

ESTYMACJA GŁĘBI W TELEWIZJI SWOBODNEGO PUNKTU WIDZENIA

DEPTH ESTIMATION IN FREE-VIEWPOINT TELEVISION

Streszczenie: Stworzenie nowej metody estymacji map głębi przeznaczonej dla systemów telewizji swobodnego widzenia jest głównym celem przedstawionych badań. W telewizji swobodnego punktu widzenia możliwości widza są rozszerzone poprzez możliwość kontroli aktualnie oglądanego przez niego punktu widzenia sceny. Nowa metoda estymacji map głębi zaproponowana przez autora składa się z trzech części: przestrzennie spójnej estymacji map głębi opartej na segmentacji widoków, metodzie zwiększenia spójności czasowej map głębi zmniejszającej złożoność obliczeniową estymacji oraz nową metodę zrównoleglenia procesu optymalizacji opartego na wykorzystaniu grafów.

Abstract: The development of a novel depth estimation method for free viewpoint television systems is the main goal of presented research. In free viewpoint television the functionalities offered to a viewer are extended by the possibility of controlling the displayed viewpoint of a scene. The novel method for depth estimation proposed by the author of the dissertation consists of three parts: the interview consistent segment-based depth estimation method, the temporal consistency enhancement that simultaneously increases temporal consistency of depth maps and reduces the complexity of estimation, and a new method of parallelisation for graph based depth estimation methods.

Słowa kluczowe: mapa głębi, segmentacja obrazu, telewizja swobodnego punktu widzenia

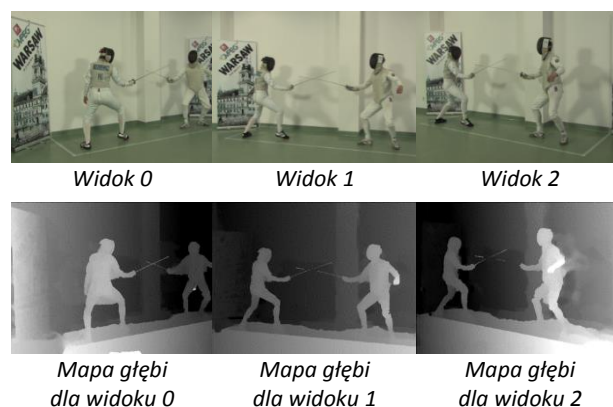
Keywords: depth map, image segmentation, free-viewpoint television

1. WSTĘP

Telewizja swobodnego punktu widzenia (ang. free-viewpoint television – FTV) [5], [11], [15] umożliwia swobodną nawigację poprzez naturalną trójwymiarową scenę. Użytkownik systemu FTV może oglądać nie tylko widoki zarejestrowane przez rzeczywiste kamery – idea swobodnej nawigacji zakłada możliwość oglądania sceny z dowolnego miejsca i kierunku obserwacji. Tak zwane widoki wirtualne mogą być umieszczone nawet między lub przed rzeczywistymi kamerami systemu.

Po pierwsze, wiele widoków sceny jest rejestrowanych synchronicznie przez kamery systemu FTV. Następnie, trójwymiarowa reprezentacja sceny jest estymowana i używana, razem z rzeczywistymi widokami, do syntezy widoków wirtualnych [1], [2]. Najczęściej wykorzystywaną reprezentacją sceny w systemach FTV jest MVD

(ang. multiview plus depth), w której scena jest reprezentowana jako zbiór widoków i odpowiadających im map głębi. Przykład takiej reprezentacji dla jednej z testowych sekwencji został przedstawiony na rys. 1. Mapa głębi to sekwencja macierzy próbek głębi dla każdego punktu danego widoku. Dla uproszczenia, macierze te są zazwyczaj wizualizowane jako obrazy w skali szarości.



Rys. 1. Przykład reprezentacji MVD dla sekwencji testowej „Poznań Fencing2”.

Mapy głębi mogą być estymowane algorytmicznie z użyciem wielu widoków zarejestrowanej sceny. Głębia punktu jest wyliczana na podstawie poszukiwania punktu odpowiadającego, czyli punktu, który reprezentuje tę samą część sceny w innym widoku. Estymacja głębi na podstawie widoków zarejestrowanych przez kamery systemu FTV jest głównym problemem rozpatrywanym przez autora.

Jakość map głębi ma bezpośredni wpływ na jakość syntezowanych widoków wirtualnych [9], a zatem również na jakość końcowego wrażenia nawigowania przez scenę. Stworzenie nowej oryginalnej metody estymacji map głębi, która zapewnia wysoką jakość widoków wirtualnych, jest głównym celem zaprezentowanych badań. W celu uniknięcia błędów widocznych w widokach wirtualnych, które mogą pogorszyć jakość swobodnej nawigacji dla końcowego użytkownika systemu FTV, metoda estymacji map głębi musi spełnić określony zbiór wymagań. Wymagania te wynikają ze struktury systemów FTV i sposobu działania syntezy widoków wirtualnych. Metoda estymacji głębi powinna charakteryzować się wysoką jakością estymowanej głębi, ze szczególnym uwzględnieniem międzywidokowej i czasowej spójności

map głębi oraz poprawnej reprezentacji krawędzi obiektów, możliwością jednoczesnego estymowania map głębi dla wszystkich widoków, oraz wszechstronnością, na którą składa się brak założeń co do liczby i rozmieszczenia kamer systemu.

