

Ewolucja czy rewolucja? Dwie dekady rozwoju metod kompresji wizji

Evolution or revolution?

Two decades of development of video compression methods

Metody hybrydowej kompresji wizji zrewolucjonizowały sposób zapisu oraz transmisji cyfrowego obrazu. Coraz wydajniejsza reprezentacja obrazu torowała drogę do szerokiego wdrożenia multimedialnych systemów kolejnych generacji, jak cyfrowa telewizja, kino domowe czy telewizja w Internecie. Jednak możliwości te nie pojawiły się z dnia na dzień. Są one wynikiem wielu lat badań, ulepszeń już istniejących rozwiązań, ale również wdrażania nowych pomysłów. W artykule przedstawiono skalę tego postępu w ciągu ostatnich dwóch dekad. Poddano ocenie kolejne generacje koderów obrazu, od MPEG-2 do nowej techniki HEVC, zarówno pod względem samych osiągnięć kompresji, jak również innych cech techniki, których nie sposób pominąć realizując jej wdrożenie.

Słowa kluczowe: kompresja danych, kodery wizyjne, kodowanie obrazu, MPEG-2, AVC, HEVC



Hybrid video compression has revolutionized the way of a video saving and transmission. More and more efficient representation of a video paved the way for the broad implementation of multimedia systems of successive generations, such as digital television, home cinema or television over the Internet. However, these opportunities did not appear from day to day. They are the result of many years of research, improvements of existing solutions, but also the implementation of new ideas. The paper presents the scale of this progress over the last 2 decades. It evaluates the successive generations of video encoders, from MPEG-2 to the new HEVC technology, both in terms of its compression efficiency, as well as other technical features that can not be omitted while putting the technology into the market.

Key words: data compression, video encoders, image coding, MPEG-2, AVC, HEVC

Transmisja czy magazynowanie obrazu ruchomego w jego oryginalnej postaci wymagają operowania na olbrzymich wręcz zbiorach danych. W przypadku danych, które opisują sekwencję obrazów o wysokiej jakości (rozdzielczość przestrzenna obrazu równa 1920 x 1080 przy 60 obrazach na sekundę), związany z nimi strumień bitowy to aż 3 Gbit/s! Jest zatem oczywiste, że praca z tak potężnym strumieniem bitów byłaby w praktyce bardzo kosztowna, a w niektórych zastosowaniach po prostu technicznie niemożliwa.

Obrazy sekwencji należy więc wcześniej poddać kompresji. To właśnie ona otwiera możliwość wydajnej transmisji i zapisu obrazu. Jej celem jest opisanie obrazów sekwencji w alternatywny, inny sposób, który nie wymaga aż tak dużych nakładów bitowych. Celem jest wielokrotna (w praktyce więcej niż sto razy) redukcja liczby bitów, które przedstawiają obrazy, przy zachowaniu ciągle zadowalającej ich jakości po kompresji i dekompresji. Ten cel przyswiera badaczom już od kilku dekad.

Z punktu widzenia tego, jak dzisiaj opisujemy obraz, szczególnie istotne są osiągnięcia ostatnich 2-3 dekad. W tym czasie rozwijano różne podejścia do wydajnej reprezentacji obrazu [1–7]. Jednak z perspektywy kompresji ruchomego obrazu największe znaczenie praktyczne zyskała jedna technika – kodowania hybrydowego z przewidywaniem treści obrazu [1–6]. W porównaniu z innymi technikami jej wydajność okazała się po prostu najlepsza.

Z tego właśnie powodu technika kompresji hybrydowej była przez lata przedmiotem licznych ulepszeń oraz modyfikacji. Kolejne odsłony tej technologii zostały ujęte w koderach obrazu kolejnych generacji, które stały się przedmiotem wielu rynkowych wdrożeń [1–6]. W artykule podjęto próbę oceny stopnia rozwoju, jakiego w ciągu ostatnich dwóch dekad doświadczyły kodery

ruchomego obrazu. Celem pracy było ponadto pokazanie różnicy w wybranych parametrach koderów kolejnych generacji. Co nas czeka w najbliższych latach? Podjęto próbę odpowiedzi również na to pytanie.

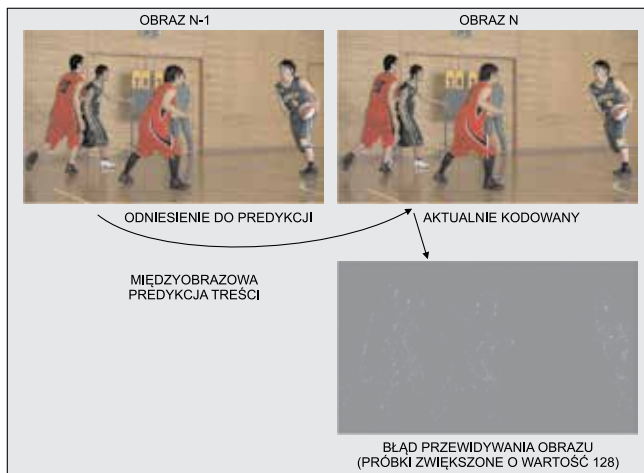
JAK DZIAŁA HYBRYDOWY KODER WIZYJNY?

Główna idea działania kodera hybrydowego jest bardzo prosta i jest konsekwentnie stosowana już od przeszło trzech dekad. Kodowanie treści obrazu realizuje się w kolejnych jego fragmentach, dzieląc wcześniej obraz na dużą liczbę niewielkich bloków. Koder nie koduje bezpośrednio treści każdego z bloków, ale stara się ją najpierw przewidzieć na podstawie innych danych, nazywanych danymi odniesienia, którymi dysponują zarówno koder, jak i dekodek obrazu.

