

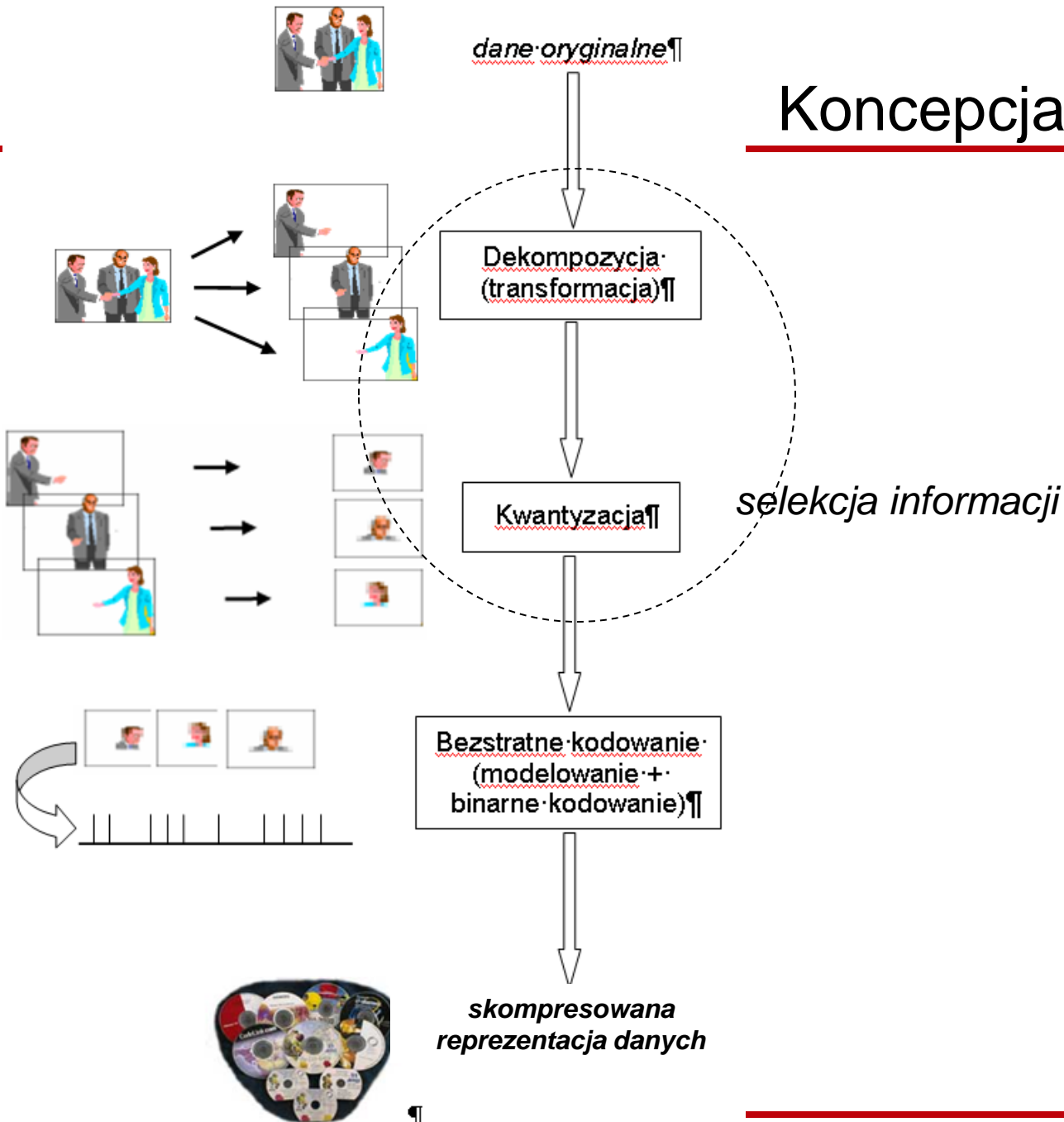
# KWANTYZACJA JAKO METODA SELEKCJI INFORMACJI

---

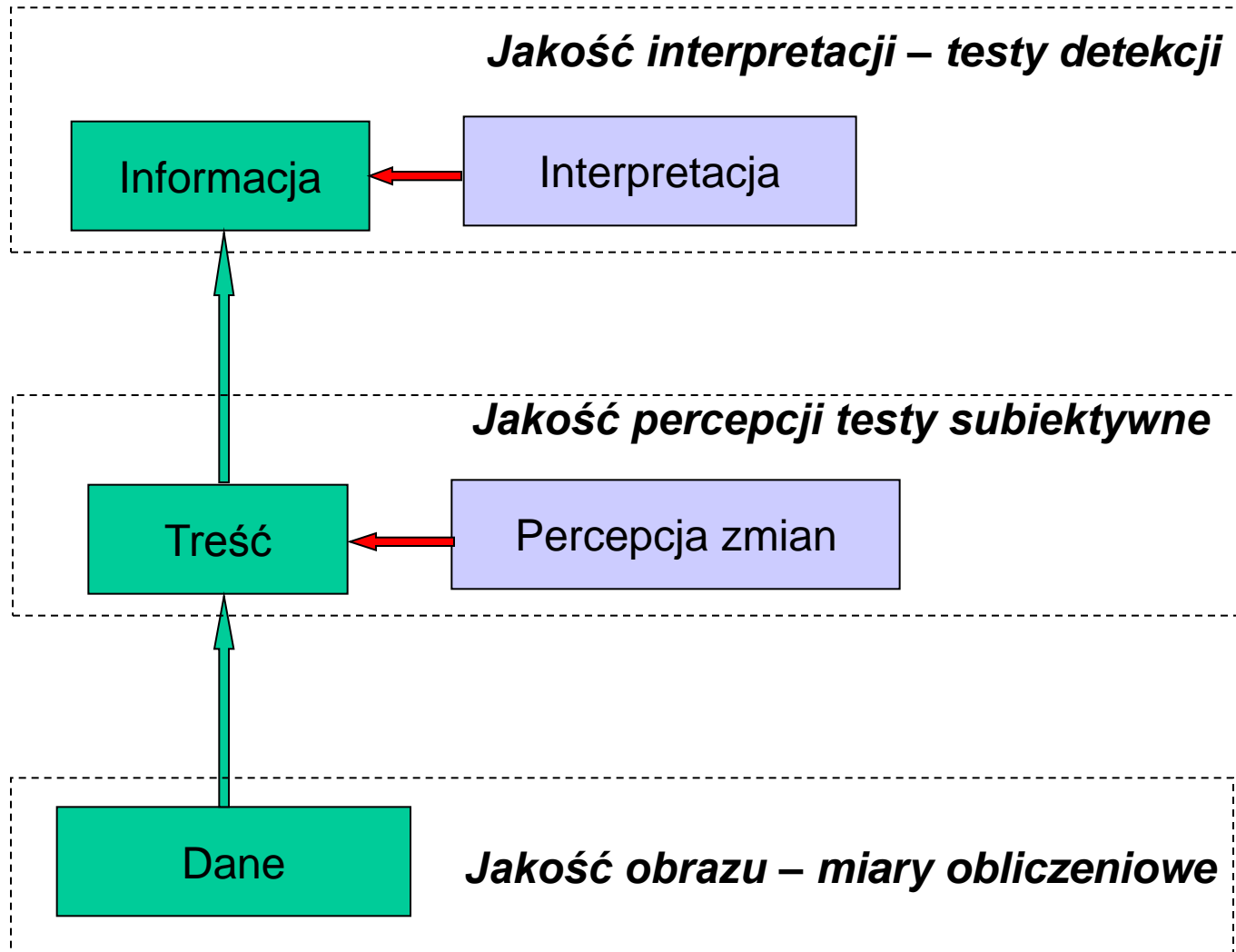
Materiały KODA, A.Przelaskowski

- Koncepcja metod z selekcją (stratnych, nieodwracalnych)
  - Podstawowe metody kwantyzacji
  - Stratna DPCM, BTC
  - Kwantyzacja wektorowa
  - TCQ
  - Testy
-

# Koncepcja podstawowa



# Odkrywanie treści, metody i ocena jakości



# Obliczeniowe miary porównawcze

- maksymalna różnica (ang. *Maximum Difference*), zwana też szczytowym błędem bezwzględnym PAE (ang. *Peak Absolute Error*):

$$MD = \max_{m,n} \{ |f(m,n) - \tilde{f}(m,n)| \};$$

$f(m,n), \tilde{f}(m,n)$ porównywane obrazy $0 \leq m < M, 0 \leq n < N$
--

- błąd średniokwadratowy (ang. *Mean Square Error*):

$$MSE = \frac{1}{MN} \sum_{m,n} [f(m,n) - \tilde{f}(m,n)]^2;$$

- szczytowy stosunek sygnału do szumu (ang. *Peak Signal to Noise Ratio*):

$$PSNR = 10 \log \frac{MN \cdot [\max_{m,n} \{f(m,n)\}]^2}{\sum_{m,n} [f(m,n) - \tilde{f}(m,n)]^2},$$

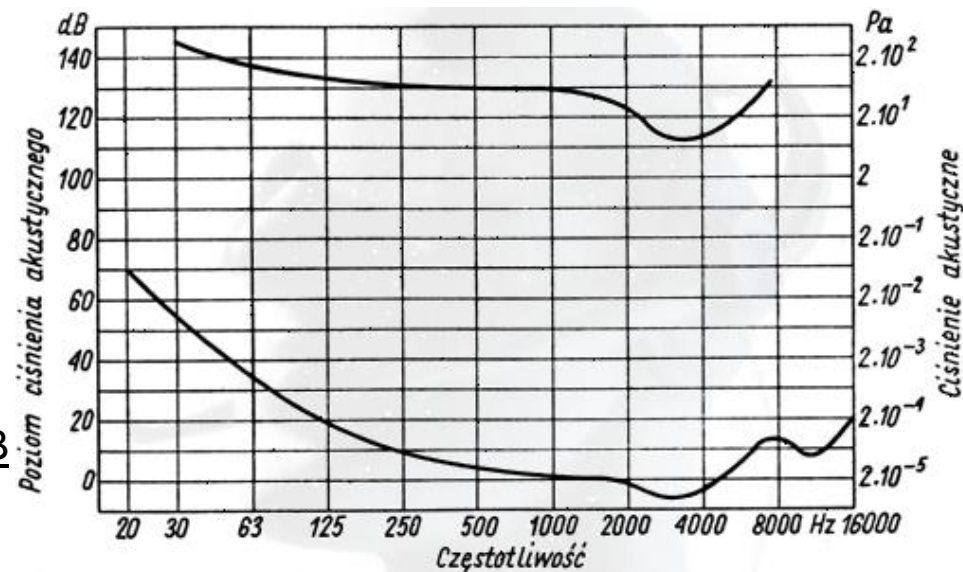
przy czym wartość  $\max_{m,n} \{f(m,n)\}$  jest zwykle ustalana na poziomie największej możliwej (a nie faktycznej) wartości funkcji jasności, np. 255 dla danych 8-bitowych;

