

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



Poznań, 21-23 czerwca 2017

PRZYSPIESZENIE ESTYMACJI RUCHU ORAZ WYBORU PODZIAŁU CU W KODERZE HEVC

ACCELERATION OF MOTION ESTIMATION AND CU DIVISION SELECTION IN A HEVC CODER

Streszczenie: W artykule przedstawione zostały wyniki prac nad wykorzystaniem dodatkowych danych uzyskanych na wcześniejszych etapach przetwarzania sekwencji wizyjnych w celu usprawnienia procesu kodowania sekwencji wizyjnych koderem HEVC. Opisane zostały metody przyspieszenia procesu wyszukiwania wektorów ruchu jak też procesu podziału bloków CTU na jednostki CU.

Abstract: This paper is a summary of the research done on the use of additional data that are obtained during the pre-processing phases of a video sequence for improving the performance of the coding process using HEVC coder. It includes the description of the method for accelerating the motion estimation process as well as the CTU block division process into CU units.

Słowa kluczowe: kodowanie HEVC, sekwencje wizyjne, wideo, estymacja ruchu, podział CU

Keywords: HEVC coding, video sequences, motion estimation, CU division

1. WSTĘP

Kodowanie sekwencji wizyjnych z użyciem kodera HEVC [2, 7] jest czasochłonne. Nieosiągalna dla wcześniejszych metod kodowania sekwencji wizyjnych efektywność kodowania okupiona jest wielokrotnym wydłużeniem czasu kodowania sekwencji, w porównaniu ze starszymi metodami. Jest to na tyle dotkliwe, że aby ułatwić rozpowszechnienie techniki HEVC, konieczne jest przyspieszenie procesu kodowania. W ostatnich dwóch latach problem przyspieszenia kodowania HEVC jest niezwykle silnie eksploatowanym polem badań naukowych i obszarem licznych wdrożeń opracowanych metod. Najważniejszym celem, stawianym przez badaczy, jest uzyskanie wyraźnego zmniejszenia złożoności obliczeniowej procesu kodowania międzyobrazowego, w tym sposobu wyznaczania wektorów ruchu oraz sposobu wyboru podziału bloków kodowania. Poza metodami wykorzystywanymi w koderze modelowym HM [5], w literaturze opisywane są również inne metody. W artykule [4] i wcześniejszym [3], opisana została metoda wyznaczania podziałów bloków kodowania na podstawie złożonej ana-

lizy statystycznej. Dane statystyczne potrzebne do prawidłowego działania tej metody wymagają jednak fazy treningowej dla każdej sekwencji. Autorzy pokazują, że ich metoda daje do 60% przyspieszenia procesu kodowania przy nieznacznym wzroście prędkości bitowej generowanego strumienia. W artykule [9] autorzy prezentują metodę przyspieszenia procesu wyboru sposobu kodowania bloków obrazu z wykorzystaniem statystycznych zależności dotyczących sąsiednich bloków, dla których wybrany już został sposób kodowania. Metoda pozwala na zakończenie procesu podziału bloku w momencie uzyskania odpowiednio niskiej wartości funkcji kosztu, uwzględniając funkcje kosztu dla sąsiednich zakodowanych bloków. Przyspieszenie uzyskiwane dzięki tej metodzie wynosi około 30% czasu kodowania, przy nieznacznej utracie jakości. W artykule [8] autorzy prezentują metodę bazującą na obliczaniu przepływu optycznego i obliczaniu wartości hierarchicznej wariancji wartości wektorów ruchu dla bloków obrazu oraz kontekstowej analizie statystycznej podziałów bloków obrazu. Uzyskiwane przyspieszenie wynosi średnio 40% czasu, przy kilkuprocentowym wzroście prędkości bitowej.

