

Andrzej KRUPICZKA

Instytut Informatyki Teoretycznej i Stosowanej PAN

ul. Bałtycka 5, 44-100 GLIWICE

## Transformacja kosinusowa o zmiennym rozmiarze bloku do kompresji obrazów cyfrowych

**Streszczenie.** W artykule przedstawiono algorytm kompresji obrazów monochromatycznych przy użyciu transformacji kosinusowej o zmiennym rozmiarze bloku. Bloki o dużych rozmiarach tworzą reprezentację obszarów obrazu o małej liczbie szczegółów, natomiast bloki o małych rozmiarach tworzą reprezentację obszarów obrazu o dużej liczbie szczegółów. Obrazy zakodowane w ten sposób charakteryzują się znaczną redukcją „efektu blokowego” i „efektu dzwonienia”. Umożliwia to osiągnięcie większej wartości kompresji obrazów i poprawę ich jakości w porównaniu z obrazami zakodowanymi z użyciem bloków o stałym rozmiarze.

### TRANSFORMÉE À COSINUS À TAILLE DE BLOC VARIABLE POUR LA COMPRESSION DES IMAGES DIGITALES

**Résumé.** L'article présente une méthode de compression d'images monochromes à la base de la transformée en cosinus avec la taille de bloc variable. Les blocs larges représentent les régions de quantité faible de détails; les blocs de taille moins grande représentent les régions d'activité plus forte. Les images codées selon cette méthode présentent une réduction de "l'effet de blocs" et de "l'effet d'oscillations". Les résultats sont caractérisés par un rapport de compression plus élevé et une qualité meilleure par rapport à ceux obtenus avec un bloc de taille fixe.

### VARIABLE BLOCK SIZE COSINE TRANSFORM FOR DIGITAL IMAGE COMPRESSION

**Summary.** A monochromatic images compression scheme based on variable block size cosine transform is demonstrated. Large blocks represent low details image areas whereas small blocks represent high details image areas. Images coded in this way show reduction of "blocking effect" and "ringing effect". Results present higher compression ratio and image quality compared to results obtained using a fixed block size.

## 1. Wprowadzenie

Dane w postaci obrazu cyfrowego powstają, gdy:

- źródło obrazu ma charakter cyfrowy;
- przetwarzane są obrazy analogowe z użyciem technik cyfrowych;
- przesyłane są obrazy za pośrednictwem cyfrowych łączy transmisyjnych, np. ISDN;
- magazynowane są obrazy przy użyciu pamięci cyfrowych, np. CD-ROM.

Źródłem obrazów jest telewizja, termowizja, technika satelitarna, technika medyczna oraz inne, ale bez względu na swoje pochodzenie obrazy zawierają dużą ilość informacji.

Obrazy monochromatyczne (czarno-białe) mają jeden kanał danych, podczas gdy kolorowe trzy kanały, z których każdy zawiera informację jednego ze składników systemu współrzędnych kolorowych, którymi mogą być trzy kolory podstawowe RGB lub sygnały luminancji i chrominancji YIQ (YUV). Obrazy satelitarne zawierają jeszcze więcej kanałów, np. LANDSAT tworzy obrazy w siedmiu pasmach częstotliwościowych. Jeśli próbkujemy obraz z częstotliwością odpowiadającą jakości konwencjonalnego obrazu telewizyjnego, to pojedynczy kadr obrazu może zawierać  $512 * 512 * 8$  bitów, czyli około 2 megabitów na kanał. Dla obrazu kolorowego o częstotliwości kadrów 25Hz, takiej jak w obrazie telewizyjnym, przy założeniu że każdy piksel zostanie skwantowany przy użyciu 8 bitów na kanał, przepływ danych wynosi około 160 megabitów na sekundę. Istnieje zatem potrzeba kompresji danych, jeśli mamy szerokopasmowe źródło obrazu i wąskopasmowy kanał transmisyjny, np. sieć komputerową, sieć telefoniczną lub ograniczoną pamięć, np. dyski czy taśmy magnetyczne.

Duży wpływ na rozwój technik kompresji obrazów miało wprowadzenie i upowszechnianie publicznych sieci cyfrowych ISDN (ang. integrated services digital network). Tworzy ona podstawę do upowszechnienia wideotelefonów oraz innych multimedialnych usług interaktywnych typu teletekst czy wideotekst.

