Instrukcja do laboratorium 5
Fonoskopia

Metody opisu sygnału mowy

Parametry czasowe, częstotliwościowe i czasowo-częstotliwościowe w ocenie sygnału akustycznego mowy.

1. Parametry sygnałów deterministycznych (dla sygnału *x(t)* będącego ciągłą funkcją czasu):

   a) pochodna sygnału: $x_p = \frac{dx(t)}{dt}$

   b) całka sygnału: $x_c = \int_{-\infty}^{+\infty} x(t)dt$

   c) wartość średnia sygnału dla sygnałów ograniczonych w przedziale <t_1,t_2>:

   $x_{śr} = \frac{\int_{t_1}^{t_2} x(t)dt}{t_2 - t_1}$ lub $x_{śr} = \frac{\int_{t_1}^{t_2} |x(t)|dt}{t_2 - t_1}$

   d) energia sygnału: $E_x = \int_{-\infty}^{+\infty} x^2(t)dt$

   e) moc średnia sygnału dla sygnałów ograniczonych w przedziale <t_1,t_2>:

   $P_x(t_1, t_2) = \frac{\int_{t_1}^{t_2} x^2(t)dt}{t_2 - t_1}$

   f) wartość skuteczna sygnału (**R**oot **M**ean **S**quare): $P_{sk} = P_{RMS} = \sqrt{P_x}$

   g) funkcja autokorelacji sygnału: $R_{xx}(\tau) = \int_{-\infty}^{+\infty} x(t)x^*(t-\tau)dt$

   h) funkcja korelacji wzajemnej sygnału *x(t)* i *y(t)*: $R_{xy}(\tau) = \int_{-\infty}^{+\infty} x(t)y(t-\tau)dt$

2. Parametry sygnałów deterministycznych (dla sygnału dyskretnego *x*, $x \in \langle 0, N-1 \rangle$ o częstotliwości próbkowania f_p):

   a) pochodna sygnału: $x_p = \frac{x_{n+1} - x_n}{\Delta t}$, $\Delta t = \frac{1}{f_p}$

   b) całka sygnału: $x_c = \Delta t \sum_{n=0}^{N-1} x_n$, $\Delta t = \frac{1}{f_p}$

   c) wartość średnia sygnału dla sygnałów ograniczonych w przedziale <n_1,n_2>:

   $x_{śr} = \frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} x_n$ lub $x_{śr} = \frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} |x_n|$

   d) energia sygnału: $E_x = \sum_{n=-\infty}^{\infty} x^2{}_n$

e) moc średnia sygnału dla sygnałów ograniczonych w przedziale <n_1,n_2>:

$$P_x = \frac{1}{n_2 - n_2 + 1} \sum_{n=n_1}^{n_2} x^2{}_n$$

dla sygnałów o nieskończonym czasie trwania: $P_x = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x^2{}_n$

dla sygnałów okresowych o okresie N: $P_x = \frac{1}{N} \sum_{n=n_0}^{n_0+(N-1)} x^2{}_n$

f) wartość skuteczna sygnału (**R**oot **M**ean **S**quare): $P_{sk} = P_{RMS} = \sqrt{P_x}$

g) funkcja autokorelacji sygnału (estymator nieobciążony):

$$R_{xx}(k) = \frac{1}{N - |k|} \sum_{n=0}^{N-1-|k|} x_n \cdot x^*_{n-k} \ , \ -N+1 \le k \le N-1$$

h) funkcja korelacji wzajemnej sygnału x_n i y_n (estymator nieobciążony):

$$R_{xy}(k) = \frac{1}{N - |k|} \sum_{n=0}^{N-1-|k|} x_n \cdot y^*_{n-k} \ , \ -N+1 \le k \le N-1$$

Wartość funkcji korelacji staje się maksymalna, gdy dwa sygnały mają identyczny kształt i nie są względem siebie przesunięte w fazie (n=0). Dlatego stosuje się ją jako miarę podobieństwa dwóch sygnałów. Zarówno w funkcji korelacji wzajemnej, jak i autokorelacji przed znakiem sumy stosuje się czynnik normalizacyjny 1/N (estymator obciążony) lub 1/(N-k) (estymator nieobciążony).

3. Parametryzacja w dziedzinie częstotliwości.

Duża liczba szczegółów, jaką niesie ze sobą analiza widmowa utrudnia interpretację i rozpoznanie istotnych informacji zawartych w tym sygnale. Dlatego podobnie jak w przypadku dziedziny czasu, wyznacza się z funkcji widma (częstotliwościowego), lub mulitiwidma (czasowo-częstotliwościowego) te cechy, które są przydatne w analizowanym problemie.

