

ZWIĘKSZANIE ROZDZIELCZOŚCI OBRAZU I MAPY GŁĘBI W CELU POPRAWY JAKOŚCI SYNTEZY WIDOKÓW WIRTUALNYCH

RESOLUTION INCREASING FOR VIRTUAL VIEW SYNTHESIS ENHANCEMENT

Streszczenie: W artykule opisano metodę poprawiającą jakość syntezowanych widoków wirtualnych w systemach telewizji swobodnego punktu widzenia. W zaproponowanej metodzie rozdzielczość zarówno widoków zarejestrowanych przez kamery systemu wielowidokowego, jak i odpowiadających im map głębi jest dwukrotnie zwiększana w celu zwiększenia dokładności rzutowania punktów do widoku wirtualnego. Metoda pozwala na zwiększenie obiektywnej i subiektywnej jakości syntezowanych widoków niezależnie od zastosowanego algorytmu syntezy.

Abstract: In the paper we described a method what increases the quality of synthesized views in free-viewpoint television systems. In the proposed method the resolution of views captured by cameras of the multiview system with corresponding depth maps is increased two times in order to increase the accuracy of projecting points to the virtual view. The method allows to increase objective and subjective quality of synthesized views, irrespectively of the synthesis algorithm.

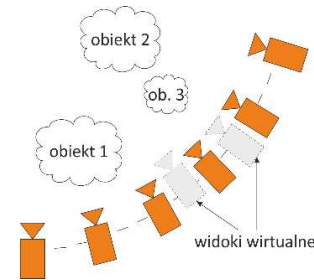
Słowa kluczowe: swobodna nawigacja, synteza widoków wirtualnych, telewizja swobodnego punktu widzenia.

Keywords: free navigation, free-viewpoint television, virtual view synthesis.

1. WSTĘP

Telewizja swobodnego punktu widzenia (ang. *Free-Viewpoint Television*, FTV) umożliwia widzowi wirtualne przemieszczanie się w zarejestrowanej przy pomocy systemu wielokamerowego scenie. Istniejące systemy wielokamerowe liczyć mogą do kilkuset kamer [14], jednakże wyklucza to ich użyteczność w kontekście rejestracji jakichkolwiek nieeksperymentalnych materiałów, jak choćby transmisji wydarzeń sportowych. Praktyczny system charakteryzuje się zdecydowanie mniejszymi rozmiarami, licząc od kilku do kilkunastu kamer [1,5,8].

Ażeby zapewnić widzowi płynną nawigację, nie można swobody jego ruchu ograniczyć wyłącznie do pozycji rzeczywistych kamer wchodzących w skład systemu wielokamerowego – należy umożliwić widzowi spoglądanie na scenę również z pozycji pośrednich. Implikuje to konieczność syntezy widoków wirtualnych, czyli stworzenia obrazów rejestrowanych przez wirtualne kamery, znajdujące się pomiędzy rzeczywistymi urządzeniami.



Rys. 1. Idea systemu telewizji swobodnego punktu widzenia

Idea systemu telewizji swobodnego punktu widzenia przedstawiona została na rys. 1, gdzie kolorem pomarańczowym oznaczono rzeczywiste kamery tworzące system wielokamerowy, a kolorem szarym przykładowe widoki wirtualne.

2. SYNTEZA WIDOKÓW WIRTUALNYCH

Najbardziej rozpowszechniona metoda syntezy widoków wirtualnych bazuje na reprezentacji „głębka – obraz” (ang. *Depth-Image-Based Rendering*, DIBR) [13]. Rozwiązanie takie jest szeroko stosowane chociażby w kompresji obrazu wielowidokowego [2], czy też w uznanym, powszechnie stosowanym oprogramowaniu odniesienia, stworzonym przez grupę normalizacyjną MPEG – VSRS (*View Synthesis Reference Software*) [12].

