

# SM@KRS

27 stycznia 2009

## 1 Advanced Video Coding

H.264/AVC jest na dzień dzisiejszy najbardziej zaawansowanym standardem kompresji sekwencji wizyjnych. W stosunku do wykorzystywanego obecnie w telewizji cyfrowej standardu MPEG-2, oferuje on podobną jakość zakodowanego materiału przy dwukrotnie mniejszych prędkościach zakodowanego strumieni. Zysk ten osiągnięto dzięki wprowadzeniu szeregu narzędzi niedostępnych w MPEG-2, wśród których do najważniejszych należą:

- estymacja ruchu z dokładnością do  $1/4$  odstępu próbkowania obrazu,
- kodowanie predykcyjne makrobloków typu Intra,
- podział makrobloków typu Inter na partycje o mniejszych rozmiarach (predykcja międzyobrazowa w blokach  $16 \times 16$ ,  $16 \times 8$ ,  $8 \times 16$ ,  $8 \times 8$ ,  $8 \times 4$ ,  $4 \times 8$  oraz  $4 \times 4$ ),
- predykcja międzyobrazowa przy użyciu kilku obrazów odniesienia,
- zastosowanie filtru deblokującego w pętli predykcji kodera,
- zaawansowana predykcja wektorów ruchu w makroblokach typu B (tryb Direct),
- kontekstowe adaptacyjne kodowanie arytmetyczne elementów strumienia binarnego (CAVLC, CABAC).

Uwaga! Filtr deblokujący zastosowany w pętli predykcji usuwa efekty blokowe pojawiające się w obrazach predykcji przy silnej kompresji. W ten sposób redukuje się również błąd predykcji.

## 2 Strumieniowe przesyłanie danych

Mozliwe typy strumieniowania:

- unicast – rodzaj transmisji, w której dokładnie jeden punkt wysyła pakiety do dokładnie jednego punktu – dla danego strumienia istnieje tylko jeden nadawca i jeden odbiorca.
- multicast – sposób dystrybucji informacji, dla którego liczba odbiorców może być dowolna. Odbiorcy są widziani dla nadawcy jako pojedynczy grupowy odbiorca danego strumienia dostępny pod jednym adresem dla danej grupy multikastowej. Multicast różni się od unicastu zasadą działania i wynikającą stąd efektywnością.
- broadcast – rozsiewczy (rozgłoszeniowy) tryb transmisji danych polegający na wysyłaniu przez jeden port (kanał informacyjny) pakietów, które powinny być odebrane przez wszystkie pozostałe porty przyłączone do danej sieci (domeny broadcastowej).

## 3 Ukrywanie błędów transmisji

### 3.1 Metody przestrzenne (wewnątrzobrazowe)

Metody przestrzenne wykorzystuje się przede wszystkim do maskowania błędów w obrazach typu I czyli pierwszych obrazach GOP. Pożądany jest jak największy stopień odzyskania informacji, gdyż niepełna rekonstrukcja w wyniku zastosowania predykcji międzyobrazowej w obrazach typu P i B może doprowadzić do propagacji błędów. Synteza wewnątrzobrazowa wykorzystuje wyłącznie przestrzenną informację z obszarów nieuszkodzonych bądź odzyskanych.

#### 3.1.1 Interpolacja dwukierunkowa

Metoda interpolacji polega na przestrzennym uśrednianiu każdego punktu w makrobloku  $x_k$  na podstawie wartości najbliższych punktów pochodzących z jego czterech sąsiednich, nieuszkodzonych makrobloków

#### 3.1.2 Interpolacja przestrzenno-pionowa SVI

Spatial vertical interpolation. Wykorzystuje ona jedynie informacje z makrobloków znajdujących się powyżej i poniżej uszkodzonego.

### 3.1.3 Metoda przestrzennej dyfuzji

Proces dyfuzji redukuje efekt rozmycia, który wprowadza interpolacja, jednakże słabo rekonstruuje obszary z wieloma szczegółami lub krawędziami i liniami innymi niż pionowe (jest to charakterystyczne dla wszystkich metod przestrzennych). Ponadto metoda ta jest stosunkowo skomplikowana obliczeniowo i jakość odzyskanego obrazu nie równoważy wprowadzonego opóźnienia.