Kompresja obrazu przedstawionego w postaci cyfrowej polega na uzyskaniu jego reprezentacji o zmniejszonej objętości poprzez redukcję informacji nadmiarowej, informacji „niepotrzebnej” i szumu zawartego w obrazie. Kompresję można podzielić na wewnątrz-kadrową (ang. intraframe coding), w której usuwa się informację nadmiarową zawartą w danej ramce obrazu i międzykadrową (ang. interframe coding), w której usuwa się informację nadmiarową zawartą między kolejnymi ramkami obrazu [1-5]. Dekompresja polega natomiast na odtwarzaniu, na podstawie posiadanej reprezentacji, obrazu oryginalnego o jakości akceptowanej dla danego zastosowania.

Wśród stosowanych metod duże znaczenie odgrywiają metody transformacyjne. Można tu wymienić następujące transformacje: Karhunen-Loevego, Hadamarda, Walsh, Fouriera, sinusową, a szczególnie kosinusową [1-4]. Transformacja kosinusowa stanowiąc dobre przybliżenie statystyki obrazu daje dużą skuteczność kompresji przy umiarkowanej złożoności obliczeniowej [1-2].

Konsekwencją zakrojonych na szeroką skalę prac z zakresu kompresji obrazów stało się ustalenie międzynarodowych standardów kompresji obrazów na podstawie transformacji kosinusowej [13-16]. Zaowocowało to pojawieniem się specjalizowanych procesorów do

dokonywania w czasie rzeczywistym transformacji kosinusowej dla obrazów, np. układ typu STV 3200 firmy SGS-Thompson.

Prowadzone nadal badania zmierzają do poprawy parametrów kompresji. Przedmiotem mojego dalszego zainteresowania jest kompresja wewnątrzkadrowa naturalnych obrazów monochromatycznych przy użyciu transformacji kosinusowej w celu ich prezentacji.

## 2. Transformacja kosinusowa

Elementem obrazu podlegającym transformacji jest najczęściej blok pikseli obrazu o rozmiarach  $4 * 4$ ,  $8 * 8$  lub  $16 * 16$ . W wyniku transformacji uzyskuje się blok danych o identycznym rozmiarze. Dane te nazywa się współczynnikami.

Poniższe równania określają matematyczną definicję prostej FDCT (ang. forward discrete cosine transform) i odwrotnej IDCT (ang. inverse discrete cosine transform) transformacji kosinusowej dla bloku pikseli obrazu o rozmiarach  $N * N$ :

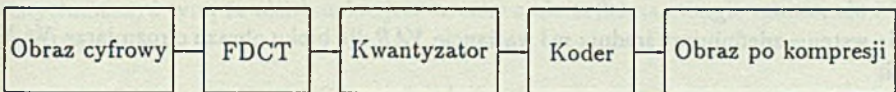
$$F(u, v) = \frac{2}{N} C(u) C(v) \left[ \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) * \cos \frac{(2x+1)u\pi}{2 * N} \cos \frac{(2y+1)v\pi}{2 * N} \right] \quad (1)$$

$$f(x, y) = \frac{2}{N} \left[ \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u) C(v) F(u, v) * \cos \frac{(2x+1)u\pi}{2 * N} \cos \frac{(2y+1)v\pi}{2 * N} \right] \quad (2)$$

gdzie:

- $f(x, y)$  - wartość stopnia szarości piksela o współrzędnych  $x, y$ ;
- $F(u, v)$  - wartość współczynnika zmiennoprądowego AC transformacji;
- $F(0, 0)$  - wartość współczynnika stałoprądowego DC transformacji;
- $C(w)$  -  $C(w)=1/\sqrt{2}$  dla  $w=0$  oraz  $C(w)=1$  dla pozostałych  $w$ .

Schemat blokowy układu kompresji na podstawie prostej dyskretnej transformacji kosinusowej FDCT przedstawiono poniżej:

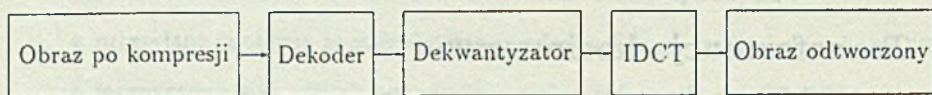


Istota zastosowania transformacji kosinusowej do kompresji obrazu polega na tym, że informacja zawarta w oryginalnym bloku pikseli swoją najistotniejszą część zawiera w małej liczbie współczynników.

Proces kompresji obrazu składa się z trzech zasadniczych etapów. W pierwszym etapie kompresji kolejno, dla każdego bloku obrazu, oblicza się za pomocą prostej dyskretnej transformacji kosinusowej, odpowiadające mu wartości współczynników. Następnie dokonuje się ich selekcji i kwantyzacji metodami adaptacyjnymi opisanymi w punkcie 4 artykułu. Selekcja polega na pominięciu tych współczynników, których wartość jest poniżej ustalonego progu. Kwantyzacji dokonuje się stosując stałą lub zmienną szerokość przedziału kwantyzacji. Ostatnim etapem kompresji jest zakodowanie skwantowanych

współczynników, przy użyciu kodu o stałej lub zmiennej długości, np. kodem Huffmana. Zakodowane współczynniki można przesłać łączem transmisyjnym lub zapisać do pamięci.

