

SYNTEZA WIDOKÓW WIRTUALNYCH W RZADKICH SYSTEMACH WIELOKAMEROWYCH DLA ZASTOSOWAŃ W SWOBODNEJ NAWIGACJI

VIRTUAL VIEW SYNTHESIS IN SPARSE MULTICAMERA SYSEMS FOR FREE NAVIGATION

Streszczenie: W referacie opisano autorską metodę syntezy widoków wirtualnych dostosowanych do sekwencji zarejestrowanych przy użyciu systemów wielokamerowych z rzadkim rozmieszczeniem kamer. Opracowana metoda syntezy umożliwia osiągnięcie dobrej jakości syntezowanych widoków wirtualnych przy jednoczesnym krótkim czasie przetwarzania, przez co znajduje ona zastosowanie w praktycznych systemach swobodnej nawigacji.

Abstract: In the paper a new virtual view synthesis method is described. Presented method is adapted for multiview sequences captured by multicamera systems with sparse camera arrangement and allows to achieve good quality virtual views with relatively short processing time. Therefore it can be applied for practical free navigation systems.

Słowa kluczowe: synteza widoków wirtualnych, system swobodnej nawigacji, system wielokamerowy.

Keywords: free-navigation system, multicamera system, virtual view synthesis.

1. WPROWADZENIE

Referat dotyczy syntezy widoków wirtualnych dla celów systemów swobodnej nawigacji, a więc systemów, w których użytkownik ma możliwość wirtualnej nawigacji w zarejestrowanej scenie [3, 20, 21]. W takim systemie widz sam może ustalać swój punkt widzenia i kierunek patrzenia. Swobodne przemieszczanie się w scenie wymaga jednak płynnego wirtualnego ruchu widza, a więc pula punktów widzenia, jaką on dysponuje nie może być ograniczona wyłącznie do widoków zarejestrowanych przez kamery systemu wielokamerowego. Widz musi mieć możliwość oglądania sceny również z innych, dowolnie wybranych pozycji (rys. 1: kolorem czarnym oznaczono rzeczywiste kamery systemu wielokamerowego, kolorem pomarańczowym – kamerę wirtualną, z której użytkownik systemu ogląda obraz i której pozycję może on w dowolny sposób zmieniać). W celu umożliwienia widzowi płynnego wirtualnego przemieszczania się w zarejestrowanej scenie należy stworzyć (zsyntezować) dodatkowe widoki (widoki wirtualne). Widoki te są tworzone na podstawie widoków rzeczywistych, czyli widoków zarejestrowanych przez rzeczywiste kamery systemu wielokamerowego. Co więcej, oczekuje się, że różnica jakości pomiędzy widokami rzeczywistymi i wirtualnymi będzie jak najmniejsza, by widz nie był w stanie stwierdzić, czy ogląda obrazy zarejestrowane przez kamery, czy też obrazy syntetyczne.



Rys. 1. Ilustracja idei swobodnej nawigacji

Synteza widoków wirtualnych w rzadkich systemach wielokamerowych jest zdecydowanie trudniejsza niż w przypadku systemów gęstych. Zwiększona trudność spowodowana jest znacznie większą odległością pomiędzy poszczególnymi kamerami lub większym kątem pomiędzy ich osiami optycznymi. Skutkuje to trzema podstawowymi problemami charakterystycznymi dla rzadkich systemów wielokamerowych [7]:

1. Mniejszy obszar sceny jest widoczny jednocześnie w różnych widokach rzeczywistych, co skutkuje powstawianiem odsłonień w widoku wirtualnym.
2. Ograniczona rozdzielczość obrazu skutkująca powstawianiem pęknięć (ang. *cracks*) – nieciągłości obiektów w widoku wirtualnym.
3. Nielambertowskie odbicia światła skutkujące różną barwą tych samych obiektów w różnych widokach rzeczywistych.

Rozwiązanie trzech wymienionych problemów jest kluczowe dla opracowania metody syntezy widoków wirtualnych dla zastosowań w swobodnej nawigacji w rzadkich systemach wielokamerowych.

2. PROPONOWANA METODA SYNTEZY WIDOKÓW WIRTUALNYCH

Zaproponowana przez autora metoda syntezy widoków wirtualnych składa się z 4 etapów [15]:

1. Przetwarzanie pojedynczych widoków rzeczywistych,
2. Łączenie informacji przerzutowanej z dwóch najbliższych widoków rzeczywistych,
3. Wypełnianie odsłonień przy użyciu informacji z dalszych widoków rzeczywistych,
4. Uzupełnianie obszarów niewidocznych w żadnym widoku rzeczywistym.