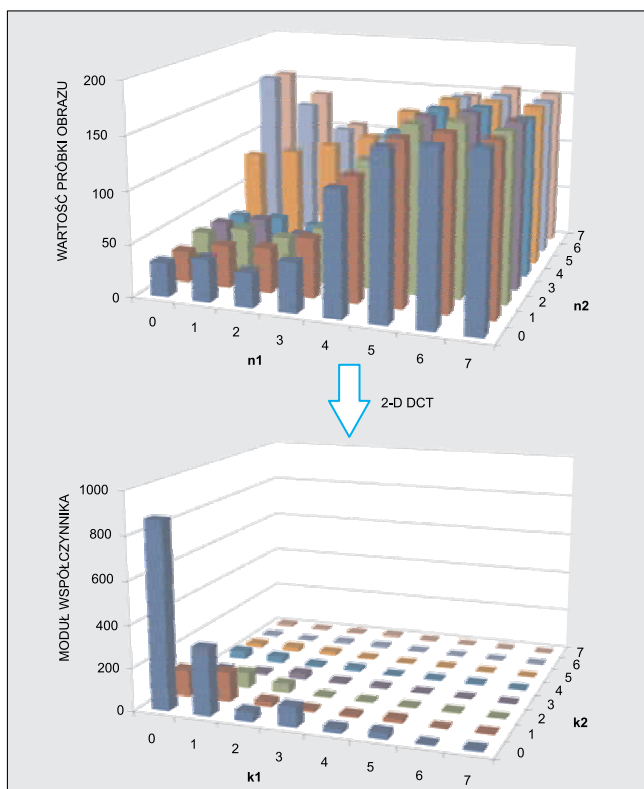
Są to więc dane bloków, które w kodowanym obrazie sąsiadują z aktualnie przetwarzanym blokiem albo dane, które pochodzą z zupełnie innych obrazów, czyli z innych chwil czasowych. Mowa jest więc tutaj odpowiednio o wewnątrz oraz międzyobrazowym przewidywaniu treści bloków. W praktyce metody te pozwalają przewidzieć treść bloków obrazu z bardzo dużą dokładnością. Co prawda, w ogólności, jest popełniany niezerowy błąd, jednak z reguły próbki błędów przyjmują bardzo małe wartości (rys. 1).

Ze względu na swoją prostotę to właśnie sygnał błędów przewidywania próbek jest przedmiotem dalszej kompresji. Opisuje się go zwykle za pomocą funkcji kosinusoidalnych o określonych częstotliwościach. Ten sposób reprezentacji sygnału błędów znany jest w literaturze jako dyskretne przekształcenie kosinusowe (*Discrete Cosine Transformation* – **DCT**) próbek błędów. Ponieważ istnieje silne podobieństwo wartości sąsiednich próbek błędów w bloku, już niewielka liczba niskoczęstotliwościowych składowych kosinusoidalnych umożliwia z dużą dokładnością reprezentowanie tego błędów (rys. 2). A mała liczba składowych przekłada się na małą liczbę bitów, którą zużyje koder do ich transmisji do dekodekera obrazu.

* Politechnika Poznańska, Wydział Elektroniki i Telekomunikacji, Katedra Telekomunikacji Multimedialnej i Mikroelektroniki, e-mail: dkarwow@multimedia.edu.pl



■ Rys. 1. Ilustracja skuteczności przewidywania treści kodowanego obrazu. Sygnał błędów przewidywania jest dużo prostszy w porównaniu z oryginalną treścią



■ Rys. 2. Blok próbek obrazu (lub blok próbek błędów predykcji treści) i jego reprezentacja za pomocą kosinusów. W przypadku takiej reprezentacji znaczące amplitudy mają tylko kosinusy niskoczęstotliwościowe

Jednak na tym możliwości wydajnej reprezentacji sygnału błędów jeszcze się nie kończą. Bardzo dużo można dodatkowo zyskać, jeśli ograniczyć się do dokładności, z jaką koder zapisuje w strumieniu danych informację o poszczególnych kosinusach. Tę dokładność ogranicza się, przeprowadzając kwantowanie amplitud składowych kosinusoidalnych. Operacja kwantowania, w przeciwieństwie do wcześniejszych operacji kodera, jest jednak przetwarzaniem stratnym. W toku tej operacji dochodzi do nieodwracalnej utraty części informacji o treści obrazu. Prowadzi to do nieuchronnej degradacji jego jakości. Ten spadek jakości można jednak minimalizować, realizując kwantowanie w taki sposób, który uwzględnia dobrze znane ograniczenia systemu widzenia człowieka w zakresie percepcji obrazu. Zgodnie z tą wiedzą kosinusy niskoczęstotliwościowe należy reprezentować dokładniej

(czyli realizować słabsze ich kwantowanie), a mniejszą dokładność opisu składowych można zastosować dla tych o wyższych częstotliwościach (czyli silniejszą kwantyzację ich amplitud). I tak to się właśnie czyni w praktyce. W przypadku każdej składowej koder realizuje dedykowany sposób kwantowania jej amplitudy.

Wynikiem dotychczas omówionych operacji jest tzw. sygnał resztkowy po kwantyzacji, przedstawiony w dziedzinie częstotliwości. Uwzględniając charakter próbek tego sygnału podlega on dalszemu kodowaniu z użyciem metod, które wydajnie reprezentują ciąg wartości. W ciągu tym niezerowe amplitudy kosinusów są przedzielone sekwencją zer o pewnej długości (np. skanowanie zig-zag współczynników przekształcenia DCT wraz z użyciem metody kodowania długości ciągów). Ponieważ finalne dane wykazują jeszcze statystyczną nadmiarowość, to ostatnią stosowaną metodą jest zawsze entropijne kodowanie danych.

