

dr inż. Piotr Ody

# PODSTAWY KOMPRESJI WIDEO

## Wprowadzenie

---

- ⦿ Dane multimedialne to przede wszystkim duże strumienie danych liczone w MB a coraz częściej w GB;
- ⦿ Mimo dynamicznego rozwoju technologii pamięci i coraz szybszych transferów danych osiąganych w systemach komunikacyjnych istnieje potrzeba stosowania wydajnych algorytmów kompresji danych multimedialnych.
- ⦿ Dzięki kompresji dostęp do danych jest szybszy.
- ⦿ Kompresja wpływa na jakość dostarczanej informacji

## Parametry a wielkość pliku

---

### ◎ obraz w SD (PAL)

- $720 \times 576$  pikseli  $\times$  25 ramek/s  $\times$  24 bity/piksel =  
= 248832000 bitów/s = 31104000 B/s =  
= 31,1MB/s -> ~1,87 GB/minutę -> ~112 GB/h

### ◎ obraz w HD

- $1920 \times 1080$  pikseli  $\times$  25 ramek/s  $\times$  24 bity/piksel =  
= 1244160000 bitów/s = 155520000 B/s =  
= 155,52MB/s -> ~9,3 GB/minutę -> ~560 GB/h

## Klasyfikacja metod kompresji

---

### ◎ Metody bezstratne

- Zakodowany strumień danych po dekompresji jest identyczny z oryginalnymi danymi przed kompresją,

### ◎ Metody stratne

- W wyniku kompresji część danych (mniej istotnych) jest bezpowrotnie tracona, dane po dekompresji nieznacznie różnią się od oryginalnych danych przed kompresją.

## Kompresja bezstratna

---

- ⦿ sygnał może być wielokrotnie poddawany kompresji i dekompresji
- ⦿ niewielki stopień kompresji – w przypadku danych multimedialnych
  - typowy stopień kompresji – 10-20%
  - maksymalny stopień kompresji – ok. 40-60%
- ⦿ dla dźwięku metody bezstratne stosowane powszechnie od momentu pojawienia się nośnika Blu-ray

## Kompresja stratna

---

- ⦿ Najczęściej wykorzystują niedoskonałości zmysłów percepcji człowieka -> kompresja perceptualna
- ⦿ rekompresja sygnału prowadzi do znaczących strat jakości
- ⦿ umożliwiają uzyskanie wysokiego stopnia kompresji
  - nawet 90% bez dostrzegalnych różnic

## Kompresja bezstratna

---

- ⦿ Kodowanie długości serii,
- ⦿ Metoda Huffmana,
- ⦿ Kodowanie arytmetyczne,
- ⦿ Metody słownikowe

## Kodowanie długości serii

---

- ⦿ Metoda nie potrzebuje danych początkowych o danych wejściowych,
- ⦿ Polega na zastępowaniu serii takich samych danych występujących po sobie rekordem danych zawierającym kod wzorca oraz liczbę jego wystąpień.
- ⦿ Przykładem takiej metody jest metoda RLE (Run-Length Encoding) stosowana do kodowania obrazów.

## Kompresji bezstratna

---

- RLE - ang. Run-Length Encoding - kodowanie długości serii;

```
aaaaaaabbbbaaaaaaa  
aaaaabbaabbbaaaaa  
aaaabbbaabbbaaaa
```

Bez kompresji: 3 x 18 bajtów = 54 bajty

Sposób kodowania:

a7b3a8 - 6 bajtów

a6b2a2b2a6 - 10 bajtów

a5b3a3b3a4 - 10 bajtów

Stopień kompresji: 2:1

## Kompresja stratna

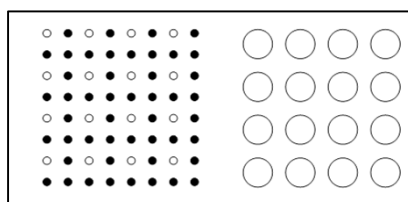
---

- Wielkość generowanych zniekształceń jest miarą jakości alg. kompresji;
- Na ogół uzyskiwane są większe stopnie kompresji;
- Algorytmy wykorzystują zaawansowane metody transformacji i kwantyzacji danych;
- Na ogół możliwe bardziej elastyczne sterowanie stopniem kompresji (jakością);

