

Tomasz Grajek, Marek Domański
Politechnika Poznańska
Wydział Elektroniki i Telekomunikacji
Katedra Telekomunikacji Multimedialnej i Mikroelektroniki

HEVC

NOWA TECHNIKA KOMPRESJI OBRAZÓW TELEWIZJI DOZOROWEJ O WIELKIEJ ROZDZIELCZOŚCI

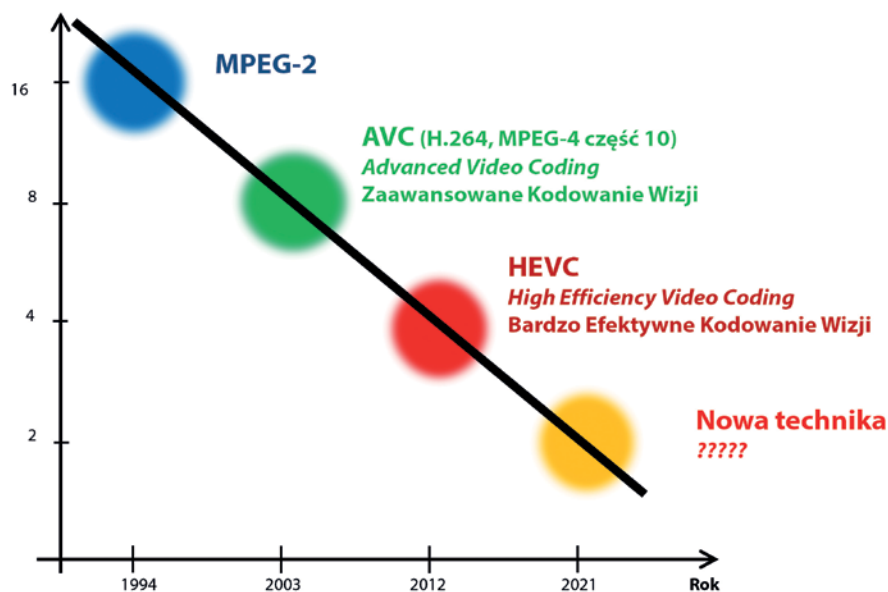
Mniej więcej co 9 lat pojawia się nowa generacja technik kompresji obrazów ruchomych, która w stosunku do poprzedniej pozwala na około dwukrotnie większą redukcję szybkości bitowej niezbędnej do przesyłania obrazu o tych samych parametrach. Dzięki temu można też przysyłać obrazy o zwiększonej rozdzielczości bez potrzeby znacznego zwiększania szybkości bitowej.

Najnowszym osiągnięciem w dziedzinie technik kompresji obrazów ruchomych jest Bardzo Wydajne Kodowanie Wizji (*High Efficiency Video Coding* – HEVC), które ujęto w normie MPEG-H część 2 (ISO/IEC IS 23008-2) Międzynarodowej Organizacji Normalizacyjnej (ISO) oraz w zaleceniu ITU-T H.265 Międzynarodowego Związku Telekomunikacyjnego (ITU) [1]. W artykule przedstawiono wybrane zagadnienia związane z przyszłymi zastosowaniami kodowania HEVC w telewizji dozorowej.

GENERACJE TECHNIK KOMPRESJI OBRAZÓW RUCHOMYCH

Wprowadzenie telewizji cyfrowej było możliwe dzięki opracowaniu techniki kompresji ujętej w normie MPEG-2. Telewizja wysokiej jakości (HDTV), odznaczająca się większą rozdzielczością obrazów, wykorzystuje głównie bardziej efektywną technikę kompresji, zwaną Zaawansowanym Kodowaniem Wizji (*Advanced Video Coding* – AVC) [2], która powstała około 9 lat po technice MPEG-2. Mniej więcej po następnych 9 latach pojawiła się omawiana w tym artykule technika HEVC, którą opracowano mając na uwadze m.in. telewizję bardzo wysokiej jakości (*Ultra High Definition Television* – UHD TV) wykorzystującą obrazy o rozmiarach 3820×2160 (tzw. format 4K). Zarówno zastąpienie kodowania MPEG-2 kodowaniem AVC, jak i zastąpienie

Prędkość bitowa dla obrazów
o dużej rozdzielczości (1920×1080 , 25 obraz/s)
w skali logarytmicznej, wyrażona w [Mb/s]



Rys. 1. Rozwój technik kodowania obrazów ruchomych

techniki AVC techniką HEVC pozwala na około dwukrotne zwiększenie redukcji szybkości bitowej przy zachowaniu niezmięnionej jakości obrazów zdekodowanych.

