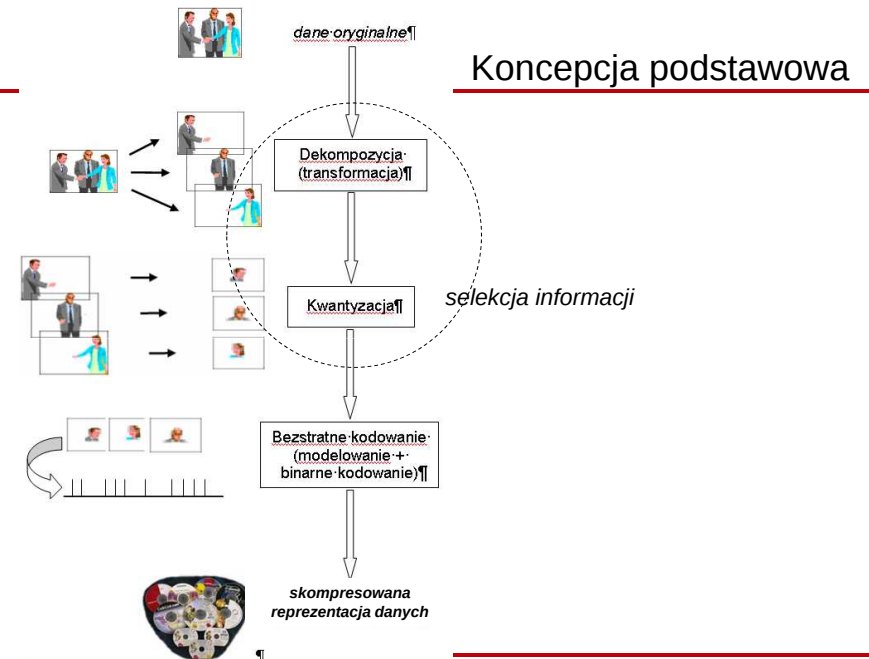


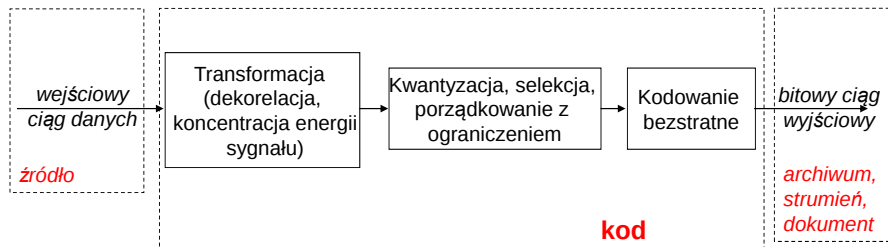
KOMPRESJA STRATNA OBRAZÓW

PODSTAWY TECHNIK MULTIMEDIALNYCH
A.Przelaskowski

- Wprowadzenie
- Kwantyzacja
- Koncepcja kodowania transformacyjnego
- DCT
- JPEG
- Falki
- JPEG2000
- Rodzina MPEG i inne standardy



Paradygmat klasyczny



Kwantyzacja

- Kwantyzacja to złożenie dwóch odwzorowań:
 - Kodera (k.prosta): nieskończony (skończony duży) zbiór wartości rzeczywistych z określonego zakresu → skończony (mały) zbiór wartości całkowitych (lub niekiedy słów kodowych, lub wartości o zdecydowanie mniejszej precyzji) – indeksów przedziałów kwantyzacji dzielących zakres wartości wejściowych
 - Dekodera (k.odwrotna): zbiór indeksów → zbiór (mały) rzeczywistych wartości rekonstruowanych (reprezentantów przedziałów kwantyzacji)

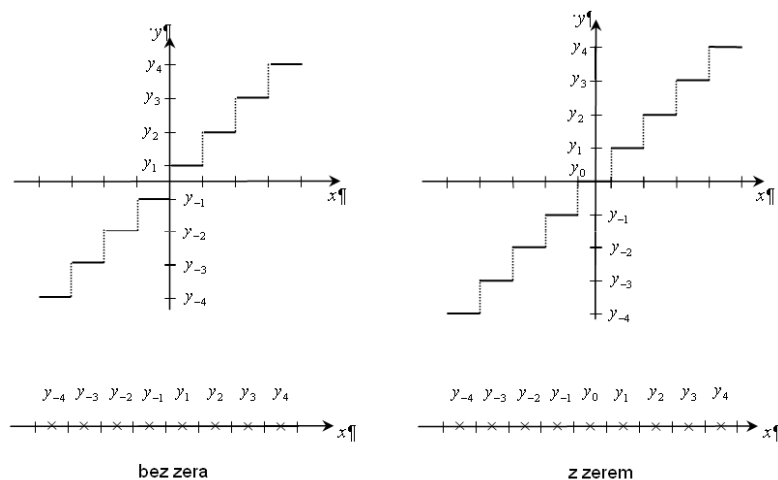
We: $X = \{x_i\}_{i=1}^K$ Wy: $\tilde{X} = \{\tilde{x}_i\}_{i=1}^K$

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

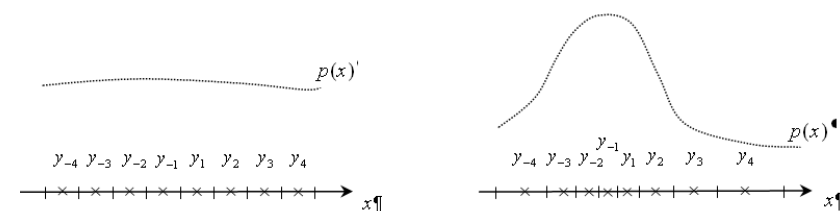
przy czym $\tilde{x}_i = y_j$ jeśli $x_i \in [\beta_{j-1}, \beta_j)$, a $\{\beta_j\}_{j=0}^M$ jest zbiorem punktów

granicznych przedziałów kwantyzacji z punktami rekonstrukcji $\{y_j\}_{j=1}^M$

Kwantyzacja równomierna



Kwantyzacja sterowana estymatą funkcji gęstości prawdopodobieństwa



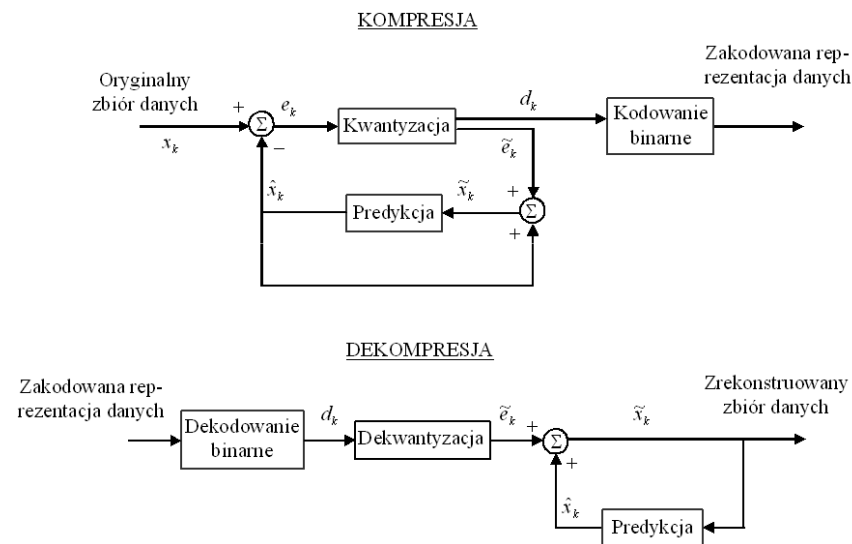
- Metoda Lloyd-Maxa (iteracyjna, zbieżna do lokalnego optimum, duży wpływ inicjalizacji, czasochłonna)
- Obowiązujące zasady centroidu i najbliższego sąsiada

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

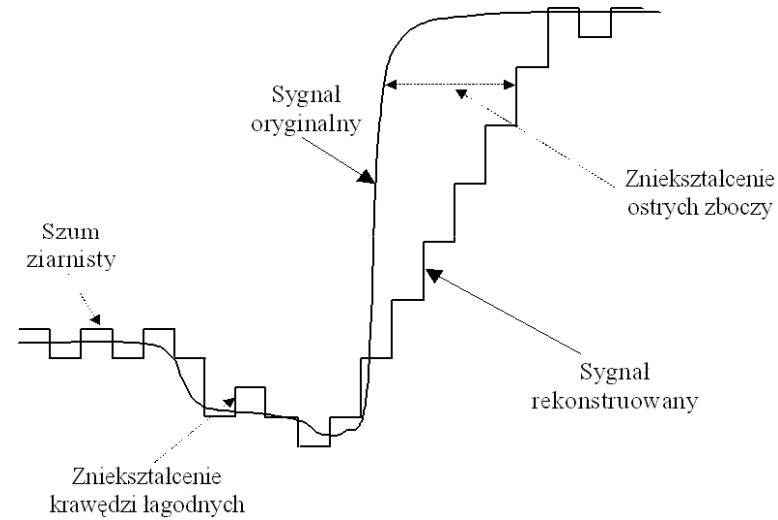
$$\beta_j = \frac{y_j + y_{j+1}}{2}$$

WYBRANE METODY KOMPRESJI STRATNEJ

Stratna DPCM



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)

