METODY PREDYKCJI

KODOWANIE DANYCH, A.Przelaskowski

- Koncepcja metod predykcyjnych
- DPCM
- Konteksty
- Metody skanowania
- Kodery obrazów
- Testy

Koncepcja

- Modelowanie: analiza funkcjonalna i teoria aproksymacji
- Funkcja przewidywania

$$\hat{x}_i = \psi(x_{i-1}, x_{i-2}, \ldots, x_1) = \psi(x_{i-1}, x_{i-2}, \ldots, x_{i-m})$$

- Kodowanie błędu predykcji

$$\epsilon_i = x_i - \hat{x}_i$$

- Model liniowy

$$\hat{x}_i = \psi(x_{i-1}, x_{i-2}, \ldots, x_{i-m}) = \sum_{k=1}^{m} \alpha_k x_{i-k}$$

- Model nieliniowy

$$\hat{x}_i = \sum_{k=1}^{m} \left(\alpha_k^{(1)} x_{i-k} + \alpha_k^{(2)} x_{i-k}^2 + \cdots + \alpha_k^{(n)} x_{i-k}^n\right)$$

DPCM (differential pulse code modulation)

- Postać modelu predykcji liniowej (kontekst, wagi)
- Adaptacja (model lokalny, globalny)
- Dobór wag: przykładowo metoda największych kwadratów

$$\varepsilon_p = \sum_{i=1}^{t} \epsilon_i^2 = \sum_{i=1}^{t} (x_i - \hat{x}_i)^2 = \sum_{i=1}^{t} \left(x_i - \sum_{k=1}^{m} \alpha_k x_{i-k}\right)^2$$

zerujemy pochodne cząstkowe

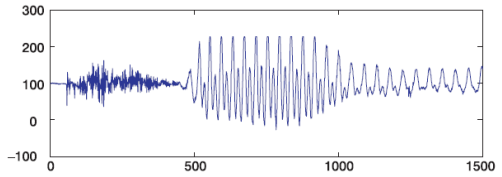
$$\frac{\partial \varepsilon_p}{\partial \alpha_1} = -2 \sum_{i=1}^{t} \left(x_{i-1}\left(x_i - \sum_{k=1}^{m} \alpha_k x_{i-k}\right)\right) = 0$$
$$\ldots$$
$$\ldots$$
$$\ldots$$
$$\frac{\partial \varepsilon_p}{\partial \alpha_m} = -2 \sum_{i=1}^{t} \left(x_{i-m}\left(x_i - \sum_{k=1}^{m} \alpha_k x_{i-k}\right)\right) = 0$$

porządkujemy

$$\sum_{i=1}^{t} \left(x_{i-1} \cdot \sum_{k=1}^{m} \alpha_k x_{i-k}\right) = \sum_{i=1}^{t} x_i \cdot x_{i-1}$$
$$\ldots$$
$$\ldots$$
$$\ldots$$
$$\sum_{i=1}^{t} \left(x_{i-m} \cdot \sum_{k=1}^{m} \alpha_k x_{i-k}\right) = \sum_{i=1}^{t} x_i \cdot x_{i-m}$$

Przykład – upraszczanie rozkładu

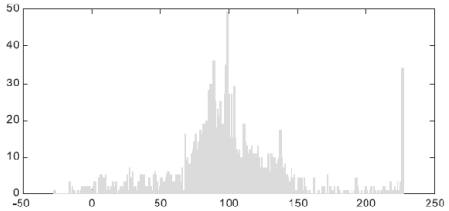
- Sygnał



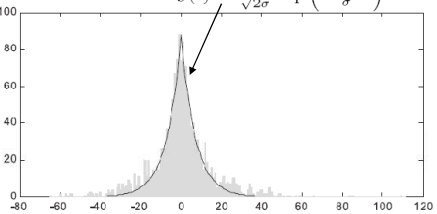
- Histogramy

aproksymacja rozkładem Laplace'a

$$h_\sigma(\epsilon) = \frac{1}{\sqrt{2}\sigma} \exp\left(\frac{-\sqrt{2}|\epsilon|}{\sigma}\right)$$