Na rys. 1 pokazano kolejne generacje technik kompresji i, dla przykładu, podano orientacyjne wielkości szybkości bitowych dla telewizji rozszerzanej. Ekstrapolacja danych z wykresu pokazanego na rys. 1 prowadzi do wniosku, że około roku 2021 można się spodziewać nowej techniki kompresji ponownie pozwalającej na dwukrotną redukcję szybkości bitowej. Zaskakujące, że to proste przewidywanie jest

zgodne z planami grup ekspertów pracujących nad nowymi technikami kompresji pod auspicjami Międzynarodowej Organizacji Normalizacyjnej i Międzynarodowego Związku Telekomunikacyjnego.

Dla obrazów telewizyjnych różnych rozmiarów można podać zależności podobne do pokazanych na rys. 1. Te zależności określające orientacyjną szybkość bitową B pozwalającą na uzyskanie zdekodowanego obrazu telewizyjnego o dobrej jakości można ująć jednym wzorem [3,4]:

$$B \approx A \times V \text{ [Mb/s]} \quad (1)$$

gdzie:

A jest współczynnikiem zależnym od techniki kompresji:

A = 4 dla MPEG-2,

A = 2 dla AVC,

A = 1 dla HEVC,

V jest współczynnikiem zależnym od formatu obrazu:

V = 1 dla SDTV (720 × 576, 25i),

V = 4 dla HDTV (1920 × 1080, 25i),

V = 16 dla UHD TV (3840 × 2160, 50p),

natomiast 25i i 50p oznaczają odpowiednio 25 obrazów/s z wybieraniem międzyliniowym („z przepłotem”) oraz 50 obrazów/s z wybieraniem kolejnoliniowym („bez przepłotu”, „obraz progresywny”).

Należy jednak zaznaczyć, że szybkość bitowa wymagana do uzyskiwania zdekodowanych obrazów bez dokuczliwych zakłóceń bardzo silnie zależy od treści obrazów i jest mniejsza dla treści łatwiejszych do kompresji (np. dla programów publicystycznych). Także obrazy telewizji dozorowej, które często obejmują dużo nieruchomego tła, przeważnie można kodować ze znacznie mniejszą szybkością bitową niż to wynika ze wzoru (1). Jednak również w tych przypadkach w przybliżeniu słuszne pozostają relacje pomiędzy szybkościami bitowymi wymaganymi dla poszczególnych technik.

W związku z tym pojawia się pytanie, jak to jest możliwe, że dla każdej kolejnej techniki udaje się zmniejszyć wymaganą szybkość bitową o połowę, nie pogarszając przy tym jakości obrazu? Wspomniane techniki kompresji obrazu ruchomego wykorzystują te same podstawowe zasady kodowania hybrydowego, a kolejne nowe techniki korzystają tylko z uzupełnień w postaci dodatkowych nowych narzędzi (przegląd technik kompresji, zwłaszcza MPEG-2 i AVC, można znaleźć w książce [5]). Ale to tak, jakby powiedzieć, że w motoryzacji od czasu Modelu T Forda do dzisiejszej Tesli praktycznie także nic się nie zmieniło. Samochody ciągle mają koła i kierownicę, a techniki kodowania predykcję z kompensacją ruchu i kodowanie wewnętrznie obrazowe. A jednak poza ogólną ideą, zarówno w przypadku motoryzacji, jak i technik kodowania, mamy do czynienia z kolejnymi zupełnie innymi rozwiązaniami. Na przestrzeni lat mnóstwo, czasem tylko drobnych usprawnień i ulepszeń powoduje, że osiąga się zupełnie inną jakość, choć główna idea pozostaje niezmienną.

