

WPLYW KOMPRESJI OBRAZÓW I MAP GŁĘBI NA SYNTEZĘ WIDOKÓW W SYSTEMIE WIELOWIDOKOWYM

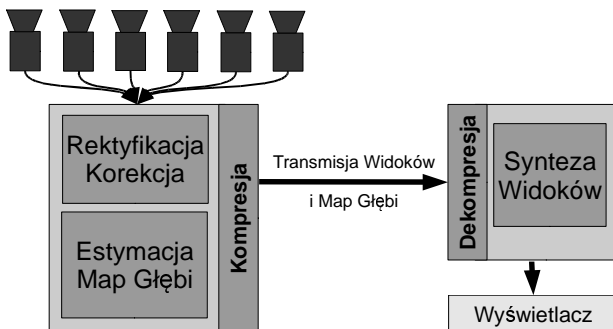
Streszczenie: W artykule zaprezentowano wyniki badań nad wpływem zniekształceń wprowadzanych przez kompresję widoków i map głębi na jakość syntezy widoku w systemach wielowidokowych. Zbadana została przydatność kodera MVC do kompresji map głębi oraz podane zostały praktyczne wnioski na temat podziału prędkości bitowej pomiędzy skompresowane strumienie obrazów i map głębi. W badaniach wykorzystano obiektywną miarę jakości PSNR, przyrównując widoki syntezy do odpowiednich widoków rzeczywistych

1. WSTĘP

Obecnie na świecie prowadzi się intensywne badania nad systemami telewizyjnymi następnej generacji. Jednym z takich systemów są systemy telewizyjnego swobodnego punktu widzenia. W systemach takich użytkownik ma możliwość obserwacji trójwymiarowych obrazów zarejestrowanej sceny z dowolnego wybranego przez siebie miejsca w przestrzeni.

Prace nad telewizją swobodnego punktu widzenia prowadzi między innymi komitet MPEG, działający w ramach organizacji ISO. W ramach jego prac opracowywane są standardy dotyczące telewizji trójwymiarowej i swobodnego punktu widzenia. W prace te zaangażowany jest już od wielu lat zespół z Katedry Telekomunikacji Multimedialnej i Mikroelektroniki Politechniki Poznańskiej, prowadzący własne badania nad nowoczesnym systemem wielowidokowym i opracowujący jego prototyp.

Kompletny system telewizyjnego swobodnego punktu widzenia składa się z bloków przedstawionych na rysunku 1.



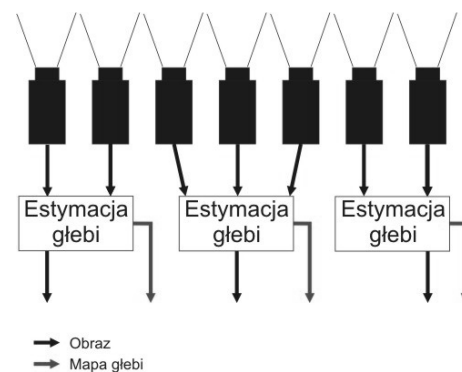
Rysunek 1. Schemat systemu telewizyjnego swobodnego punktu widzenia

Część akwizycyjna składa się z szeregu kamer dokonujących jednoczesnej rejestracji obrazu sceny z różnych miejsc w przestrzeni. Obecnie najczęściej przyjmuje się ustawienie kamer w jednej linii. W module estyma-

cji map głębi dokonuje się wyznaczenia informacji dodatkowej w postaci map głębi, umożliwiającej trójwymiarową reprezentację sceny. Mapy głębi określają odległość poszczególnych punktów sceny, obserwowanych przez kamery, od linii ustawienia kamer. Informacje te wykorzystywane są w procesie kompresji i syntezy nowych widoków. Kompresja zapewnia dostosowanie strumienia danych do prędkości transmisji możliwej do osiągnięcia w danym kanale transmisyjnym. Widoki i informacje dodatkowe rekonstruowane są w odbiorniku. Moduł syntezy widoków tworzy żądany przez widza widok sceny który następnie wyświetlany jest na wyświetlaczu trójwymiarowym – autostereoskopowym lub polaryzacyjnym.

