



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

**Analiza porównawcza algorytmów kodowania sygnału wizyjnego
MPEG-2 i MPEG-4 w celu określenia liczby programów w
multipleksie**

Warszawa, 2005



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

Wykonawcy pracy:

Kierownik pracy - dr inż. Alina Karwowska-Lamparska
mgr inż. Andrzej Chudziński
mgr inż. Justyn Połujan



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

Spis treści:

Wprowadzenie	4
1. Ogólne właściwości standardu MPEG-2.....	4
2. Ogólne właściwości standardu MPEG-4 H.264/AVC.....	12
3. Zasadnicze różnice między MPEG-2 a AVC/H.264.....	19
4. Zakres stosowania standardów	23
5. Opłaty licencyjne	24
6. Producenci sprzętu.....	25
7. Wnioski.....	26
8. Literatura	29



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

Wprowadzenie

Genezą pracy była prośba p. Przewodniczącej K.R.R. i T. o przeprowadzenie analizy porównawczej standardów kodowania MPEG 2 i H.264/AVC (MPEG 4) zawarta w piśmie z dnia 30 września 2005.

Przeprowadzona analiza dotyczy wyłącznie aspektów technicznych standardów i nie obejmuje analizy ekonomicznej opracowania urządzeń.

Analiza została przeprowadzona przede wszystkim z punktu widzenia liczby programów, które można umieścić w multipleksie.

1. Ogólne właściwości standardu MPEG-2

Standard MPEG-2 jest pierwszym standardem cyfrowym opracowanym pod kątem zastosowania w telewizji programowej. Określa on metodę kompresji i kodowania sygnału wizyjnego, fonii i danych dodatkowych.

w standardzie MPEG można transmitować zarówno obrazy wytwarzane w standardzie europejskim 625 linii / 50 Hz jak i w amerykańskim 525 linii / 60 Hz, dopuszczalne są również różne formaty obrazu w tym 4:3 i 16:9, wybieranie może być międzyliniowe lub kolejno liniowe.

Standardy MPEG należą do metod nieodwracalnych tzn. takich, których część informacji nieistotnych w odtwarzanym obrazie jest bezpowrotnie tracona w procesie kodowania. Metoda kompresji stosowana w standardzie wykorzystuje:

- korelację przestrzenną (wewnątrz obrazową);
- korelację czasową;
- właściwości ludzkiego oka;
- właściwości statystyczne programu.



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

Metoda jest oparta na kodowaniu hybrydowym, opartym na wewnątrzpolowej transformacji kosinusowej (DCT) i międzypolowym kodowaniu z prognozowaniem i kompensacją ruchu.

W podlegający kompresji sygnałe analogowym wykorzystując właściwości oka ludzkiego stosuje się ponad dwukrotne ograniczenie pasma chrominancji w stosunku do pasma luminancji. Możliwe są też trzy przypadki położenia próbek sygnałów różnicowych w stosunku do próbek sygnału luminancji: 4 : 2 : 0; 4 : 2 : 2; 4 : 4 : 4.

Dla przypadku: 4 : 2 : 0 – częstotliwość próbkowania sygnałów różnicowych, w obu kierunkach poziomym i pionowym, jest dwukrotnie mniejsza od częstotliwości próbkowania sygnału luminancji, a linie zawierające próbki chrominancji znajdują się pomiędzy liniami próbek luminancji.

W przypadku: 4 : 2 : 2 – częstotliwość próbkowania sygnałów różnicowych w kierunku poziomym jest dwukrotnie mniejsza od częstotliwości próbkowania sygnału luminancji, a w kierunku pionowym częstotliwość próbkowania luminancji i chrominancji są jednakowe.

W przypadku: 4 : 4 : 4 – częstotliwości próbkowania sygnałów luminancji i chrominancji w obu kierunkach są jednakowe.

Korelacje przestrzenną (wewnątrzobrazową) wykorzystuje się dzięki zastosowaniu dyskretnej transformacji kosinusowej DCT.

Dyskretna transformacja kosinusowa (DTC – *Discrete Cosine Transform*) jest linearną transformacją dwuwymiarową, mającą następujące zalety:

- wykorzystywanie w bardzo dużym stopniu korelacji pomiędzy elementami obrazu,
- zgrupowanie współczynników o znaczących amplitudach w ograniczonej części transformowanej płaszczyzny ,
- transformacja jest rzeczywista, a jej podstawowymi funkcjami są funkcje sinusoidalne,
- proces odrzucania współczynników lub modyfikacji i kwantowania ich amplitud jest bardzo podobny do procesu filtracji linearnej w obecności szumu.



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

Podstawowe procesy zachodzące przy dyskretnej transformacji kosinusowej (DCT) są następujące. Przesyłany obraz jest podzielony na małe podobrazy o wymiarach 8×8 elementów zależnie od zastosowania. Elementy podobrazu są następnie próbkowane i przesyłane do kodera dyskretnej transformacji kosinusowej kolejno dla wszystkich podobrazów. Transformacja jest przeprowadzana dla każdego bloku indywidualnie. Podobraz 8×8 próbek jest więc przetransformowany na blok 8×8 współczynników (transformant). Współczynniki w bloku transformant 8×8 reprezentują oryginalny podobraz 8×8 elementów obrazu w dziedzinie częstotliwości. Proces transformacji powoduje zgromadzenie większości informacji z obrazu oryginalnego w jednym współczynniku transformacji kosinusowej. Poziom tego współczynnika jest duży, natomiast poziomy pozostałych są małe.

Po transformacji współczynniki (transformanty) podlegają procesowi kwantowania. Dla poszczególnych współczynników są stosowane inne sposoby kwantowania w zależności od położenia danego współczynnika w podobrazie (tj. w zależności od tego jaką częstotliwość przestrzenną on reprezentuje). Kwantowanie może być linearne lub nielinearne. W przypadku współczynników odpowiadających większym częstotliwościom są tolerowane większe zniekształcenia kwantowania, a więc większe przedziały kwantowania, ponieważ zniekształcenia o tych częstotliwościach są mniej widoczne od zniekształceń o małych częstotliwościach (np. w podobrazach zawierających duże powierzchnie). Podobnie zniekształcenia kwantowania sygnału chrominancji są mniej przykre dla oka niż zniekształcenia sygnału luminancji. Zauważalność zniekształceń kwantowania zależy więc od rodzaju kodowanego podobrazu.

Przed kwantowaniem współczynniki mogą być znormalizowane przez ich wariancję. W celu zmniejszenia zniekształceń współczynniki przed procesem kwantowania są podzielone przez odpowiednie funkcje wagowe. Szczególnie szeroki przedział kwantowania wprowadza się w pobliżu zera, co powoduje w rezultacie zwiększenie liczby współczynników, którym zostaje przypisana wartość zero.