## Podpróbkowanie

- ⊙ jest stratne i polega na redukcji liczby bitów opisujących obraz
  - skutkiem ubocznym jest równoczesne obniżenie jakości tego obrazu
- ⊙ wykorzystuje zdolności ludzkiej percepcji do wypełnienia luk pomiędzy sąsiednimi pikselami.
- ⊙ dekodery mogą również wstawiać piksele usunięte w wyniku podpróbkowania (interpolacja)
- ⊙ rodzaje podpróbkowania:
  - w kopii obrazu wykorzystana jest tylko część pikseli z obrazu oryginalnego
  - grupa kilku pikseli zastępowana jest wartością średnią

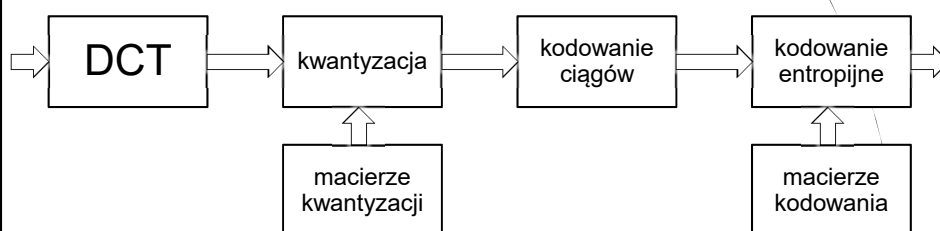
## Podpróbkowanie -przykłady



## Formaty obrazu - JPEG (\*.jpg)

- ◎ Joint Photographic Experts Group
  - najpopularniejszy format kompresji obrazów (przede wszystkim zdjęć)
  - opracowany na przełomie lat 80-tych i 90-tych
- ◎ typowo kompresja stratna
  - transformata DCT (bloki 8x8 pikseli)
  - **zerowanie składowych odpowiadających za wyższe częstotliwości (kwantyzacja)**
  - kodowanie Huffmana
- ◎ działanie w systemie Y Cb Cr (najczęściej 4:2:2)

## Kompresja JPEG - algorytm



- ◎ operacja DCT jest bezstratna
- ◎ kwantyzacja współczynników transformaty kosinusowej zmniejsza liczbę bitów, ale jednocześnie obniża jakość obrazu
- ◎ skwantowane wsp. transformaty są kodowane analogicznie do RLE
- ◎ kodowanie entropijne - najczęściej Huffmana

## Kodowanie transformacyjne

- ⊙ przekształcenie przestrzennej reprezentacji obrazu w dziedzinę częstotliwości
- ⊙ w procesach przetwarzania obrazów stosuje się dyskretną transformację kosinusową DCT (ang. Discrete Cosine Transform)

### ⊙ DCT

$$F(u, v) = \frac{1}{4} \sum_i^{16} \sum_j^{16} C(u) \cdot C(v) \cdot \cos \frac{(2i+1)u\pi}{16} \cdot \cos \frac{(2j+1)v\pi}{16} \cdot f(i, j)$$

$$C(\xi) = \begin{cases} \frac{1}{\sqrt{2}} & \text{dla } \xi = 0 \\ 1 & \text{dla } \xi \neq 0 \end{cases}$$

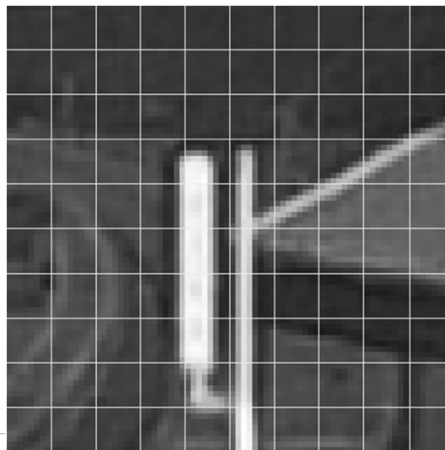
### ⊙ IDCT

$$f(i, j) = \frac{1}{4} \sum_i^{16} \sum_j^{16} C(u) \cdot C(v) \cdot \cos \frac{(2i+1)u\pi}{16} \cdot \cos \frac{(2j+1)v\pi}{16} \cdot F(u, v)$$

$$C(\xi) = \begin{cases} \frac{1}{\sqrt{2}} & \text{dla } \xi = 0 \\ 1 & \text{dla } \xi \neq 0 \end{cases}$$