Schemat blokowy układu dekompresji na podstawie odwrotnej dyskretnej transformacji kosinusowej IDCT przedstawiono poniżej:



Proces dekompresji obrazu składa się również z trzech etapów. W pierwszym etapie dekompresji odebrane łączem transmisyjnym lub odczytane z pamięci informacje są dekodowane. Następnie dokonuje się ich dekwantyzacji, czyli odtwarza się prawdziwe wartości współczynników transformacji. Ostatnim etapem dekompresji jest obliczenie, na podstawie wartości współczynników, kolejno dla każdego bloku obrazu wartości wszystkich pikseli obrazu w danym bloku za pomocą odwrotnej dyskretnej transformacji kosinusowej.

### 3. Parametry kompresji obrazu

#### 3.1. Wartość kompresji

Wartość kompresji określa skuteczność kompresji podając, ile razy została zmniejszona objętość informacji zawierająca treść obrazu. Najczęściej podaje się ją w jednostkach ilości bitów potrzebnych do zakodowania pojedynczego piksela obrazu. Dla przykładu, jeśli wartość kompresji wynosi 20, a jeden piksel obrazu oryginalnego był zakodowany na 8 bitach, to odpowiadająca wartość kompresji wynosi  $8/20 = 0.4$  bita/piksel.

#### 3.2. Jakość obrazu

Na wstępie zdefiniujmy średnią  $m$  i wariancję  $VAR$  dla bloku obrazu o rozmiarze  $N * N$  pikseli:

$$m = \frac{1}{N * N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \quad (3)$$

$$VAR = \frac{1}{N * N} \left[ \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} [f(x, y) - m]^2 \right] \quad (4)$$

gdzie:

$f(x, y)$  - wartość stopnia szarości piksela o współrzędnych  $x, y$ .

W szczególności blokiem pikseli może być cały obraz o rozdzielczości  $M * M$  pikseli.

Ogólnie akceptowaną obiektywną miarą jakości obrazu jest błąd średniokwadratowy  $MSE$  (ang. mean square error) zdefiniowany dla obrazu o rozdzielczości  $M * M$  pikseli jako

$$MSE = \frac{1}{M * M} \left[ \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} [f(x, y) - f'(x, y)]^2 \right] \quad (5)$$

gdzie:

- $f(x, y)$  - wartość stopnia szarości piksela o wsp.  $x, y$  obrazu oryginalnego,  
 $f'(x, y)$  - wartość stopnia szarości piksela o wsp.  $x, y$  obrazu odtworzonego.

Po znormalizowaniu  $MSE$  uzyskuje się błąd znormalizowany  $NMSE$  (ang. normalized  $MSE$ ) lub procentowe zniekształcenie obrazu  $distortion$

$$NMSE = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{M-1} [f(x, y) - f'(x, y)]^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{M-1} [f(x, y)]^2} \quad (6)$$

$$distortion = NMSE * 100\% \quad (7)$$

Bezpośrednio związanymi z  $NMSE$  i  $MSE$  są logarytmiczne stosunki sygnału do szumu  $SNR$  (ang. signal to noise ratio) i  $PSNR$  (ang. peak to peak signal to noise ratio)

$$SNR = -10 \log(NMSE) \text{ dB} \quad (8)$$

$$PSNR = 10 \log\left(\frac{255^2}{MSE}\right) \text{ dB} \quad (9)$$

Powyższe miary, mimo ich powszechnej akceptacji, mają ograniczoną przydatność do oceny jakości obrazu ze względu na to, że nie uwzględniają mechanizmów subiektywnej oceny jakości obrazu. Te same wartości  $MSE$  dla dwóch obrazów nie gwarantują ich jednakowej subiektywnej jakości. Dla przykładu dodanie do każdego piksela obrazu pewnej stałej wartości, przy założeniu że zakres wartości nie zostanie przekroczony, spowoduje otrzymanie obrazu o tej samej jakości co oryginalny, ale błąd  $MSE$  zostanie powiększony o kwadrat dodanej wartości. W [9] pokazano, że zniekształcenia wynikające z „efektu blokowego” są 10-krotnie bardziej widoczne niż biały szum o tej samej energii. Można to wytłumaczyć tym, że oko ludzkie jest wrażliwe nie tylko na energię szumu, ale i jego strukturę.