BARDZO WYDAJNE KODOWANIE WIZJI

Różne badania eksperymentalne wskazują, że kompresja obrazów ruchomych z wykorzystaniem techniki HEVC zamiast bardzo popularnej obecnie techniki kodowania AVC pozwala na około dwukrotne zmniejszenie szybkości bitowej [6-8]. HEVC, podobnie jak wiele wcześniejszych technik kompresji, wykorzystuje ideę hybrydowego kodowania sygnałów wizyjnych. Każdy obraz jest dzielony na nienachodzące na siebie fragmenty (bloki próbek), a zawartość każdego z nich jest przewidywana (dokonuje się predykcji) na podstawie fragmentów (bloków) obrazu za-



Rys. 2. Przykładowe obrazy z sekwencji testowych typowych dla telewizji dozorowej użytych w eksperymentach



kodowanych wcześniej i pochodzących albo z tego samego obrazu (kodowanie wewnątrzobrazowe), albo z innych obrazów (kodowanie międzyobrazowe). Różnica pomiędzy faktyczną treścią każdego fragmentu obrazu a jego przewidywaną wersją jest nazywana błędem predykcji i podlega transformacji kosinusowej. Wyliczone w ten sposób współczynniki transformaty, po skwantowaniu, są kodowane w koderze entropijnym i zapisywane do strumienia danych. Dokładny opis rozwiązań zastosowanych w technice HEVC można znaleźć w książkach [6, 7].

Choć powyższy prosty schemat kodowania nie zmienił się od ponad 20 lat, to każda nowa technika kodowania wprowadzała do niego wiele istotnych ulepszeń. Ulepszenia zastosowane w HEVC pozwalają na tak dobrą predykcję, że zazwyczaj dla dużych części kolejnych obrazów nie trzeba przysyłać żadnych informacji. W zależności od treści obrazu koder HEVC dzieli obraz na bloki próbek o różnych rozmiarach i dla każdego bloku dobiera tryb kodowania. Dla poszczególnych obrazów i fragmentów obrazów koder HEVC wyszukuje takie warianty kodowania, które najsilniej redukują liczbę bitów, a jednocześnie wprowadzają możliwie najmniejsze zakłócenia obrazu.

Liczba możliwych do wyboru wariantów podziału na bloki i kodowania bloków jest olbrzymia. Dlatego pełne wykorzystanie możliwości tkwiących w technice HEVC wymaga zastosowania w koderze olbrzymiej mocy obliczeniowej, wielokrotnie większej niż w koderach AVC. Z tego powodu na razie koder HEVC są bardzo kosztownymi systemami komputerowymi o znacznym poborze mocy zasilania. Można się jednak spodziewać, że wkrótce pojawią się układy scalone realizujące dobre koder HEVC w jednej strukturze półprzewodnikowej. Dostępność takich układów na rynku i ich nieuchronnie spadające ceny umożliwią również

szerokie zastosowanie techniki HEVC w dozorze wizyjnym, o czym już dzisiaj powinni zacząć myśleć producenci urządzeń i oprogramowania dla dozorów wizyjnego. Powinni oni też brać pod uwagę, że pierwsze, przeważnie jeszcze niedopracowane, tanie realizacje koderów HEVC nie będą pozwalały na pełne wykorzystanie możliwości tkwiących w technice HEVC, co jednak prawdopodobnie szybko się poprawi.