W pierwszym etapie zaproponowanej metody syntezy widoków wirtualnych wszystkie N widoków rzeczywistych przetwarzane jest niezależnie, czego wynikiem jest N pośrednich widoków wirtualnych, zawierających punkty przerzutowane z pojedynczych widoków rzeczywistych.

Po stworzeniu pośrednich widoków wirtualnych dla wszystkich widoków rzeczywistych, dokonywane jest łączenie pośrednich widoków wyznaczonych dla najbliższego lewego i najbliższego prawego widoku (w celu wyboru najbliższych widoków rzeczywistych opracowano uniwersalny sposób wyboru sąsiednich kamer rzeczywistych działający poprawnie dla dowolnego rozmieszczenia kamer). Aby uniknąć skokowej zmiany barwy podczas wirtualnej nawigacji, mieszanie barwy punktu przerzutowanego z lewego i prawego widoku realizowane jest poprzez średnią ważoną odległością pomiędzy kamerą wirtualną, a obiema najbliższymi kamerami rzeczywistymi.

Widok wirtualny stworzony na podstawie dwóch najbliższych widoków rzeczywistych zawiera obszary odsłonięte, czyli obszary, które nie były widoczne ani w najbliższym lewym, ani w najbliższym prawym widoku rzeczywistym. W metodach opisanych w literaturze odsłonięcia są najczęściej uzupełniane na podstawie otoczenia, co skutkuje błędami i pogorszeniem jakości syntezowanych widoków. W zaproponowanej metodzie syntezy widoków wirtualnych obszary odsłonięte są wypełniane przy wykorzystaniu informacji z pozostałych widoków rzeczywistych, co umożliwia znaczącą poprawę jakości syntezowanych widoków. Uzupełnianie treści na podstawie otoczenia przeprowadzane jest jedynie dla obszarów niewidocznych w żadnym widoku rzeczywistym.

Wypełnianie odsłonięć przy użyciu dalszych widoków rzeczywistych rozwiązuje pierwszy z trzech problemów wypisanych w punkcie pierwszym. Aby rozwiązać dwa pozostałe problemy, opracowano dodatkowe techniki przetwarzania widoków wirtualnych, a mianowicie technikę rzutowania punktów wymuszającą ciągłość obiektów w widoku wirtualnym oraz dwie techniki korekcji niespójności barwnej widoku wirtualnego. Ponadto, zaproponowano technikę wstępnego zwiększania rozdzielczości obrazów wejściowych oraz techniki eliminacji błędów spowodowanych rozmyciem krawędzi w widokach rzeczywistych i odpowiadających im mapach głębi.

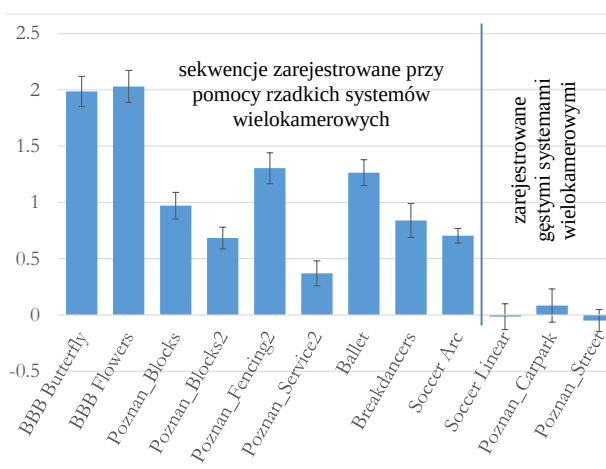
Szczegóły zaproponowanej metody opisano w [5].

3. OCENA EFEKTYWNOŚCI METODY

Opracowaną metodę syntezy widoków wirtualnych (MVS, *MultiView Synthesis*) [11] porównano z metodą odniesienia (zaimplementowaną w oprogramowaniu VSRS, *View Synthesis Reference Software*), rozwijaną przez międzynarodową grupę ekspertów MPEG [19]. Oceniano zarówno jakość mierzoną w sposób obiektywny (przy użyciu miary PSNR, tabela 1), jak i subiektywny (44 uczestników, metoda *Pair Comparison*, rekomendowana przez ITU-T [18], rys. 2). Porównania dokonano na zbiorze 12 zróżnicowanych sekwencji wielowidokowych.

