

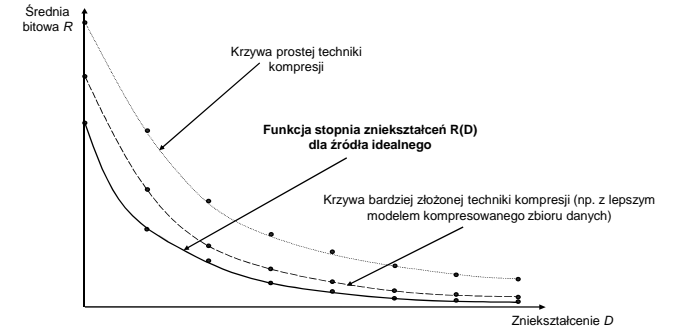
# DOSKONALENIE METOD KOMPRESJI STRATNEJ

Materiały KODA, A, Przelaskowski

- Podstawy teorii optymalizacji metod
- Transformacyjne metody falkowe
- Nowa koncepcja efektywnego kodera (np. JPEG2000)
- Sposób realizacji: przekształcenie + porządkowanie informacji
- Wieloskalowe metody kompresji

# Podstawy teorii stopnia zniekształceń źródeł informacji

- Krzywa R-D



- Przybliżenie X przez Y

$$D = \sum_{i=0}^{N_1-1} \sum_{j=0}^{N_2-1} d(x_i, y_j) P(x_i, y_j) = \sum_{i=0}^{N_1-1} \sum_{j=0}^{N_2-1} d(x_i, y_j) P(x_i) P(y_j | x_i)$$

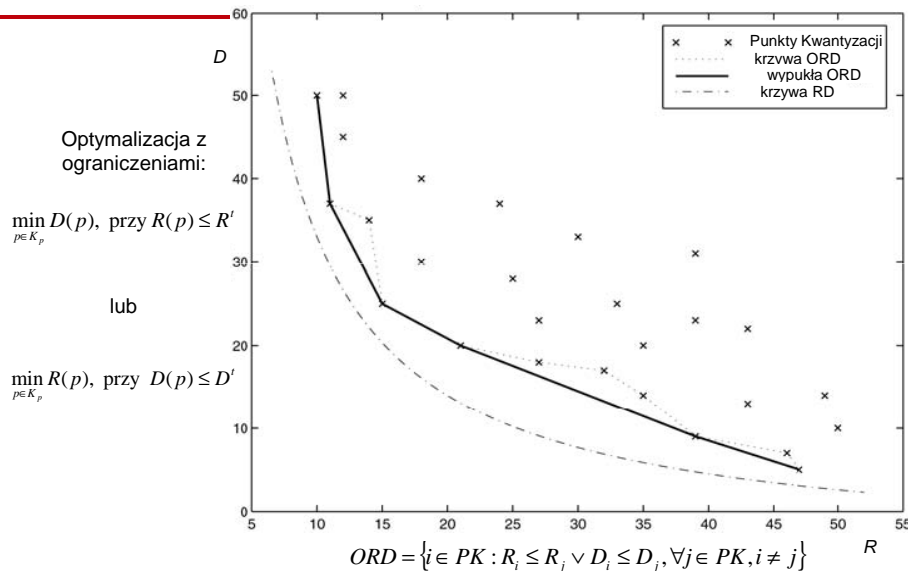
- Optymalizacja

$$R(D) = \min_{\{P(y_j | x_i)\} \in \Lambda} I(X; Y) \quad I(X; Y) = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} P(x_i, y_j) \log \left[ \frac{P(x_i, y_j)}{P(x_i)P(y_j)} \right]$$

$\Lambda = \{P(y_j | x_i)\}$  takich, że  $D(\{P(y_j | x_i)\}) \leq D'$

$$I(X; Y) = H(X) + H(Y) - H(X, Y) = H(X) - H(X | Y) = H(Y) - H(Y | X)$$

## Podstawy R-D: Operacyjna krzywa RD

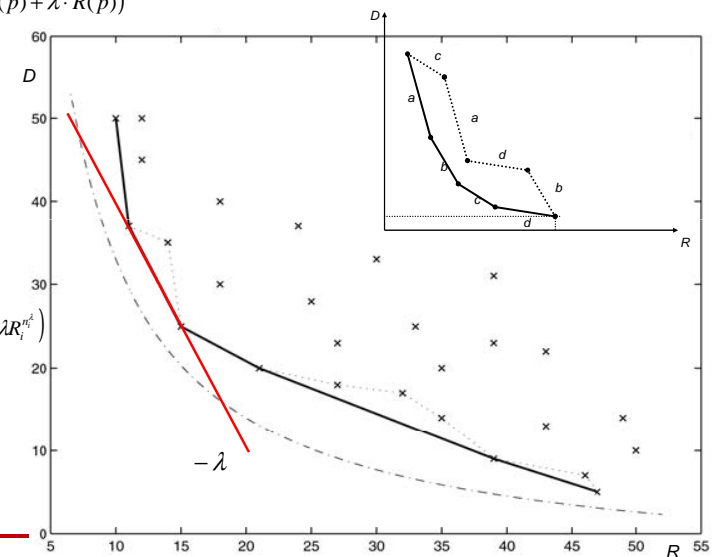


## Podstawy R-D: optymalizacja bez ograniczeń Lagrange'a

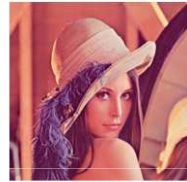
$$\tilde{p}(\lambda) = \arg \min_{p \in K_p} (D(p) + \lambda \cdot R(p))$$

$$\lambda = - \frac{\Delta D}{\Delta R}$$

$$D(\lambda) + \lambda R(\lambda) = \sum_i (D_i^{n_i} + \lambda R_i^{n_i})$$



## Paradygmat JPEG



Kompresja/Dekompresja



### Opcje kodera:

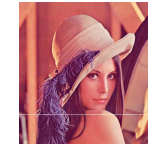
- → przestrzeń kolorów
- → kwantyzacja
- → entropijne kodowanie
- → przetwarzanie wstępne

Brak opcji dekodera (ustalony w koderze porządek i sposób rekonstrukcji, bez żadnych możliwości wyboru)

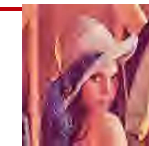
## Paradygmat JPEG2000

### Opcje kodera:

- format danych: obrazy ze skalą szarości i obrazy binarne
- dzielenie na części (podobrazy, ang. *tiling*)
- kompresja stratna-bezstratna
- pozostałe jak w paradygmacie JPEG



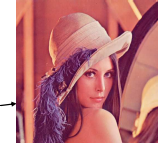
Kompresja/Dekompresja



skalowanie jakości



odtworzenie wybranego komponentu



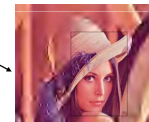
kompresja bezstratna



skalowanie rozdzielczości



rekonstrukcja wybranego ROI



progresa z wybranym ROI

### Opcje dekodera:

- rozdzielczość obrazu
- jakość obrazu
- ustalanie dokładnego rozmiaru reprezentacji
- bitowej dekodowanej informacji
- składniki (komponenty)
- regiony zainteresowań
- kompresja stratna-bezstratna

