

ANALIZA WIDMOWA – notatki do wykładów

- Wprowadzenie – str. 1
- Nieparametryczne metody analizy widmowej: periodogram i korelogram – str. 5
- Momenty jako parametry zmiennych losowych – str. 7
- Momenty widmowe – str. 8
- Identyfikacja mówcy z zastosowaniem momentów widmowych – str. 10
- Wysokorozdzielcza parametryczna estymacja WGM – widma gęstości mocy – str. 13
- Estymacja WGM na podstawie modelu AR – str. 15
- Literatura uzupełniająca z analizy widmowej – str. 19
- Pytania z analizy widmowej dla Akustyki Mowy – str. 20

ANALIZA WIDMOWA

Wprowadzenie

Analiza widmowa to każda metoda przetwarzania sygnału opisująca zawartość częstotliwościową sygnału. Przekształcenie Fouriera jest podstawą matematyczną wiążącą sygnał czasowy lub przestrzenny, lub model takiego sygnału, z jego reprezentacją w dziedzinie częstotliwości, czyli z widmem.

Widmo sygnału analogowego deterministycznego $x(t)$ to

$$X(j\Omega) = \int_{-\infty}^{\infty} x(t)e^{-j\Omega t} dt \quad (1.1)$$

gdzie pulsacja (częstotliwość kątowna) $\Omega = 2\pi F$ jest wyrażana w rad/s, F to częstotliwość w Hz, a $j^2 = -1$. Mamy tu do czynienia z \mathbf{F} – transformacją sygnału $x(t)$, której wynikiem jest transformata $X(j\Omega)$, czyli widmo sygnału $x(t)$.

Widmo gęstości energii (WGE) to

$$E(\Omega) = |X(j\Omega)|^2 \quad (1.2)$$

(po angielsku ESD – energy spectral density). WGE to również wynik \mathbf{F} – transformacji funkcji autokorelacji $r(\tau)$ sygnału $x(t)$

$$r(\tau) = \int_{-\infty}^{\infty} x(t)x^*(t-\tau)dt = r^*(-\tau); \quad r(0) = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad (1.3)$$

(przy czym dla sygnałów o wartościach rzeczywistych $r(\tau)$ jest rzeczywista i parzysta), tj.

$$E(\Omega) = \int_{-\infty}^{\infty} r(\tau) e^{-j\Omega\tau} d\tau \quad (1.4)$$

a całkowita energia sygnału, ze wzoru Parsevala, to

$$E = r(0) = \int_{-\infty}^{\infty} |x(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} E(\Omega) d\Omega \quad (1.5a)$$

i gdy sygnał jest rzeczywisty (o wartościach rzeczywistych) to

$$E = \frac{1}{\pi} \int_0^{\infty} E(\Omega) d\Omega \quad (1.5b)$$

Powyższe jest słuszne dla sygnałów o skończonej energii:

$$E = \int_{-\infty}^{\infty} |x(t)|^2 dt < \infty \quad (1.6)$$

Uwaga! Jeżeli po prawej stronie we wzorze (1.5a) posłużymy się częstotliwością F zamiast pulsacją Ω to $E = \int_{-\infty}^{\infty} E(F) dF$ czyli pozbywamy się współczynnika $1/(2\pi)$. Podobnie w innych tego typu wzorach.

Dla sygnałów dyskretnych operuje się dyskretno-czasową transformacją Fouriera (DTFT). Wówczas, dla ciągu $x[n]$, mamy widmo

$$X(j\omega) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}; \quad \omega = \Omega / F_s; \quad F_s - \text{częstotliwość próbkowania w Hz} \quad (1.7)$$

Tu ω jest wyrażana w rad/Sa (w radianach na próbkę); $\omega = 2\pi f$; $\omega \in [-\pi, \pi)$, $f \in [-1/2, 1/2)$. WGE sygnału $x[n]$ to

$$E(\omega) = |X(j\omega)|^2 \quad (1.8)$$

lub

$$E(\omega) = \sum_{k=-\infty}^{\infty} r[k] e^{-j\omega k} \quad (\text{DTFT z } r[k]) \quad (1.9)$$

gdzie

$$r[k] = \sum_{n=-\infty}^{\infty} x[n] x^*[n-k] \quad (\text{funkcja autokorelacji – dyskretna}) \quad (1.10)$$

Widmo gęstości mocy (WGM) dla sygnału deterministycznego $x(t)$ to

$$P(F) = \lim_{T \rightarrow \infty} \left(\frac{|X_T(jF)|^2}{T} \right) \quad (1.11)$$

(po angielsku PSD – power spectral density). Widmo to opisuje rozkład mocy sygnału wzdłuż osi częstotliwości. WGM jest bardziej użyteczna niż WGE, ponieważ w rozwiązywaniu problemów komunikacji generalnie wykorzystuje się modele oparte na

mocy sygnału. We wzorze (1.11) $P(F)$ jest wyrażana w jednostkach wat na Hz, a $X_T(jF)$ to \mathbf{F} – transformacja sygnału uciętego (centralny fragment sygnału – ma ograniczoną energię)

$$x_T(t) = \begin{cases} x(t), & -T/2 < t < T/2 \\ 0, & -\infty < t < -T/2, \quad T/2 < t < \infty \end{cases} \quad (1.12)$$

Znormalizowana (względem 1 ohma) moc średnia

$$P = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} |x(t)|^2 dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} |x_T(t)|^2 dt \quad (1.13)$$

Na mocy twierdzenia Parsewała znormalizowana moc średnia

$$P = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} |X(F)|^2 dF = \int_{-\infty}^{\infty} \left(\lim_{T \rightarrow \infty} \frac{|X_T(F)|^2}{T} \right) dF \quad (1.14)$$

Zmienna podcałkowa po prawej stronie (1.14) to PSD. Jest to zawsze nieujemna funkcja częstotliwości, niewrażliwa na widmo fazowe sygnału $x(t)$. Rzecz jasna, znormalizowana moc średnia to

$$P = \langle |x(t)|^2 \rangle = \int_{-\infty}^{\infty} P(F) dF \quad (1.15)$$

A moc średnia sygnału o wartościach rzeczywistych zawarta w paśmie $\Delta F = F_2 - F_1$ jest równa

$$P(\Delta F) = \lim_{T \rightarrow \infty} \frac{1}{T} \left[\int_{-F_2}^{-F_1} |X(F)|^2 dF + \int_{F_1}^{F_2} |X(F)|^2 dF \right] = 2 \int_{F_1}^{F_2} \left(\lim_{T \rightarrow \infty} \frac{|X_T(F)|^2}{T} \right) dF \quad (1.16)$$

Powyżej mieliśmy do czynienia z tzw. bezpośrednią metodą obliczania PSD. Ten sam wynik otrzymuje się za pomocą metody autokorelacyjnej. Autokorelacja przebiegu fizycznego jest definiowana jako

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x^*(t-\tau) dt \quad (1.17)$$

Twierdzenie Wienera-Chinczyna mówi, że pomiędzy WGM a funkcją autokorelacji zachodzi związek (para transformat Fouriera)

$$R(\tau) \Leftrightarrow P(F) \quad (1.18)$$

gdzie $P(F)$ jest \mathbf{F} – transformatą funkcji autokorelacji $R(\tau)$. Zatem drugą metodą obliczania WGM, tj. $P(F)$, jest metoda polegająca na obliczeniu funkcji autokorelacji, a następnie obliczeniu transformaty Fouriera tej funkcji.