### 3.1.4 Metoda przestrzennego podziału z pasowaniem bloków

Split match error concealment. Do określenia podobieństw obszarów (pasowania bloków) można wykorzystać podejście oparte na średniej bezwzględnej różnicy MAD (ang. mean absolute difference) pomiędzy estymowanymi sąsiednimi blokami (blok górny i blok dolny). W momencie określenia najbardziej pasujących bloków przeprowadza się proces rekonstrukcji jedną z trzech metod:

- (a) kopiowania,
- (b) dyfuzji,
- (c) interpolacji

zawartości obrazu na podstawie dopasowanych bloków w uszkodzonych obszarach.

## 3.2 Metody czasowe (międzyobrazowe)

### 3.2.1 Metoda zerowego ruchu ZM (ang. zero motion)

W metodzie ZM dokonuje się tzw. „czasowego zastąpienia”, to znaczy uszkodzony makroblok zastępuje się makroblokiem z tej samej pozycji  $(i, j)$  uprzednio zdekodowanej ramki.

### 3.2.2 Metoda czasowego pasowania bloków w przód/wstecz

Temporal, forward–backward block matching EC metod.

Metoda czasowego pasowania bloków zapewnia wierniejsze odtworzenie utraconej zawartości obrazu w stosunku do poprzednie zaprezentowanych metod. Ze względu na wprowadzoną kompensację ruchu efekt blokowy praktycznie nie występuje. W obszarach gdzie się pojawia, stosuje się również

delikatną filtrację. Wadą metody jest natomiast duża złożoność obliczeniowa. Poszukiwanie najlepiej pasującego bloku wprowadza dość znaczne opóźnienia. Jednym ze sposobów ich zmniejszenia jest np. zastąpienie metody pełnego poszukiwania algorytmem poszukiwania logarytmicznego.

## 4 Strumień transportowy MPEG-2

### 4.1 Tablice Program Specific Information (PSI)

- PAT (Program Association Table). Zawiera listę programów dostępnych w strumieniu transportowym. Każdy z tych programów jest identyfikowany przez 16-bitową wartość `program_number`. Każdy program jest powiązany z PIDem jego PMT.

Pod `program_number = 0x0000` znajduje się PID wskazujący NIT. Jeżeli informacji tej tam nie ma, używana jest domyślna wartość `0x0010`.

Pakiety zawierające PAT zawsze mają PID `0x0000`.

- PMT (Program Map Table) zawiera informacje o programie – dla każdego programu jest jedno PMT. Zawiera PIDy i inne informacje dotyczące strumieni składowych programu
- CAT (Conditional Access Table) zawiera informacje dostępu warunkowego. Dokładny format niespecyfikowany w standardzie.
- NIT (Network Information Table) zawiera m.in. informacje o sieci i przesyłanych w niej strumieniach transportowych. Dokładny format niespecyfikowany w standardzie.
- TDT (Time and Date Table) – data i czas UTC.

## 5 Bazy danych i MPEG-7

### 5.1 Założenia MPEG-7

- Wymienialność metadanych – niezależność od formatu danych i typu mediów
- Obiektość – opis obiektów multimedialnych (idea MPEG-4)
- Rozszerzalność – dodawanie kolejnych elementów opisu
- Zróżnicowanie poziomów abstrakcji opisu

- Ograniczona objętość (bitowa) opisu

## 5.2 Dekryptory

### 5.2.1 Deskryptory koloru

- DCD (Dominant Colour Descriptor) – opis kolorów dominujących, do 8 najważniejszych kolorów w obrazie
- SCD (Scalable Colour Descriptor) – skalowalny deskryptor koloru (histogram kolorów obraz)
- CLD (Color Layout Descriptor) – opis barwy tła obrazu
- CSD (Color Structure Descriptor) – przestrzenne rozmieszczenie barw w obrazie
- Deskryptor przestrzeni koloru (Color Space)
- Deskryptor kwantyzacji koloru (Color Quantization)

### 5.2.2 Deskryptory kształtu

- RSD (Region-based Shape Descriptor) – opis grupy punktów stanowiących region
- CSD (Contour-based Shape Descriptor) – opis krzywizny konturu obiektu
- 3D SD (3-D Shape Descriptor) – opis kształtu obiektów trójwymiarowych

### 5.2.3 Deskryptory tekstury

- HTD (Homogeneous Texture Descriptor) – opis jednolitych obszarów tekstury
- TBD (Texture Browning Descriptor) – opis kierunkowości, regularności i ziarnistości
- EHD (Edge Histogram Descriptor) – histogram krawędzi tekstur

## 6 Ciekawostki

...