### Dodatkowe możliwości schematu kompresji:

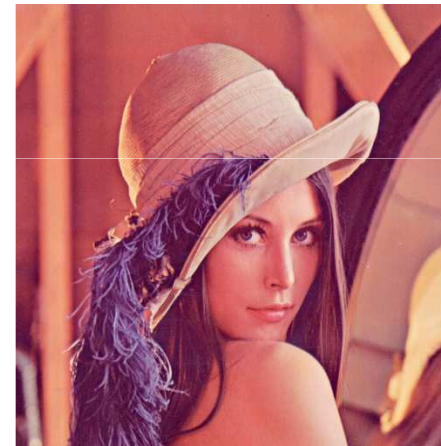
- zwiększona odporność na zakłócenia
- adnotacje o prawach autorskich
- XML-owa struktura informacji dodatkowych
- osadzony strumień bitowy (progressywne dekodowanie i skalowalność SNR)
- reprezentacja wielorozdzielcza
- bezpośredni dostęp do kodowanego strumienia i przetwarzanie

## Zobaczmy to



## Skalowanie jakości

Oryginał



Przybliżenie



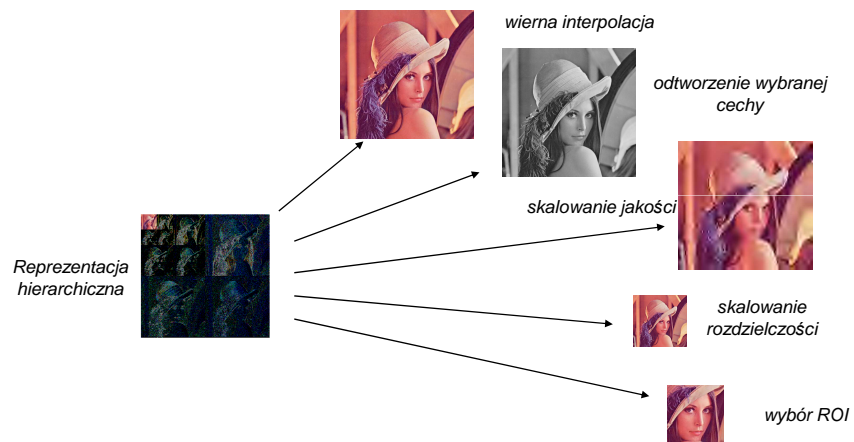
## Przykład reprezentacji obrazu



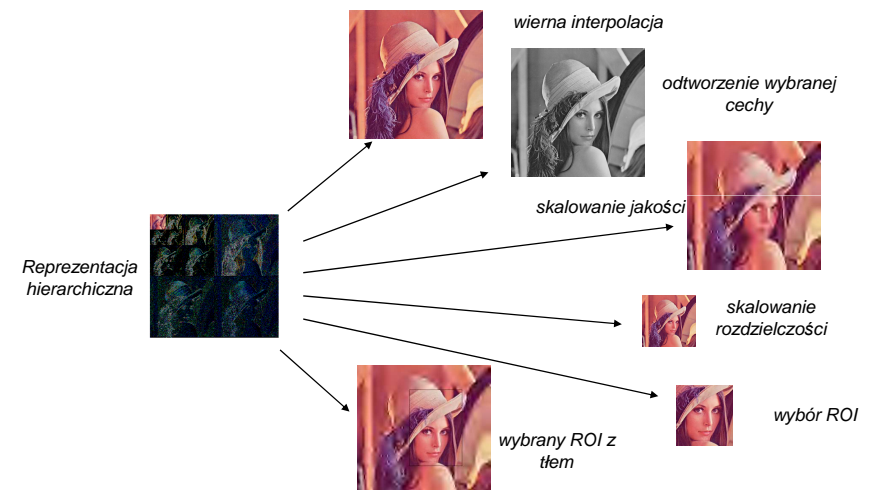
## Skalowanie rozdzielczości



## Przykład reprezentacji obrazu



## Przykład elastycznej reprezentacji obrazu



## Wybrany ROI z tłem

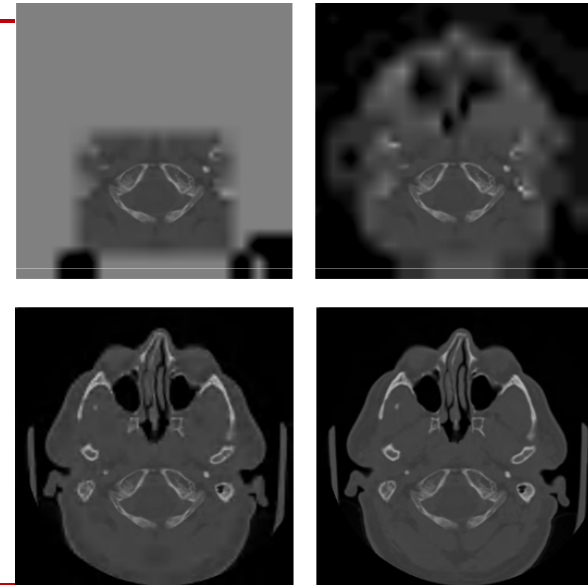
Oryginał



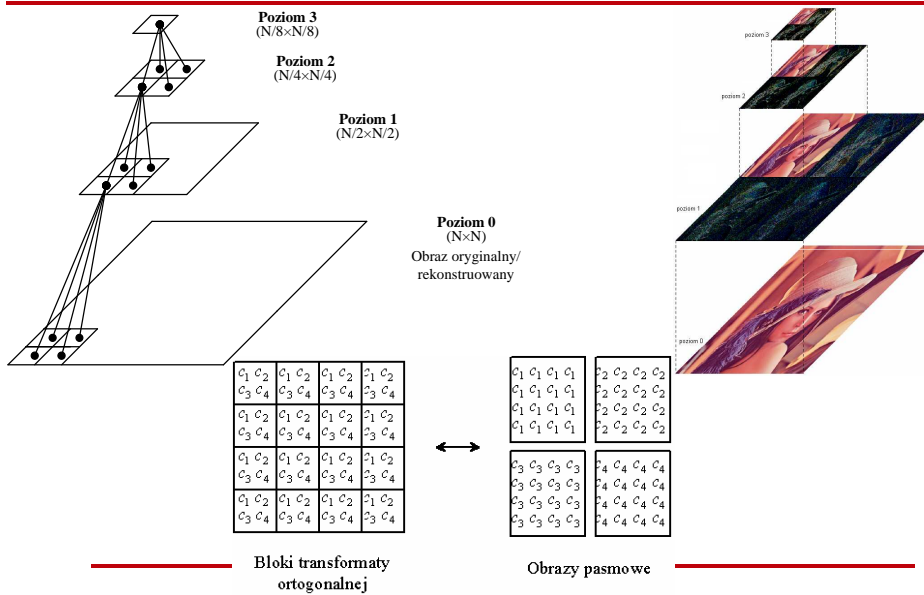
Przybliżenie



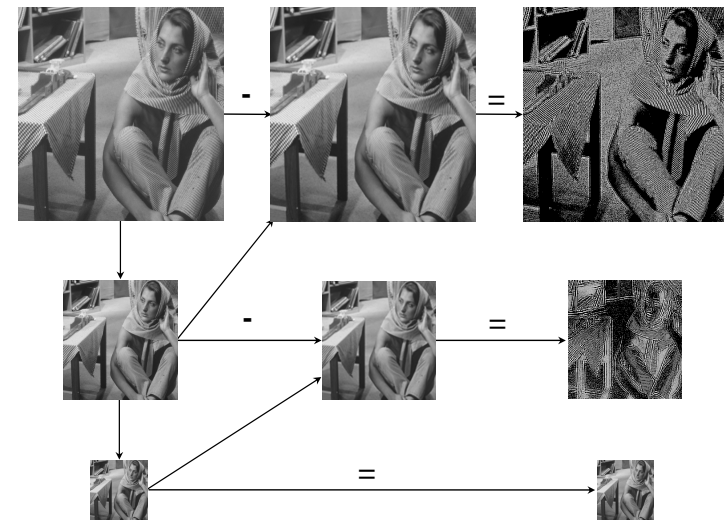
## Interaktywna transmisja (JPIP)