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

Dla większości transformowanych podobrazów tylko niewielka część współczynników jest znacząco różna od zera i jedynie te współczynniki muszą być kodowane i przesyłane. W przypadku typowych obrazów telewizyjnych amplitudy współczynników dotyczących wyższych częstotliwości przestrzennych są zwykle równe w przybliżeniu zera. Liczba ich, podobnie jak liczba znaczących współczynników zależy od treści podobrazów. Po kwantowaniu współczynniki są grupowane w strumień danych. Stosuje się wówczas specjalne metody wyboru współczynników, zwane metodami klasyfikacji podobrazów (*blok classification*), polegające na wybieraniu współczynników wzdłuż linii ukośnych tzw. „*zigzag scanning*”.

Wykorzystując właściwości statystyczne sygnału współczynniki dyskretnej transformacji kosinusowej są kodowane ze zmienną długością słowa. Kody o zmiennej długości słowa (VLC) przypisują słowa o różnej długości sygnałom wytworzonym w koderze, czyli w źródle sygnału cyfrowego, zwanym alfabetem źródła (*source alphabet*). Ogólna zasada kodowania o zmiennej długości słowa (VLC) polega na przypisaniu każdemu symbolowi słowa kodowanego liczby bitów odwrotnie proporcjonalnej do prawdopodobieństwa jego występowania. Wówczas symbole występujące częściej są kodowane z mniejszą liczbą bitów, a symbole występujące rzadziej – z większą. Pozwala to na zmniejszenie średniej liczby bitów na symbol. Dekoder musi w tym przypadku prawidłowo dekodować zakodowane symbole. W związku z tym kod musi spełniać tzw. regułę przedrostka (*prefix rule*) tzn. żadne ze słów kodowych nie może być równe przedrostkowi drugiego. Przykładem kodu o zmiennej długości słowa jest kod Huffmana, stosowany bardzo często w technice wizyjnej.

Korelację czasową sygnału wykorzystuje się przez zastosowanie zasady prognozowania z kompensacją ruchu, która polega na oszacowaniu ruchu różnych obiektów pomiędzy jednym polem i drugim (lub pomiędzy kolejnymi obrazami) i tworzenia prognozy w kierunku ruchu. Podstawowym elementem tej metody jest sposób oszacowania przemieszczania się elementów (ruchu). Oszacowanie to jest najczęściej oparte na informacjach kodowanych poprzednio.



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

Stosowana metoda, tzw. dopasowywanie bloków, polega na określeniu zależności pomiędzy ruchomymi częściami obrazu i obrazu nadawanego poprzednio. Obraz jest podzielony na bloki, a kompresję ruchu przeprowadza się dwuetapowo. W pierwszym etapie następuje oszacowanie ruchu, tj. przeszukiwanie poprzednio nadawanego obrazu, znalezienie bloku odpowiadającego danemu blokowi i wykonanie ortogonalnego rzutu na ten obraz analizowanego bloku. W drugim etapie jest przeprowadzana kompensacja ruchu tj. obliczenie wektora przemieszczenia analizowanego bloku (pomiędzy poprzednim jego położeniem i jego rzutem ortogonalnym) i wykorzystanie go do tworzenia prognozy. Metoda ta wymaga przesyłania informacji o wektorze przemieszczenia dla każdego bloku pociąga więc za sobą zwiększenie prędkości bitowej przesyłanego sygnału. Nie jest natomiast konieczne przeprowadzenie w dekodерze dodatkowych obliczeń kompensacji ruchu.

W standardzie MPEG - 2 obrazy są połączone w grupy o ustalonej strukturze dla całej sekwencji. Grupy zawierają określoną liczbę obrazów. Dopuszczalne są trzy sposoby kodowania sygnałów poszczególnych obrazów w grupie:

- Obrazy typu I (kodowane wewnątrzobrazowo), w których prognozę tworzy się tylko z wykorzystaniem zawartych w nich informacji, tj. położonych na tych samych lub sąsiednich liniach wybierania. Obrazy typu I muszą wystąpić na początku grupy obrazów; zapewniają jednak stosunkowo niewielki stopień kompresji..
- Obrazy typu P (kodowane z prognozowaniem międzyobrazowym), w których prognozę tworzy się z wykorzystaniem informacji zawartych we wcześniejszym obrazie (może to być, ale nie musi, poprzedzający obraz) i informacji o przemieszczeniu elementów danego obrazu w stosunku do elementów wcześniejszego obrazu (wektorze ruchu). Zapewniają one większy stopień kompresji niż typu I. Obrazy typu P mogą być odniesieniem dla innych obrazów typu P lub obrazów typu B.
- Obrazy typu B (kodowane z prognozowaniem dwukierunkowym) w których prognozę tworzy się podobnie jak w przypadku obrazów typu P, z tym, że odniesieniem dla nich są dwa obrazy – obraz wcześniejszy i obraz późniejszy. Zapewniają one



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

największy stopień kompresji, ale nie mogą być odniesieniem dla żadnego innego obrazu.

Standard MPEG - 2 nie narzuca konkretnej struktury sygnału wizyjnego; liczba obrazów poszczególnych typów w grupie obrazów zależy od konkretnej realizacji kodera.

W przypadku kodowania typu I cały obraz jest dzielony na rozłączne makrobloki, zawierające 4 bloki próbek sygnału luminancji oraz 2, 4 lub 8 bloków próbek sygnałów różnicowych, odpowiadających przestrzennie próbkom sygnału luminancji. Wszystkie bloki zawierające $8 \times 8 = 64$ próbki są niezależnie przekształcane przy wykorzystaniu dyskretnej transformacji kosinusowej oraz kodowania o zmiennej długości słowa (VLC). Obrazy typu P są kodowane różnicowo (DPCM) z prognozowaniem międzyobrazowym, przy czym prognoza jest tworzona z kompensacją ruchu metodą przesunięcia bloków. Błąd prognozy jest następnie kodowany z wykorzystaniem DCT.

W większości przypadków kodowanie obrazów typu P zapewnia większą kompresję danych niż kodowanie obrazów typu I. Standard MPEG określa jedynie format w jakim mają być przesyłane wektory ruchu i zakodowane elementy obrazu. Metoda wyznaczania wektorów ruchu jest natomiast nieokreślona i może być w każdym koderze inna.

Metoda kodowania obrazów typu B jest podobna do metody kodowania obrazów typu P. Różnica polega na tym, że zarówno tworzenie prognozy, jak i kompensację ruchu, przeprowadza się na podstawie dwóch obrazów: wcześniejszego i późniejszego. Ponieważ do odtworzenia obrazu typu B dekodery musi znać zawartość obydwu obrazów stanowiących dla niego odniesienie a więc zawartość obrazu wcześniejszego i późniejszego musi więc nastąpić zmiana kolejności nadawania obrazów. Najpierw są wysyłane dwa obrazy odniesienia, a później sam obraz typu B. Dla przechowywania obrazów odniesienia dekodery musi zawierać pamięć buforową o odpowiednio dużej pojemności. W takim przypadku uzyskuje się większą kompresję danych, niż w przypadku obrazów typu P.



