

Zaawansowana kompresja cyfrowych sygnałów wizyjnych – standard AVC/H.264

MAREK DOMAŃSKI, TOMASZ GRAJEK, JAROSŁAW MAREK
Politechnika Poznańska, Zakład Telekomunikacji Multimedialnej i Radioelektroniki

Prezentujemy Państwu tekst będący przedłużeniem serii artykułów [1-4] opublikowanych w „Systemach Alarmowych” w latach 2000-2001. W artykułach tych przedstawiono popularne techniki i standardy kompresji danych reprezentujących obrazy. W szczególności opisano podstawową technikę kompresji obrazów ruchomych, czyli technikę hybrydowego kodowania sygnałów wizyjnych [3].

Już po opublikowaniu w „Systemach Alarmowych” tekstów o kompresji pojawiła się nowa grupa technik, które można nazwać zaawansowanymi technikami kompresji. Ten artykuł opisuje jedną z nich – technikę Zaawansowanego Kodowania Sygnałów Wizyjnych (Advanced Video Coding), która została znormalizowana w finalizowanej obecnie części 10. standardu MPEG-4 (ISO/IEC IS 14496-10 [5]) oraz w zaleceniu H.264 przyjmowanym przez Międzynarodową Unię Telekomunikacyjną (ITU).

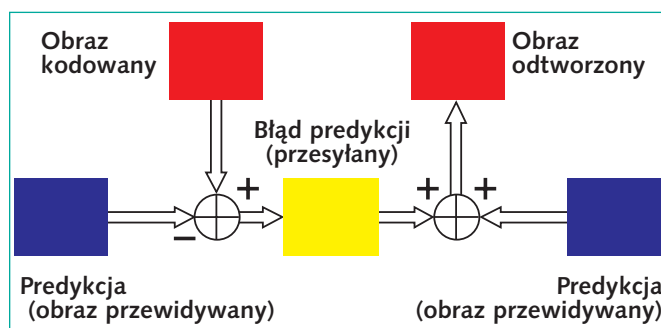
Klasyczne kodowanie hybrydowe sekwencji wizyjnych

Wszystkie praktycznie stosowane odmiany kompresji obrazu ruchomego wykorzystują tę samą podstawową metodę **hybrydowego kodowania sygnałów wizyjnych** (patrz też [3]).

Cechą tej metody jest duży stopień adaptacji algorytmów kompresji do lokalnych właściwości obrazu. W zasadzie we wszystkich praktycznie stosowanych technikach kompresji obrazu ruchomego przyjęto, że jednostką, w której dokonuje się adaptacji, jest **makroblok**. Makroblok to 256 próbek luminancji tworzących kwadrat o rozmiarze (16×16) punktów oraz odpowiednia liczba próbek chrominancji zajmujących ten sam obszar obrazu, co wspomniane 256 próbek luminancji. Dla schematu próbkowania $4 : 2 : 0$ [1], w którym na każdą próbkę chrominancji przypadają cztery próbki luminancji, makroblok składa się z 256 próbek lu-

minancji oraz z (8×8) próbek składowej chrominancji C_B i z (8×8) próbek składowej chrominancji C_R . Kodowany obraz jest dzielony na całkowitą liczbę makrobloków, które są rozdzielne, czyli nie nakładają się na siebie.

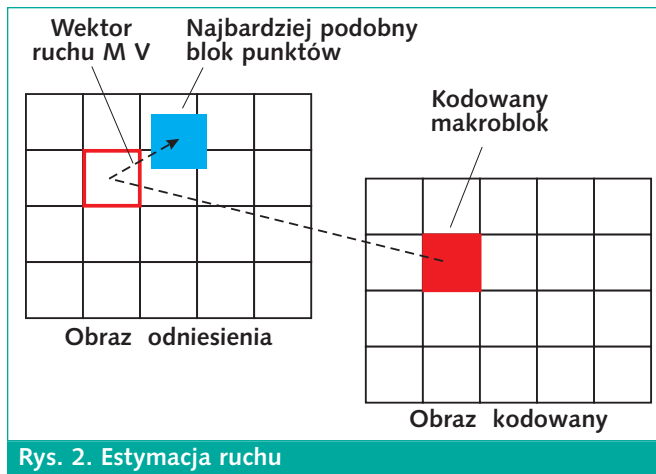
Bardzo istotnym elementem tej metody jest **predykcja międzyobrazowa** (przewidywanie obrazów). Koder (w nadajniku) i dekoder (w odbiorniku) przewidują treść kolejnego kodowanego obrazu na podstawie obrazów poprzednio zakodowanych i zdekodowanych. Predykcja jest wykonywana dokładnie tak samo w koderze i dekodерze. Dzięki temu dekoder może odtworzyć zakodowany obraz po otrzymaniu błędu predykcji, wyznaczonego w koderze jako różnica między oryginalnym obrazem kodowanym a jego predykcją (rys. 1). Zamiast przesyłania kolejnych obrazów przesyła się błędy predykcji, które przy dobrej predykcji zawierają mało informacji i mogą być zakodowane za pomocą niewielu bitów.