2. ZASADA DZIAŁANIA PROPONOWANYCH ALGORYTMÓW

Cechą wyróżniającą propozycję przedstawioną w artykule, niespotykaną w źródłach literaturowych, jest znaczące ograniczenie procesu wyznaczania wektorów ruchu w celu skrócenia czasu kodowania na podstawie wyznaczonego uprzednio pola ruchu. Zamiast dokonywać wyznaczania wektorów ruchu dla każdego fragmentu obrazu niezależnie, w koderze wykorzystuje się raz wyznaczone pole ruchu (w blokach 4x4). Takie pole ruchu można uzyskać jako jeden z rodzajów danych powstających przy pewnych operacjach przetwarzania wstępnego, jak na przykład usuwanie szumu z sekwencji. W takim wypadku nie jest konieczne ponoszenie żadnych dodatkowych nakładów obliczeniowych w celu wykorzystania metod opisywanych w artykule. Dzięki wykorzystaniu tych danych możliwe jest zredukowanie czasu przetwarzania sekwencji poprzez uniknięcie konieczności wykonywania pewnych obliczeń, związanych z estymacją ruchu w koderze oraz z wyborem sposobu kodowania kolejnych bloków

CTU. Są to dwa najbardziej czasochłonne etapy kompresji sekwencji wizyjnej w standardzie HEVC [1]. Złożoność tych procesów znacznie przewyższa złożoność analogicznych procesów w koderze poprzedniej generacji – AVC. Wyznaczone pole ruchu można wykorzystać w procesie kodowania do dwóch celów – wykorzystać dane bezpośrednio lub po niewielkich dodatkowych obliczeniach jako wektory ruchu oraz do ogólnego wyznaczenia sposobu podziału jednostki CTU (*Coding Tree Unit*). Podstawowe informacje o wyszukiwaniu wektorów ruchu i wyznaczaniu podziałów CTU zostaną opisane dalej.

3. KODOWANIE OBRAZU W KODERZE HEVC

3.1. Podział obrazu na jednostki CTU

W trakcie kodowania koderem HEVC, obraz jest dzielony na nienakładające się kwadratowe bloki nazywane jednostkami CTU (*Coding Tree Unit*) o ustalonej wielkości (z reguły 64x64 punkty) [7]. Jednostki CTU zawierają między innymi jednostkę CTB (*Coding Tree Block*), która z kolei zawiera jedną lub więcej jednostek CU (*Coding Unit*) zorganizowanych w strukturę drzewa czwórkowego. Jednostki CU podlegają podziałom na 4 mniejsze jednostki CU. Najmniejsza jednostka CU może mieć rozmiar 8x8 punktów. Koder wielokrotnie koduje dany fragment obrazu z różnym podziałem na jednostki CU i ostatecznie wybiera układ jednostek CU zapewniający jak najwyższą jakość przy jak najniższym zużyciu bitów, zgodnie z przyjętą funkcją optymalizacyjną.

3.2. Wektory ruchu

Każda jednostka CU zawiera co najmniej jedną jednostkę PU (*Prediction Unit*), dozwolone są następujące kombinacje jednostek PU w zależności od ich rozmiaru (względem CU): brak podziału, podział na cztery kwadratowe jednostki PU, podział na dwie prostokątne jednostki PU w pionie lub poziomie, asymetryczny podział na dwie prostokątne jednostki PU w pionie lub poziomie, zatem dla jednostki CU o rozmiarze 16x16 punktów PU mogą mieć rozmiar 16x16, 4 bloki 8x8, 2 bloki 8x16, 2 bloki 16x8 punktów, 4x16 + 12x16, 12x16 + 4x16, 16x4 + 16x12 lub 16x12 + 16x4. Dla każdej jednostki PU trzeba wyznaczyć jeden lub więcej wektorów. Jest to niezwykle czasochłonne, gdyż musi to być powtarzane dla każdego z możliwych podziałów jednostki CTU na CU i PU. W tym kontekście estymacja ruchu w koderze HEVC jest znacznie bardziej złożona niż dla koderów AVC, choćby ze względu na dużo więcej koniecznych do sprawdzenia sposobów podziału. W koderze HEVC, dla każdej jednostki CTU możliwych jest około 100 wariantów podziału na jednostki CU.