Istnieją dwa schematy kompresji obrazów telewizyjnego wielowidokowego. Pierwszy schemat polega na kodowaniu wszystkich zarejestrowanych widoków sceny. W tym celu obecnie prowadzi się intensywne badania nad standardem kompresji sekwencji wielowidokowych MVC [1]. W pracach tych próbuje się wykorzystać powszechnie znane mechanizmy kompresji hybrydowej, adaptując mechanizmy skalowalności i predykcji z kompensacją ruchu do zastosowań telewizyjnego wielowidokowego.

Drugi schemat polega na wysyłaniu tylko wybranych widoków wraz z informacją dodatkową zawartą w postaci map głębi. Pozostałe widoki rekonstruowane są w odbiorniku. Schemat ten ma szereg zalet, między innymi możliwość łatwej syntezy dowolnego widoku, która jest nieodzowna w systemach telewizyjnego swobodnego punktu widzenia. Dodatkowo możliwość syntezy dowolnego widoku umożliwia wybranie niewielu widoków do transmisji (rysunek 2) i poprzez to zmniejszenie wysłanego strumienia danych.



Rysunek 2. Schemat metody redukcji liczby widoków do przestania

Badania autorów dotyczą wykorzystania drugiego scenariusza. W tym wypadku konieczne jest określenie

wpływu kompresji danych dodatkowych (map głębi) na jakość syntezy obrazów. Ważne jest też określenie jaki procent prędkości transmisji powinien być przeznaczony dla obrazów, a jaki dla informacji dodatkowych w postaci map głębi. Istotną częścią badań jest również zbadanie wpływu decymacji map głębi na proces syntezy widoków. Autorzy przeprowadzili badania nad tymi zagadnieniami i w ich wyniku opracowali metodę wyznaczania optymalnego stosunku podziału strumienia danych pomiędzy obrazy i informacje dodatkowe.

2. KOMPRESJA MAP GŁĘBI

System akwizycji obrazu może składać się z dużej liczby kamer, lecz do przesłania można wybrać jedynie ograniczoną liczbę widoków. Konieczne jest więc opracowanie metod pozwalających na generowanie dodatkowych widoków wirtualnych, potrzebnych do prezentacji w odbiorniku. Konieczne jest zatem przesłanie, wraz z obrazami z kamer, odpowiadających im map głębi. To z kolei rodzi konieczność opracowania efektywnej metody kompresji map głębi, których charakterystyki są nieco odmienne od charakterystyk obrazu. Najistotniejsze różnice zostały wymienione w tabeli 1.

Tabela 1. Porównanie własności map głębi i obrazów

Mapy Głębni	Widoki
tylko jedna składowa	trzy składowe koloru
uboga tekstura	bogactwo tekstur
wyraźne kontury	kontury rozmyte
istotna dokładna wartość	dopuszczalne zmiany wartości

Wśród opisanych do tej pory w literaturze metod kompresji map głębi można wymienić te zawarte w pracach [2] [3] [4]. W niniejszej pracy autorzy skupili się na podejściu wykorzystującym do kompresji map głębi ten sam algorytm, który jest wykorzystywany w kompresji obrazów sekwencji wielowidokowych, czyli opracowywany w ramach prac standaryzacyjnych komitetu MPEG kodek MVC [5]. Podstawową zaletą takiego podejścia jest wykorzystanie istniejącego oprogramowania jak również możliwość zintegrowania procesu kompresji widoków oraz map głębi. Celem pracy jest określenie optymalnego stosunku w jakim całkowita prędkość transmisji powinna być dzielona pomiędzy transmisję widoków i odpowiadających im map głębi. Dodatkowym celem jest zbadanie wpływu zniekształceń wprowadzanych przez wybraną przez autorów metodę kompresji map głębi na jakość syntezowanych w odbiorniku widoków wirtualnych. Podstawowym kryterium przydatności algorytmów kompresji map głębi wybranym przez autorów jest jakość syntezowanych widoków. Wybrane kryterium wynika bezpośrednio z idei telewizji swobodnego widoku, gdzie widzowi prezentowane są jedynie widoki syntezowane i to ich jakość jak najważniejsza.