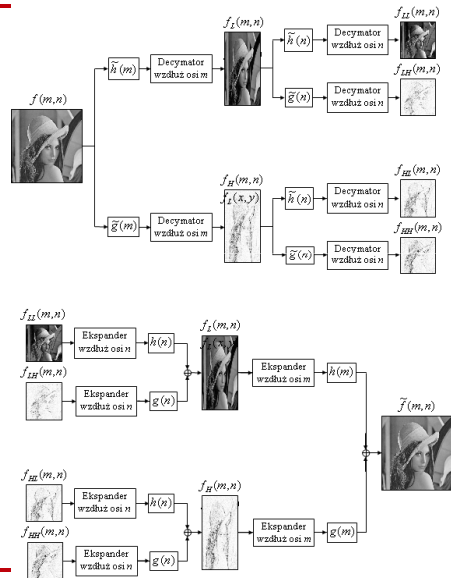
## SPOSÓB: hierarchia skal, podpasma i lokalne upakowanie



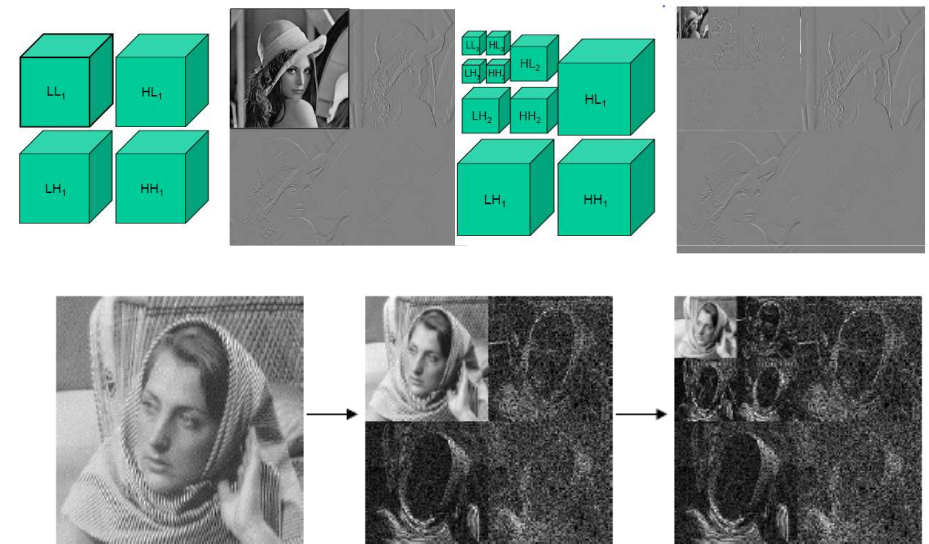
## Hierarchiczna piramida



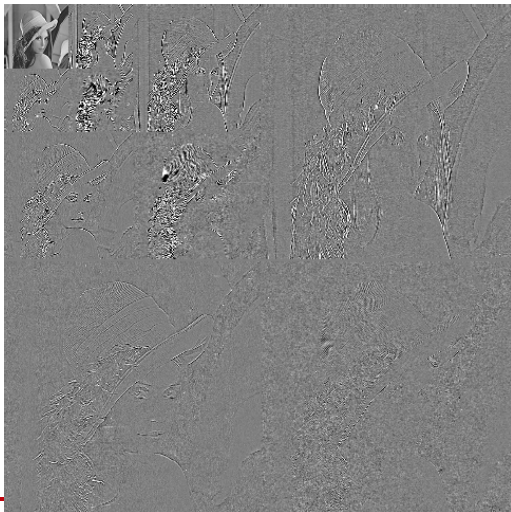
## Realizacja dekompozycji falkowej



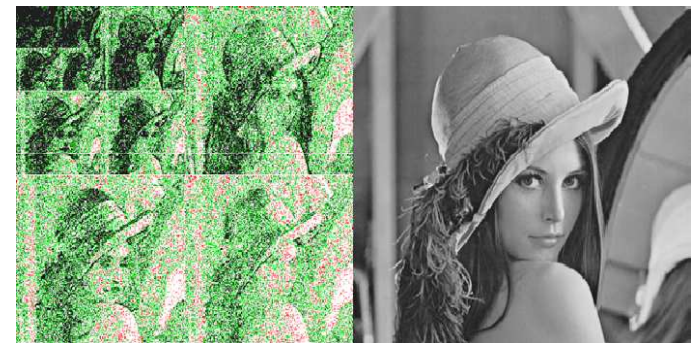
## Dekompozycja falkowa



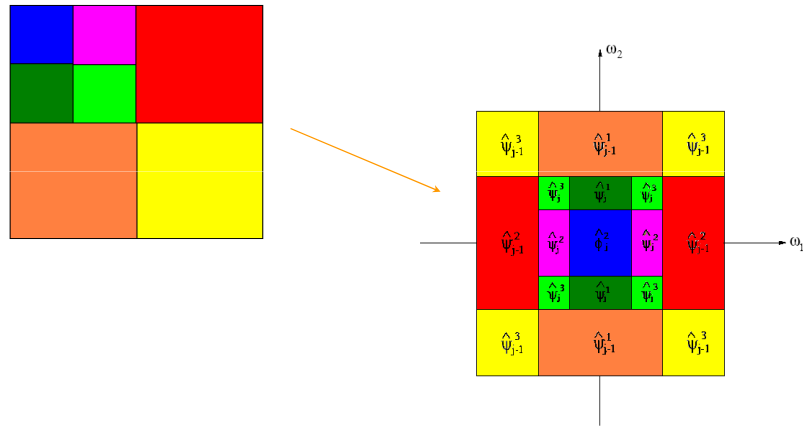
## Dekompozycja ...