- średnia różnica (ang. *Average Difference*):

$$AD = \frac{1}{MN} \sum_{m,n} |f(m,n) - \tilde{f}(m,n)|;$$

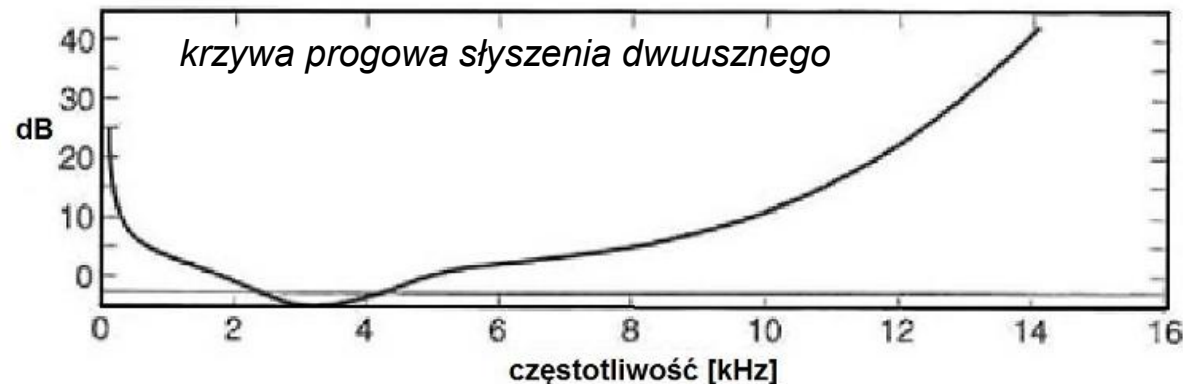
# Percepcja dźwięku - cd

- Czułość ucha zależy od częstotliwości odbieranego dźwięku - zakres najlepszej słyszalności to przedział częstotliwości od ok. 1 do 5kHz, przy czym minimum czułości wypada na około 3,5kHz
- Zakres natężenia odbieranych dźwięków (tj. dynamika słuchu) wynosi ok. 120dB, przy czym
  - dolna granica słyszalności (próg czułości słuchu) odpowiada minimalnej wartości ciśnienia akustycznego, przy której można uzyskać ledwie postrzegalne wrażenie słyszenia dźwięku
  - górna granica słyszalności jest określona przez minimalną wartość ciśnienia akustycznego dźwięku, przy której ucho reaguje bólem.- dźwięki o poziomie natężenia przekraczającym 140dB mogą uszkodzić słuch
  - najmniejsza zmiana poziomu ciśnienia powodująca zmianę wrażenia głośności, tzw. próg różnicy głośności, wynosi około 0,2 dB dla tonu 1000 Hz
  - maksymalna dynamika mowy to 50dB
  - dynamika natężenia dźwięku mierzona w czasie kilkugodzinnego koncertu dużej orkiestry symfonicznej nie przekracza 80dB
  - dynamika poszczególnych instrumentów nie przekracza 50 dB



# Efekty ludzkiej percepcji

- Precedens - krótkie dźwięki (impulsy, transjenty) docierające w małym odstępie czasu (1 - 5 ms dla impulsów tonu, do 40 ms dla impulsów mowy lub muzyki) słyszane są jako jeden dźwięk, którego lokalizacja jest zdeterminowana przez kierunek pierwszego
  - dźwięk wtórny ma niewielki wpływ na lokalizację, jeśli dociera z kierunku bardzo odległego od kierunku dźwięku pierwotnego
  - jeśli odstęp czasu między impulsami jest mniejszy od 1 ms, to efekt precedensu nie występuje, a źródło lokalizowane jest pomiędzy kierunkiem dźwięku pierwotnego i wtórnego.
  - efekt ten zanika, gdy poziom dźwięku wtórnego przewyższa poziom pierwotnego o 10-15 dB
- Próg słyszenia - ucho ludzkie nie reaguje na dźwięki o natężeniu mniejszym od pewnej wartości zwanej progiem słyszenia dwuuszynego, przy czym wartość progowa zależy od częstotliwości dźwięku – największa czułość słyszenia występuje w zakresie od 2 kHz do 4 kHz (dźwięki usytuowane na krzywej progowej są ledwie słyszalne)



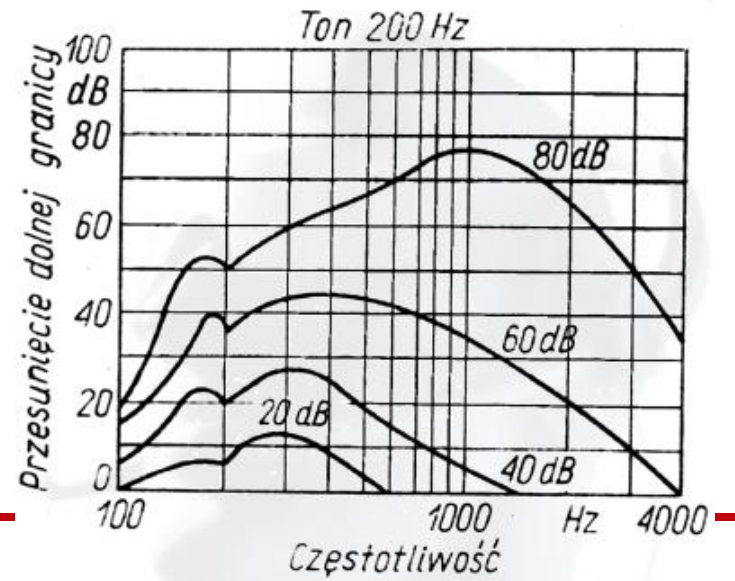
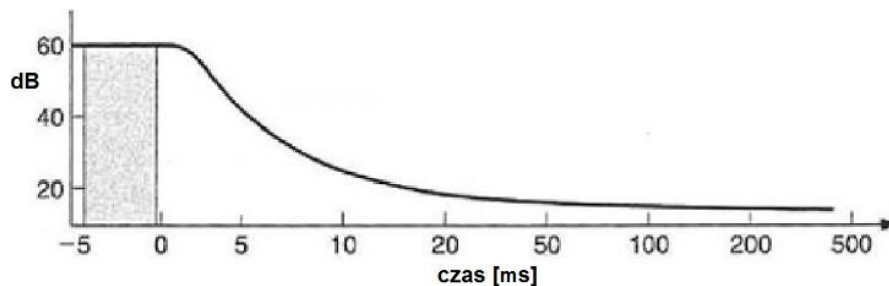
Słuchacz nie odczuwa pogorszenia jakości percepcji sygnału dźwiękowego, jeśli z widma tego sygnału zostaną usunięte składowe, których poziom jest niższy od progu słyszenia

# Efekty maskowania (zagłuszenia)

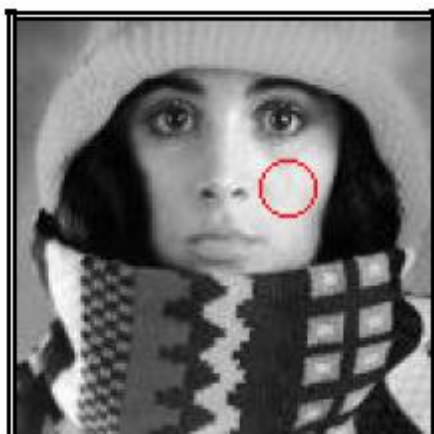
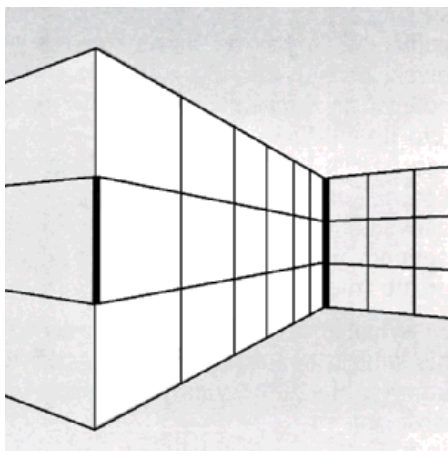
■ Chodzi o zjawisko podniesienia progu słyszalności dźwięku maskowanego wskutek obecności dźwięku maskującego (maskera)

- dźwięk maskowany może być wtedy słyszany słabo albo wcale
- zjawisko to jest ściśle związane z adaptacją układu słuchowego
- rozróżnia się maskowanie równoczesne i nierównoczesne, związane odpowiednio z właściwościami częstotliwościowymi i czasowymi dźwięków
- maskowanie równoczesne polega na tym, że dźwięk maskowany występuje w najbliższym czasowo sąsiedztwie (przy jednoczesnej percepcji) tonu maskującego; skuteczność maskowania zależy od natężenia dźwięków: maskującego i maskowanego oraz wzajemnej relacji ich widm częstotliwościowych;
- najsilniej maskowane są dźwięki o zbliżonych częstotliwościach i częstotliwościach wyższych

*zmiana krzywej progowej wskutek maskowania równoczesnego tonem 200 Hz, przy różnych poziomach natężenia dźwięku*

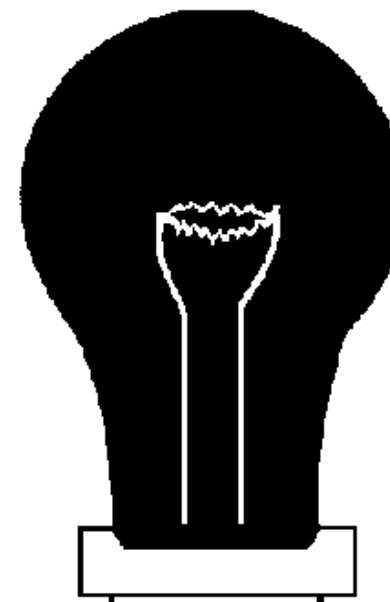
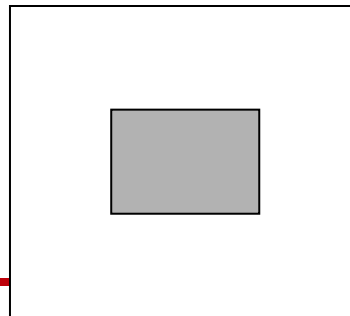
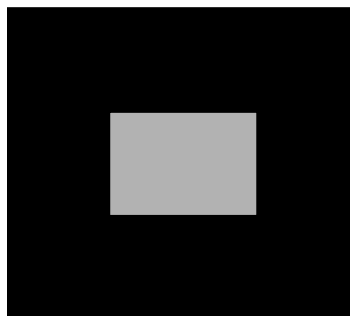
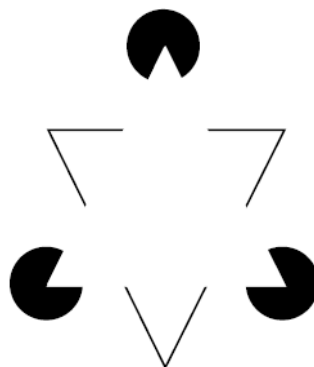
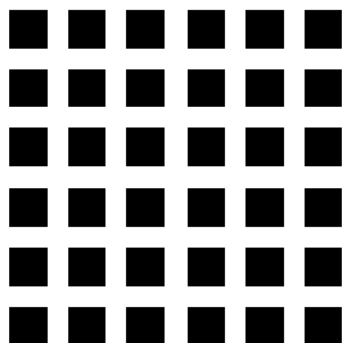


# HVS, jakość, wiarygodność



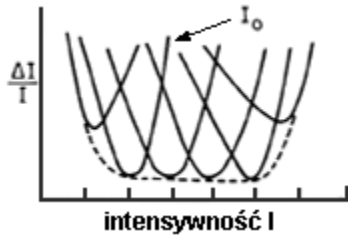
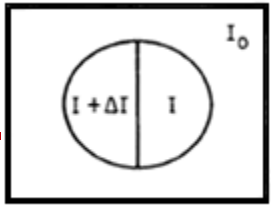
<i>Statistic</i>	<i>Image</i>	<i>ROI</i>
Average	137.7	219.3
Standard Deviation	49.5	4.0
Minimum	56	202
Median	141	220
Maximum	241	226
Mode	62	220
SNR (db)	NA	33.3

iluzje





# Czułość kontrastu



0%

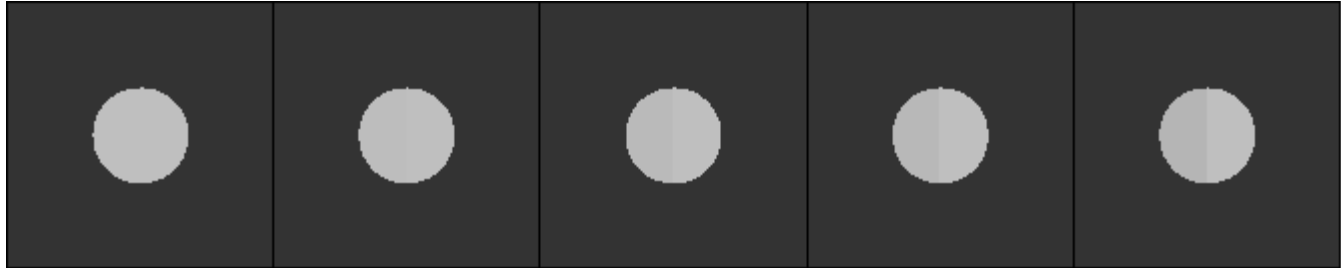
1%

2%

3%

4%

*tło różne*



*tło jak lewa  
połówka*



JND 4% (górze) i 2% (dół)



*lewa część jednolitego paska  
jaśniejsza od drugiej*

# Testy subiektywne

- Absolutne, porównawcze
- Za pomocą skali:

$$S = \frac{\sum_{k=1}^K (s_k n_k)}{\sum_{k=1}^K n_k}$$

gdzie  $K$  – liczba kategorii w przyjętej skali ocen,  $s_k$  - wartość oceny związanej z  $k$ -tą kategorią,  $n_k$  - liczba ocen z danej kategorii.


