Transformaty

w kompresji sygnałów

KODERY TRANSFORMATY

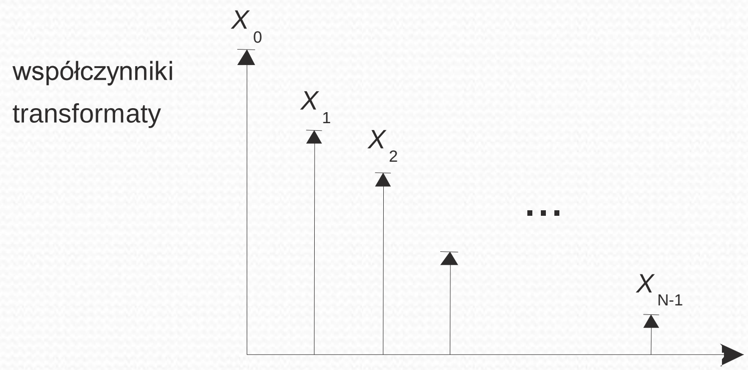
$$\overline{X} = \overline{W}\,\overline{x} \qquad \begin{vmatrix} X_0 \\ X_1 \\ \vdots \\ X_{N-1} \end{vmatrix} = \overline{W}_{NxN} \begin{vmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{vmatrix}$$

- obliczenie transformaty $\overline{X} = \overline{W}\,\overline{x}$
- kwantowanie transformaty (wykorzystanie koncentracji energii w niewielkiej liczbie współczynników transformaty)

$$\overline{X} \;\; \text{->} \;\; \overline{X}^*$$

- obliczenie transformaty odwrotnej $\overline{x}^* = \overline{W}^{-1}\,\overline{X}^*$

KODERY TRANSFORMATY



Kwantowanie współczynników transformaty:

1. skalarne (można wykorzystać kwantyzatory o różnym zakresie pracy i liczbie poziomów) $L_i = 2^{b_i}$ - np. MP3

2. wektorowe w podzakresach częstotliwości (z różną rozdzielczością) - np. kodery ATRAC

Kwantyzator skalarny

Kwantyzator skalarny:
rozdzielczość b bit/próbkę,
$L=2^b$ poziomów kwantyzacji, rozmieszczonych optymalnie.
Sygnał x_n o rozkładzie gaussowskim, wartość średnia =0, wariancja σ_x^2.

Moc szumów kwantyzacji przy bezpośrednim kwantowaniu próbek x_n :

$$\sigma_e^2(b) = c(1)\,\sigma_x^2\,2^{-2b}$$

gdzie $c(1) = \frac{\sqrt{3}}{2}\pi$ (dla b>>1)

Stosunek mocy sygnału do szumu kwantyzacji:

$$SNR = \sigma_x^2 / \sigma_e^2(b) = \frac{1}{c(1)}2^{2b} \approx 0.367\ 2^{2b}$$

$$SNR[dB] = 10\log_{10}[\sigma_x^2 / \sigma_e^2(b)] \approx 10\log_{10}(0.367) + 20b\ \log_{10} 2 \approx 6b - 4.3dB$$

Transformata unitarna

Transformata T: macierz NxN:

y=Tx   (gdzie x: N-wymiarowy wektor próbek sygnału x_n),
y=$(y_1, y_2,\ldots,y_N)^t$ – przetransformowany wektor próbek.

Transformata jest ortogonalna (unitarna): $T^t T = T T^t = I$, gdzie t - transpozycja, I - macierz jednostkowa.

Transformata unitarna zachowuje wariancję:

$$\sigma_y^2 = \tfrac{1}{N} E(y^t y) = \tfrac{1}{N} E(x^t T^t T x) = \tfrac{1}{N} E(x^t x) = \sigma_x^2$$

a więc szum kwantyzacji składowych wektora y pojawi się z tą samą mocą w odtworzonym (za pomocą transformaty odwrotnej T^t) sygnale x.