#### 4. Przegląd metod adaptacyjnych

Metody adaptacyjne, jak już wspomniano, mają na celu uzależnienie ilości kodowanych współczynników oraz dokładności ich zapamiętania od lokalnych właściwości obrazu, głównie od omówionej poniżej tzw. wartości aktywności przestrzennej (ang. spatial activity). Pojęcie aktywności przestrzennej wewnątrz bloku jest miarą liczby i wartości zmian stopnia szarości w bloku, a tym samym miarą ilości szczegółów w tym bloku. Dla bloku o jednakowej wartości stopnia szarości wszystkich pikseli aktywność przestrzenna wynosi 0, natomiast im więcej szczegółów jest zawartych w danym bloku, tym aktywność przestrzenna ma większą wartość. Aktywność przestrzenna przyjmuje wartość maksymalną, jeśli połowa pikseli bloku przyjmuje wartość maksymalną, a połowa wartość

minimalną. Przegląd metod adaptacyjnych zawiera [10]. Praca [1] wymienia następujące metody adaptacji:

- kodowanie progowe (ang. *threshold coding*). Polega na zapamiętaniu wartości tych współczynników, których wartości są większe od wcześniej określonego progu. Wadą tej metody jest konieczność zapamiętania, które ze współczynników wybrano;
- kodowanie pasmowe (ang. *zonal coding*). Polega na zapamiętaniu współczynników w porządku wzrastającej częstotliwości przestrzennej lub w inny określony sposób oraz oznaczenia końca bloku, gdy pozostałe współczynniki są poniżej określonego progu;
- klasyfikacja bloków (ang. *block classification*). Polega na klasyfikacji każdego bloku do jednej z kilku klas w zależności od wartości aktywności przestrzennej w bloku (tj. ilości szczegółów obrazu) mierzonej sumą kwadratów współczynników transformacji, a następnie zaprojektowaniu dla każdej z klas takiego kodu, że zniekształcenia obrazu są podobne dla wszystkich klas. Szczegóły dotyczące tej metody można znaleźć w [11].

[2] podaje jeszcze inne metody adaptacji:

- zapamiętanie wartości tych współczynników, których suma kwadratów wartości stanowi zadany procent sumy kwadratów wartości wszystkich współczynników;
- za pomocą algorytmu rozpoznawania wzorców (ang. *pattern recognition algorithm*) zaklasyfikować każdy z bloków do jednej z trzech klas w zależności od średniego stopnia szarości bloku i ilości szczegółów obrazu w bloku:
  1. klasa o dużej ilości szczegółów,
  2. klasa o małej ilości szczegółów i blok ciemniejszy od średniego stopnia szarości całego obrazu,
  3. klasa o małej ilości szczegółów i blok jaśniejszy od średniego stopnia szarości całego obrazu.

Dla klasy o dużej liczbie szczegółów zachować dwa razy więcej współczynników niż dla bloków o małej ilości szczegółów. Dla klasy o małej ilości szczegółów i ciemniejszej od średniego stopnia szarości całego obrazu zachować więcej współczynników ze względu na większą czułość oka na błędy w ciemniejszych obszarach.

## 5. Kompresja obrazu blokiem o stałym rozmiarze

Ważnym problemem pojawiającym się przy stosowaniu transformacji kosinusowej dla osiągnięcia dużych wartości kompresji jest tzw. „efekt blokowy” (ang. *block effect, tile effect*) [7-9]. Jest on widoczny w postaci krawędzi bloków. Wynika on z tego, że każdy z bloków obrazu jest kodowany niezależnie i zniekształcenia wprowadzane przez kodowanie są nieciągłe na granicy bloków. Nieciągłości te są najbardziej widoczne na obszarach

o małej wartości aktywności przestrzennej, tj. wolno zmieniającej się wartości luminancji. Drugą niedoskonałością obrazów kodowanych za pomocą transformacji kosinusowej jest tzw. „efekt dzwonienia” (ang. ringing effect). Powstaje on wtedy, gdy w danym przetwarzanym bloku obrazu sąsiadują obszar o małej wartości aktywności przestrzennej z obszarem o dużej wartości aktywności przestrzennej zawierającym np. „ostrą” krawędź. Objawia się to zaburzeniem jednorodności obszaru o małej aktywności przestrzennej w postaci fali stopni szarości. Wynika on z odrzucenia współczynników zmiennoprądowych o małych wartościach w celu osiągnięcia dużej wartości kompresji.