Kategoria $k$	Wartość skali ocen $s_k$	Opis słowny charakteryzujący jakość obrazów
1	5.	Wyśmienita
2	4.	Dobra
3	3.	Średnia
4	2.	Słaba
5	1.	Zła

Kategoria $k$	Wartość skali ocen $s_k$	Opis słowny charakteryzujący jakość obrazów
1	3.	Zdecydowanie (bezwzględnie) lepsza
2	2.	Wyraźnie lepsza
3	1.	Nieznacznie lepsza
4	0.	Porównywalna z oryginałem
5	-1.	Nieznacznie gorsza
6	-2.	Wyraźnie gorsza
7	-3.	Zdecydowanie (bezwzględnie) gorsza

## Give your mark! (SCASJ method)

Please, choose your opinion about the quality of the LEFT picture compared to the quality of the RIGHT picture (for example, choosing -2 or -3 means that the LEFT picture is slightly worse than the RIGHT one).

Much worse    Worse    Slightly worse    The same    Slightly better    Better    Much better



-3    -2    -1    0    1    2    3

Circles symbolize your opinion on left and right video correspondingly. Red circle means that video is bad, and green means that video is good.

Your choice: 2

# Test SCASJ



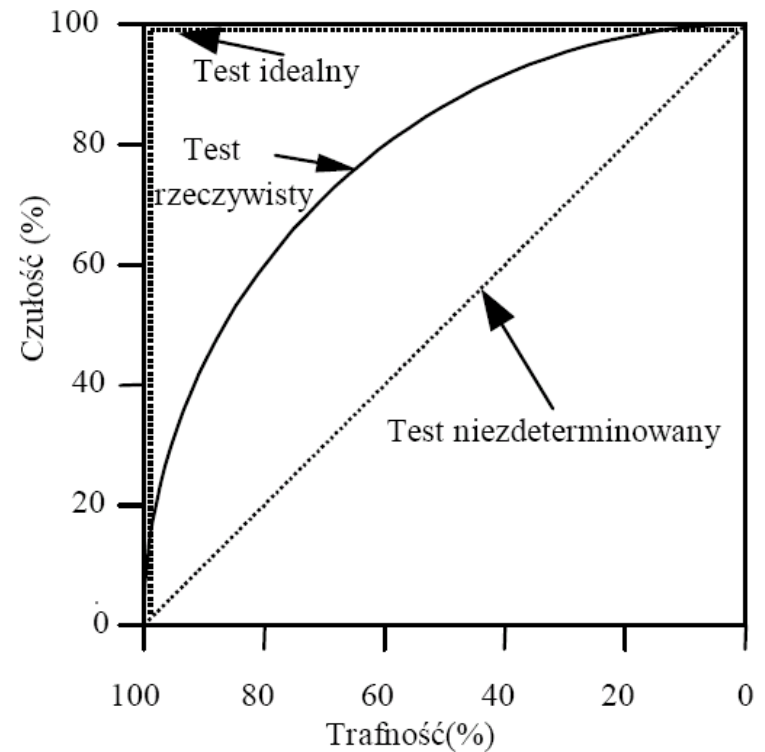
Test controls  
Status: Paused  
Expert: a  
Task name: ocena1  
Test number: 2 of 9



Hot keys  
Vote against left frame: left\_arrow  
Vote against right frame: right\_arrow  
Play/Pause: space, mouse L  
3 seconds back: backspace  
Restart sequence: Ctrl + R

# Testy detekcji (użytkowe)

- Reprezentatywny zbiór danych testowych
- Detekcja obiektów znaczących
- Czułość, specyficzność
- Krzywa ROC
- Weryfikacja hipotez statystycznych
- Uwarunkowania
  - dużą złożoność i czasochłonność
  - maksymalizacja obiektywizmu
  - przestrzeganie realiów użytkownika



# Przykłady: czułość

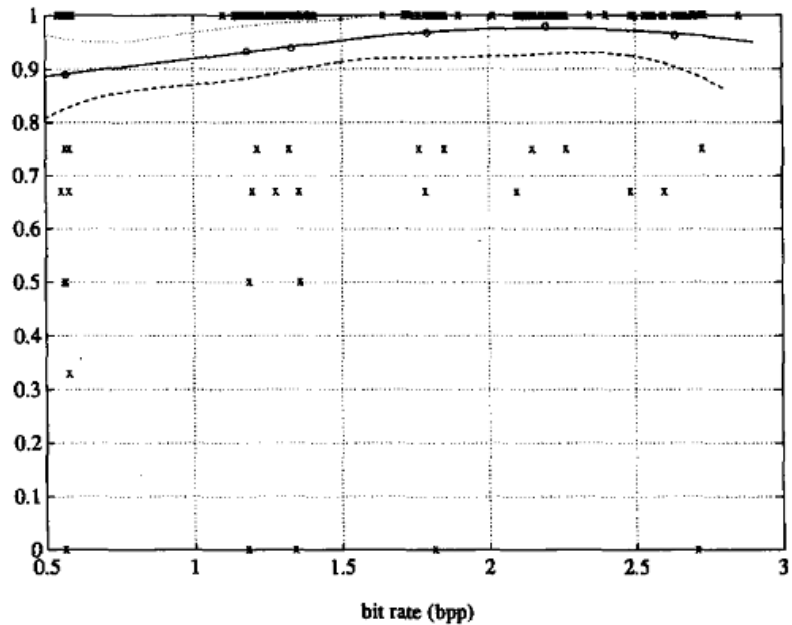


Fig. 8. Lung sensitivity: rms = 0.177.

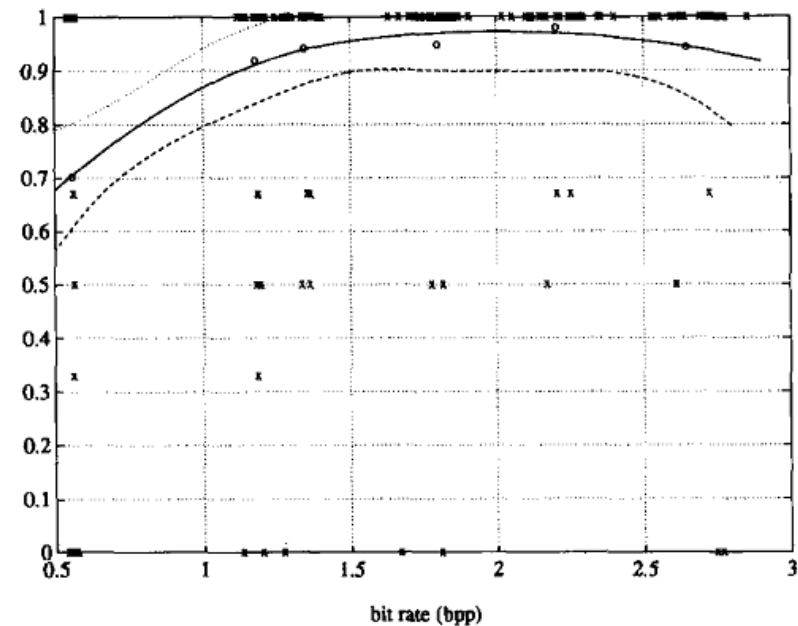


Fig. 10. Mediastinum sensitivity: rms = 0.243.

P. Cosman, R. Gray, R. Olshen Evaluating Quality of Compressed Medical Images: SNR, Subjective Rating, and Diagnostic Accuracy. PROCEEDINGS OF THE IEEE, VOL. 82, NO. 6, JUNE 1994

# Kwantyzacja

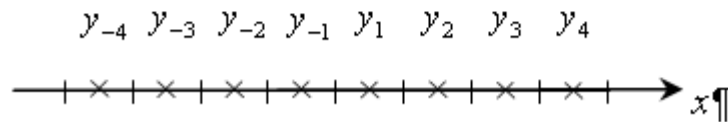
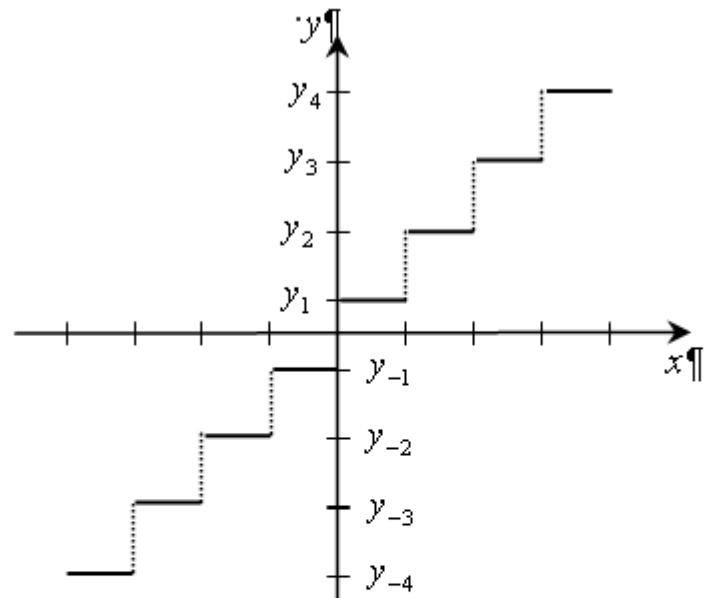
- Kwantyzacja to złożenie dwóch odwzorowań:
  - Kodera (k.prosta): nieskończony (skończony, duży alfabet) zbiór wartości (rzeczywistych) z określonego zakresu  $\rightarrow$  skończony (mały) zbiór wartości skwantowanych - całkowitych (lub wartości o zdecydowanie mniejszej precyzji) lub indeksów przedziałów kwantyzacji, dzielących zakres wartości wejściowych; indeksy te są zastępowane niekiedy słowami kodowymi
  - Dekodera (k.odwrotna): zbiór indeksów (wartości skwantowanych)  $\rightarrow$  zbiór (mały) rzeczywistych (całkowitych) wartości rekonstruowanych (reprezentantów przedziałów kwantyzacji)

$$\text{We: } X = \{x_i\}_{i=1}^K \quad \text{Wy: } \tilde{X} = \{\tilde{x}_i\}_{i=1}^K$$

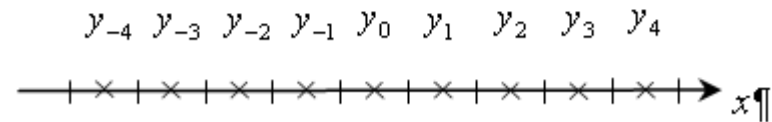
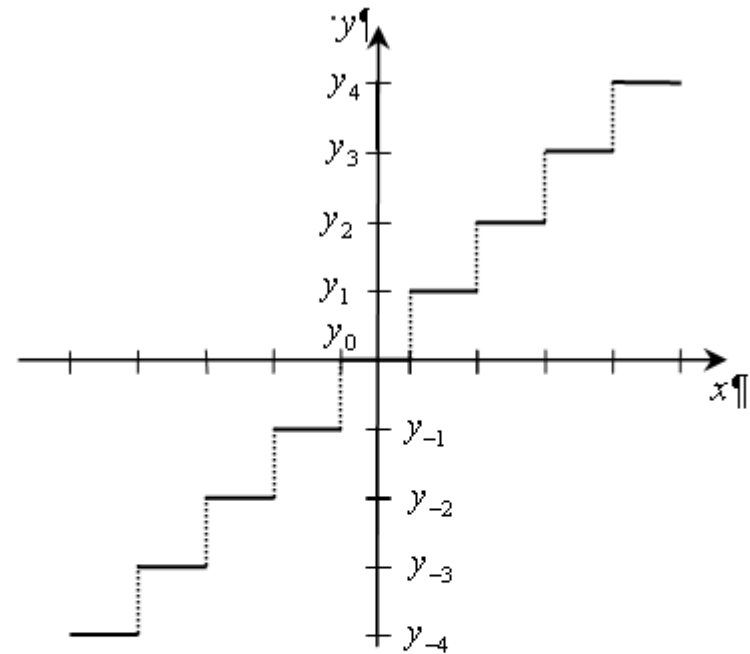
$$\text{Kwantyzacja: } Q: \mathbb{R} \rightarrow \mathbb{R}, \text{ tak że } Q(x_i) = \tilde{x}_i \text{ z błędem: } D_Q = \frac{1}{K} \sum_{i=1}^K (x_i - \tilde{x}_i)^2$$

przy czym  $\tilde{x}_i = y_j$  jeśli  $x_i \in [\beta_{j-1}, \beta_j)$ , a  $\{\beta_j\}_{j=0}^M$  jest zbiorem punktów granicznych przedziałów kwantyzacji z punktami rekonstrukcji  $\{y_j\}_{j=1}^M$