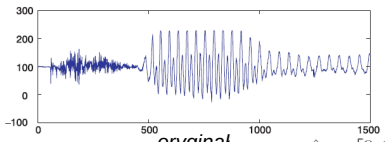
h.sygnału



h.błędu predykcji

$$\hat{x}_i = \lceil 0,89x_{i-1} + 0,63x_{i-2} - 0,54x_{i-3} \rfloor$$

Przykład – zwiększanie rzędu modelu



oryginał

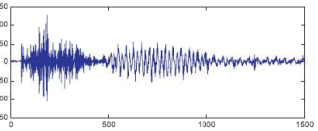
$\sigma^2_{we} = 1930,4$

BR=8bps

a) BR=6,66bps

$\hat{x}_i = \lceil 0,98x_{i-1} \rfloor$

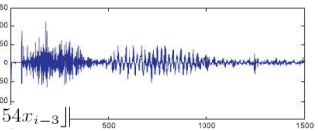
$\sigma^2_{\varepsilon_{(m=1)}} = 398$



b) BR=6,28bps

$\sigma^2_{\varepsilon_{(3)}} = 256,4$

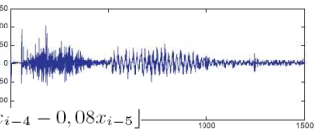
$\hat{x}_i = \lceil 0,88x_{i-1} + 0,63x_{i-2} - 0,54x_{i-3} \rfloor$



c) BR=6,23bps

$\sigma^2_{\varepsilon_{(5)}} = 234,4$

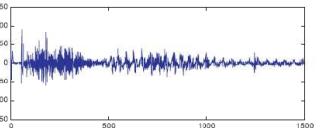
$\hat{x}_i = \lceil 0,74x_{i-1} + 0,75x_{i-2} - 0,27x_{i-3} - 0,17x_{i-4} - 0,08x_{i-5} \rfloor$



d) BR=5,96bps

m=10

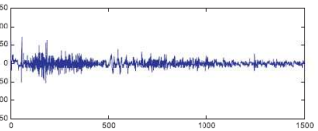
$\sigma^2_{\varepsilon_{(10)}} = 169,4$



e) BR=5,81bps

m=150

$\sigma^2_{\varepsilon_{(150)}} = 106,3$



Optymalizacja modelu predykcji

- Proste wskazówki:
  - nieobciążanie estymatora

$$\mathrm{E}\{\epsilon_i\} = \mathrm{E}\{x_i - \hat{x}_i\} = \mathrm{E}\{x_i - \sum_{k=1}^{m} \alpha_k \cdot x_{i-k}\} = \mathrm{E}\{x_i\} - \sum_{k-1}^{m} \alpha_k \cdot \mathrm{E}\{x_{i-k}\}$$

przy założeniu $\mathrm{E}\{x_i\} = \mathrm{E}\{x_{i-1}\} = \mathrm{E}\{x\}$ mamy $\mathrm{E}\{x\}(1 - \sum_{k=1}^{m} \alpha_k) = 0$

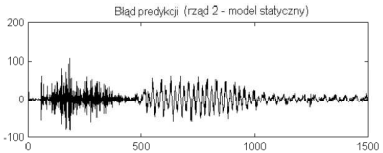
co daje warunki

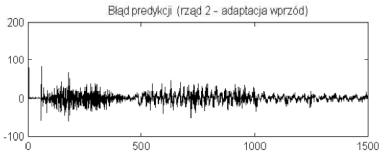
$$\mathrm{E}\{x\} = 0 \quad \text{lub} \quad \sum_{k=1}^{m} \alpha_k = 1$$

- ZAOKRĄGLANIE BŁĘDU!
- Metoda minimalizacji błędu nie uwzględnia entropii !?!
- Lepsze rozwiązania: przełączanie małych, typowych modeli liniowych, interpolacja, ‘drobna nieliniowość’

Adaptacja

- Dobór α_k metodą średniokwadratową zakłada stacjonarność kodowanego źródła
- Podział na bloki formowane dynamicznie
- Adaptacja wprzód





	Statyczny model predykcji		Adaptacyjny model predykcji (w przód)	
Rząd modelu predykcji	Wariancja błędu	Średnia bitowa	Wariancja błędu	Średnia bitowa
1	398,0	6,66	394,5	6,66
2	379,7	6,71	166,4	5,98
3	256,4	6,28	150,8	5,87
5	234,4	6,23	140,9	5,82
10	169,4	5,96	122,9	5,70
20	162,7	5,94	118,0	5,66