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

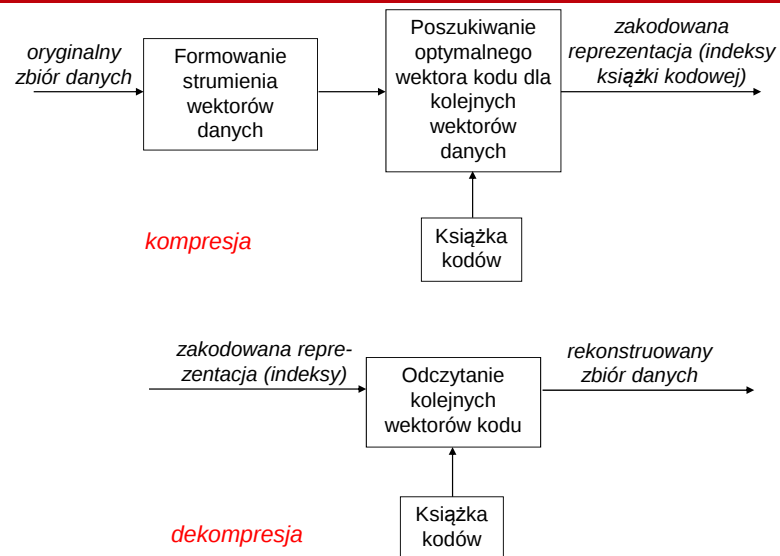
Wyznaczenie map bitowych

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

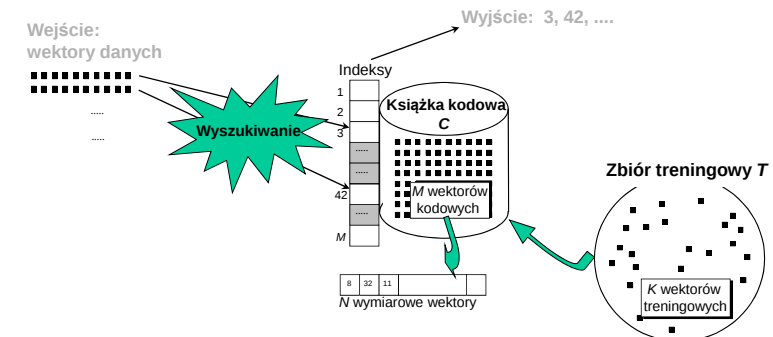
Rekonstrukcja

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

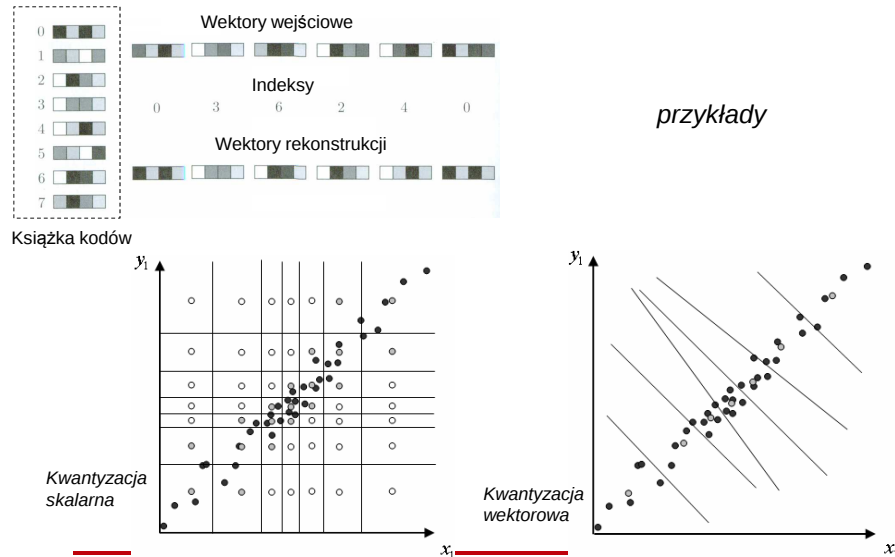
Kwantyzacja wektorowa VQ (metoda kompresji)



VQ bardziej obrazowo



VQ - wyjaśnienia

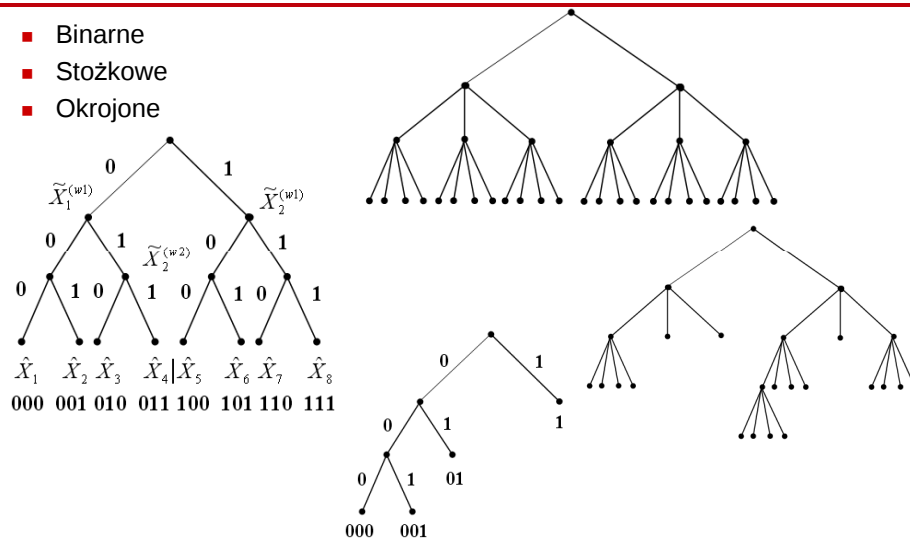


VQ - zagadnienia

- Budowa książki kodowej (kraty-nośniki, metoda LBG)
- Na bazie LBG: Inicjalizacja (losowa, grupowanie, rozdzielanie)
- Struktura książki kodowej (drzewa, adaptacja, klasyfikacja)
- Wykorzystanie książki kodowej do kompresji
- Formowanie, kodowanie indeksów
- Ograniczenia metody (złożoność, brak optymalnych metod wyznaczania książki, małe stopnie kompresji)

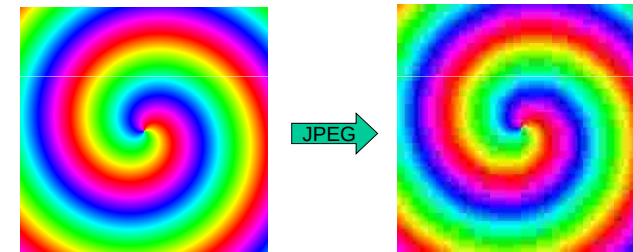
KK - drzewo

- Binarne
- Stożkowe
- Okrojone



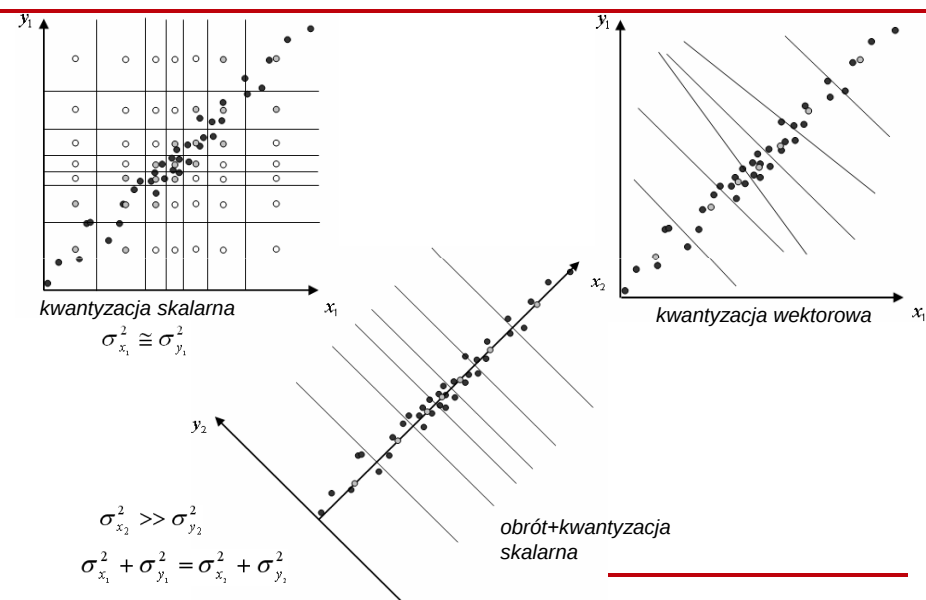
Metody kodowania transformacyjnego

- Uproszczenie kwantyzacji
- Dobór transformacji
- Metoda kwantyzacji z kodowaniem



Napis Testowy Napis Testowy

Koncepcja kodowania: uproszczenie kwantyzacji



Transformacja

Pożądane cechy:

- Dekorelacja danych (usuwanie zależności statystycznych), upakowanie energii wzdłuż wybranych kierunków (wokół elementów bazy)
- Dobrana baza transformacji (dopasowana lokalnie do sygnału, elastyczna, zupełna)
- Szybka implementacja

Transformacja DCT (2W)

prosta

$$k(u, v) = \frac{1}{\sqrt{2n}} c_u c_v \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} f(x, y) \cos \frac{(x + \frac{1}{2})u\pi}{n} \cos \frac{(y + \frac{1}{2})v\pi}{n}$$

$$c_u, c_v = \frac{1}{\sqrt{2}} \text{ dla } u, v = 0; \quad c_u, c_v = 1 \text{ w p.p.}$$

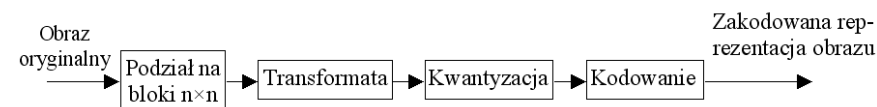
odwrotna

$$f(x, y) = \frac{1}{\sqrt{2n}} \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} c_u c_v k(u, v) \cos \frac{(x + \frac{1}{2})u\pi}{n} \cos \frac{(y + \frac{1}{2})v\pi}{n}$$

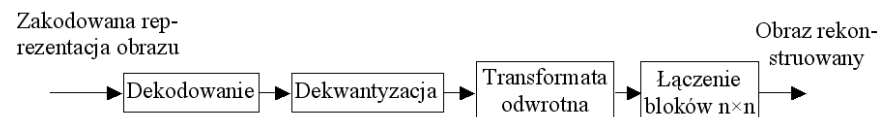
Stosowana m, in , w JPEG

Paradygmat TRANSFORMACYJNEGO KODOWANIA

KOMPRESJA



DEKOMPRESJA



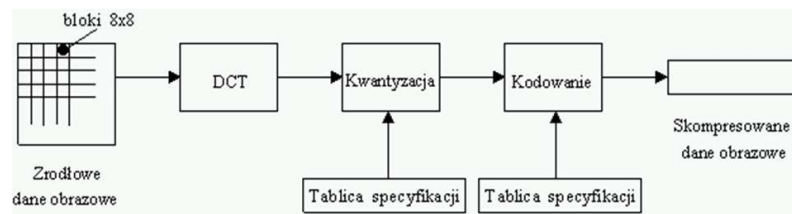
Selekcja próbek

blok obrazu	109	101	91	79	80	85	84	89	784.2	6.630	-6.45	-5.46	2.25	0.000	-3.82	-2.80
	87	87	89	99	98	91	91	93	-34.9	10.89	-0.46	5.318	4.209	3.641	3.270	-2.62
	100	102	109	127	112	99	87	84	0.726	10.11	12.68	-0.72	2.956	-5.22	-0.91	3.229
	84	89	87	88	83	90	100	94	-19.8	2.843	21.02	7.299	-5.43	-7.31	7.198	2.446
strefowa selekcja próbek	100	101	108	112	110	114	113	109	9.000	-2.50	20.10	1.802	-8.00	5.061	-0.48	1.331
	103	103	95	89	89	89	82	86	-21.7	-3.61	8.453	-1.76	-5.13	-3.66	0.128	0.899
	84	97	100	100	108	103	106	112	2.405	37.31	-0.16	3.307	2.564	0.964	2.072	-3.61
	111	115	111	113	107	100	109	107	40.72	-3.57	-18.2	-1.70	3.892	0.977	2.199	-0.53
rekonstrukcje bloku	784	7	-6	-5	2	0	0	0	784	0	0	0	0	0	0	0
	-35	11	0	5	0	0	0	0	-35	11	0	0	0	0	0	0
	1	10	13	0	0	0	0	0	0	10	13	0	0	0	0	0
	-20	3	0	0	0	0	0	0	-20	0	21	0	0	0	0	0
DCT	9	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0
	0	0	0	0	0	0	0	0	-22	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	37	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	41	0	-18	0	0	0	0	0
progowa selekcja próbek (próg=10)	99	96	93	90	88	86	85	85	106	99	89	81	77	80	85	89
	97	96	94	93	92	90	89	89	87	89	91	93	95	97	97	98
	96	97	98	100	99	97	95	94	100	106	114	118	114	103	90	82
	95	97	101	104	104	101	99	98	85	85	85	86	88	91	94	96
DCT	93	96	100	103	102	99	97	97	100	102	104	107	110	112	113	113
	93	96	99	100	98	96	96	97	102	99	94	90	88	88	89	91
	100	102	103	103	100	99	101	103	90	93	99	104	108	110	109	109
	109	109	109	108	106	105	108	112	114	112	108	106	104	105	106	107

