

Olgierd Stankiewicz*
Krzysztof Wegner
Damian Karwowski
Jakub Stankowski
Krzysztof Klimaszewski
Tomasz Grajek
Katedra Telekomunikacji Multimedialnej i Mikroelektroniki
Wydział Elektroniki i Telekomunikacji
Politechnika Poznańska
ul. Polanka 3, 60-965 Poznań
*ostank@multimedia.edu.pl



Poznań, 21-23 czerwca 2017

STEROWANIE KODEREM HEVC WYKORZYSTUJĄCE USUWANIE SZUMÓW

MODE SELECTION IN HEVC USING NOISE REMOVAL

Streszczenie: Artykuł dotyczy optymalizacji kodera HEVC. Proponowana jest nowa metoda wyboru trybów kodowania, w której koder HEVC kodujący sekwencję wejściową sterowany jest przez drugi koder, kompresujący sekwencję odszumioną. Sposób realizacji sterowania zapewnia, iż łączna złożoność obu koderów jest niewiele większa niż złożoność pojedynczego kodera. Zaproponowaną metodę zaimplementowano i przebadano eksperymentalnie. Uzyskane wyniki pokazują, iż umożliwia ona około 1,5% redukcję strumienia bitowego.

Abstract: The article is related to HEVC encoder optimization. A novel encoder-side mode selection method is proposed. It exploits information about selected modes from a second encoder, which in parallel compresses a denoised version of the input sequence. The realization of the mode selection method ensures that the complexity of both encoders is negligibly larger than complexity of a single encoder. The proposed method has been implemented and assessed experimentally. The results show that it enables up to 1.5% reduction of the bitstream.

Słowa kluczowe: HEVC, sterowanie koderem, optymalizacja wyboru trybów, kodowanie, redukcja szumu.

Keywords: HEVC, encoder control, mode optimization, encoding, noise reduction.

1. WSTĘP

Postęp w rozwoju technik kompresji pozwala coraz bardziej efektywnie reprezentować obraz ruchomy. Najnowszą metodą kodowania opisaną w normach ISO/IEC oraz ITU jest technika HEVC [8]. Względem poprzedniej techniki AVC, HEVC pozwala uzyskać około 50% redukcję strumienia przy zachowaniu tej samej jakości [6]. Tak duży wzrost efektywności wynika ze znacznego rozwoju narzędzi kompresji dostępnymi dla kodera. Przykładowo, wyliczyć można [3], iż dla jednostki kodowania LCU (Largest Coding Unit) koder HEVC ma do wyboru około $48 \cdot 10^{18}$ różnych kombinacji podziałów CU, TU, PU, 33 kierunki predykcji intra, oraz dwie listy referencyjne ze zmiennymi obrazami w predykcji inter. Z tego względu wybór trybów podczas kodowania jest jeszcze bardziej złożony niż w przypadku wcześniejszych technik, np. AVC. Wśród wszystkich dostępnych

możliwości, koder poszukuje optymalnego zestawu trybów, które zmaksymalizowałyby stosunek jakości-przepływność. Przejrzenie wszystkich możliwości w skali całego obrazu jest praktycznie niemożliwe i dlatego najczęściej wybór trybów optymalizuje się jedynie lokalnie na poziomie jednej jednostki LCU. Podejście takie jest zastosowane np. w oprogramowaniu modelowym HEVC grupy JCT-VC – HM [5].

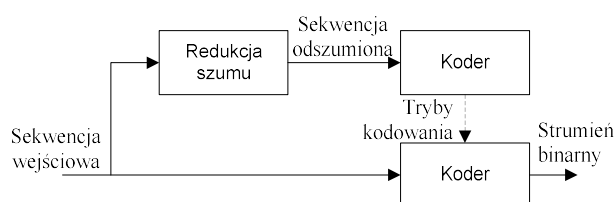
Najszerszym zastosowaniem koderów wizyjnych jest kodowanie sekwencji naturalnych, rejestrowanych kamerami. Sekwencje takie charakteryzuje występowanie w nich szumu, w sposób nieunikniony powstającego w układach elektronicznych kamery. Szum jest w większości przypadków niepożądanym zakłóceniem, degradującym jakość sekwencji. Z kolei z perspektywy wyboru trybów, występowanie szumu w obrazie powoduje, że w danej jednostce CU wybierana jest taka kombinacja trybów predykcji i współczynników transformaty, która najlepiej reprezentuje sygnał z szumem. Optymalizacja jest zatem zakłócona przez niepożądane składniki kodowanego sygnału. Dodatkowo – zakodowany i zrekonstruowany fragment obrazu z szumem nie jest najlepszym źródłem predykcji dla kolejnego fragmentu z szumem, gdyż szum w kolejnych fragmentach obrazu jest słabo skorelowany. Zjawisko to jest problemem zarówno w predykcji wewnątrzobrazowym jak i międzyobrazowym. Jednocześnie, koder HEVC posiada tak finezyjne możliwości predykcji, że błędem predykcji jest w praktyce sam szum (szum ma charakter losowy i nie może być przewidziany na podstawie treści obrazu). W praktyce oznacza to, że dla dużych fragmentów obrazu koder często koduje jedynie błąd predykcji wynikający z istnienia szumu, niepotrzebnie tracąc bity na reprezentację szumu.