Poniżej przedstawiona zostanie krótka charakterystyka obrazów kodowanych blokiem o stałym rozmiarze:

- kompresja blokiem o rozmiarze  $4 \times 4$ . Obraz charakteryzuje się bardzo dobrą jakością subiektywną. Około 70% zapamiętanej informacji stanowią współczynniki stałoprądowe DC. Konsekwencją tego są widoczne „efekty blokowe”, w obszarach obrazu o wolno zmieniającej się luminancji, które są kodowane jedynie przy użyciu wartości średniej bloku. Obszary o dużej wartości aktywności przestrzennej charakteryzują się zachowaniem wysokiej wierności odwzorowania krawędzi i szczegółów. Obraz nie zawiera „efektu dzwonienia” ze względu na małą wielkość bloku. Wartość kompresji ok. 0.8 bita na piksel,
- kompresja blokiem o rozmiarze  $8 \times 8$ . Obraz charakteryzuje się dobrą jakością subiektywną. Około 60% zapamiętanej informacji stanowią współczynniki stałoprądowe DC. Są widoczne „efekty blokowe”, w obszarach obrazu o wolno zmieniającej się luminancji, będące skutkiem niezależnego kodowania poszczególnych bloków. Obszary o dużej wartości aktywności przestrzennej charakteryzują się zachowaniem dobrej wierności odwzorowania krawędzi z pominięciem drobnych szczegółów. Obraz zawiera „efekt dzwonienia”. Wartość kompresji ok. 0.3 bita na piksel,
- kompresja blokiem o rozmiarze  $16 \times 16$ . Obraz charakteryzuje się umiarkowaną jakością subiektywną. Około 50% zapamiętanej informacji stanowią współczynniki stałoprądowe DC. Wyraźnie są widoczne „efekty blokowe” w obszarach obrazu o wolno zmieniającej się luminancji, będące skutkiem niezależnego kodowania poszczególnych bloków. Obszary o dużej wartości aktywności przestrzennej charakteryzują się zachowaniem dobrej wierności odwzorowania krawędzi z pominięciem szczegółów. Obraz zawiera wyraźnie „efekt dzwonienia”. Wartość kompresji ok. 0.15 bita na piksel.

## 6. Kompresja obrazu blokiem o zmiennym rozmiarze

Idea kompresji obrazu blokiem o zmiennym rozmiarze ma na celu wykorzystanie bloków o małym rozmiarze do zakodowania obszarów o dużej wartości aktywności przestrzennej, aby zredukować „efekt dzwonienia” oraz bloków o dużym rozmiarze do zakodowania obszarów o małej wartości aktywności przestrzennej o odpowiednio niskim progu

dla współczynników zmiennoprądowych, aby zminimalizować „efekt blokowy” i uzyskać wystarczającą wartość kompresji.

W [12] przedstawiono kodler wykorzystujący blok o zmiennym rozmiarze do kompresji obrazów cyfrowych. Do tego celu użyto bloków prostokątnych o 9 różnych rozmiarach. Zaprezentowane wyniki potwierdzają powyższą ideę. Brak dostępu do użytych tam obrazów źródłowych uniemożliwia porównanie wyników z niżej omówionym zaimplementowanym algorytmem.

Zrealizowany algorytm wykorzystuje bloki kwadratowe o trzech różnych rozmiarach. Składa się on z następujących etapów:

- Obliczyć wariancję dla bloku o największym rozmiarze;
- Porównać jej wartość z wartością pierwszego zadanego progu;
- Jeśli wartość wariancji nie przekracza zadanego progu, zakodować blok współczynnikami, których wartość jest większa od wartości innego zadanego progu;
- Jeśli wartość wariancji przekracza zadany próg, podzielić blok na cztery mniejsze. Dla każdego z mniejszych bloków ponownie obliczyć wartość wariancji i porównać jej wartość z wartością drugiego zadanego progu;
- Jeśli wartość wariancji mniejszego bloku nie przekracza zadanego progu, zakodować blok współczynnikami, których wartość jest większa od wartości innego zadanego progu;
- Jeśli wartość wariancji mniejszego bloku znowu przekracza zadany próg, podzielić blok na cztery bloki o najmniejszym rozmiarze, nie podlegające już dalszemu podziałowi i są kodowane współczynnikami, których wartość jest większa od wartości innego zadanego progu.

## 7. Prezentacja wyników

Do celów kompresji wykorzystano typowy obraz naturalny zawierający twarz o rozdzielczości  $512 \times 512$  pikseli, z których każdy zakodowano na 1 bajcie, czyli o 256 stopniach szarości przedstawiony na rysunku 1.