STANDARDY JPEG

Standard JPEG (szczególna realizacja paradygmatu TK)

- Podstawowy proces kodowania (ang. *baseline process*).
- Rozszerzony, bazujący na DCT proces kodowania (ang. *extended DCT-based process*).
- Bezstratny proces kodowania (ang. *lossless process*).
- Hierarchiczny proces kodowania (ang. *hierarchical process*).

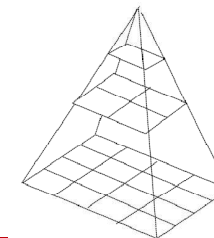


Właściwości kodeka JPEG

- Progresja kodowania



- Kodowanie hierarchiczne (z podpróbkowaniem)



Kwantyzacja i kodowanie w JPEG

składowa stała

percepcyjna tablica kwantyzacji

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Kwantyzacja/selekcja:

dzielenie współczynników DCT $k(u,v)$ przez element tablicy kwantyzacji o współrzędnych (u,v) i zaokrąglanie do liczby całkowitej

współczynniki DCT

310	38	-18	7	-40	12	35	25
103	-7	-37	-26	7	-21	21	-45
-84	6	-23	25	-12	33	43	-32
-47	14	-50	27	49	-61	20	6
25	-54	-32	2	-43	-2	23	-70
-5	-135	-2	45	5	-17	26	-18
28	103	22	-8	-4	17	-11	-16
95	-16	16	-3	1	27	-22	-18

tablica kwantyzacji

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

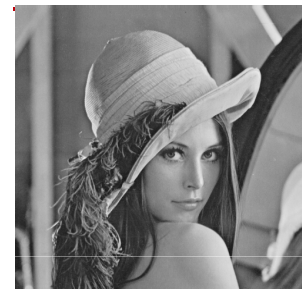
skwantowane współczynniki DCT

19	3	1	0	1	0	0	0
-8	0	1	1	1	0	0	0
-6	0	1	1	1	0	0	0
-3	0	1	1	1	0	0	0
1	-2	0	0	0	0	0	0
0	3	0	0	0	0	0	0
0	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0

Kodowanie:

przebieganie zygzak, różnicowo składowa stała bloków (predykcja), kod Huffmana-RLE lub kod arytmetyczny

Efekty JPEGowe



Oryginał



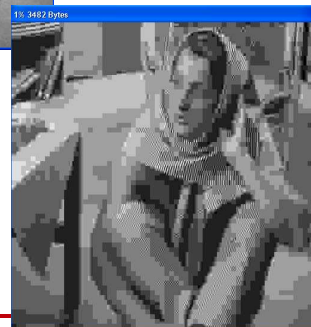
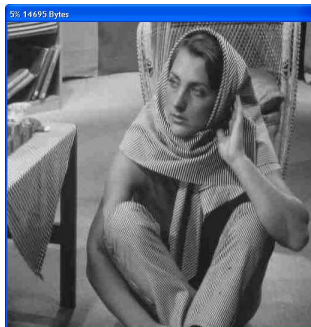
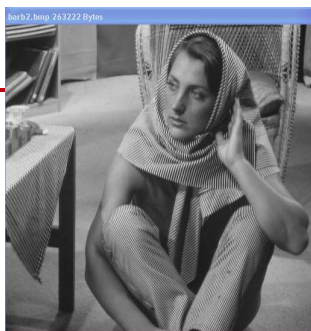
12:1



43:1



Efekty JPEGowe



Ćwiczenie z JPEG

<http://www.sfu.ca/~cjenning/toybox/hjpeg/index.html>

- Transform of each channel. Notice that you can make out the patterns of little 8x8 boxes. Remember that the upper-left corner of those boxes is the low-frequency part, the other corner is high-frequency. Brighter pixels mean more information, i.e. more of the image was of that kind of frequency.
- Now choose the next drop down item (Quantized DCTs). Whoosh! The lower corners of each of the 8x8 boxes just went dark. The values in each box were just divided by the values in the corresponding cells (throwing away any remainders) of the quantization tables (Step 3), so they now have less information. The next drop down box shows the reconstructed DCTs, i.e. what we get back when we decompress the image by multiplying by the values in the tables. Compare the original DCT of the colour planes to the reconstructed versions. A lot of colour information has been lost, but it is not very noticeable in the final decompressed picture.

JPEG and Hierarchical JPEG Demo

1. Choose a sample image:

2. Choose a chroma subsampling format: ☒ None (4:4:4) ☐ Quartered (4:2:0)

3. Choose a quality setting or... Low High

...create custom quantization tables:

Luminance								Chrominance							
28	21	24	24	32	42	87	129	19	21	23	30	39	62	114	105
17	24	28	39	66	98	140	170	28	33	42	51	100	114	156	176
42	46	71	91	122	145	185	201	71	104	102	156	195	186	217	179
91	107	123	143	185	203	215	185	109	98	100	111	138	165	181	177

Zoom Level: x y

R G B

Y Cb Cr

0 0 0

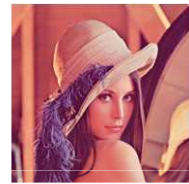
The first row of monitors shows the input image. The second row shows: Quantized DCTs

Data Values from Current 8 x 8 Data Block

40	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

: applet requires that your browser support at least the Java 2 runtime. If the applet above doesn't work properly, this is almost certainly the blem. You can download the latest Java runtime from Sun to correct this.

Paradygmat JPEG



Kompresja/Dekompresja



Opcje kodera:

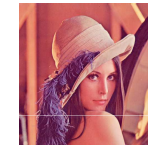
- przestrzeń kolorów
- kwantyzacja
- entropijne kodowanie
- przetwarzanie wstępne

Brak opcji dekodera: (ustalony w kodzie 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

progresja z wybranym ROI

rekonstrukcja wybranego ROI

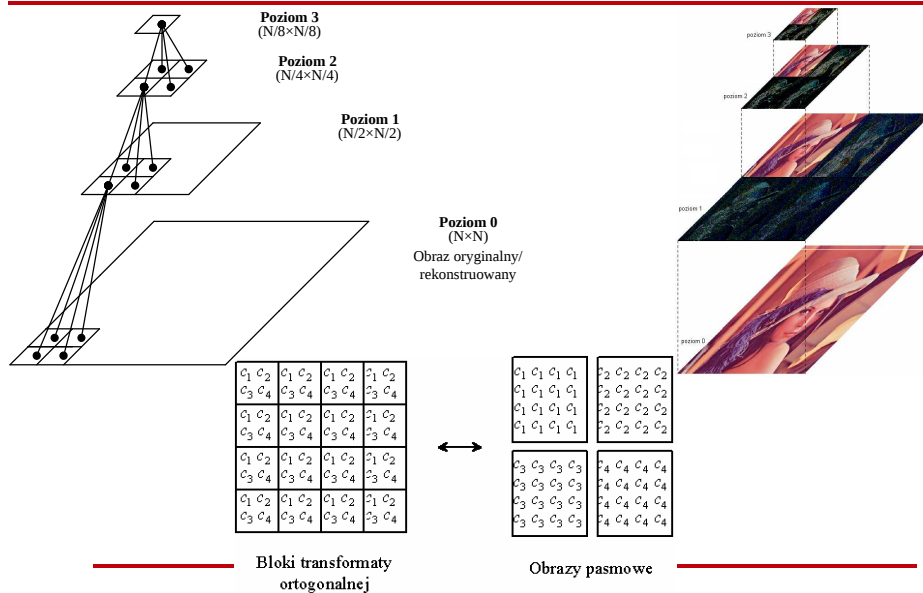
Opcje dekodera:

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

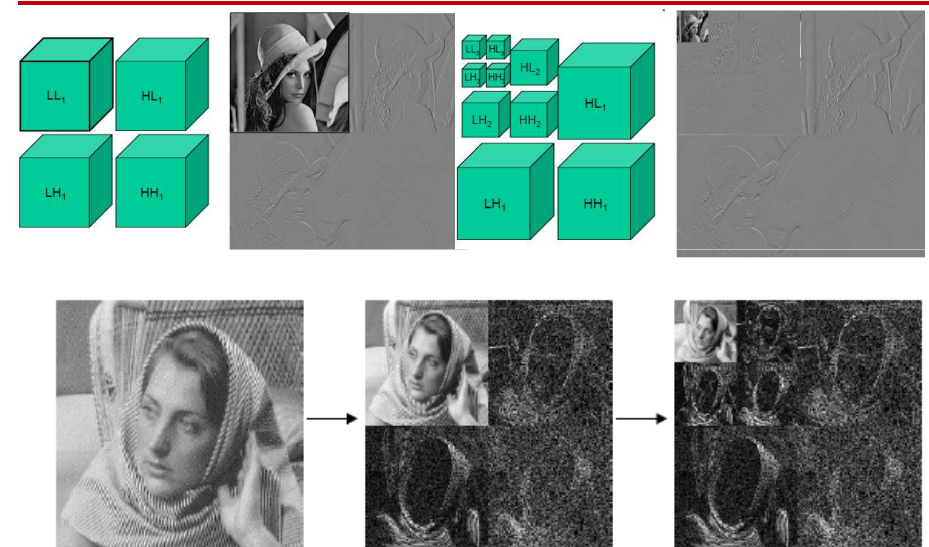
Dodatkowe możliwości schematu kompresji:

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

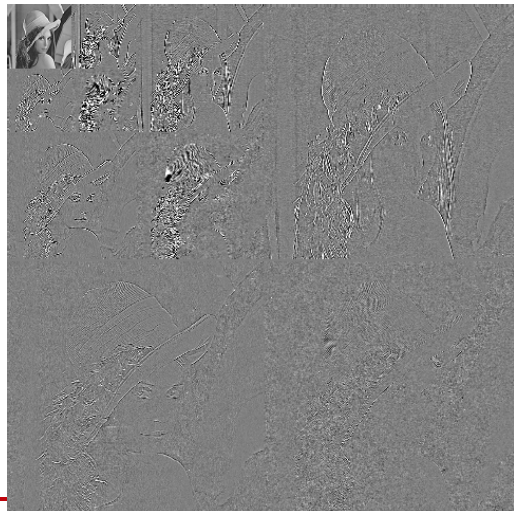
SPOSÓB: hierarchia skal, podpasma i lokalne upakowanie



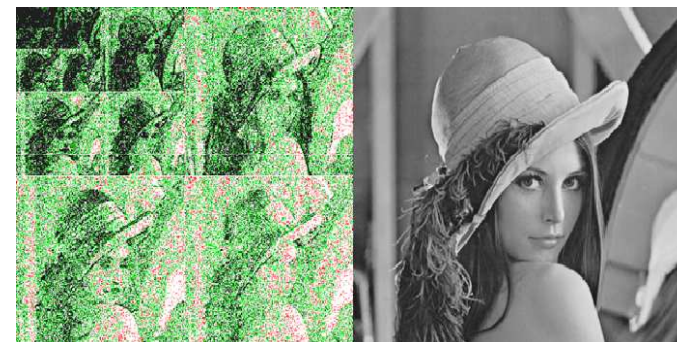
Dekompozycja falkowa



Dekompozycja ...