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

Standard MPEG-2 może być wykorzystywany dla kodowania obrazów o różnej rozdzielczości przy zastosowaniu różnych wariantów kompresji sygnałów. W tym celu przyjęto dwa podstawowe pojęcia „*poziom*” (Level) oraz „*profil*” (Profile).

Pojęcie „poziom” jest związane z algorytmami wybierania, a mianowicie przyjęto:

- dla telewizji o dużej rozdzielczości obrazu i 1920 próbek na linii tzw. „poziom wysoki 1920”;
 - dla telewizji o dużej rozdzielczości obrazu i 1440 próbek na linii tzw. „poziom wysoki 1440”;
- liczba linii czynnych dla obu tych poziomów wynosi 1152;
- dla telewizji konwencjonalnej tzw. „poziom główny” o rozdzielczości 720 punktów x 576 linii, który może być również wykorzystywany dla systemów o poprawionej jakości (o rozszerzonym formacie obrazu);
 - dla telewizji o małej rozdzielczości obrazu (352 punkty x 288 linii) tzw. „poziom niski”.
- Sygnałami wejściowymi są zawsze sygnały składowe telewizji kolorowej.

Dla każdego z tych poziomów można stosować różne warianty metody kompresji sygnału, pozwalające na uzyskanie różnych prędkości bitowych. Parametry te nazwano profilem.

W systemie MPEG przyjęto pięć niżej wymienionych podstawowych profili (Profils).

- Profil prosty, wykorzystuje dyskretną transformację kosinusową, kodowanie z prognozowaniem oraz kompensację ruchu, przy czym sygnały różnicowe kolorowości obrazu są kodowane sekwencyjnie, co drugą linię (standard 4: 2: 0).
- Profil główny, wykorzystuje tę samą metodę kodowania, lecz dwa rodzaje prognozowania (tzw. ramka –B), co zapewnia lepszą jakość sygnału przy tej samej prędkości bitowej. Sygnały różnicowe są w tym przypadku również kodowane sekwencyjnie.
- Profil skalowany szumowo (SNR scalable). Dane wizyjne są podzielone na części : sygnał podstawowy i sygnał podwyższający jakość. Sygnał podstawowy wytwarza obraz o zmniejszonym stosunku sygnału do szumu i wymaga znacznie mniejszej



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

prędkości bitowej, może więc być przesyłany przez tory transmisyjne o mniejszej przepustowości. Ponadto, jako bardziej odporny może być odbierany przy zmniejszonym natężeniu pola. Dodanie do sygnału podstawowego sygnału podwyższającego jakość poprawia stosunek sygnału do szumu w obrazie. Zapewnia to przy tej samej prędkości bitowej lepszą jakość odtwarzanego obrazu niż w profilu głównym. Sygnały różnicowe są w tym przypadku również kodowane sekwencyjnie.

- Profil skalowany przestrzennie. Jest zbudowany podobnie jak profil skalowany szumowo, lecz wykorzystuje inną metodę podziału danych, a mianowicie w zależności od rozdzielczości obrazu. Sygnał podstawowy odpowiada obrazowi o zmniejszonej rozdzielczości. Obraz pełnej rozdzielczości otrzymuje się przez zsumowanie tego sygnału z sygnałem poprawiającym jakość. W profilu tym sygnały różnicowe są również kodowane sekwencyjnie, wymaga on jednak prędkości bitowej około 10 ÷ 20% większej niż profil główny. Profil skalowany przestrzennie umożliwia odbiór sygnałów słabszych, a więc w gorszych warunkach odbioru, przy gorszej rozdzielczości obrazu, co pozwala na stosowanie większych zasięgów stacji nadawczych.
- Profil wysoki. Stosowane są te same metody co w przypadku profilu skalowanego przestrzennie, lecz jednocześnie kodowanie sygnałów różnicowych kolorowości obrazu. Powoduje to zwiększenie prędkości bitowej, co daje dalszą poprawę jakości. Jednakże w przypadku wymaganej ograniczonej prędkości bitowej jakość odtwarzanego obrazu pogarsza się.

Prędkość bitowa przesyłanego sygnału zależy od kombinacji poziom/profil wynosi dla telewizji konwencjonalnej (SDTV) 2 ÷ 4 Mbit/s a dla telewizji dużej rozdzielności (HDTV) 8 ÷ 10 Mbit/s.

Potwierdzona wynikami badań prędkość bitowa, przy transmisjach eksperymentalnych sygnałów SDTV kodowanych wg. standardu MPEG-2 przy jakości odtwarzanego obrazu równorzędnej jakości obrazu analogowego, wynosi średnio 3,5 Mb/s.



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

Minimalne wartości prędkości bitowej dla standardu MPEG-2 i jakości SDTV zależą od treści przesyłanego obrazu i wynoszą około 3,5 Mb/s, natomiast dla obrazów krytycznych (o dużej zawartości ruchu) do 4,5 Mb/s.

2. Ogólne właściwości standardu MPEG-4 H.264/AVC

Szeroko stosowany w praktyce MPEG-2 został ulepszony w roku 1991 przez grupę MPEG poprzez wprowadzenie standardu MPEG-4 zgodnego z normą ISO/IEC 14496 cz. 2. Standard ten wykorzystuje bardziej zaawansowane techniki kompresji oraz wiele dodatkowych narzędzi umożliwiających kodowanie i manipulowanie mediami cyfrowymi

Jądro rozwiązania jest oparte na znanym modelu kodowania hybrydowego DPCM/OCT a podstawowe funkcje modelu są wspomagane, jak na przykład zwiększenia wydajności kompresji, niezawodności transmisji, kodowanie niezależnych obiektów sceny wizyjnej, kompresja oparta na siatce i ożywianie twarzy i modeli ludzkich. Jest to standard bardzo skomplikowany wykorzystujący skomplikowany software. Oczywiście pojedyncze zastosowanie nie wymaga zastosowania wszystkich możliwości standardu. Standard MPEG-4 opisuje wiele profili grupujących narzędzia dla poszczególnych zastosowań a między innymi wydajne kodowanie ramek wizyjnych, kodowanie wizyjne dla zawodnych sieci transmisyjnych, kodowanie i manipulacja oparte na obiekcie. kodowanie syntetycznych i hybrydowych, syntetyczno-naturalnych scen i interaktywne zastosowania wizyjne. Jednakże z punktu widzenia zastosowań najbardziej potrzebne są narzędzia profilu kodowania prostego i ulepszanego prostego oraz wydajnego kodowania prostokątnych ramek. wymagania takie spełnia standard H.264/AVC wynaleziony w 2003r. w celu zastąpienia standardu MPEG-4.

Opublikowany w ISO/IEC 14496 cz. 10 i zaleceniu ITU-T 264.

Podobnie jak koder standardu MPEG-2, koder standardu H.264/AVC jest koderem hybrydowym z kompensacją ruchu.