Konieczność spełnienia wszystkich zaprezentowanych wymagań czyni bardzo trudnym stworzenie wszechstronnej metody estymacji map głębi, którą można by z powodzeniem użyć dla wszystkich systemów FTV. Z tego powodu, przeprowadzone badania znacząco wpływają na dalszy rozwój systemów FTV. Istotność takich badań można zauważyć w mnogości i różnorodności innych uznanych metod estymacji głębi zaprezentowanych w ciągu ostatnich dwóch lat (np. [6], [10], [17]).

Niniejszy referat stanowi podsumowanie osiągnięć opisanych w pracy doktorskiej autora [7].

2. ZAPROPONOWANA METODA ESTYMACJI MAP GŁĘBI

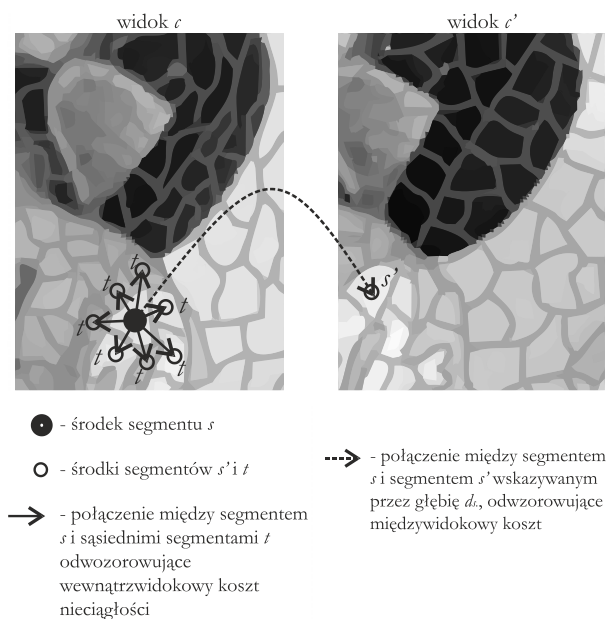
Oryginalność zaproponowanej metody estymacji głębi, oraz jej szczególna użyteczność dla systemów telewizji swobodnego punktu widzenia, są rezultatem łącznego wykorzystania następujących rozwiązań:

- 1) Estymacja głębi jest przeprowadzana dla segmentów, a nie dla punktów widoków wejściowych, przez co czas estymacji może być zmniejszony bez konieczności zmniejszania rozdzielczości estymowanych map głębi.
- 2) Przedstawienie oryginalnej funkcji kosztu, dedykowanej dla estymacji głębi bazującej na segmentacji – estymacja jest przeprowadzana dla wszystkich widoków i umożliwia estymację map głębi spójnych międzywidosko, bez założeń dotyczących rozmieszczenia kamer.
- 3) Estymacja głębi jest przeprowadzana dla segmentów otrzymywanych niezależnie dla wszystkich widoków.
- 4) Proponowana oryginalna metoda poprawy spójności czasowej map głębi wykorzystuje głęboko wyestymowane dla poprzednich ramek podczas estymacji nowej ramki mapy głębi. Zaproponowane narzędzie zwiększa spójność czasową map głębi i zmniejsza czas estymacji.
- 5) Proponowana metoda estymacji map głębi wykorzystuje oryginalną metodę zrównoleglania obliczeń przeznaczoną dla estymacji głębi bazującej na optymalizowaniu funkcji kosztu z użyciem grafów.

2.1. Estymacja głębi bazująca na segmentacji

Estymacja głębi w proponowanej metodzie bazuje na minimalizacji funkcji kosztu, na którą składają się dwa elementy: wewnątrzwidokowy koszt nieciągłości głębi $V_{(s,t)}$ i międzywidokowy koszt pasowania $M_{(s,s')}$ odpowiedzialny za międzywidokową spójność map głębi. Lokalne minimum zaproponowanej funkcji kosztu jest wyznaczane poprzez użycie algorytmu graph cut [4]. Inaczej niż w znanych metodach estymacji map głębi opartych na wykorzystaniu grafów, w których wierzchołki grafu reprezentują każdy punkt widoków wejściowych, w proponowanej metodzie każdy wierzchołek odpowiada jednemu segmentowi (rys. 2). Wierzchołki są łączone

dwoma typami połączeń, które reprezentują międzywidokowy koszt pasowania i wewnątrzwidokowy koszt nieciągłości.



Rys. 2. Wizualizacja wewnątrzwidokowego kosztu nieciągłości głębi i międzywidokowego kosztu pasowania dla segmentu s przy estymacji przeprowadzanej dla dwóch widoków.