Parametry kształtu:
Posiadając określone widmo sygnału G(t,f) można wyznaczyć parametry opisujące jego kształt:

Moment widmowy m-tego rzędu:

$$M_m(t) = \sum_{i=0}^{N} |G(t, f_i)| [f_i]^m$$

gdzie: G(t,f) – widmo czasowo-częstotliwościowe w chwili czasowej t,
f_i – częstotliwość i-tego pasma w analizie częstotliwościowej,
f_0, f_N – częstotliwość dolna, częstotliwość górna - odpowiednio w dyskretnej skali częstotliwości ograniczają pasmo wyznaczenia momentu widmowego.

Moment zerowy:

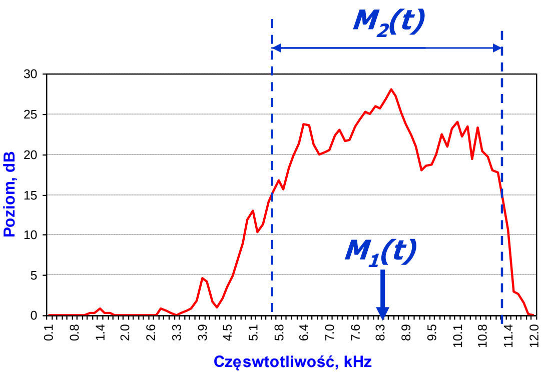
$$M_0(t) = \sum_{i=0}^{N} |G(t, f_i)|$$

Moment zerowego rzędu M_0(t) używany do normalizacji momentów wyższych rzędów.
Wygodnie jest, bowiem stosować momenty unormowane:

$$M_u(t) = \frac{M_m(t)}{M_0(t)}$$

Moment unormowany pierwszego rzędu:

$$M_1(t) = \frac{\sum_{i=0}^{N} |G(t, f_i)| \cdot f_i}{M_0(t)}$$



Moment $M_1(t)$ może być interpretowany jako środek ciężkości widma (średnia częstotliwość ważona). Jest używany we wzorach do obliczeń momentów centralnych wyższych rzędów.

Moment unormowany centralny drugiego rzędu:

$$M_2(t) = \frac{\sum_{i=0}^{N} |G(t, f_i)| \cdot [f_i - M_1(t)]^2}{M_0(t)}$$

Moment $M_2(t)$ może być interpretowany jako kwadrat szerokości widma.

Moment unormowany centralny trzeciego rzędu:

$$M_3(t) = \frac{\sum_{i=0}^{N} |G(t, f_i)| \cdot [f_i - M_1(t)]^3}{M_0(t)}$$

Moment unormowany centralny czwartego rzędu:

$$M_4(t) = \frac{\sum_{i=0}^{N} |G(t, f_i)| \cdot [f_i - M_1(t)]^4}{M_0(t)}$$

Parametr miary niesymetrii widma – skośność (ang. skewness):

skośność $= \frac{M_3(t)}{M_2(t)^3}$

Moment $M_3(t)$ oraz $M_2(t)$ wykorzystuję się do obliczenia tzw. skośności czyli miary niesymetrii widma.

Parametr miary spłaszczenia widma (kurtoza):

kurtoza $= \frac{M_4(t)}{M_2(t)^2}$

Moment $M_4(t)$ oraz $M_2(t)$ wykorzystuję się do obliczenia tzw. kurtozy czyli miary spłaszczenia widma co można rozumieć jako miarę harmoniczności sygnału.

Mniej przydatne są momenty widmowe wyższych rzędów, nawet unormowane, gdyż są one ze sobą skorelowane.

4. Pomocne funkcje programu MATLAB:

   diff, sum, mean, min, max, rms, xcorr**, **spectralCentroid, spectralSpread, spectralSkewness, spectralKurtosis, spectralEntropy, spectralFlatness

   Opis parametrów spektralnych MATLAB:
   https://www.mathworks.com/help/audio/ug/spectral-descriptors.html?s_tid=srchtitle_Spectral%20Descriptors_1

Zadania do wykonania:

1) Dla zdania {dziś jest ładna pogoda} utworzyć skrypt realizujący:
   a) ramkowanie sygnału o szerokości 20÷30 ms. Obliczyć dla każdej ramki sygnału parametry w dziedzinie czasu: całka sygnału, energia sygnału, moc sygnału, rms sygnału, wartość minimalna, wartość maksymalna sygnału. Przedstawić wynik graficznie.
   b) ramkowanie sygnału o szerokości 20÷30 ms. Obliczyć dla każdej ramki sygnału parametry w dziedzinie czasowo-częstotliwościowej: momenty widmowe, skośność kurtoza. Przedstawić wynik graficznie.