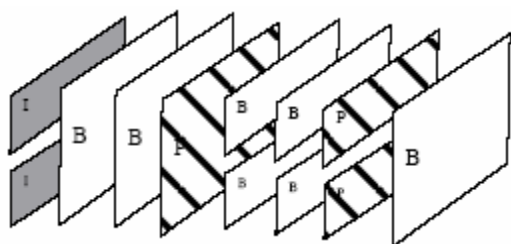
Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

W standardzie AVC wprowadzono pewne innowacje do sposobu kodowania obrazów w sekwencji wizyjnej:

- Każdy obraz w sekwencji może być zakodowany niezależnie jako RAMKA (FRAME), POLA (FIELDS) oraz w trybie adaptacyjnym MBAFF;
- Może być więcej niż jeden obraz odniesienia zarówno na liście odniesienia wstecz jak i w przód;
- Obrazy mogą być kodowane praktycznie w dowolnej kolejności jak przedstawiono na rysunku poniżej:



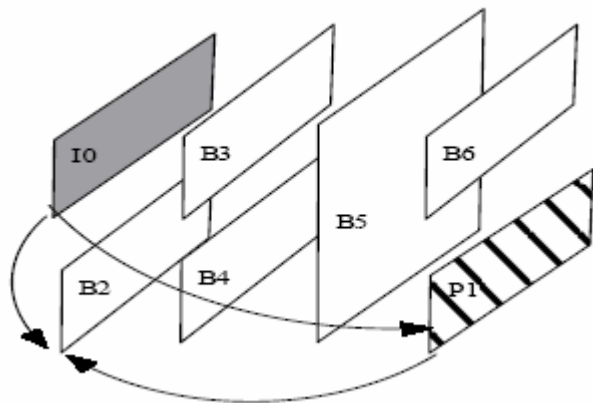
Wprowadzono zarządzanie listą obrazów odniesienia. Lista ta ma dwie części: krótko i długoterminową. Na liście krótkoterminowej znajdują się obrazy ostatnio zdekodowane natomiast na liście długoterminowej mogą się znaleźć obrazy wskazane przez specjalne znaczniki w strumieniu. Zdefiniowane znaczniki strumienia AVC umożliwiają precyzyjne sterowanie zawartością obu list. Poniżej przedstawiono przykład kolejności kodowania pól w sekwencji. Numery obok typu pola/ramki informują o kolejności kodowania.



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095



Wybór kolejności kodowania kolejnych obrazów ma istotny wpływ na stopień kompresji. Dzięki temu możemy np. drugie pole pierwszego obrazu (obrazu dostępowego INTRA) zakodować jako pole B. w efekcie możemy uzyskać nawet 25% wzrost stopnia kompresji wynikający tylko i wyłącznie ze zmiany kolejności kodowania i dopuszczeniu różnego typu pól w jednym obrazie (Rys.3)

W standardzie AVC wprowadzono także kodowanie adaptacyjne POLE/RAMKA na poziomie makrobloku (MBAFF). W takim trybie kodowania obrazu sąsiadujące w pionie makrobloki łączone są w pary. Każda taka para może zostać zakodowana jako dwa makrobloki z przeplotem lub jako progresywne.

W AVC wartości próbek obrazu jak i wektory ruchu wyliczane są na podstawie prognozy danych zawartych w sąsiadujących blokach 4x4. W trybie MBAFF wyznaczenie sąsiedztwa jest dość złożone, ponieważ każdy z sąsiadów (prawy, górny, lewy-górny i prawy-górny) może być zakodowany w innym trybie, a zatem istnieją 4 kombinacje dla każdego z czterech z sąsiedztwa. Tak, więc, aby określić konteksty potrzebne do zakodowania bieżącego makrobloku należy znać typ makrobloku sąsiadującego. Tryb MBAFF pozwala adaptować się lokalnie do struktury obrazu. Tam gdzie nie ma przeplotu, obraz jest gładki i nie ma tekstur przeważnie używany jest tryb FRAME, w innych przypadkach a szczególnie na krawędziach obiektów stosuje się tryb FIELD. Kodowanie obrazu z adaptacją trybu na poziomie makrobloku daje największy współczynnik kompresji.



Centralne Laboratorium Badawcze

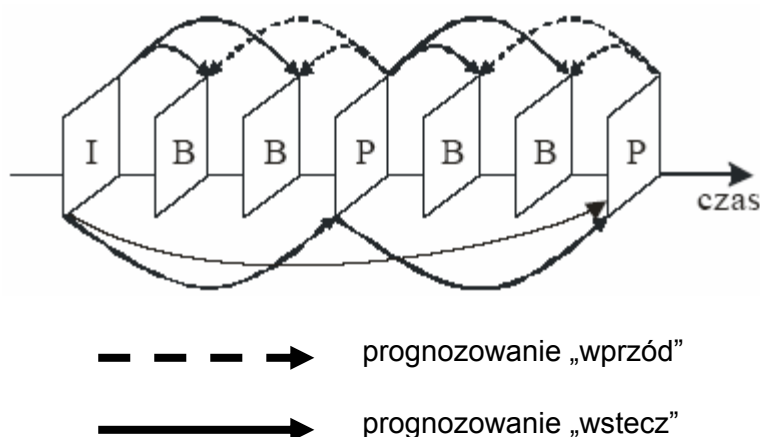
Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

W standardzie AVC/H.264 obrazy kodowane w trybie międzyobrazowym (P i B) rekonstruowane są przy użyciu jednego lub więcej obrazów odniesienia, występujących w sekwencji wcześniej lub później w stosunku do bieżącego obrazu.

Każdy z makrobloków może zostać podzielony w procesie kodowania na prostokątne fragmenty o rozmiarach 16×8 , 8×16 , 8×8 , 8×4 , 4×8 oraz 4×4 próbki luminancji.

Mechanizm *międzyobrazowego prognozowania z kompensacją ruchu* przedstawiono poniżej:



Do rekonstrukcji makrobloku w takim obrazie konieczne jest przesłanie w zakodowanym strumieniu binarnym indeksów obrazów odniesienia oraz wektorów ruchu dla wszystkich fragmentów makrobloku. Ponieważ pole wektorów ruchu cechuje się lokalną korelacją, dlatego kodowaniu entropijnemu podlegają jedynie błędy wektorów określonych na podstawie prognozy, co owocuje znacznym zmniejszeniem strumienia wektorów ruchu.

W standardzie AVC zdefiniowano dwa ogólne schematy prognozy wektorów ruchu: prognoza kierunkową dla makrobloków podzielonych na prostokąty o rozmiarze 16×8 i 8×16 oraz adaptacyjną prognozę medianową dla pozostałych przypadków.