3.3. Sposoby estymacji ruchu w HEVC

Norma HEVC nie opisuje sposobu estymacji wektorów ruchu, a jedynie opisuje sposób kodowania danych o ruchu wewnątrz strumienia danych. Możliwe jest zatem użycie w koderze dowolnej metody wyznaczenia wektorów ruchu. Oprogramowanie modelowe koderów techniki HEVC udostępnia dwa algorytmy poszukiwania wektorów ruchu. Pierwszy to pełne (*full search*) przeszukiwanie zdefiniowanego obszaru (na przykład $\pm 64 \times 64$ punkty obrazu od pozycji startowej przeszukiwania). Jest to metoda

niezwykle czasochłonna, lecz dająca najlepsze dopasowanie w ramach obszaru przeszukiwania. Alternatywnym algorytmem poszukiwania, wykorzystywanym przy domyślnych ustawieniach koderów, jest tak zwany algorytm TZ search [6]. Jest to złożona modyfikacja algorytmu przeszukiwania logarytmicznego, zawierająca warunki zakończenia przeszukiwania zależne od wartości danych wejściowych. Jest to zatem algorytm o czasie przeszukiwania zależnym od danych. Algorytm pozwala na wielokrotne skrócenie czasu potrzebnego na odnalezienie wektorów ruchu, przy zachowaniu bardzo dobrej dokładności i efektywności kompresji, mimo, że algorytm ten nie gwarantuje odnalezienia optymalnych wektorów ruchu. Estymacja ruchu w oprogramowaniu modelowym koderów HEVC wykonywana jest z dokładnością do $\frac{1}{4}$ próbki obrazu, jednak jest to proces hierarchiczny – po odnalezieniu wektora ruchu z dokładnością do pełnego okresu próbkowania (*full pel*) obrazu następuje przeszukiwanie jego bezpośredniego sąsiedztwa, dzięki czemu uzyskuje się wektor o dokładności do $\frac{1}{2}$ punktu, a następnie w ten sam sposób wektor ruchu jest uściślany do $\frac{1}{4}$ punktu obrazu. Na etapie estymacji pół- i ćwierć-punktowej wykorzystywany jest obraz interpolowany przy użyciu zdefiniowanych filtrów interpolacyjnych.

4. PROPOZYCJA PRZYSPIESZONEJ METODY ESTYMACJI RUCHU

Pierwszą proponowaną modyfikacją jest zastąpienie estymacji wektora ruchu prowadzonej niezależnie dla każdego możliwej jednostki PU przez użycie wektora ruchu odczytanego bezpośrednio z raz wyznaczonego pola ruchu. Odczytany wektor ruchu musi być odpowiednio przemnożony, aby umożliwić poprawną predykcję dla różnych ramek znajdujących się na listach odniesienia. W celu zwiększenia dokładności wyznaczonego pola ruchu podczas kodowania danej jednostki PU dokonywane jest bardzo proste doszukiwanie wektora ruchu o dokładności $\frac{1}{4}$ punktu obrazu. Ponieważ jest wykonywane tylko dla jednego wektora ruchu, ma niewielki wpływ na prędkość kodowania.

5. OGRANICZENIE GŁĘBOKOŚCI PRZESZUKIWANIA DZREWA PODZIAŁÓW CTU

Jednym z etapów kompresji, które zajmują znaczącą część czasu kodowania koderem techniki HEVC jest wybór sposobu podziału jednostki CTU na CU, oraz dalej - jednostki CU na jednostki PU. Analizując powody dla których dana jednostka CU powinna zostać podzielona na wiele jednostek PU można skonkludować, iż fragmenty obrazu, które zawierają obiekty które poruszają się niezależnie od siebie powinny zostać zakodowane jako oddzielne jednostki predykcji PU. Z analizy pola ruchu można pośrednio wyciągnąć wnioski na temat sposobu poruszania się obiektów. Jeśli fragmenty rozważanej jednostki CU mają przypisaną znacząco różne wektory ruchu (zarówno pod względem amplitudy jak i kierunku), można domniemywać iż reprezentują różne obiekty i wydajniej będzie zakodować je jako oddzielne jednostki predykcji PU. W związku z tym, analizując wyznaczone pole ruchu, można stworzyć a priori (przed kodowaniem)