EWOLUCJA TECHNIKI KODOWANIA HYBRYDOWEGO

Ten wydawałoby się dziecinnie prosty sposób działania kodera obrazu doczekał się w ciągu lat wielu rynkowych wdrożeń. Kodery MPEG-1 (koniec lat 80.), MPEG-2 (1994 rok), AVC (2003 rok), HEVC (2013 rok) – wszystkie one wykorzystują przedstawiony wcześniej schemat kodowania danych. Mogłoby się więc wydawać, że w podejściu do kodowania ruchomego obrazu nic się nie zmieniło lub niewiele, wszystko jest „po staremu”. Obserwując kodery z „lotu ptaka” można nawet tak powiedzieć. Jednak – jak to zwykle bywa – diabeł tkwi w szczegółach. Dokonując dokładniejszego porównania koderów, takiego na niższym poziomie, z łatwością można zobaczyć ogromne wręcz różnice w ich działaniu. Dotyczą one trzech fundamentalnych kwestii:

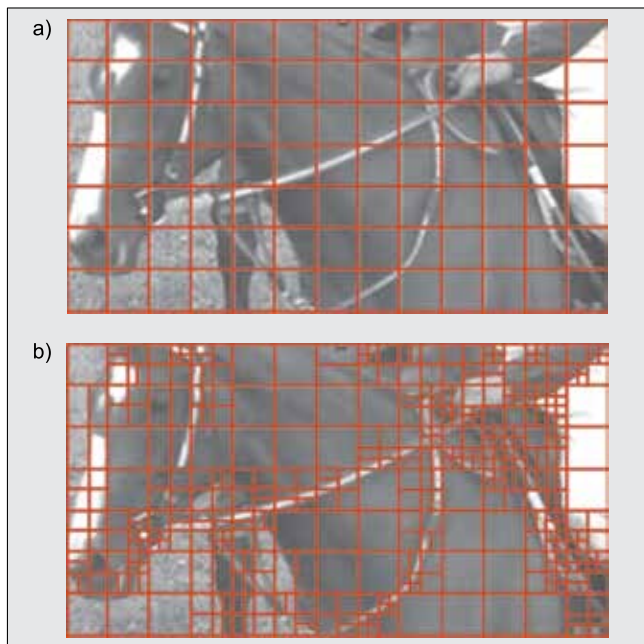
- realizowanego przez koder sposobu podziału obrazu na bloki, w których jest realizowana kompresja danych,
- łącznej liczby wszystkich trybów kodowania, którymi dysponuje koder obrazu,
- stosowanych przez koder rozwiązań w zakresie poprawy jakości obrazu po kompresji.

PODZIAŁ OBRAZU NA BLOKI

W przypadku starszych koderów, jak MPEG-2, rozmiar bloków, na które dzielony był obraz, pozostawał stały. Przewidywanie danych realizowano w blokach o wielkości 16x16 próbek składowej luminancji (i to tylko w kodowaniu międzyobrazowym), podczas gdy kodowanie transformacyjne – w blokach o rozmiarze 8x8. Kodery te nie miały więc żadnej zdolności adaptacji rozmiaru kodowanego bloku do charakteru treści, która podlegała kompresji (rys. 3). Dlatego podział z wykorzystaniem ustalonej z góry, równomiernej siatki bloków prowadził do słabszych wyników kompresji w niektórych fragmentach obrazu.

Wspomniana zdolność stała się cechą znaną dopiero nowszych koderów, jak AVC czy HEVC (rys. 3). Oprócz bloków 16x16 (dla predykcji próbek) oraz 8x8 (dla kodowania transformacyjnego) koder AVC wprowadził dodatkowe rozmiary bloków: 16x8, 8x16, 8x8, 8x4, 4x8, 4x4 na potrzeby predykcji danych oraz 4x4 dla potrzeb kodowania transformacyjnego. W tym obszarze jeszcze głębszą zmianę można zaobserwować w nowym koderze HEVC. Tutaj, w porównaniu z tym, co dotyczy kodera AVC, istnieją dodatkowo bloki o rozmiarach 64x64, 32x32, 16x16, 8x8, z dalszą możliwością ich podziału na dwa równe bloki (podział w kierunku pionowym lub poziomym), na cztery bloki, a z wyłączeniem bloków 8x8 również z niesymetrycznym podziałem bloku na dwa (w kierunku pionowym lub poziomym), w stosunku 1/4 i 3/4. Tak jest w przypadku samej predykcji próbek. Kodowanie transformacyjne kodera HEVC ma do dyspozycji bloki 4x4, 8x8, 16x16, 32x32.

Wprowadzone zmiany dały możliwość podziału obrazu z użyciem nierównomiernej siatki bloków, dostosowanej do treści kodowanego obrazu, poprawiając tym samym w sposób istotny efektywność kodowania danych.



■ Rys. 3. Stały (a) oraz zmienny (b) rozmiar bloków obrazu, stosowany odpowiednio w koderach starszych (np. MPEG-2) i nowszych (np. AVC czy HEVC)

LICZBA DOSTĘPNYCH TRYBÓW KOMPRESJI

Stały rozmiar bloków obrazu w kodekach starszych, jak np. kodek MPEG-2, przekłada się na bardzo ubogi zakres trybów kompresji, którym operuje koder obrazu. Widać to zarówno w przypadku kodowania wewnątrzobrazowego, jak i międzyobrazowego.

W trybie kompresji wewnątrzobrazowej starsze kodery w ogóle nie przewidywały treści kodowanych bloków. Kodowanie transformatowe było od razu realizowane na blokach próbek, a nie na blokach z próbkami błędu predykcji. Takie podejście znacznie uprościło działanie kodera, ale odbyło się to kosztem efektywności kompresji danych obrazu. Istotne ograniczenie występuje również w trybie kompresji międzyobrazowej. Tutaj przewidywanie treści odbywa się tylko w blokach 16x16, bez możliwości zastosowania bloków większych oraz mniejszych. Wszystkie te uproszczenia powodują, że liczba możliwych wariantów zakodowania treści fragmentu obrazu jest bardzo niewielka.