Przykład 1. Obliczanie WGM sinusoidy metodą autokorelacyjną

Weźmy sygnał

$$x(t) = A \sin(\Omega_0 t), \quad -\infty < t < \infty \quad (1.19)$$

Autokorelacją jest

$$\begin{aligned}
R(\tau) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x^*(t-\tau)dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} A^2 \sin(\Omega_0 t) \sin(\Omega_0(t-\tau))dt = \\
&= \frac{A^2}{2} \cos(\Omega_0 \tau) \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} dt - \frac{A^2}{2} \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \cos(2\Omega_0 t - \Omega_0 \tau) dt
\end{aligned} \tag{1.20}$$

ponieważ

$$\sin \alpha \sin \beta = \frac{1}{2} [\cos(\alpha - \beta) - \cos(\alpha + \beta)] \tag{1.21}$$

Wzór (1.20) sprowadza się do

$$R(\tau) = \frac{A^2}{2} \cos(\Omega_0 \tau) \tag{1.22}$$

Zatem WGM, czyli transformata Fouriera funkcji (1.22) wynosi

$$P(F) = \frac{A^2}{4} [\delta(F - F_0) + \delta(F + F_0)] \tag{1.23}$$

Znormalizowana moc średnia

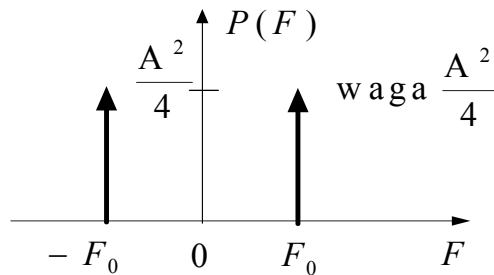
$$P = \int_{-\infty}^{\infty} \frac{A^2}{4} [\delta(F - F_0) + \delta(F + F_0)] dF = \frac{A^2}{2} \tag{1.24}$$

Wartość ta zgadza się ze znanym wynikiem na znormalizowaną moc sinusoidy

$$P = (A/\sqrt{2})^2 = A^2/2 \tag{1.25}$$

gdzie w nawiasie występuje wartość skuteczna (inaczej średniokwadratowa) sinusoidy.

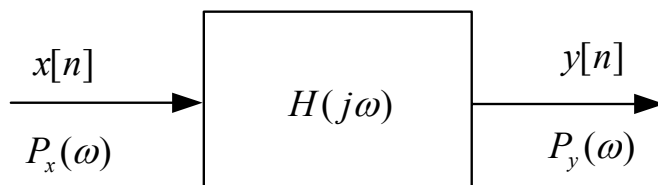
Sygnaly $A \sin(\Omega_0 t)$ i $A \cos(\Omega_0 t)$ mają tę samą WGM i tę samą funkcję autokorelacji, ponieważ faza nie ma wpływu na WGM. Można to sprawdzić obliczając WGM dla $A \cos(\Omega_0 t)$ za pomocą tej samej procedury, jakiej użyliśmy wcześniej do obliczenia WGM dla $A \sin(\Omega_0 t)$.



Rys. 1. WGM sinusoidy.

Związek pomiędzy WGM na wejściu i wyjściu systemu liniowego o charakterystyce amplitudowo-fazowej $H(j\omega)$

$$P_y(\omega) = |H(j\omega)|^2 P_x(\omega) \tag{1.26}$$



Rys. 2. Przenoszenie WGM przez dyskretny układ liniowy.

WGM dla dyskretnych sygnałów stochastycznych mówi jaką moc średnią zawiera sygnał w funkcji częstotliwości. Zakłada się, że sygnał o ograniczonej mocy średniej ma wartość oczekiwaną (to jest wartość uśrednioną po zbiorze realizacji)

$$E\{x[n]\} = 0 \quad (1.27)$$

Jeżeli ciąg autokorelacji wyrazimy za pomocą wzoru

$$r[k] = E\{x[n]x^*[n-k]\} = r^*[-k] \quad (1.28)$$

wówczas WGM to DTFT z $r[k]$

$$P(\omega) = \sum_{k=-\infty}^{\infty} r[k]e^{-j\omega k} \quad (1.29)$$

i odwrotnie, IDTFT z $P(\omega)$

$$r[k] = \frac{1}{2\pi} \int_{-\pi}^{\pi} P(\omega) e^{j\omega k} d\omega \quad (1.30)$$

Albo, zakładając, że ciąg autokorelacji dostatecznie szybko maleje

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=-\infty}^{\infty} |k| |r[k]| = 0 \quad (1.31)$$

możemy wykorzystać inną definicję WGM

$$P(\omega) = \lim_{N \rightarrow \infty} E \left\{ \frac{1}{N} \left| \sum_{k=-\infty}^{\infty} x[n] e^{-j\omega n} \right|^2 \right\} \quad (1.32)$$

prowadzącą do tych samych wyników jak (1.29).

Nieparametryczne metody analizy widmowej: periodogram i korelogram

Metoda **periodogramu** (należąca do metod nieparametrycznych) oparta jest na definicji (1.32). Formalnie, problem estymacji WGM formułuje się następująco.

Na podstawie skończonej obserwacji $\{x[1], x[2], \dots, x[N]\}$ procesu losowego stacjonarnego, oblicz estymatę $P(\omega)$ widma gęstości mocy dla $-\pi \leq \omega < \pi$.

Pomijając operator wartości oczekiwanej w (1.32) oraz przejście do granicy, których nie można wykonać, gdy jedyną dostępną informacją jest zbiór próbek sygnału $\{x[n]\}_{n=1}^N$, otrzymujemy estymator $\hat{P}(\omega)$ periodogramu $P(\omega)$ Shustera

$$\hat{P}(\omega) = \frac{1}{N} \left| \sum_{n=1}^N x[n] e^{-j\omega n} \right|^2 \quad (1.33)$$

którego pierwszym zastosowaniem (Shuster) było wykrywanie ukrytych okresowości (ang. *hidden periodicities*, gdzie okres – *period*) w ciągu próbek czasowych. Stąd nazwa: periodogram.

Periodogram to wykres (diagram), który pokazuje najważniejsze okresowe regularności występujące w sygnale. Piki na diagramie odpowiadają okresom (cyklom), które najbardziej korelują z danymi. Interpretacją periodogramu zajmuje się analiza widmowa.