Rys. 1. Wykorzystanie predykcji międzyobrazowej

Dokładność predykcji ma więc kluczowe znaczenie dla uzyskania silnej kompresji. We współczesnych koderach stosuje się **predykcję z kompensacją ruchu**. Najpierw przez porównanie składowych luminancji kolejnych obrazów w koderze wyznacza się ruch poszczególnych fragmentów obrazu. Informację o ruchu zapisuje się w postaci **wektorów ruchu** reprezentujących przesunięcie fragmentu obrazu w czasie dzielącym wyświetlanie tych obrazów. W procesie **estymacji ruchu** wektory ruchu dla każdego makrobloku kodowanego obrazu wyznacza się przez poszukiwanie najbardziej podobnego do tego makrobloku bloku punktów luminancji w **obrazie odniesienia**, czyli w zdekodowanym obrazie wykorzystywanym w predykcji

międzyobrazowej. Przesunięcie między położeniem makrobloku a najbardziej podobnym blokiem w obrazie odniesienia wyznacza wektor ruchu (rys. 2). W praktyce za parę najbardziej podobnych bloków punktów luminancji przyjmuje się tę, dla której suma wartości bezwzględnych różnic we wszystkich odpowiednich punktach jest najmniejsza. Pomimo tak prymitywnego zdefiniowania miary podobieństwa bloków punktów, estymacja ruchu jest najbardziej pracochłonną operacją w koderze i może wymagać większej liczby operacji, niż wszystkie pozostałe operacje w algorytmie kompresji razem wzięte.



Rys. 2. Estymacja ruchu

Wektory ruchu, czyli informację o ruchu przesyła się do dekodera. W ten sposób ta informacja jest dostępna w procesie predykcji zarówno w koderze, jak i w dekodерze. Sama predykcja z kompensacją ruchu polega na przewidywaniu treści kolejnego obrazu, z uwzględnieniem ruchu poszczególnych fragmentów obrazu. Przewidywana treść obrazu kodowanego jest tworzona z fragmentów obrazów odniesienia, czyli wybranych obrazów, które już poprzednio zostały zakodowane i zdekodowane, a więc są dostępne zarówno w koderze, jak i w dekodерze. Te fragmenty obrazów odniesienia są przemieszczane zgodnie z wektorami ruchu i w ten sposób układa się predykcję obrazu kodowanego (rys. 2). Należy podkreślić, że wektory ruchu wskazują dla każdego makrobloku blok najbardziej podobnych próbek luminancji i niekoniecznie reprezentują faktyczny ruch, chociaż bardzo często rzeczywiście modelują ruch w scenie wizyjnej.

Koder przesyła do dekodera dwie zasadnicze grupy danych:

- informację o błędzie predykcji,
- wektory ruchu.

Błąd predykcji jest też obrazem, chociaż jego treść nie przypomina treści obrazów scen naturalnych. Błąd predykcji jest poddawany kompresji techniką podobną do stosowanej w standardzie JPEG [2]:

- makrobloki dzieli się na bloki o rozmiarze (8×8) punktów jednej składowej (luminacji lub chrominancji),
- dla każdego bloku wyznacza się jego widmo za pomocą szybkiego przekształcenia kosinusowego (*DCT – Discrete Cosine Transform*),

- próbki widma są kwantowane za pomocą kwantyzatorów dobranych indywidualnie do poszczególnych częstotliwości przestrzennych,
- niezerowe próbki widma są kodowane przez zapisanie ich położenia i wartości,
- położenia i wartości niezerowych próbek widma są efektywnie reprezentowane za pomocą kodowania Huffmana lub kodowania arytmetycznego.