Najpierw poddano go kompresji z użyciem bloku o stałym rozmiarze  $4 \times 4$  oraz  $8 \times 8$ , a następnie z użyciem bloków o trzech rozmiarach odpowiednio  $4 \times 4$ ,  $8 \times 8$  i  $16 \times 16$  oraz  $8 \times 8$ ,  $16 \times 16$  i  $32 \times 32$ . Uzyskano następujące wyniki:

Blok	Typ bloku	Kompresja[bit/piksel]	Bajtów	DC bajtów	PSNR[dB]
$4 \times 4$	stały	0.69	22755	16384	32.35
$4 \times 4$	zmienny	0.54	17578	3211	34.81
$8 \times 8$	stały	0.34	11030	4096	32.30
$8 \times 8$	zmienny	0.34	11271	1708	33.86





Rys. 1. Obraz oryginalny, rozdzielczość 512 \* 512 pikseli

Fig. 1. Original image, resolution 512 \* 512 pixels

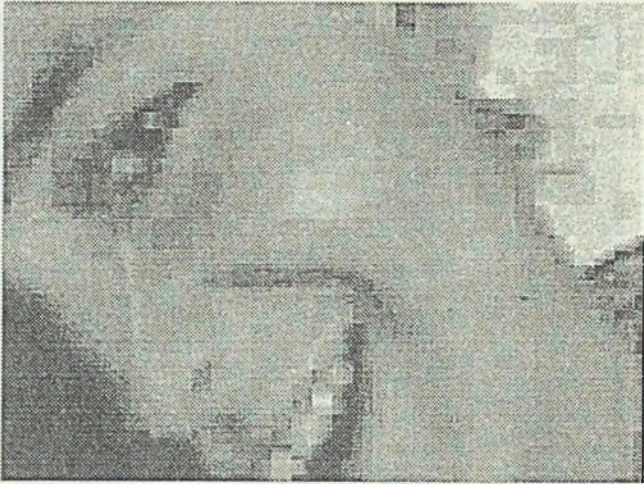
W celu wyeksponowania niedoskonałości zakodowanego obrazu wybrany jego fragment powiększono czterokrotnie i przedstawiono na rysunkach 2-5 stosując metodę roztrząsania opisaną w artykule pt. „Obrazy w stopniach szarości na drukarce laserowej”.

Przedstawione wyniki potwierdzają zalety użycia bloku o zmiennym rozmiarze do kompresji obrazów cyfrowych w porównaniu z blokami o stałym rozmiarze w odniesieniu zarówno co do wartości kompresji, jak i jakości obrazu. Wynika to z faktu użycia mniejszej ilości bloków (określonej ilości współczynników stałoprądowych DC bajtów), co redukuje „efekty blokowe”, a zaoszczędzone w ten sposób bity mogą być przypisane współczynnikom zmiennoprądowym AC poprawiając jakość odwzorowania szczegółów obrazu.

## 8. Wnioski

Możliwa jest dalsza poprawa parametrów kompresji poprzez doświadczalny dobór wartości progów dla obliczanych wariancji bloków decydujących o tym, czy dany blok należy podzielić na cztery mniejsze oraz wartości progów i szerokości przedziałów kwantyzacji dla współczynników zmiennoprądowych. Oczywisty jest fakt, że dla bloków o większym rozmiarze, reprezentujących obszary o małej aktywności przestrzennej zakres zmienności wartości współczynników zmiennoprądowych jest ograniczony i zależy od dobranej wartości progów dla wariancji.

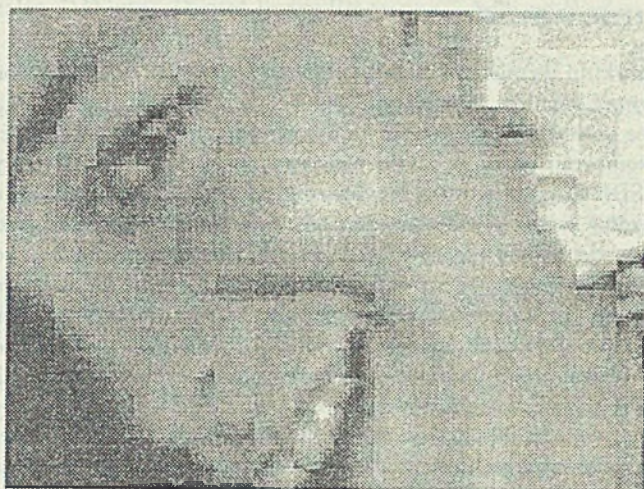
Przeprowadzone obserwacje potwierdziły ograniczoną przydatność parametru PSNR do oceny jakości obrazu. Uwidoczniło się to szczególnie w obszarach o małej wartości aktywności przestrzennej w postaci „efektów blokowych” po podwyższeniu wartości progów dla wariancji bloków o największym rozmiarze powodując gorszą subiektywną ocenę jakości obrazu w porównaniu z obrazem otrzymanym po kompresji blokiem o stałym rozmiarze mimo wyższej wartości PSNR.