Jeżeli zastosujemy ważenie sygnału $\{x[n]\}_{n=1}^N$ za pomocą okna $\{w[n]\}_{n=1}^N$, to otrzymujemy estymator periodogramu ważonego, zwanego też zmodyfikowanym

$$\hat{P}(\omega) = \frac{\frac{1}{N} \left| \sum_{n=1}^N x[n] w[n] e^{-j\omega n} \right|^2}{\frac{1}{N} \sum_{n=1}^N |w[n]|^2} \quad (1.34)$$

W każdym przypadku do obliczania periodogramu używa się N -punktowej FFT, by obliczyć $\hat{P}(\omega)/F$ (w dyskretnych punktach pulsacji ω), gdzie za F podstawiamy 2π , gdy nie zadajemy częstotliwości próbkowania, a postawiamy F_s , gdy zadajemy częstotliwość próbkowania. Powyższy algorytm działa w MATLABie pod nazwą `periodogram`. Zwraca on estymatę (1.33) lub (1.34).

W MATLABie działają również inne estymatory WGM, ale należące do parametrycznych. Są to:

- ◆ `pburg` działający wg metody Burga,
- ◆ `pcov` działający wg metody kowariancji,
- ◆ `peig` działający wg metody wektorów własnych,
- ◆ `pncov` działający wg zmodyfikowanej metody kowariancji,
- ◆ `pmusic` wykorzystujący algorytm MUSIC (Multiple Signal Classification),
- ◆ `pwelch` działający wg metody Welch,
- ◆ `pyulear` działający wg metody AR (autoregresji) Yule-Walkera.

Wszystkie te algorytmy, należące do wysokorozdzielczej parametrycznej estymacji WGM, są szczegółowo opisane w `Help` do MATLABa i opatrzone przykładami i ilustracjami.

Na podstawie definicji WGM jako DTFT z ciągu autokorelacji (1.29) estymatorem widmowym w postaci **korelogramu** jest

$$\hat{P}_c(\omega) = \sum_{k=-(N-1)}^{N-1} \hat{r}[k] e^{-j\omega k} \quad (1.35)$$

gdzie $\hat{r}[k]$ oznacza estymatę $r[k]$ otrzymaną na podstawie dostępnych próbek $\{x[1], x[2], \dots, x[N]\}$. Jeżeli o tym sygnale zakładamy jedynie to, że jest on stacjonarny, to możemy znaleźć $\hat{r}[k]$ potrzebne w (1.35) dwoma sposobami: jako estymator standardowy nieobciążony

$$\hat{r}[k] = \frac{1}{N-k} \sum_{n=k+1}^N x[n]x^*[n-k], \quad 0 \leq k \leq N-1 \quad (1.36)$$

i jako estymator standardowy obciążony (lepszy dla średnich i dużych wartości parametru k)

$$\hat{r}[k] = \frac{1}{N} \sum_{n=k+1}^N x[n]x^*[n-k], \quad 0 \leq k \leq N-1 \quad (1.37)$$

przy czym, dla indeksów ujemnych, korzystamy z właściwości

$$\hat{r}[-k] = \hat{r}^*[k] \quad (1.38)$$

W praktyce obliczenia periodogramu i korelogramu wykonuje się za pomocą FFT zamiast DTFT wymagającej continuum wartości częstotliwości. Stosuje się też uzupełnianie ciągu $\{x[n]\}_{n=1}^N$ zerami, gdy N nie jest potęgą liczby 2.

Momenty jako parametry zmiennych losowych

W poniższych wzorach zastosowano następujące oznaczenia [3]

- ξ – zmienna losowa,
- $f_{\xi}(x)$ – funkcja gęstości prawdopodobieństwa zmiennej losowej ξ ,
- x – wartość zmiennej losowej w zbiorze $\xi(\varepsilon) \subseteq R$ (inkluzja zbiorów),
- ε – zbiór zwany przestrzenią wyników doświadczenia lub przestrzenią zdarzeń elementarnych; elementy $e \in \varepsilon$ tego zbioru są nazywane wynikami doświadczenia lub zdarzeniami elementarnymi.

Moment zwykły rzędu r zmiennej losowej ξ to liczba

$$\overline{\xi^r} = \int_{-\infty}^{\infty} x^r f_{\xi}(x) dx$$

Moment centralny rzędu r zmiennej losowej ξ to liczba

$$\overline{(\xi - \bar{\xi})^r} = \int_{-\infty}^{\infty} (x - \bar{\xi})^r f_{\xi}(x) dx$$

W przypadku zmiennej losowej dyskretnej ξ o skończonej liczbie realizacji x_k ; $k = 0, 1, \dots, N-1$; powyższe wyrażenia przyjmują postać

$$\overline{\xi^r} = \sum_{k=0}^{N-1} x_k^r P_{\xi}(x_k)$$

$$\overline{(\xi - \bar{\xi})^r} = \sum_{k=0}^{N-1} (x_k - \bar{\xi})^r P_{\xi}(x_k)$$

gdzie $\{P_{\xi}(x_k); k = 0, 1, \dots, N-1\}$ to zbiór prawdopodobieństw przyporządkowanych tym realizacjom.

Momenty unormowane otrzymujemy dzieląc momenty zwykłe i centralne przez $\overline{\xi^0} = \overline{\xi^r} \Big|_{r=0}$. Momentami najważniejszymi i najczęściej wykorzystywanymi do opisu zmiennych losowych są:

- $\bar{\xi}$ – moment zwykły rzędu 1, nazywany wartością oczekiwaną, wartością średnią lub nadzieją matematyczną zmiennej losowej ξ i oznaczany również $E(\xi)$ lub m_{ξ} , lub $M(0)$,
- $\overline{\xi^2}$ – moment zwykły rzędu 2, nazywany wartością średniokwadratową zmiennej losowej ξ i oznaczany także $E(\xi^2)$,
- $\overline{(\xi - \bar{\xi})^2}$ – moment centralny rzędu 2 lub inaczej wariancja zmiennej losowej ξ i oznaczany najczęściej σ_{ξ}^2 .

Często też wykorzystywany jest również parametr $\sqrt{\sigma_{\xi}^2}$ - pierwiastek z wariancji oznaczany przez σ_{ξ} , nazywany odchyleniem standardowym (dyspersją) zmiennej losowej ξ . (Dyspersja – inaczej rozrzut.)

Momenty widmowe

Momenty widmowe to **deskryptory** cech sygnału. Wykorzystuje się je do badania **podobieństwa** pomiędzy różnymi obiektami. Oblicza się je najczęściej na podstawie estymaty WGM widma sygnału.

