

## ROZDZIAŁ 14. Kompresja wideo

Zasadniczym tematem tego rozdziału jest kompresja sekwencji obrazów na sposób stratny. Możliwość redukcji pewnej części najmniej potrzebnej informacji z obrazów pozwala na efektywne zastosowanie metod upraszczających sceny kolejnych obrazów, poszukujące adekwatnego modelu o niewielu parametrach, które mogą być łatwo adaptowane do zmian treści obrazów w sekwencji. Trzeba także pamiętać o dodatkowych warunkach stawianych algorytmom przez zastosowania transmisyjne i bazo-danowe. Łatwość dostępu do wybranych fragmentów sekwencji, odporność na błędy transmisji, duża efektywność kompresji przy małej złożoności algorytmu i niewielkiej czasochłonności są kluczowymi wymaganiami stawianymi przed nowoczesnymi koderami obrazów wideo i skojarzonego z nimi często dźwięku.

### 14.1. Kompresja obrazów wideo

Przedmiotem rozważań będą typowe obrazy wideo, głównie telekonferencyjne, filmowe, z serią scen naturalnych, w tym z możliwymi elementami grafiki komputerowej.

### 14.2. Kompresja sekwencji obrazów medycznych

Przykład kompresji sekwencji obrazów medycznych stawia przed algorytmami stratnymi szczególne wymagania. Konieczność zachowanie wartości diagnostycznej obrazów pozwala jedynie na ograniczoną redukcję informacji z obrazów przy stałej kontroli jakości każdego z obrazów w sekwencji. Źródłem informacji diagnostycznej może być bowiem każdy pojedynczy obraz w czasie niezależnej analizy.

#### **Kompresja badań 3-D**

W przypadku badań tomograficznych (CT, MRI), w których rejestrowany jest szereg obrazów kolejnych warstw prezentowanych narządów, można przeprowadzić dodatkową eliminację występującej wówczas często międzyobrazowej nadmiarowości przy pomocy tzw. międzyobrazowych (ang. interframe) technik kompresji. Pozwala to zwykle na skuteczniejszą kompresję w stosunku do metod jednoobrazowych (ang. intraframe) przetwarzających oddzielnie każdy z obrazów, kosztem jednak bardziej złożonych algorytmów i większej czasochłonności procesów kompresji i dekompresji.

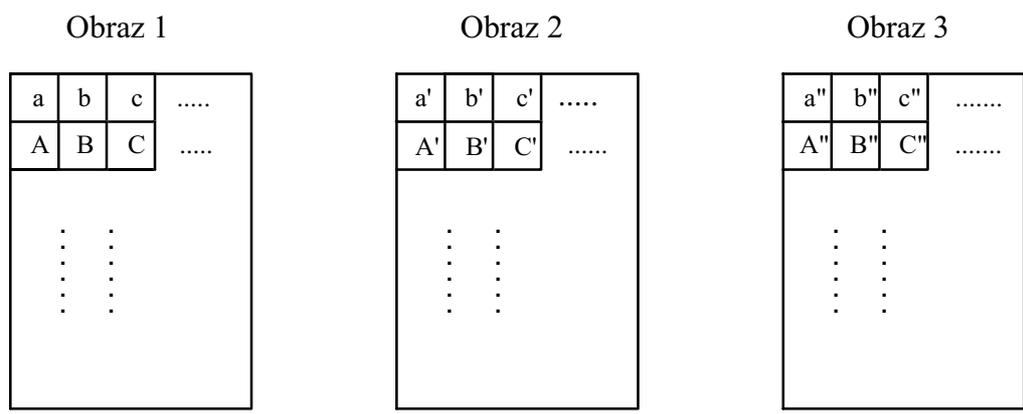
Chan [1,2] zastosował schemat 3-D transformacyjnego kodowania do kompresji ośmiu obrazów angiograficznych DSA  $512 \times 512 \times 8$  bitów (szybkość rejestracji 4 obrazy/s). Technika pełnoobjętościowej 3-D DCT z kwantyzacją i metodą przydziału bitów pozwoliła uzyskać prawie dwukrotnie większą efektywność od 2-D DCT przy błędzie średniokwadratowym MSE w granicach 1.5-2.5. Z kolei kompresja przy pomocy 3-D DCT obrazów rentgenowskich CT głowy nie wprowadziła wizualnie zauważalnych zmian w rekonstruowanych obrazach (przy stopniach kompresji od 6 do 15) i zwiększyła o około 50% wartości uzyskiwanych stopni kompresji w stosunku do 2-D DCT, przy takim samym poziomie zniekształceń. W przypadku kompresji obrazów MR zastosowanie algorytmu 3-D zamiast 2-D DCT nie zwiększyło efektywności, co autorzy tłumaczyli niższym poziomem korelacji pomiędzy sąsiednimi kadratami.