Odwrotną sytuację można zaobserwować w rozwiązaniach nowszych, np. w technikach AVC czy HEVC. Tutaj paleta dostępnych trybów kompresji bloków jest znacznie rozszerzona. W tych koderach kodowanie transformatowe w trybie wewnątrzobrazowym poprzedza już przewidywanie wartości próbek obrazu, co poprawia oczywiście osiągi kodera. Nie dosyć, że predykcja próbek może zostać przeprowadzona wieloma różnymi sposobami (maksymalnie 9 sposobów w koderze AVC i aż 35 sposobów w nowym koderze HEVC), to jeszcze koder decyduje o tym, w jak dużym bloku najlepiej jest ją zrobić (dla przypomnienia – są to bloki o rozmiarze 16x16, 8x8 i 4x4 w koderze AVC i bloki 64x64, 32x32, 16x16, 8x8 i 4x4 w koderze HEVC). Sumarycznie daje to więc ogromną wręcz liczbę możliwych kombinacji zakodowania fragmentu obrazu, np. o wielkości 64x64 próbki. Podobna sytuacja występuje również w trybie kompresji międzyobrazowej. Zakres rozmiarów bloków, w których można przewidywać dane, znacznie się zwiększyła – od jednego bloku 16x16 w koderze MPEG-2 do bloków 64x64, 32x32, 16x16, 8x8, z możliwością dalszego ich podziału na dwa (podział w pionie lub w poziomie na dwa bloki równe lub analogicznie, ale z podziałem na niesymetryczne bloki¹⁾

¹⁾ Podział na niesymetryczne bloki nie jest dozwolony w blok 8x8.

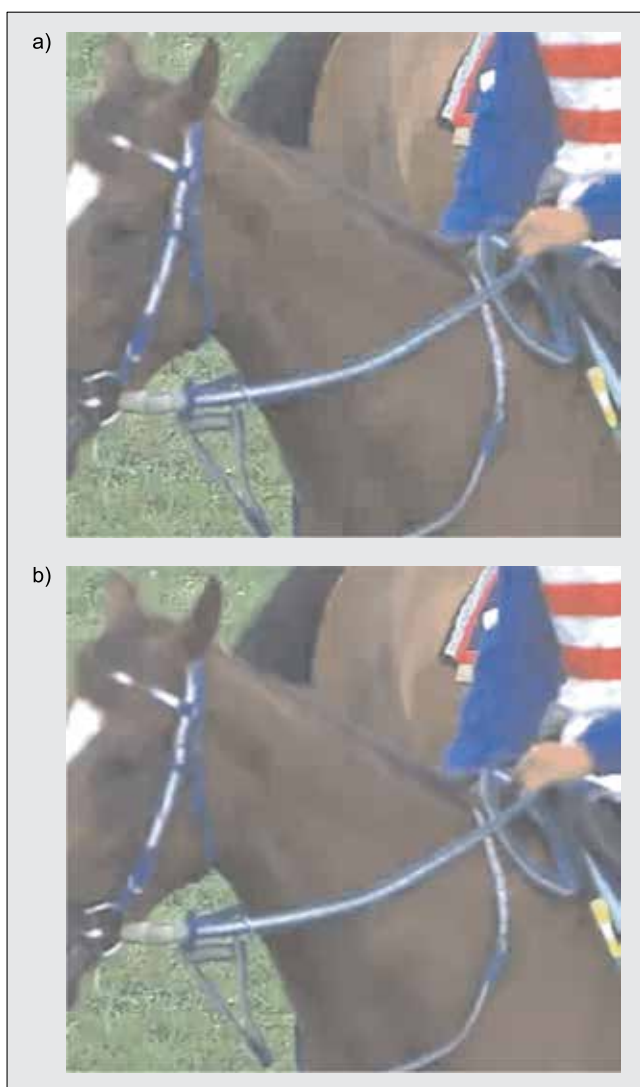
lub cztery jednakowe, mniejsze bloki w koderze HEVC. Podobnie jak w przypadku wewnątrzobrazowej kompresji daje to ogromną liczbę możliwych sposobów zakodowania fragmentu obrazu.

W nowszych koderach (AVC, HEVC) nowością są ponadto specjalne tryby kompresji, których nie sposób znaleźć w starszych koderach. Są to tryby SKIP oraz DIRECT, których zastosowanie umożliwia ekstremalnie wydajną kompresję tych fragmentów obrazu, dla których trafnie można przewidzieć odpowiednio dane o ruchu w sekwencji oraz treść kodowanego bloku (czyli błąd przewidywania wartości próbek jest zerowy) lub samą tylko informację o ruchu.

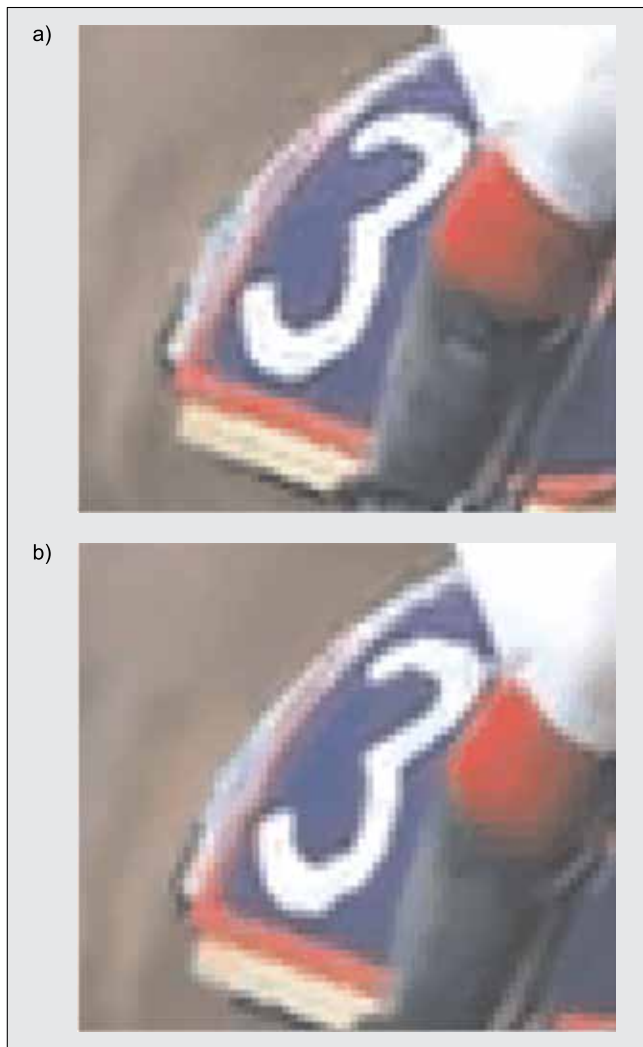
POPRAWA JAKOŚCI OBRAZÓW PO KOMPRESJI

Kwantowanie amplitud kosinusów może doprowadzić do wystąpienia w obrazie wielu widocznych zniekształceń, z których najbardziej znamienne to tzw. efekt blokowy oraz efekt dzwonienia.