Tab. 1. Porównanie jakości mierzonej obiektywnie

Sekwencja (+ źródło)		PSNR [dB]	
		VSRS	MVS
BBB Butterfly	[17]	31,25	33,58
BBB Flowers	[17]	22,62	25,97
Poznan_Blocks	[2]	24,78	25,21
Poznan_Blocks2	[4]	29,60	31,62
Poznan_Fencing2	[4]	29,62	30,69
Poznan_Service2	[4]	24,68	25,48
Ballet	[22]	29,21	30,10
Breakdancers	[22]	30,62	31,64
Soccer Arc	[16]	20,51	21,21
Soccer Linear	[16]	34,18	34,28
Poznan_Carpark	[1]	33,50	33,89
Poznan_Street	[1]	35,48	35,72
Średnio		28,84	29,95



Rys. 2. Poprawa jakości mierzonej subiektywnie (względem VSRS); 95% przedział ufności; skala: [-3, 3]



Rys. 3. Fragmenty widoków wirtualnych zsintezowanych przy użyciu VSRS (po lewej) i MVS (po prawej); sekwencje: BBB Flowers, Poznan_Fencing2, Soccer Arc

Jak pokazano, dla wszystkich sekwencji testowych osiągnięto poprawę jakości mierzonej przy użyciu PSNR. Ponadto, dla wszystkich sekwencji zarejestrowanych przy pomocy rzadkich systemów wielokamerowych uzyskano statystycznie istotną poprawę jakości.

Na rys. 3 fragmenty widoków wirtualnych stworzonych przy pomocy metody odniesienia porównano z fragmentami widoków zsintezowanych przy użyciu zaproponowanej metody. Znacząca poprawa jakości spowodowana została głównie dzięki wypełnianiu odsłonięć informacją z dalszych widoków rzeczywistych (górny rząd), rzutowaniem punktów uwzględniającym ciągłość obiektów (środkowy rząd) i korekcją niespójności barwnej widoku wirtualnego (dolny rząd).

4. SYNTEZA CZASU RZECZYWISTEGO

Poprawa jakości syntezowanych widoków została okupiona niewiele dłuższym czasem syntezy 1 widoku wirtualnego (średnio 3,5 sekundy w porównaniu do 2,1 sekundy dla VSRS). Jednakże, aby umożliwić realizację praktycznego systemu swobodnej nawigacji, synteza widoków wirtualnych musi działać w czasie rzeczywistym. Aby spełnić ten warunek, dokonano szeregu uproszczeń opracowanej metody umożliwiających znaczącą redukcję czasu syntezy przy jednoczesnym niewielkim pogorszeniu jakości [13].

Tab. 2. Porównanie czasu syntezy widoku wirtualnego

Sekwencja	Czas syntezy 1 ramki [ms]		
	VSRS	rtMVS 1 wątek	rtMVS 4 wątki
BBB Butterfly	1275	42	24
BBB Flowers	1488	47	28
Poznan_Blocks	3764	97	54
Poznan_Blocks2	2738	89	50
Poznan_Fencing2	2219	89	49
Poznan_Service2	2189	91	52
Ballet	989	34	19
Breakdancers	745	34	21
Soccer Arc	2379	90	49
Soccer Linear	1524	56	33
Poznan_Carpark	2151	74	40
Poznan_Street	1883	74	42
Średnio	1945	68	38

Tab. 3. Porównanie jakości mierzonej obiektywnie

Sekwencja	PSNR [dB]	
	VSRS	rtMVS
BBB Butterfly	31,25	32,39
BBB Flowers	22,62	24,76
Poznan_Blocks	24,78	24,91
Poznan_Blocks2	29,60	30,30
Poznan_Fencing2	29,62	29,68
Poznan_Service2	24,68	24,80
Ballet	29,21	29,26
Breakdancers	30,62	31,28
Soccer Arc	20,51	20,69
Soccer Linear	34,18	34,14
Poznan_Carpark	33,50	33,62
Poznan_Street	35,48	35,48
Średnio	28,84	29,28

Jak pokazano w tabeli 2, w przypadku implementacji 4-wątkowej [8], opracowana metoda (rtMVS, *real-time MultiView Synthesis*) umożliwia synteze widoku wirtualnego w czasie rzeczywistym (procesor: Intel i7-4770). Jednocześnie, jakość syntezowanych widoków jest wyraźnie lepsza, niż dla metody odniesienia (tabela 3).