Rekonstrukcja ...

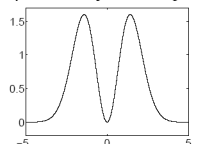


Baza falkowa: dopasowana do cech sygnału (obrazu?)

Przykładowe falki:

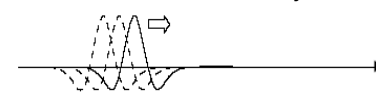


falka: meksykański kapelusz

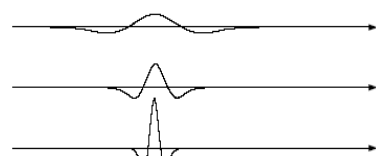


moduł trans. Fouriera tej falki

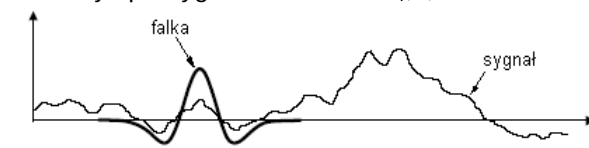
Baza falkowa: translacja



skalowanie

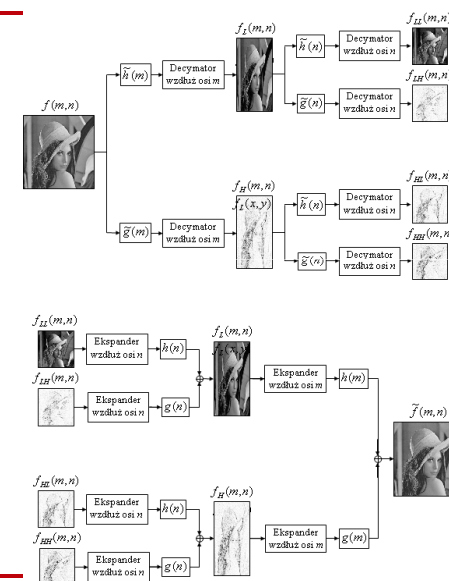


Falkowy opis sygnału:

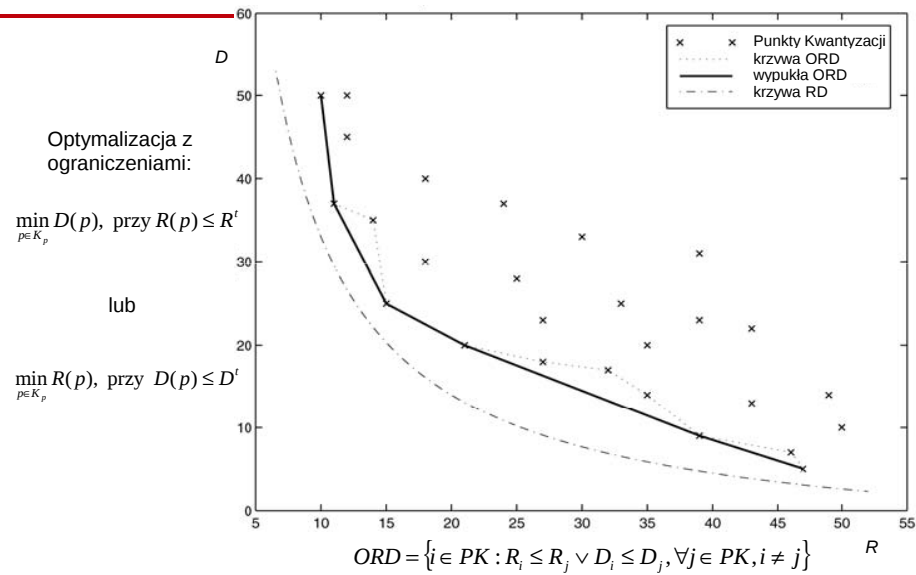


$$\psi^{s,x}(t) = \frac{1}{\sqrt{|s|}} \psi\left(\frac{t-x}{s}\right), \quad s \neq 0$$