Zysk transformaty

Kwantujemy próbki przetransformowanego wektora, używając N kwantyzatorów o rozdzielczości b_1, b_2,…b_N. Średnio mamy $b = \frac{1}{N}\sum_{i=1}^{N} b_i$ bitów na próbkę.

Średnia moc błędu kwantyzacji $\sigma_e^2(T,b) = \frac{c(1)}{N}\sum_{k=1}^{N}\sigma_k^2 2^{-2b_k}$ (ma być jak najmniejsza)

gdzie σ_k^2 średnia moc k-tej składowej wektora y. Średnia arytmetyczna nie jest mniejsza niż geometryczna:

$$\sigma_e^2(T,b) = \frac{c(1)}{N}\sum_{k=1}^{N}\sigma_k^2 2^{-2b_k} \geq c(1)\left[\prod_{k=1}^{N}\sigma_k^{2/N}\right]2^{-2\Sigma\, b_k/N} = c(1)\,2^{-2b}\prod_{k=1}^{N}\sigma_k^{2/N}$$

Równość (i najmniejszy średni błąd kwantowania) wystąpi gdy składniki sumy będą identyczne:

tzn. dla każdego k $\sigma_k^2 2^{-2b_k} = 2^{-2b}\prod_{k=1}^{N}\sigma_k^{2/N} = const$

Obliczamy stąd optymalny rozdział bitów (rozdzielczości) pomiędzy kwantyzatory:

$$\sigma_k^2 2^{-2b_k} = const \qquad \sigma_k^2[dB] - 6\,b_k = const'$$

$$b_k = \tfrac{1}{6}\sigma_k^2[dB] + const''$$

Zysk transformaty

Dla optymalnego rozdziału bitów: $$\sigma_e^2(T,b) = \frac{c(1)}{N}\sum_{m=1}^{N}\left[2^{-2b}\prod_{k=1}^{N}\sigma_k^{2/N}\right] = c(1)\,2^{-2b}\prod_{k=1}^{N}\sigma_k^{2/N}$$

Zysk ze skalarnego kwantowania współczynników transformaty w stosunku do skalarnego kwantowania próbek sygnału x:

$$\frac{\sigma_e^2(b)}{\sigma_e^2(T,b)} = \frac{c(1)\sigma_x^2\,2^{-2b}}{c(1)\,2^{-2b}\prod_{k=1}^{N}\sigma_k^{2/N}} = \frac{\sigma_x^2}{\prod_{k=1}^{N}\sigma_k^{2/N}} = \frac{\frac{1}{N}\sum_{k=1}^{N}\sigma_k^2}{\prod_{k=1}^{N}\sigma_k^{2/N}}$$

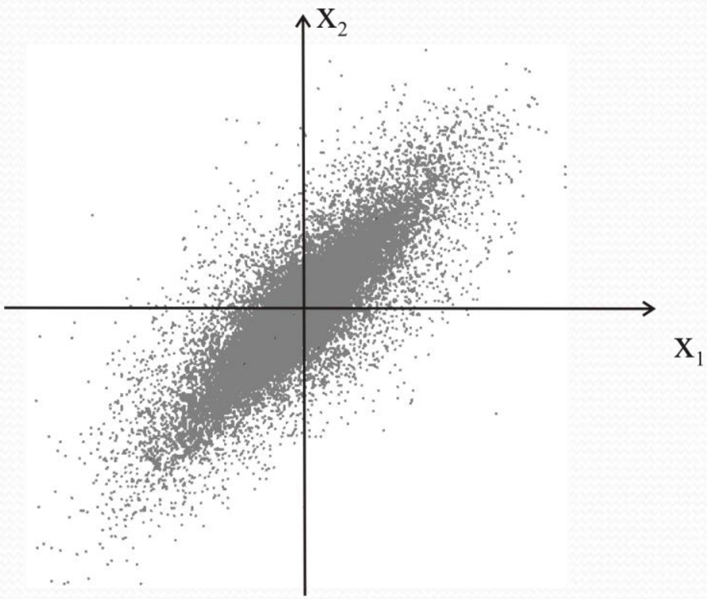
gdyż $\sigma_x^2 = \sigma_y^2 = \frac{1}{N}\sum_{k=1}^{N}\sigma_k^2$