Wskutek silnego kwantowania amplitud kosinusów, realizowanego niezależnie w poszczególnych blokach obrazu, na granicy tych bloków może dojść do sytuacji, w której wartości próbek zmieniają się gwałtownie, w sposób skokowy. W zakodowanym obrazie objawia się to pojawieniem charakterystycznych kwadratów. Dodatkowym typem zniekształcenia, który się w tej sytuacji pojawi, jest błędne odwzorowanie krawędzi obrazu, jako skutek usunięcia (lub silnego stłumienia) z częstotliwościowej reprezentacji sygnału wysokoczęstotliwościowych składowych kosinusoidalnych. Wokół krawędzi obrazu pojawiają się zafalowania, tętnienia. Opisywane zniekształcenia można zobaczyć na zamieszczonych dwóch rysunkach (rys. 4 i rys. 5).



■ Rys. 4. Ilustracja efektu blokowego w obrazie (a); rzeczywiste możliwości redukcji tego efektu (b)



■ Rys. 5. Ilustracja efektu „tętnień” wokół krawędzi obrazu (a); rzeczywiste możliwości redukcji tego efektu (b)

Zniekształcenia te mocno pogarszają jakość zakodowanych obrazów. Powodują również spadek efektywności samej kompresji – obrazy po kompresji, w których obecne są wspomniane zniekształcenia, bardziej różnią się od kodowanych obrazów, co pogarsza skuteczność przewidywania treści. Pomimo tego w starszych koderach obrazu, w tym MPEG-2, nie była podejmowana z tymi zniekształceniami żadna walka. Nie były podejmowane próby choćby drobnej ich redukcji.

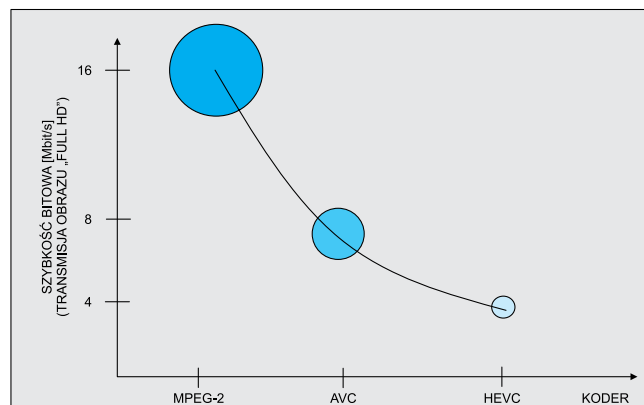
Inaczej jest w przypadku najnowszych technik kompresji. W koderze AVC zdefiniowano filtr, nazywany filtrem deblokującym, który ogranicza występowanie efektu blokowego. W koderze HEVC występuje dodatkowo inna filtracja, która ogranicza efekt dzwonięcia na krawędziach. Działanie każdego z tych dwóch filtrów znacznie poprawia subiektywną jakość obrazów i przyczynia się do poprawy rezultatów kompresji (w przypadku każdego z filtrów o kilka procent). Ujemną stroną użycia filtrów jest większa złożoność koderów i dekodera obrazu.

EWOLUCJA EFEKTYWNOŚCI REPREZENTACJI OBRAZÓW

Można powiedzieć, że z perspektywy efektywności kompresji danych w ciągu ostatnich 20 lat dokonała się rewolucja możliwości reprezentacji obrazu. Przeszło 20 lat wysiłków naukowców i inżynierów umożliwiło zmniejszenie objętości danych, które opisują obrazy, aż 4-krotnie, nie pogarszając przy tym jakości obrazów po kompresji!

I tak w przypadku legendarnego koderów MPEG-2 transmisja obrazu dobrej jakości o przestrzennej rozdzielczości full HD (1920x1080 próbek w obrazie) wymaga strumienia danych o szybkości około 16 Mbit/s. Prawie 10 lat badań nad ulepszeniem techniki MPEG-2 (lata 1994 – 2003) umożliwiło zredukowanie tej szybkości 2-krotnie, czyli w omawianym przypadku do wartości 8 Mbit/s. Takie właśnie możliwości zapewnia następcą techniki MPEG-2, czyli kodowanie AVC. Dzięki temu postępowi, jak również dokonującym się równolegle rozwojowi technologii transmisji sygnału, możliwe się stało szersze wykorzystanie po roku 2003 cyfrowej telewizji (większa liczba programów dla odbiorców), w tym transmisja sygnału o znacznie lepszej jakości (większa rozdzielczość przestrzenna obrazów).

Jednak na tym postęp się nie zakończył. Po roku 2003 w dalszym ciągu ulepszano metody kompresji. Prace te były prowadzone głównie w kontekście reprezentacji obrazów o wysokiej i ultrawysokiej rozdzielczości (przestrzenna rozdzielczość obrazów odpowiednio 1920x1080 i 3840x2160 próbek). W efekcie udało się poprawić efektywność koderów AVC prawie dwukrotnie! Wyniki tych prac leżą u podstaw nowej technologii HEVC (*High Efficiency Video Coding*) kompresji obrazu. Zamiast więc szybkości bitowej strumienia zakodowanych danych 8 Mbit/s (w przypadku kodowania AVC) otrzymuje się strumień około 4 Mbit/s, stosując najnowszą technikę kompresji obrazów. Takie możliwości kompresji obrazu otwierają już drogę do szerokiego wdrożenia systemów nowej generacji, w których rozdzielczość stosowanych obrazów jest wyższa niż full HD. Postęp rezultatów kompresji, który się dokonał w ciągu ostatnich dwóch dekad, dodatkowo przedstawiono na rys. 6.