mapę podziałów na jednostki CU i PU. W takim przypadku koder nie dokonuje czasochłonnego wielokrotnego sprawdzania wariantów podziału na bloki (jednostki), a stosuje bezpośrednio podział obrazu zgodnie z wyznaczoną wcześniej mapą podziałów. Stworzenie takiej mapy podziałów wymaga opracowania algorytmu, który, na podstawie znajomości wartości w mapie pola ruchu dla obszaru odpowiadającego obszarowi analizowanego CTU, opracuje najlepszy przewidywany sposób podziału na bloki CU oraz podział każdego z bloków CU na bloki PU. W proponowanym algorytmie wykorzystanych może być wiele różnych sposobów określenia, czy dany blok CTU lub CU powinien zostać podzielony, czy nie. Jednym z możliwych rozwiązań jest obliczenie wariancji wartości pola ruchu w obszarze obejmowanym przez rozważane CTU lub CU. Przekroczenie wartości progowej oznacza konieczność przeprowadzenia podziału bloku. Kwestia doboru sposobu podziału CTU jest złożona, zależy nie tylko od wartości punktów obszaru CTU, lecz również od innych czynników, takich jak kontekst kodowania, efektywność zakodowania informacji o poszczególnych podziałach oraz efektywność kodowania odpowiednich wektorów ruchu. Dla każdego CTU wybierany jest optymalny sposób podziału z punktu widzenia wielu kryteriów. W proponowanej metodzie zakłada się, że sposób podziału CTU na CU jest zależny głównie od wartości pola ruchu dla danego CTU. Jest to oczywiście znaczne uproszczenie względem optymalizacji wykonywanej w koderze. W badaniach przyjęto, że podziały CTU i CU dopuszczane będą w przypadku, gdy wariancja wartości próbek pola ruchu przyporządkowanych danemu CTU lub CU przekracza określoną wartość progową. W takiej sytuacji koder przeprowadza obliczenia dla aktualnej wielkości CU jak i dla mniejszych jednostek CU. Decyzje o kolejnych podziałach dokonywane są w analogiczny sposób. Na podstawie wyników badań można zaobserwować, że metoda taka prowadzi do znacznego ograniczenia czasu kompresji sekwencji. Zysk wynika przede wszystkim ze znacznego ograniczenia liczby przypadków, dla których koder sprawdza kodowanie dla małych podziałów, choć zysk z zakodowania bloku z użyciem mniejszych CU jest nieznaczny. Ograniczając głębokość przeszukiwania w drzewie podziałów dla CTU uzyskujemy, kosztem zmniejszenia efektywności kodowania, znaczne przyspieszenie procesu kodowania.

6. METODOLOGIA BADAŃ I WYNIKI

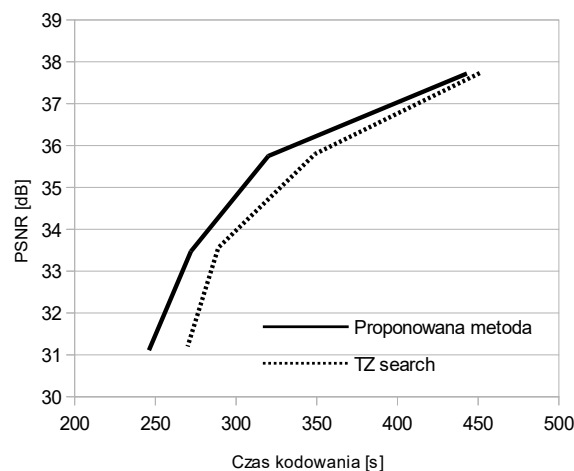
W celu przeprowadzenia eksperymentalnej weryfikacji proponowanych metod, przeprowadzone zostały testy z wykorzystaniem sekwencji testowych rekomendowanych przez zespół ekspertów ISO/IEC oraz ITU, zespół, który opracował technikę HEVC. Na potrzeby badań wykorzystane zostały sekwencje w rozdzielczości 1920x1080 punktów: „Cactus” oraz „BasketballDrive”. Dla sekwencji tych dostępne były mapy pola ruchu (oddzielnie dla współrzędnych poziomych i pionowych wektorów ruchu) o rozdzielczości 480x270 punktów, powstałe w ramach innych badań dotyczących usuwania szumu z sekwencji wizyjnych. Metoda została zaimplementowana w koderze modelowym HM 13.0 [5], a testy zostały przeprowadzone na komputerze z procesorem Intel i7 taktowanym zegarem 3,4GHz, wyposażonym w