Dekoder odtwarza błąd predykcji, składając poszczególne bloki z kosinusoid o częstotliwościach określonych przez położenie próbki widma wewnątrz bloku i o amplitudach zdefiniowanych przez wartości wspomnianych niezerowych próbek widma. Próbki widma stanowią zazwyczaj największą część strumienia binarnego, reprezentującego cyfrowy sygnał wizyjny. Typowy udział próbek widma błędu predykcji w całości strumienia to 60 – 70%.

Strumień binarny zawiera ponadto różne informacje pomocnicze oraz wektory ruchu, których składowe również są poddawane kompresji za pomocą kodowania Huffmana lub kodowania arytmetycznego.

Obrazy zakodowane w omówionym wyżej trybie międzyobrazowym nazywa się **obrazami typu P**, jeżeli były kodowane z wykorzystaniem jednego tylko obrazu, który poprzedza obraz kodowany w kolejności akwizycji i prezentacji. Proces predykcji międzyobrazowej może wykorzystywać także obraz, który co prawda będzie prezentowany po obrazie kodowanym, ale został zakodowany i przesłany przed nim. Wymaga to oczywiście zróżnicowania kolejności kodowania i kolejności prezentacji. Niektóre obrazy bywają kodowane jako **obrazy typu B**, czyli obrazy kodowane z wykorzystaniem dwóch obrazów odniesienia – jednego prezentowanego przed obrazem kodowanym i jednego prezentowanego w sekwencji wizyjnej po obrazie kodowanym. Dla makrobloków w obrazach typu B najbardziej podobne bloki próbek są wybierane z obu obrazów odniesienia albo są wyznaczone jako średnie z bloków z obu obrazów odniesienia.

Przedstawionego powyżej **kodowania międzyobrazowego** z wykorzystaniem predykcji z kompensacją ruchu nie można zastosować w pierwszym obrazie sekwencji, ani w żadnym innym obrazie, od którego rozpoczyna się dekodowanie. Dla takich obrazów nie można utworzyć predykcji międzyobrazowej i dlatego takie obrazy koduje się w **trybie wewnątrzobrazowym**, czyli w taki sam sposób jak obrazy nieruchome, a więc podobnie jak sam błąd predykcji. Takie obrazy są nazywane **obrazami typu I**.

Dla niektórych makrobloków obrazów typu P lub B też nie daje się utworzyć dobrej predykcji. Dotyczy to np. obszarów tła odślanianych przez poruszający się obiekt. Podobnie trudno znaleźć dobrą predykcję na brzegach poruszających się obiektów, dla obiektów wirujących itp. Dla odpowiednich makrobloków też stosuje się wtedy kodowanie wewnątrzobrazowe i takie makrobloki są nazywane **makroblokami typu I**.

Standardy kompresji

Dla zapewnienia kompatybilności koderów i dekoderów pochodzących od różnych producentów konieczna jest standaryzacja semantyki i składni strumienia binarnego, produkowanego przez koder i interpretowanego przez dekoder.

Należy podkreślić, że standardy kompresji nie definiują działania kodera, a tylko zasadę jednoznacznego dekodowania strumienia binarnego. Różne kodery tego samego standardu mogą się znacznie różnić efektywnością kompresji w zależności od zastosowanych algorytmów estymacji ruchu, dokładności wyznaczania wektorów ruchu, zakresu poszukiwania wektorów ruchu, algorytmu wyboru kwantyzatorów w poszczególnych makroblokach, algorytmu wyboru wariantu kodowania indywidualnych makrobloków itp.

Kilka standardów kompresji zostało przyjętych przez Międzynarodową Organizację Normalizacyjną (ISO – Interna-

tional Organization for Standardization). Te standardy zostały opracowane przez grupy robocze MPEG (*Moving Picture Experts Group*) lub JPEG (*Joint Photographic Experts Group*). Ponadto kompresję obrazów regulują odpowiednie zalecenia Międzynarodowej Unii Telekomunikacyjnej (ITU – *International Telecommunication Union*). Uproszczony przegląd podstawowych standardów przedstawiono w tabeli 1.

Dalsze nasze rozważania będą dotyczyć standardu AVC (*Advanced Video Coding*) stanowiącego część 10. standardu MPEG-4 (ISO/IEC IS 14496), który już w części 2. definiuje kompresję obrazu klasyczną techniką, bardzo bliską technice H.263+, uzupełnioną jednak o narzędzia do kodowania obiektów multimedialnych o dowolnych kształtach oraz liczne narzędzia służące do reprezentacji obiektów hybrydowych naturalnych i syntetycznych. Należy zaznaczyć, że techniki kompresji zdefiniowane w częściach 2. i 10. standardu MPEG-4 są niekompatybilne pod względem składni strumienia binarnego.

