpolację. Niejednokrotnie sygnały audio pochodzące z różnych urządzeń są inaczej próbkowane.

Próbkowanie fali dźwiękowej



Rys. 1. Proces cyfryzacji sygnału analogowego

$T = 1/8000 \text{ sec} = 125 \mu\text{sec}$ 8 kHz częstotliwości próbkowania
 $T = 1/10000 = 100 \mu\text{sec}$ 10 kHz częstotliwości próbkowania
 $T = 1/16000 \text{ sec} = 67 \mu\text{sec}$ 20 kHz częstotliwości próbkowania

Twierdzenie o próbkowaniu

Jeżeli sygnał $x_a(t)$ ma ograniczoną pasmowo transformację Fouriera tak, że $X_a(j\Omega) = 0$ dla $|\Omega| \geq 2\pi F_N$, więc $x_a(t)$ może być jednoznacznie zrekonstruowany z próbek $x_a(nT)$, $-\infty < n < \infty$, jeżeli $1/T \geq 2 F_N$ ($F_s \geq 2F_N$)

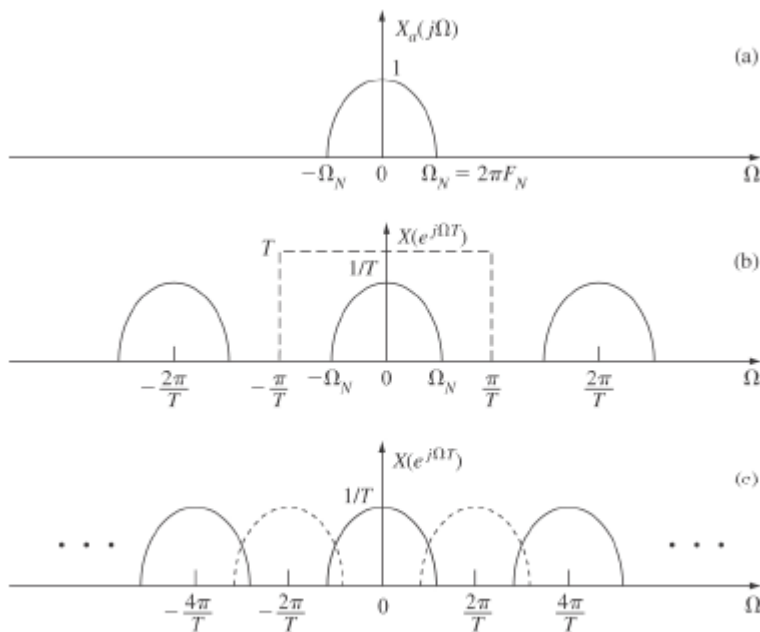
Równania próbkowania

$$x_a(t) \longleftrightarrow X_a(j\Omega) = \int_{-\infty}^{\infty} x_a(t) e^{-j\Omega t} dt$$

$$x[n] \longleftrightarrow X(e^{j\Omega T}) = \sum_{n=-\infty}^{\infty} x_a(nT) e^{-j\Omega nT}$$

$$X(e^{j\Omega T}) = \frac{1}{T} \sum_{k=-\infty}^{\infty} X_a(j\Omega + j2\pi k / T)$$

Interpretacja częstotliwościowa teorii próbkowania



Rys.2. Częstotliwościowa interpretacja częstotliwości próbkowania

Odzyskanie sygnału z próbek

Jeżeli $1/T > 2 F_N$ transformata Fouriera ciągu próbek jest proporcjonalna do transformaty Fouriera sygnału oryginalnego w przedziale podstawowym tzn.