32GB RAM. W trakcie trwania testów komputer nie był obciążony innymi zadaniami, a jednowątkowa aplikacja koderka miała do wyłącznej dyspozycji jeden rdzeń procesora. Wyniki uzyskiwane przy wykorzystaniu metody przyspieszonej estymacji ruchu dają wyraźne przyspieszenie w porównaniu z metodą pełnego przeszukiwania. W porównaniu z algorytmem TZ search metoda daje nieco mniej wyraźne przyspieszenie. W tabeli 1 poniżej przedstawione są przykładowe wyniki.

Tab. 1. Przykładowe wyniki działania koderka z przyspieszoną estymacją ruchu

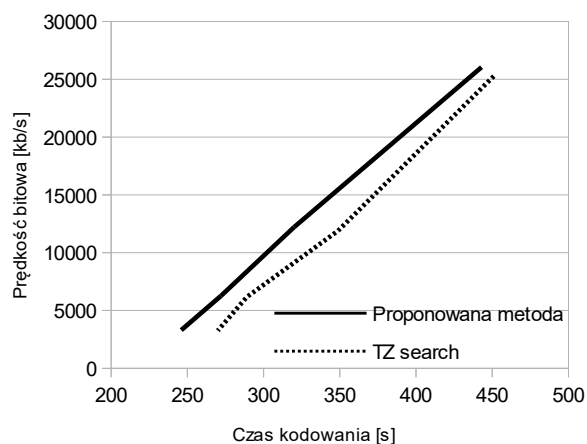
Metoda estymacji	Czas kodowania [s]	PSNR [dB]	Prędkość bitowa [kb/s]
Sekwencja „Cactus”, QP = 30			
Pełne przeszukiwanie	3370	35,81	11 913
TZ search	349	35,80	11 913
Proponowana metoda	320	35,75	12 206

Przykładowe wyniki pokazują, że możliwe jest uzyskanie około 10% redukcji czasu kodowania (w porównaniu do szybkiej metody estymacji ruchu TZ search), przy pomijalnie małej stracie jakości i nieznacznym zwiększeniu prędkości bitowej strumienia (około 3%). Warto zauważyć, że stworzony w ten sposób strumień bitowy jest całkowicie zgodny z normą i daje się poprawnie dekodować. Bardziej ogólne dane dotyczące zaproponowanej metody pokazane są na rysunkach 1 i 2.

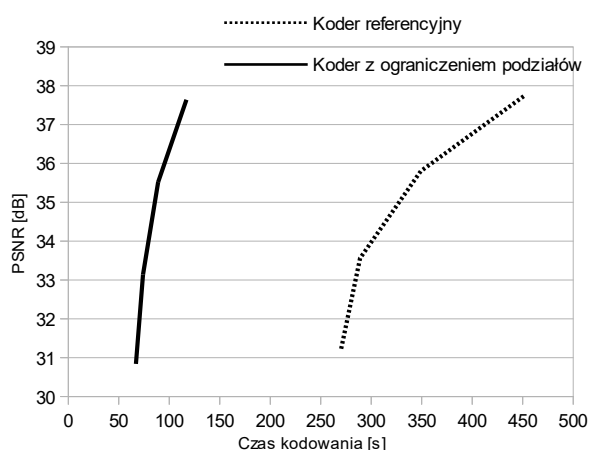


Rys. 1. Porównanie metod estymacji ruchu – jakość sekwencji zakodowanej.