# Kwantyzacja równomierna



bez zera



z zerem

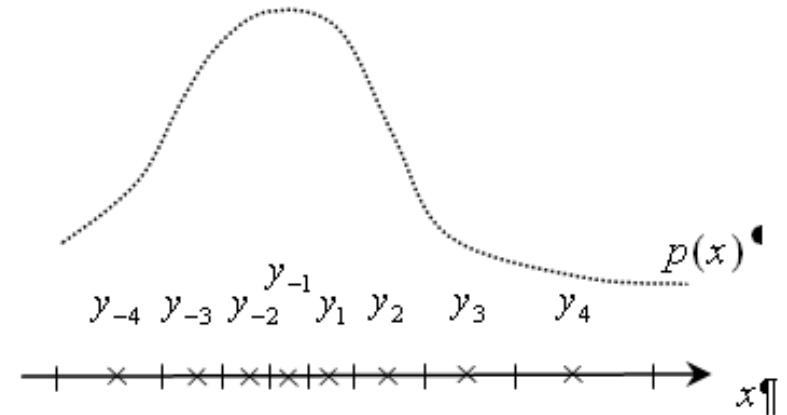
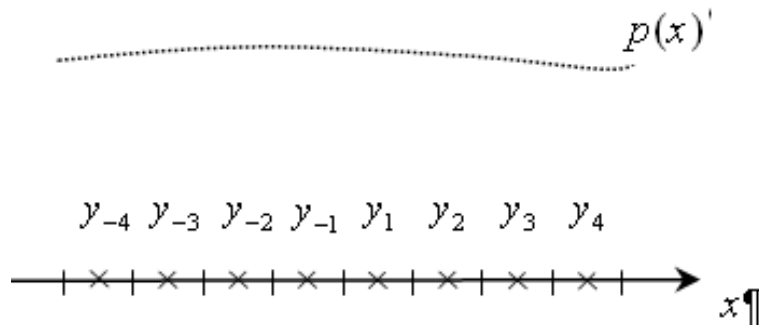
---

# **METODY KWANTYZACJI SKALARNEJ**

---



# Kwantyzacja nierównomierna - sterowana estymatą funkcji gęstości prawdopodobieństwa



- Metoda Lloyd-Maxa (iteracyjna, zbieżna do lokalnego optimum, duży wpływ inicjalizacji, czasochłonna)
- Obowiązujące zasady centroidu i najbliższego sąsiada

$$y_j = \frac{\int_{\beta_{j-1}}^{\beta_j} x f_X(x) dx}{\int_{\beta_{j-1}}^{\beta_j} f_X(x) dx}$$

$$\beta_j = \frac{y_j + y_{j+1}}{2}$$

# Kwantyzatory równomierne rozkładów nierównomiernych

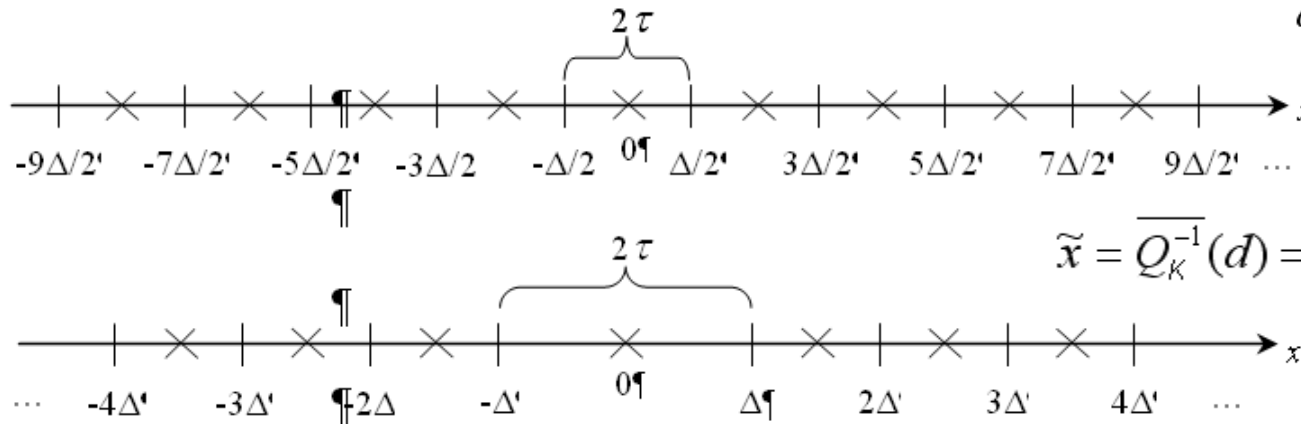
Liczba poziomów kwantyzacji $M$	Źródło o rozkładzie Gaussa ( $x=0, \sigma =1$ )	Źródło o rozkładzie Laplace'a ( $x=0, \sigma =1$ )
	Szerokość przedziału kwantyzacji $\Delta$	Szerokość przedziału kwantyzacji $\Delta$
2	1,5960	1,4140
4	0,9957	1,0873
6	0,7334	0,8707
8	0,5860	0,7309
10	0,4908	0,6334
12	0,4238	0,5613
14	0,3739	0,5055
16	0,3352	0,4609
32	0,1881	0,2799

# Kwantyzatory nierównomierne

$j$	Źródło o rozkładzie Gaussa ( $x=0, \sigma=1$ )				Źródło o rozkładzie Laplace'a ( $x=0, \sigma=1$ )				← Liczba przedziałów kwantyzacji
	4	8	16	32	4	8	16	32	
1	0,4528	0,2451	0,1284	0,0660	0,4198	0,2334	0,1240	0,0641	} Wartości punktów rekonstrukcji $y_j$
2	1,5104	0,7561	0,3882	0,1983	1,8340	0,8330	0,4049	0,2005	
3		1,3440	0,6569	0,3318		1,6725	0,7288	0,3463	
4		2,1420	0,9426	0,4672		3,0868	1,1111	0,5028	
5			1,2565	0,6056			1,5780	0,6718	
6			1,6183	0,7479			2,1776	0,8555	
7			2,0693	0,8954			3,0171	1,0565	
8			2,7328	1,0497			4,4314	1,2786	
9				1,2128				1,5266	
10				1,3874				1,8076	
11				1,5773				2,1314	
12				1,7883				2,5138	
13				2,0298				2,9807	
14				2,3188				3,5803	
15				2,6912				4,4199	
16				3,2616				5,8314	

# Uprozczone kwantyzatory

- Z regulowanym przedziałem zerowym



$$d = Q_k(x) = \text{znak}(x) \left\lfloor \frac{|x| + 0.5\Delta}{\Delta} \right\rfloor$$

$-0.5 \leq \delta < 0.5$

$$\tilde{x} = Q_k^{-1}(d) = \begin{cases} 0, & d = 0 \\ \text{znak}(d)(|d| + \delta)\Delta, & d \neq 0 \end{cases}$$

$0 \leq \delta < 1$

uogólniając

$$d = \begin{cases} 0, & |x| < -k\Delta \\ \text{znak}(x) \left\lfloor \frac{|x| + k\Delta}{\Delta} \right\rfloor, & |x| \geq -k\Delta \end{cases}$$

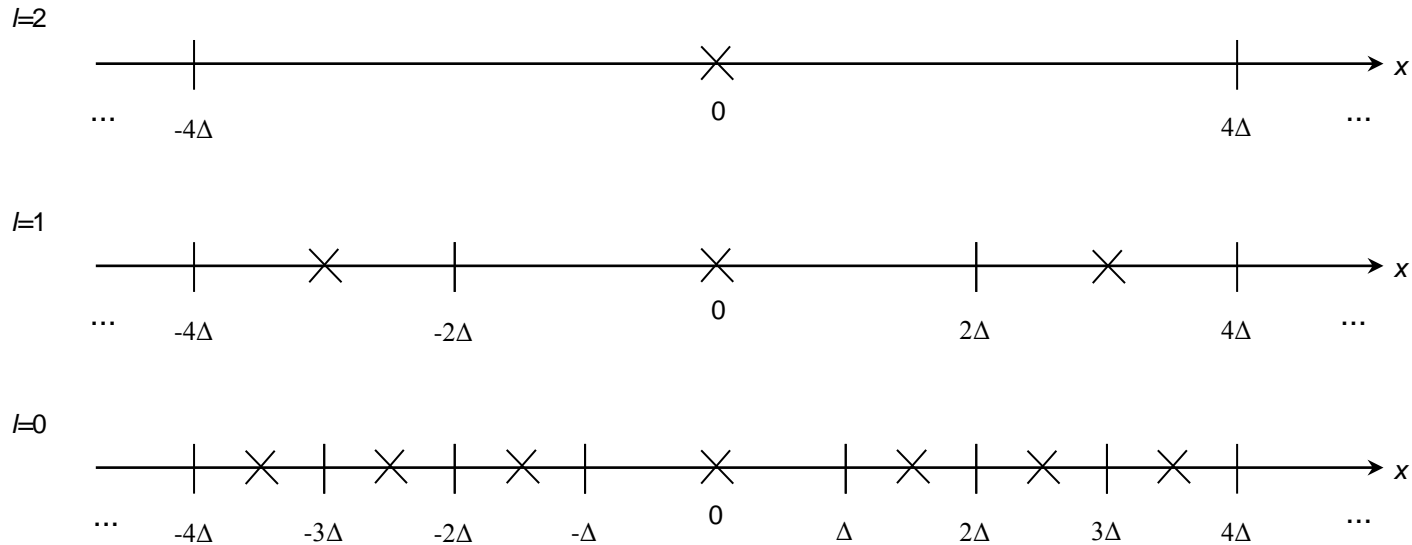
$$\tilde{x} = \begin{cases} 0, & d = 0 \\ \text{znak}(d)(|d| - k + \delta)\Delta, & d \neq 0 \end{cases}$$

$$d = \text{znak}(x) \left\lfloor \frac{|x|}{\Delta} \right\rfloor$$

szerokość przedziału zerowego  $2(1-k)\Delta$ , gdzie  $k \in (-1, 1)$

# Uproszczone kwantyzatory

- Sukcesywna aproksymacja (progowanie lub dostawianie bitów)