Adaptacja wstecz

- Błąd (model rzędu 1)

$$\varepsilon_i^2 = (x_i - \alpha_1 \cdot x_{i-1})^2$$

- Pochodna

$$\frac{\partial \varepsilon_i}{\partial \alpha_1} = -2(x_i - \alpha_1 \cdot x_{i-1}) \cdot x_{i-1}$$

- Przyczynowa modyfikacja współczynników

$$\alpha_1^{(i+1)} = \alpha_1^{(i)} - \beta' \frac{\partial \varepsilon_i}{\partial \alpha_1}$$

$$\alpha_1^{(i+1)} = \alpha_1^{(i)} + 2\beta' \cdot (x_i - \alpha_1 \cdot x_{i-1}) \cdot x_{i-1} = \alpha_1^{(i)} + \beta \cdot \epsilon_i \cdot x_{i-1}$$

- Ogólniej, ze zróżnicowaną szybkością adaptacji

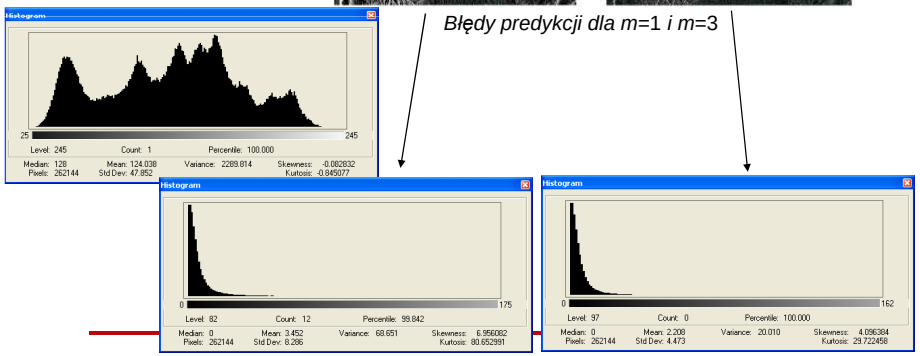
$$\alpha_k^{(i+1)} = \alpha_k^{(i)} + \beta_k \cdot \epsilon_i \cdot x_{i-k} \quad \text{dla} \quad k = 1, \ldots, m$$

- Ograniczona efektywność takich rozwiązań

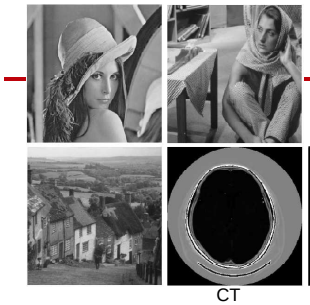
Przykład z obrazem

oryginał

Błędy predykcji dla m=1 i m=3



Rezultaty predykcji



CT CT'

kodowanie arytmetyczne rzędu 1

Porządkowanie pikseli	Lena	Barbara	Goldhill	CT
Po wierszach	5,28	6,40	5,50	1,98

	Obrazy testowe					
Metoda	Lena	Barbara	Goldhill	CT	CT'	Średnio
Kontekst $m = 3$	4,55	5,28	4,98	2,00	1,51	*3,66*
Kontekst $m = 4$	4,48	5,23	4,95	2,02	1,48	*3,63*
Kontekst $m = 5$	4,48	5,17	4,95	2,12	1,51	*3,65*
Kontekst $m = 8$	4,49	5,15	4,88	2,14	1,53	*3,64*
Kontekst $m = 12$	4,50	5,15	4,88	2,14	1,53	*3,64*