## Rekonstrukcja ...



# Falkowa dekompozycja obrazów



# Nowe modele informacji: Analiza funkcjonalna

Sygnal skończony o ograniczonej energii

ciągły  $\rightarrow s(t) \in L^2$ , czyli  $\int_{-\infty}^{+\infty} s^2(t) dt < +\infty$

dyskretny  $\rightarrow s \in l^2$ , czyli  $\sum_n s_n^2 < +\infty$

**ANALIZA FUNKCJONALNA**

Interpolacja sygnału za pomocą bazy funkcji  $\phi$

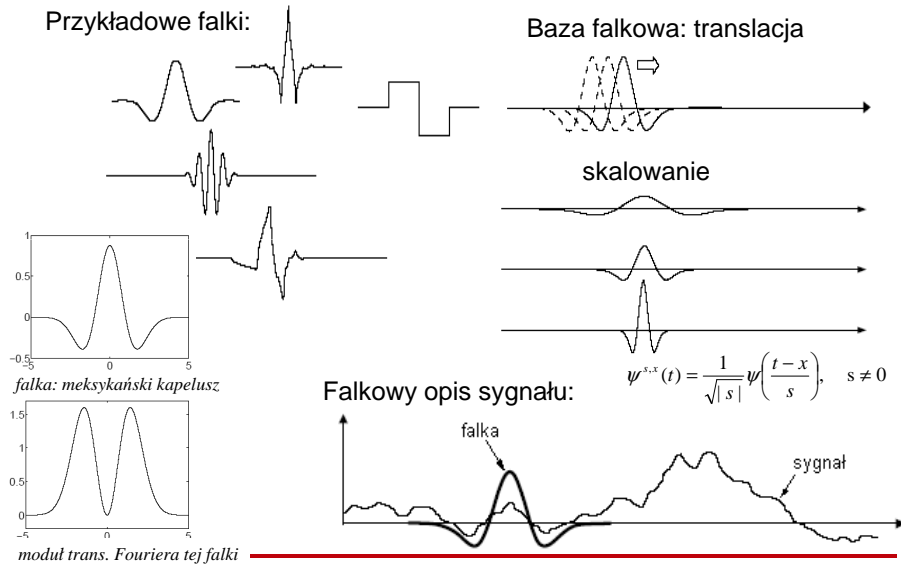
$$s = \sum_n a_n \phi_n, \text{ gdzie } a_n = \langle s, \phi_n \rangle = \int s_n \phi_n$$

Aproksymacja za pomocą mało licznej bazy (oszczędna reprezentacja):

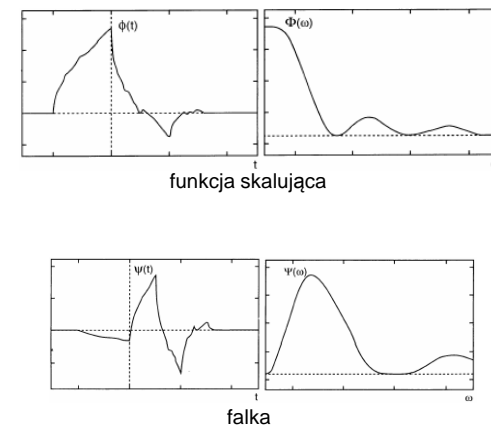
liniowa  $s = \sum_{n=1}^M a_n \phi_n$

nieliniowa  $s = \sum_{n \in A_M} a_n \phi_n$   
 – co do modułu - współczynników  
 ( $A_M$  jest zbiorem  $M$  największych)

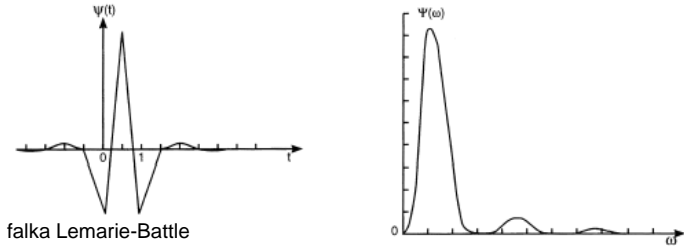
# Baza falkowa: dopasowana do cech sygnału (obrazu?)



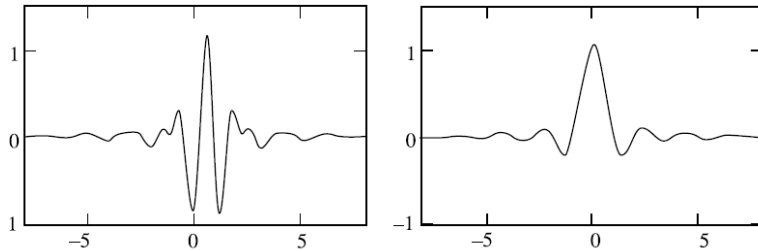
# Falki Daubechies



## Inne falki



falka Lemarie-Battle



falka Meyera wraz z funkcją skalującą

## Analiza wielorozdzielcza

$\phi(t)$  - funkcja skalująca, analogicznie baza  $\phi_{m,n}(t) = 2^{-m/2}\phi(2^{-m}t - n)$

Dekompozycja  $f$ :

$$f(t) = \sum_{n \in \mathbb{Z}} \langle m, \phi_{m,n} \rangle \phi_{m,n}(t) + \sum_{k \leq m, n \in \mathbb{Z}} \langle f, \psi_{k,n} \rangle \psi_{k,n}(t)$$

Sposób obliczania współczynników transformaty jest następujący:

$$a_{m,n}(f) = \sum_k h_{k-2n} a_{m-1,k}(f) \quad c_{m,n}(f) = \sum_k g_{k-2n} a_{m-1,k}(f)$$

współczynniki skalujące

współczynniki falkowe

Rekonstrukcja:

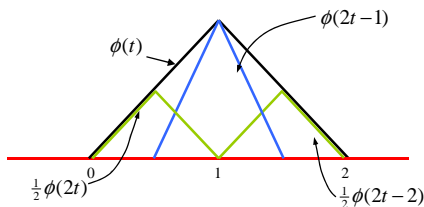
$$a_{m-1,n}(f) = \sum_k [h_{n-2k} a_{m,k}(f) + g_{n-2k} c_{m,k}(f)]$$

## Filtry

Filtry sprzężone z falkami (funkcjami skalującymi):

$$\phi(t) = \sqrt{2} \sum_n h_n \phi(2t - n)$$

$$\psi(t) = \sqrt{2} \sum_n g_n \phi(2t - n)$$



$$h_0 = h_2 = \frac{1}{2\sqrt{2}}, h_1 = \frac{1}{\sqrt{2}}$$

## Polecam zabawę z falkami

MAGING WEB DEMONSTRATIONS  
IOMEDICAL IMAGING GROUP

Fractional Spline Wavelet Transform

<http://bigwww.epfl.ch/demo/fractsplines/demoprep.html>

Filter: B-spline  
Degree: 3,14  
Shift: 0,00

Scaling function  
Wavelet function

Filter: B-spline  
Degree: 1,41  
Shift: 0,00

Scaling function  
Wavelet function

MRI Close Direct Transform Inverse Transform

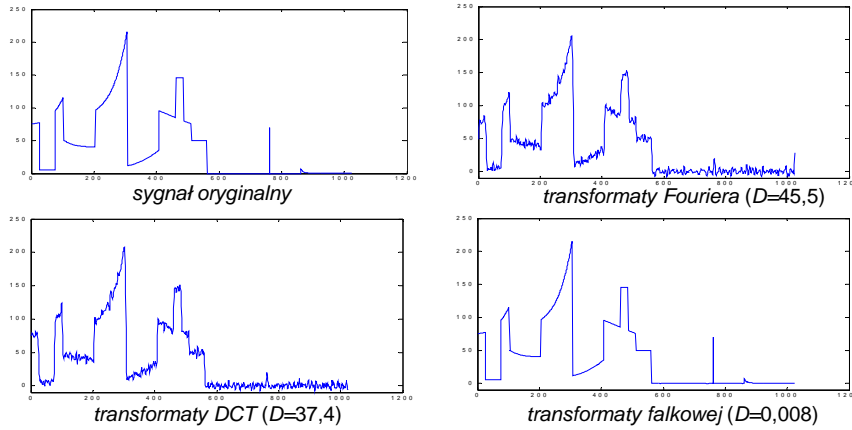
MRI Close Direct Transform Inverse Transform