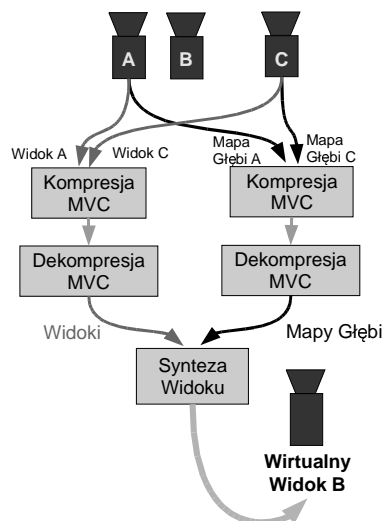
Przeprowadzone w trakcie badań eksperymenty dowiodły, że efektywną metodą zmniejszenia prędkości transmisji przypadającej na mapy głębi przy zachowaniu niezmięnionej jakości syntezowanego widoku jest decymacja mapy głębi. Doświadczenia pokazują, iż dla dwu-

krotnej decymacji (w obu osiach obrazu) mapy głębi uzyskać można ogólną poprawę wyników. Zwiększenie stopnia decymacji do czterech powoduje jednak już ich pogorszenie, zatem w dalszych badaniach użyta została dwukrotna decymacja.

3. UKŁAD DOŚWIADCZALNY

W badaniach wykorzystane zostały standardowe wielowidokowe sekwencje testowe [6][7][8] dostarczane dzięki uprzejmości ich twórców na potrzeby badań komitetu MPEG. Przebieg eksperymentów kształtuje się następująco. Widoki z dwóch wybranych kamer poddawane są kompresji za pomocą kodaera MVC. Podobnie kompresowane są odpowiadające tym widokom mapy głębi, stworzone z wykorzystaniem oprogramowania referencyjnego DERS [9], dostarczonego przez Nagoya University, następnie zmodyfikowanego [10]. Użyto bardzo szerokiego zakresu wartości parametru QP przy kompresji zarówno widoków (w dalszej części nazywanego QP) jak i map głębi (w dalszej części nazywanego QD), aby wyraźnie ustalić wpływ kompresji obu składników na jakość syntezy widoku. Syntezy widoku dokonano za pomocą oprogramowania referencyjnego VSRS [9], dostarczonego przez Nagoya University. Wynikowe widoki porównywane są z oryginalnymi widokami, uzyskanymi z rzeczywistej kamery umieszczonej dokładnie w punkcie dla którego generowany jest widok syntetyczny w odbiorniku.