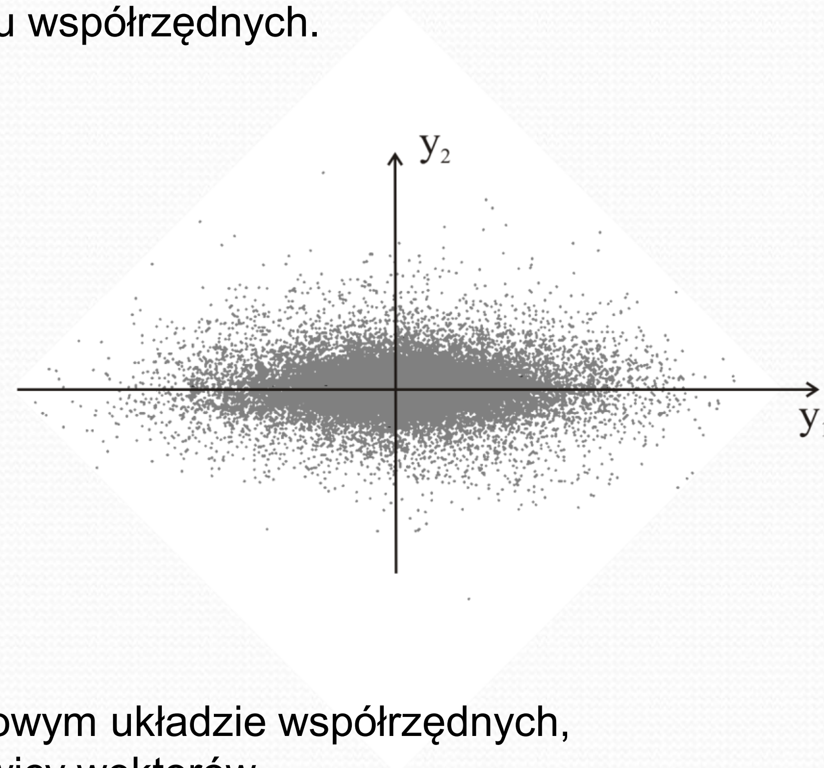
Zysk zależy od transformaty i nie może być < 1 (średnia arytmetyczna niemniejsza od średniej geometrycznej)

Wygramy tym więcej im większy będzie rozrzut wartości wariancji współczynników transformaty $y_1, y_2, \ldots, y_N$

Jak powiększyć zysk transformaty?

Transformata ortogonalna T: $T T^t = T^t T = I$ (I – macierz jednostkowa)
Wiersze T tworzą nową bazę ortogonalną.
Transformacja y=Tx to zmiana układu współrzędnych.

