

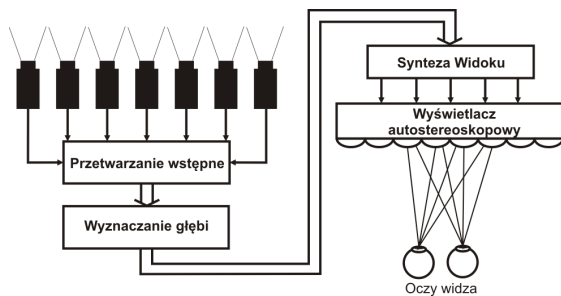
ANALIZA DOKŁADNOŚCI SYNTEZY OBRAZU W SYSTEMACH TELEWIZJI WIELOWIDOKOWEJ

Streszczenie: Artykuł dotyczy problematyki związanej z nowo rozwijającą się ideą telewizji wielowidokowej oraz telewizji swobodnego punktu widzenia. Na przykładzie systemu opracowywanego w Katedrze Telekomunikacji Multimedialnej i Mikroelektroniki Politechniki Poznańskiej przedstawiono ogólną koncepcję takich systemów oraz na jej tle przedstawiono zagadnienie dokładności w algorytmach dotyczących syntezy obrazów. W szczególności, omówiono współczesny algorytm syntezy wykorzystujący mapy rozbieżności a także algorytm służący do ich estymacji. Dokonano analizy i wyciągnięto wnioski dotyczące wpływu dokładności tych algorytmów na działanie całego systemu. Badania przeprowadzono z wykorzystaniem standardowych wielowidokowych sekwencji wizyjnych.

1. WSTĘP

Współcześnie prowadzi się badania na stworzeniem nowych systemów telewizyjnych, oferujących nie tylko wyższą jakość dostarczanego obrazu i dźwięku, ale również nową funkcjonalność. Jedną z cech pożądanых dla takich systemów jest możliwość dostarczania treści trójwymiarowych. Przykładowym zastosowaniem dla takiego systemu jest telewizja swobodnego punktu widzenia (*Free-view Television – FTV*), umożliwiająca użytkownikowi obserwację sceny z wybranego przez niego miejsca, oraz wyświetlanie obrazu z zachowaniem wrażenia głębi na monitorach polaryzacyjnych i autostereoskopowych [1].

Oba z tych zastosowań, wymagają opracowania efektywnych, technicznych możliwości dostarczania wielu widoków. Typowo, system składa się z szeregu kamer dokonujących rejestracji sekwencji związanych z kolejnymi widokami. Najczęściej [2] przyjmuje się liniowe ułożenie kamer z zachowaniem stałych odstępów pomiędzy obiektywami (Rys. 2).



Rys. 1. Diagram systemu telewizji stereowizyjnej

Ze względu na trudności lub wręcz niemożliwość uzyskania identycznych parametrów w wszystkich kamerach (balansu bieli, ostrości, czasu ekspozycji), jak również problemy związane z właściwą kalibracją pozycji kamer, konieczne jest dokonanie przetwarzania wstępnego przed

jakąkolwiek dalszą obróbką zebranego materiału. Zazwyczaj dokonywany jest wtedy proces rektyfikacji oraz wyrównania poziomu oświetlenia i wprowadzana jest korekta balansu bieli [3].

Ponieważ przetwarzanie wszystkich możliwych widoków byłoby nieefektywne, usiłuje się przesyłać jedynie wybrane widoki, pozostałe syntezując za pomocą informacji dodatkowej. Najczęściej wykorzystywane są tzw. mapy głębi, reprezentujące odległości pomiędzy poszczególnymi punktami sceny a wybraną kamerą. Informacje takie wraz z obrazami widoków umożliwiają syntezę nowego widoku sceny z wybranego miejsca w przestrzeni.

Najbardziej powszechną metodą pozyskiwania map głębi jest wyznaczanie mapy tzw. rozbieżności, która jest odwrotnie proporcjonalna do głębi [4]. Mapę rozbieżności estymuje się badając przesunięcia obiektów pomiędzy sąsiadującymi widokami. Dla poprawy jakości mapy rozbieżności, współczesne techniki [5] posiłkują się algorytmami optymalizacyjnymi takimi jak algorytm propagacji wierzeń [6] (*BP-Belief propagation*) lub algorytm cięcia grafów (*GC-Graph Cuts*) [7].