Aby zdefiniować momenty widmowe, w powyższych wzorach zastępujemy x_k przez częstotliwość f_k , $k = 0, 1, \dots, N-1$ oraz $P_{\xi}(x_k)$ przez widmo gęstości mocy P_k lub jego estymatę liczone za pomocą N -punktowej DFT (FFT). Wówczas wprowadzając oznaczenie

$$M(r) = \overline{\xi^r} \Big|_{f_k = x_k, P_k = P_{\xi}(x_k) \text{ lub } \hat{P}_k = P_{\xi}(x_k)} \quad (1.39)$$

otrzymujemy moment widmowy r -tego rzędu

$$M(r) = \sum_{k=0}^{N-1} P_k f_k^r \quad (1.40)$$

i moment widmowy unormowany r -tego rzędu

$$M_u(r) = \frac{M(r)}{M(0)} \quad (1.41)$$

gdzie $M(0) = \sum_{k=0}^{N-1} P_k$ to moment widmowy rzędu zerowego mający zastosowanie normalizujące i oznacza moc sygnału, i moment widmowy unormowany centralny r -tego rzędu

$$M_{uc}(r) = \sum_{k=0}^{N-1} \frac{P_k [f_k - M_u(1)]^r}{M(0)} \quad (1.42)$$

gdzie

$$M_u(1) = \sum_{k=0}^{N-1} \frac{P_k f_k}{M(0)} \quad (1.43)$$

to moment widmowy unormowany pierwszego rzędu używany do obliczania momentów widmowych centralnych wyższych rzędów – ma interpretację środka ciężkości widma.

A teraz dalsze szczególne przypadki momentów widmowych. Moment widmowy unormowany centralny drugiego rzędu – ma interpretację kwadratu szerokości pasma częstotliwości zajmowanych przez sygnał

$$M_{uc}(2) = \sum_{k=0}^{N-1} \frac{P_k [f_k - M_u(1)]^2}{M(0)} \quad (1.44)$$

Moment widmowy unormowany centralny trzeciego rzędu to niesymetria widma, inaczej skośność (ang. *skewness*)

$$M_{uc}(3) = \sum_{k=0}^{N-1} \frac{P_k [f_k - M_u(1)]^3}{M(0)} \quad (1.45)$$

Parametr widmowy będący miarą splaszczenia widma (popularnie kurtoza)

$$kurtosis = \frac{M_{uc}(4)}{[M_{uc}(2)]^2} \quad (1.46)$$

Można też inaczej zapisać

$$kurtosis = \frac{1}{N} \sum_{k=0}^{N-1} \frac{(x_k - \bar{x})^4}{\sigma_x^4} \quad (1.47)$$

gdzie x_k to k -ta obserwacja spośród dostępnych N obserwacji, \bar{x} to średnia arytmetyczna ze wszystkich N próbek, σ_x – to odchylenie standardowe (dyspersja, rozrzut) liczone na podstawie obserwacji.

Parametrem służącym jako miara płaskości widma (ang. SFM – *Spectral Flatness Measure*) jest (średnia geometryczna przez średnią arytmetyczną)

$$\text{SFM} = 10 \log_{10} \frac{\frac{1}{N/2} \sqrt{\prod_{k=0}^{N/2-1} P_k}}{\frac{1}{N/2} \sum_{k=0}^{N/2-1} P_k} \quad (1.48)$$

przy czym P_k to widmowa gęstość mocy (WGM) obliczona za pomocą N -punktowej DFT (FFT).

Momenty widmowe można także obliczać dla fragmentów widma. Wówczas należy w powyższych wzorach zmienić zakresy sumowania (i iloczynu dla SFM) na numery próbek widma odpowiadające częstotliwościom: f_d – dolnej i f_g – górnej, widma sygnału. Przykładowo, pierwszy moment widmowy znormalizowany $M_u(1)$ – środek ciężkości widma, liczony w zakresie pomiędzy dwoma kolejnymi minimami obwiedni widma, może być interpretowany jako częstotliwość **formantu**, znajdującego się w tym paśmie częstotliwości.

Poniżej jest zamieszczony przykład zastosowania momentów widmowych w praktyce.

Identyfikacja mówcy z zastosowaniem momentów widmowych

Zgromadź zbiory nagrań głosów. Aby sklasyfikować szczególny, docelowy głos spośród głosów w zbiorze odniesienia (referencyjnym), należy [1] zgromadzić głosy o nazwie angielskiej *isochunks* (jednakowe kawałki). *Isochunk* jest definiowany jako segment mowy, w którym mówca, wymawiając go, za każdym razem koniecznie stosuje tę samą wymowę, a więc brzmi on tak samo. Aby móc taki segment wykorzystać, w bazie nagrań głosów musi egzystować więcej niż jeden taki segment dla każdego mówcy. Im więcej, tym lepiej z punktu widzenia jego weryfikacji/identyfikacji na podstawie porównania z nagraniami zapisanymi w bazie.

Isochunk może mieć dowolną długość, jednak nie powinien zawierać przerw. Chodzi o to, by miał on z zasady taką samą reprezentację, przy której może on różnić się dla poszczególnych mówców sposobami jego artykulacji. (Artykulacja – ruchy i układ narządów mowy (np. języka, warg, podniebienia miękkiego, więzadeł głosowych) podczas wymawiania poszczególnych głosek.) Czynnikiem krytycznym w selekcji segmentu mowy jako *isochunk*-a jest, by wybrany ciąg brzmiał za każdym razem tak podobnie, jak to tylko możliwe dla danego mówcy i niepodobnie, jak to tylko możliwe, dla innych mówców.

Po wyselekcjonowaniu *isochunks* i ich wydobyciu, następnie, celem wydobycia z nich cech dystynktywnych przydatnych do klasyfikacji, można zastosować następujący algorytm [1], sporządzony w Szwecji na zamówienie banku, a oparty na metodzie momentów widmowych opracowanej w North Carolina State University, USA, którego kolejne kroki zestawiono poniżej.

1. Oblicz DFT (ang. *Discrete Fourier Transform*) w oknie o długości N .
2. Odrzuć część urojoną.
3. Przesuń o jedną próbkę i powtórz N -krotnie kroki 1 i 2.
4. Weź średnią z N transformat i przeskaluj przez pierwiastek trzeciego stopnia. To zmniejszy wpływ pierwszego formantu.
5. Wykonaj interpolację wynikowej średniej za pomocą sześciennego splinu (ang. *cubic spline*), patrz MATLAB – *cubic spline interpolation*, aby otrzymać synchroniczne, ciągłe widmo wysokości dźwięku (ang. *pitch*).
6. Scałkuj ciągłe widmo wysokości dźwięku od 0 do 4000 Hz. To zwróci masę widma

$$mass = \int_0^{4000} S(f)df \quad (1.49)$$

7. Podziel widmo przez jego masę. Dostarczy to poniższej funkcji gęstości prawdopodobieństwa i sprawi, że powierzchnia pod krzywą widma będzie miała

wartość jeden, a dla wszystkich pod-powierzchni będzie to wartość z zakresu od 0 do 1.

$$P(f) = \frac{S(f)}{mass} \quad (1.50)$$

8. Scałkuj funkcję gęstości prawdopodobieństwa przemnożoną przez częstotliwość. Otrzymasz pierwszy moment tj. wartość oczekiwaną