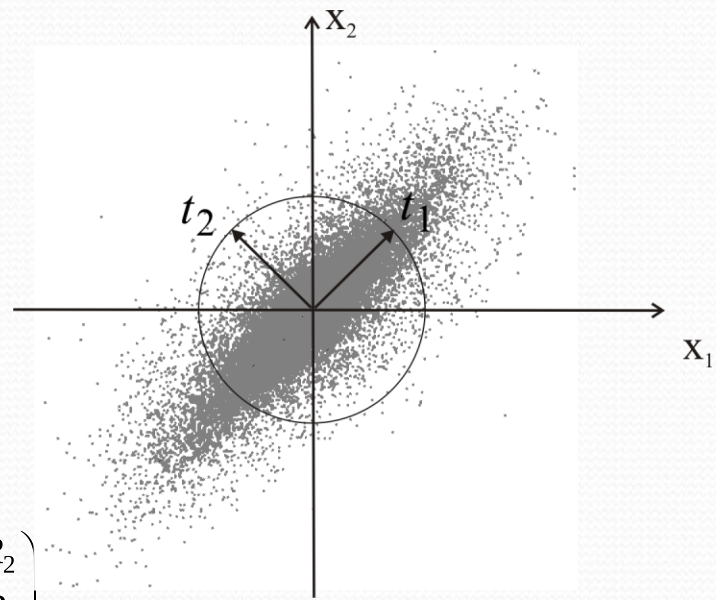




Należy powiększyć rozrzut wariancji w nowym układzie współrzędnych, tzn. wykorzystać osie elipsoidalnej mgławicy wektorów.
W tym układzie współrzędnych składowe wektora y (współczynniki transformaty) nie będą skorelowane.

Jak znaleźć dekorelującą transformatę?

Osie nowego układu współrzędnych to wektory własne macierzy kowariancji sygnału.



Macierz kowariancji (tu równa macierzy korelacji, gdyż wartość średnia próbek sygnału = 0)

$$R(\bar{x}) = E(\bar{x}\,\bar{x}^t)$$

Np. dla N=3:

$$R(\bar{x}) = E\begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix}\begin{pmatrix} x_0 & x_1 & x_2 \end{pmatrix} = \begin{pmatrix} E(x_0^2) & E(x_0x_1) & E(x_0x_2) \\ E(x_1x_0) & E(x_1^2) & E(x_1x_2) \\ E(x_2x_0) & E(x_2x_1) & E(x_2^2) \end{pmatrix} = \begin{pmatrix} R_0 & R_1 & R_2 \\ R_1 & R_0 & R_1 \\ R_2 & R_1 & R_0 \end{pmatrix}$$

gdzie $R_i = E(x_n x_{n-i})$ - współczynnik autokorelacji

Gdy sygnał jest nieskorelowany, wówczas $R_i = E(x_n x_{n-i}) = 0, \quad i = 1,2,\ldots$

Właściwości dekorelujące ma transformata Karhunena- Loevego (KLT)

Wektory własne macierzy kowariancji tworzą wiersze macierzy transformaty Karhunena-Loevego, która dekoreluje sygnał i maksymalizuje zysk transformaty, zapewniając największą wartość SNR po odpowiednim rozdziale bitów.

Macierz kowariancji $R(\bar{x}) = E[\bar{x}\,\bar{x}^t]$ (równa macierzy korelacji, gdyż $E[\bar{x}] = 0$)

$\bar{x}$ – N-wymiarowy wektor próbek sygnału

Wektor własny t_k : $R(\bar{x})\,t_k = \lambda_k\,t_k$, wartość własna λ_k

Transformata KLT: $\bar{T} = [t_1, t_{2,} \ldots, t_N]^t$

$$R(\bar{x})\bar{T}^t = [\lambda_1 t_1, \lambda_2 t_{2,} \ldots, \lambda_N t_N] = [t_1, t_2, \ldots, t_N]\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{bmatrix} = \bar{T}^t\,\Lambda$$

$\bar{y} = \bar{T}\,\bar{x}$ – wektor po transformacji

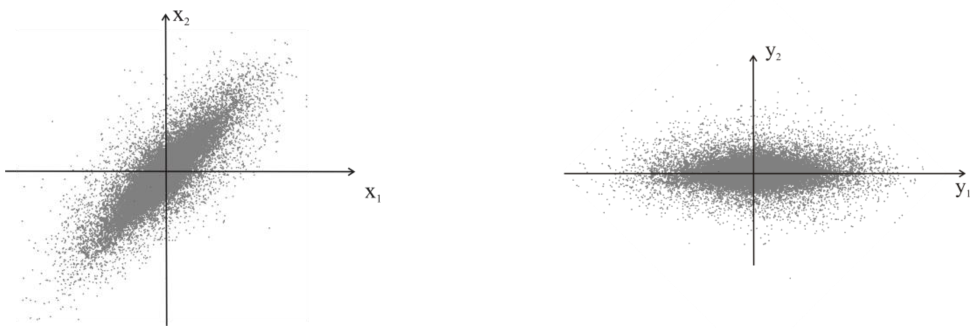
Jego macierz korelacji:

$$R(\bar{y}) = E[\bar{y}\,\bar{y}^t] = E[\bar{T}\bar{x}\bar{x}^t\bar{T}^t] = \bar{T}\,R(\bar{x})\bar{T}^t = \bar{T}\bar{T}^t\Lambda$$

Wektory własne są ortogonalne, więc $\bar{T}\,\bar{T}^t = I$

(I - macierz jednostkowa)

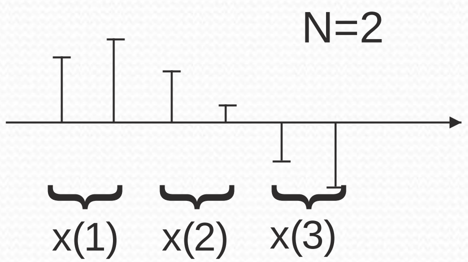
$\Rightarrow \quad R(\bar{y}) = \Lambda$ $\equiv$ KLT dekoreluje sygnał



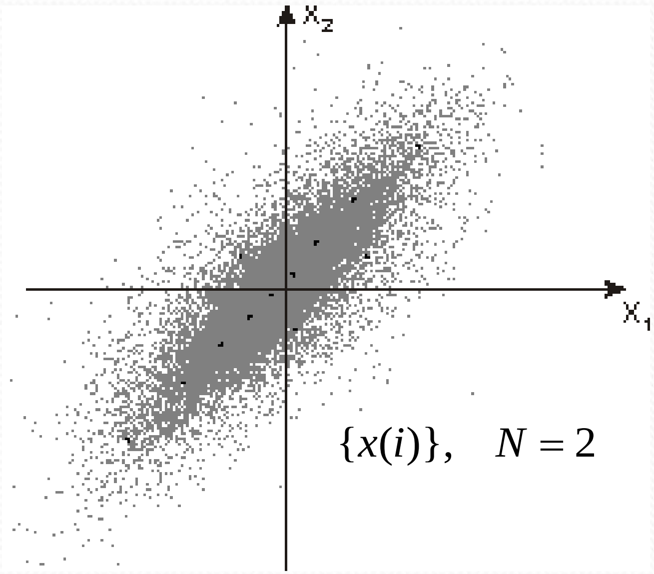
Zysk transformaty KLT zależy od rozrzutu wartości własnych

(wartości własne to moce składowych wektora $\bar{y} = \bar{T}\,\bar{x}$)