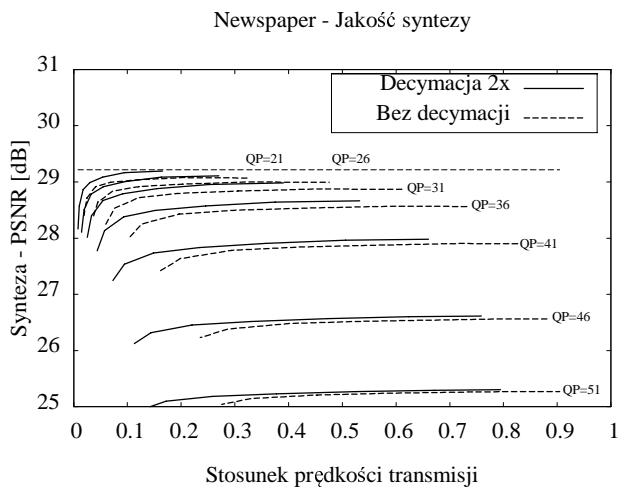
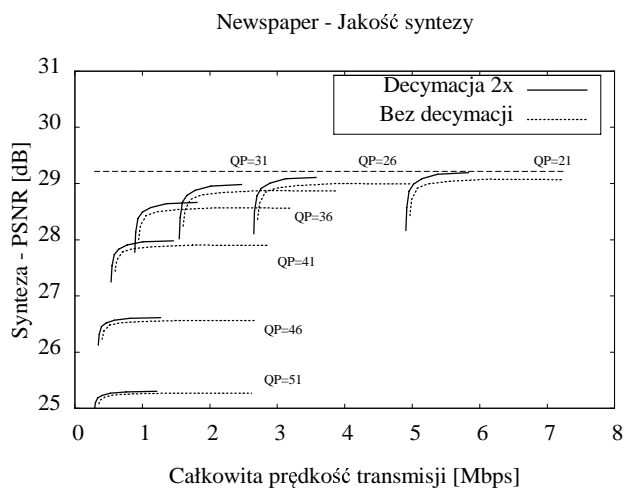
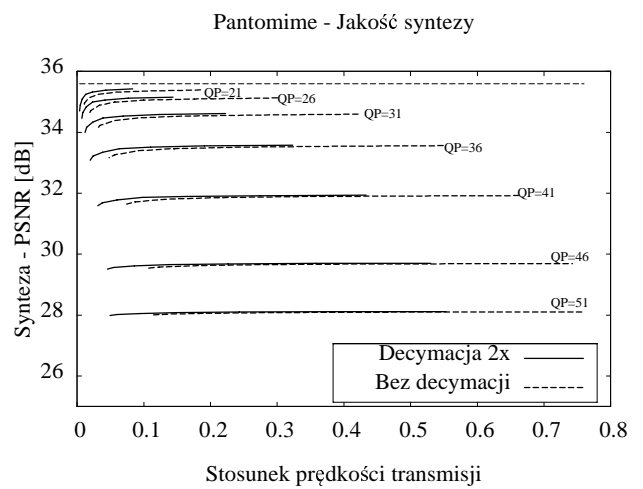
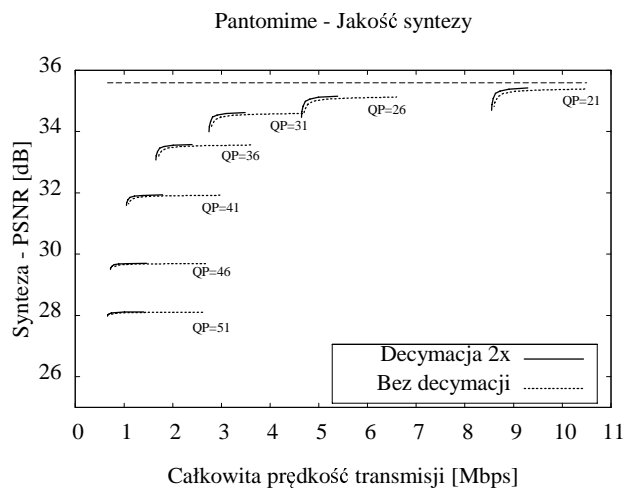
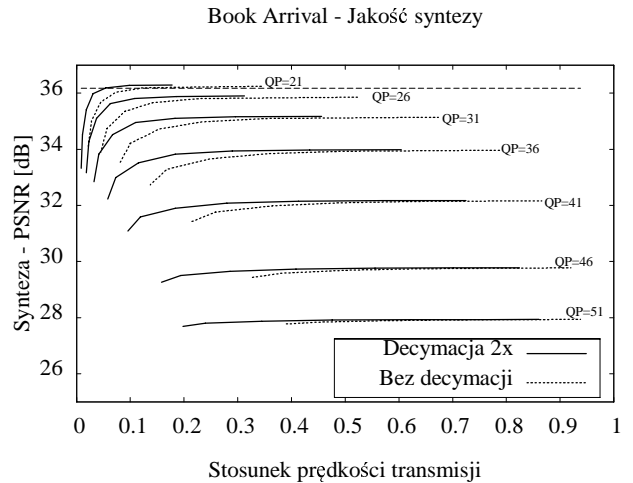
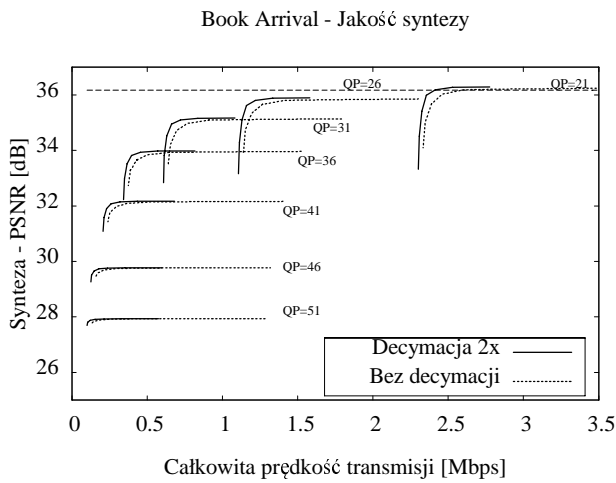
Zastosowany w badaniach schemat przepływu danych przedstawiony jest na rysunku 3.