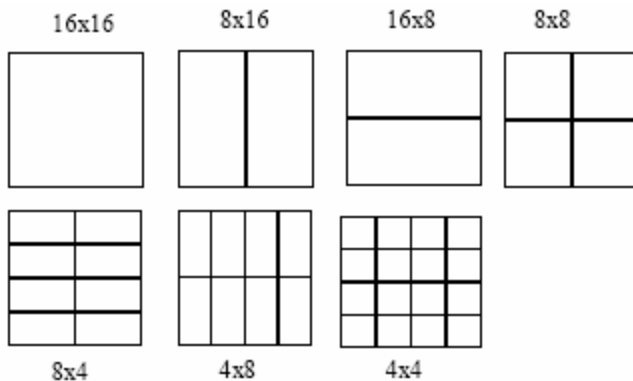


Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

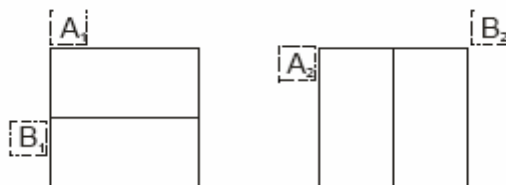
Przykładowe podziały makrobloku na partycje w standardzie AVC przedstawiono poniżej.



Tryby podziału makrobloku na prostokąty o rozmiarach 16×8 i 8×16 zostały wprowadzone specjalnie do reprezentacji fragmentów obrazu o wyraźnym kierunkowym charakterze ruchu, tam gdzie występuje krawędź pomiędzy poruszającymi się obiektami, bądź też poruszającym się obiektem a statycznym tłem.

W trybie 16×8 prognozę wektora ruchu dla górnego fragmentu makrobloku jest wektor ruchu z górnego sąsiedniego bloku (A1), zaś prognoza wektora ruchu dla dolnego fragmentu makrobloku jest wektor ruchu z lewego sąsiada (B1 - Rys. 7).

W trybie 8×16 wektor ruchu dla lewego fragmentu makrobloku jest przewidywany na podstawie lewego sąsiedniego bloku (A2), zaś wektor ruchu dla prawego fragmentu makrobloku – z górnego prawego bloku (B2). Kierunkowa prognoza wektorów ruchu przedstawiono poniżej:



--- bloki używane do prognozowania wektorów ruchu
———— kodowane prostokąty o rozmiarach 8×16 i 16×8



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

Powyższy schemat prognozy ma zastosowanie tylko w przypadku, gdy sąsiednie bloki używają tego samego obrazu odniesienia. W przeciwnym razie, prognozowanie przełącza się w tryb prognozy medianowej, opisany poniżej. Prognoza medianowa wykorzystuje wektory ruchu trzech bloków sąsiadujących z kodowanym fragmentem makrobloku (Rys. 8), standardowo są to wektory z bloków A, B i C, chyba, że obraz odniesienia bloku C jest różny od obrazu odniesienia bieżącego bloku, wówczas wektor ruchu bloku C zastępowany jest wektorem ruchu bloku D. Następnie prognoza przebiega kolejno według następującego schematu:

- jeśli którykolwiek z bloków A, B, C/D kodowany jest w trybie wewnątrzobrazowym, występuje poza krawędzią obrazu lub nie używa wektora ruchu dla danego trybu prognozy („wprzód”/”wstecz”), wówczas oznacza się go jako blok używający nieistniejącego obrazu odniesienia,
- jeśli jedynym dostępnym z punktu widzenia prognozy jest blok A, wynikiem prognozy jest jego wektor ruchu,
- jeśli tylko jeden spośród bloków A, B, C/D używa tego samego obrazu odniesienia co bieżący fragment makrobloku, wynikiem prognozy jest wektor ruchu tego bloku,
- wynikiem prognozy jest wektor o składowych będących medianą wektorów ruchu bloków A, B, C/D.

W H.264/AVC stosuje się tryby kodowania makrobloku z pomijaniem całości (SKIP) lub części danych (DIRECT). Tryby kodowania makrobloków pomijanych SKIP i DIRECT są szczególnymi trybami prognozy z kompensacją ruchu, występującymi w obrazach typu P i B. Dla tych makrobloków w strumieniu binarnym nie są przesyłane żadne dane (SKIP), lub przesyłane są wyłącznie współczynniki transformaty błędów prognozy (DIRECT).

Wektory ruchu oraz obrazy odniesienia dla makrobloków pomijanych są wyznaczone przy pomocy danych z makrobloków sąsiednich. Tryby te pozwalają bardzo efektywnie reprezentować fragmenty kodowanych obrazu, w których ruch ma charakter stacjonarny w sensie czasowym lub przestrzennym. Dla makrobloków typu



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

SKIP wysłana jest tylko informacja o liczbie kolejno występujących po sobie takich makrobloków, zaś tryb DIRECT sygnalizowany jest niezależnie dla każdego makrobloku.

W obrazach typu P dozwolone są tylko makrobloki typu SKIP, zaś w obrazach typu B – makrobloki typu SKIP lub DIRECT. W obrazach typu B możliwe są dwa rodzaje makrobloków pomijanych: z przestrzenną lub czasową prognozą wektorów ruchu. Rodzaj ten sygnalizowany jest niezależnie dla każdej warstwy makrobloków, zatem może być adaptacyjnie przełączany w zależności od treści sekwencji.

W obrazach P kompensacja ruchu dla makrobloków pomijanych odbywa się dla bloków 16×16 próbek luminancji, zaś w obrazach typu B niezależnie dla bloków 4×4 lub 8×8 próbek luminancji.

Podczas prognozy wektorów ruchu dla makrobloków SKIP w obrazach typu P najpierw sprawdzane jest czy wszystkie sąsiednie bloki (A,B i C) używają do prognozy międzyobrazowej obrazu odniesienia o indeksie zerowym i czy ich wektory ruchu są zerowe. W takim przypadku, przyjmuje się, że we danym fragmencie obrazu nie występuje żaden ruch i wówczas próbki danego makrobloku są kopią odpowiednich próbek z obrazu odniesienia, bez kompensacji ruchu.

Jeżeli którykolwiek z sąsiednich makrobloków nie spełnia powyższego warunku (tzn. jest kodowany w trybie wewnątrzobrazowym lub ma niezerowe wektory ruchu) wówczas dokonywana jest standardowa prognoza medianowa, opisana w punkcie 4 a wektor wynikowy tej prognozy staje się wektorem ruchu dla danego makrobloku. Następnie dokonywana jest kompensacja ruchu.

Standard H.264/AVC jest standardem wizyjnym, natomiast standardy kodowania dźwięku towarzyszących są takie same jak dla standardu MPEG-2. Standard H.264/AVC dotyczy wyłącznie kodowania źródłowego natomiast jest przesyłany za pomocą strumienia transportowego MPEG-2 TS podobnie jak w MPEG-2.