Inny sposób eliminacji nadmiarowości międzyobrazowej proponują Lee i inni []. Biorąc pod uwagę fakt, iż zasadnicze różnice pomiędzy kolejnymi 2-D warstwami związane są z ruchem obiektu, pracą lub ruchem 'kamery' oraz zmianami w anatomii ciała, proponują oni estymację przemieszczenia pomiędzy kolejnymi obrazami. Można wówczas, kodując jedynie informacje o przemieszczeniach oraz różnice porównywalnych obszarów w kolejnych obrazach, znacznie zredukować tę nadmiarowość, skuteczniej niż w 3-D DCT. Stosowana metoda estymacji jest metodą ekstrapolacyjną. Polega ona na znajdowaniu dla każdego bloku  $8 \times 8$  obrazu najlepszego przybliżenia z obrazu poprzedniego (nie oryginalnego lecz kompresowanego/ dekompresowanego, by zapobiec akumulacji błędów estymacji i kwantyzacji), poprzez przeszukiwanie regionu  $15 \times 15$  z kryterium minimalizacji błędów średniokwadratowego (stosowanego ze względów czasowych zamiast minimalizacji korelacji skrośnej przy założeniu, że błąd przybliżenia rośnie monotonicznie w miarę oddalania się od najlepszego przybliżenia). Wektory przemieszczenia są zapisywane przy pomocy ośmiu bitów na blok (po cztery bity na przesunięcie pionowe i poziome), które są następnie entropijnie kodowane i oddzielnie zapisywane (wysyłane). Natomiast w blokach wpisywane są różnicowe wartości pomiędzy danymi oryginalnymi a najlepszym przybliżeniem z obrazu poprzedniego, które są następnie transformowane za pomocą blokowej ( $8 \times 8$ ) 2-D DCT, a współczynniki tej transformaty podlegają kwantyzacji (optymalna równomierna kwantyzacja) oraz kodowaniu (składowa stała-DPCM, zmienne-RLE, a potem kodowanie Huffmana). Osiągane stopnie kompresji obrazów rentgenowskich CT głowy (przy zastosowaniu opisanej techniki) są około 5% większe dla warstw o grubości 5mm i około 10% dla warstw 3mm, niż w przypadku użycia 2-D DCT oddzielnie dla każdego obrazu. Ponadto efekty blokowe, występujące przy wyższych stopniach kompresji, są mniejsze przy estymacji przesunięć wskutek mniejszej korelacji pomiędzy wartościami różnicowymi obrazu, niż pomiędzy wartościami pikseli w obrazie oryginalnym.

**PRZYKŁAD 14.1** By zwiększyć skuteczność kompresji sekwencji obrazów skorelowanych ze sobą (obrazy tomograficzne CT, MR, dynamiczne badania scyntygraficzne itp.), zaproponowano metodę przybliżania wartości współczynników w bloku  $8 \times 8$  po transformacji DCT poprzez wartości odpowiednich współczynników z analogicznego bloku obrazu poprzedniego w danej sekwencji i kodowania różnic między wartością rzeczywistą i przybliżoną.

Postanowiono więc wykorzystać podobieństwo pomiędzy odpowiednimi zbiorami danych nie w dziedzinie obrazu, lecz w dziedzinie transformaty, co wobec postępującego ujednolicenia rozkładu wartości współczynników w blokach wskutek coraz silniejszej kwantyzacji (redukcja szumów) może zwiększyć skuteczność predykcji. Zasadniczą ideę tej metody przedstawia rys. 14.1.