Grupa standardów	Standardy	Charakterystyczne cechy
Standardy kompresji wewnątrzobrazowej (tylko obrazy I)	M-JPEG (ISO IS 10918-1, ITU-T Rec. T.81) M-JPEG 2000 (ISO IS 15444-3) DV (IEC 61834) DVPRO (SMPTE 306M) CCIR 721 (ITU-T Rec. J.80)	mała efektywność kompresji (często mniej niż 1:10), proste kodery, indywidualne dekodowanie obrazów sekwencji, bezpośredni dostęp do każdego obrazu
Standardy kompresji pierwszej generacji	H.120 (ITU-T Rec. H.120) CCIR 723 (ITU-T Rec. J.81)	proste techniki z wykorzystaniem kodowania międzyobrazowego o małej efektywności kompresji, już nie są stosowane
Standardy kompresji drugiej generacji	H.261 (ITU-T Rec. H.261) MPEG-1 (ISO IS 11172-2)	ograniczony wybór trybów kodowania makrobloków, proste kodowanie Huffmana, jeden wektor ruchu na makroblok typu P, obecnie już przestarzałe
Standardy kompresji trzeciej generacji	MPEG-2 (ISO IS 13818-2, ITU-T Rec. H.262) H.263 / H.263+ (ITU-T Rec. H.263) MPEG-4 (ISO IS 14496-2) DivX	pewna poprawa efektywności kompresji i znaczne poszerzenie właściwości funkcjonalnych, zwiększony wybór trybów kodowania makrobloków, opcje pozwalające na kodowanie sygnałów z wybieraniem międzyliniowym (z przeplotem), zwiększony wybór kwantyzatorów, alternatywne warianty kodowania Huffmana, a czasem także kodowania arytmetycznego, opcjonalnie więcej niż jeden wektor ruchu na makroblok typu P, obecnie dominujące techniki (telewizja cyfrowa, płyty DVD, Internet)
Standardy kompresji czwartej generacji	AVC (ISO DIS 14496-10, ITU-T Draft Rec. H.264) Windows Media 9/10	wiele usprawnień prowadzących do znacznej poprawy efektywności kompresji, techniki obecnie wprowadzane do szerokiego zastosowania

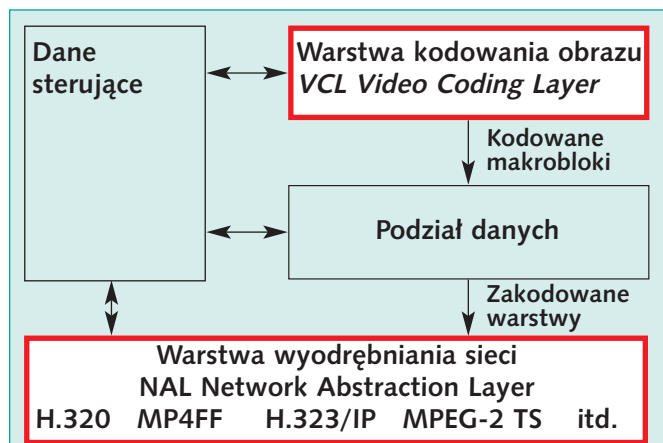
Tabela 1. Podstawowe standardy kompresji cyfrowych sygnałów wizyjnych

Technika kompresji zdefiniowana w standardzie AVC (część 10. standardu MPEG-4) także wykorzystuje omówioną wcześniej metodę kodowania hybrydowego z międzyobrazową predykcją z kompensacją ruchu, ale z licznymi istotnymi nowymi ulepszeniami i modyfikacjami, które spowodowały znaczącą poprawę efektywności kompresji w stosunku do powszechnie stosowanych dotychczas metod.

Koder AVC

W odróżnieniu od istniejących standardów, w AVC wyraźnie rozrózono dwie części (rys. 3):

- warstwę kodowania obrazu VCL (*Video Coding Layer*), która jest odpowiedzialna za wytwarzanie efektywnej reprezentacji obrazu,
- warstwę wyodrębniania sieci NAL (*Network Adaptation Layer*), która formatuje reprezentację VCL i zapewnia dodatkowe informacje potrzebne dla różnych systemów transportowych i nośników danych.