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

3. Zasadnicze różnice między MPEG-2 a AVC/H.264

Do podstawowych różnic pomiędzy standardem MPEG-2 a standardem H.264/AVC należą:

1. adaptacyjny podział makrobloków od 16x16 do 4x4 próbek luminancji na blok
2. kodowanie adaptacyjne sekwencji wizyjnych – na poziomie obrazu (PAFF) oraz na poziomie makrobloków (MBAFF);
3. specjalne tryby rekonstrukcji makrobloków z pomijaniem danych (SKIP i DIRECT)
4. zaawansowane prognozowanie wewnątrz obrazowe
5. zaawansowane prognozowanie międzyobrazowe z większą niż jeden liczbą obrazów odniesienia
6. transformata całkowitoliczbową błędów prognozowania w blokach o wielkości 4x4 próbki
7. kontekstowe kodowanie entropijne – kodowanie o zmiennej długości (VLC) lub kontekstowym adaptacyjnym koderem adaptacyjnym (CABAC)
8. transformata całkowitoliczbową błędów prognozy w blokach o wielkości 4x4 próbki,
9. specjalne tryby rekonstrukcji makrobloków z pomijaniem danych (SKIP i DIRECT),
10. zaawansowane narzędzia do kodowania sekwencji wizyjnych z przeplotem (PAFF oraz MBAFF)
11. wieloobrazowa kompensacja ruchu przy wykorzystaniu jako odniesienie poprzednio kodowanych obrazów, znacznie bardziej elastycznie niż w poprzednich standardach, tj. umożliwiającą stosowanie 32 obrazów odniesienia w pewnych przypadkach (w przeciwieństwie do poprzednich standardów gdzie typowo stosowano jeden lub w przypadku obrazów typu



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

- dwa. Ta szczególna cecha daje niewielkie korzyści w ograniczeniu prędkości i poprawie jakości wielu scen. W pewnych rodzajach scen np. scen zawierających gwałtowne powtarzające się błyski lub scen zawierających cięcia lub duże fragmenty jednorodnego tła pozwala na znaczne obniżenie prędkości bitowej.
12. Kompensacja ruchu w blokach o zmiennych rozmiarach (VBSMC) przy rozmiarach bloków tak dużych jak 16x16 i tak małych jak 4x4, umożliwiającą bardzo precyzyjną segmentację rejonów ruchomych.
 13. Dokładność do 1/4 piksela dla kompensacji ruchu umożliwiającą bardzo precyzyjne opisanie przemieszczanie się powierzchni ruchomych dla chrominancji rozdzielczość jest typowo mniejsza o połowę (patrz 4:2:0) wobec czego dokładność kompensacji ruchu jest obniżona do 1/8 piksela.
 14. Ważone prognozowanie umożliwiające w koderze wybór określonego skalowania i offsetu przy przeprowadzaniu kompensacji ruchu i zapewniające znaczną korzyść w warunkach specjalnych takich ściemnianie, rozjaśnianie oraz przejścia między rozjaśnianiem i ściemnianiem
 15. Struktura par makrobloków umożliwiającą stosowanie makrobloków o rozmiarach 16x16 w trybie pola, podczas gdy w MPEG-2 jest 16x8.
 16. Odpowiednia filtracja dla uzyskania prognoz próbek sygnału luminancji wynoszących 1/2 elementu obrazu, w celu zmniejszenia efektu przeplatania i zwiększenia ostrości obrazu.
 17. Zastosowanie pętli kodowania filtru, który zapobiega powstawaniu zniekształceń wywołanych efektem podziału na bloki powstające podczas kompresji obrazów opartej na transformacie DTC.
 18. Kodowanie adaptacyjne w stosunku do kontekstu o zmiennej długości (CAVLC), które jest mniej skomplikowanym, alternatywnym do CABAC dla kodowania wartości współczynników transformacji. Chociaż mniej skomplikowany niż CABAC, CAVLC jest bardziej dokładny i bardziej



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

dokładny aniżeli metody stosowane typowo dla kodowania współczynników innych poprzednich projektach.

19. Wspólne o prostej i skomplikowanej strukturze techniki kodowania o zmiennej długości (CLC) dla wielu elementów składowych niekodowanych metodami CABAC lub CAVLC, odnosi się to do kodowania wykładniczego Golomb (Exp-Golomb).
20. Określenie warstwy sieciowej abstrakcyjnej (NAL) umożliwiającej, że te same składniki wizyjne mogą być stosowane w wielu otoczeniach sieciowych, takich jak zbiory parametrów sekwencji (SPS) i zbiory parametrów obrazu (PPS) co zapewnia większą moc i elastyczność niż poprzednie projekty standardów.
21. Część przełączająca (zwana częścią SP i SJ) jest to cecha, która umożliwia koderowi zarządzanie dekodowaniem, aby przełączył się na przychodzący strumień wizyjny dla takich celów jak przełączanie prędkości bitowej strumienia i pracy w "trybie pułapka" gdy dekodowanie wskoczy na środek strumienia wizyjnego wykorzystując cechę SP/SJ może uzyskać dokładne dopasowanie do dekodowanego obrazu w tym położeniu w strumieniu wizyjnym niezależnie czy stosuje się różne obrazy (lub brak obrazu) jako odniesienia przed przełączeniem.
22. Dokładne dopasowanie transformacji bloku 4 x 4 (podobnego do znanej transformaty DCT) i w przypadku nowych profili "wysokich" FREXT zdolność koderów adaptacyjnego wyboru pomiędzy wymiarami transformowanych bloków 4 x 4 i 8 x 8 dla uzyskania całościowej pracy transformacji.
23. Wtórna transformata Hadamarda przeprowadzana na współczynnikach DC pierwotnej transformaty przestrzennej (dla współczynników chrominancji i również w specjalnych przypadkach współczynników luminancji w celu uzyskania lepszej kompresji gładkich powierzchni).
24. Prognoza przestrzenna z krawędzi sąsiednich bloków dla kodowania "wewnętrznego" (w odróżnieniu od prognozowania tylko DC w MPEG 2



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

- część 2 i prognozowania współczynników transformacji w H.263+ i MPEG 2 część 2).
25. Kodowanie arytmetyczne binarne adaptywne w stosunku do kontekstu (CABAC), co jest wygodną techniką bezstratnej kompresji elementów składowych strumienia wizyjnego przy znajomości prawdopodobieństwa wystąpienia elementów składowych w danym kontekście.
 26. Elastyczne zarządzanie mikroblokami (FMO, znane także jako grupy segmentów) i przypadkowe zarządzanie segmentami (ASO) co jest techniką rekonstrukcji zarządzania reprezentacją podstawowych rejonów (zwanymi makroblokami) w obrazach. Typowo zaprojektowane własności silnej utraty błędu FMO i ASO mogą być również stosowane dla innych celów.
 27. Podział danych (DP) właściwość umożliwiająca oddzielenie bardziej ważnych i mniej ważnych elementów składowych w różne pakiety, umożliwiającą zastosowanie zróżnicowanej protekcji błędu (UEP) i innych rodzajów korekcji błędów.
 28. Nadmiarowe segmenty RS własność silnej utraty błędów pozwalająca, aby koder wysłał ekstra reprezentacje rejonu obrazu) typowo dla mniejszej wierności co może być stosowane jeśli pierwotna reprezentacja jest uszkodzona lub stracona.
 29. Prosty proces automatyczny dla zapobiegania przypadkowej emulacja kodów startowych, stanowiących specjalną sekwencję bitów w kodowanych danych, co umożliwia swobodny dostęp do strumienia bitów i odtworzenia ustawienia bajtów w systemie, który mógł utracić synchronizację bajtów
 30. Dodatkowa ulepszająca informacja (SEI) i informacja użyteczności wizji (VUI), które stanowią dodatkowe informacje, które mogą być wprowadzone do strumienia bitów w celu ulepszenia zastosowania wizji dla wielu różnych przeznaczeń.
 31. Obrazy pomocnicze, które mogą być stosowane dla takich celów jak łączenie ze znakami alfanumerycznymi.