Kwantyzator wektorowy
VQ - vector quantizer

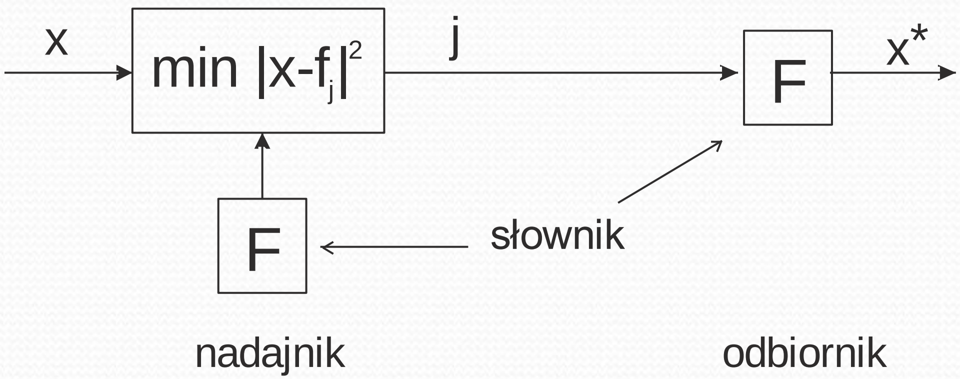


{x(i)} - zbiór wektorów
N - wymiarowych

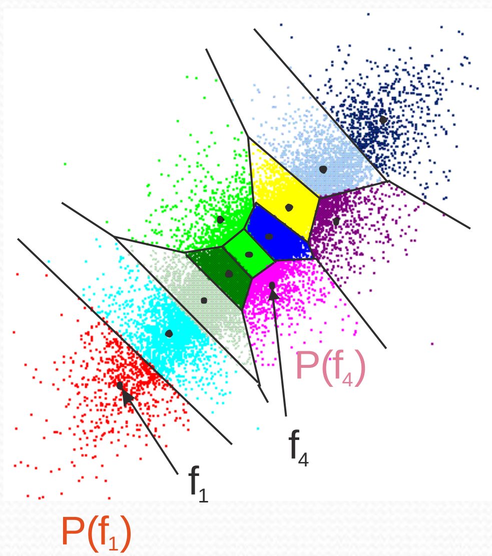


F: $\{f_1,...,f_L\}$ - słownik (książka kodowa)
zbiór L wektorów wzorcowych,
odpowiedników poziomów kwantowania

$L=2^{bN}$ b - rozdzielczość [bit/próbkę]



Georgy Voronoi 1868-1908, matematyk ukraiński, profesor UW.



$P(f_1),\ldots,P(f_L)$:

komórki Voronoi'a

N-wymiarowe odpowiedniki przedziałów kwantyzacji

Oszacowanie zniekształceń wprowadzanych przez N-dim VQ o rozdzielczości b bit/próbkę, dla skorelowanego sygnału gaussowskiego:

$$\sigma_e^2 \approx c(N)\left[\det R(x)\right]^{1/N} \; 2^{-2b}$$

gdzie $c(1) = \frac{\sqrt{3}}{2}\pi, \ldots\ldots, c(N) \approx \frac{c(1)-1}{N} + 1$

dla $N \to \infty, \quad c(N) \to 1$

R(x) - macierz korelacji N-dim wektorów sygnału $R(x) = E[x\,x^T] = \begin{matrix} R_0 & R_1 & \cdots & R_{N-1} \\ R_1 & R_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_1 \\ R_{N-1} & \cdots & R_1 & R_0 \end{matrix}$

np. dla sygnału gaussowskiego nieskorelowanego:

$$R_0 = \sigma_x^2, \;\; R_1 = R_2 = \cdots = 0, \qquad \det R(x) = [\sigma_x^2]^N$$

$$\sigma_e^2 \approx c(N)\;\sigma_x^2\; 2^{-2b}$$

Kwantyzacja wektorowa przetransformowanego sygnału

$$y = Tx \qquad \sigma_y^2 = \sigma_x^2$$

$$R^N(y) = E(yy^t) = T\,E(xx^t)\,T^t = T\,R^N(x)\,T^t$$

$$\det R^N(y) = \det(TT^t)\det R^N(x) = \det I \det R^N(x) = \det R^N(x)$$

Kwantując wektorowo *N* – wymiarowy wektor y otrzymujemy błąd kwantyzacji

$$\sigma_e^2(T,b,N) \approx c(N)\left[\det R(y)\right]^{1/N} 2^{-2b} = c(N)\left[\det R(x)\right]^{1/N} 2^{-2b}$$

Jest on taki sam jak błąd kwantowania wektorowego *N* – wymiarowego wektora x przed transformacją. Transformata nic tu nie wnosi, zysk = 1.

Kwantyzacja wektorowa przetransformowanego sygnału, c.d.