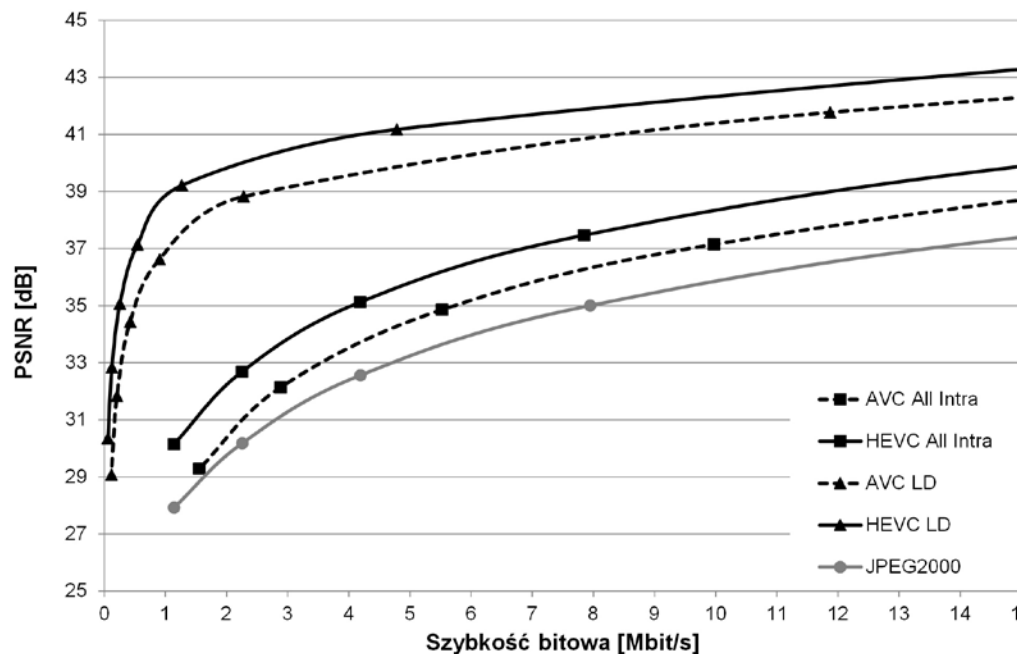
Natomiast dekodery HEVC w zasadzie nie są bardziej skomplikowane niż dekodery AVC i mogą być realizowane przez programy uruchamiane nawet na małych komputerach przenośnych.

EFEKTYWNOŚĆ KOMPRESJI I PORÓWNANIE Z INNYMI TECHNIKAMI

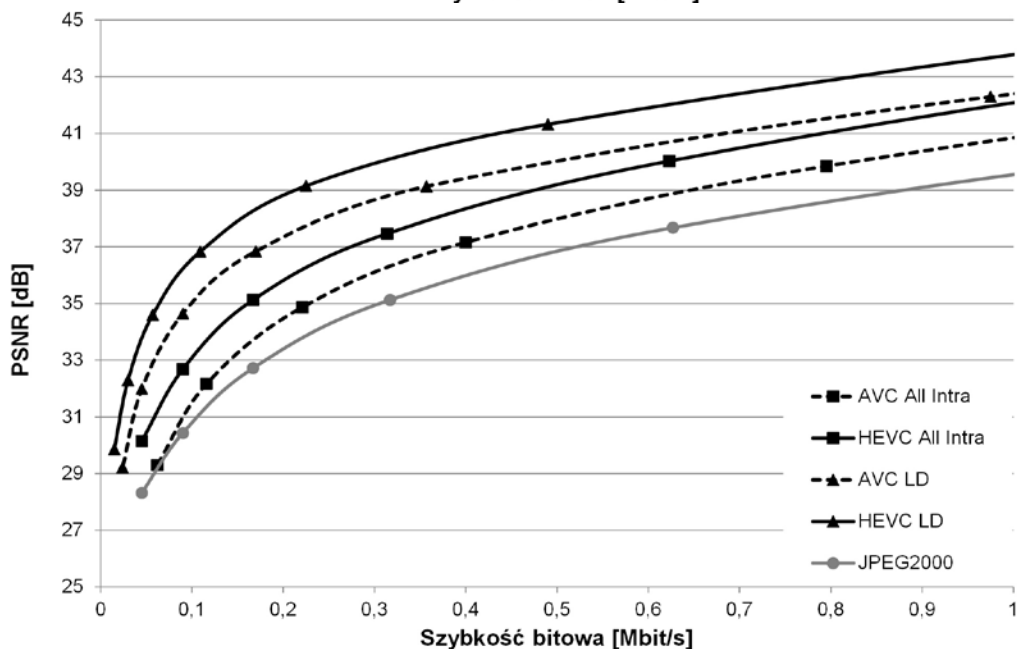
Jak już wspomniano, ulepszenia i zmiany wprowadzone w HEVC pozwalają na zmniejszenie szybkości bitowej o około połowę dla większości sekwencji wizyjnych, przy zachowaniu podobnej jakości zakodowanych obrazów [8]. Większość wyników przedstawionych w literaturze dotyczy telewizji rozświetlonej, natomiast brakuje prac związanych z telewizją dozorową. Dlatego w Katedrze Telekomunikacji Multimedialnej i Mikroelektroniki Politechniki Poznańskiej przeprowadzono eksperymenty z sekwencjami wizyjnymi typowymi dla telewizji dozorowej (rys. 2). W badaniach tych porównano modele koderów AVC, HEVC oraz JPEG2000. Dla AVC wykorzystano oprogramowanie JM_18.4 [9], dla HEVC – HM_15.0 [10], a dla JPEG2000 – OpenJPEG_2.1.0 [11]. Modele te pozwalają na osiągnięcie stosunku szybkości bitowej do jakości zakodowanych obrazów bliskiego maksymalnemu możliwemu do uzyskania dla danej techniki kompresji.