Input: original  
Wavelet coefficients at iteration:3

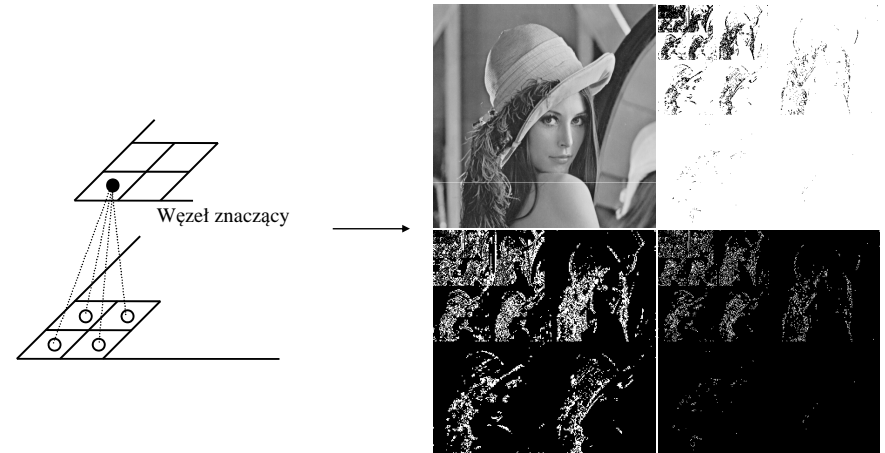
Input: original  
Wavelet coefficients at iteration:3

# Efekt: efektywna aproksymacja nieliniowa

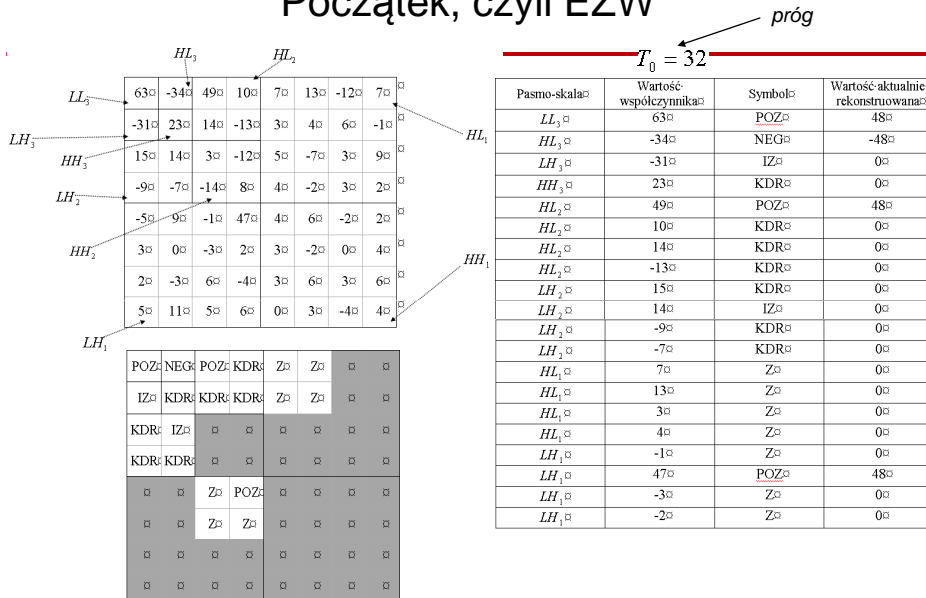
Oszczędna reprezentacja: aproksymacja z 15 % współczynników uporządkowanych nierosnąco



# Kodowanie: drzewa zer



# Początek, czyli EZW





# Testy JPEG2000 –JPEG



# Testy JPEG2000 – JPEG

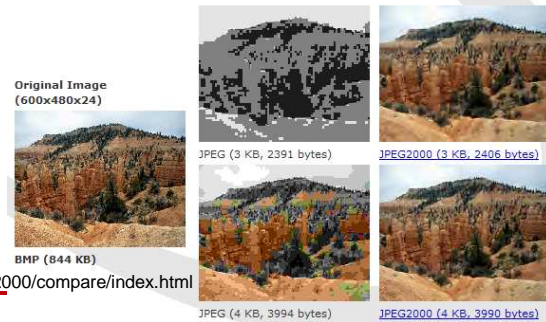


Compression Ratio => 310:1 (top row), 130:1 (bottom row)



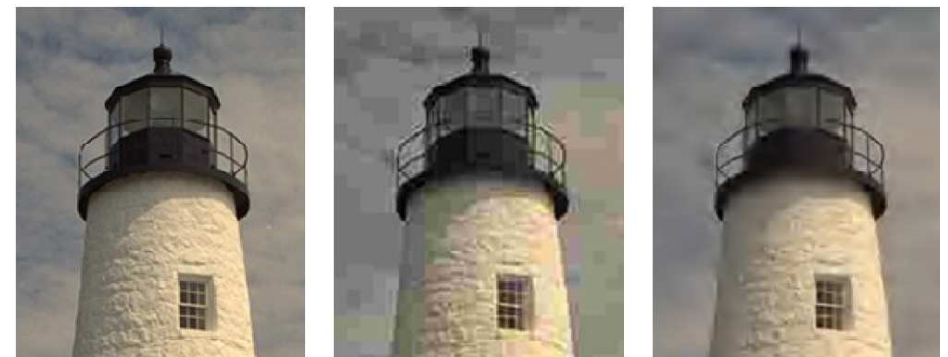
## JPEG vs JPEG2000

Compression Ratio => 350:1 (top row), 215:1 (bottom row)