*Rys. 2. Obraz po dekompresji, blok stały 4 \* 4, 0.69bit/piksel, PSNR 32.35dB*  
*Fig. 2. Decompressed image, fixed block 4 \* 4, 0.69bit/pixel, PSNR 32.35dB*



*Rys. 3. Obraz po dekompresji, blok zmienny 4 \* 4, 0.54bit/piksel, PSNR 34.81dB*  
*Fig. 3. Decompressed image, variable block 4 \* 4, 0.54bit/pixel, PSNR 34.81dB*



Rys. 4. Obraz po dekompresji, blok stały  $8 \times 8$ ,  $0.34\text{bit/piksel}$ , PSNR 32.30dB  
Fig. 4. Decompressed image, fixed block  $8 \times 8$ ,  $0.34\text{bit/pixel}$ , PSNR 32.30dB



Rys. 5. Obraz po dekompresji, blok zmienny  $8 \times 8$ ,  $0.34\text{bit/piksel}$ , PSNR 33.86dB  
Fig. 5. Decompressed image, variable block  $8 \times 8$ ,  $0.34\text{bit/pixel}$ , PSNR 33.86dB

## 9. Podsumowanie

Na wstępie artykułu wymieniono źródła obrazów cyfrowych oraz omówiono wymagania na przepustowość kanałów transmisyjnych.

Wymagania te powodują konieczność kompresji obrazów cyfrowych. Duży wpływ na rozwój technik kompresji obrazów miało wprowadzenie i upowszechnianie publicznych sieci cyfrowych ISDN (ang. integrated services digital network) oraz ustalenie międzynarodowych standardów kompresji obrazów na podstawie transformacji kosinusowej [13-16].

W artykule przedstawiono algorytm kompresji obrazów monochromatycznych przy użyciu transformacji kosinusowej o zmiennym rozmiarze bloku. Istota zastosowania transformacji do kompresji obrazu polega na tym, że informacja zawarta w oryginalnym bloku pikseli swoją najistotniejszą część zawiera w małej liczbie danych zwanych współczynnikami [2]. Z kolei zdefiniowano parametry kompresji: wartość kompresji i jakość obrazu określoną wartością stosunku sygnału do szumu (ang. PSNR).

Proces kompresji obrazu składa się z trzech zasadniczych etapów. W pierwszym etapie kompresji, kolejno dla każdego bloku obrazu, oblicza się za pomocą transformacji kosinusowej, odpowiadające mu wartości współczynników. Następnie dokonuje się ich selekcji i kwantyzacji metodami adaptacyjnymi. Ostatnim etapem kompresji jest zakodowanie skwantowanych współczynników przy użyciu kodu o stałej lub zmiennej długości, np. kodem Huffmana.

Z kolei omówiono niedoskonałości obrazu powstające jako wynik kompresji w postaci tzw. „efektu blokowego” i „efektu dzwonienia” oraz scharakteryzowano cechy obrazów zakodowanych przy użyciu transformacji kosinusowej o stałym rozmiarze bloku. Aby zredukować powyższe niedoskonałości, zaimplementowano algorytm kompresji obrazu za pomocą transformacji kosinusowej o zmiennym rozmiarze bloku [12]. Bloki o dużych rozmiarach tworzą reprezentację obszarów obrazu o małej liczbie szczegółów, natomiast bloki o małych rozmiarach tworzą reprezentację obszarów obrazu o dużej liczbie szczegółów. Obrazy zakodowane w ten sposób charakteryzują się znaczną redukcją „efektu blokowego” i „efektu dzwonienia”. Umożliwia to osiągnięcie większej wartości kompresji obrazów i poprawę ich jakości w porównaniu z obrazami zakodowanymi z użyciem bloków o stałym rozmiarze.

Do celów kompresji wykorzystano typowy obraz naturalny zawierający twarz o rozdzielczości  $512 \times 512$  pikseli, z których każdy zakodowano na 1 bajcie, czyli o 256 stopniach szarości przedstawiony na rysunku 1. Najpierw poddano go kompresji z użyciem bloku o stałym rozmiarze  $4 \times 4$  (rysunek 2) oraz  $8 \times 8$  (rysunek 3), a następnie z użyciem bloków o trzech rozmiarach odpowiednio  $4 \times 4$ ,  $8 \times 8$  i  $16 \times 16$  (rysunek 4) oraz  $8 \times 8$ ,  $16 \times 16$  i  $32 \times 32$  (rysunek 5).

