

Sygnały audio

1 Parametry dźwięków

Dźwięk może być opisany przy pomocy szeregów parametrów (czas trwania, średnia, energia, etc.). Do najczęściej spotykanych parametrów dźwięku zaliczamy:

- wysokość dźwięku f_0 (ang. *pitch*). Jest to niemierzalna wielkość powiązana z częstotliwością danego dźwięku. Do estymacji jej wartości używa się nieliniowej skali melowej (1000 meli to wysokość odpowiadająca częstotliwości 1 [kHz] i poziomemu natężeniu 40 [dB]) [3]. Empiryczną zależność wysokości dźwięku od częstotliwości przedstawiono na równ. 1 [1]:

$$mel = 1127 \cdot \ln \left(1 + \frac{700}{f} \right) = 2959 \cdot \log_{10} \left(1 + \frac{700}{f} \right) \quad (1)$$

- energia sygnału:

$$E_x = \sum_{n=-\infty}^{+\infty} |x[n]|^2 \quad (2)$$

- środek ciężkości sygnału (unormowany moment bezwładności pierwszego rzędu):

$$m_{1n} = \frac{\sum_{n=-\infty}^{+\infty} n \cdot x[n]}{\sum_{n=-\infty}^{+\infty} x[n]} \quad (3)$$

- obwiednia sygnału;
- kurtoza - mówi o wpływie wartości skrajnych (np. dla $Kurt > 0$ - jest on większy niż dla rozkładu normalnego) (μ - średnia, σ - odchylenie standardowe):

$$Kurt = \frac{\frac{1}{N} \sum_{n=1}^N (x[n] - \mu)^4}{\sigma^4} - 3 \quad (4)$$

- gęstość przejść przez zero;
- czas trwania.

2 AudioFeatureExtractor

Polecenie `parametry = audioFeatureExtractor('parametr', 'wartość', ...)` służy do obliczania wartości parametrów sygnału audio w oknie o określonej szerokości. Polecenie wymaga podania wartości kilku podstawowych danych wejściowych oraz określenia zakresu danych wyjściowych. Do danych wejściowych zaliczamy:

- `SampleRate` - częstotliwość próbkowania w Hz. Domyślnie 44.1 kHz;
- `Window` - okna w analizie częstotliwościowej. Domyślnie jest to `hamming(1024)`;
- `OverlapLength` - ilość próbek nakładania się kolejnych okien. Domyślnie 512.
- `SpectralDescriptorInput` - określenia w jakiej skali mają być liczone wartości parametrów (liniowa - domyślna, melowa, barkowa lub erbowa).

Wartości parametrów, które chcemy uzyskać wybieramy zmieniając wartość flagi danego parametru z domyślnego `false` na `true`. Do obliczalnych parametrów zaliczamy m.in.:

- `pitch` - wysokość dźwięku;
- `spectralCentroid` - unormowany moment bezwładności;
- `spectralCrest` - widmowy współczynnik szczytu;
- `spectralDecrease` - nachylenie (spadek) widma amplitudowego;

- `spectralEntropy` - entropia widma;
- `spectralFlatness` - płaskość widma;
- `spectralKurtosis` - kurtoza;
- `spectralSkewness` - skośność rozkładu;
- `spectralSpread`

Wartości parametrów uzyskujemy korzystając z poleceń: `info(parametry)` (kolejność parametrów w zmiennej wyjściowej) i `extract(parametry, sygnał)`

3 Cepstrum

Cepstrum sygnału to odwrotna transformata Fouriera z logarytmu rzeczywistego lub zespolonego widma amplitudowego (równ. 5). Wykorzystanie rzeczywistego logarytmu pozwala na pominięcie wpływu fazy (przesunięcia). Tym samym jednak nie jest możliwe odtworzenie sygnału pierwotnego.

$$C_P(x(t)) = \left| F^{-1} \left(\log |F(x(t))|^2 \right) \right|^2 \quad (5)$$

W MatLAB operacja ta jest dostępna z użyciem komend:

- rzeczywiste cepstrum: `rceps(sygnał)`
- zespolone cepstrum: `cceps(sygnał)`
- odwrotne zespolone cepstrum: `icceps(sygnał)`