<http://www.imagepower.com/technology/jpeg2000/compare/index.html>

## JPEG vs JPEG2000



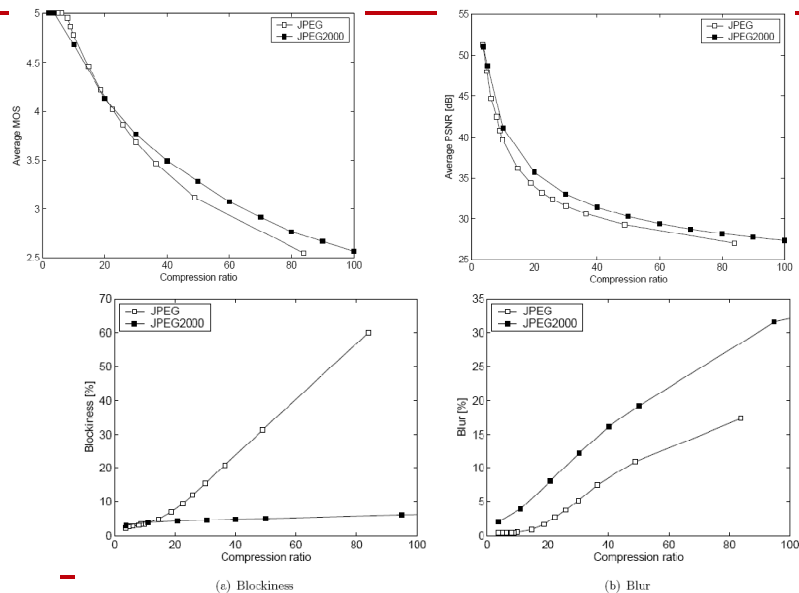
(a) Original image

(b) JPEG-encoded image

(c) JPEG2000-encoded image

F. Ebrahimi, M. Chamik, S. Winkler, JPEG vs. JPEG2000: An Objective Comparison of Image Encoding Quality, *Proc. SPIE Applications of Digital Image Processing*, 5558:300-308, 2004.

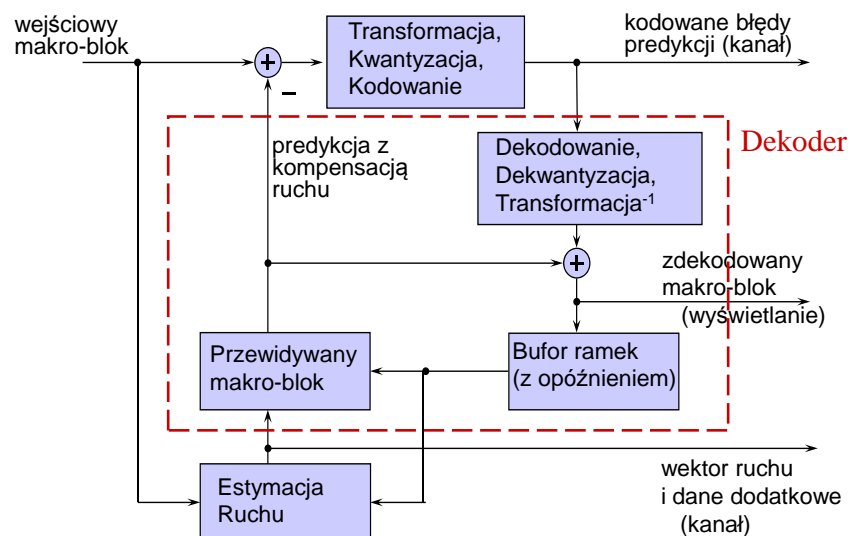
## JPEG vs JPEG2000



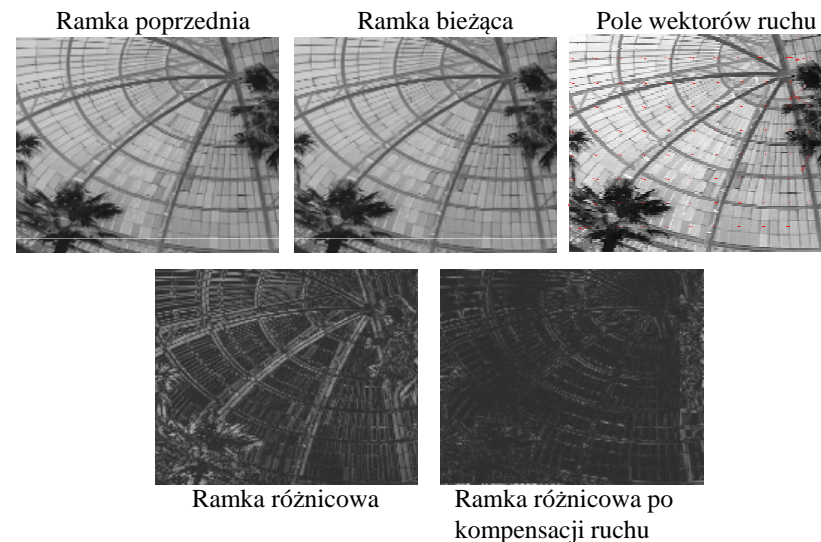
## Testy efektywności (wartości PSNR)

Technika kompresji	Lena		Barbara		Goldhill	
	0.25	0.5	0.25	0.5	0.25	0.5
PVQ	34.08	37.18	30.12	33.89	30.86	33.46
ECTCQ	33.90	37.10	29.66	33.51	30.67	33.35
KVQ-TCQ	34.31	37.67	-	-	-	-
SPIHT	34.11	37.21	29.36	33.07	30.56	33.13
ECECOW	34.81	37.92	28.85	32.69	-	-
RDE	34.20	37.20	29.10	33.10	-	-
SFQ	34.35	37.41	29.67	33.51	30.71	33.37
C/B	34.57	37.52	28.75	32.64	30.80	33.53
PACC	34.53	37.51	28.65	32.54	30.84	33.51
PC-AUTQ	34.46	37.56	-	-	30.78	33.46
EQ	34.57	37.68	-	32.87	30.76	33.44
TCSFQ=SFQ+TCQ+ECECOW	34.76	37.87	<b>30.60</b>	<b>34.48</b>	<b>30.98</b>	<b>33.75</b>
FD	<b>34.89</b>	<b>38.02</b>	29.21	33.06	-	-
SC&CE	34.72	37.75	28.76	32.80	-	-
MBWT	34.60	37.58	30.17	33.95	<b>30.99</b>	33.60
Stack-run	33.80	36.89	28.90	32.66	-	-
Koder indeksowy	33.44	36.59	-	-	-	-
Koder morfologiczny	34.12	37.18	-	-	30.53	33.15
JPEG2000	34.23	37.32	29.35	33.11	30.70	33.28

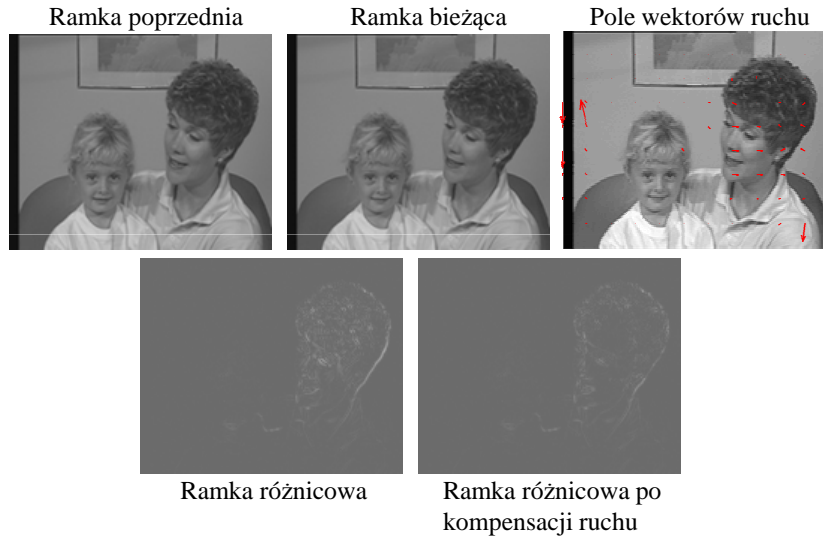
## Hybrydowa metoda kompresji, blokowa (DCT) z kompensacją ruchu