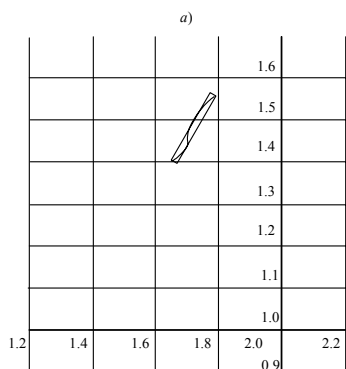
$$m_1 = \bar{x} = \int_0^{4000} f P(f) df \quad (1.51)$$

9. Scałkuj kwadrat różnicy pomiędzy pierwszym momentem i iloczynem częstotliwości przez funkcję gęstości prawdopodobieństwa. Otrzymasz drugi moment, to jest dyspersję wokół wartości oczekiwanej

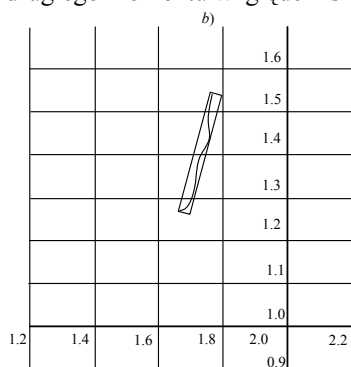
$$m_2 = \sigma^2 = \int_0^{4000} (f - m_1)^2 P(f) df \quad (1.52)$$

10. Powtórz kroki od 1 do 9, gdy liczba próbek pozostawionych w segmencie mowy jest większa od $3N$.
11. Przeskaluj pierwszy i drugi moment, odpowiednio, przez 10^{-3} i 10^{-6} .
12. Sporządź wykres drugiego momentu (oś pionowa) w funkcji pierwszego (oś pozioma). To utworzy ślad (ang. *track*).
13. Nałóż na *track* otaczający go prostokąt o minimalnej wielkości – MER od ang. *minimal enclosing rectangle*. Końcowe wyniki do klasyfikacji są wydobywane z MER.
14. Wydobądź z MER następujące atrybuty: wartości minimalną i maksymalną na osiach poziomej i pionowej (każda z nich reprezentuje jeden narożnik), długości boków po stronie dłuższej i krótszej, współrzędne środkowego punktu prostokąta na osi poziomej i pionowej i kąt orientacji prostokąta względem osi poziomej.

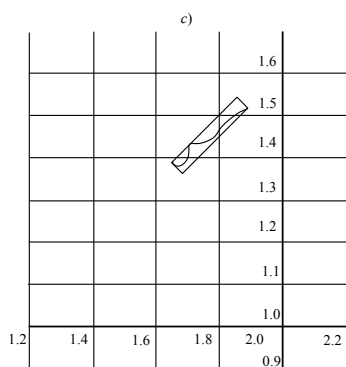
Przykłady i trzy wykresy momentów: drugiego względem pierwszego, zawierające MER są dostępne w [1] i pokazane na rys. 3, 4 i 5. Rozpatrzono tam głos – segment wypowiedzi – znanego szwedzkiego polityka oraz segment tej samej wypowiedzi ale wypowiedzianej przez profesjonalnego imitatora (naśladowcę), a następnie przez imitatora jego naturalnym głosem. Do klasyfikacji stosowano z sukcesem dyskryminanty z p. 14 (patrz wyżej) oraz miary oparte na odległości Mahalanobisa wprowadzonej w 1936. (patrz Wikipedia). Osiągnięto podobne wyniki dla obu tych metod. Pozwalały one na pozytywną ocenę wyników porównania oryginalnego naturalnego nagrania z imitacjami. Metoda momentów widmowych okazała się niewrażliwa na imitacje oryginalnej wypowiedzi.



Rys. 3. Wykres śladu (ang. *track*) uformowanego przez wykreślenie pierwszego i drugiego momentu względem siebie. Głos słynnego szwedzkiego polityka.



Rys. 4. Wykres śladu (ang. *track*) uformowanego przez wykreślenie pierwszego i drugiego momentu względem siebie. Głos jednego z profesjonalnych imitatorów podczas imitacji słynnego szwedzkiego polityka. Ta sama wypowiedź, jak na rys. 3.



Rys. 5. Wykres śladu (ang. *track*) uformowanego przez wykreślenie pierwszego i drugiego momentu względem siebie. Naturalny głos imitatora z rys. 3 podczas imitacji słynnego szwedzkiego polityka. Ta sama wypowiedź, jak na rys. 3.

Wysokorozdzielcza parametryczna estymacja WGM – widma gęstości mocy

Podstawowe założenie: badany sygnał traktowany jest jako proces stochastyczny. Proces ten jest generowany przez model matematyczny zależny od małej liczby parametrów. Poszukiwane WGM determinują parametry modelu. Parametryczna estymacja WGM składa się z trzech etapów.

1. Wybór odpowiedniego modelu parametrycznego (na podstawie wiedzy *a priori* o procesie).
2. Estymacja parametrów (identyfikacja) modelu.
3. Obliczanie widma na podstawie tak otrzymanych parametrów modelu.

Najczęściej stosowane modele parametryczne (inaczej innowacyjne):

- AR – autoregresywny, autoregresji,
- MA – ruchomej średniej, ang. *moving average*,
- ARMA – autoregresji i ruchomej średniej, mieszany
- suma cisoid (zespolonych sinusoid) z addytywnym szumem białym.

Model AR jest filtrem cyfrowym rekursywnym (IIR od ang. *Infinite Impulse Response*). Jego transmitancja nie ma zer poza punktem $z = 0$. Model kształtuje widmo wyłącznie za pomocą biegunów. Stąd ma angielską nazwę *all-pole model*. Równanie różnicowe (r.r.) – algorytm – modelu AR rzędu p , $AR(p)$ to

$$x[n] = -\sum_{m=1}^p a_m x[n-m] + e[n] \quad (2.1)$$

gdzie $x[n]$ to (dostępna) obserwacja wyjścia modelu, a_1, a_2, \dots, a_p – parametry modelu, a $e[n]$ – nieznanne losowe pobudzenie modelu (nie dostępne), innowacja, jest realizacją białego szumu gaussowskiego o zerowej wartości średniej i nieznannej mocy (wariancji) σ^2 .

Model MA jest filtrem cyfrowym nierekursywnym (FIR od ang. *Finite Impulse Response*) rzędu q , $MA(q)$ nie mającym biegunów (ang. *all-zero model*). Równanie różnicowe (r.r.) – algorytm – modelu $MA(q)$ to

$$x[n] = \sum_{m=0}^q b_m e[n-m] \quad (2.2)$$

gdzie $b_0 = 1, b_1, b_2, \dots, b_q$ – parametry modelu, inaczej współczynniki ruchomej średniej.