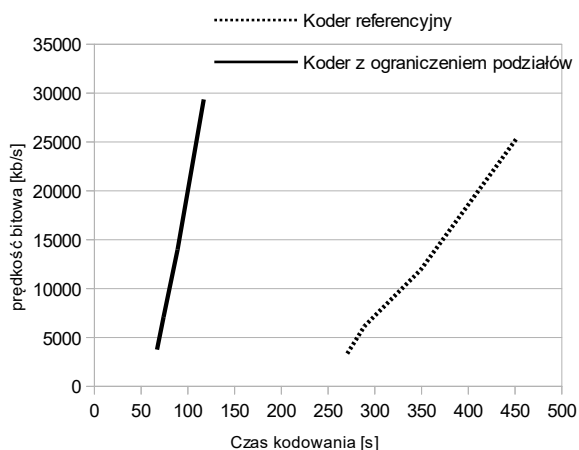
Z kolei wyniki kompresji dla przypadku, w którym wykorzystana została metoda ograniczenia głębokości przeszukiwania drzewa podziałów CTU przedstawione są na rysunkach 3 i 4. Zaprezentowane dane świadczą o możliwości uzyskania aż około czterokrotnego przyspieszenia procesu kompresji sekwencji wizyjnej. Niestety, przyspieszenie to wiąże się ze spadkiem jakości sekwencji o około 1dB dla takiej samej prędkości bitowej strumienia wyjściowego w porównaniu ze standardowym koderem.



Rys. 2. Porównanie metod estymacji ruchu – prędkość bitowa sekwencji zakodowanej



Rys. 3. Wyniki metody ograniczenia głębokości podziałów – jakość sekwencji zakodowanej.



Rys. 4. Wyniki metody ograniczenia głębokości podziałów – prędkość bitowa sekwencji zakodowanej

7. KIERUNKI PRZYSZŁYCH BADAŃ

Zaprezentowane w artykule wyniki badań stanowią dobrą podstawę do dalszych badań, ukierunkowanych na opracowanie szybkich metod kompresji sekwencji wizyjnych z wykorzystaniem kodera HEVC. Metoda wykorzystująca pole ruchu do przyspieszenia kodowania przy niezmięnionej procedurze wyboru sposobu podziału CTU daje ograniczone zyski i już w obecnym stanie daje poprawę bliską maksymalnej, możliwej do osiągnięcia w tej

metodzie. Odmianą jest jednak sytuacja w odniesieniu do metody wykorzystującej pole ruchu do określenia sposobu podziału jednostki CTU na mniejsze jednostki CU. Możliwe jest bowiem uzyskanie z jej wykorzystaniem kodera, dla którego można w łatwy sposób regulować stopień przyspieszenia kodowania. Duże przyspieszenia owocować będą spadkiem efektywności kodowania (gorsza jakość sekwencji rekonstruowanej lub większa prędkość bitowa strumienia), jednak dla mniejszych przyspieszeń należy oczekiwać, że spadek jakości będzie niewielki i możliwy do zaakceptowania.

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] 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.
- [2] ISO/IEC, ITU-T. 2013. Norma „High Efficiency Video Coding (HEVC)”, *ISO/IEC 23008-2 (MPEG-H Part 2) / ITU-T Rec. H.265*.
- [3] Mallikarachchi T., Fernando A., Arachchi H. K. 2014. „Effective coding unit size decision based on motion homogeneity classification for HEVC inter prediction”. *Proceedings of 2014 IEEE International Conference on Image Processing (ICIP)*: 3691-3695.
- [4] Mallikarachchi T., Talagala D. S., Arachchi H. K., Fernando A. 2016. „Content-Adaptive Feature-Based CU Size Prediction for Fast Low-Delay Video Encoding in HEVC” *IEEE Transactions on Circuits and Systems for Video Technology*, przyjęty do publikacji.
- [5] Oprogramowanie modelowe HEVC, dostępne online <https://hevc.hhi.fraunhofer.de/svn/>
- [6] Stankowski J., Karwowski D., Klimaszewski K., Wegner K., Stankiewicz O., Grajek T. 2016. „Analysis of the complexity of the HEVC motion estimation”. *Proceedings of 23th International Conference on Systems, Signals and Image Processing (IWSSIP 2016)*.
- [7] 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.
- [8] Xiong J., Li H., Wu Q., Meng F. 2014. „A Fast HEVC Inter CU Selection Method Based on Pyramid Motion Divergence”. *IEEE Transactions on Multimedia*, 16 (2): 559-564.
- [9] Zhang X., Zhu W., Wang C., Zhang H. 2015. „A fast Coding Tree Unit depth Prediction for HEVC”. *Proceedings of 2015 8th International Congress on Image and Signal Processing (CISP)*: 114-118.