Badania przeprowadzono dla dwóch skrajnych częstotliwości obrazów spotykanych w telewi-

a) Sekwencje zarejestrowane w trybie 25 obrazów na sekundę



b) Sekwencje zarejestrowane w trybie 1 obrazu na sekundę



Rys. 3. Porównanie efektywności kompresji JPEG2000, AVC i HEVC dla sekwencji wizyjnej z rys. 2a (rozmiar obrazu 1920 × 1080).

Dla koderów AVC i HEVC uwzględniono dwa warianty pracy: All_Intra – kolejne obrazy kodowane niezależnie od siebie; LD – tryb kodowania o małym opóźnieniu kodowania z wykorzystaniem kodowania międzyobrazowego

zji dozorowej: 25 obrazów/s i 1 obraz/s. Kodery HEVC i AVC pracowały w dwóch konfiguracjach pozwalających na ograniczenie czasu dekodowania (dla realizacji za pomocą szybkich układów) do kilku – kilkunastu milisekund:

- All Intra – obrazy w sekwencji są kodowane wewnątrzobrazowo i niezależnie od siebie,
- LD – co 2 sekundy (dla 25 obrazów/s) lub co 4 sekundy (dla 1 obrazu/s) jeden obraz jest kodowany w trybie wewnątrzobrazowym, wszystkie pozostałe wykorzystują informacje o wcześniej zakodowanych obrazach (są kodowane w trybie międzyobrazowym).

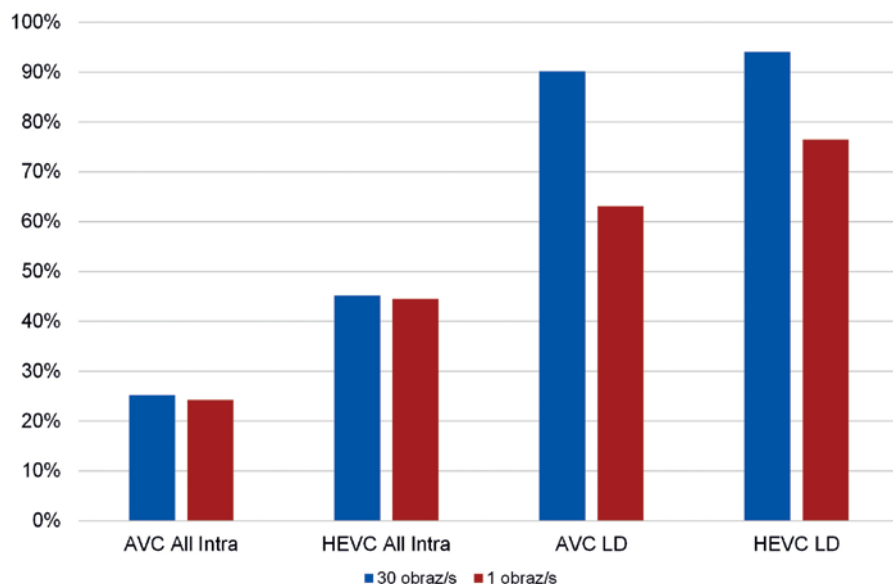
Jakość mierzono za pomocą współczynnika PSNR [5], który jest miarą błędu średniokwa-

dratowego 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.