4 Transformata Hilberta

Ciągła transformata Hilberta jest splotem sygnału $x_r(t)$ z odpowiedzią impulsową $\frac{1}{\pi t}$, która jest dana wzorem:

$$x_i(t) = H[x_r(t)] = \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{x_r(\tau)}{t - \tau} d\tau \quad (6)$$

Natomiast odwrotna Transformata Hilberta jest splotem z $\frac{-1}{\pi t}$:

$$x_r(t) = H^{-1}[x_i(t)] = \frac{-1}{\pi} \int_{-\pi}^{\pi} \frac{x_i(\tau)}{t - \tau} d\tau \quad (7)$$

W domenie częstotliwości odpowiada to mnożeniu widma sygnału z filtrem o następującej charakterystyce:

$$H(i\omega) = -H^{-1}(i\omega) = \begin{cases} -i & \omega > 0 \\ 0 & \omega = 0 \\ i & \omega < 0 \end{cases} \quad (8)$$

Sumę sygnału oraz jego transformaty Hilberta nazywamy sygnałem analitycznym. Dyskretna Transformata Hilberta dana jest wzorem ($\Omega = \frac{2\pi f}{f_S}$):

$$x_i(n) = \sum_{k=-\infty}^{\infty} H_{n-k} x_r(k) \quad (9)$$

Dyskretna transformata Hilberta służy m.in. do chwilowej demodulacji fazy (sygnał radiowy AM) lub chwilowej demodulacji przesunięcia fazowego.

5 Modulacja AM / FM

Sygnal może zostać przesłany falami korzystając z modulacji amplitudy (AM) lub częstotliwości (FM). W obu przypadkach mamy falę nośną o określonej częstotliwości f_c i amplitudzie (zwykle $A=1.0$).

W przypadku AM zmieniamy tylko amplitudę fali nośnej mnożąc ją przez amplitudę nadawanego sygnału. Demodulacja polega na wyliczeniu modułu (widma amplitudowego) sygnału analitycznego.

W przypadku FM modulujemy częstotliwość dodając do niej wzmocnioną całkę z przesyłanego sygnału. Demodulacja polega na policzeniu pochodnej z sygnału i późniejszej filtracji dolnoprzepustowej.

6 Zadania

1. Wczytaj sygnał *mtlb.mat*. Dodaj do niego echo z przesunięciem 0.2 sekundy i 0.35 sekundy osłabione odpowiednio 0.7 razy w stosunku do poprzedniego wybrzmienia. Korzystając z domeny cepstrum zidentyfikuj echo, a następnie korzystając z filtru IIR usuń je z sygnału. Spróbuj zlokalizować echo korzystając z autokorelacji.
2. Policz transformatę Hilberta iloczynu dwóch sygnałów harmonicznym o amplitudzie 1.0 i częstotliwościach $f_1 = 10$ Hz i $f_2 = 18$ Hz. Załóż: $t \in \langle 0, 10 \rangle$ s, $F_S = 100$ Hz. Wyświetl sygnał, jego widmo amplitudowe transformaty Fouriera oraz widmo amplitudowe transformaty Hilberta.
Następnie zmień częstotliwość f_2 na liniową zmienną od 6 do 14 Hz i powtórz czynności.
3. Dany jest sygnał o $t \in \langle 0, 10 \rangle$ i $F_s = 1KHz$. Korzystając z fali nośnej o $F_c = 100Hz$ i $Amp = 1.0$ dokonaj modulacji AM/FM i demodulacji wiadomości w postaci:
 - sygnału trójkątnego (amp.=1, tw = 5 s, szer. = 4 s);
 - sygnału harmonicznego (amp.=1, f = 20 Hz).

Załącz wzmocnienie = 100 oraz użyj LP Butterwortha ($f_0 = 30$ Hz i $N = 5$). Sprawdź wpływ częstotliwości próbkowania oraz częstotliwości fali nośnej na jakość odtwarzanego sygnału.

Literatura

- [1] Ciota, Z. *Metody przetwarzania sygnałów akustycznych w komputerowej analizie mowy*. Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2010.
- [2] Izydorczyk, J. i inni. *MATLAB i podstawy telekomunikacji*. Helion, Gliwice, 2017.
- [3] Makowski, R. *Automatyczne rozpoznawanie mowy - wybrane zagadnienia*. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, 2011.
- [4] Zieliński, T.P., (2009) *Cyfrowe Przetwarzanie Sygnałów. Od teorii do zastosowań*. WKŁ, Warszawa