$$\begin{aligned}
 2\pi / T - \Omega_N &> \Omega_N && \text{W celu uniknięcia zjawiska} \\
 \Rightarrow 2\pi / T &> 2\Omega_N && \text{aliasingu} \\
 \Rightarrow F_s = 1/T &> 2F_N && \text{konieczne jest spełnienie warunku}
 \end{aligned}$$

$$X(e^{j\Omega T}) = \frac{1}{T} X_a(j\Omega), \quad |\Omega| < \frac{\pi}{T}$$

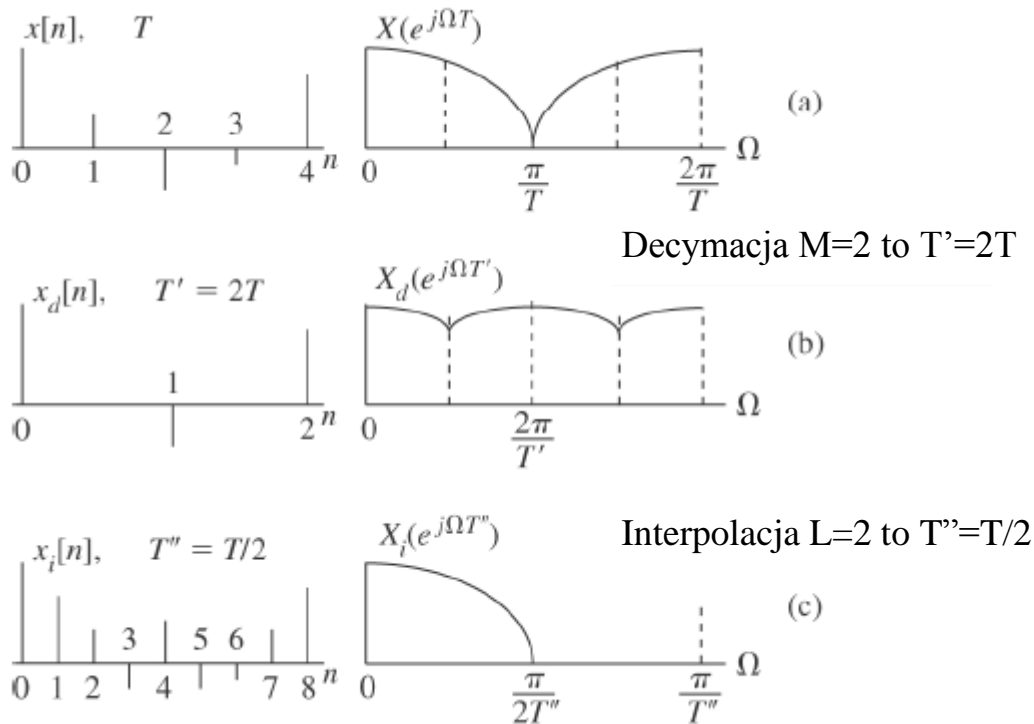
Odzyskanie sygnału z próbek jest zgodne z regułą

$$x_a(t) = \sum_{n=-\infty}^{\infty} x_a(nT) \left[\frac{\sin(\pi(t-nT)/T)}{\pi(t-nT)/T} \right]$$

Jest to równoważne filtracji dolnoprzepustowej (skrót LPF) filtrem o paśmie zawartym w przedziale π/T .

Często zachodzi potrzeba konwersji częstotliwości próbkowania w aplikacjach medialnych np. próbkowanie z aplikacji CD (44.06 kHz) do aplikacji DAT (48 kHz). Do konwersji wykorzystywane są operacje decymacji i interpolacji.

Operacja decymacji i interpolacji.



Rys.3 Graficzna interpretacja operacji decymacji (widmo ulega ściśnieniu) i interpolacji (widmo ulega rozciągnięciu).

Operacja decymacji jest wykorzystywana do redukcji częstotliwości próbkowania o czynnik M .

Operacja interpolacji jest wykorzystywana do zwiększenia częstotliwości próbkowania o czynnik L .