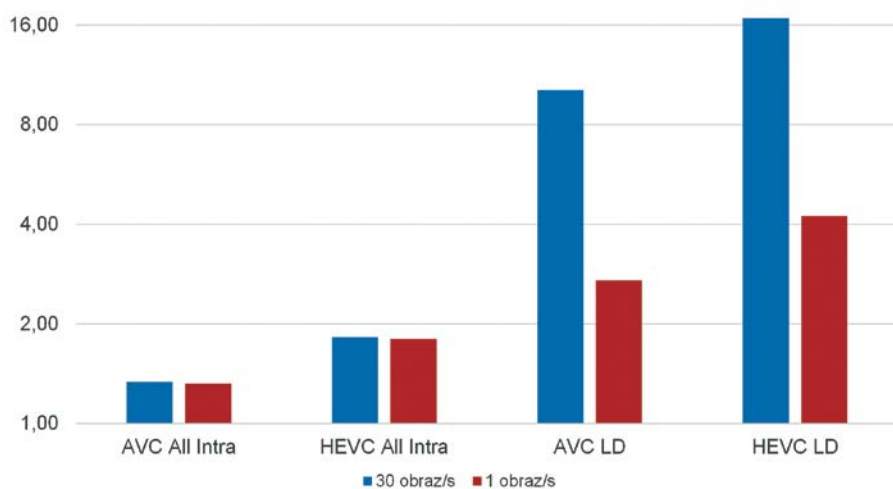
Wyniki otrzymane dla przykładowej sekwencji wizyjnej z rys. 2a przedstawiono na rys. 3. O „lepszej” (bardziej efektywnej) technice kompresji mówimy wtedy, gdy na wykresie krzywa związana z tą techniką kompresji leży powyżej i na lewo od innych krzywych. W celu ilościowego porównania efektywności kompresji różnych technik kodowania obliczono miarę Bjøntegaarda [5]. Miara ta określa średnią zmianę szybkości bitowej zakodowanego strumienia spowodowaną zastosowaniem

nowej techniki, przy zachowaniu niezmięnionej jakości zakodowanych obrazów. Na rys. 4 przedstawiono wyniki uśrednione dla wspomnianych pięciu sekwencji wizyjnych. Jako technikę odniesienia przyjęto technikę JPEG2000, często stosowaną w systemach dozorowych. Dodatkowo na rys. 5 pokazano stopień kompresji uzyskany dla AVC i HEVC w stosunku do JPEG2000.

Uzyskane wyniki potwierdzają, że HEVC pozwala na znaczną redukcję szybkości bitowej. Szczególnie duże zyski wynikają z zastosowania predykcji międzyobrazowej (konfiguracja LD), czyli takiej, w której kolejno kodowane obrazy mogą korzystać z informacji z wcześniej zakodowanych obrazów. Co więcej, w przypadku konfiguracji



Rys. 4. Redukcja szybkości bitowej w porównaniu do techniki JPEG2000. Redukcja o 50% oznacza, że nowa technika kodowania wymaga dwukrotnie mniejszej szybkości bitowej w porównaniu do JPEG2000. Wyniki uśredniono dla wszystkich sekwencji testowych



Rys. 5. Wzrost stopnia kompresji w porównaniu do techniki JPEG2000. Wyniki uśredniono dla wszystkich badanych sekwencji testowych.

kodera LD wyraźnie widać wpływ częstotliwości obrazów na efektywność kompresji (słupki czerwone i niebieskie).

Dużo lepiej można przewidzieć treść kolejnych obrazów w sekwencji w przypadku, gdy obrazy są rejestrowane z małymi odstępami w czasie (25 obrazów/s) niż w przypadku, gdy ta odległość w czasie wzrasta (1 obraz/s). W związku z tym redukcja szybkości bitowej w stosunku do techniki JPEG2000 jest silniejsza dla sekwencji zarejestrowanej z 25 obrazami/s niż z tylko 1 obrazem na sekundę.