■ Rys. 6. Efektywność kompresji trzech koderów wizyjnych

JAK ZMIENIŁA SIĘ ZŁOŻONOŚĆ KODOWANIA OBRAZÓW?

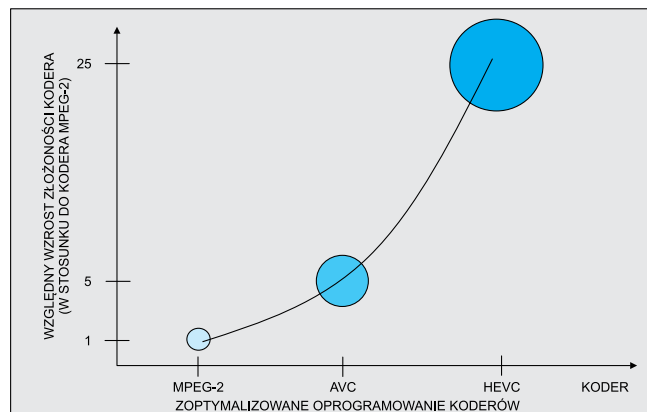
Wzrost efektywności kompresji obrazów ma jednak również swoją cenę. Jest nią coraz większa złożoność kodowania i dekodowania obrazów.

Wdrożenie w nowych koderach bardziej skomplikowanych narzędzi przewidywania i kodowania treści, ale również wprowadzenie zupełnie nowych rozwiązań, jak chociażby filtracje obrazu, które poprawiają jakość, istotnie wydłużają czas kodowania i dekodowania obrazów. Żeby uzmysłowić sobie, jak duży jest to przyrost, warto prześledzić parametry trzech koderów obrazu: MPEG-2, AVC oraz najnowszego koderów HEVC.

Zoptymalizowana na prędkość działania programowa wersja koderów MPEG-2 umożliwi na współczesnym procesorze (jeden wątek procesora Core i7, o częstotliwości zegara 3,4 GHz) kodowanie około 100 obrazów w czasie jednej sekundy. Dotyczy to kompresji obrazów full HD, czyli takich, których rozdzielczość przestrzenna wynosi 1920 x 1080 próbek obrazu. Praktyka pokazuje, że analogiczna realizacja nowszego koderów AVC umożliwi zakodowanie już tylko około 20 obrazów w ciągu sekundy. W sto-

sunku do kodera MPEG-2 wzrost złożoności jest więc aż 5-krotny! Ale najnowszy koder HEVC jest jeszcze bardziej złożony niż AVC. W przypadku oprogramowania nowego kodera, w którym zastosowano podobne techniki optymalizacji kodu, czas kodowania obrazów wydłuży się aż 5-krotnie, w porównaniu z koderem AVC, czyli aż 25-krotnie w stosunku do starszego kodera MPEG-2! Jak więc widać na rys. 7, dokonujący się w ciągu wielu lat wzrost złożoności kodowania obrazów ma charakter wykładniczy. Mniej więcej co dekadę przychodzi nam powitać nową technologię kompresji obrazów, której wymagania obliczeniowe, w stosunku do technologii poprzedniej, są aż kilka razy większe.

W tym miejscu można się oczywiście zastanawiać, z czego wynika tak duży wzrost złożoności koderów? Czy jest to wynik



■ Rys. 7. Porównanie złożoności trzech koderów ruchomego obrazu: MPEG-2, AVC i HEVC. Wyniki złożoności dla trzech zoptymalizowanych programowych realizacji koderów

użycia w koderze obrazu trudniejszych obliczeniowo rozwiązań? Oczywiście tak, ale nie tylko to stanowi powód. Zdaniem autora, zasadniczą przyczyną tkwi w liczbie trybów kodowania treści obrazu, którymi dysponuje dany koder. Starsze kodery, jak np. MPEG-2, były relatywnie bardzo proste i miały relatywnie mało różnych trybów. Wybór spośród nich tego najlepszego (dla danego fragmentu obrazu) odbywał się więc w krótkim czasie. Ale jeśli liczba możliwych sposobów zakodowania fragmentu jest bardzo duża (a tak jest w przypadku najnowszych koderów, jak HEVC), to bezpośrednio przekłada się to na znacznie dłuższy czas podejmowania decyzji, który z trybów najlepiej zastosować. Żeby się przekonać, który z nich jest najlepszy w danym fragmencie obrazu, koder sprawdza różne możliwe sposoby jego kompresji, co w przypadku szerszej palety dostępnych trybów znacznie wydłuża czas kodowania obrazów.

REALIZACJA OPROGRAMOWANIA KODEKA – JAK ZMIENIŁA SIĘ TRUDNOŚĆ?

Liczba dostępnych w kodeku trybów kodowania treści, jak również stopień skomplikowania poszczególnych trybów, mają bezpośrednie przełożenie na wielkość oprogramowania kodera i dekodera obrazów. Starsze kodeki obrazu, np. MPEG-2, korzystają z prostych metod kodowania danych, a i sama paleta dostępnych trybów jest w tych koderach (jak wcześniej powiedziano) bardzo ograniczona. W efekcie oprogramowanie starszych kodeków jest stosunkowo nieduże, tzn. nie przekracza 20 tysięcy linii kodu programu, opracowanego w języku programowania C/C++. Typowa implementacja kodeka MPEG-2 ma tych linii nie więcej niż 15 tysięcy.

Jednak najnowsze kodeki obrazu, np. HEVC, to już całkowicie inna historia. Dostępnych narzędzi kompresji jest w nich znacznie więcej. Wiele z nich realizuje dodatkowo dużo bardziej złożone algorytmy przetwarzania danych, co też ma swoje odzwierciedlenie w wielkości ich oprogramowania. Wynikiem tych różnic jest

oprogramowanie nowego kodeka, które w porównaniu z realizacją starszych kodeków jest aż kilka razy większe! W porównaniu z programem kodeka MPEG-2 program nowego kodeka HEVC zawiera 6 razy więcej linii kodu programu, czyli w sumie jest to około 90 tysięcy linii kodu! (patrzy rys. 8).