Często stosowanym podejściem do rozwiązania tego problemu [7] jest odszumienie sekwencji przed kodowaniem. W takim przypadku koder dokonuje decyzji o trybach na podstawie samej treści, a nie szumu, jednak informacja o szumie nie jest w ogóle kodowana. Nie jest to dobre podejście, gdyż szum jest istotny percepcyjnie, zaś jego brak powoduje u odbiorców poczucie nienaturalności obrazu. Ponadto, proces odszumiania obarczony jest błędem i w jego trakcie usunięte mogą zostać istotne

elementy z obrazu, niebędące szumem, np. drobne detale itp. Z literatury nie są znane techniki, które wykorzystywałyby odszumianie sekwencji dla wyboru trybów a jednocześnie kodowałyby informacje o szumie.

2. PROPONOWANA IDEA

W niniejszej pracy proponowane jest wykorzystanie nowatorskiego podejścia hybrydowego. Jego celem jest uniknięcie problemu z bezpośrednim kodowaniem sekwencji odszumionej, które prowadzi do utraty informacji o szumie, a jednocześnie zapewnienie lepszego doboru trybów, nieobciążonego występowaniem szumu. Główna idea propozycji polega na tym, że wybór trybów w koderze dokonywany jest na podstawie sekwencji pozbawionej szumu (odszumionej), jednakże faktycznie kodowany jest obraz zawierający szum i to próbki błędu predykcji tej sekwencji zostają umieszczone w finalnym strumieniu produkowanym przez koder.



Rys. 1. Ogólny schemat proponowanego kodowania ze sterowaniem trybami kodowania sekwencji odszumionej.

Realizacja proponowanej metody (Rys. 1) wykorzystuje dwa zmodyfikowane kodery pracujące na dwóch wersjach kodowanej sekwencji: wejściowej i odszumionej. Pierwszym koderem jest pełen koder z wyłączonym blokiem kodowania entropijnego, który kodując sekwencje odszumioną podejmuje decyzje dotyczącą trybów kodowania. Tryby, podobnie jak wspomniano we wstępie, optymalizowane są lokalnie na poziomie jednostek LCU, jednakże dzięki pracy na obrazie odszumionym, decyzje kodera nie są obciążone występowaniem szumu. Drugim koderem jest uproszczony koder, który nie wykonuje kosztownego poszukiwania trybów kodowania a jedynie koduje sekwencję wejściową dokładnie tymi samymi trybami, które wyznaczone zostały przez pierwszy koder. Tak więc tryby wybrane przez koder pracujący na sekwencji odszumionej przekazywane są do kodera pracującego na sekwencji wejściowej w wersji oryginalnej. W wyniku tego procesu powstaje strumień reprezentujący sekwencję oryginalną, jednakże wykorzystujący tryby kodowania pochodzące z kodowania sekwencji odszumionej.

3. IMPLEMENTACJA

Aby umożliwić sprawdzenie poprawności i ocenę skuteczności zaproponowanej metody, zaimplementowano ją dokonując modyfikacji oprogramowania modelowego HM [HM] (wersja 13.0) techniki HEVC opracowanego przez grupę JCT-VC. Bezpośrednia realizacja proponowanej metody (Rys. 1) prowadziłaby do wykorzystania dwóch koderów do kodowania pojedynczej

sekwencji. Taka nadmiarowość nie jest konieczna, co jest omówione poniżej.