Technika ta pozwala wyraźnie zwiększyć skuteczność kompresji skorelowanych obrazów w stosunku do jednoobrazowych metod kompresji przy niewielkim wzroście złożoności algorytmu (niski koszt czasowy i sprzętowy).



Kolejność kodowania: a, b, c, .....

- a' - a, b' - b, c' - c, ....
- a'' - a', b'' - b', c'' - c', ....
- A, B, C, ....
- A' - A, B' - B, C' - C, ....
- A'' - A', B'' - B', C'' - C', ....
- .....

Rys. 14.1. Schemat kodowania współczynników DCT z kolejnych bloków w sekwencji obrazów.

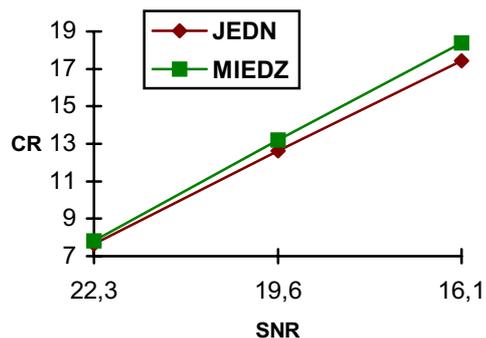
Poniżej przedstawiono wyniki testów oceny tak prostej metody kompresji sekwencji obrazów - rysunki 14.2 - 14.5. Metodę kompresji międzyobrazowej oznaczono na wykresach jako MIEDZ. Symbol JEDN dotyczy stopni kompresji uzyskanych poprzez oddzielną, niezależną kompresję każdego obrazu (metoda jednoobrazowa).

Wykresy z literką b) dla każdej grupy obrazów przedstawiają procentowy zysk (p[%]) wynikający z zastosowania międzyobrazowego algorytmu kompresji, który obliczany jest dla poszczególnych poziomów zniekształceń obrazu rekonstruowanego, określonych przez stosunek sygnału do szumów SNR, następująco:

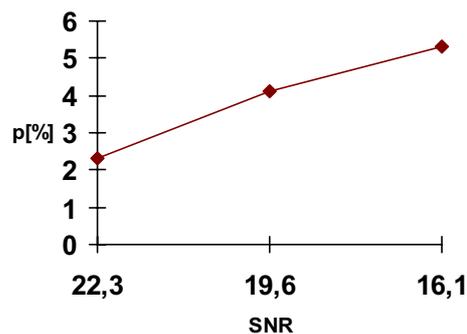
$$p[\%] = \frac{l_{jedn} - l_{miedz}}{l_{jedn}} \cdot 100\%, \quad (14.1)$$

gdzie  $l_{jedn}$  - długość skompresowanego pliku przy pomocy algorytmu jednoobrazowego,  $l_{miedz}$  - długość skompresowanego pliku według algorytmu międzyobrazowego.

W przypadku obrazów skorelowanych przestrzennie kompresowano sekwencje pięciu obrazów CT i MR, natomiast w przypadku obrazów skorelowanych czasowo sekwencje pięciu obrazów MR i szesnastu obrazów scyntygraficznych.

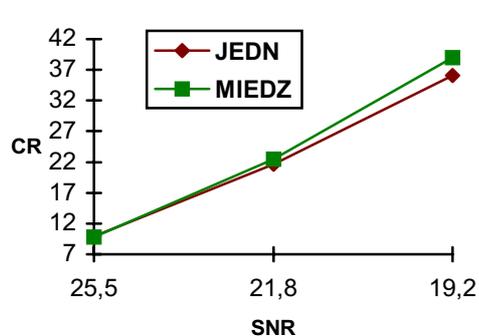


a)

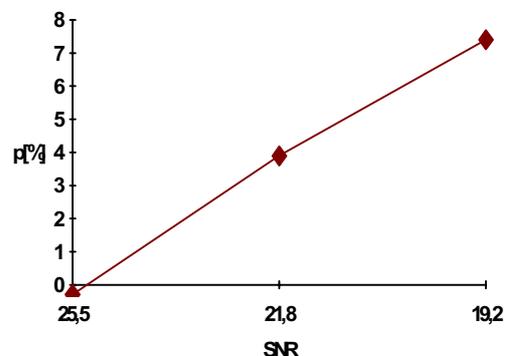


b)

Rys. 14.2. Efektywność kompresji sekwencji obrazów; a) stopnie kompresji CR oraz b) procentowy zysk p[%], uzyskane w kompresji obrazów CT skorelowanych przestrzennie.