Realizacja dekompozycji falkowej

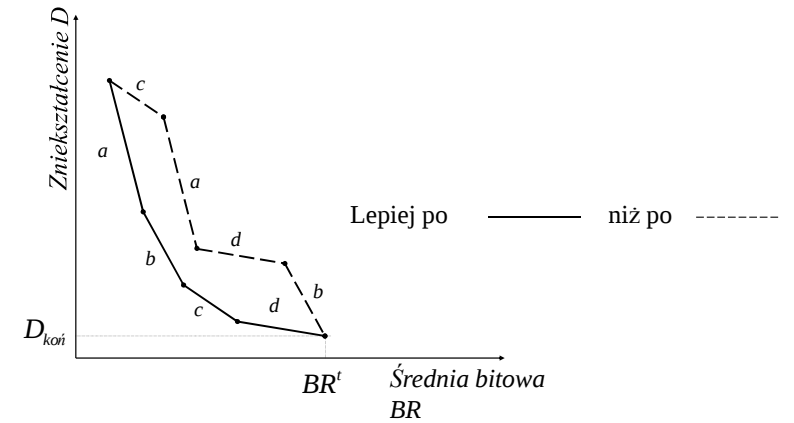


Podstawy R-D: Operacyjna krzywa RD



Lepiej, czyli porządkowanie informacji w JPEG2000

Optymalizacja w sensie R-D, czyli progowanie z kontekstem kodowania



Testy

JPEG2000

JPEG

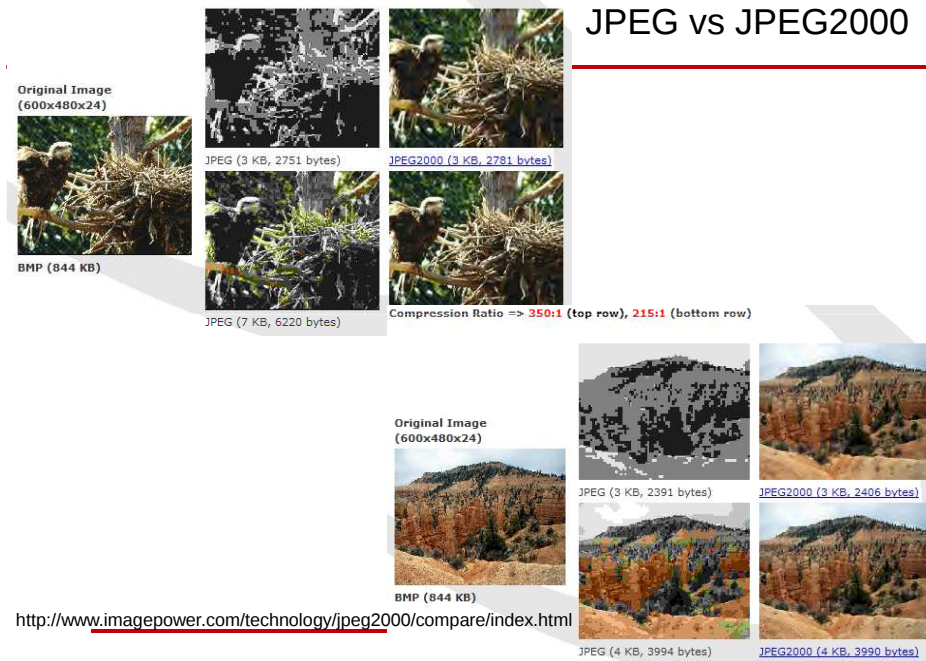


0.25bpp

Testy JPEG2000 – JPEG



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



JPEG vs JPEG2000



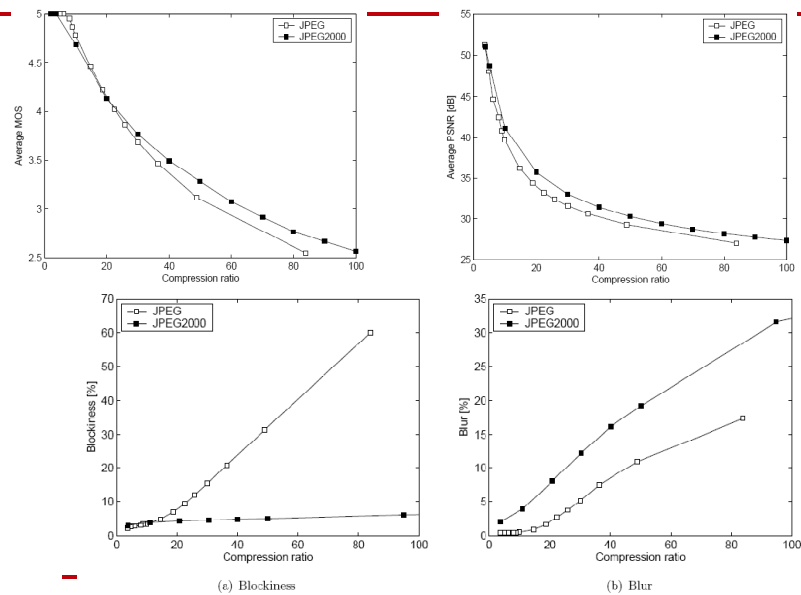
(a) Original image

(b) JPEG-encoded image

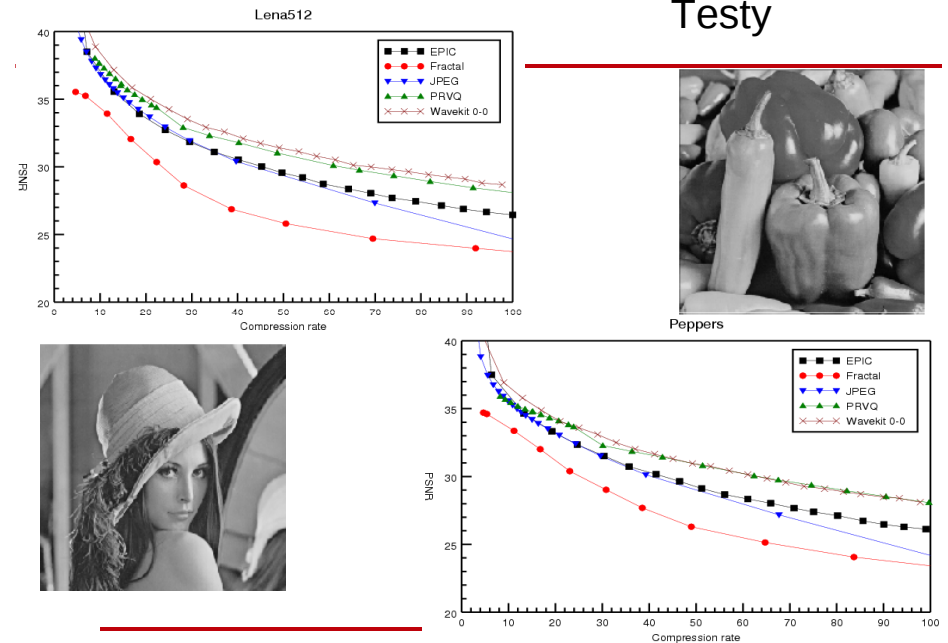
(c) JPEG2000-encoded image

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



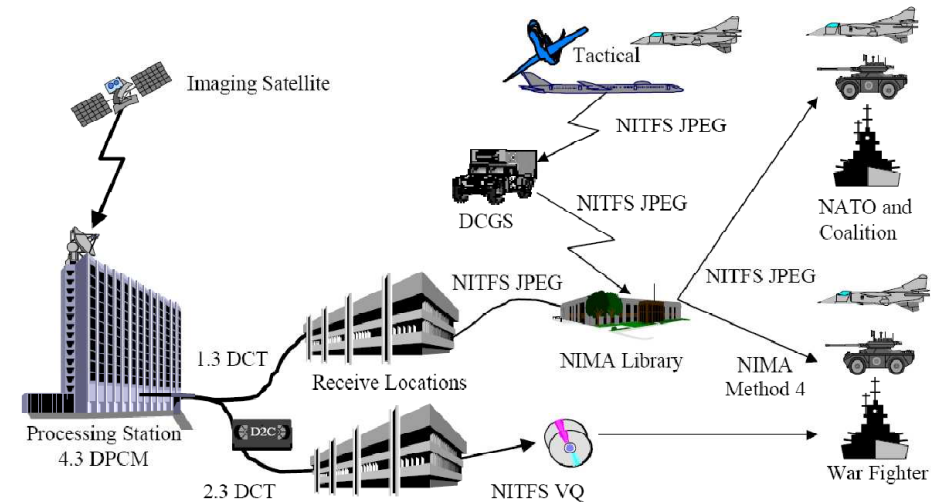
Zastosowania teleinformatyczne

- Interaktywna transmisja (JPIP)
- Transmisja bezprzewodowa (JPWL)
- Kompresja sekwencji obrazów - kino domowe (Motion JPEG2000)
- Kompresja danych 3W (JP3D)
- Ochrona danych (JPSEC)

Zapotrzebowanie na kompresję ...



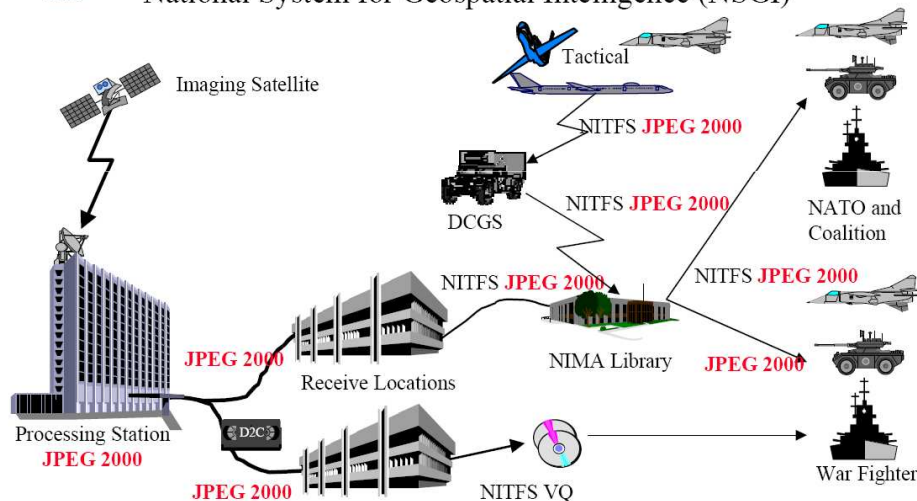
National System for Geospatial Intelligence (NSGI)



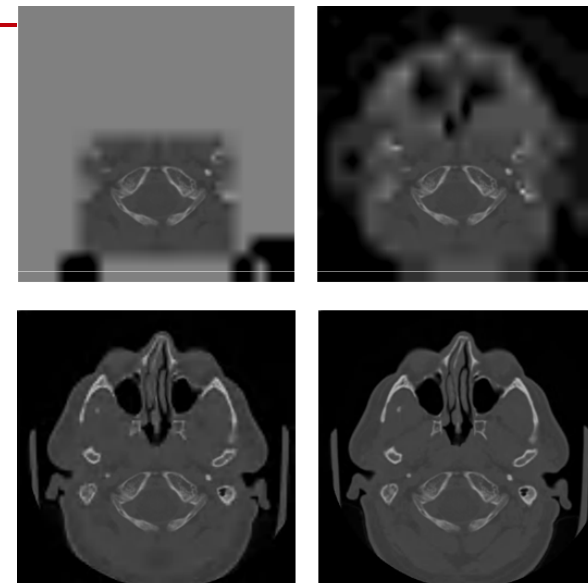
Standard uniwersalny



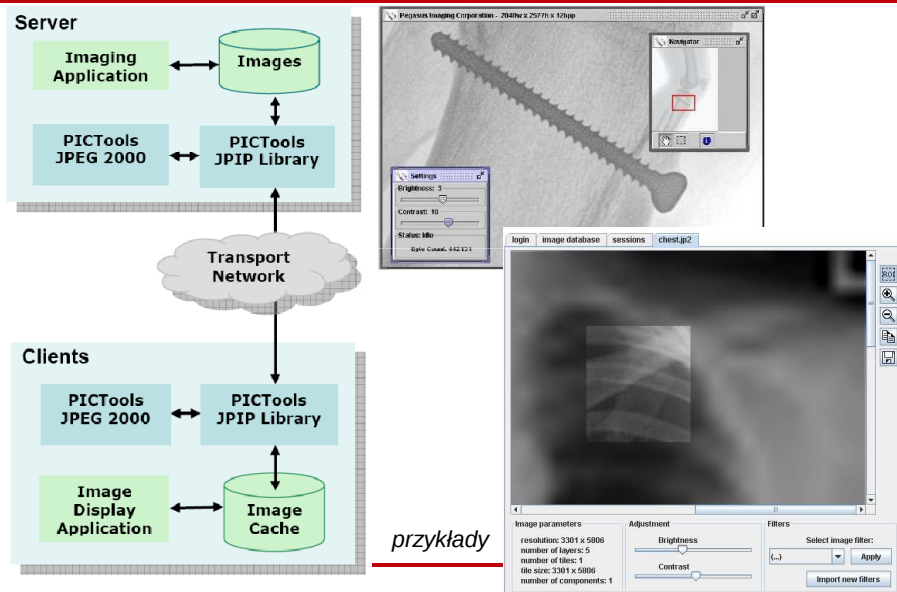
National System for Geospatial Intelligence (NSGI)



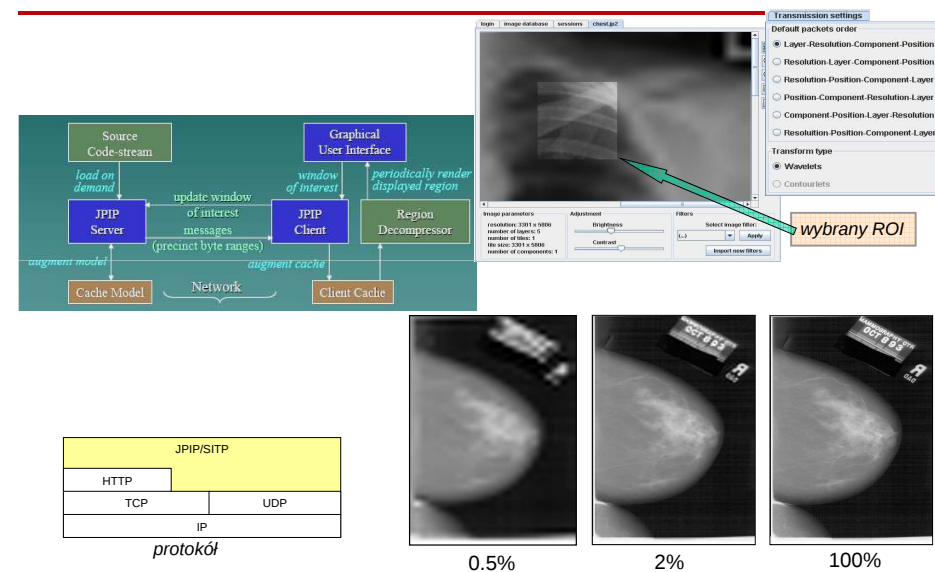
Interaktywna transmisja (JPIP)