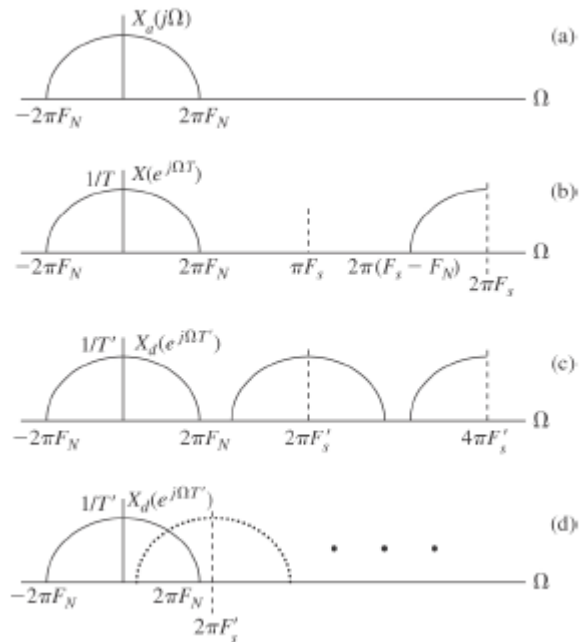
Decymacja powinna być wykonywana bez zjawiska aliasingu

$$x[n] = x_a(nT) \leftrightarrow X_a(j\Omega) = 0, \quad |\Omega| \geq 2\pi F_N \quad (a)$$

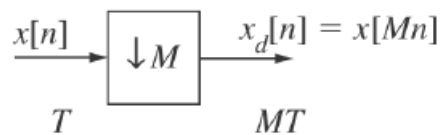
$$F_s = \frac{1}{T} \geq 2F_N$$

$$X(e^{j\Omega T}) = \frac{1}{T} X_a(j\Omega), \quad |\Omega| < \frac{\pi}{T} \quad (b)$$

$$X(e^{j\Omega T}) = 0, \quad 2\pi F_N \leq |\Omega| \leq 2\pi(F_s - F_N)$$



Należy zredukować częstotliwość próbkowania o czynnik $M \geq 2$ (zwiększyć okres próbkowania). Należy obliczyć nowy sygnał pozostawiając co M -tą próbkę sygnału oryginalnego



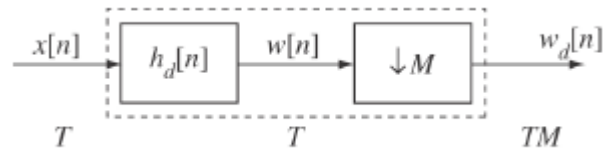
$$x_d[n] = x_a(nT')$$

$$F'_s = 1/T' = 1/(MT) = F_s / M \quad \text{Nowa zredukowana częstotliwość próbkowania}$$

$$H_d(e^{j\omega}) = \begin{cases} 1 & |\omega| < \pi / M \\ 0 & \pi / M < |\omega| \leq \pi \end{cases}$$

Ażby uzyskać widmo sygnału $x_d[n]$ bez aliasingu musimy zapewnić warunek, aby największa częstotliwość widma sygnału $x[n]$ była nie większa od $F_s / (2M)$.

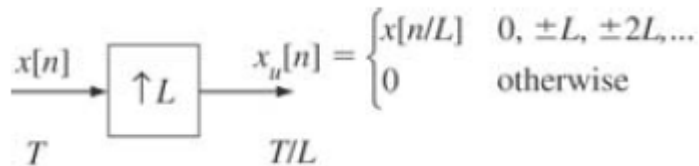
Wykonujemy wstępnie filtrację dolnoprzepustową filtrem o charakterystyce $H_d(e^{j\omega})$



ograniczając w ten sposób widmo sygnału podlegającego decymacji a następnie dokonujemy pobrania co M-tej próbki.

Operacja interpolacji

Sygnał oryginalny $x[n]$, próbkowany z okresem próbkowania, T , jest nadpróbkowany z nowym okresem T'' tworząc sygnał $x_u[n]$ z okresem próbkowania $T''=T/L$, czyli zwiększamy częstotliwość próbkowania o czynnik L .