## Kompresja JPEG - algorytm

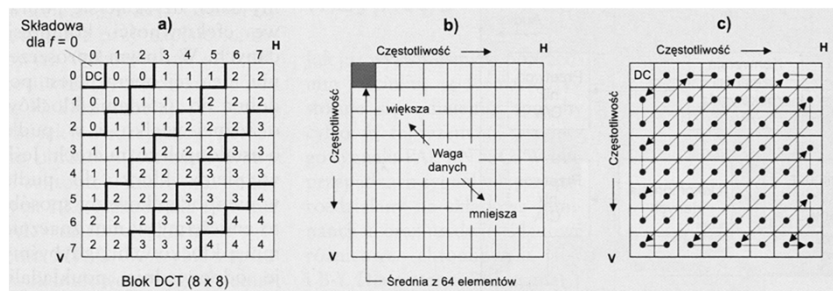
- ⊙ Obraz dzielony jest na bloki 8x8
  - upraszcza operacje
  - umożliwia zrównoleglenie obliczeń





# Kompresja JPEG

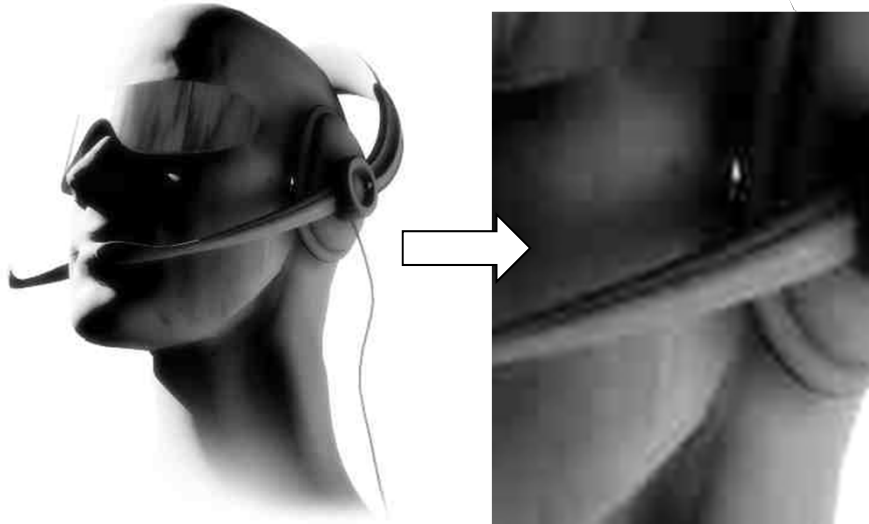
- ⦿ porządkowanie wg sekwencji “zygzakowatej” zapewnia, że współczynniki DCT o niższej częstotliwości (większe prawdopodobieństwo, iż wartości są niezerowe) są grupowane przed współczynnikami DCT o większej częstotliwości (większe prawdopodobieństwo, iż wartości są zerowe).



## Formaty obrazu – JPEG (\*.jpg)

- ⦿ możliwość kodowania progresywnego
  - kodowanie skalowalne jakościowo
  - najpierw przesyłana składowa stała i współ. niskoczęstotliwościowe, następnie współ. odpowiedzialne za wyższe częstotliwości
- ⦿ możliwość kodowania hierarchicznego
  - kodowanie skalowalne przestrzennie
  - stopniowe zwiększanie rozdzielczości przesyłanego obrazka
- ⦿ możliwość użycia trybu bezstratnego
- ⦿ możliwość zapisu dodatkowych informacji
  - profil kolorów
  - EXIF