Model ARMA jest filtrem cyfrowym IIR rzędu $\max(q, p)$. Równanie różnicowe – algorytm – modelu ARMA to kombinacja liniowa (2.1) i (2.2)

$$x[n] = -\sum_{m=1}^p a_m x[n-m] + \sum_{m=0}^q b_m e[n-m] \quad (2.3)$$

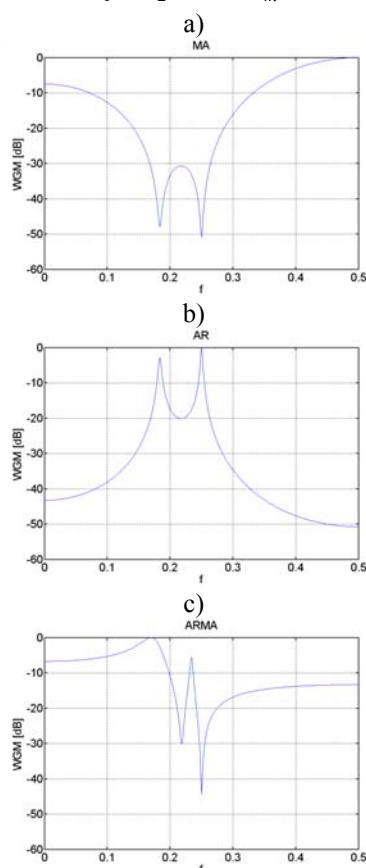
Parametry tego modelu nazywają się tak samo, jak parametry jego składowych AR i MA.

Model „suma zespolonych sinusoid w białym szumie” ma algorytm

$$x[n] = \sum_{i=1}^m A_i \exp(j\omega_i n) + e[n] \quad (2.4)$$

gdzie A_1, A_2, \dots, A_m – amplitudy zespolone (parametry modelu), a $e[n]$ – realizacja białego szumu gaussowskiego o zerowej wartości średniej i nieznaney mocy (wariancji) σ^2 (tak samo, jak wyżej).

Ogólnie parametrami modelu są zarówno zespolone amplitudy i pulsacje sinusoid. Te ostatnie spełniają związek $-\pi < \omega_1 < \omega_2 < \dots < \omega_m < \pi$.



Rys. 6. Typowe widma modeli parametrycznych: a) MA, b) AR i c) ARMA; dla celów porównawczych maksima znormalizowano do 0 dB.

Zera transmitancji MA to: $\pm j0.99$ i $-0.4 \pm j0.9$, stąd $H_{MA}(z) = (z^2 + 0.99^2)[(z + 0.4)^2 + 0.9^2]$, bieguny transmitancji AR to $\pm j0.99$ i $-0.4 \pm j0.9$, stąd $H_{AR}(z) = \frac{1}{H_{MA}(z)} = \frac{1}{(z^2 + 0.99^2)[(z + 0.4)^2 + 0.9^2]}$, a dla transmitancji ARMA przyjęto na powyższym rysunku położenie zer: $\pm j0.995$ i $-0.2 \pm j0.995$, położenie biegunów: $-0.1 \pm j0.98$ i $-0.5 \pm j0.999$, przy czym $H_{ARMA}(z) = H_{AR}(z)H_{MA}(z)$. Czytelnikowi pozostawiamy znalezienie tej transmitancji.

Estymacja WGM na podstawie modelu AR

Model AR stosowany jest szeroko w technice fonicznej, np. w modelu artykulacji mowy, w modelowaniu niektórych instrumentów muzycznych i in. Modeluje się również charakterystyki torów transmisyjnych, bezpośrednio na podstawie odbieranego sygnału.

Jeżeli $x[n]$ jest procesem AR rzędu p ($\{x[n]\} \in \text{AR}(p)$), to jego WGM określone jest wzorem

$$P_{\text{AR}(p)}(\omega) = \frac{\sigma^2}{\left| 1 + \sum_{m=1}^p a_m \exp(-j\omega m) \right|^2} \quad (2.5)$$

Aby znaleźć WGM trzeba znać σ^2 i a_m , $m = 1, 2, \dots, p$. Jeżeli obie strony równania (2.1) pomnożymy przez $x^*[n-l]$, $l = 0, 1, \dots$ i uśrednimy po zbiorze realizacji operatorem wartości oczekiwanej E , to otrzymamy

$$\underbrace{E\{x[n]x^*[n-l]\}}_{\substack{\text{autokorelacja} \\ r[l]}} = - \underbrace{\sum_{m=1}^p a_m E\{x[n-m]x^*[n-l]\}}_{\substack{\text{autokorelacja} \\ r[l-m]}} + \underbrace{E\{e[n]x^*[n-l]\}}_{=0} \quad (2.6)$$

Zerowanie się ostatniego składnika oznacza niezależność. Z (2.6) wynikają równania Youle-Walkera

$$r[l] = \begin{cases} - \sum_{m=1}^p a_m r[l-m], & l \neq 0 \\ - \sum_{m=1}^p a_m r[l-m] + \sigma^2, & l = 0 \end{cases} \quad (2.7)$$

Metoda autokorelacyjna

Dysponując estymatą ciągu autokorelacji (por. z (1.10))

$$\hat{r}[l] = \frac{1}{N} \sum_{k=0}^{N-1} x[k]x^*[k-l] \quad (2.8)$$

gdzie N to liczba próbek ciągu obserwacji $x[n]$ i gdzie $l = -(p-1), \dots, 0, \dots, (p-1)$, znajdujemy parametry modelu $\text{AR}(p)$ rozwiązując równanie macierzowe

$$\hat{\mathbf{R}}\hat{\mathbf{a}} = -\hat{\mathbf{r}} \quad (2.9)$$

gdzie

$\hat{\mathbf{a}} = [\hat{a}_1 \ \hat{a}_2 \ \dots \ \hat{a}_p]^T$ wektor współczynników autokorelacji (T oznacza tu transpozycję wektora)

$$\hat{\mathbf{R}} = \begin{bmatrix} \hat{r}[0] & \hat{r}[-1] & \hat{r}[-2] & \dots & \hat{r}[-p+1] \\ \hat{r}[1] & \hat{r}[0] & \hat{r}[-1] & \dots & \hat{r}[-p+2] \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \hat{r}[p-1] & \hat{r}[p-2] & \hat{r}[p-3] & \dots & \hat{r}[0] \end{bmatrix}$$

$\hat{\mathbf{r}} = [\hat{r}[1] \hat{r}[2] \cdots \hat{r}[p]]^T$ – wektor autokorelacji.