Jak łatwo zauważyć, koder pierwszy (kodujący sekwencję odszumioną) nie generuje strumienia binarnego, tak więc koder ten nie potrzebuje bloku kodowania entropijnego. W przypadku techniki HEVC, kodowanie entropijne CABAC jest bardzo złożone i odpowiada za znaczną część obliczeń prowadzonych przez koder. Z drugiej strony, drugi koder (kodujący sekwencję oryginalną) pozbawiony jest mechanizmu poszukiwania trybów, gdyż wybór trybów dokonywany jest przez koder pierwszy. Poszukiwanie trybów, które zawiera w sobie między innymi poszukiwanie wektorów ruchu, stanowi pod względem złożoności praktycznie drugą, komplementarną część do kodowania entropijnego. Można więc powiedzieć, że sumaryczna złożoność obu wykorzystywanych w propozycji koderów jest zbliżona do złożoności pojedynczego, pełnego kodera. Do realizacji redukcji szumu w pracy wykorzystano istniejący pakiet oprogramowania mv-tools [4]. Umożliwia on uzyskanie odszumionej sekwencji wizyjnej z kompensacją ruchu w czasie zbliżonym do rzeczywistego.

4. WYNIKI EKSPERYMENTALNE

Weryfikację eksperymentalną proponowanej techniki przeprowadzono z wykorzystaniem zbioru sekwencji testowych opracowanego przez grupę MPEG / JCT-VC podczas prac nad techniką HEVC [CTC]. Sekwencje te należą do pięciu klas (A,B,C,D,E) w zależności od rozdzielczości, a co za tym idzie – wymaganej przepływności (Tab. 1):

Tab. 1. Opis wykorzystanych sekwencji testowych

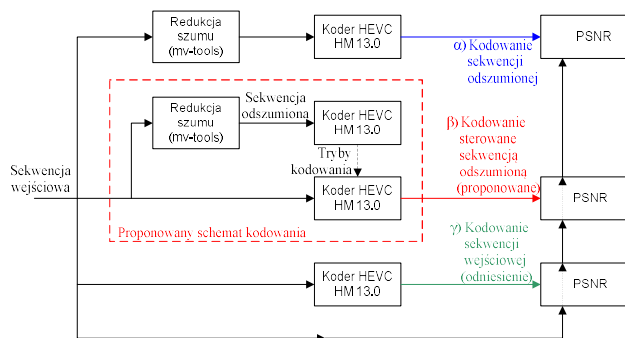
Klasa	Liczba sekwencji	Rozdzielczość
A	3	2560×1600
B	5	1920×1080
C	4	832×480
D	4	416×240
E	3	1280×720

Przykładową ramkę z jednej z wykorzystanych sekwencji (BQTerrace) [2] przedstawiono na Rys. 2a. Na Rys 2bc przedstawiony jest fragment tej samej ramki w wersji oryginalnej (zawierającej szum) i w wersji odszumionej. Aby zwiększyć czytelność różnic, wybrany fragment został powiększony, zaś kontrast został sztucznie zwiększony. Eksperymenty przeprowadzono zgodnie ze schematem przedstawionym na Rys. 3. Każda z sekwencji została zakodowana z wykorzystaniem indeksów QP=[22,27,32,37,42,47] w trzech wariantach: α , β oraz γ . W wariacie (γ) sekwencja wejściowa jest kodowana w wersji oryginalnej przez niezmodyfikowany koder HM 13.0. Wariant ten jest odniesieniem dla pozostałych wyników. W wariacie (α) sekwencja wejściowa jest odszumiana a następnie kodowana przez niezmodyfikowany koder HM 13.0. Wariant ten odzwierciedla stan wiedzy znany z literatury. Wariant (β) jest proponowanym kodowaniem sekwencji wejściowej z wykorzystaniem trybów kodowania sekwencji odszumionej. Wykorzystywane są tutaj dwa zmodyfikowane