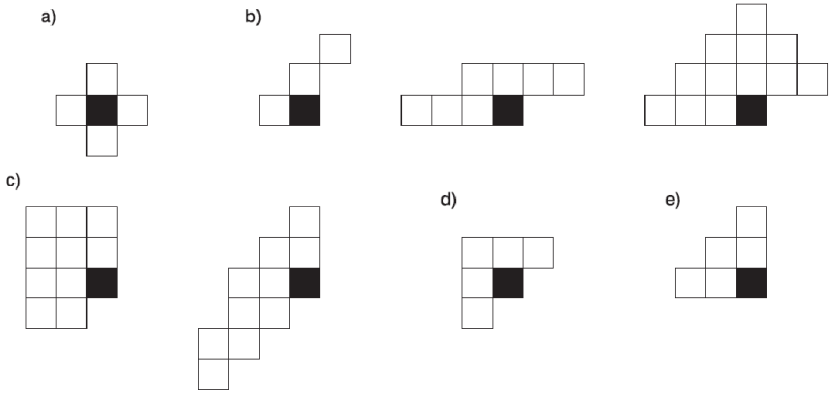
Użyte modele predykcji rzędu 3

f_{gl}	f_g
f_w	f

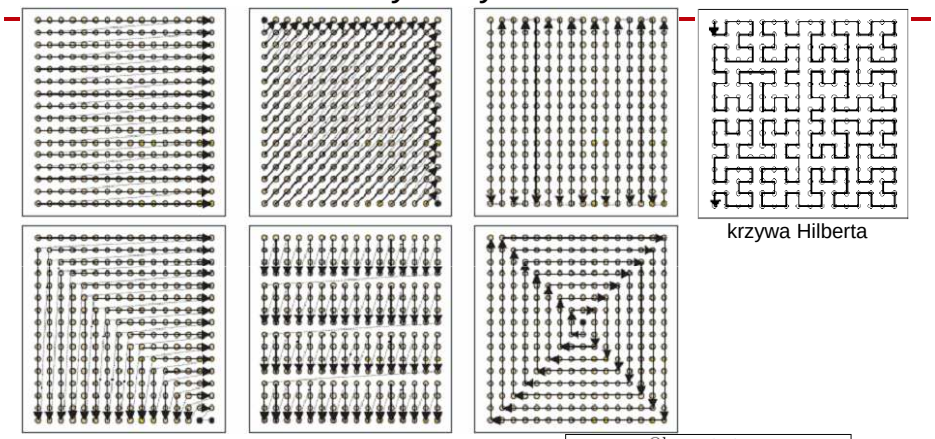
współczynniki wyznaczono metodą średniokwadratową

Obraz testowy	Lena			Barbara			Goldhill		
Współczynniki modelu rzędu 3	α_w	α_{gl}	α_g	α_w	α_{gl}	α_g	α_w	α_{gl}	α_g
	0,7	−0,5	0,8	0,25	−0,06	0,8	0,81	−0,53	0,72
Obraz testowy	CT			CT'					
Współczynniki modelu rzędu 3	α_w	α_{gl}	α_g	α_w	α_{gl}	α_g			
	0,53	−0,11	0,57	0,79	−0,65	0,87			

Przykładowe konteksty obrazowe



Alternatywa: różne metody skanowania + kodowanie arytmetyczne



krzywa Hilberta

predykcja optymalna

4,48	5,15	4,88	1,48

	Obrazy testowe				
Porządkowanie pikseli	Lena	Barbara	Goldhill	CT	*średnio*
Po wierszach	5,28	6,40	5,50	1,98	*4,79*
Po kolumnach	4,87	5,86	5,45	1,90	*4,52*
Naprzemienne	5,03	6,22	5,42	1,98	*4,66*
Spiralne	4,97	6,04	5,38	1,77	*4,54*
Po krzywej Hilberta	5,06	6,17	5,47	1,93	*4,66*
Według operatora HD	4,79	5,79	5,31	1,89	*4,45*

Alternatywa: ‘predykcja stochastyczna’

- Kwantyzacja kontekstu kodera arytmetycznego rzędu *m*

$$\check{c} = \left\lfloor \frac{\sum_{p,q\in\Omega} \alpha(p,q) f(p,q)}{\Delta} \right\rfloor$$

- Takie same konteksty jak w predykcji funkcjonalnej

	Obrazy testowe			
Rozmiar sąsiedztwa kwantowanego kontekstu	Lena ($\Delta = 4$)	Barbara ($\Delta = 6$)	Goldhill ($\Delta = 5$)	CT' ($\Delta = 1$)
$m_\Omega = 3$	4,61	5,56	4,98	1,49
$m_\Omega = 4$	4,56	5,51	4,95	1,48
$m_\Omega = 5$	4,55	5,39	4,95	1,41
$m_\Omega = 8$	4,53	5,37	4,94	1,36

przypomnienie – predykcja i skanowanie

Skanowanie optymalne	4,87	5,86	5,38	1,84
Kontekst $m = 3$	4,55	5,28	4,98	1,51
Kontekst $m = 4$	4,48	5,23	4,95	1,48
Kontekst $m = 5$	4,48	5,17	4,95	1,51
Kontekst $m = 8$	4,49	5,15	4,88	1,53
Kontekst $m = 12$	4,50	5,15	4,88	1,53