Rys. 3. Struktura kodera AVC (według [7])

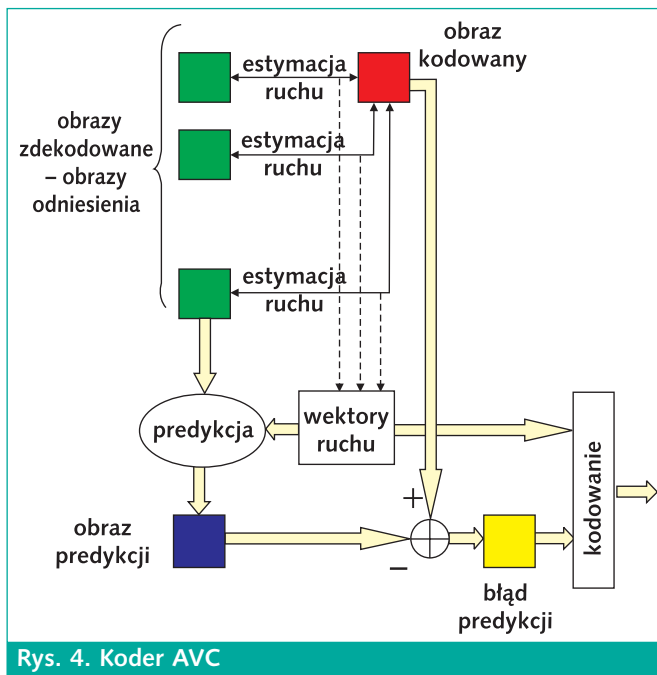
Warstwa wyodrębniania sieci rozmieszcza dane wizyjne i dane sterujące w pakietach NAL, których struktura może być dostosowana do specyfiki poszczególnych nośników danych lub sieci telekomunikacyjnych (RTP/IP, format plików MP4, protokoły H. 32x, strumień transportowy MPEG-2 itd.).

Algorytm kompresji AVC

W warstwie kodowania obrazu wprowadzono wiele ulepszeń i modyfikacji. Poniżej wymienimy niektóre z nich. Odróżniają one technikę AVC od starszych standardów wykorzystujących również hybrydowe kodowanie z międzyobrazową predykcją z kompensacją ruchu. Należy jednak podkreślić, że ogólna struktura algorytmu kompresji jest podobna do przedstawionej powyżej (rys. 4).

Następujące usprawnienia poprawiły efektywność predykcji z kompensacją ruchu:

- **Zmienny rozmiar bloków dla kompensacji ruchu.** Kompensacji ruchu można dokonywać także dla części makrobloków, którym wtedy przypisywane są osobne wektory ruchu. Najmniejsze bloki dla kompensacji ruchu mają rozmiar (4 × 4) punkty luminancji. W ten sposób kompen-

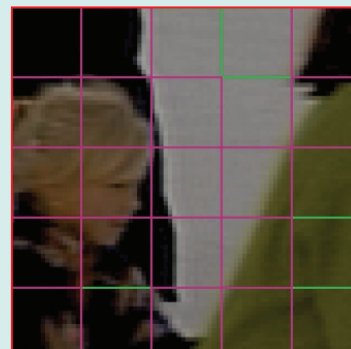


Rys. 4. Koder AVC

sacja ruchu może być lepiej dostosowana do kształtu poruszających się obiektów (rys. 5 i 6). Dzięki temu uzyskuje się mniejsze błędy predykcji, które mogą być przedstawione za pomocą niewielu bitów.

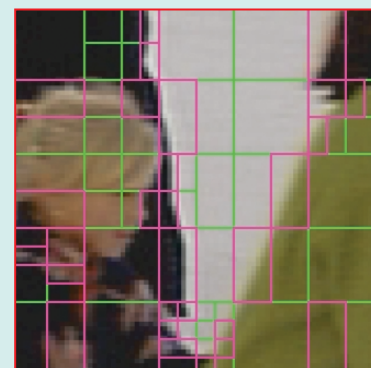
Rys. 5. Klasyczna kompensacja ruchu dokonywana makroblokami o rozmiarze (16 × 16) punktów luminancji np. w standardzie MPEG-2

Liniami fioletowymi zaznaczono makrobloki na granicach ruchomych obiektów. Te makrobloki zostaną prawdopodobnie zakodowane kosztowną techniką kodowania wewnątrzobrazowego.



Rys. 6. Standard AVC: Elastyczny podział makrobloków na małe bloki dla kompensacji ruchu

Fioletowymi liniami zaznaczono bloki na krawędziach obiektów. Kodowanie tych bloków jest znowu bardziej kosztowne, ale teraz takie bloki zawierają mniej punktów. Większą część powierzchni obrazu zajmują zaznaczone na zielono bloki, dla których predykcja z kompensacją ruchu może być przeprowadzona z bardzo małymi blokami predykcji.