Sygnał $x_u[n]$ powtarza próbki sygnału oryginalnego dla $n=0, \pm L, \pm 2L, \dots$. Pozostałe próbki są zerowe.

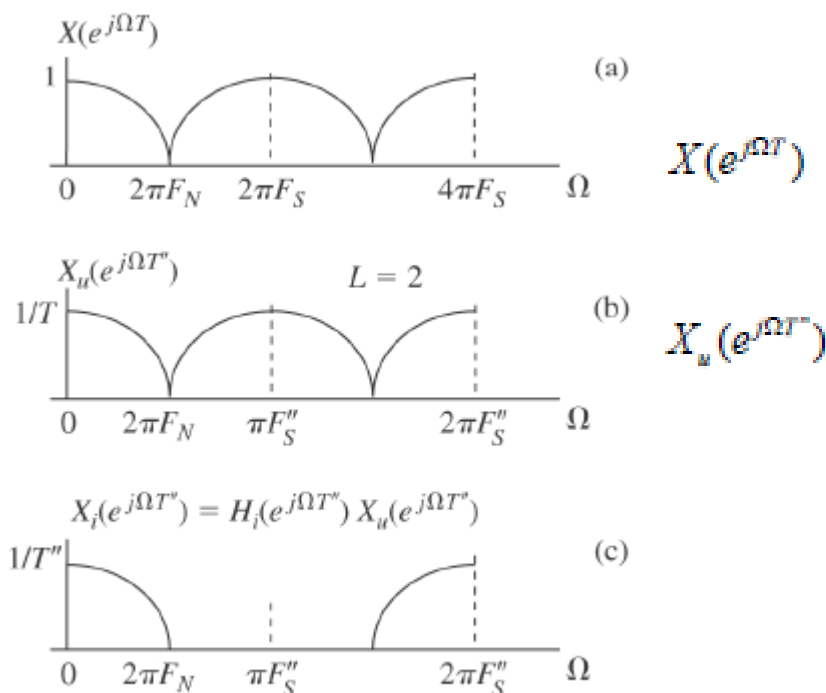
Transformata Fouriera sygnału $x_u[n]$ jest równa :

$$X_u(e^{j\omega}) = X(e^{j\omega L})$$

$$X_u(e^{j\Omega T''}) = X(e^{j\Omega T'' L}) = X(e^{j\Omega T})$$

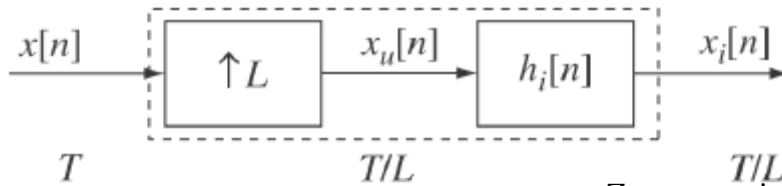
Oznacza to, że transformata $X_u(e^{j\Omega T''})$ ma dwa okresy powtarzania widma, wynikające z nadpróbkowania $2\pi/T''$ i próbkowania sygnału analogowego $2\pi/T$.

Dla ustawień $L = 2$, $T'' = T/2$ mamy



$$X_i(e^{j\omega}) = \begin{cases} (1/T^M) X(e^{j\omega L}) & 0 \leq \omega < \pi/L \\ 0 & \pi/L \leq \omega \leq \pi \end{cases}$$