Przykłady: standard JPEG (kompresja bezstratna)

f_{gl}	f_g
f_w	f

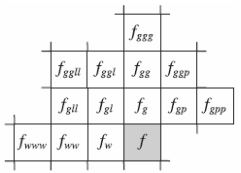
Numer	Funkcja przewidywania
0	$\hat{f} = 0$
1	$\hat{f} = f_w$
2	$\hat{f} = f_g$
3	$\hat{f} = f_{gl}$
4	$\hat{f} = f_w + f_g - f_{gl}$
5	$\hat{f} = f_w + \lfloor (f_g - f_{gl})/2 \rfloor$
6	$\hat{f} = f_g + \lfloor (f_w - f_{gl})/2 \rfloor$
7	$\hat{f} = \lfloor (f_w + f_g)/2 \rfloor$

Standard PNG

f_{gl}	f_g
f_w	f

Numer	Funkcja przewidywania
1	$\hat{f} = 0$
2	$\hat{f} = f_w$
3	$\hat{f} = f_g$
4	$\hat{f} = \lfloor (f_w + f_g)/2 \rfloor$
5	$\hat{f} = \begin{cases} f_w, & p_w = \min(p_w, p_g, p_{gl}) \\ f_g, & p_g = \min(p_w, p_g, p_{gl}) \\ f_{gl}, & \text{wpp} \end{cases}$ gdzie $p_w = \lvert f_g - f_{gl} \rvert, p_g = \lvert f_w - f_{gl} \rvert, p_{gl} = \lvert f_w + f_g - 2f_{gl} \rvert$

Predyktory MED/MAP i z gradientem GAP



$$\hat{f} = \begin{cases} \min(f_w, f_g), & f_{gl} \geq \max(f_w, f_g) \\ \max(f_w, f_g), & f_{gl} \leq \min(f_w, f_g) \\ f_w + f_g - f_{gl}, & \text{wpp} \end{cases}$$

MED/MAP

$$\nabla_h = |f_w - f_{ww}| + |f_g - f_{gl}| + |f_g - f_{gp}|$$
$$\nabla_v = |f_w - f_{gl}| + |f_g - f_{gg}| + |f_{gp} - f_{ggp}|$$

$$\hat{f} = \begin{cases} f_{\boldsymbol{w}}, & \nabla_v - \nabla_h > 80 \text{ (ostra krawędź pozioma)} \\ \lfloor (f_w + \tilde{f})/2 \rfloor, & 80 \geq \nabla_v - \nabla_h > 32 \text{ (krawędź pozioma)} \\ \lfloor (f_w + 3\tilde{f})/4 \rfloor, & 32 \geq \nabla_v - \nabla_h > 8 \text{ (słaba krawędź pozioma)} \\ f_g, & \nabla_v - \nabla_h < -80 \text{ (ostra krawędź pionowa)} \\ \lfloor (f_g + \tilde{f})/2 \rfloor, & -80 \leq \nabla_v - \nabla_h < -32 \text{ (krawędź pionowa)} \\ \lfloor (f_g + 3\tilde{f})/4 \rfloor, & -32 \leq \nabla_v - \nabla_h < -8 \text{ (słaba krawędź pionowa)} \\ \tilde{f}, & \text{wpp} \end{cases}$$

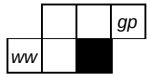
$$\tilde{f} = \lfloor (f_w + f_g)/2 \rfloor + \lfloor (f_{gp} - f_{gl})/4 \rfloor$$

GAP

Predyktory Grahama, DARC, ALCM

Grahama

$$\hat{f} = \begin{cases} f_g, & |f_{gl} - f_g| > |f_{gl} - f_w| \\ f_w, & \text{wpp} \end{cases}$$



ALCM

adaptacja wstecz, początkowe wagi są jednakowe

- jeśli $\hat{f} < f$, waga piksela o największej wartości funkcji jasności w kontekście modelu predykcji jest zwiększana o 1/256, a waga sąsiedniego piksela o najmniejszej wartości - zmniejszana
- jeśli $\hat{f} > f$, waga piksela o największej wartości funkcji jasności w kontekście modelu predykcji jest zmniejszana o 1/256, a waga sąsiedniego piksela o najmniejszej wartości - zwiększana

priorytet: f_{gl}, f_g, f_{gp}, f_w, f_{ww}

DARC

$$\hat{f} = \lceil \alpha f_{gl} + (1 - \alpha) f_g \rfloor, \quad \alpha = \frac{\nabla_v}{\nabla_h + \nabla_v}$$
$$\nabla_h = |f_{gl} - f_g| \text{ oraz } \nabla_v = |f_{gl} - f_w|$$