• **Estymacja ruchu z dokładnością do 1/4 odstepu próbkowania obrazu.** Duża dokładność wektorów ruchu wymaga znacznie bardziej złożonej estymacji ruchu, ale pozwala znowu na bardziej precyzyjną predykcję z kompensacją ruchu.

• **Wykorzystywanie wielu obrazów odniesienia** pozwala na wykorzystanie pamięci długookresowej, przydatnej w predykcji obszarów odślawianych.

• **Tryb „bezpośredni” (*direct*)** pozwala na predykcję z kompensacją ruchu dla bloków, dla których nie przesyła się wektorów ruchu.

• **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. Ponadto wprowadzono jeszcze inne ulepszenia algorytmu kompresji:

• **Kierunkowa predykcja wewnątrzobrazowa dla makrobloków kodowanych wewnątrzobrazowo.** Bardzo kosztowne kodowanie makrobloków typu I zostało usprawnione dzięki wykorzystaniu wzajemnej korelacji sąsiednich bloków punktów.

• **Transformacja operująca na krótkich słowach.** Dotychczas stosowana transformacja kosinusowa została przybliżona transformacją o współczynnikach zapisywanych za pomocą niewielu bitów. Umożliwia to realizację transformacji na stałoprzecinkowych procesorach 16-bitowych i pozwala na łatwe zapewnienie dokładnego dopasowania transformacji prostej i odwrotnej.

• **Transformacja na małych blokach.** Dotychczas stosowana transformacja kosinusowa operowała na blokach o rozmiarze (8×8) próbek luminancji lub chrominancji. Transformacja stosowana w AVC operuje na małych blokach (4×4) próbek luminancji oraz (2×2) próbek chrominancji, co pozwala na lepszą adaptację do lokalnych właściwości obrazu. Ponadto dla składowych stałych można dokonywać jeszcze powtórnej transformacji redukującej wzajemną korelację składowych stałych sąsiednich bloków.

• **Adaptacyjne kodowanie entropijne.** Dwie metody adaptacyjnego kodowania entropijnego pozwalają na modelowanie rozkładów prawdopodobieństw warunkowych i mogą być alternatywnie stosowane w koderach AVC:

CAVLC (*Context-Adaptive Variable Length Coding*) – adaptacyjne kodowanie ze zmienną długością słowa,

CABAC (*Context-Based Binary Arithmetic Coding*) – bardziej skomplikowane adaptacyjne kodowanie arytmetyczne, pozwalające na uzyskanie silniejszej kompresji.

Profile i poziomy koderów i dekodeków AVC

Podobnie jak dla koderów MPEG-2 i MPEG-4, standard AVC definiuje pewne profile określające podzbiory składni strumienia binarnego, odpowiednie dla specyficznych zastosowań. Obecnie [5] standard definiuje siedem różnych profili. Podobnie jak w MPEG-2, zdefiniowano profil główny (*main profile*) do zastosowań telewizyjnych.

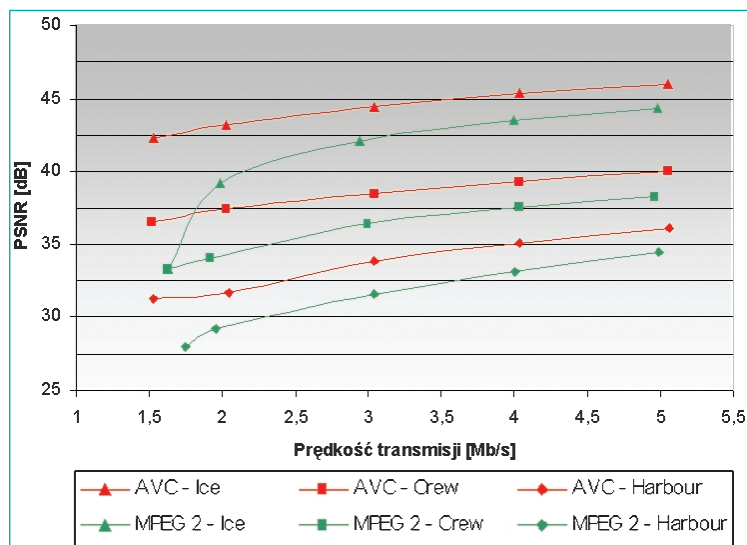
Ponadto definiuje się również poziom kodera lub dekodeka. Poziom określa m.in. maksymalną liczbę makrobloków na sekundę, maksymalną objętość obrazu (liczoną w bajtach), maksymalny rozmiar bufora danych reprezentujących obraz, maksymalną prędkość transmisji i maksymalny zakres wartości składowej pionowej wektorów ruchu. Standard [5] definiuje obecnie 17 różnych poziomów.