Rozwiązanie:

$\hat{\mathbf{a}} = -\hat{\mathbf{R}}^{-1}\hat{\mathbf{r}}$, przy czym

$$\sigma^2 = \hat{r}[0] + \sum_{m=1}^p \hat{a}_m \hat{r}^*[m].$$

Równania Youle-Walkera dla $x[n]$ – stacjonarnego procesu AR(p) o ciągu autokorelacji $r_x[l]$

$$r_x[l] = \begin{cases} -\sum_{m=1}^p a_m r_x[l-m], & l = 1, 2, \dots \\ -\sum_{m=1}^p a_m r_x[l-m] + \sigma^2, & l = 0 \\ r_x^*[-l], & l = -1, -2, \dots \end{cases} \quad (2.10)$$

gdzie σ^2 – to moc ciągu innowacji (szumu białego). Inaczej w postaci macierzowej

$$\begin{bmatrix} r_x[0] & r_x[-1] & \cdots & r_x[-p] \\ r_x[1] & r_x[0] & \cdots & r_x[-p+1] \\ \vdots & \vdots & \cdots & \vdots \\ r_x[p] & r_x[p-1] & \cdots & r_x[0] \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \text{układ równań normalnych}$$

Macierz autokorelacji $(p+1) \times (p+1)$ jest macierzą Toeplitza. Jest to dodatnio określona macierz Hermite'a. Szybkie odwracanie tej macierzy uzyskuje się za pomocą algorytmu Levinsona-Durbina.

W praktyce, dysponując obserwacją $x[n]$, najpierw estymujemy ciąg autokorelacji $\{\hat{r}_x[n]\}_{-L}^L$. Następnie formujemy macierz estymat autokorelacji $(p+1) \times (p+1)$, $p \ll L$.

W końcu obliczamy wektor współczynników $\{\hat{a}_m\}_1^p$. Są one nazywane współczynnikami liniowej predykcji (ang. LPC – *linear predictive coding/coefficients*).

Rząd modelu p , to w przybliżeniu liczba zespolonych sinusoid wykrywanych przez $P_{AR(p)}(\omega)$ (2.5).

Model AR(p) stacjonarnego ciągu próbek $x[n]$

Równanie różnicowe procesu autoregresji AR(p) rzędu p

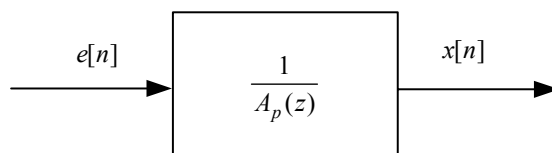
$$x[n] + \sum_{m=1}^p a_m x[n-m] = e[n] \quad (5.1)$$

albo

$$\sum_{m=0}^p a_m x[n-m] = e[n]$$

gdzie zawsze $a_1 = 1$,

$e[n]$ – stacjonarny szum biały – ciąg innowacji. Schemat blokowy modelu $AR(p)$ pokazano poniżej, gdzie $x[n] = e[n] - \sum_{m=1}^p a_m x[n-m]$ i $A_p(z) = \sum_{m=0}^p a_m z^{-m}$; $\{a_m\}_1^p$ – współczynniki modelu $AR(p)$.



Rys. 7. Schemat blokowy modelu $AR(p)$.

Uwaga: **predyktor** działa wg algorytmu

$$\hat{x}[n] = -\sum_{m=1}^p a_m x[n-m] = -a_1 x[n-1] - a_2 x[n-2] - \dots - a_p x[n-p]$$

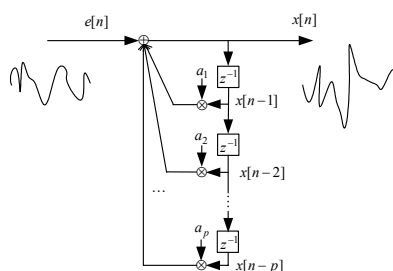
z błędem predykcji

$$e[n] = x[n] - \hat{x}[n] = \sum_{m=0}^p a_m x[n-m].$$

Filtr rekursywny $AR(p)$ – stabilny, przyczynowy (tj. realizowalny w czasie rzeczywistym) i odwracalny

Model $AR(p)$ o strukturze transwersalnej (bezpośredniej) ma transmitancję $H(z) = \frac{X(z)}{E(z)} = \frac{1}{1 + \sum_{m=1}^p a_m z^{-m}}$. Filtr do niego odwrotny operuje transmitancją

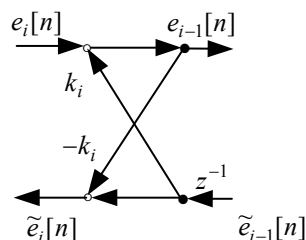
$$H^{-1}(z) = 1 + \sum_{m=1}^p a_m z^{-m}.$$



Rys. 8. Struktura transwersalna modelu $AR(p)$.

Struktura kratowa (ang. *lattice*) Markela i Graya modelu $AR(p)$. Algorytm i -tego ogniwa kraty

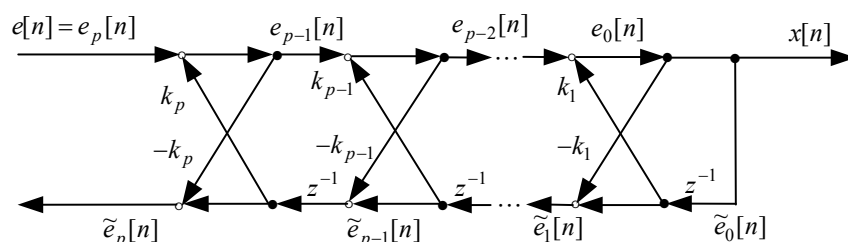
$$\begin{aligned} e_{i-1}[n] &= e_i[n] + k_i \tilde{e}_{i-1}[n-1] \\ \tilde{e}_i[n] &= -k_i e_{i-1}[n] + \tilde{e}_{i-1}[n-1] \end{aligned}$$



Rys. 9. Pojedyncze ogniwo struktury kratowej modelu AR(p).

Na rys. 9 oznaczono sygnały na wyjściach sumatorów. Cały filtr kratowy, pokazany na rys. 10, ma transmitancję

$$H(z) = \frac{X(z)}{E(z)} = \frac{1}{A_p(z)} = \frac{1}{1 + \sum_{m=1}^p a_m z^{-m}}$$



Rys. 10. Struktura kratowa (ang. *lattice*) Markela i Graya modelu AR(p).

Algorytm Levinsona-Durбина obliczania współczynników k_i , $i = 1, 2, \dots, p$ na podstawie współczynników a_m , $m = 1, 2, \dots, p$

- (1) $a_m^{(p)} = a_m$
- (2) $k_p = a_p^{(p)}$
- (3) $k_i = a_i^{(i)}$
- (4) $a_m^{(i-1)} = \frac{a_m^{(i)} + k_i a_{i-m}^{(i)}}{1 - k_i^2}$; $m = 1, 2, \dots, (i-1)$

Rekursję powtarza się dla $i = p, p-1, \dots, 2, 1$, a więc dla malejących indeksów i , aż do obliczenia wszystkich k_i , $i = 1, 2, \dots, p$.