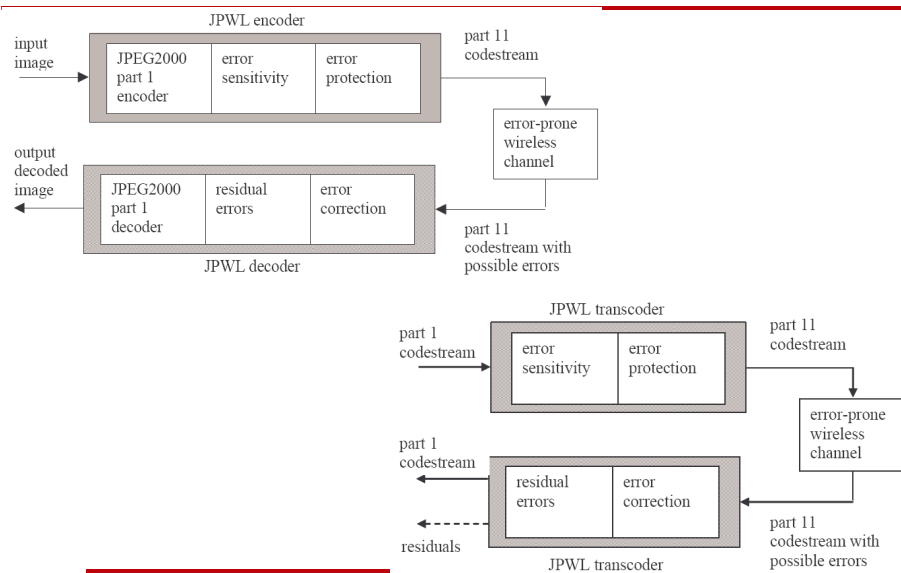
JPIP



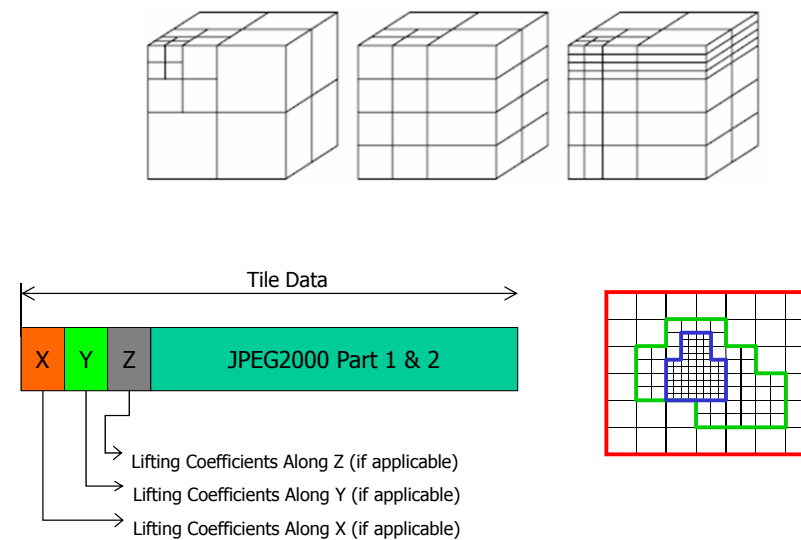
JPIP (przykłady)



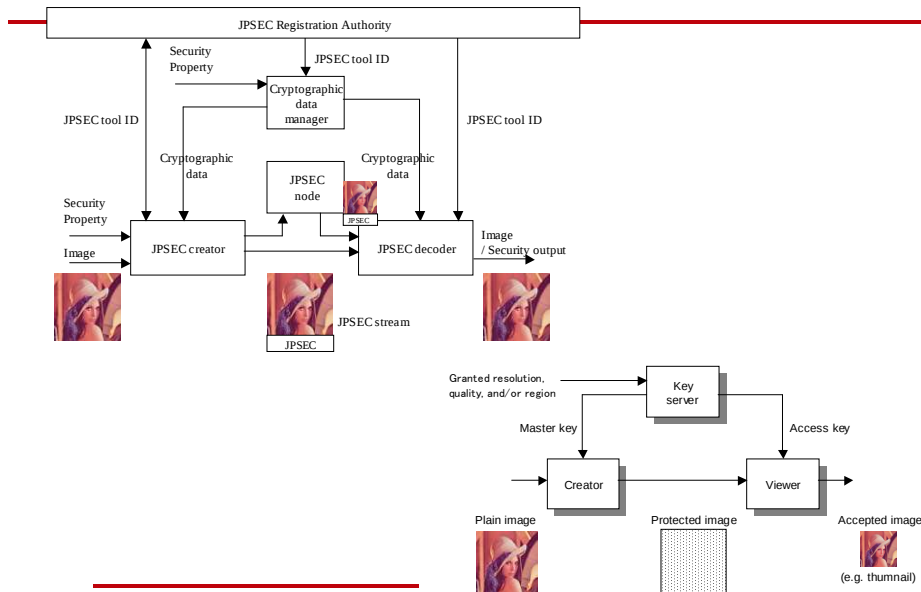
JPWL



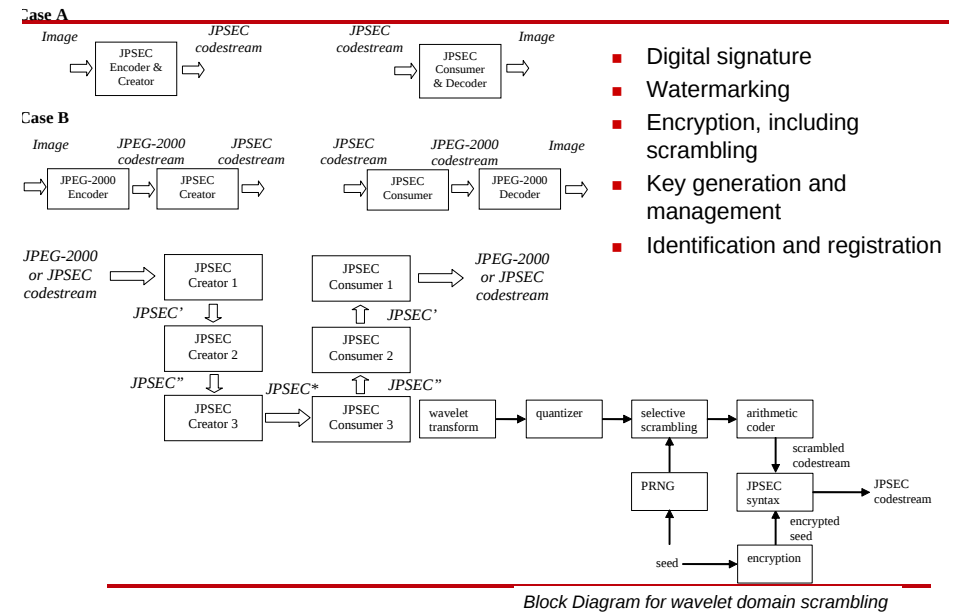
JP3D



JPSEC



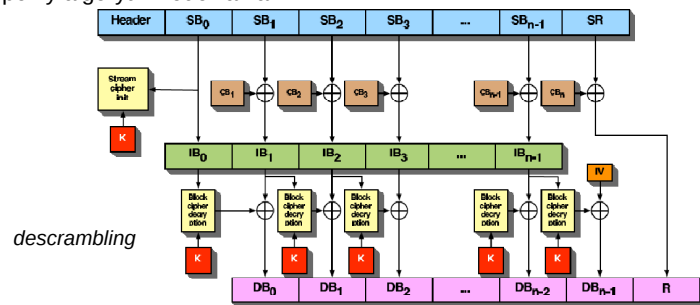
JPSEC



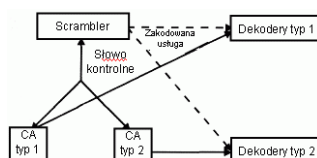
Block Diagram for wavelet domain scrambling

Scrambling (szyfrowane kodowanie)