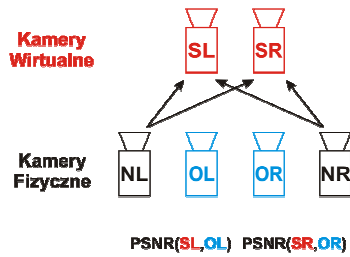
Ostatnim krokiem przetwarzania jest synteza nowych widoków dokonywana w odbiorniku. Na podstawie wybranych wcześniej widoków i odpowiadających im map rozbieżności wyznacza się obrazy odpowiadające nowym widokom wirtualnym. Obrazy te trafiają do monitora stereowizyjnego (najczęściej autostereoskopowego) umożliwiającego wyświetlanie widoku żądanego przez widza.

Jakość syntezowanego wirtualnego widoku jest związana z dokładnością mapy rozbieżności i zależy zarówno od algorytmu jej wyznaczania jak i algorytmu syntezy widoku. Współcześnie, istnieje niedomiar badań dotyczących analizy wpływu każdego z tych etapów na jakość syntezy.

Niniejsza praca poświęcona jest analizie wpływu dokładności w algorytmie estymacji map rozbieżności oraz w algorytmie syntezy widoków, celem określenia najbardziej opłacalnych kierunków dalszych badań.



Rys. 2. Fotografia systemu akwizycji tworzonego w KTMiM Politechniki Poznańskiej.



Rys. 5. Konfiguracja kamer dla pomiaru jakości.

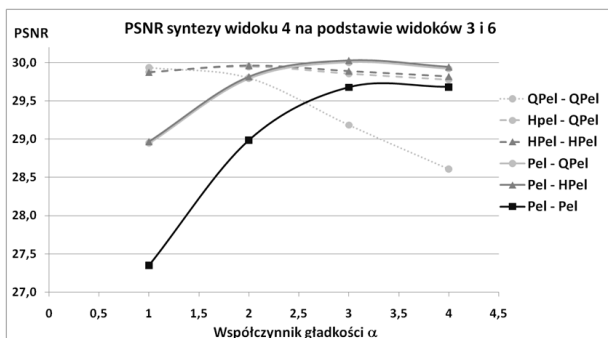
3. WYNIKI

Eksperymenty przeprowadzono z wykorzystaniem wielowidokowych sekwencji testowych wybranych przez grupę ekspertów w zakresie telewizji wielowidokowej MPEG-FTV. Przedstawiane wyniki dotyczą sekwencji ‘Alt Moabit’, ‘Book Arrival’ (HHI [11]), ‘Newspaper’ (GIST [12]) oraz ‘Lovebird 1’ (ETRI [13]).

Dla sprawdzenia wpływu dokładności podpunktowej na jakość działania systemu przebadano różne tryby pracy punktowej i podpunktowej: pełnopunktowej (Pel), pół-punktowej (HPel), oraz ćwierć punktowej (QPel), niezależnie dla estymacji map rozbieżności i syntezy widoków. Badania przeprowadzono w zakresie współczynnika gładkości $\alpha \in \{1, 4\}$.

Tab. 1. *Uśredniony poziom PSNR syntezowanych widoków (lewego i prawego) dla sekwencji ‘Newspaper’ z różnymi nastawami dokładności.*

Jakość syntezy [PSNR]		Dokładność estymacji rozbieżności		
		Pel	HPel	QPel
Dokładność syntezy	Pel	29,6761	-	-
	HPel	30,0276	29,8861	-
	QPel	30,0023	29,8539	28,2948



Rys. 6. Wykresy PSNR zsyntezowanego widoku wirtualnego dla sekwencji ‘Newspaper’ dla różnych nastaw dokładności (dokładność estymacji rozbieżności – dokładność syntezy).

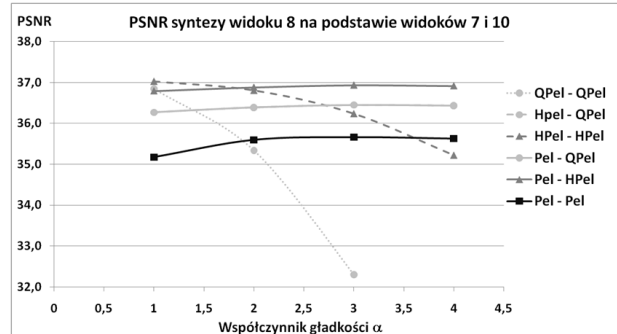
W przypadku sekwencji ‘Newspaper’, dokładność estymacji rozbieżności ma znaczenie jedynie dla niewielkich wartości współczynnika gładkości (Rys. 6, Tabela 1). Jednocześnie, można zauważyć, że wartości te zapewniają relatywnie niższy PSNR w związku z czym nie są zalecane dla tej sekwencji. Dla wartości optymalnych ($\alpha > 3$) jedynie podwyższanie dokładności syntezy przynosi wymierny zysk.