Przeplot, interpolacja

1	1	1	1	1	1	1	1
4	4	4	4	4	4	4	4
3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4
2	2	2	2	2	2	2	2
4	4	4	4	4	4	4	4
3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4

GIF

1	6	4	6	2	6	4	6
7	7	7	7	7	7	7	7
5	6	5	6	5	6	5	6
7	7	7	7	7	7	7	7
3	6	4	6	3	6	4	6
7	7	7	7	7	7	7	7
5	6	5	6	5	6	5	6
7	7	7	7	7	7	7	7

PNG

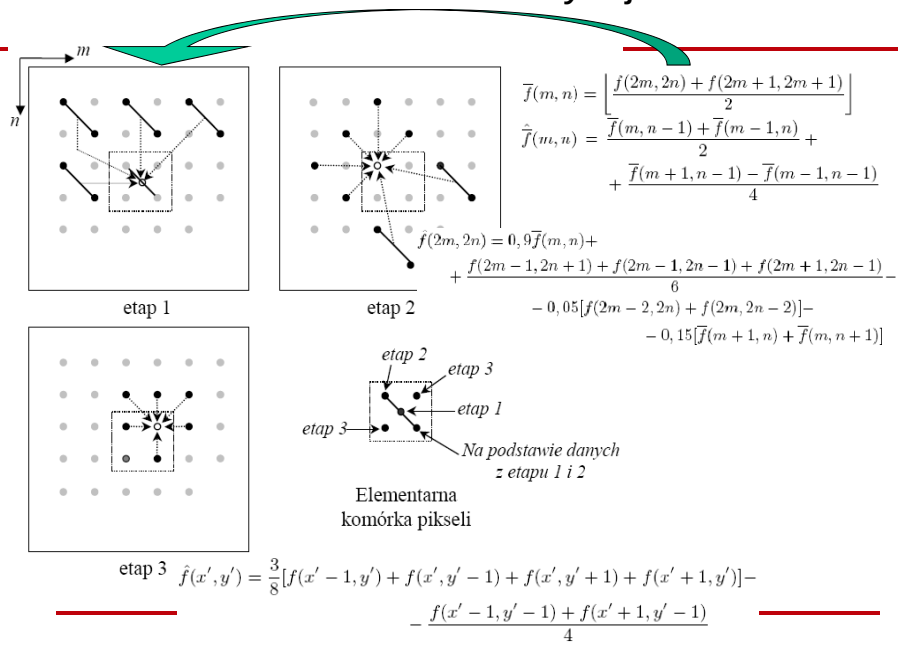
HINT

A	$\mathcal{E}$	C	$\mathcal{E}$	A	$\mathcal{E}$	C	$\mathcal{E}$	A
$\mathcal{E}$	D	$\mathcal{E}$	D	$\mathcal{E}$	D	$\mathcal{E}$	D	$\mathcal{E}$
C	$\mathcal{E}$	B	$\mathcal{E}$	C	$\mathcal{E}$	B	$\mathcal{E}$	C
$\mathcal{E}$	D	$\mathcal{E}$	D	$\mathcal{E}$	D	$\mathcal{E}$	D	$\mathcal{E}$
A	$\mathcal{E}$	C	$\mathcal{E}$	A	$\mathcal{E}$	C	$\mathcal{E}$	A
$\mathcal{E}$	D	$\mathcal{E}$	D	$\mathcal{E}$	D	$\mathcal{E}$	D	$\mathcal{E}$
C	$\mathcal{E}$	B	$\mathcal{E}$	C	$\mathcal{E}$	B	$\mathcal{E}$	C
$\mathcal{E}$	D	$\mathcal{E}$	D	$\mathcal{E}$	D	$\mathcal{E}$	D	$\mathcal{E}$
A	$\mathcal{E}$	C	$\mathcal{E}$	A	$\mathcal{E}$	C	$\mathcal{E}$	A