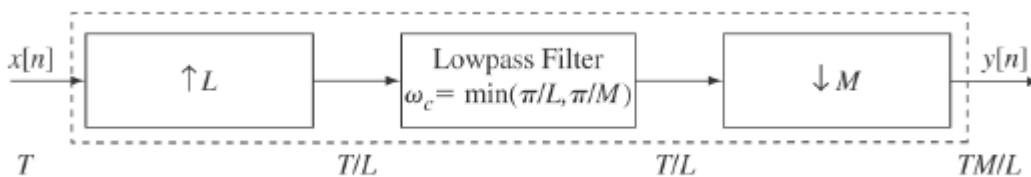
$$H_i(e^{j\omega}) = \begin{cases} L & |\omega| < \pi/L \\ 0 & \pi/L \leq |\omega| \leq \pi \end{cases}$$



Zastosowanie filtra dolnoprzepustowego o częstotliwości ucięcia $2\pi F_N = \pi/T$ pozwala na usunięcie replik widma sygnału oryginalnego.

$$x_i[n] = x_a(nT^M) = x_a(nT/L)$$

Jeżeli zachodzi potrzeba ułamkowego okresu próbkowania $T' = MT/L \Rightarrow$ to dokonujemy interpolacji o współczynnik L i decymacji o współczynnik M.



Zadania do wykonania.

1. Badanie procesu konwersji częstotliwości

1.1. Badania decymacji

Wprowadź plik 123.wav

Sygnal 123.wav jest próbkowany z częstotliwością 8000 kHz

Wykorzystaj polecenie

```
[x,FS,NBITS]=wavread('123.wav');
```

Utwórz AUDIO PLAYER obiekt dla sygnału x wykorzystując częstotliwość próbkowania FS.

```
xplayer = audioplayer(x, FS);
```

Wykorzystaj metodę play przypisaną do obiektu AUDIO PLAYER i odsłuchaj rozważany sygnał mowy

```
play(xplayer)
```

Zaprojektuj filtr dokonujący decymacji o współczynnik $M=2$, $M=4$, wykorzystując filtr decymujący obliczony poleceniem

```
hm = mfilt.firfracdecim(l,m,num)
```

- Wykreśl charakterystyki częstotliwościowe zaprojektowanych filtrów.
- Wykreśl sygnał oryginalny i sygnał po konwersji.
- Oceń czy następuje opóźnienie (zniekształcenie) wprowadzane przez proces filtracji dokonującej konwersji.

Pełna postać wywołania funkcji `hm = mfilt.firfracdecim(l,m,num)`, zawiera współczynniki filtru FIR zawarte w wektorze `num`. W wywołaniu funkcji bez argumentu `num`, filtr jest ustawiany domyślnie jako filtr dolnoprzepustowy z częstotliwością ucięcia pasma równą $\pi/\max(l,m)$.

Funkcja `mfilt.firfracdecim` pozwala na decymację o współczynnik ułamkowy gdzie l jest współczynnikiem interpolacji, m współczynnikiem decymacji, `num` jest wektorem zawierającym współczynniki filtru FIR dolnoprzepustowego wykorzystanego w decymacji. Specyfikując zarówno współczynnik interpolacji l jak i współczynnik m decymacji można dokonać dowolnej decymacji. Wymierny współczynnik decymacji wymaga najpierw interpolacji a następnie pożądanej wartości decymacji aby uzyskać właściwą częstotliwość próbkowania. Należy ustawiać l mniejsze niż m .

1.2. Załóżmy, że należy nadać do odbiorców dźwięk zapisany z częstotliwością próbkowania 48 kHz a urządzenia nadawcze są przystosowane do przesyłania dźwięku z częstotliwością 32 kHz. Dokonaj właściwej konwersji.

Wprowadź plik `audio48`
`load audio48;`

`Audio48` jest plikiem z rozszerzeniem `.mat`. i zawiera `signal48kHz` o częstotliwości próbkowania 48 kHz.

Odsłuchaj wprowadzony sygnał.
`p48 = audioplayer(signal48kHz,Fs48);`
`play(p48);`

Zaprojektuj proces konwersji z 48 kHz do 32 kHz.