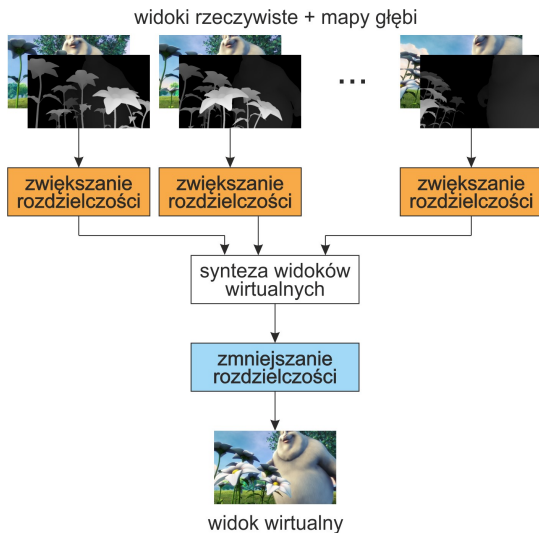
Metoda odniesienia stworzona została na potrzeby prostych systemów wielokamerowych, w których to kamery ustawione są równolegle, a przetwarzanie obrazu jest zdecydowanie prostsze. Pomimo zaadaptowania VSRS do dowolnego (np. łukowego) ustawienia kamer, istnieją lepsze metody syntezy, jak choćby zaproponowana przez zespół z Politechniki Poznańskiej MVS (*Multiview Synthesis*) [4].

3. PROPONOWANA METODA

W artykule proponujemy inne podejście, zwiększające jakość syntezowanych widoków bez ingerencji w sam algorytm syntezy. Co istotne, zaproponowana metoda pozwala na zwiększenie jakości niezależnie od tego, jaki algorytm syntezy zostanie zastosowany, a więc może

zostać z powodzeniem użyta przez badaczy rozwijających własne metody syntezy.

Zaproponowano dodanie przetwarzania wstępnego, w czasie którego rozdzielczość widoków rzeczywistych i odpowiadających im map głębi jest podwajana (zarówno w pionie, jak i w poziomie). Następnie, na zwiększonych obrazach wejściowych przeprowadzana jest synteza widoku wirtualnego (również o zwiększonej rozdzielczości). W ostatnim kroku zsyntezowany widok wirtualny jest zmniejszany do oryginalnych rozmiarów.



Rys. 2. Algorytm syntezy widoków wirtualnych ze zwiększeniem rozdzielczości widoków i map głębi

Algorytm syntezy wraz z zaproponowanymi zmianami został przedstawiony na rys. 2. Dodane etapy przetwarzania wstępnego i końcowego oznaczono odpowiednio kolorem pomarańczowym i niebieskim. Dla przykładu, w przypadku widoków wejściowych brakujące próbki są interpolowane w sposób liniowy. Taki sam sposób interpolacji głębi byłby niewłaściwy, generowałby bowiem nieistniejące wartości głębi na krawędziach obiektów, co wiązałoby się z powstaniem irytujących widza artefaktów w syntezywanym obrazie. Z tej przyczyny zdecydowano się zwiększać rozdzielczość głębi w inny sposób – kopiując wartość głębi z sąsiedniego punktu (z góry / z lewej).

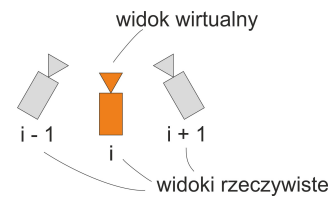
4. OPIS EKSPERYMENTU

W celu pokazania, iż zaproponowana metoda pozwala zwiększyć jakość syntezywanego widoku wirtualnego, zdecydowano się przeprowadzić eksperyment dla dwóch algorytmów syntezy: VSRS [12] i MVS [4].

Zestaw sekwencji testowych zawierał 8 sekwencji wielowidokowych. 2 z nich są sekwencjami syntetycznymi, wygenerowanymi komputerowo, z dostępną perfekcyjną mapą głębi – Big Buck Bunny Flowers oraz Big

Buck Bunny Butterfly (obie z lukowo rozmieszczonymi kamerami). 6 pozostałych sekwencji testowych zarejestrowano przy użyciu rzeczywistych systemów wielokamerowych: Ballet, Breakdancers [15]; Poznań Blocks2, Poznań Fencing2 [3]; Soccer Arc [6]; oraz jedyna w zbiorze sekwencja z liniowym ustawieniem kamer – Soccer Linear [7].

W ramach eksperymentu przebadano jakość obiektywną i subiektywną. Do zmierzenia jakości obiektywnej zastosowano miarę PSNR. Sposób estymacji PSNR przedstawiono na rys. 3. Dla każdej sekwencji testowej na podstawie rzeczywistych widoków $i - 1$ oraz $i + 1$ syntezywanego widoku wirtualnego w pozycji widoku rzeczywistego i , który stanowił widok odniesienia dla wyliczenia PSNR.