Efekt: progresywne kodowanie

uzupełnianie brakującej informacji

Predykcja w CALIC



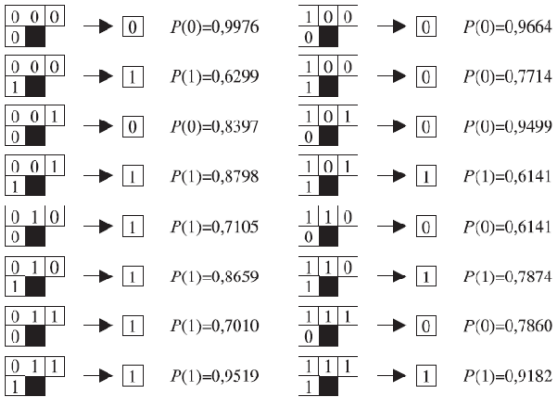
$$\overline{f}(m,n) = \left\lfloor \frac{\hat{f}(2m,2n) + f(2m+1,2m+1)}{2} \right\rfloor$$

$$\hat{\overline{f}}(m,n) = \frac{\overline{f}(m,n-1) + \overline{f}(m-1,n)}{2} + \frac{\overline{f}(m+1,n-1) - \overline{f}(m-1,n-1)}{4}$$

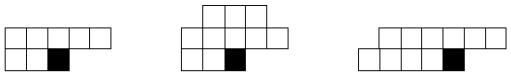
$$\hat{f}(2m,2n) = 0,9\overline{f}(m,n) + \frac{f(2m-1,2n+1) + f(2m-1,2n-1) + f(2m+1,2n-1)}{6} - 0,05[f(2m-2,2n) + f(2m,2n-2)] - 0,15[\overline{f}(m+1,n) + \overline{f}(m,n+1)]$$

etap 3 $\hat{f}(x',y') = \frac{3}{8}[f(x'-1,y') + f(x',y'-1) + f(x',y'+1) + f(x'+1,y')] - \frac{f(x'-1,y'-1) + f(x'+1,y'-1)}{4}$

Predykcja obrazów czarno-białych



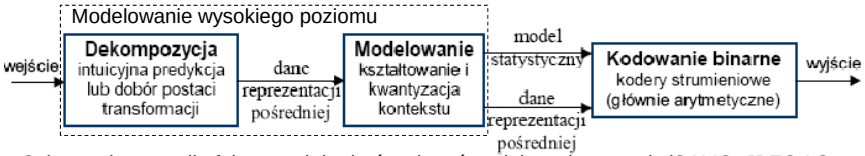
przykładowe modele predykcji



konteksty z JBIG

Mapy bitowe Leny

Paradygmaty bezstratnego kodowania obrazów



Schemat kompresji efektywnych koderów obrazów wielopoziomowych (CALIC, JPEG-LS, falkowe itd.)

Schemat kompresji binarnych koderów obrazów wielopoziomowych (np. JBIG)

Archiwizery (modelowanie statystyczne)

Testy

Obrazy	Metody bezstratnej kompresji obrazów							
	CALIC	JPEG-LS	APT	SPIHT	JPEG 2000	JBIG	JB2	BAC
131 mammogramów, 1000x1000-4500x4500 pikseli, 12-14 bpp	6,64	6,68	7,14	6,69	7,86	7,09	7,31	6,78
234 CT, 256x256-756x749 pikseli, 8-12 bpp	3,91	4,02	4,76	4,25	4,31	4,67	5,04	4,73
245 MR, 256x256-1024x1024 pikseli, 8-12 bpp	7,11	7,24	7,84	7,21	7,41	7,73	8,27	7,61
126 USG, 128x128-640x480 pikseli, 8 bpp	2,78	3,10	3,17	3,47	3,41	3,14	3,40	2,39
149 radiogramy (CR,DR,skanowane), 187x475-4000x5200 pikseli, 8-15 bpp	5,91	6,07	3,08	5,75	5,00	3,72	6,93	3,14
105 NM[2] (PET, SPECT, scyntygrafia), 64x64-256x1024 pikseli, 5-12 bpp	2,90	2,89	3,69	–	3,55	3,29	4,15	3,17
Średnio (990 obrazów)	*5,12*	*5,24*	*5,27*	–	*5,46*	*5,27*	*6,12*	*5,01*
Średnio (bez NM)	*5,38*	*5,52*	*5,46*	*5,58*	*5,69*	*5,50*	*6,36*	*5,24*

Testy kompresji 29 obrazów naturalnych:
4,55 bpp (JPEG-LS), 4,60 bpp (JPEG2000), 4,75 bpp (JBIG), 4,35 bpp (CALIC), 4,36 (BAC) oraz 4,1 bpp (WinRK v. 2.1.6)