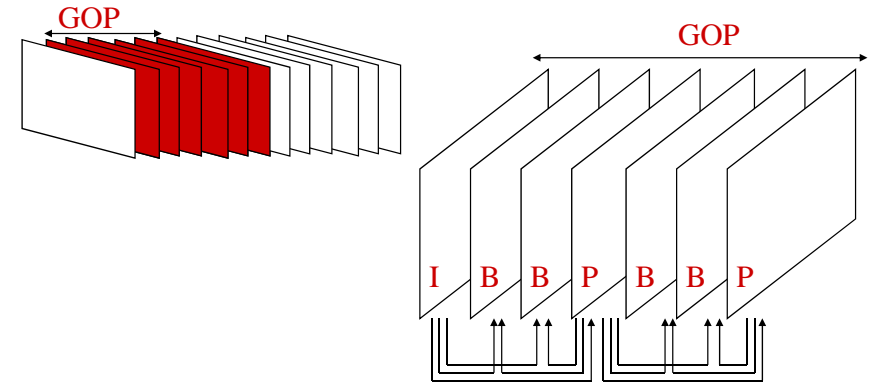
## Przykład 1



## Przykład 2



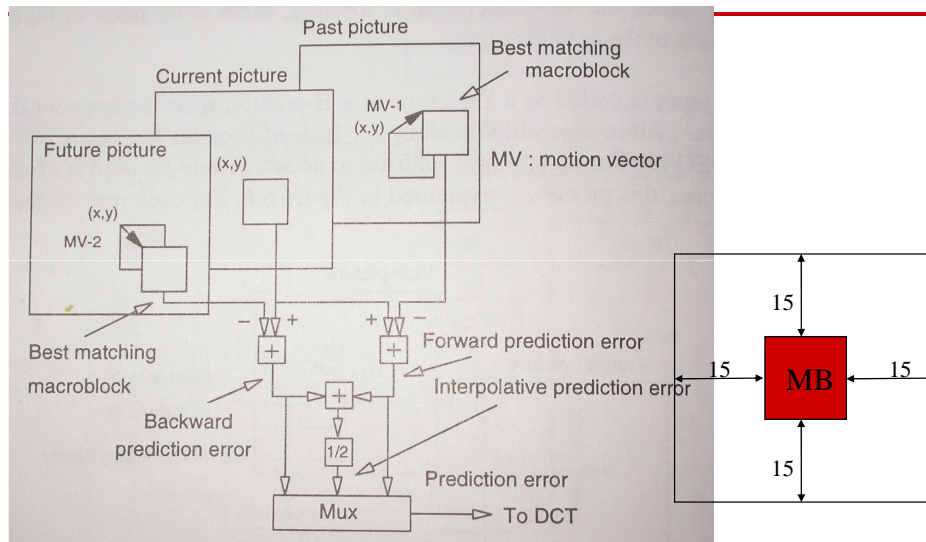
## Predycja INTER (międzyobrazowa)



Porządek kodowania: I P B B P B B

Porządek wyświetlania: I B B P B B P

## Estymacja ruchu

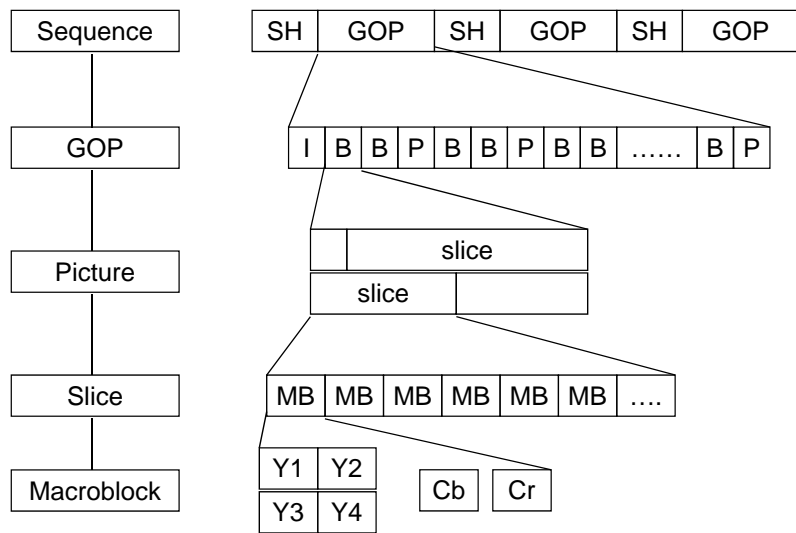


Makroblok 16x16

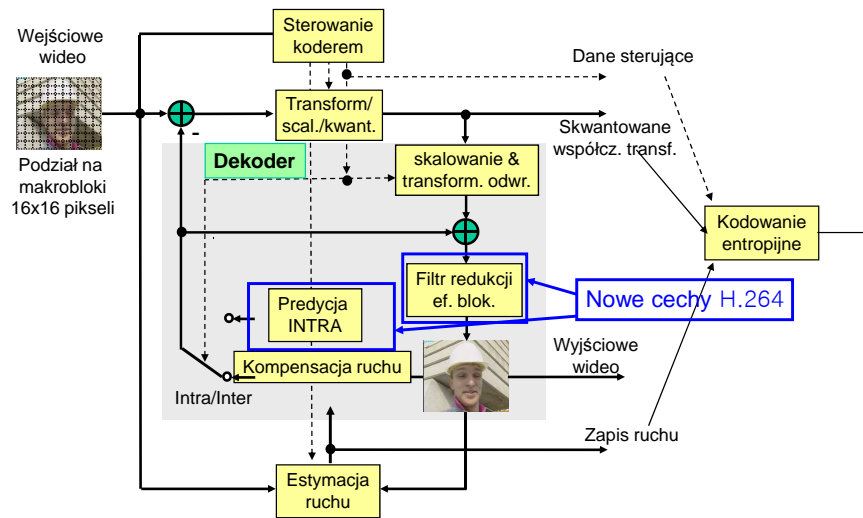
## Standardy rodziny MPEG i pokrewne

ITU-T Standard	H.261 (Version 1)	H.261 (Version 2)	H.263	H.263+	H.263++				
Joint ITU-T/MPEG Standards	H.262/MPEG-2		H.264/MPEG-4 AVC						
MPEG Standard	MPEG-1	MPEG-4 (Version 1)	MPEG-4 (Version 2)						
	1988	1990	1992	1994	1996	1998	2000	2002	2004