Rysunek 3. Schemat przepływu danych w eksperymencie

4. WYNIKI

Uzyskane przez autorów wyniki przedstawione są na rysunkach 4, 5 i 6. Przedstawione na nich są wykresy zależności jakości wyrażonej w mierze PSNR syntezowanych widoków w funkcji całkowitej prędkości transmisji dla całego strumienia (2 widoki i 2 mapy głębi) rysunek 4 i w funkcji stosunku prędkości transmisji zdefiniowanego jak we wzorze (1) rysunek 5.



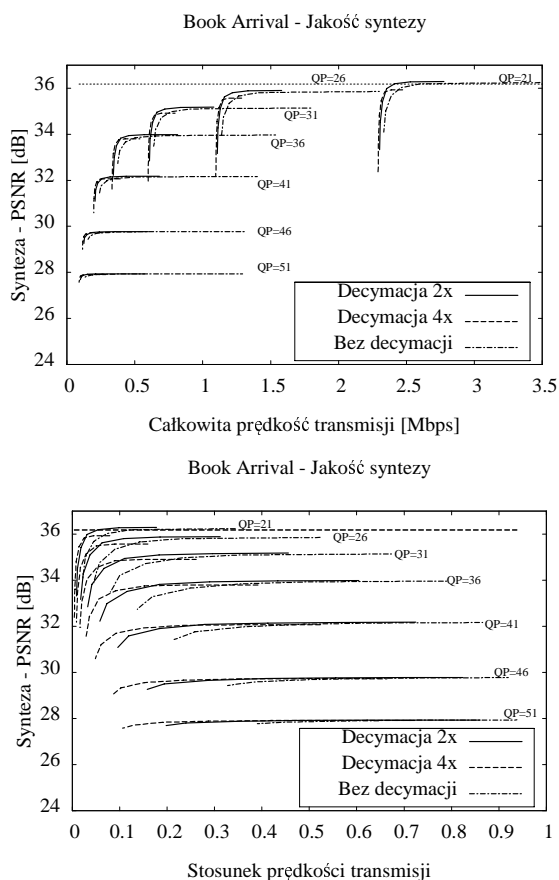
Rysunek 4. Wykresy zależności jakości syntezy widoku od wartości całkowitej prędkości transmisji (2 widoki + 2 mapy głębi). Pozioma przerywana linia oznacza poziom jakości uzyskany z użyciem nieskompresowanych widoków i map głębi

Rysunek 5. Wykresy zależności jakości syntezy widoku od stosunku prędkości transmisji opisanego wzorem (1). Pozioma przerywana linia oznacza poziom jakości uzyskany z użyciem nieskompresowanych widoków i map głębi

$$\text{stosunek prędkości} = \frac{\text{pr. transmisji mapy głębi}}{\text{pr. transmisji widoku} + \text{pr. transmisji mapy głębi}} \quad (1)$$

Pozioma, przerywana linia na wykresach odpowiada jakości widoku syntezerowanego z użyciem nieskompresowanych widoków i nieskompresowanych map głębi. Kolejne krzywe obrazują jakości syntezerowanych obrazów przy zmieniających się wartościach parametrów QP i QD kompresji. Przy kolejnych krzywych podana jest wartość parametru QP przy jakim otrzymane zostały dane krzywe. Przemierzając się wzdłuż pojedynczej krzywej zmienia się wartość parametru QD.

Daje się zauważyć, że zmiany parametru QD poniżej pewnej wartości progowej (im mniejsza wartość parametru QD tym wyższa jakość) nie wpływają zauważalnie na jakość obrazu syntezerowanego. Z kolei widoczny jest duży wpływ jakości skompresowanych widoków na wyniki syntezy. Ogólnym wnioskiem jest więc obserwacja, że w większości przypadków znacznie ważniejsza jest jakość skompresowanych widoków niż jakość map głębi. Kolejnym ważnym spostrzeżeniem jest fakt poprawy wyników w przypadku użycia zdecydowanych dwukrotnie map głębi. Najwyraźniej różnicę między mapami głębi zdecydowanymi i niezdecydowanymi widać dla wartości QD większych niż wspomniana wartość progowa. Zwiększenie stopnia decymacji do 4 powoduje znaczny spadek jakości syntezerowanego obrazu, co uwidocznione zostało na rysunku 6.