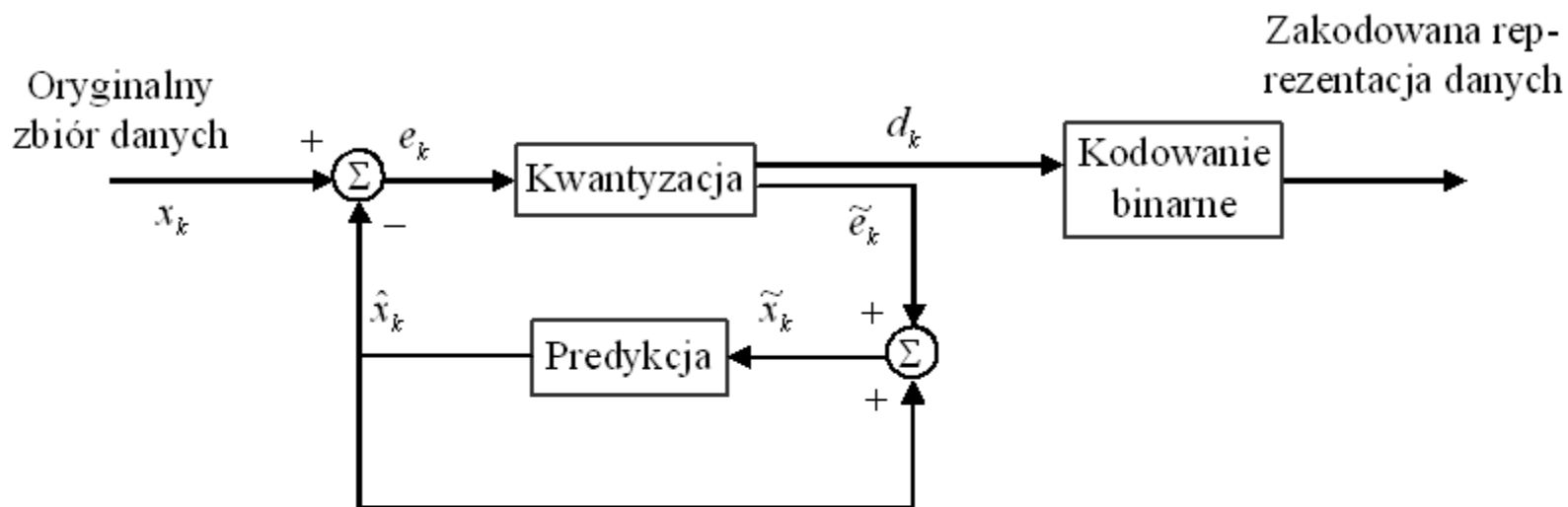
---

# **PROSTE METODY KOMPRESJI Z KWANTYZACJĄ**

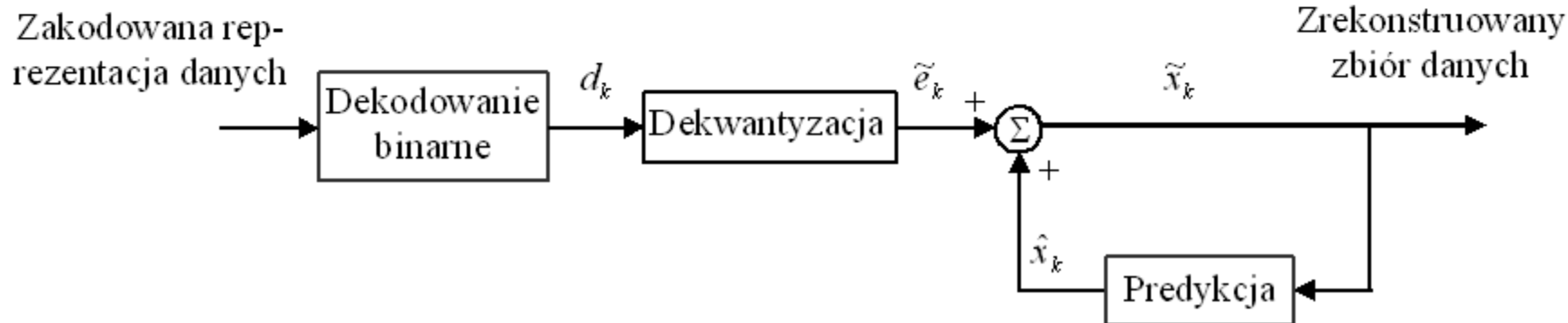
---

# Stratna DPCM

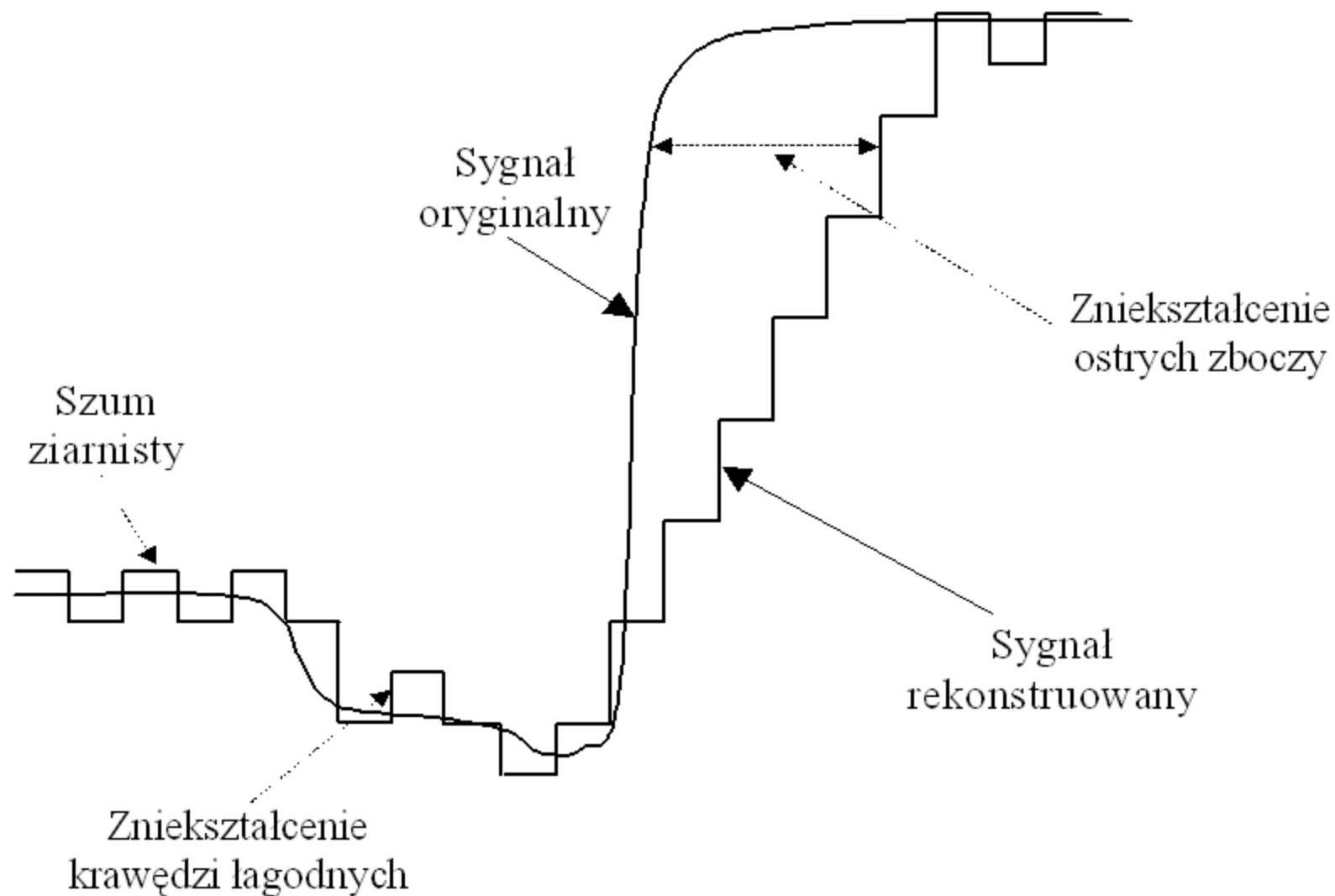
## KOMPRESJA



## DEKOMPRESJA



# Typowe zniekształcenia stratnej DPCM





# Blokowa kwantyzacja dwupoziomowa BTC

- Podział na bloki

$$X_1 = \begin{bmatrix} 45 & 48 & 46 & 47 \\ 44 & 45 & 44 & 46 \\ 45 & 47 & 43 & 45 \\ 47 & 46 & 44 & 47 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 145 & 148 & 146 & 146 \\ 44 & 145 & 144 & 146 \\ 45 & 47 & 143 & 45 \\ 47 & 46 & 45 & 47 \end{bmatrix}, \quad X_3 = \begin{bmatrix} 49 & 53 & 57 & 59 \\ 48 & 51 & 54 & 53 \\ 48 & 47 & 52 & 50 \\ 47 & 46 & 49 & 50 \end{bmatrix}$$

- Segmentacja progowa, charakterystyka obszarów (w. średnie)

$$t_1 = 46, \quad t_2 = 96, \quad t_3 = 51 \quad y_1^L = 44, \quad y_1^H = 47, \quad y_2^L = 46, \quad y_2^H = 145, \quad y_3^L = 48, \quad y_3^H = 54$$

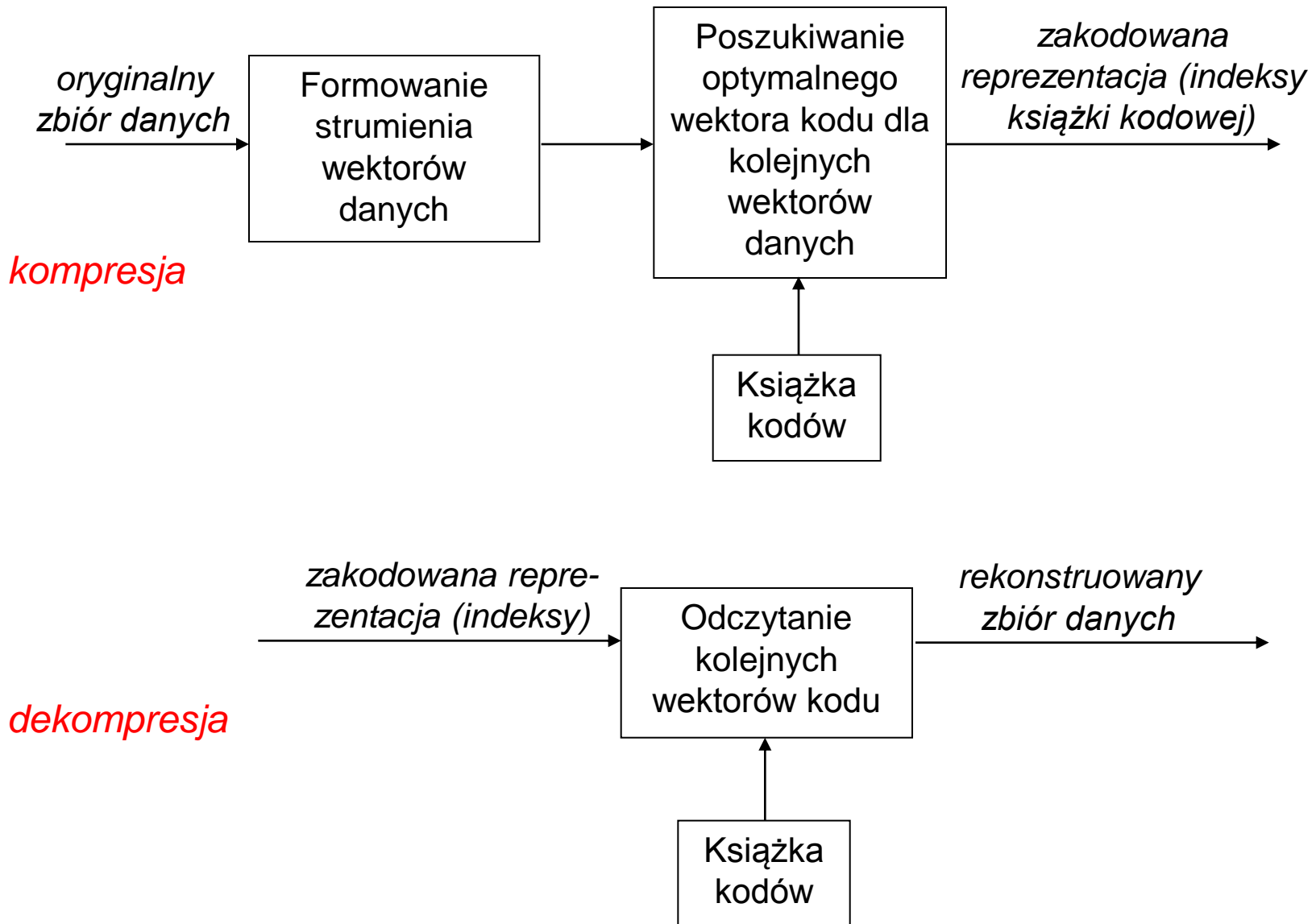
- Wyznaczenie map bitowych

$$M_1 = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad M_3 = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- Rekonstrukcja

$$\tilde{X}_1 = \begin{bmatrix} 44 & 47 & 47 & 47 \\ 44 & 44 & 44 & 47 \\ 44 & 47 & 44 & 44 \\ 47 & 47 & 44 & 47 \end{bmatrix}, \quad \tilde{X}_2 = \begin{bmatrix} 145 & 145 & 145 & 145 \\ 46 & 145 & 145 & 145 \\ 46 & 46 & 145 & 46 \\ 46 & 46 & 46 & 46 \end{bmatrix}, \quad \tilde{X}_3 = \begin{bmatrix} 48 & 54 & 54 & 54 \\ 48 & 54 & 54 & 54 \\ 48 & 48 & 54 & 48 \\ 48 & 48 & 48 & 48 \end{bmatrix}$$

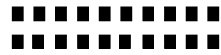
# Kwantyzacja wektorowa VQ (metoda kompresji)



# VQ bardziej obrazowo

koncepcja

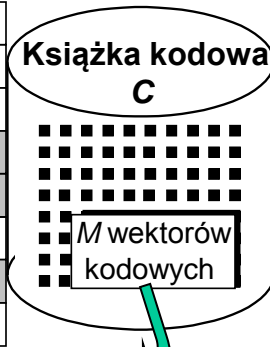
Wejście:  
wektory danych



....  
....