PODSUMOWANIE

Zastosowanie techniki HEVC zamiast którejkolwiek techniki dotychczas stosowanej do kompresji obrazów telewizyjki dozorowej

doprowadzi do znacznej redukcji szybkości bitowej nawet przy zachowaniu bardzo małych wartości opóźnienia kodowania, co wiąże się z zachowaniem podczas kodowania niezmięnionej kolejności obrazów. Eksperymenty z pięcioma sekwencjami dozorowymi, w których zarejestrowano dosyć dużo ruchu obiektów na nieruchomym tle, pokazały, że redukcja szybkości bitowej w stosunku do JPEG2000 jest około 16-krotna, a w stosunku do AVC osiąga prawie 50%, co jest w przybliżeniu zgodne ze wzorem 1. Gdy zostają wyłączone mechanizmy predykcji międzyobrazowej, a więc gdy kolejne obrazy są w koderze HEVC kodowane całkowicie niezależnie, tak jak w technikach JPEG i JPEG2000, to ta redukcja nadal wynosi ponad 40% w stosunku

do JPEG2000 i około 20% w stosunku do AVC z wyłączonym trybem międzyobrazowym.

Eksperymenty pokazują, że ruchome obrazy dozorowe HD (1920 × 1080, 25 obrazów/s) ze statycznym tłem i dużą liczbą szybko poruszających się obiektów można kodować z szybkością bitową poniżej 4 Mb/s przy zachowaniu dobrej jakości obrazów zdekodowanych i bardzo małego opóźnienia kodowania wynikającego z czasu wykonywania algorytmu.

Przedstawione uwagi dotyczące kompresji HEVC prowadzą do wniosku, że w niedalekiej przyszłości należy się spodziewać szerokiego zastosowania tej techniki w systemach dozorowych, a zwłaszcza w tych, w których będzie się wprowadzać kamery wytwarzające obrazy o liczbach punktów znacznie przekraczających 2 miliony.

Literatura:

- [1] ISO/IEC International Standard 23008-2: 2015 (ed. 2), [także:] ITU-T Recommendation H.265: April 2015, High efficiency video coding.
- [2] M. Domański, T. Grajek, J. Marek, *Zaawansowana kompresja cyfrowych sygnałów wizyjnych – standard AVC/H.264*, Systemy Alarmowe, nr 2, 2005, str. 2-7.
- [3] M. Domański, „Approximate video bitrate estimation for television services”, ISO/IEC JTC1/SC29/WG11 (MPEG), Doc. M36571, Warszawa June 2015, [także:] http://www.multimedia.edu.pl/publications/2015/MPEG_m36571.pdf.
- [4] M. Domański, A. Dziembowski, T. Grajek, A. Grzelka, Ł. Kowalski, M. Kurc, A. Łuczak, D. Mieloch, R. Ratajczak, J. Samelak, O. Stankiewicz, J. Stankowski, K. Wegner, „Methods of high efficiency compression for transmission of spatial representation of motion scenes”, IEEE Int. Conf. Multimedia and Expo, Torino, July 2015. [także:] http://multimedia.edu.pl/publications/2015/icme2015_EU_project_opus.pdf
- [5] M. Domański, *Obraz cyfrowy. Reprezentacja, kompresja, podstawy przetwarzania. Standardy JPEG i MPEG*, Wydawnictwa Komunikacji i Łączności, Warszawa, 2010.
- [6] V. Sze, M. Budagavi, G.J. Sullivan, *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2015.
- [7] M. Wien, *High Efficiency Video Coding: Coding Tools and Specification*, Springer, 2015.
- [8] J.R. Ohm, G.J. Sullivan, H. Schwarz, T.K. Tan, T. Wiegand, *Comparison of the coding efficiency of video coding standards – Including High Efficiency Video Coding (HEVC)”. IEEE Transactions on Circuits and Systems for Video Technology*, nr 22(12), 2012, str. 1669–1684.
- [9] H.264/MPEG-4 AVC Reference Software, Joint Model 18.4, <http://iphome.hhi.de/suehring/tml/download/jm18.4.zip>.
- [10] HM 15.0 Reference Software, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-15.0.
- [11] Open JPEG, An open-source JPEG 2000 codec written in C, <http://www.openjpeg.org>.