Efektywność kompresji i porównanie z innymi technikami

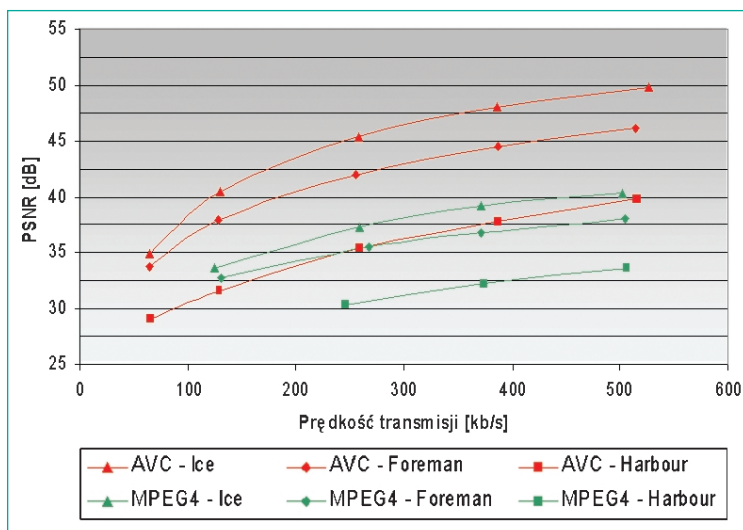
Kodery AVC pozwalają na uzyskiwanie dla większości sekwencji ponad dwukrotnie mniejszych prędkości transmisji w porównaniu ze współczesnymi koderami MPEG-2, wytwarzającymi strumienie binarne reprezentujące obraz o podobnej jakości [7]. W Zakładzie Telekomunikacji Multimedialnej i Radioelektroniki Politechniki Poznańskiej przeprowadzono badania porównawcze, w których badano implementację programową kodera MPEG-2 oraz referencyjne oprogramowanie AVC przygotowane przez komitet MPEG. Otrzymane wyniki (rys. 7) również wskazują na ponad dwukrotnie mniejszą prędkość transmisji uzyskiwaną dzięki stosowaniu kodera AVC.

W badaniach tych jakość mierzono za pomocą współczynnika PSNR [1], który jest miarą błędu średniokwadratowego, wprowadzanego przez kompresję. Wartości współczynnika PSNR powyżej 40 dB wyrażają bardzo dobrą jakość obrazu pozbawionego wyraźnych zakłóceń. Natomiast wartości współczynnika PSNR poniżej 30 dB wskazują na mierną jakość obrazu.

Badania porównawcze w Zakładzie Telekomunikacji Multimedialnej i Radioelektroniki Politechniki Poznańskiej przeprowadzono także dla pary koderów AVC i MPEG-4 (koder zgodny



Rys. 7. Porównanie efektywności kompresji AVC i MPEG-2 dla trzech sekwencji wizyjnych (standardowa europejska rozdzielczość telewizyjna, 4:2:0, 30 obraz/s)



Rys. 8. Porównanie efektywności kompresji AVC i MPEG-4 (część 2. standardu) dla trzech testowych sekwencji wizyjnych (rozdzielczość QCIF (176 × 144) punktów luminancji, 4 : 2 : 0, 15 obrazów/s)

z częścią 2. standardu, dostępny w Internecie koder firmy Microsoft wersji 3). Również w tym przypadku uzyskane wyniki (rys. 8) wskazują na znacznie lepszą efektywność kompresji za pomocą koderów AVC.

Podobne wyniki można znaleźć też w literaturze [6, 7], która podaje, że dla szerokiego zakresu prędkości transmisji i rozdzielczości obrazu zastosowanie koderów AVC zamiast koderów H.263+ lub MPEG-4 (według części 2. standardu) pozwala na redukcję prędkości transmisji o kilkadziesiąt procent (zazwyczaj o 30 – 70%), przy zachowaniu tej samej jakości obrazu.