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

32. Numeracja ramek, własność która umożliwia tworzenie „sub-sekwencji” (umożliwiająca skalarność czasową przy opcjonalnym włączeniu obrazów ekstra pomiędzy inne obrazy), detekcję i ukrycie strat całego obrazu (które może wystąpić wskutek utarty pakietu w sieci lub błędów kanałowych).
33. Obniżenie rzędu obrazu – własność, która służy do utrzymania porządkowania obrazów i wartości próbek w obrazach dekodowanych wydzielonych z informacji czasowej (umożliwiająca, że informacja czasowa jest przesyłana i kontrolowana/zmieniana oddzielnie przez system bez oddziaływania na zawartość dekodowanego obrazu).

4. Zakres stosowania standardów

Standard MPEG-2 znalazł zastosowanie w wielu dziedzinach transmisji, przetwarzania i przechowywania treści telewizyjnych,

- telewizja programowa satelitarna DVB-S, kablowa DVB-C i naziemna DVB-T
- telewizja wysokiej rozdzielczości HDTV w ograniczonym zakresie
- telewizja płatna
- zapis na nośnikach DVD,

Standard AVC/H.264 ma jeszcze szersze zastosowanie w wielu dziedzinach transmisji, przetwarzania i przechowywania treści telewizyjnych:

- telewizja programowa satelitarna DVB-S, kablowa DVB-C i naziemna DVB-T
- telewizja wysokiej rozdzielczości HDTV
- telewizja płatna
- telewizja w telefonii trzeciej generacji
- zastosowania militarne w NATO
- telewizja w mediach strumieniowych
- telekonferencje
- zapis na nośnikach HD-DVD, Blu-ray Disk
- usługi wideo na życzenie w sieciach internetowych
- w piątej generacji iPod do odtwarzania sekwencji wizyjnych firmy Apple



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

- w przenośnych stacjach gier Playstation Portale firmy Sony Standard AVC/H.264 zastosowano w telewizji naziemnej we Francji, USA, Rosji, Niemczech i Korei Płd. We Francji standard AVC/H.264 zastosowano dla telewizji HDTV i w programach płatnych natomiast w Niemczech i Rosji dla HDTV.

W Japonii wprowadzono usługę odbioru telewizji w ruchu ISDB-T w sieciach głównych dostawców:

- NHK
- Tokyo Broadcasting System (TBS)
- Nippon Television (NTV)
- TV Asahi
- Fuji TV
- TV Tokyo

Bezpośredni odbiór z satelity zapewniają następujący nadawcy:

- DirecTV (w USA)
- Dish Network (w USA)
- Euro1080 (w Europie)
- Premiere (w Niemczech)
- ProSieben HD & Sat1 HD (w Niemczech)
- BSkyB (w Zjednoczonym Królestwie i Irlandii)

5. Opłaty licencyjne

Tak, jak m. in. w przypadku MPEG-2, Części 1 i 2 oraz MPEG-4 Część 2, od sprzedawców produktów i usług H.264/AVC oczekuje się płacenia tantiem za licencje na opatentowane technologie wykorzystywane w ich produktach. Podstawowym źródłem licencji na patenty stosowane w niniejszym standardzie jest prywatna organizacja znana jako MPEG-LA, LLC (niezwiązana w żaden sposób z organizacją standaryzacji MPEG, lecz administrująca także patentami na MPEG-2 Part 1 Systems, MPEG-2 Part 2 Video, MPEG-4 Part 2 Video i inne technologie. Via Licensing zarządza także patentami H.264. Niektórym posiadaczom patentów nie



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

wolno przyłączyć się do żadnego z dwóch ww. właścicieli. (Właściciele licencji generalnie nie występują przeciwko patentom stron trzecich i nie mogą zmusić posiadaczy patentów do przyłączenia się do nich). Sytuacja ta spowodowała niechęć do przyjęcia H.264 przez niektóre instytucje potencjalnie przyjmujące ten standard i może spowodować przyjęcie alternatywnych kodeków, na które opłaty licencyjne i ryzyko sporów sądowych są niższe.

6. Producenci sprzętu

Kodery AVC/H.264 są produkowane między innymi przez następujące firmy:

- Tanberg
- Polycom
- Mainconcept

Obwody scalone dedykowane do Set-Top-Box'ów oraz Set-Top-Box'y dostarczają między innymi następujące firmy:

- Broadcom BCM7411
- Conexant CX2418X
- Sigma Designs SMP8630, EM8622L, and EM8624L
- STMicroelectronics STB7100, NOMADIK (STn 8800/8810/8815 series)
- WISchip DeCypher 8100
- ADB ADVANCED DIGITAL BROADCAST

Firma ADP Polska jest gotowa do rozpoczęcia natychmiast masowej produkcji dekodерów multisystemowych MPEG-2 i AVC/H.264 a produkcji dekodерów jednosystemowych AVC/H.264 za pół roku. Firma RWT Radom produkująca dekodery MPEG-2 będzie gotowa do produkcji dekodерów AVC/H.264 za rok. Wszystkie firmy wstrzymują się z podjęciem produkcji do czasu zapadnięcia decyzji wyboru standardu w Polsce.