Zysk można osiągnąć, gdy przetransformujemy duży blok ($\overline{N}$) próbek,
a próbki transformaty podzielimy na *M* wektorów *N*=wymiarowych.
Każdy z tych wektorów będziemy kwantować osobnym kwantyzatorem wektorowym
(odrębny słownik i liczba wektorów w słowniku)

$$\overline{y} = \overline{T}\overline{x}, \quad \overline{y} = \begin{matrix} \overline{y}^{1} \\ \vdots \\ \overline{y}^{M} \end{matrix}$$

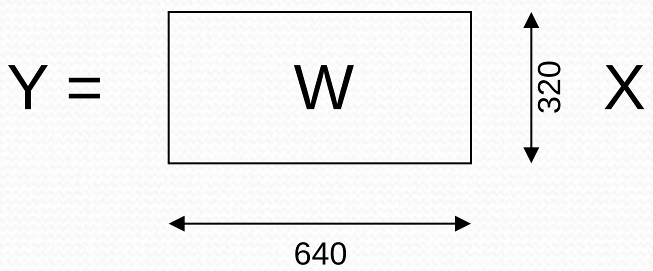
Przy optymalnym rozdziale bitów między kwantyzatory, zysk transformaty
(w odniesieniu do kwantyzacji wektorowej *N*=wymiarowych wektorów *x*) wyniesie

$$\frac{\sigma_e^2(b,N)}{\sigma_e^2(\overline{T},b,N)} = \frac{c(N)[\det R^N(x)]^{1/N}\ 2^{-2b}}{c(N)\prod_{k=1}^{M}[\det R^N(\overline{y}^k)]^{1/NM}\, 2^{-2b}}$$

Kwantyzacja wektorowa przetransformowanego sygnału, c.d.

Przykład: koder G.722.1

- Częstotliwość próbkowania 16 kHz
- Pasmo sygnału akustycznego 7 kHz
- Transformata MLT (MDCT)