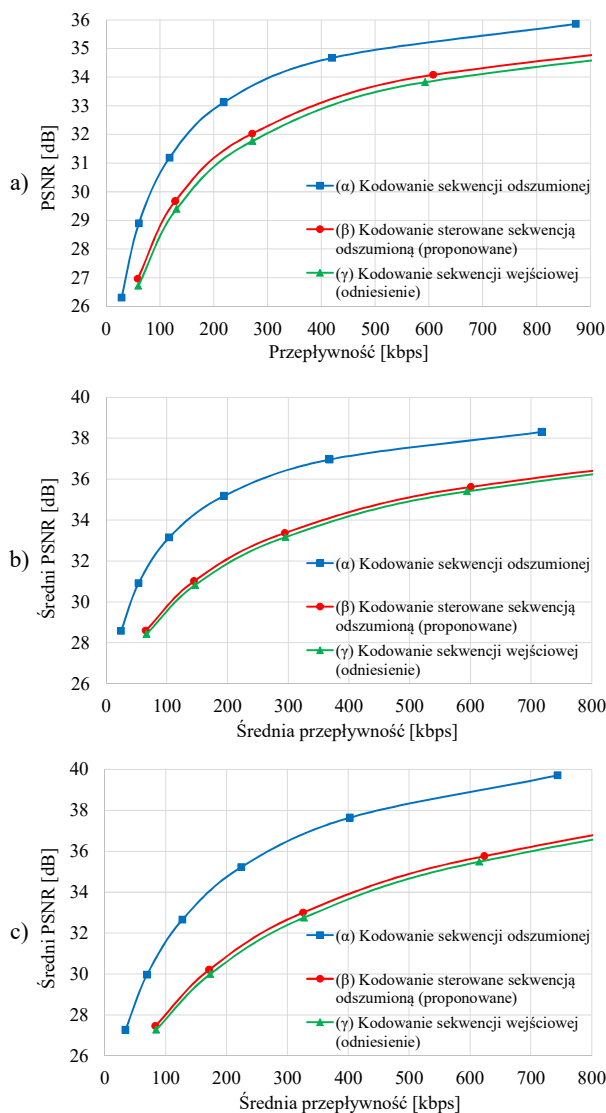
kodery bazujące na HM 13.0, których sumaryczna złożoność jest zbliżona do jednego, niezmodyfikowanego kodera HM 13.0. Wyniki pracy koderów w każdym z wariantów, w postaci zrekonstruowanego sygnału, porównano z oryginalną sekwencją wejściową (z szumem) z pomocą miary obiektywnej PSNR.



Rys. 2. Przykładowa ramka z wykorzystanej sekwencji BQTerrace (a), zbliżony kadr ze zwiększonym kontrastem (b) oraz jego wersja odszumiona (c).



Rys. 3. Schemat przeprowadzonych badań eksperymentalnych.



Rys. 4. Porównanie wydajności kodowania sterowanego sekwencją odszumioną (β - proponowane) z bezpośrednim kodowaniem sekwencji wejściowej (γ - odniesienie) i kodowaniem sekwencji odszumionej (α). Krzywe przepływność-zniekształcenie:

a) dla przykładowej sekwencji z klasy B – BQTerrace, b) uśrednione po wszystkich sekwencjach klasy B, c) uśrednione po wszystkich sekwencjach testowych.

Wybrane z uzyskanych wyników, w postaci krzywych zniekształcenie-przepływność, zostały przedstawione na Rys. 3abc, kolejno dla przykładowej sekwencji z klasy B (BQTerrace), średniej dla sekwencji z klasy B, oraz średniej dla wszystkich rozpatrywanych sekwencji testowych. Wyznaczono również stopień redukcji wymaganej prędkości bitowej względem wariantu odniesienia (γ). W tym celu posłużono się miarą Bjøntegaard'a $\Delta(\text{BD-RATE})$ [1] dla QP 27.42 (Tab. 2). Ujemne wartości $\Delta(\text{BD-RATE})$ wskazują na redukcję prędkości bitowej. Zauważyć można, że uzyskane wyniki są spójne dla wszystkich klas. Uzyskiwana redukcja prędkości bitowej dla proponowanej metody to około 1,5%. Znana z literatury metoda kodowania sekwencji odszumionej (a więc ztracająca informację o szumie) przynosi znacznie większą oszczędność prędkości bitowej sięgając średnio około 60%. Wskazuje to, iż proponowane

sterowanie wykorzystujące odszumioną wersję sekwencji wejściowej bliższe jest optymalnemu wyborowi trybów kodowania, niż sterowanie wykorzystujące sekwencję wejściową zawierającą szum, jednakże nie pozwala na uzyskanie redukcji prędkości bitowej porównywalnej z przypadkiem całkowitego pominięcia szumu.

Tab. 2. Stopień redukcji wymaganej prędkości bitowej, przedstawiony jako delty Bjøntegaard'a $\Delta(BD-RATE)$ [1]