Rys. 3. Schemat wyznaczania PSNR dla widoku i

W celu wyznaczenia jakości subiektywnej syntezywanego widoku wirtualnego zdecydowano się użyć metody oceny podobieństwa (ang. *similarity judgement*) [10,11]. Metoda ta umożliwia bezpośrednie porównanie jakości dwóch obrazów i stwierdzenie, który z nich ma lepszą jakość oraz jak bardzo jakości te się różnią. Dla każdej sekwencji testowej porównywano widok zsyntezowany bez zwiększania rozdzielczości z widokiem wyliczonym przy zastosowaniu zaproponowanej metody. Kolejność wyświetlania obu obrazów była losowa.

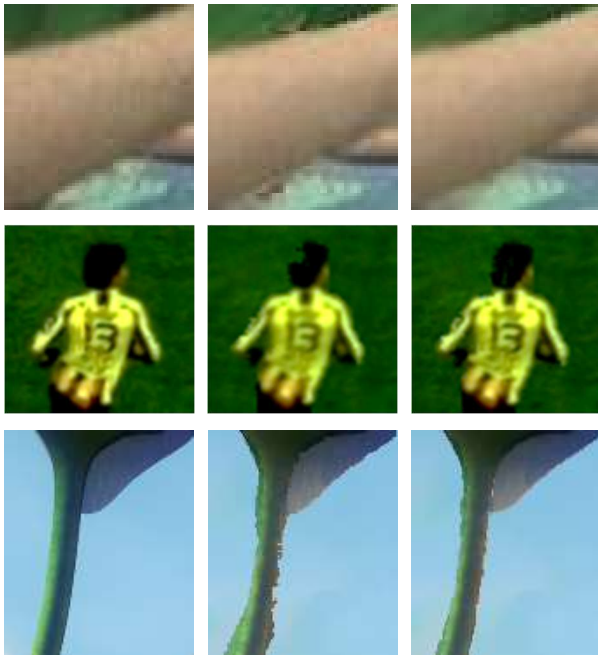
W testach subiektywnych uczestniczyło 10 ekspertów. Każdy z nich oceniał, który z wyświetlonych widoków ma lepszą jakość, a także – w skali od 0 do 5 – jak bardzo różni się ich jakość, gdzie 0 oznaczało brak widocznych różnic a 5 – znacząco wyższą jakość jednego z obrazów.

5. WYNIKI EKSPERYMENTALNE

Na rys. 4 przedstawiono fragmenty widoków wirtualnych zsyntezowanych bez zwiększania rozdzielczości (środkowa kolumna) oraz ze zwiększaniem rozdzielczości (kolumna prawa). Dla porównania, w lewej kolumnie przedstawiono fragmenty widoków odniesienia.

Kolejno, przedstawiono fragmenty z sekwencji Poznań Blocks2, Soccer Linear oraz Big Buck Bunny Flowers Arc.

Jak pokazano, zastosowanie zaproponowanej metody zwiększania rozdzielczości widoków wejściowych wraz z odpowiadającymi im mapami głębi pozwala na zmniejszenie artefaktów w syntezywanym obrazie (ucięta głowa piłkarza, zwiększając jednocześnie jakość odwzorowania krawędzi obiektów (lepiej odwzorowana łądyga)).



Rys. 4. Porównanie fragmentów widoków odniesienia (lewa kolumna) z fragmentami widoków zszyntezowanych bez zwiększania rozdzielczości (środkowa kolumna) i ze zwiększaniem rozdzielczości (kolumna prawa)

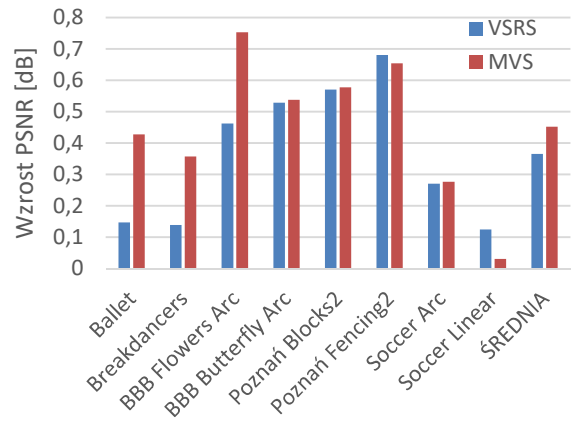
W tabeli 1 przedstawiono wartości PSNR dla wszystkich sekwencji testowych, wyliczone dla obu metod syntezy, bez zwiększania i ze zwiększaniem rozdzielczości widoków i map głębi.