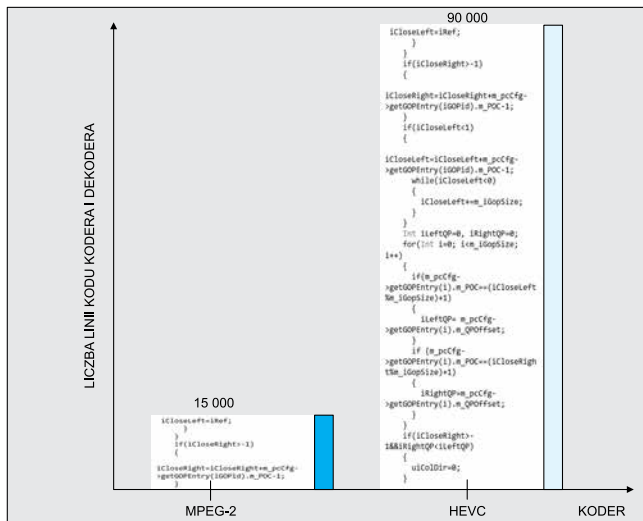
Większa objętość oprogramowania najnowszych kodeków wpływa oczywiście na czas, jaki jest potrzebny do opracowania oraz przetestowania kodu programu. Na pierwszy rzut oka mogłoby się wydawać, że w stosunku do starszych kodeków czas ten w przypadku nowszych kodeków będzie proporcjonalnie dłuższy, zgodnie z różnicą wielkości kodu kodeka starszego i nowszego. Zatem na opracowanie programu kodeka HEVC poświęci się około 6 razy więcej czasu w porównaniu z programem kodeka MPEG-2. Jednak doświadczenie autora pokazuje, że tak prosto się to nie skaluje. W praktyce ten czas jest znacznie dłuższy. Oprócz implementacji samych narzędzi kompresji, których – jak powiedziano – jest w nowszych kodekach znacznie więcej, istnieje jeszcze kwestia doprowadzenia do właściwego ich współdziałania, sprawnej i niezawodnej wymiany danych pomiędzy poszczególnymi technikami, co – jak pokazuje praktyka – również wymaga bardzo dużo czasu. Z dość dużą dokładnością można przewidywać, że w stosunku do implementacji kodeka MPEG-2 realizacja programu kodera i dekodera HEVC zajmie aż 10 razy więcej czasu! Przynajmniej tak będzie w przypadku realizacji programowej kodeka. W przypadku realizacji sprzętowej różnica czasu implementacji może być jeszcze bardziej pogłębiona. Podane wartości należy traktować jako pewien szacunek, jednak – zdaniem autora – dobrze obrazujący rzeczywistą skalę trudności wykonania kodu programu. Wiadomo przecież, że na czas ten wpływ mają również przyjęte w oprogramowaniu sposoby sterowania pracą kodera oraz użyte techniki optymalizacji kodu programu.

Ale warto podać choćby przybliżone terminy realizacji oprogramowania kodeka. Wykwalifikowany 2-3 osobowy zespół programistów potrafi wykonać od zera pełne oprogramowanie kodera i dekodera MPEG-2 w ciągu 2-3 miesięcy. Prawdopodobnie dodatkowy czas byłby jeszcze jednak potrzebny do optymalizacji opracowanego oprogramowania. W przypadku nowego kodeka HEVC będzie to w praktyce ponad 20 miesięcy! Jeszcze raz podkreślimy, takie byłyby mniej więcej czasy realizacji kodu przez doświadczonych programistów. Prace te wymagają więc dużej wiedzy i niemałych nakładów finansowych. Z tych powodów bardzo mało jest na świecie ośrodków, które podejmują to zadanie, szczególnie w odniesieniu do sprzętowych realizacji kodeków.

JAKA JEST NAJBLIŻSZA PRZYSZŁOŚĆ?

Wysokie osiągnięcia nowej techniki HEVC nie stanowią przyczynku do przerwania dalszych badań w zakresie kompresji. Takie prace są oczywiście nadal prowadzone. I to już od momentu ustanowienia międzynarodowego zalecenia dla techniki HEVC. Intensywne prace zmierzające do pojawienia się następcy technologii HEVC prowadzą obecnie grupy ekspertów ISO/IEC MPEG (*Moving Picture Experts Group*) oraz ITU-T VCEG (*Video Coding Experts Group*). Ich zwińczeniem ma być nowa technologia kompresji (nazywana roboczo **FVC** – *Future Video Coding*), której można się spodziewać około roku 2020.

Chociaż obecnie nie wiadomo dokładnie, jakie rozwiązania zostaną ostatecznie przyjęte, to wiele wskazuje na to, że będzie to w dalszym ciągu ulepszenie idei kompresji hybrydowej. Nowe rozmiary bloków obrazu, większa liczba predyktorów treści (dotyczy przewidywania wewnątrz- i międzyobrazowego), dokładniejsza reprezentacja informacji o ruchu, przyjęcie w koderze obrazu bardziej złożonego modelu ruchu obiektów w kodowanej scenie (nie tylko prosty, translacyjny ruch obiektów na płaszczyźnie obrazu, ale również zmiany treści zgodnie z wybranymi przekształceniami afinicznymi) to są zmiany, których można się spodziewać. Ostanie badania w tym zakresie wskazują na możliwość zwiększenia efektywności kompresji obrazów o około 40%



■ Rys. 8. Porównanie wielkości oprogramowania kodeków MPEG-2 i HEVC. Różnica wielkości oprogramowania, wyrażana liczbą linii kodu programu jest aż 6-krotna!

(w porównaniu z koderem HEVC), przy więcej niż 10-krotnym zwiększeniu złożoności koderów obrazów (również w porównaniu do złożoności koderów HEVC) [8].