## LITERATURA

- [1] A.N.Netravali, B.G.Haskell, „Digital Pictures. Representation and Compression,” AT&T Bell Laboratories 1988.
- [2] K.R.Rao, P.Yip, „Discrete Cosine Transform. Algorithms, Advantages, Applications,” Academic Press, Inc. 1990.

- [3] R.Gonzalez, P.Wintz, „Digital Image Processing,” Addison-Wesley 1987.
- [4] R.J.Clarke, „Transform Coding of Images,” Academic Press, Inc. 1985.
- [5] C.-T.Chen, T.R.Hsing, „Review Digital coding techniques for visual communications,” Journal of Visual Comm. and Image Representation, Vol. 2, No. 1, March, pp. 1-16, 1991.
- [6] K.Hosaka, „A new picture quality evaluation method,” Proc. International Picture Coding Symposium, Tokyo, Japan, April 1986.
- [7] P.M.Farrelle, „Recursive Block Coding for Image Data Compression,” Springer-Verlag New York Inc. 1991.
- [8] H.C.Reeve, J.S.Lim, „Reduction of blocking effect in image coding,” Proc. ICASSP 83, 1212-1215, Boston, 1983.
- [9] M.Miyahara, K.Kotani, „Block distortion in orthogonal transform coding- analysis, minimization, and distortion measure,” IEEE Trans. Commun., vol. COM-33, 90-96, Jan. 1985.
- [10] A.Habibi, „Survey of adaptive image coding techniques,” IEEE Trans. Commun., vol. COM-25, 1275-1284, Nov. 1977.
- [11] W.-H.Chen, H.Smith, „Adaptive coding of monochrome and color images,” IEEE Trans. Commun., vol. COM-25, 1285-1292, Nov. 1977.
- [12] I.Dinstein, K.Rose, A.Heiman, „Variable block size DCT image coder,” IEEE Trans. Commun., vol. COM-38, pp.2073-2078, Nov 1990.
- [13] „Digital Compression and Coding of Continuous-Tone Still Images,” Part 1, Requirements and Guidelines.  
ISO/IEC JTC1 Committee Draft 10918-1, Feb. 1991.
- [14] „Digital Compression and Coding of Continuous-Tone Still Images,” Part 2, Compliance Testing.  
ISO/IEC JTC1 Committee Draft 10918-2, Summer 1991.
- [15] „Video Codec for Audiovisual Services at  $p * 64\text{kbits/s}$ ,”  
CCITT Recommendation H.261, CDM XV-R 37-E, International Telegraph and Telephone Consultative Committee (CCITT), Aug. 1990.
- [16] „Coding of moving pictures and associated audio.”  
Committee Draft of Standard ISO 11172: ISO/MPEG 90/176, Dec 1990.

Recenzent: Prof. dr hab. inż. Andrzej Grzywak

Wpłynęło do Redakcji dnia 31.08.1992

## Abstract

The introductory section of the paper classifies digital image sources and states capacity requirements for image transmission channels.

Channel capacity limitations make it necessary to compress digital images. The development of image compression techniques has been boosted by the advent of Integrated Services Digital Networks and by the establishment of international standards of image compression based on the cosine transform [13-16].

The paper presents a monochrome image compression algorithm using variable block size cosine transform. The use of the transform for image compression relies on the fact that most of the essential information contained in an original block of pixels is represented by a small set of numbers called coefficients [2]. Two compression parameters are defined: the compression ratio and image quality expressed as signal to noise ratio.

Image compression is performed in three stages. At the first stage coefficients are computed for each image block. The second stage uses adaptive methods to select the essential coefficients and quantize them. The last stage consists in coding the quantized coefficients by a fixed-length or variable-length (e.g. Huffman) code.

Next are described the so-called "block effect" and "ringing effect", image faults resulting from the compression process. Image quality obtained by fixed block size compression is discussed. To reduce these artifacts, an image compression algorithm was implemented using variable block size cosine transform [12]. Large blocks are used to represent low-activity areas, while small blocks represent busy regions. Images coded in this way show much weaker "block effect" and "ringing effect". This allows higher compression ratios and better image quality to be obtained as compared to images coded at fixed block size.

As an example, a typical  $512 * 512 * 8bit$  (256 graylevels) image of a human face was compressed, first using fixed block size  $4 * 4$  (Fig. 2) and  $8 * 8$  (Fig. 3), then using three different block sizes  $4 * 4$ ,  $8 * 8$  and  $16 * 16$  (Fig. 4) and  $8 * 8$ ,  $16 * 16$  i  $32 * 32$  (Fig. 5).