5. PODSUMOWANIE

Najistotniejszym osiągnięciem jest opracowanie dostosowanej do systemów z rzadkim rozmieszczeniem kamer metody syntezy widoków wirtualnych, w której do syntezy widoku wirtualnego wykorzystywana jest informacja pochodząca z wielu widoków rzeczywistych. W opracowanej metodzie obszary nieprzerzutowane z dwóch sąsiednich widoków rzeczywistych wypełniane są z widoków dalszych.

Analizując rezultaty dla dwunastu zróżnicowanych sekwencji testowych wykazano, iż jakość widoków wirtualnych syntezowanych w zaproponowany sposób jest lepsza, niż jakość widoków syntezowanych przy użyciu oprogramowania odniesienia. Średnia poprawa wartości PSNR wyniosła niemal 1,2 dB. Co ważne, dla wszystkich sekwencji zarejestrowanych przy użyciu rzadkich systemów wielokamerowych uzyskano statystycznie istotną poprawę jakości mierzonej w sposób subiektywny.

Ponadto, opracowano kilka technik poprawiających jakość syntezowanych widoków:

1. Filtrację rozmytych krawędzi w widokach rzeczywistych opartą na analizie gradientów [6],
2. Technikę filtracji rozmytych krawędzi w rzeczywistych mapach głębi [6],
3. Wymuszającą ciągłość obiektów technikę rzutowania punktów z widoków rzeczywistych [11],
4. Technikę adaptacyjnej korekcji niespójności barwnej, wyrównującą lokalną charakterystykę obszarów przerzutowanych z różnych widoków rzeczywistych [9],
5. Szybką korekcję niespójności barwnej, wyrównującą globalną charakterystykę barwną obszarów rzutowanych z różnych widoków rzeczywistych [12],
6. Technikę wstępnego zwiększania rozdzielczości widoków rzeczywistych i map głębi [10, 14].

Niewątpliwym osiągnięciem jest również opracowanie algorytmu syntezy widoków wirtualnych [8, 13] działającego w czasie rzeczywistym bez konieczności zastosowania kart graficznych czy programowalnych układów FPGA. Co istotne, zaproponowany algorytm umożliwia synteze widoków mających średnio o 0,5 dB większą wartość PSNR niż widoki syntezowane przy użyciu oprogramowania odniesienia VSRS przy jednoczesnym 51-krotnym skróceniu czasu obliczeń.

Podsumowując, wykazano, iż jest możliwe stworzenie efektywnej metody syntezy widoków wirtualnych dla rzadkich systemów wielokamerowych, a tym samym możliwe jest stworzenie systemu swobodnej nawigacji, w którym użytkownik może wirtualnie poruszać się w scenie zarejestrowanej przy pomocy rzadkiego systemu wielokamerowego.

PODZIĘKOWANIA

Praca finansowana ze środków przyznanych przez Ministerstwo Nauki i Szkolnictwa Wyższego na działalność statutową polegającą na prowadzeniu badań naukowych lub prac rozwojowych oraz zadań z nimi związanych, służących rozwojowi młodych naukowców oraz uczestników studiów doktoranckich.