- Wykreśl charakterystyki częstotliwościowe i czasowe filtru FIR dokonującego właściwej konwersji.
- Wykreśl sygnał oryginalny ($FS=48$ kHz) i sygnał po konwersji ($FS=32$ kHz).
- Oceń czy następuje opóźnienie (zniekształcenie) wprowadzane przez proces filtracji dokonującej konwersji.

Odsłuchaj sygnał po konwersji częstotliwościowej za pomocą audio player, `play`.

1.3. Badania interpolacji.

Dokonaj konwersji zwiększającej częstotliwość próbkowania z $FS=2$ kHz do $FS=16$ kHz dla sygnału `123.wav`.

- Wykreśl sygnał oryginalny (FS=2 kHz) i sygnał po konwersji (FS 16 kHz).
- Oceń czy następuje opóźnienie (zniekształcenie) wprowadzane przez proces filtracji dokonującej konwersji.

2. Ocena procesu maskowania tonów

Wprowadź pliki sin_200.wav, sin_220.wav, sin_300.wav.

```
[x,Fs,N]=wavread('sin_200.wav');
[x2,Fs,N]=wavread('sin_300.wav');
x2=x1/15;
```

Sygnał x jest traktowany jako masker, sygnał x2 jest sygnałem maskowanym.
Odsłuchaj sygnały:

```
wavplay(x)
wavplay(x2)
wavplay(x+x2)
```

- Utwórz obiekty do sygnałów x, x+x₂ i odsłuchaj za pomocą funkcji play.
- Oceń czy występuje zjawisko maskowania sygnałów.

3. Obrazowa reprezentacja częstotliwościowa sygnału audio - STFT.

3.1 Wykorzystaj sygnał e8.wav c8.wav.

```
Okno1=hamming(60)
Okno2=hamming(200)
```

Wyznacz STFT (spektrogram) dla wczytanych głosek. W Matlabie w toolbox-ie Signal Processing mamy funkcję SPECTROGRAM

```
S = SPECTROGRAM(X)
S = SPECTROGRAM(X,WINDOW)
[S,F,T] = SPECTROGRAM(...)
S = SPECTROGRAM(X,WINDOW,NOVERLAP)
S = SPECTROGRAM(X,WINDOW,NOVERLAP,NFFT)
S = SPECTROGRAM(X,WINDOW,NOVERLAP,NFFT,Fs)
[S,F,T] = SPECTROGRAM(...)
```

Jaką informację o sygnale mowy można uzyskać ze spektrogramu przy dwóch różnych długościach okna?

3.2 wyznacz pitch sygnału mowy (Częstotliwość tonu krtaniowego F_0) wykorzystując FFT głoski dźwięcznej. Ton krtaniowy jest pierwszym maksimum widma o ile następne maksima są wielokrotnościami pierwszego maksimum).

Wyznacz STFT dla sygnału akustycznego niebędącego sygnałem mowy np. SOUND24.wav i skomentuj różnice spektrogramu dla sygnału mowy i sygnału akustycznego.

Literatura:

- [1] V. Oppenheim, R. W. Schafer: Cyfrowe przetwarzanie sygnałów, Oppenheim WKŁ, Warszawa, 1979.
- [2] A. V. Oppenheim (red): Sygnały cyfrowe, przetwarzanie i zastosowania. WNT, 1982
- [3] Rabiner, L.R., and R.W. Schafer, Digital Processing of Speech Signals, Prentice-Hall, Englewood Cliffs, NJ, 1978
- [4] Strona internetowa kursu L. Rabinera Digital Speech processing, <http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/>.
- [5] B. Mrozek, Z. Mrozek, Matlab uniwersalne środowisko do obliczeń naukowo-technicznych. Wprowadzenie do programowania, Wyd. Go, Kraków 1994.
- [6] T. Zieliński, Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań, WKŁ, Warszawa 2005.
- [7] Matlab, Toolbox: Signal Processing, Filter Design.