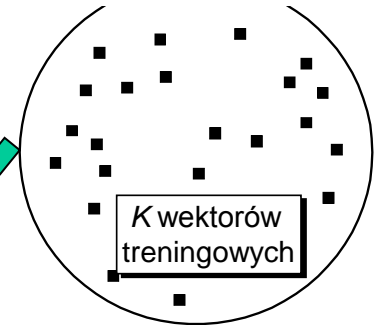


Indeksy



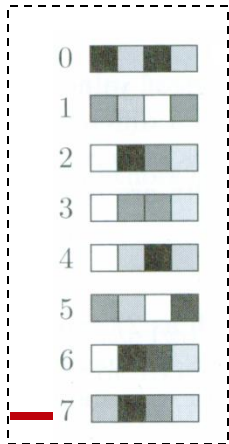
Wyjście: 3, 42, ....

Zbiór treningowy  $T$



przykład

Książka kodów



Wektory wejściowe



Indeksy

0 3 6 2 4 0

Wektory rekonstrukcji

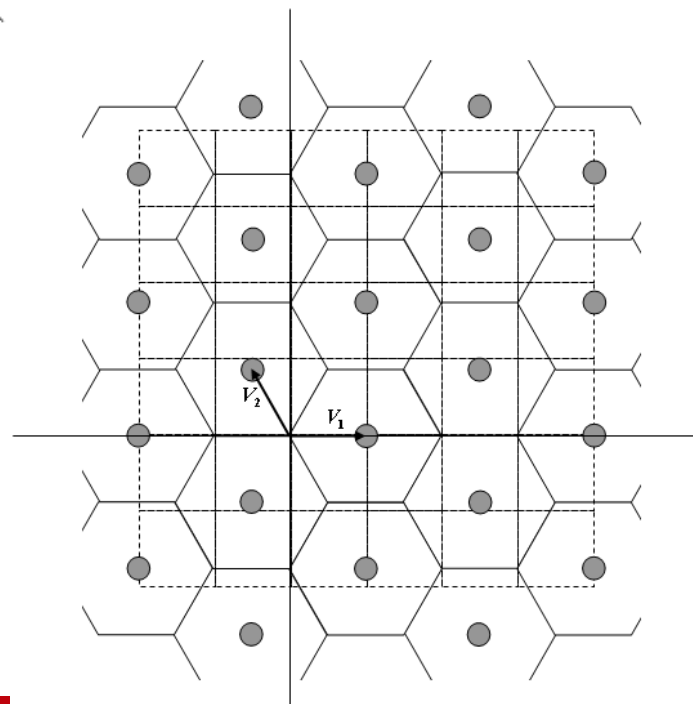
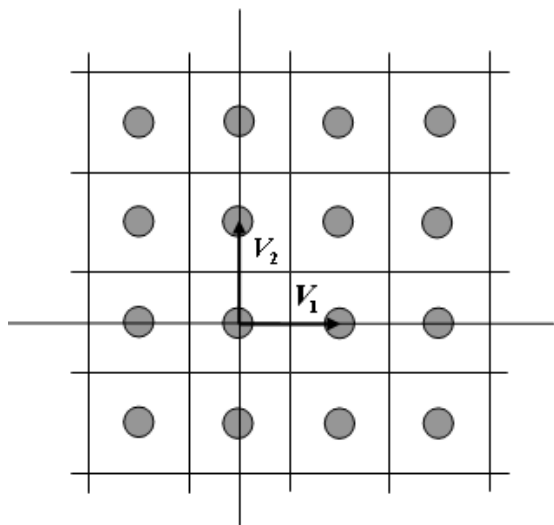
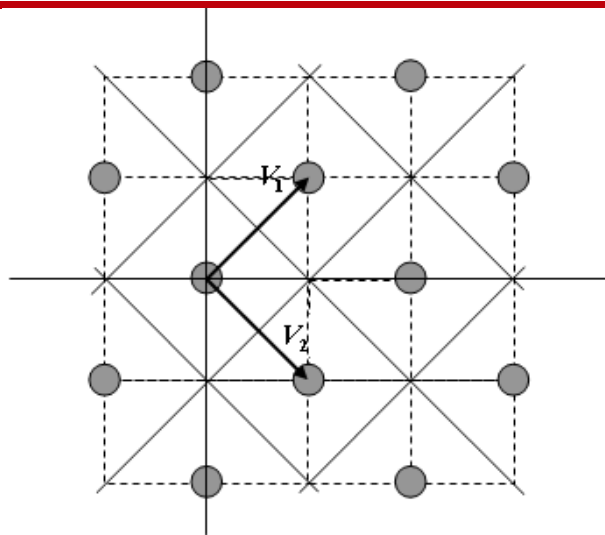


# VQ - problemy

---

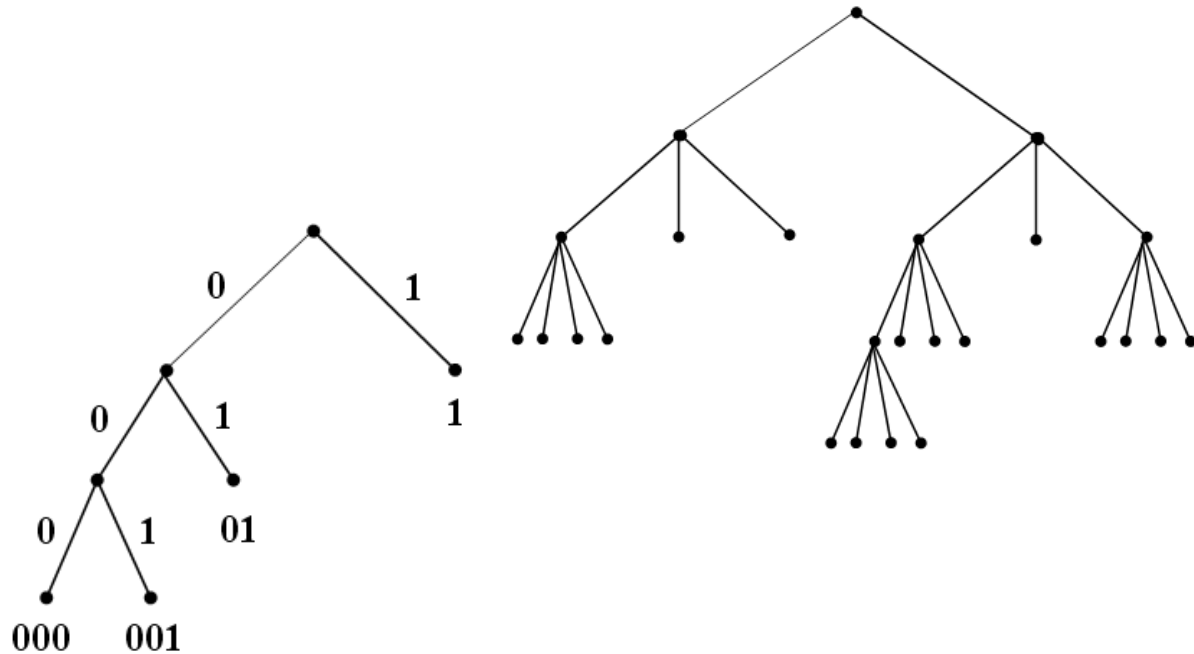
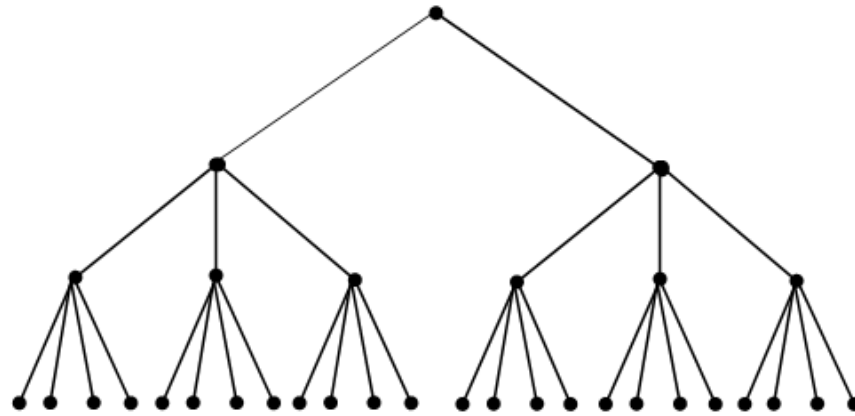
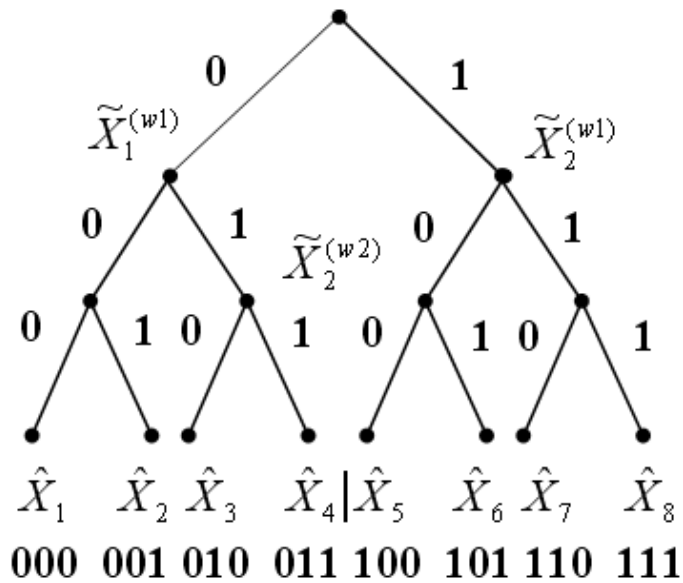
- Budowa książki kodowej
    - kraty-nośniki
    - metoda LBG
    - inicjalizacja LBG (losowa, grupowanie, rozdzielanie)
    - adaptacja (lokalne książki)
  - Struktura książki kodowej - drzewa
  - Wykorzystanie książki kodowej do kompresji
    - konstrukcja wektorów (dziedzina)
    - przeszukiwanie (struktury- drzewa)
    - formowanie i kodowanie indeksów
  - Ograniczenia metody
    - złożoność (uczenie książek)
    - brak optymalnych metod wyznaczania książki
    - efektywność jedynie w zakresie małych stopni kompresji (ograniczenie wymiaru wektorów praktycznie do 16)
-