Ostatnie dwie dekady to czas ogromnego postępu w zakresie możliwości reprezentacji ruchomego obrazu. Motorem tego postępu były badania ośrodków naukowych i przemysłowych, rosnąca moc obliczeniowa procesorów i układów scalonych, ale również coraz wyższe oczekiwania samych użytkowników co do jakości dostarczanego im obrazu.

Takie oczekiwania przeszło 20 lat temu spełniła technologia MPEG-2, przynajmniej w niemałym zakresie. W historii rozwoju metod kompresji jej znaczenie jest szczególne, ponieważ stała się tą technologią, która przyczyniła się do sukcesów, jakim w tamtym czasie było rynkowe wdrożenie cyfrowej telewizji oraz kina domowego.

Ale jak pokazał czas, osiągnięcia początku lat 90. to był dopiero początek burzliwego rozwoju ery multimediów, który trwa przecież do dzisiaj. Można powiedzieć, że motor napędowy do dalszego postępu pozostał ten sam. Dalsze wspólne wysiłki badaczy nad jeszcze lepszymi rozwiązaniami, niż te znane MPEG-2, dały owoce w postaci nowych, znacznie wydajniejszych technik kompresji, odpowiednio AVC (rok 2003) i HEVC (rok 2013). W toku 20 lat badań efektywność reprezentacji obrazów udało się zwiększyć aż 4-krotnie, zwiększając jednak nakłady obliczeniowe na kompresję obrazu aż 25 razy!

Zatem dokonany w czasie tych lat postęp w zakresie kompresji to ewolucja czy rewolucja? A może ani jedno, ani drugie? Zdaniem autora, to zależy, jak się na tę kwestię spojrzy. Skrajna

opinia może być taka (i niektórzy rzeczywiście tak uważają), że skoro kolejne generacje koderów opierają się na tej samej idei kompresji danych (bo przecież cały czas następuje kodowanie obrazów według idei kompresji hybrydowej), to w istocie nic się nie zmieniło bądź zmieniło niewiele. Gdyby zgodzić się z tą opinią, to również trzeba byłoby zgodzić się z tezą, że nie ma większej różnicy pomiędzy samochodami Syrena 105 i najnowszym modelem Tesli. Każdy z tych samochodów ma koła, kierownicę, silnik, czyli jeżdżą według tej samej ogólnej zasady. Łatwo się zgodzić, że teza dotycząca samochodów nie jest prawdziwa. Między nimi jest przecież ogromna różnica. I bardzo podobnie jest w przypadku koderów generacji starszej i najnowszej. Widziana z lotu ptaka główna idea działania wydaje się taka sama, ale szczegóły są bardzo różne. W ogromny sposób przekłada się to na osiągnięcia koderów, złożoność ich oprogramowania, złożoność kodowania obrazów, co autor stara się wykazać w tym artykule.

Można więc powiedzieć, że w toku ewolucyjnych zmian działania koderów osiągnięto nowe techniki kompresji. W stosunku do rozwiązań starszych rewolucjonizują one wręcz wyniki samej kompresji, jak również parametry koderów, jak jego złożoność i wielkość implementacji. Zdaniem autora, tak jest, jeśli dokona się bezpośredniego porównania ze sobą technik kompresji MPEG-2 i HEVC, które dzieli odstęp czasowy 20 lat.

Praca finansowana ze środków przyznanych przez Ministerstwo Nauki i Szkolnictwa Wyższego na działalność statutową w roku 2018 w ramach zadania: *Teoria i algorytmy wielowymiarowego przetwarzania sygnałów.*

LITERATURA

- [1] ISO/IEC 11172-2, Coding of Moving Pictures and Associated Audio for Digital Storage Media up to about 1.5 Mbit/s, Part 2: Video. (MPEG-1), November 1993.
- [2] ISO/IEC 13818-2 and ITU-T Rec. H.262, Generic Coding of Moving Pictures and Associated Audio Information – Part 2: Video. (MPEG-2), November 1994.
- [3] ISO/IEC 14496-10, Generic Coding of Audio-Visual Objects, Part 10: Advanced Video Coding, March 2006.
- [4] Wiegand T., G. J. Sullivan, G. Bjontegaard, A. Luthra, „Overview of the H.264/AVC Video Coding Standard”. *IEEE Transactions on Circuits and Systems for Video Technology* 13(7), 560-576, 2003.
- [5] ISO/IEC and ITU-T, High Efficiency Video Coding (HEVC), ISO/IEC 23008-2 (MPEG-H Part 2) / ITU-T Rec. H.265, April 2013.
- [6] Sullivan G. J., J.-R. Ohm, W.-J. Han, T. Wiegand, „Overview of the High Efficiency Video Coding (HEVC) Standard”. *IEEE Transactions on Circuits and Systems for Video Technology* 22(12), 1649-1668, 2012.
- [7] Ohm J.-R., *Multimedia Communication Technology. Representation, Transmission and Identification of Multimedia Signals*, Springer. 2004.
- [8] Huanbang Chen i inni, Description of SDR, HDR and 360° video coding technology proposal by Huawei, GoPro, HiSilicon, and Samsung, Document JVET-J0025, 10th Meeting: San Diego, US, 10–20 Apr. 2018.

SYMBOLE TEMATYKI ARTYKUŁÓW

ZAGADNIENIA OGÓLNE	SIĘCI TELEKOMUNIKACYJNE	USŁUGI MULTIMEDIA	ELEMENTY UKŁADY METODY	TELETRANSMISJA	TELEFONIA	TELEINFORMACYKA	OPTOTELEKOMUNIKACJA	RADIOKOMUNIKACJA RADIOFONIA TELEWIZJA	POMIARY	EKONOMIKA PRAWO	SPONSOROWANE
GENERAL PROBLEMS	COMMUNICATION NETWORKS	SERVICES MULTIMEDIA	COMPONENTS DEVICES METHODS	TELETRANSMISSION	TELEPHONY	COMPUTER SCIENCE	OPTOCOMMUNICATIONS	RADIOCOMMUNICATION RADIO TELEVISION	MEASUREMENTS	ECONOMICS LAW	SPONSORING