Przykład 2.

$$\text{Dla } H(z) = \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3}} = \frac{1}{1 + 0.9z^{-1} - 0.64z^{-2} + 0.57z^{-3}}, \quad N = 3$$

dostajemy

$$\begin{aligned}
z(1) \quad & a_1^{(3)} = a_1 = 0.9, \quad a_2^{(3)} = a_2 = -0.64, \quad a_3^{(3)} = a_3 = 0.576 \\
z(2) \quad & i = p = 3, \quad k_{p3} = a_3^{(3)} = 0.576 \\
z(4) \quad & m = 1, \quad a_1^{(2)} = \frac{a_1^{(3)} + k_3 a_2^{(3)}}{1 - k_3^2} = \frac{0.9 + 0.576 \times (-0.64)}{1 - 0.576^2} \cong 0.7952 \\
z(4) \quad & m = 2, \quad a_2^{(2)} = \frac{a_2^{(3)} + k_3 a_1^{(3)}}{1 - k_3^2} = \frac{-0.64 + 0.576 \times 0.9}{1 - 0.576^2} \cong -0.182 \\
z(3) \quad & i = p - 1 = 2, \quad k_2 = a_2^{(2)} = -0.182 \\
z(4) \quad & m = 1, \quad a_1^{(1)} = \frac{a_1^{(2)} + k_2 a_1^{(2)}}{1 - k_2^2} = \frac{0.7952 - 0.182 \times 0.7952}{1 - (-0.182)^2} \cong 0.673 \\
z(3) \quad & i = p - 2 = 1, \quad k_1 = a_1^{(1)} = 0.673
\end{aligned}$$

Struktura kratowa ma bardzo dobre właściwości numeryczne. Współczynniki tej struktury, k_i , $i = 1, 2, \dots, p$ nazywa się współczynnikami odbicia (od ang. *reflection coefficients*). Jeszcze inna nazwa to PARCORy (od ang. *partial correlation coefficients*). Są to współczynniki o wartościach $|k_i| < 1$, pozwalające na tworzenie stabilnych filtrów. Nie wymagają one żmudnego przeprojektowania całej struktury kratowej, gdy trzeba zwiększyć liczbę ogniw kraty. Takiej właściwości nie ma struktura transwersalna.

Sygnał mowy można modelować za pomocą struktury kratowej, podobnie jak sygnały instrumentów muzycznych czy trakt głosowy. Można znaleźć WGM tak wygenerowanych sygnałów oraz transmitancję traktu głosowego. W technice fonicznej bardzo często są wykorzystywane struktury kratowe z podanych wyżej powodów. Filtr kratowy wykorzystuje się np. do rekonstrukcji sygnału fonicznego obciążonego szumem i zakłóceniami, metodą interpolacji adaptacyjnej, opierając się na założeniu, że sygnał akustyczny można zamodelować jako proces autoregresywny [2].

Literatura uzupełniająca z analizy widmowej

- [1] E.J. Ericsson, L. F. Cepeda, R. D. Rodman, K. P. H. Sullivan, D. F. McAllister, D. Bitzer and P. Arroway: Robustness of Spectral Moments: a Study using Voice Immitations, Proc. of the 10th Australian International Conference on Speech Science and Technology, Macquarie University, Sydney, December 8 to 10, 2004, pp. 259-264. Dostępne w Internecie.
- [2] A. Czyżewski: Dźwięk cyfrowy. Wybrane zagadnienia teoretyczne, technologia, zastosowania. Akademicka Oficyna Wydawnicza EXIT, Warszawa 1998, p. 4.3 str. 156 – Predykcja liniowa, filtry kratowe, też na str. 177, filtry formantowe – str. 185.
- [3] J. Szabatin: Podstawy teorii sygnałów. WKiŁ, Warszawa 1982 i późniejsze wydania, moc zdefiniowana na str. 27, momenty jako parametry zmiennych losowych – str. 147, widmo gęstości energii i widmo gęstości mocy str. 305 i 307.
- [4] L. Rutkowski: Filtry adaptacyjne i adaptacyjne przetwarzanie sygnałów. Teoria i zastosowania. WNT, Warszawa 1994, rozdz. 3, str. 49 – Zagadnienia liniowej predykcji, str. 55 – właściwości macierzy autokowariancji.
- [5] A. Wojnar: Teoria sygnałów. WNT, Warszawa 1980, widmo gęstości energii str. 23, momenty procesu stacjonarnego str. 33, szerokość widma str. 56.
- [6] J. Sobkowski: Częstotliwościowa analiza sygnałów. Wydawnictwo Ministerstwa Obrony Narodowej, Warszawa 1975, Widmo gęstości energii i widmo gęstości mocy

- p. 5.2.5, str. 170, momenty jako charakterystyki procesów przypadkowych – p. 9.1.1, str. 437.
- [7] W. Borodziejewicz, K. Jaszczak: Cyfrowe przetwarzanie sygnałów. WNT, Warszawa 1987, Durbin – str. 131 w rozdziale pt. Liniowa predykcja sygnału mowy.
- [8] S. W. Smith: Cyfrowe przetwarzanie sygnałów. Praktyczny poradnik dla inżynierów i naukowców. Wydawnictwo BTC, Warszawa 2003, str. 351 – Synteza i rozpoznawanie mowy,
- [9] T. Zieliński: Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań. WKŁ, Warszawa 2005, momenty na str. 6 w tablicy, korelacja str. 14, periodogram i modelowanie AR, MA i ARMA ale bez struktur, WGM. Też str. 447, STFT str. 455.

Pytania z analizy widmowej dla Akustyki Mowy

1. Czym zajmuje się analiza widmowa sygnałów mowy? Co to widmo gęstości mocy (WGM) i jak się to widmo oblicza?
2. Jaki jest związek pomiędzy WGM na wyjściu i na wejściu systemu dyskretnego? Objaśnij użyte symbole.
3. Zdefiniuj periodogram. Objaśnij jak się oblicza periodogram Shustera i w jakim celu się go stosuje.
4. Do czego wykorzystuje się momenty widmowe i na jakiej podstawie się je oblicza?
5. Zapisz formuły na najważniejsze momenty widmowe. Objaśnij użyte symbole.
6. Podaj interpretację momentów widmowych unormowanych, centralnych, rzędu drugiego i trzeciego.
7. Co to jest miara płaskości widma (ang. *Spectral Flatness Measure – SFM*)? Jak można ją obliczyć dla fragmentu widma?
8. Dla jakich sygnałów używamy nieparametryczną, a dla jakich parametryczną estymację widma? Wymień etapy parametrycznej estymacji widma.
9. Wymień najczęściej stosowane modele parametryczne i naszkicuj typowe widma dla modeli parametrycznych.
10. Objaśnij w jaki sposób wykonuje się estymację WGM na podstawie modelu autoregresji.
11. Narysuj modele autoregresji o strukturach: transwersalnej i kratowej. Objaśnij użyte symbole.
12. Jakie właściwości ma struktura kratowa modelu autoregresji i gdzie znajduje zastosowanie?

E. Hermanowicz
29.03.2011