# KK regularne (kratki) z 'podpróbkowaniem'

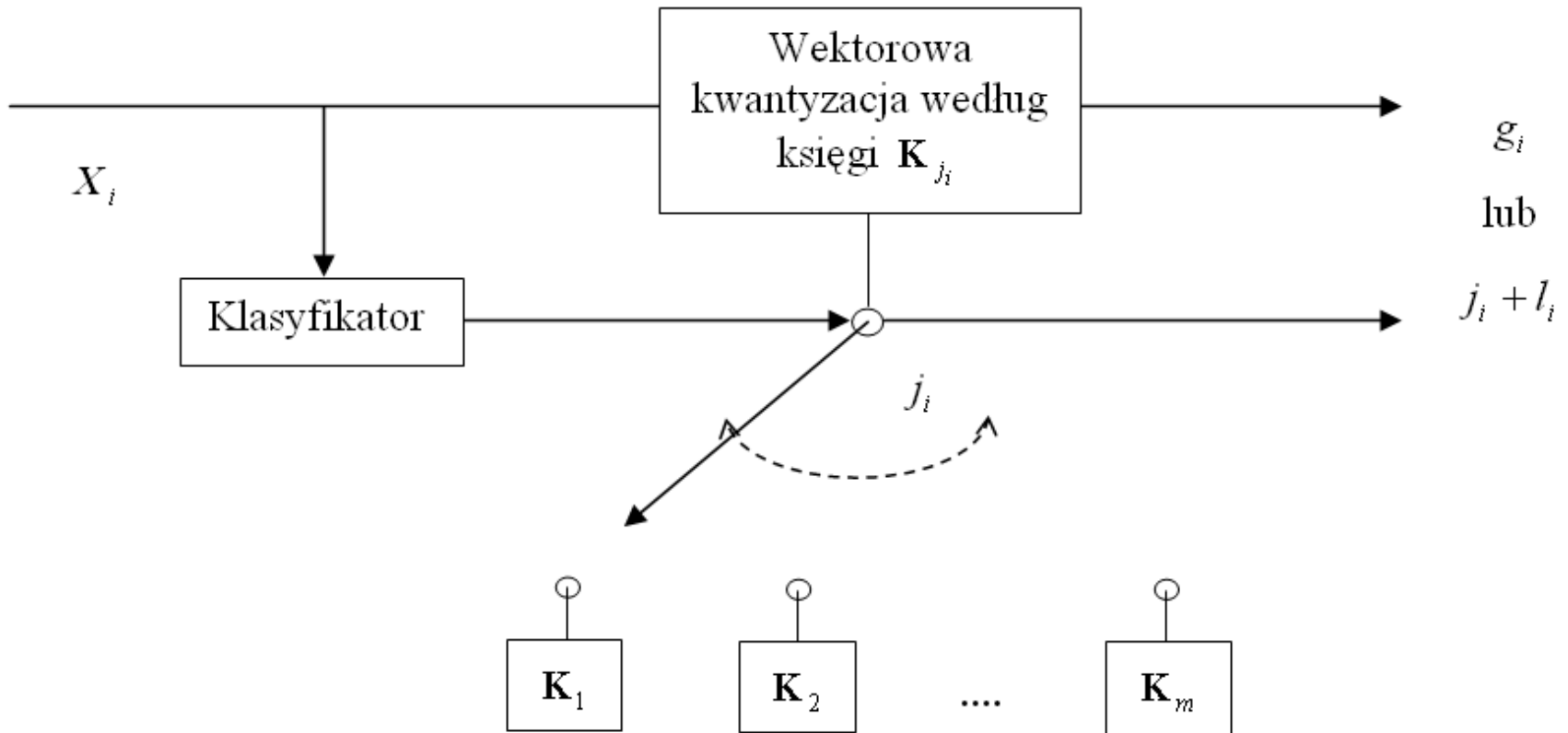


# KK - drzewo

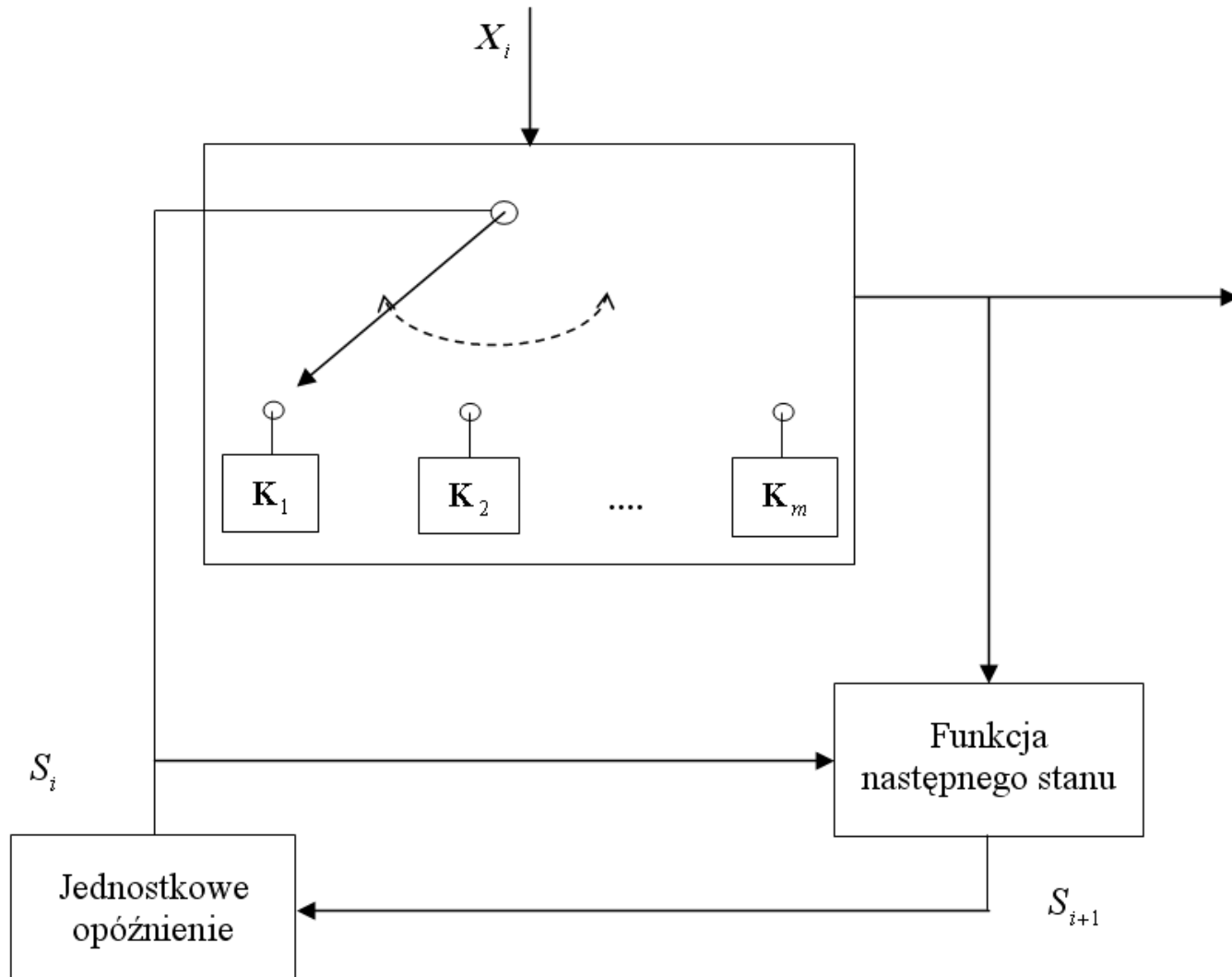
- Binarne
- Stożkowe
- Okrojone



# VQ z klasyfikacją

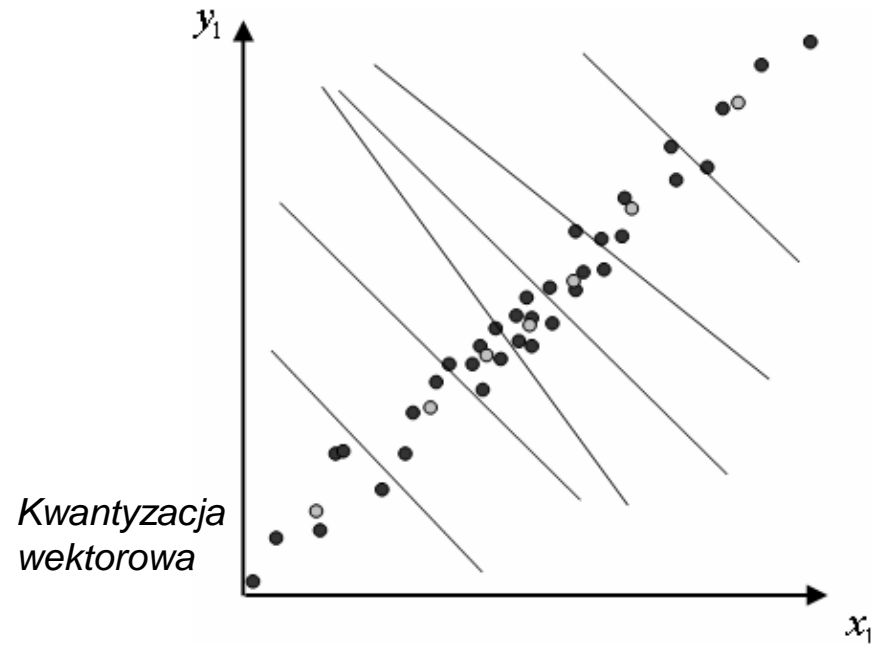
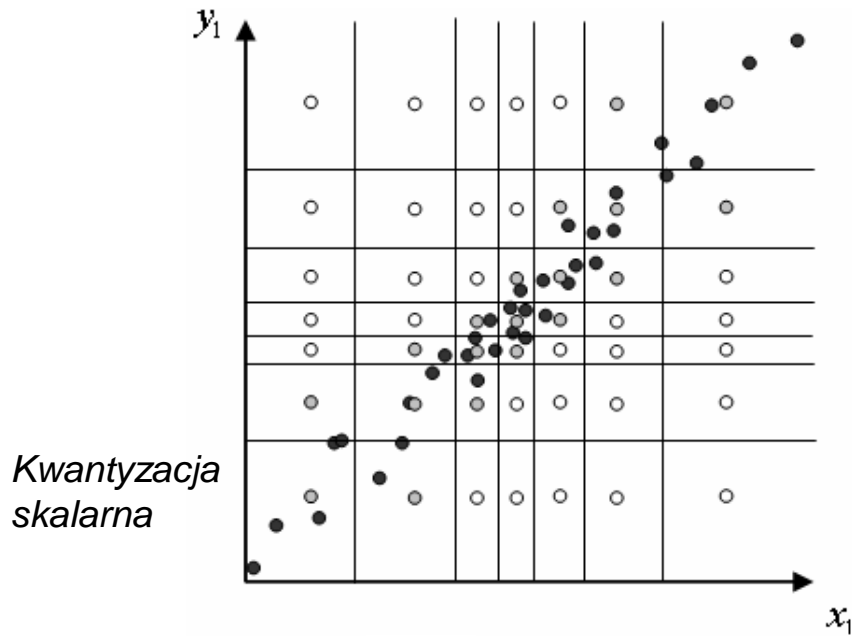