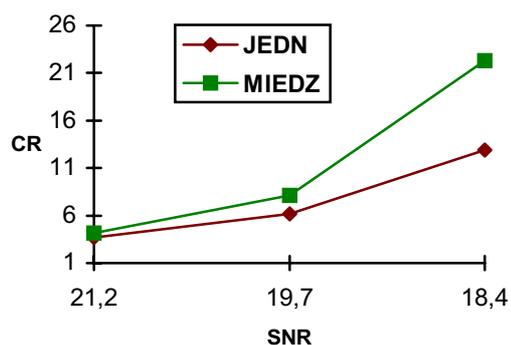


a)

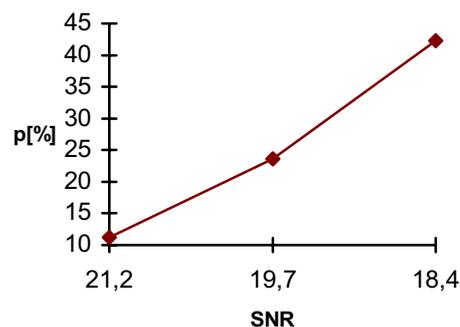


b)

Rys. 14.3. Efektywność kompresji sekwencji obrazów; a) stopnie kompresji CR oraz b) procentowy zysk p[%], uzyskane w kompresji obrazów MR skorelowanych przestrzennie.

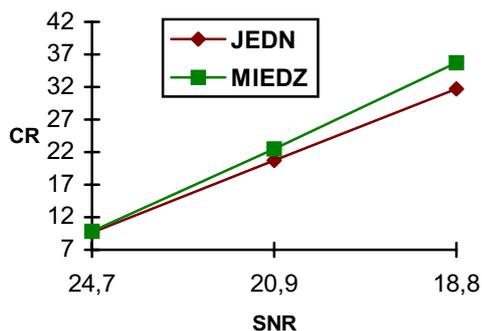


a)

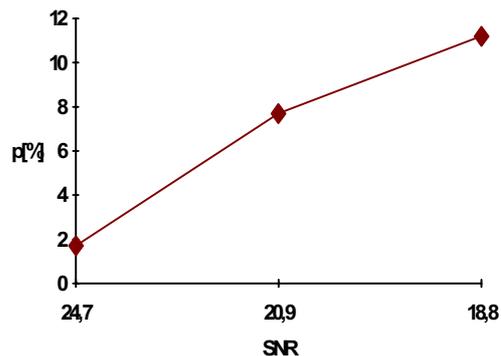


b)

Rys. 14.4. Efektywność kompresji sekwencji obrazów; a) stopnie kompresji CR oraz b) procentowy zysk p[%], uzyskane w kompresji obrazów scyntygraficznych skorelowanych czasowo.



a)



b)

Rys. 14.5. Efektywność kompresji sekwencji obrazów; a) stopnie kompresji CR oraz b) procentowy zysk  $p[\%]$ , uzyskane w kompresji obrazów MR skorelowanych czasowo.

Przedstawione wykresy pokazują znaczną efektywność tak prostego algorytmu. W porównaniu z metodami jednoobrazowymi długość reprezentacji sekwencji obrazów uległa skróceniu nawet o ponad 40%.

Obrazy skorelowane czasowo są bardziej podatne na kodowanie międzyobrazowe, co (wobec takiej samej jakości obrazów MR) związane jest przede wszystkim ze znacznie większą korelacją pomiędzy kolejnymi obrazami.

Uzyskane wyniki poprawy skuteczności kompresji są porównywalne w wynikami osiąganymi z wykorzystaniem bardziej rozbudowanych algorytmów, prezentowanymi w literaturze, a więc potwierdzają dużą przydatność przedstawionego algorytmu międzyobrazowego ze względu na jego prostotę i niskie koszty czasowo-sprzętowe.

### 14.3. Kompresja dźwięku