Tab. 1. Jakość obiektywna widoków wirtualnych zszyntezowanych za pomocą algorytmu VSRS i MVS, bez zwiększania rozdzielczości (kolumna A) i ze zwiększaniem rozdzielczości (kolumna B)

Sekwencja	PSNR [dB]			
	VSRS		MVS	
	A	B	A	B
Ballet	31,05	31,14	31,74	32,14
Breakdancers	31,47	31,69	31,73	32,26
BBB Flowers	22,68	23,13	25,81	26,49
BBB Butterfly	32,04	32,71	33,56	34,27
Poznań Blocks2	29,25	29,92	29,61	30,28
Poznań Fencing2	28,33	28,97	28,82	29,58
Soccer Arc	22,79	23,06	23,48	23,76
Soccer Linear	35,00	35,12	34,98	35,01
Średnia	29,08	29,45	29,97	30,42

We wszystkich przetestowanych przypadkach zastosowanie zwiększania rozdzielczości pozwoliło zwiększyć obiektywną jakość syntezy widoków. Wzrost PSNR wynikający z zastosowania proponowanej metody przedstawiony został na rys. 5.

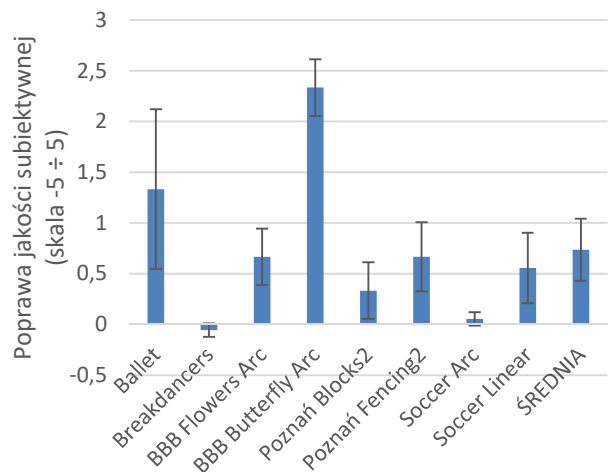
Uśredniając wyniki dla wszystkich sekwencji testowych, zastosowanie zwiększania rozdzielczości pozwoliło na zwiększenie jakości o 0,41 dB (odpowiednio o 0,37 dB dla oprogramowania odniesienia i 0,45 dla algorytmu MVS).



Rys. 5. Wpływ zwiększania rozdzielczości na jakość obiektywną

Testy jakości subiektywnej przeprowadzono wyłącznie dla jednego algorytmu syntezy – MVS. Na rys. 6. przedstawiono poprawę jakości subiektywnej wynikającą z zastosowania proponowanej metody. Dla siedmiu spośród ośmiu sekwencji testowych zwiększenie rozdzielczości pozwala zwiększyć subiektywną jakość syntezy widoków, przy czym dla 6 z nich poprawa jakości jest statystycznie istotna (przy założeniu stopnia istotności 5%).

W przypadku sekwencji Breakdancers oraz Soccer Arc nie można bez zwiększania grupy testowej stwierdzić, czy zwiększenie rozdzielczości wpływa na subiektywną jakość syntezy. W przypadku sekwencji Soccer Arc głównym problemem jest bardzo niska jakość map głębi – niezależnie od zastosowanego algorytmu syntezy, generowane widoki wirtualne charakteryzują się dużą ilością artefaktów, przez co ewentualne różnice pomiędzy dwoma zszyntezowanymi widokami pozostają niezauważone. Sekwencja Breakdancers jest natomiast sekwencją niskiej jakości – widoki rzeczywiste są ciemne, nie ma w nich ostrych, wyraźnych krawędzi – co przekłada się na brak poprawy jakości subiektywnej.



Rys. 6. Wpływ zwiększania rozdzielczości na jakość subiektywną

Średnio, zastosowanie metody zwiększania rozdzielczości pozwala na uzyskanie jakości lepszej o 0,75 w skali od -5 do 5.

6. PODSUMOWANIE

W artykule przedstawiono efektywną metodę zwiększenia jakości syntezy widoków wirtualnych. Zaproponowana metoda polega na dodaniu etapu przetwarzania wstępnego przed syntezą widoków wirtualnych, w którym to rozdzielczość widoków wejściowych i odpowiadających im map głębi jest podwajana.