# VQ z pamięcią



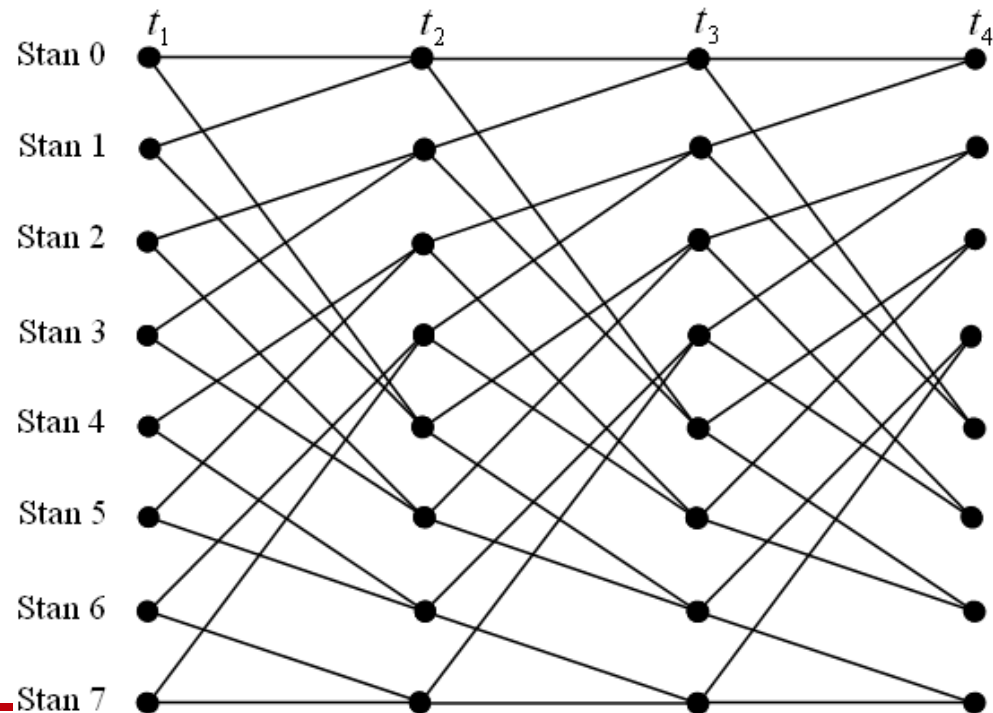


# VQ vs SQ



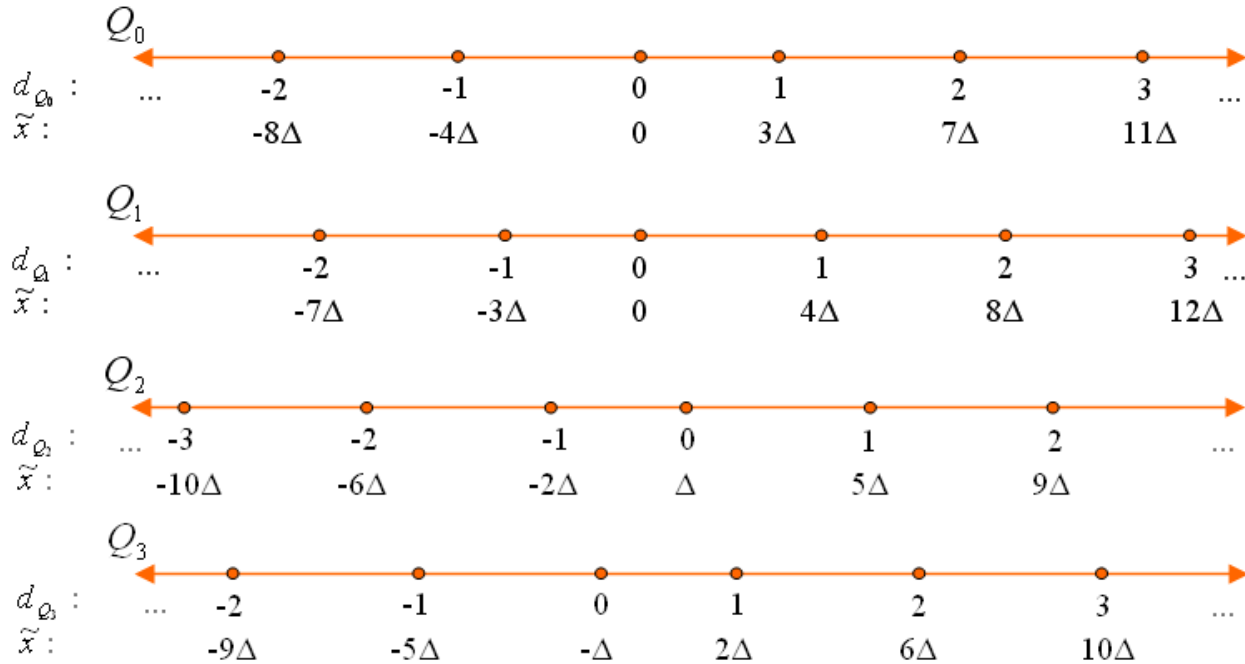
# Kwantyzacja kontekstowa skalarna: z kodowaniem kraty (TCQ)

- kilka kwantyzatorów skalarnych (o przesuniętych przedziałach) połączonych w unie, diagram przejść pomiędzy stanami (kwantyzacja kontekstowa)
- $R$  bitowe indeksy pozwalają użyć  $2^{R+1}$  przedziałów kwantyzacji
- przykład: krata o ośmiu stanach, 2 uniach i 4 kwantyzatorach ( $R-1$  bitowych, wtedy indeks= $R-1$  bitowy indeks kwantyzatora + 1 bit przejścia w kracie



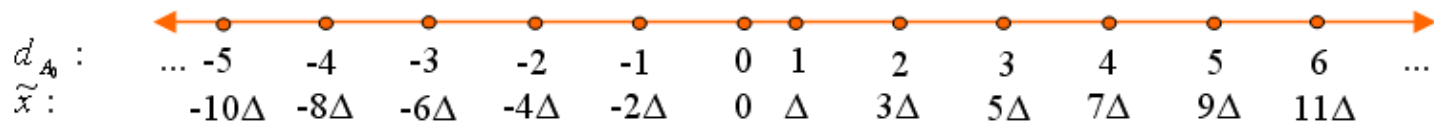
# TCQ

## Kwantyzatory

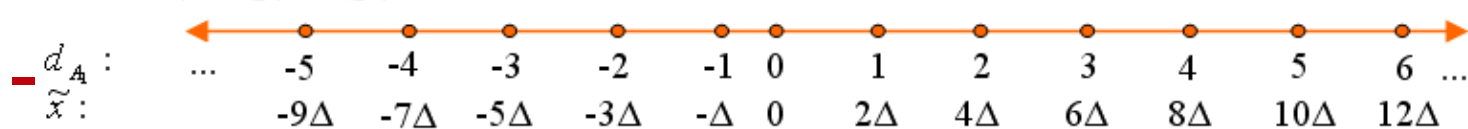


## Unie

$$A_0 = Q_0 \cup Q_2$$

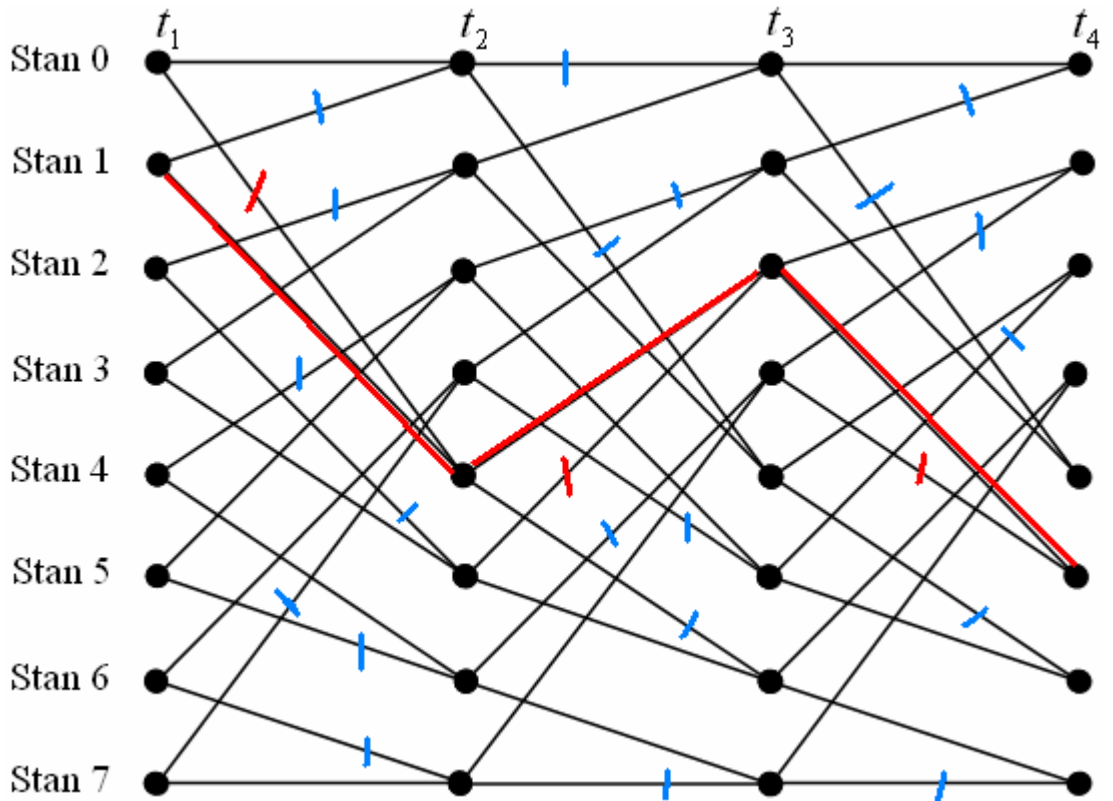
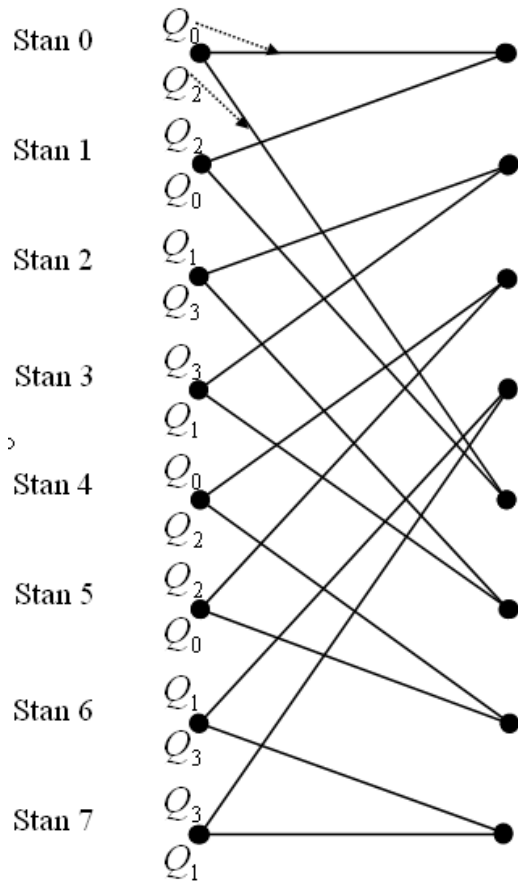


$$A_1 = Q_1 \cup Q_3$$



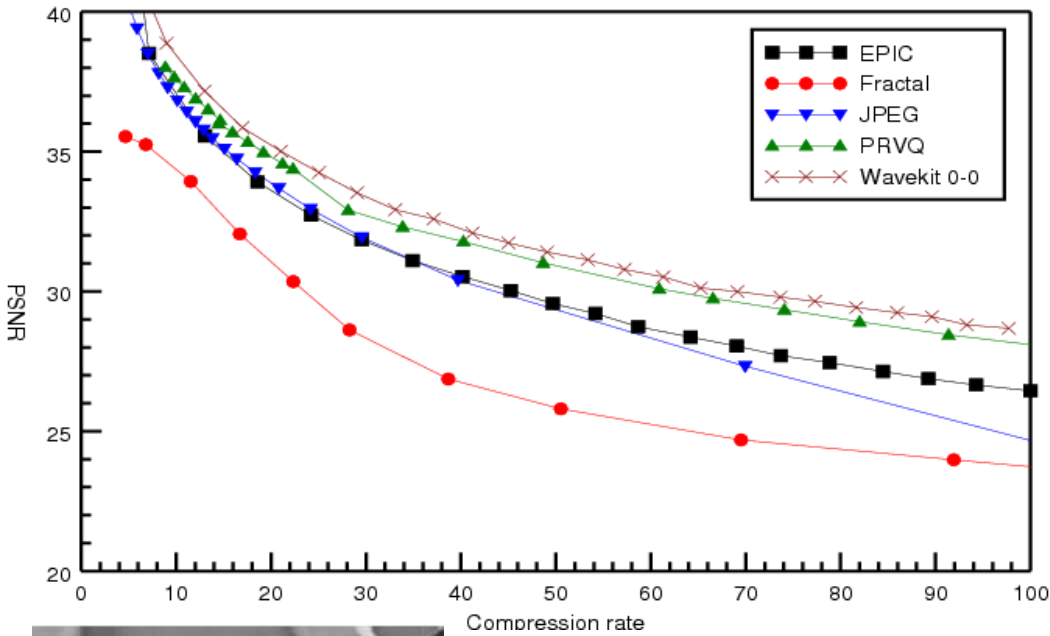
# TCQ

- Analiza wsteczna – poszukiwanie drogi o najmniejszym błędzie kwantyzacji



# Testy

Lena512



Peppers