## Formaty obrazu – JPEG (\*.jpg)



## Kompresja obrazu ruchomego

- ⦿ sygnał wizyjny można traktować jako sygnał o czterech wymiarach:
  - atrybuty pojedynczego piksela
  - rozdzielczość w poziomie
  - rozdzielczość w pionie
  - czas
  
- ⦿ każdy z tych wymiarów może podlegać kompresji

## Kompresja perceptualna

---

- ⦿ można wyeliminować z sygnału część informacji, nie powodując jednocześnie pogorszenia jego subiektywnej jakości, bo oko nie jest doskonałym przetwornikiem
  - oko jest bardziej wrażliwe na zmiany luminancji niż chrominancji
  - przy gwałtownych zmianach obrazu, oko nie dostrzega wszystkich detali

## Kompresja obrazu ruchomego

---

- ⦿ obraz ruchomy (animacja, wideo) powstaje z sekwencji obrazów statycznych
- ⦿ w procesie kompresji wykorzystuje się
  - metody kompresji obrazów statycznych
  - metody kompresji wykorzystujące właściwości ruchu obiektów w kolejnych ramkach obrazu ruchomego.
- ⦿ dwie płaszczyzny kompresji obrazu ruchomego:
  - kompresja wewnątrzramkowa,
  - kompresja międzyramkowa.

## Kompresja wewnątrzramkowa

---

- ⊙ polega na redukcji nadmiaru informacji przestrzennej w obrębie jednej ramki (ang. *spatial redundancy reduction*)
- ⊙ służy głównie do kompresji pojedynczych obrazów nieruchomych
- ⊙ może służyć do kodowania pojedynczych ramek w sekwencji sygnału wizyjnego.
- ⊙ **opiera się na analogicznych algorytmach, jak w przypadku obrazów nieruchomych**

## Kompresja międzyramkowa

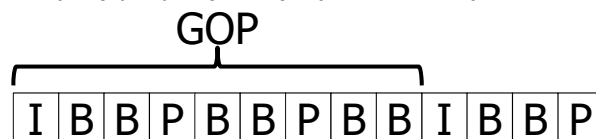
---

- ⊙ kolejne obrazy w sekwencji niewiele się od siebie różnią –możliwa jest redukcja nadmiaru informacji czasowej między kolejnymi ramkami sygnału wizyjnego
- ⊙ redukcja nadmiaru informacji czasowej polega na wyszukaniu różnic między kolejnymi ramkami i odpowiednim ich kodowaniu
- ⊙ metody kompresji międzyramkowej:
  - kodowanie różnicowe (ang. difference coding),
  - blokowe kodowanie różnicowe (ang. block based difference coding),
  - kompensacja ruchu (ang. motion compensation).

## Struktura ramek w MPEG

### ● ramki typu I (*intra frames*)

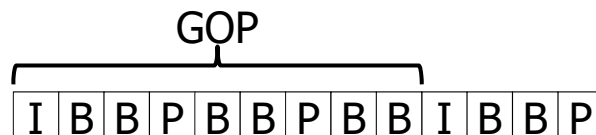
- zakodowane podobnie do JPEG, oparte na DCT
- używane jako swobodny punkt dostępu do strumienia danych MPEG
- rozpoczynają każdy GOP
- charakteryzują się najmniejszym współczynnikiem kompresji



## Struktura ramek w MPEG

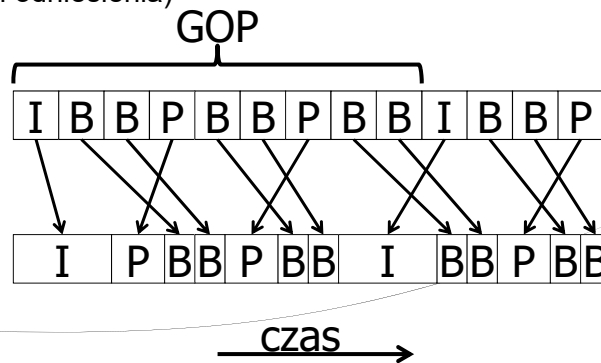
### ● ramki typu P (*predicted frames*)