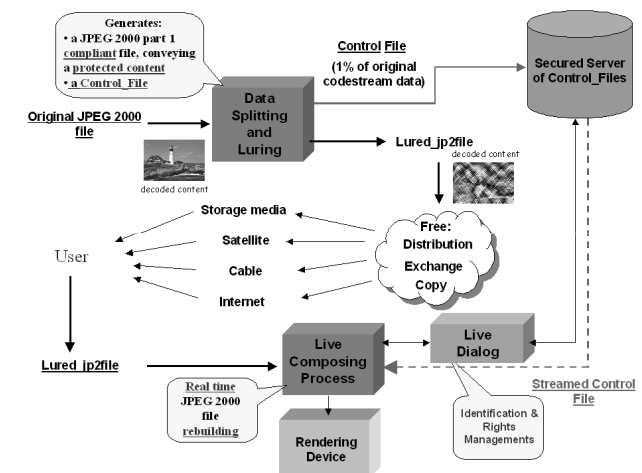
- Wspólny algorytm kodowania



- Różne systemy warunkowego dostępu (CA) z generacją słów kontrolnych



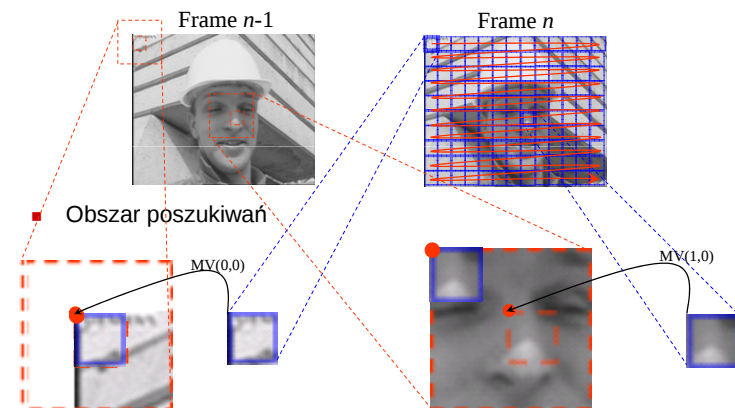
JPSEC



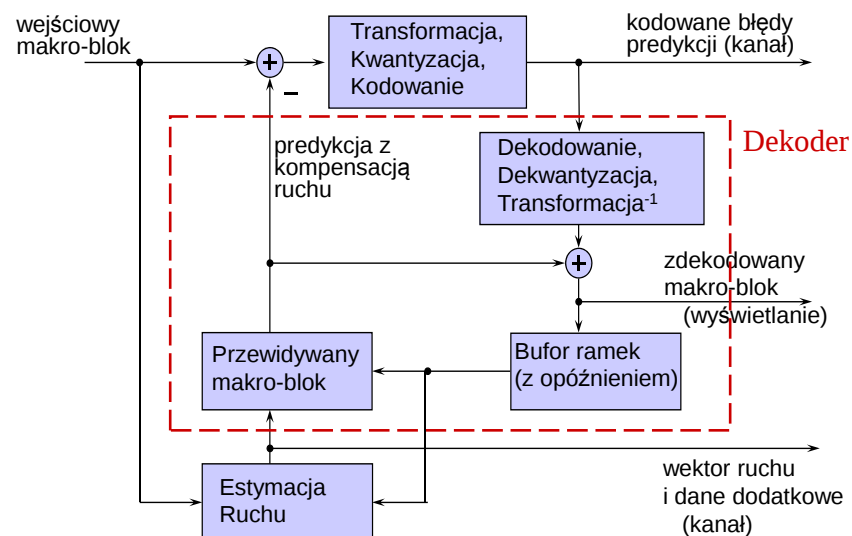
KODEKI WIDEO

Estymacja ruchu

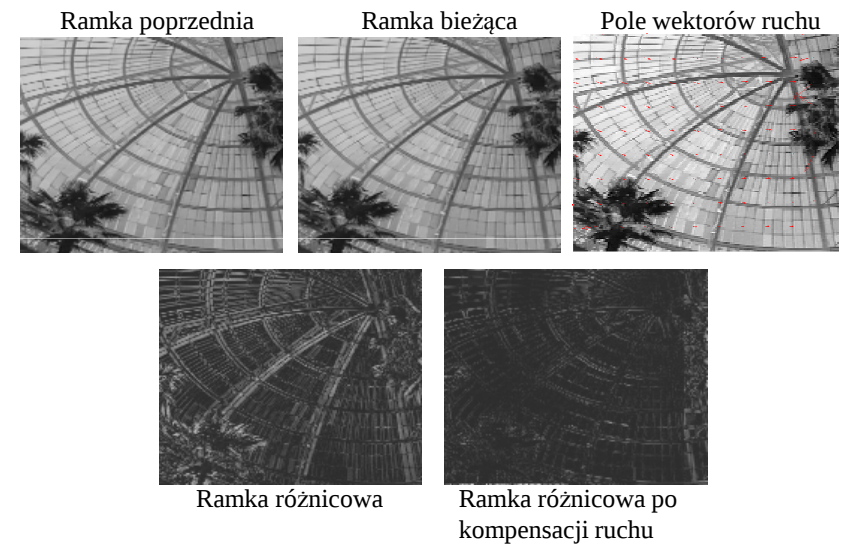
- Po wierszach



Hybrydowa metoda kompresji, blokowa (DCT) z kompensacją ruchu



Przykład 1



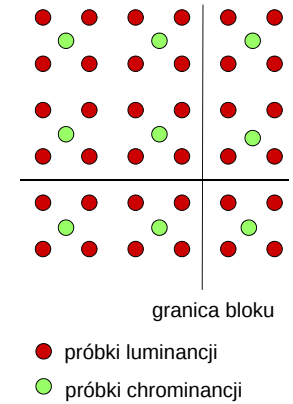
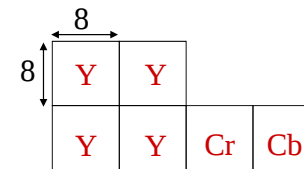
Przykład 2



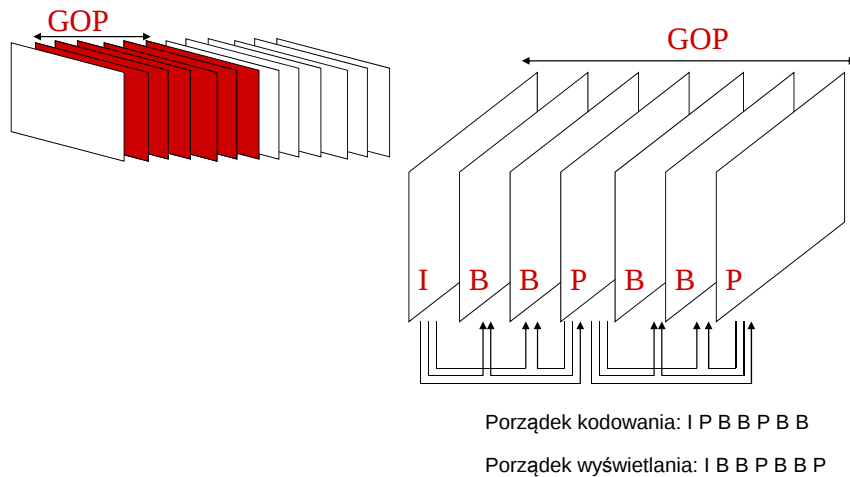
Format ramek

■ CIF

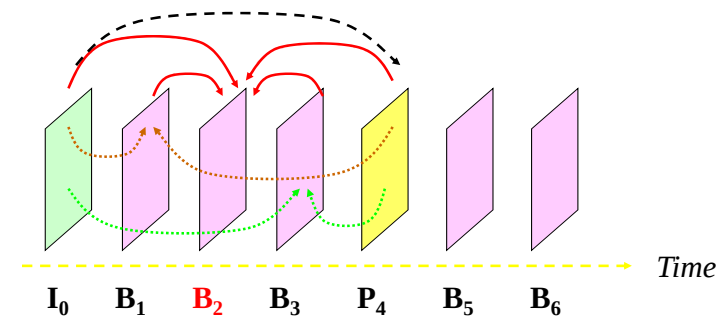
- Common Intermediate Format
- 4:2:0 Y Cr Cb format, progresja (bez przeplotu)
- 352 x 288 dla ramek luminancji, 176 x 144 dla ramek chrominancji



Predykcja INTER (międzyobrazowa)

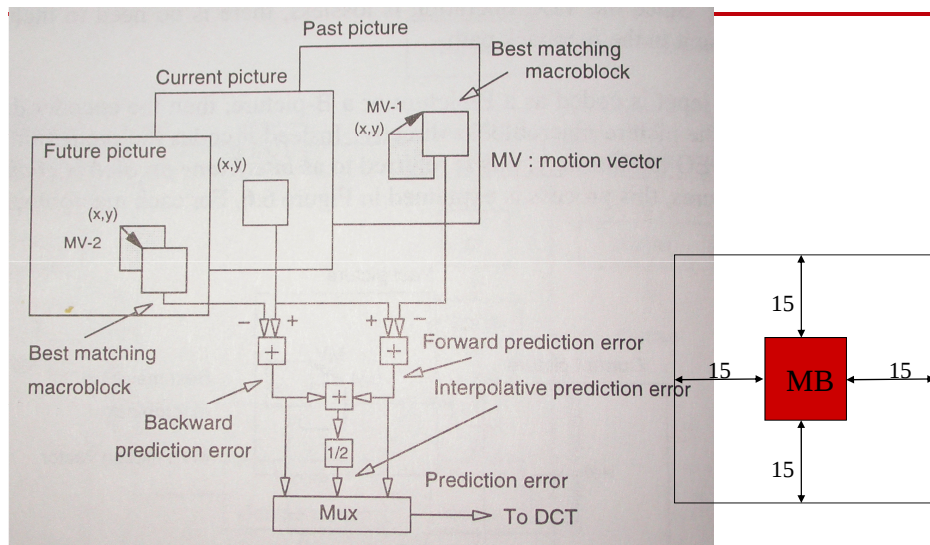


Referencyjne ramki B



- Playback order: I_0 B_1 B_2 B_3 P_4 B_5 B_6
- Bitstream order: I_0 P_4 B_1 B_3 B_2 P_8 B_5

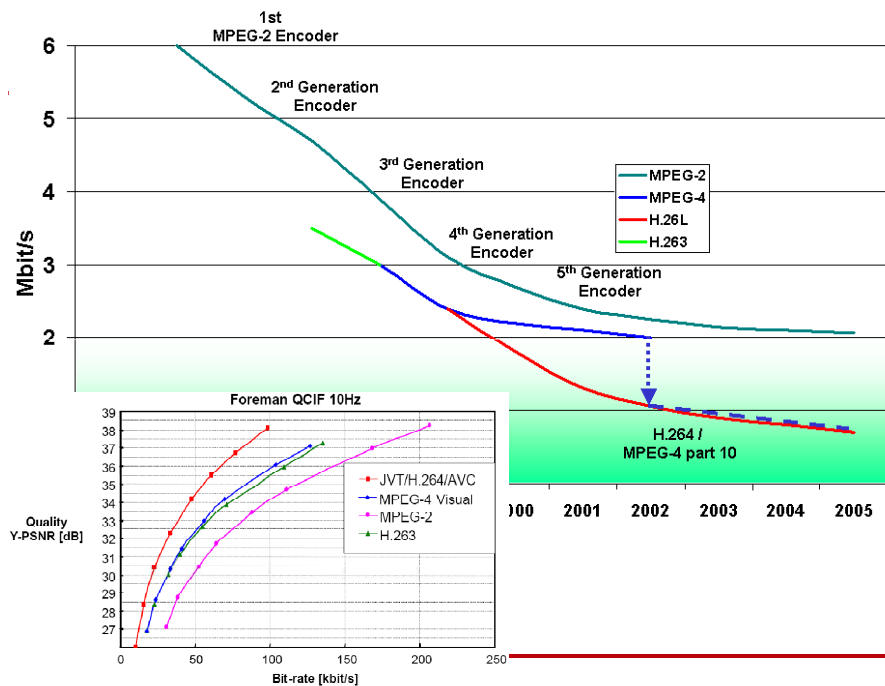
Estymacja ruchu



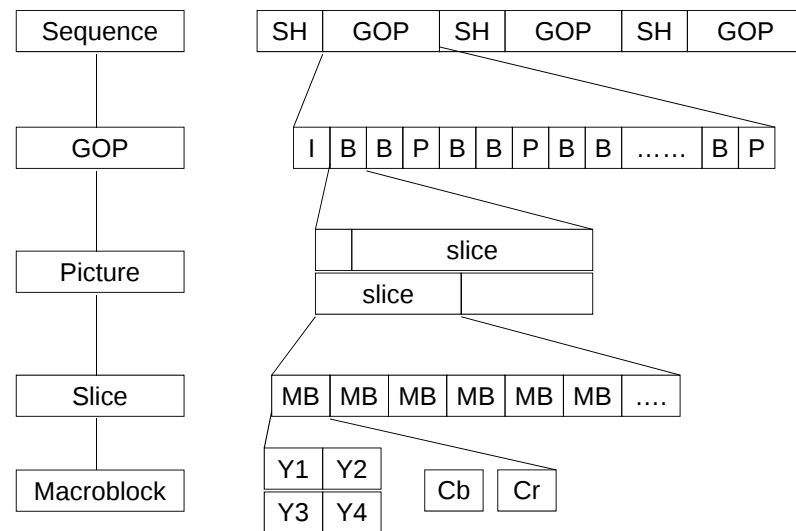
Makroblok 16x16

Standardy rodziny MPEG i pokrewne

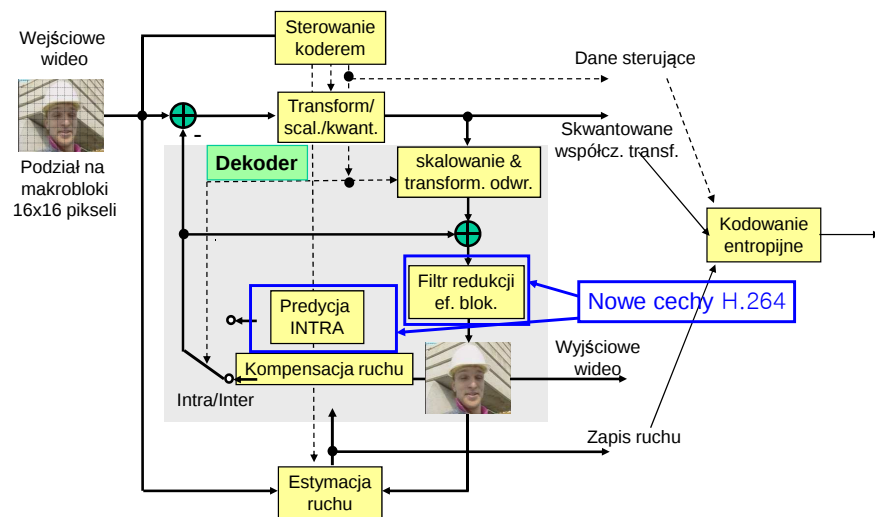
ITU-T Standard	H.261 (Version 1) H.261 (Version 2) H.263 H.263+ H.263++								
Joint ITU-T/MPEG Standards	H.262/MPEG-2 H.264/MPEG-4 AVC								
MPEG Standard	MPEG-1 MPEG-4 (Version 1) MPEG-4 (Version 2)								
	1988	1990	1992	1994	1996	1998	2000	2002	2004