Klasa	Sekwencja	$\Delta(BD-RATE)$	
		Proponowane (β) vs odniesienie (γ)	Odszumione (α) vs odniesienie (γ)
A	Traffic	-1,77%	-58,93%
A	PeopleOnStreet	-1,44%	-64,21%
A	SteamLocomotive	-1,31%	-20,62%
Średnia - klasa A		-1,51%	-47,92%
B	Kimono1	-1,01%	-62,80%
B	ParkScene	-1,50%	-67,47%
B	Cactus	-1,80%	-61,02%
B	BQTerrace	-2,40%	-49,08%
B	BasketballDrive	-1,74%	-78,37%
Średnia - klasa B		-1,69%	-63,75%
C	RaceHorses	-1,44%	-63,09%
C	BQMall	-1,43%	-67,04%
C	PartyScene	-1,61%	-49,09%
C	BasketballDrill	-1,88%	-63,37%
Średnia - klasa C		-1,59%	-60,65%
D	RaceHorsesLow	-0,30%	-58,30%
D	BQSquare	-1,88%	-56,77%
D	BlowingBubbles	-1,70%	-62,31%
D	BasketballPass	-1,21%	-72,50%
Średnia - klasa D		-1,27%	-62,47%
E	FourPeople	-1,88%	-56,35%
E	Johnny	-2,50%	-56,08%
E	KristenAndSara	-2,03%	-55,95%
Średnia - klasa E		-2,14%	-56,13%
Średnia - całość		-1,62%	-59,12%

Tab. 3. Czas kodowania pojedynczej ramki

Klasa	(γ) sam koder (odniesienie)	(α) odszu- mianie+koder	(β) odszumia- nie+dwa kodery (proponowane)
A	130,1	133,4	133,7
B	153,3	155,1	155,4
C	39,5	41,2	41,3
D	12,7	13,5	13,9
E	40,4	41,7	42,1
Średnia	72,0	73,8	74,2

Warto nadmienić, iż dzięki zastosowaniu optymalizacji wspomnianych w punkcie 3, złożoność proponowanego rozwiązania jest praktycznie taka sama jak złożoność pojedynczego kodera HM, co pokazano w Tab. 3.

5. PODSUMOWANIE

W artykule przedstawiono nowatorską metodę sterowania wyborem trybów w koderze. Metoda ta wykorzystuje informacje o wybranych trybach pochodzące z drugiego kodera, który równolegle dokonuje kodowania odszumionej wersji sekwencji wejściowej. Zaproponowaną metodę zaimplementowano w koderze HM (wersja 13.0) będącym oprogramowaniem modelowym

MPEG/JCT-VC dla techniki kodowania HEVC. Z pomocą stworzonej implementacji, eksperymentalnie zbadano skuteczność proponowanej metody z wykorzystaniem zbioru sekwencji testowych MPEG, szeroko stosowanego w literaturze. Uzyskane wyniki pokazują, iż zaproponowana metoda umożliwia około 1,5% redukcję strumienia bitowego. Wskazuje to, iż sterowanie wykorzystujące odszumioną wersję sekwencji wejściowej bliższe jest optymalnemu wyborowi trybów kodowania, niż sterowanie wykorzystujące sekwencję wejściową zawierającą szum.

6. INFORMACJA

Przedstawione wyniki są rezultatem prac badawczych prowadzonych w ramach projektu LIDER Narodowego Centrum Badań i Rozwoju NCBiR. Numer projektu: LIDER/023/541/L-4/12/NCBR.

LITERATURA

- [1] Bjøntegaard G. 2001. „Calculation of average PSNR differences between RD curves”. ITU-T SG16/Q6, Doc. VCEG-M33.
- [2] Bossen F.. 2012. „Common test conditions and software reference configurations”. Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Doc. JCTVC-J1100.
- [3] Bossen F., Bross B., Suhring K., Flynn D. 2012. „HEVC Complexity and Implementation Analysis”. IEEE Transactions on Circuits and Systems for Video Technology, 22 (12) : 1685–1696.
- [4] Fizick M., Balakhnin A., Schniede T. dostęp: 01.12.2013. “Mv-tools web-page”, <http://avisynth.org.ru/mvtools/mvtools.html>.
- [5] „HM 15.0 Reference Software”. 2014. 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-13.0.
- [6] Ohm J.-R., Sullivan G. J., Schwarz H., Tan T.K., Wiegand T.. 2012. „Comparison of the coding efficiency of video coding standards - Including High Efficiency Video Coding (HEVC)”. IEEE Transactions on Circuits and Systems for Video Technology, 22 (12) : 1669–1684.
- [7] Song B.C., Chun K.W. 2004. “Motion-compensated temporal prefiltering for noise reduction in a video encoder”. 2004 International Conference on Image Processing, ICIP '04: 1221-1224.
- [8] Sullivan G.J., Ohm J.-R., Han W.-J., Wiegand T.. 2012. „Overview of the High Efficiency Video Coding (HEVC) Standard”. IEEE Transactions on Circuits and Systems for Video Technology, 22 (12) : 1649–1668.