- zakodowane przy użyciu kodowania predykcyjnego w przód
- ramka odniesienia (I lub P) nie musi bezpośrednio poprzedzać danej ramki
- współczynnik kompresji ramek typu P jest znacznie większy niż dla ramek typu I

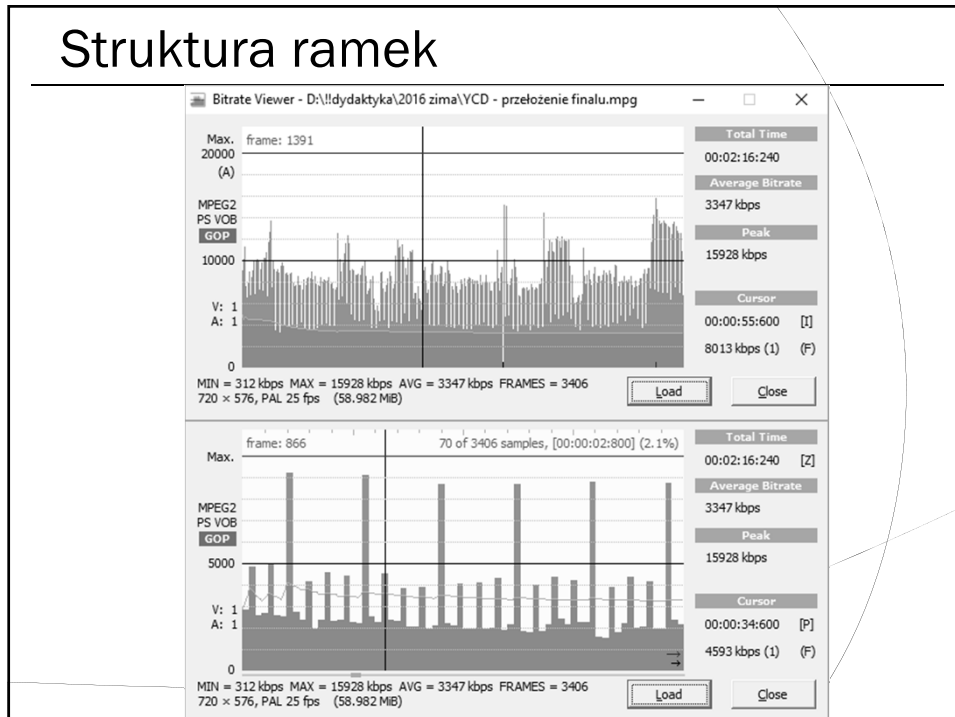


## Struktura ramek w MPEG

- ◎ ramki typu B (*bidirectional frames*)
  - zakodowane przy użyciu dwóch ramek referencyjnych: byłej i przyszłej (I lub P)
  - współczynnik kompresji dla ramek typu B osiąga największe wartości
  - konieczna zmiana kolejności transmisji obrazów (najpierw ramki odniesienia)

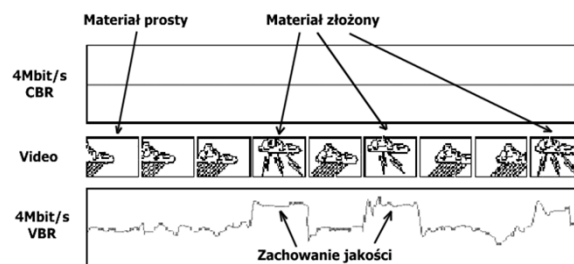


## Struktura ramek



## Przepływność

- ⊙ CBR (*Constant Bit Rate*) - utrzymywana jest stała przepływność niezależnie od materiału wejściowego
- ⊙ VBR (*Variable Bit Rate*) - chwilowa przepływność dostosowywana jest do sygnału wejściowego
  - parametrem może być jakość sygnału po kompresji, wielkość pliku lub średnia przepływność (ABR)



## Dla zainteresowanych

- ⊙ Marek Domański, „Obraz cyfrowy”, WKiŁ, Warszawa 2010.
- ⊙ John C. Russ, „The Image Processing Handbook”, CRC Press, 2007.
- ⊙ Nigel Chapman, Janny Chapman, „Digital Multimedia”, Wiley, 2009.