# Struktura danych w MPEG

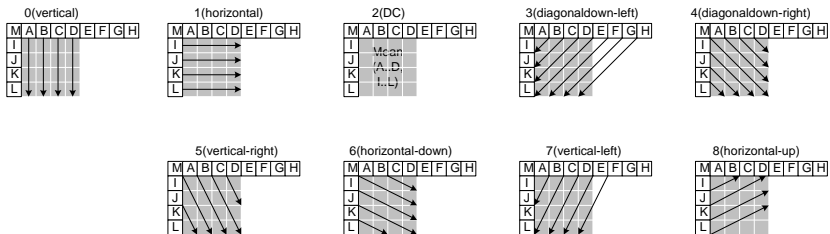


# Struktura kodeka h.264

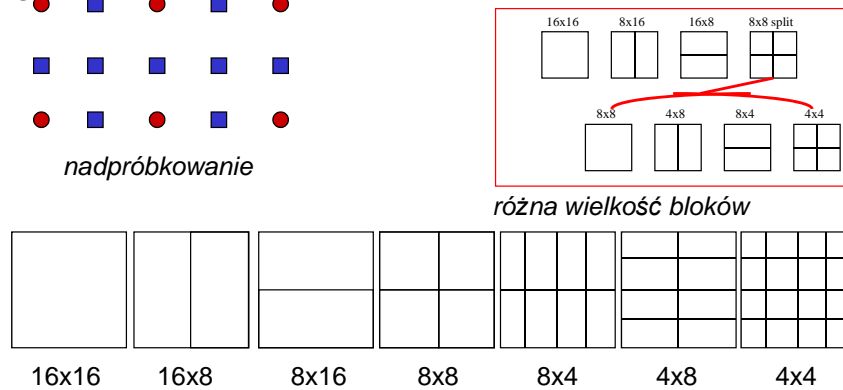


# Predycja INTRA

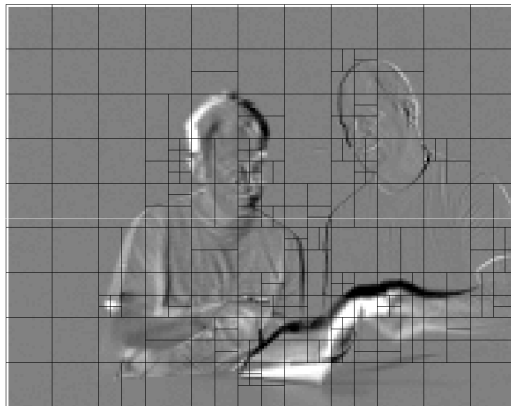
- 4 mody luminancji 16x16
- 9 modów luminancji 4x4
- 4 mody chrominancji 8x8
- Znacząca złożoność kodowania



# Estymacja ruchu z nadpróbkowaniem

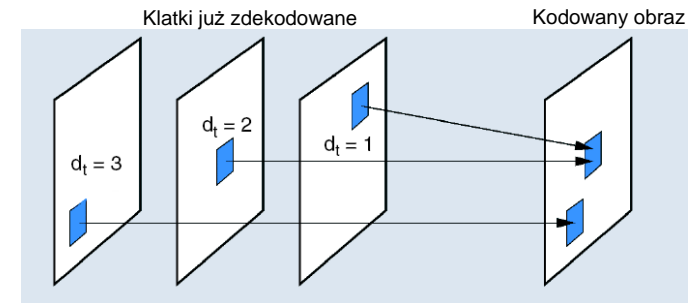


## Przykład podziału na makrobloki



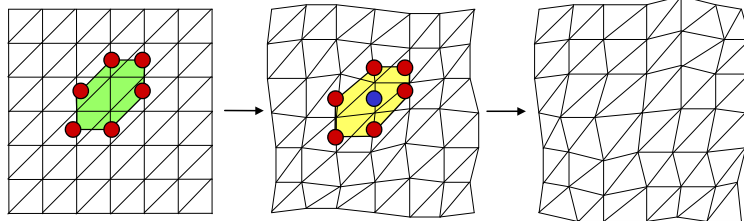
## Kompensacja ruchu

- Wiele obrazów referencyjnych
  - Długości wagi
  - Kierunek ruchu nieistotny
  - Klatki B mogą być referencyjne



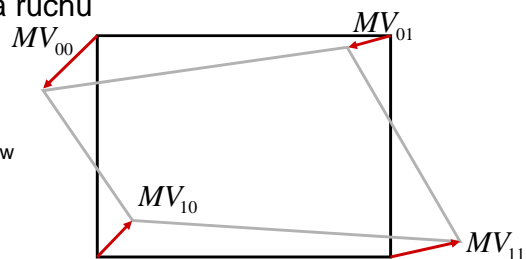
## Estymacja ruchu

- Nieregularne siatki do modelowania obiektów w ruchu



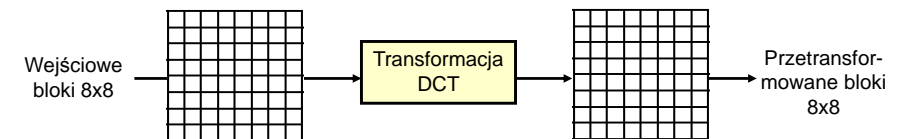
- Globalna estymacja ruchu

Estymacja dużych fragmentów obrazu z uwzględnieniem zmieniającej się perspektywy lub skali obiektów

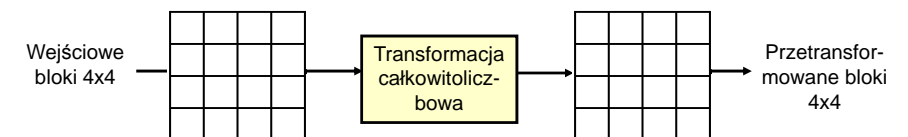


## Nowy rodzaj transformacji

- MPEG-2 / MPEG-4



- MPEG-4 AVC



## Transformacje w AVC

$$T_{4 \times 4} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

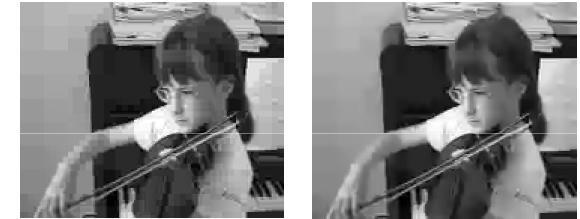
Rozszerzona na 8x8 dla Chroma za pomocą transformacji Hadamarda 2x2 współczynników DC  
Rozszerzona na 16x16 dla Luma Intra16x16 za pomocą transformacji Hadamarda 4x4 współczynników DC

$$T_{8 \times 8} = \begin{bmatrix} 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ 12 & 10 & 6 & 3 & -3 & -6 & -10 & -12 \\ 8 & 4 & -4 & -8 & -8 & -4 & 4 & 8 \\ 10 & -3 & -12 & -6 & 6 & 12 & 3 & -10 \\ 8 & -8 & -8 & 8 & 8 & -8 & -8 & 8 \\ 6 & -12 & 3 & 10 & -10 & -3 & 12 & -6 \\ 4 & -8 & 8 & -4 & -4 & 8 & -8 & 4 \\ 3 & -6 & 10 & -12 & 12 & -10 & 6 & -3 \end{bmatrix}$$

$$H_{2 \times 2} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad H_{4 \times 4} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

## Adaptacyjny filtr redukcji efektów blokowych

- Filtracja zależy od Intra/Inter, wektorów ruchu, różnic pomiędzy pikselami, QP
- Redukcja średniej bitowej o 6~9%
- Poprawa subiektywnej jakości oraz PSNR dekodowanych klatek



Bez filtru

Z filtrem