Struktura danych w MPEG

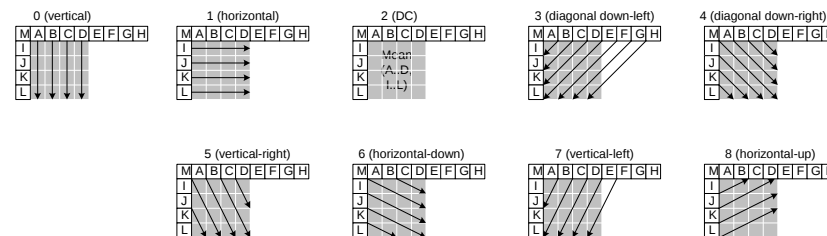


Struktura kodeka h.264

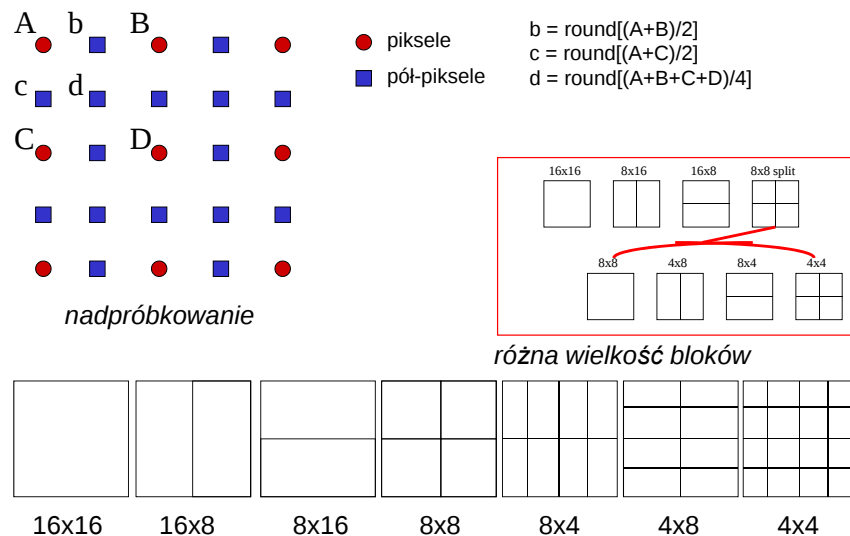


Predycja INTRA

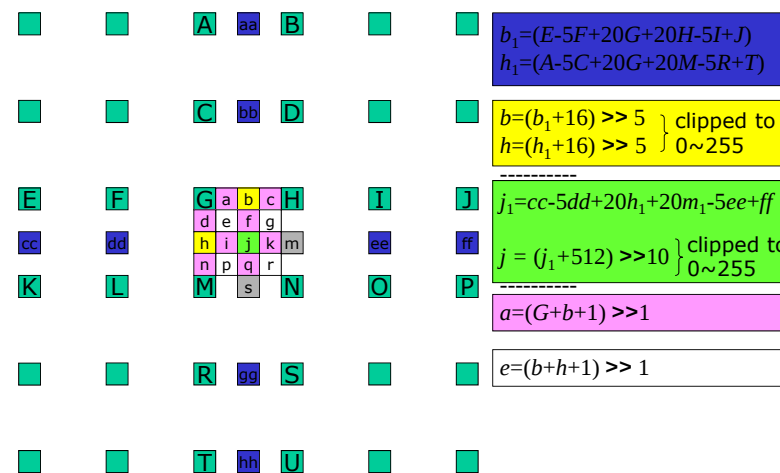
- 4 tryby luminancji 16x16
- 9 trybów luminancji 4x4
- 4 tryby chrominancji 8x8
- Znacząca złożoność kodowania



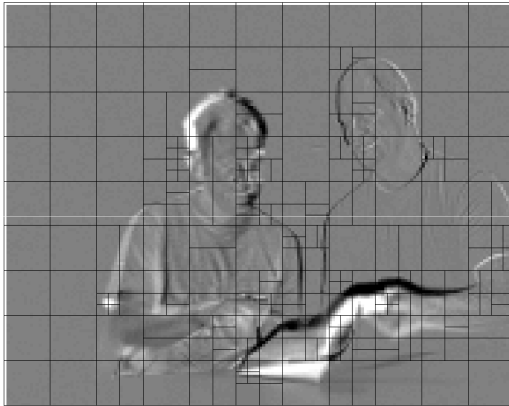
Estymacja ruchu z nadpróbkowaniem



Dokładność kompensacji ruchu



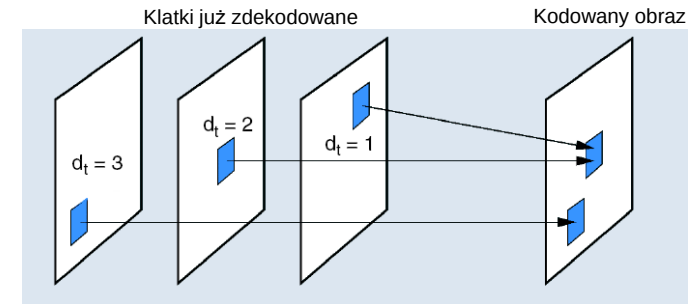
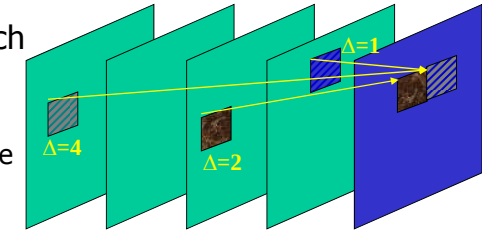
Przykład podziału na makrobloki



Kompensacja ruchu

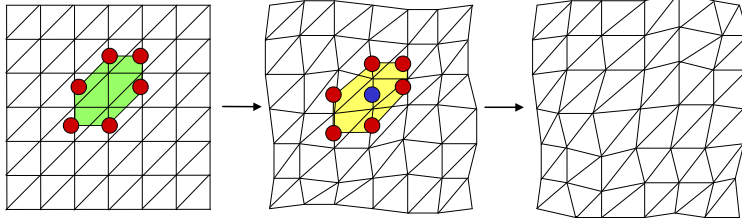
Wiele obrazów referencyjnych

- Dowolne wagi
- Kierunek ruchu nieistotny
- Klatki B mogą być referencyjne



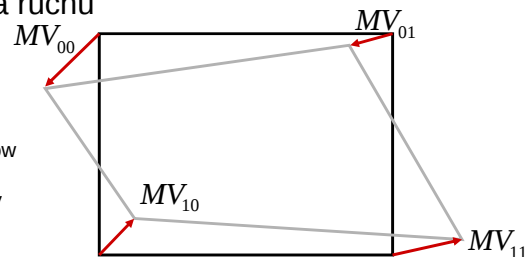
Estymacja ruchu

Nieregularne siatki do modelowania obiektów w ruchu



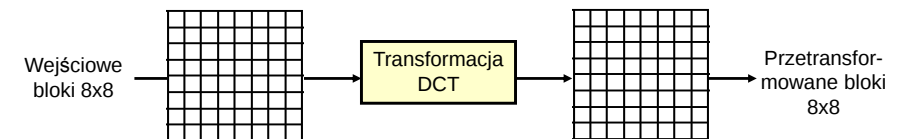
Globalna estymacja ruchu

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

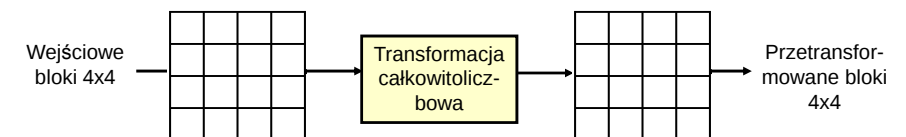


Nowy rodzaj transformacji

MPEG-2 / MPEG-4



MPEG-4 AVC



Transformacje w AVC

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

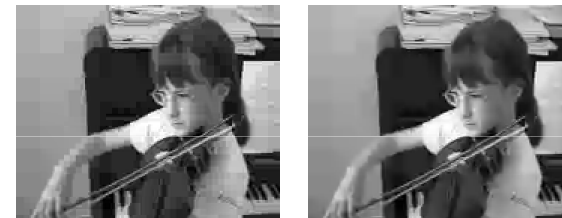
Rozszerzona na 8x8 dla Chroma za pomocą transformacji Hadamarda 2x2 współczynników DC
Rozszerzona na 16x16 dla Luma Intra16x16 za pomocą transformacji Hadamarda 4x4 współczynników DC

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

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

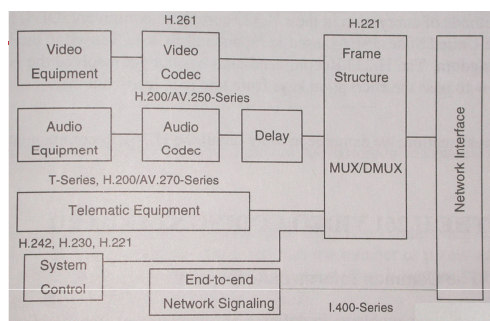
Adaptacyjny filtr redukcji efektów blokowych

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



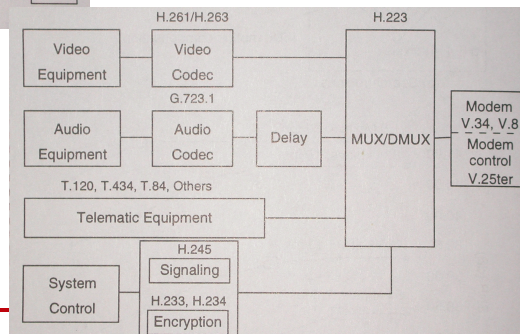
Bez filtru

Z filtrem



H.320

Standardy multimedialne



H.324