Najlepsze wyniki dla sekwencji ‘Book arrival’ otrzymuje się wykorzystując tryb Pel-HPel (dokładność punktowa estymacji i pół-punktowa syntezy) (Rys. 7, Tabela 2). Wykorzystywanie większej dokładności pod-

punktowej w estymacji rozbieżności nie jest umotywowane.

Tab. 2. *Uśredniony poziom PSNR syntezowanych widoków (lewego i prawego) dla sekwencji ‘Book arrival’ z różnymi nastawami dokładności.*

Jakość syntezy [PSNR]		Dokładność estymacji rozbieżności		
		Pel	HPel	QPel
Dokładność syntezy	Pel	35,6590	-	-
	HPel	36,9301	36,2315	-
	QPel	36,4471	36,1465	32,2936

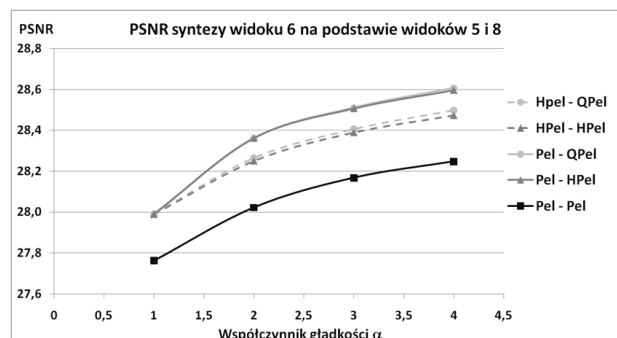


Rys. 7. Wykresy PSNR zsyntezowanego widoku wirtualnego dla sekwencji ‘Book Arrival’ dla różnych nastaw dokładności (dokładność estymacji rozbieżności – dokładność syntezy).

Wyniki dla sekwencji ‘Lovebird 1’ (Rys. 8, Tab. 3) są zbliżone do przypadku ‘Newspaper’. Dla optymalnej wartości współczynnika $\alpha > 3$, wykorzystanie podpunktowej dokładności estymacji rozbieżności przynosi spadek PSNR, spowodowany najprawdopodobniej rozmyciem obrazu i pogorszeniem estymacji rozbieżności na rozmytych krawędziach.

Tab. 3. *Uśredniony poziom PSNR syntezowanych widoków (lewego i prawego) dla sekwencji ‘Lovebird 1’ z różnymi nastawami dokładności.*

Jakość syntezy [PSNR]		Dokładność estymacji rozbieżności		
		Pel	HPel	QPel
Dokładność syntezy	Pel	28,1675	-	-
	HPel	28,5054	28,3882	-
	QPel	28,5104	28,4053	28,2105



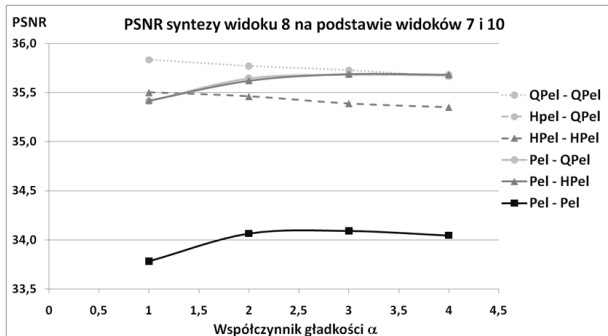
Rys. 8. Wykresy PSNR zsyntezowanego widoku wirtualnego dla sekwencji ‘Lovebird 1’ dla różnych nastaw dokładności (dokładność estymacji rozbieżności – dokładność syntezy).

Sekwencja ‘Alt Moabit’ wykazuje się niewrażliwością na zastosowany współczynnik gładkości α (Rys. 9, Tabela 4). Niezależnie od jego wartości, także w przypadku tej sekwencji, nie odnotowano korzystnego wpływu wykorzystania dokładności podpunktowej w algorytmie estymacji rozbieżności.

Obserwowane zyski pochodzą z zastosowanej dokładności podpunktowej w algorytmie syntezy.

Tab. 4. Uśredniony poziom PSNR syntezowanych widoków (lewego i prawego) dla sekwencji 'Alt Moabit' z różnymi nastawami dokładności.

Jakość syntezy [PSNR]		Dokładność estymacji rozbieżności		
		Pel	HPel	QPel
Dokładność syntezy	Pel	34,0911	-	-
	HPel	35,6875	35,3875	-
	QPel	35,6817	35,3665	35,7263



Rys. 9. Wykresy PSNR zsyntezowanego widoku wirtualnego dla sekwencji 'Alt Moabit' dla różnych nastaw dokładności (dokładność estymacji rozbieżności – dokładność syntezy).