Rysunek 6 Zależność jakości syntezy widoku dla sekwencji Book Arrival od całkowitej prędkości transmisji i od stosunku prędkości transmisji przy różnych wartościach stopnia decymacji

5. WNIOSKI

Uzyskane wyniki pozwalają na wysnucie wniosku, że w przypadku przesyłania widoków wraz z mapami głębi w systemach wielokamerowych znacznie większy wpływ na finalną jakość syntezerowanych w odbiorniku widoków ma jakość przesyłanych widoków, natomiast jakość mapy głębi jest mniej istotna. Okazuje się również, że przy transmisji mapy głębi dobre rezultaty otrzymuje się poprzez dokonanie dwukrotnej decymacji mapy głębi, co powoduje zauważalne zmniejszenie wymagań co do prędkości transmisji przeznaczanej na przesłanie mapy głębi, lecz nie powoduje pogorszenia jakości syntezerowanych widoków. Podział prędkości transmisji pomiędzy widoki i mapy głębi powinien zawierać się w przedziale 10%-30% dla mapy głębi, jeśli nie jest ona decymowana i 5%-20% dla decymowanej mapy głębi. To z kolei jest jasną przesłanką dla stosowania dwukrotnej decymacji głębi w procesie jej kompresji.

SPIS LITERATURY

- [1] ISO/IEC 14496-10:2008/FDAM 1:2008(E), Information technology – Coding of audio-visual objects – part 10: Advanced Video Coding, Amendment 1: Multiview Video Coding.
- [2] P. Merkle, Y. Morvan, A. Smolic, D. Farin, K. Müller, P.H.N. de With, T. Wiegand, “The Effect of Depth Compression on Multiview Rendering Quality”, Proceedings of 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, str. 245-248, 28-30 maja 2008.
- [3] P. Merkle, A. Smolic, K. Müller, and T. Wiegand, “Multi-view Video Plus Depth Representation and Coding”, ICIP 2007, IEEE International Conference on Image Processing, San Antonio, TX, USA, wrzesień 2007.
- [4] Y. Morvana, P.H.N. de Witha, D. Farina, “Platelet-Based Coding of Depth Maps for the Transmission of Multiview Images”, Proceedings of SPIE, Stereoscopic Displays and Applications, vol. 6055 str. 93-100, styczeń 2006, San Jose (CA), USA
- [5] P. Pandit, Y. Chen, S. Yea, “Text of ISO/IEC 14496-5:2001/PDAM 15 Reference Software for Multiview Video Coding”, ISO/IEC JTC1/SC29/WG11 MPEG 2008/W9974 Hanower, Niemcy, lipiec 2008.
- [6] I. Feldmann, M. Mueller, F. Zilly, R. Tanger, K. Mueller, A. Smolic, P. Kauff, T. Wiegand, “HHI Test Material for 3D Video”, ISO/IEC JTC1/SC29/WG11 MPEG 2008/M15413, Archamps, Francja, kwiecień 2008.
- [7] Yo-Sung Ho, Eun-Kyung Lee, Cheon Lee, “Multiview Video Test Sequence and Camera Parameters”, ISO/IEC JTC1/SC29/WG11 MPEG 2008/M15419, Archamps, Francja, kwiecień 2008.
- [8] M. Tanimoto, T. Fujii, N. Fukushima, “1D Parallel Test Sequences for MPEG-FTV”, ISO/IEC JTC1/SC29/WG11 MPEG2008/M15378, Archamps, Francja, kwiecień 2008.
- [9] M. Tanimoto, T. Fujii, K. Suzuki, N. Fukushima, Y. Mori, “Reference Softwares for Depth Estimation and View Synthesis”, ISO/IEC JTC1/SC29/WG11 MPEG 2008/M15377, Archamps, Francja, kwiecień 2008.
- [10] Sang-Beom Lee, Yo-Sung Ho, “Enhancement of Temporal Consistency for Multi-view Depth Map Estimation”, ISO/IEC JTC1/SC29/WG11 MPEG 2008/M15594, Hanower, Niemcy, lipiec 2008.