LITERATURA

- [1] Domański Marek, Grajek Toamsz, Klimaszewski Krzysztof, Kurc Maciej, Stankiewicz Olgierd, Stankowski Jakub, Wegner Krzysztof. 2009. „Poznań multiview video test sequences and camera parameters”. ISO/IEC JTC1/SC29/WG11 MPEG2009, M17050, Xi'an, Chiny.
- [2] Domański Marek, Dziembowski Adrian, Kuehn Agnieszka, Kurc Maciej, Łuczak Adam, Mieloch Dawid, Siast Jakub, Stankiewicz Olgierd, Wegner Krzysztof. 2014. „Poznan Blocks – a multiview video test sequence and camera parameters for free viewpoint television”. ISO/IEC JTC1/SC29/WG11 MPEG2014, M32243, San Jose, Stany Zjednoczone.
- [3] Domański Marek, Dziembowski Adrian, Kuehn Agnieszka, Mieloch Dawid. 2014. „Telewizja swobodnego punktu widzenia – nowa usługa czy przyszłość wizja?”. Przegląd Telekomunikacyjny, nr 8-9/2014: 734-737.
- [4] Domański Marek, Dziembowski Adrian, Grzelka Adam, Łuczak 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 MPEG2016, M38247, Genewa, Szwajcaria.
- [5] Dziembowski Adrian. 2018. „Synteza widoków wirtualnych w rzadkich systemach wielokamerowych dla zastosowań w swobodnej nawigacji”. Rozprawa doktorska, Politechnika Poznańska.
- [6] Dziembowski Adrian, Domański Marek. 2018. „View and depth preprocessing for view synthesis enhancement”. International Journal of Electronics and Telecommunications, tom 64, nr 3: 269-275.
- [7] Dziembowski Adrian, Samelak Jarosław, Domański Marek. 2018. „View selection for virtual view synthesis in free navigation systems”. International Conference on Signals and Electronic Systems, ICSES 2018, Kraków.
- [8] Dziembowski Adrian, Stankowski Jakub. 2018. „Real-time CPU-based virtual view synthesis”. International Conference on Signals and Electronic Systems, ICSES 2018, Kraków.
- [9] Dziembowski Adrian, Domański Marek. 2018. „Adaptive color correction method in virtual view synthesis”. 3DTV Conference 2018, Sztokholm/Helsinki, Szwecja/Finlandia.
- [10] Dziembowski Adrian, Grzelka Adam, Mieloch Dawid, Stankiewicz Olgierd, Domański Marek. 2017. „Enhancing view synthesis with image and depth map upsampling”. International Conference on Systems, Signals and Image Processing, IWSSIP 2017, Poznań.
- [11] Dziembowski Adrian, Grzelka Adam, Mieloch Dawid, Stankiewicz Olgierd, Wegner Krzysztof, Domański Marek. 2016. „Multiview Synthesis – improved view synthesis for virtual navigation”. 32nd Picture Coding Symposium, PCS 2016, Norymberga, Niemcy.
- [12] Dziembowski Adrian, Stankiewicz Olgierd. 2018. „[MPEG-I Visual] Fast color correction technique for view synthesis”. ISO/IEC JTC1/SC29/WG11 MPEG2018, M43694, Ljubljana, Słowenia.
- [13] Dziembowski Adrian, Stankowski Jakub. 2018. „Szybka synteza widoków wirtualnych w systemach telewizji swobodnego punktu widzenia”. Przegląd Telekomunikacyjny, nr 6/2018: 330-333.
- [14] Dziembowski Adrian, Grzelka Adam, Mieloch Dawid. 2017. „Zwiększanie rozdzielczości obrazu i mapy głębi w celu poprawy jakości syntezy widoków wirtualnych”. Przegląd Telekomunikacyjny, tom 86, nr 6/2017: 405-408.
- [15] Dziembowski Adrian, Grzelka Adam, Mieloch Dawid. 2016. „Wielowidokowa synteza w systemach telewizji swobodnego punktu widzenia”. Przegląd Telekomunikacyjny, tom 88, nr 6/2016: 233-236.
- [16] Goorts Patrick. 2014. „Real-time adaptive plane sweeping for free viewpoint navigation in soccer scenes”. Rozprawa doktorska, Hasselt University, Belgia.
- [17] Kovacs Peter Tamas, Fekete Adam, Lackner Kristof, Adhikarla Vamsi Kiran, Zare Alireza, Balogh Tibor. 2015. „[FTV AHG] Big Buck Bunny light-field test sequence”. ISO/IEC JTC1/SC29/WG11 MPEG2013, M35721, Genewa, Szwajcaria.
- [18] Rec. ITU-T. 2008. „Subjective video quality assessment methods for multimedia applications”. Recommendation ITU-T P.910.
- [19] Senoh Takanori, Yamamoto Kenji, Tetsutani Nobuji, Yasuda Hiroshi, Wegner Krzysztof. 2017. „View Synthesis Reference Software (VSRS) 4.2 with improved inpainting and hole filling”. ISO/IEC JTC1/SC29/WG11 MPEG2017, M40657, Hobart, Australia.
- [20] Stankiewicz Olgierd, Domański Marek, Dziembowski Adrian, Grzelka Adam, Mieloch Dawid, Samelak Jarosław. 2018. „A free-viewpoint television system for horizontal virtual navigation”. IEEE Transactions on Multimedia, tom 20, nr 8/2018: 2182-2195.
- [21] Tanimoto Masayuki. 2010. „FTV (Free-Viewpoint TV)”. IEEE International Conference on Image Processing, ICIP 2010, Hongkong, Chiny.
- [22] Zitnick C. Lawrence, Kang Sing Bing, Uyttendaele Matthew, Winder Simon, Szeliski Richard. 2004. „High-quality video view interpolation using a layered representation”. ACM Transactions on Graphics, tom 3, nr 23: 600-608.