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

7. Wnioski

1. Standard AVC/H.264 zapewnia przy takiej samej subiektywnie ocenianej jakości obrazu uzyskanie dla jednego programu telewizyjnego prędkości bitowej o połowę mniejszej niż standard MPEG-2. Jest to powodowane podstawowymi różnicami rozwiązania pomiędzy tymi dwoma standardami
2. Obydwa standardy są oparte o rozwiązania softwerowe, które przy obecnym stanie techniki są co kilka lat wymieniane na nowocześniejsze. Biorąc to pod uwagę standard MPEG-2 staje się już przestarzały.
3. Metody kodowania w obydwu standardach MPEG-2 i MPEG-4 są metodami nieodwracalnymi tzn. informacje stracone w procesie kodowania zostają stracone bezpowrotnie i nie można ich odtworzyć, co ma wpływ na jakość odtwarzanych obrazów.
4. Standard AVC/H.264 wchodzi coraz szerzej do eksploatacji, szczególnie dla HDTV i telewizji płatnej (rozdz. 5)
5. Przy rozważaniu zawartości multipleksu należy uwzględnić, że oprócz sygnałów wizyjnych należy przesyłać sygnały fonii towarzyszącej i danych. Sygnały fonii stereofonicznej wymagają prędkości bitowej 192 kb/s. Fonia w systemie Dolby Surround wymaga prędkości 256 kb/s a fonia AC3 – prędkości 448 kb/s. Ponadto w multipleksie przesyłane są sygnały teletekstu o prędkości średnio 200 kb/s, sygnały EPG o prędkości około 500 kb/s oraz inne usługi dodane (np. aplikacje interaktywne) co najmniej kilka set KB/s do 1Mb/s. Dodatkowo trzeba uwzględnić, że w multipleksie należy przesyłać:
 - jednocześnie fonię dwukanałową i fonię AC3, dla kilku programów
 - programy radiowe Polskiego RadiaProgram TVP3, ze względu na konieczność transmisji regionalnych bloków programowych powinien być przesyłany poza multipleksem statystycznym, co obniża możliwości kompresji multipleksu.
6. Na przykład przy multipleksie 20 Mb/s i przesyłaniu sygnałów kodowanych MPEG-2 zakładamy przesyłanie 4 sygnałów wizyjnych o prędkościach średnio



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

4 Mb/s co zajmuje 16 Mb/s. Sygnały foniczne i inne sygnały nadawane obligatoryjnie zajmą około 2 Mb/s. Na przesłanie piątego programu o zadowalającej jakości nie ma już miejsca. Przy przesyłaniu sygnałów wizyjnych kodowanych w standardzie AVC/H.264 w takim multipleksie można przesłać około dwukrotnie więcej sygnałów wizyjnych.

Na podstawie rozważań podanych w rozdziałach 1, 2 i 3 ustosunkowując się do pytań p. Przewodniczącej KRRiT zawartych w piśmie KRRiT-073-14/053/05-4 z dn. 9 listopada 2005r. można stwierdzić, że:

ad 1. Potwierdzona wynikami badań, przy transmisjach eksperymentalnych sygnałów SDTV kodowanych wg. standardu MPEG-2 przy jakości odtwarzanego obrazu równorzędnej jakości obrazu analogowego wynosi 3,5 Mb/s. Prędkość bitowa dla obrazów kodowanych wg. standardu MPEG-4 jest przy takiej samej jakości odtwarzanego obrazu jest w przybliżeniu dwukrotnie mniejsza.

ad 2. Nie jest jasne co KRRiT rozumie przez słowa: „bezpieczniejszy w sensie technicznym” – czy dotyczy to niezawodności, czy jakości odtwarzania, czy nowoczesności rozwiązania czy też zabezpieczenia płatnych programów. Jeśli chodzi o niezawodność i jakość oba standardy są równorzędne natomiast standard MPEG-2 w chwili obecnej staje się standardem przestarzałym, natomiast standard MPEG-4 w wersji AVC/H.264 zawiera nowoczesne rozwiązania.

ad 3. Kodowanie HDTV wymaga większej prędkości bitowej niż SDTV. Kodeki HDTV AVC/H.264 różnią się prędkością działania i pojemnością pamięci.

ad 4. Minimalne wartości prędkości bitowej dla standardu MPEG-2 i jakości SDTV zależą od treści przesyłanego obrazu i wynoszą około 3,5 Mb/s, natomiast dla obrazów krytycznych (o dużej zawartości ruchu) do 4,5 Mb/s. Dla standardu MPEG-4 wersja AVC/H.264, dzięki wprowadzeniu nowych technik kompresji uzyskuje się znaczące zmniejszenie wynikowego strumienia bitowego do 1,7 Mb/s – dla obrazów krytycznych 2,2 Mb/s przy subiektywnej ocenie obrazu równej 4.

ad 5. Stan dostępności chipsetów AVC/H.264 przedstawiono rozdziale 6.



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

ad 6. Niezawodność kodeków nie zależy od stosowanego standardu tylko od technologii wytwarzania chipsetów. Należy przypuszczać, że technologia wytwarzania chipsetów AVC/H.264 jest bardziej nowoczesna niż MPEG-2 (co najmniej o 8 lat młodsza).

ad 7. Według producentów dekoderek (np. firmy ADB) cena dekodera multistandardowego MPEG-2 i AVC/H.264 będzie w początkowym okresie większa o 30% od ceny dekodera MPEG-2. Cena dostępnego za pół roku dekodera jednostandardowego AVC/H.264 nie będzie wyższa niż dekodera MPEG-2.

Podsumowując można stwierdzić:

- Standard AVC/H.264 jest standardem przyszłościowym i jako taki znajduje szersze zastosowanie podczas gdy standard MPEG-2 staje się standardem przestarzałym i będzie stopniowo wypierany przez rozwiązania nowocześniejsze, dające znacznie szersze możliwości.
- Co raz większa liczba producentów sprzętu odbiorczego już produkuje lub rozpoczyna produkcję dekoderek AVC/H.264 (w tym również RWT Radom planuje rozpoczęcie dekodera AVC/H.264 w połowie roku 2006).
- Przewodniczący Grupy Studiów 6 (radiofonia i telewizja) ITU-R zaleca wykorzystanie standardu AVC/H.264 w telewizyjnych systemach emisyjnych
- Standard AVC/H.264 umożliwia przesyłanie w multipleksie dwukrotnie większej liczby programów telewizyjnych,

Ponieważ w Polsce nie jest jeszcze wprowadzana emisja cyfrowa wydaje się celowe wprowadzenie standardu AVC/H.264 do próbnych emisji celem oceny jego właściwości transmisyjnych.



Centralne Laboratorium Badawcze

Projekt badawczy finansowany ze środków działalności statutowej

109/05; 09300095

8. Literatura

- [1] ISO/IEC JTC1/SC 29/WG 11 „ISO/IEC 14496 10 Advanced Video Coding 3rd edition“, Redmond, lipiec 2004
- [2] T. Wiegand, G.J. Sullivan, G.Bjontegaard A. Luthura “Overview of the H264/AVC Video Coding standard”, *IEEE Trans. on Circuits and Systems for Video Technology* vol. 13, lipiec 2003
- [3] Rafał Lange, Adam Łuczak: Zawansowane narzędzia kompresji sekwencji wizyjnych w koderze standardu AVC/H.264“, *KKRRiT Kraków 15+17 czerwca 2005*
- [4] Sviatoslav Buniaev „HDTV In Russian Federation” IBU
- [5] Jean-Pierre Dardayrol.... „Television numerique: enjeux et perspectives en 2005” GTI Rapport nr II-B.2
- [6] An EBU „route map” to high definition (HD). EBU – Tech 3298. Geneva, 2004