4. PODSUMOWANIE

Przeprowadzone badania wpływu dokładności estymacji rozbieżności i procesu syntezy pokazują, że jakość syntezowanych obrazów zależy głównie od dokładności, z jaką dokonywana jest synteza obrazu.

Dla wykorzystanych metod, odzwierciedlających obecny postęp badań w tej dziedzinie, w żadnej przebadanej sekwencji, estymacja rozbieżności z wyższą dokładnością podpunktową nie przynosi wyraźnych zysków. Zachowanie to jest niezgodne z intuicyjnym pojmowaniem dokładności podpunktowej.

W rozpatrywanym przypadku, podpunktowa estymacja rozbieżności wiąże się z początkową interpolacją obrazu w celu zwiększenia jego rozdzielczości, która powoduje powstawanie rozmytych krawędzi obiektów. Wyznaczenie kosztu dopasowania elementów sceny w otoczeniu takich krawędzi jest kłopotliwe, co wpływa niekorzystnie na estymację rozbieżności także otaczających elementów.

Mimo, iż w dokładność podpunktową w syntezie obrazu również wiąże się z interpolacją oraz zwiększeniem rozdzielczości przetworzonego obrazu, to w przeciwieństwie do estymacji rozbieżności, przynosi wymierny zysk. Dzieje się tak, gdyż natura procesu syntezy, polega na przesuwaniu fragmentów obrazu referencyjnego z wybraną dokładnością. Pozwala to na syntezę ostrych krawędzi z wyższą dokładnością położenia, zaś nie wprowadza dodatkowego rozmycia. Końcowo, rozmycie wprowadzone przez interpolację jest odwracane przez decymację do rozmiaru oryginalnego widoku.

Wyniki te prowadzą do wniosku, że aktualnie stosowane techniki uzyskiwania dokładności podpunktowej w estymacji rozbieżności są niewystarczające, a co za tym idzie - podpunktowa

estymacja rozbieżności stanowi interesujący obszar dla dalszych badań. Można przewidywać, że zysk w tej dziedzinie można uzyskać poprzez zastosowanie bardziej zaawansowanych algorytmów interpolacji, adaptacyjnie dopasowujących się do treści obrazu i zachowujących krawędzie.

Praca finansowana ze środków na naukę w latach 2007-2009 jako projekt „badawczy rozwojowy”.

SPIS LITERATURY

- [1] A. Smolic, K. Müller, P. Kauff, T. Wiegand, "Considerations about the Vision of a 3D Video Standard", ISO/IEC JTC1/SC29/WG11, MPEG 2008/M16130, Lozanna, Szwajcaria, I.2009.
- [2] "Call for Contributions for FTV Test Material", MPEG 2007/N9468, Shenzhen, Chiny, X.2007.
- [3] I. Lee, Y. Lung, B. Kim, S. KO. "An Advanced Video "Camera System with Robust AF, AE, and AWB Control", IEEE Trans on Consumer Electronics. roz. 47, nr. 3. str. 694-699, IV.2001.
- [4] "Description of Exploration Experiments in 3D Video" wzór (1), MPEG 2008/N9596, Antalya, Turcja, I.2008
- [5] D. Scharstein, R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", International Journal of Computer Vision 2002.
- [6] J.S. Yedidia, W.T. Freeman, Y. Weiss, "Understanding Belief Propagation and Its Generalizations", Exploring Artificial Intelligence in the New Millennium, Roz. 8, str. 239-236, I.2003 (Science & Technology Books).
- [7] V. Kolmogorov, R. Zabih, "Multi-camera scene reconstruction via graph cuts", European Conference on Computer Vision 2002.
- [8] J. Sun, H. Y. Shum, N. N. Zheng, "Stereo matching using belief propagation", European Conference on Computer Vision 2002.
- [9] O. Stankiewicz, K. Wegner, "Depth Map Estimation Software version 3", MPEG 2008/M15540, Hannover, Niemcy, VI.2008.
- [10] M. Gotfryd, K. Wegner, M. Domański, "View synthesis software and assessment of its performance", ISO/IEC JTC1/SC29/WG11, MPEG 2008/M15672, Hannover, Niemcy, VI.2008.
- [11] I. Feldmann, M. Müller, F. Zilly, R. Tanger, A. Smolic, P. Kauff, T. Wiegand, „HHI Test Material for 3D Video", ISO/IEC JTC1/SC29/WG11, MPEG 2008/M15413, Archamps, Francja, IV.2008.
- [12] 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, IV.2008.
- [13] M. Tanimoto, T. Fujii, N. Fukushima, "1D Parallel Test Sequences for MPEG-FTV", MPEG 2008/M15378, Archamps, Francja, IV.2008.