Podsumowując można stwierdzić, że przy zastosowaniu koderów AVC prędkość transmisji około 1,5 Mb/s zazwyczaj wystarcza do uzyskania dobrej jakości standardowego obrazu telewizyjnego. To oznacza, że zwykłe łącza telefoniczne z wykorzystaniem popularnej techniki ADSL, a także łącza UMTS mogą być stosowane do przesyłania dobrego sygnału telewizji cyfrowej.

Dla sekwencji QCIF [(176 × 144) punkty luminancji, 15 obrazów/s] uzyskanie bardzo dobrej jakości wymaga prędkości transmisji od około 30 kb/s (dla sekwencji z jedną siedzącą osobą mówiącą do kamery) do około 400 kb/s (dla sekwencji o skomplikowanej treści i znacznym ruchu).

Kodowanie skalowalne

Kodowanie skalowalne umożliwia dekodowanie pewnej części strumienia. Odebranie tylko części strumienia binarnego pozwala na zdekodowanie przynajmniej obrazu o gorszej jakości i/lub mniejszej rozdzielczości. Jeżeli w sieci zrealizowano pewną hierarchię pakietów, to w pewnych sytuacjach do odbiornika docierają tylko nieuszkodzone pakiety o wyższym priorytecie, pozwalające na zdekodowanie obrazu o gorszej jakości.

Obecnie standard AVC nie określa kodowania skalowalnego, jednak w wielu ośrodkach, w tym także w Zakładzie Telekomunikacji Multimedialnej i Radioelektroniki Politechniki Poznańskiej trwają koordynowane przez komitet MPEG prace mające na celu ustalenie standardu kodowania skalowalnego, rozwijającego standard AVC.

Uwagi końcowe

Standard AVC pozwala na istotną redukcję prędkości transmisji. Pełne wykorzystanie możliwości standardu AVC wymaga dokonywania w koderze optymalnych wyborów między licznymi możliwymi trybami kodowania. Określenie optymalnych algorytmów sterowania koderem, a więc algorytmów wyboru trybów i wartości parametrów, jest jeszcze przedmiotem badań naukowych w wielu ośrodkach badawczych na świecie. Zwłaszcza kodery AVC są bardzo złożone i wykonują znacznie więcej operacji niż popularnie stosowane dotychczas kodery. Natomiast złożoność dekodatorów jest w stosunku do złożoności koderów przynajmniej o rząd wielkości mniejsza. Złożoność dekodera AVC jest niewiele większa od złożoności obecnie stosowanych dekodatorów.

Prace przeprowadzone w Zakładzie Telekomunikacji Multimedialnej i Radioelektroniki Politechniki Poznańskiej pokazują, że procesor Pentium IV z zegarem 3 GHz może realizować uproszczony koder AVC dla sygnałów telewizyjnych, jeżeli krytyczne fragmenty zostały przygotowane jako programy niskiego poziomu, wykorzystujące funkcje SSE procesora Pentium IV.

Obecnie wiele firm pracuje nad przygotowaniem układów scalonych realizujących kodery i dekodery AVC.

Zastosowanie kodowania AVC może mieć w przyszłości duże znaczenie dla systemów nadzoru wizyjnego, gdyż zmniejszenie prędkości transmisji obrazów wydatnie obniży koszty systemów nadzoru. Kodery umieszczone w obudowach inteligentnych kamer pozwolą na bardzo efektywną transmisję sygnałów wizyjnych w systemach rozproszonych. Zastosowanie efektywnego kodowania AVC może się przyczynić do rewizji poglądów dotyczących konieczności budowy sieci szerokopasmowych (np. światłowodowych).

Literatura

- [1] M. Domański, *Kompresja cyfrowych sygnałów wizyjnych w systemach nadzoru, Część I. Wprowadzenie*. Systemy Alarmowe nr 1, 2000, s. 26-29
- [2] Jak wyżej, *Część II. JPEG i M-JPEG*. Systemy Alarmowe nr 2, 2000, s. 18-22
- [3] Jak wyżej, *Część III. MPEG*. Systemy Alarmowe nr 5, 2000, s. 33-42
- [4] Jak wyżej, *Część IV. Kodowanie falkowe*. Systemy Alarmowe nr 2, 2001, s. 28-31
- [5] ISO/IEC JTC1/SC29/WG11, Doc. N6540, Text of ISO/IEC 14496-10 *Advanced Video Coding, 3rd Edition*, July 2004
- [6] I. Richardson, *H.264 and MPEG-4 video compression*, Wiley 2003
- [7] *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, nr 7, July 2003, Special issue on the H.264/AVC video coding standard