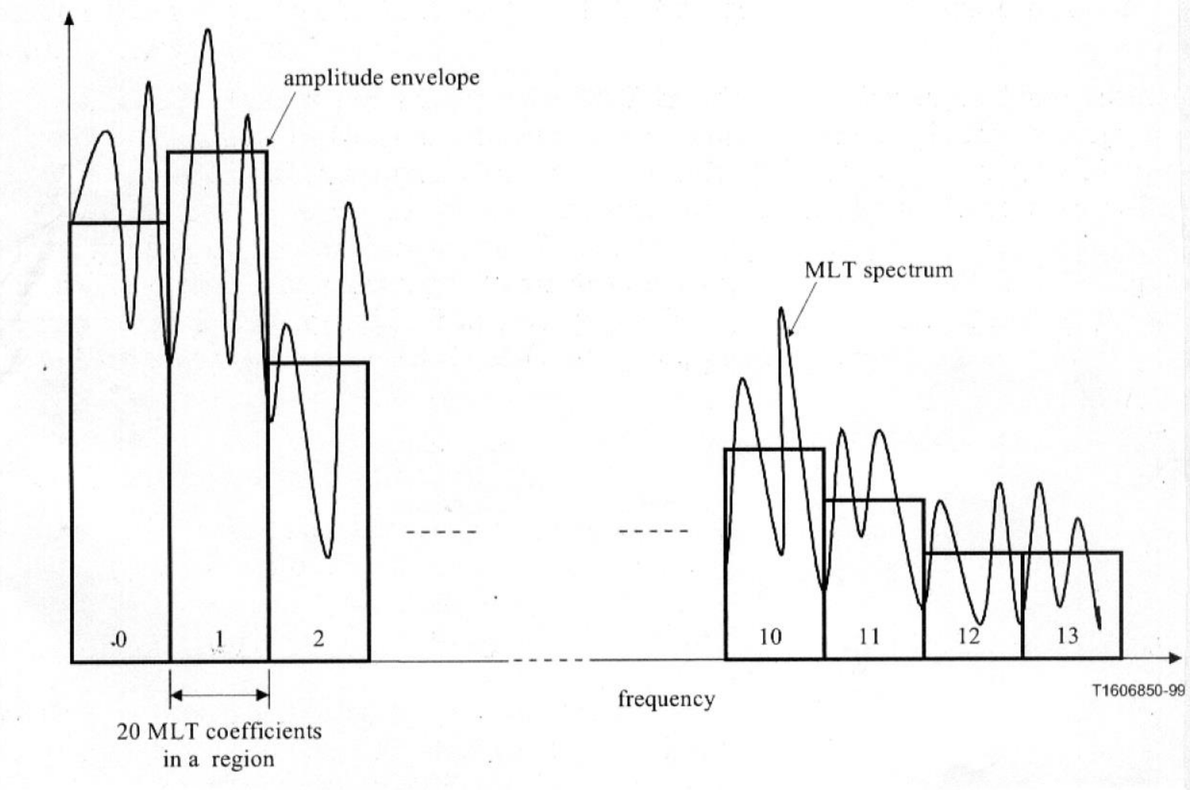


- 14 podpasm po 20 próbek transformaty (szerokość podpasma = 0.5 kHz)
  - podpasm 7-7.5 kHz i 7.5-8 kHz nie koduje się
- Przepływności 24 i 32 kbit/s

Kwantyzacja wektorowa przetransformowanego sygnału, c.d.

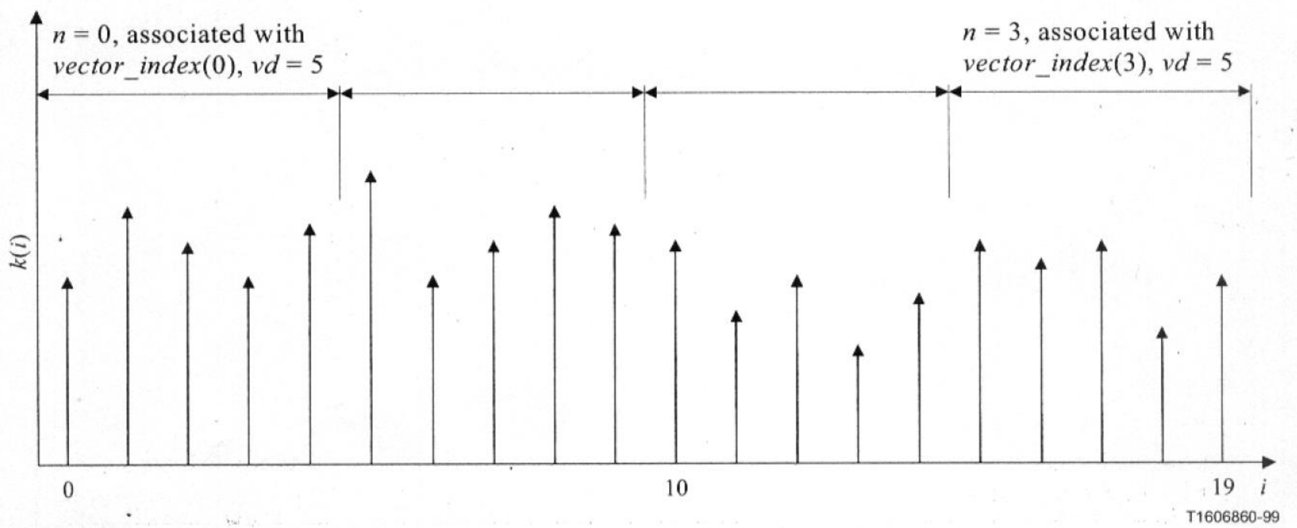
14 podpasm

W każdym podpaśmie oblicza się obwiednię widma amplitudy



Kwantyzacja wektorowa przetransformowanego sygnału, c.d.

Każde podpasmo dzieli się na vpr wektorów o wymiarze vd (vpr · vd = 20).
Słownik kwantyzatora wektorowego liczy u wektorów.



NOTE – Each vector represents *vd* quantized MLT coefficients.