Zaproponowana metoda pozwala na zwiększenie jakości syntezy dla dowolnego algorytmu syntezy bazującego na reprezentacji „głębia – obraz” (DIBR).

W ramach dalszych prac przebadane zostaną bardziej zaawansowane metody zwiększania rozdzielczości, niż przedstawione w artykule. Niemniej jednak, nawet zastosowanie prostych metod (liniowa interpolacja w przypadku obrazu i interpolacja zerowego rzędu dla mapy głębi) pozwala na zwiększenie PSNR syntezowanych widoków o 0,4 dB, a jakości subiektywnej o 0,75 w skali od -5 do 5.

PODZIĘKOWANIA

Praca sfinansowana ze środków Narodowego Centrum Badań i Rozwoju w ramach Umowy nr TANGO1/266710/NCBR/2015.

LITERATURA

- [1] Domański Marek, Dziembowski Adrian, Kuehn Agnieszka, Mieloch Dawid. 2014. „Telewizja swobodnego punktu widzenia – nowa usługa czy futurystyczna wizja?”. *Przegląd Telekomunikacyjny* 8-9/2014: 734-737.
- [2] Domański Marek, Stankiewicz Olgierd, Wegner Krzysztof, Kurc Maciej, Konieczny Jacek, Siast Jakub, Stankowski Jakub, Ratajczak Robert, Grajek Tomasz. 2013. „High Efficiency 3D Video Coding Using New Tools Based on View Synthesis”. *IEEE Trans. Image Proc.*, 22/2013: 3517-3527.
- [3] Domański Marek, Dziembowski Adrian, Grzelka Adam, Mieloch Dawid, Stankiewicz Olgierd, Wegner Krzysztof. 2016. „Multiview test video sequences for free navigation exploration obtained using pairs of cameras”. *ISO/IEC JTC1/SC29/WG11, Doc. MPEG M38247*.
- [4] Dziembowski Adrian, Grzelka Adam, Mieloch Dawid. 2016. „Wielowidokowa synteza w systemach telewizji swobodnego punktu widzenia”. *Przegląd Telekomunikacyjny* 6/2016: 233-236.
- [5] Dziembowski Adrian, Kuehn Agnieszka, Łuczak Adam, Mieloch Dawid, Wegner Krzysztof. 2014. „Realizacja eksperymentalnego systemu telewizji swobodnego punktu widzenia z łukowym ustawieniem kamer”. *Przegląd Telekomunikacyjny* 6/2014: 161-164.
- [6] Goorts P., Maesen S., Dumont M., Rogmans S., Bekaert P. 2014. „Free viewpoint video for soccer using histogram-based validity maps in plane sweeping”. *Int. Conf. Computer Vision Theory Appl.*
- [7] Goorts P., Dumont M., Rogmans S., Bekaert P. 2012. „An end-to-end system for free viewpoint video for smooth camera transitions”, *International Conf. 3D Imaging*.
- [8] Goorts P., L. Jorissen, G. Lafruit. 2014. “[FTV AHG] EE2 results on Soccer-Linear1 with foreground/background segmentation”. *ISO/IEC JTC1/SC29/WG11, Doc. M34310*.
- [9] Kovacs Peter. 2015. “[FTV AHG] Big Buck Bunny light-field test sequences”, *ISO/IEC JTC1/SC29/WG11, Doc. MPEG M35721*.
- [10] Mantiuk R., Tomaszewska A., Mantiuk R. 1981. „Comparison of four subjective methods for image quality assessment”. *Computer Graphics Forum* 0: 1 – 13.
- [11] de Ridder H., Majoor G. 1990. „Numerical category scaling: an efficient method for assessing digital image coding impairments”. *Proc. of SPIE* 1249: 65.
- [12] Stankiewicz Olgierd, Wegner Krzysztof, Tanimoto Masayuki, Domański Marek. 2013. „Enhanced view synthesis reference software (VSRS) for Free-viewpoint Television”, *ISO/IEC JTC 1/SC 29/WG 11, MPEG M31520*.
- [13] Sun Z., Jung Ch. 2015. „Real-Time Depth-Image-Based Rendering on GPU”. *International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery*: 324 – 328.
- [14] Tanimoto M. 2010. „FTV (Free-Viewpoint TV)”. *Proc. of 2010 IEEE 17th Int. Conf. on Image Proc.*
- [15] Zitnick L., Kang S.B., Uyttendaele M., Winder S., Szeliski R. 2004. „High-quality video view interpolation using a layered representation”, *ACM SIGGRAPH*.