Zaproponowana metoda oparta na segmentach zmniejsza zatem liczbę wierzchołków w grafie w porównaniu do estymacji opartej na punktach. Złożoność optymalizacji jest zależna od liczby wierzchołków i połączeń między nimi. Umożliwia to zmniejszenie czasu wyznaczania mapy głębi. Ponadto, mapy głębi są nadal estymowane w tej samej rozdzielczości co wejściowe widoki, a z uwagi na własności segmentów, krawędzie obiektów w mapach głębi odpowiadają krawędzom widocznym w widokach wejściowych.

Użycie segmentacji obrazu w estymacji map głębi jest spotykane również w innych znanych metodach (m.in. [6], [10]). Co wyróżnia metodę zaprezentowaną przez autora, to estymacja głębi w całości oparta na segmentacji obrazu. W innych metodach na jednym z etapów estymacji konieczna jest optymalizacja przeprowadzana dla poszczególnych punktów obrazu. Brak konieczności przeprowadzania takiej optymalizacji daje znaczące zmniejszenie złożoności proponowanej metody względem innych metod bazujących na segmentacji obrazu.

W celu uzyskania międzywidokowej spójności estymowanych map głębi, koszt pasowania potencjalnie odpowiadających sobie punktów w sąsiednich widokach nie jest obliczany niezależnie dla każdego widoku. Konwencjonalny koszt pasowania został zastąpiony międzywidokowym kosztem pasowania $M_{(s,s')}(d_s)$, zdefiniowanym pomiędzy parą segmentów s i s' odpowiadających sobie dla aktualnie rozpatrywanej głębi d_s . Dzięki temu spójność międzywidokowa głębi zapewniana jest nie poprzez dodatkowe przetwarzanie map głębi, ale od razu przy estymacji głębi. Zaproponowany międzywidokowy koszt pasowania nie wymaga aby segmentacja była międzywi-

dokowo spójna dla wejściowych widoków, dlatego segmentacja jest dokonywana niezależnie dla każdego widoku, co dodatkowo zmniejsza całkowitą złożoność zaprezentowanej metody estymacji map głębi.

Wewnątrzwidokowy koszt nieciągłości głębi jest obliczany między wszystkimi sąsiadującymi ze sobą segmentami wewnątrz tego samego widoku z użyciem podobieństwa sąsiadujących segmentów. Jak zaprezentował autor w [8], użycie zmiennego współczynnika gładkości obliczanego na podstawie podobieństwa sąsiednich segmentów jednocześnie powoduje zwiększenie jakości estymowanych map głębi oraz zmniejsza czas estymacji.

2.2. Metoda poprawy spójności czasowej

W praktycznych systemach telewizji swobodnego punktu widzenia pozycje kamer nie zmieniają się w trakcie akwizycji sekwencji. Z tego powodu, w takiej sekwencji jedynie mała część obrazu znacząco zmienia się w kolejnych ramkach obrazu. W proponowanej metodzie poprawy spójności czasowej map głębi nowa wartość głębi jest estymowana tylko dla segmentów reprezentujących fragmenty sceny, które znacząco zmieniły się (kolorystycznie) w porównaniu z poprzednią lub pierwszą ramką zarejestrowanej sekwencji.

Proponowana metoda umożliwia oznaczenie segmentu jako niezmienionego względem poprzednich ramek. Taki segment jest ciągle wykorzystywany w obliczaniu minimalizowanej funkcji celu, jednak nie jest reprezentowany przez żaden wierzchołek w strukturze optymalizowanego grafu. Zmniejsza to liczbę wierzchołków grafu, zmniejszając czas estymacji, oraz jednocześnie zwiększając spójność czasową estymowanych map głębi.

Dla pierwszej ramki mapy głębi, określanej jako ramka głębi „typu I” (poprzez analogię do terminologii kompresji sekwencji wizyjnych, w której ramka I jest kompresowana bez użycia predykcji na podstawie poprzednich ramek sekwencji), estymacja jest przeprowadzana dla wszystkich segmentów. Dalsze ramki (oznaczane jako ramki głębi „typu P”) wykorzystują głębię wyestymowaną dla poprzednich ramek głębi typu P i ramki głębi typu I.

Wprowadzenie dwóch ramek odniesienia ma korzystny wpływ na jakość wirtualnej nawigacji. Po pierwsze, użycie głębi z ramki głębi typu P pozwala na użycie wcześniej wyestymowanej głębi dla obiektów, które zmieniły swoją pozycję w trakcie trwania sekwencji (np. dla obiektów, które pojawiły się w sekwencji w pewnym momencie, a następnie nie zmieniały swojej pozycji). Po drugie, użycie głębi z ramki głębi typu I minimalizuje migotanie głębi obiektów tła.

Spójność czasowa w innych metodach może być uzyskana poprzez dodatkowe przetwarzanie map głębi [16]. Metody te bazują najczęściej na segmentacji tła, przez co nie zwiększają spójności czasowej obiektów usytuowanych blisko kamer. Inne metody bazują na odszumianiu wejściowych sekwencji [12]. Zaletą takiego podejścia jest niezależnienie od wybranej późniejszej metody estymacji głębi, jednakże, dodatkowy krok (odszumianie) zwiększa całkowity czas przetwarzania sekwencji. Me-

toda poprawy spójności czasowej map głębi zaproponowana przez autora, w przeciwieństwie do powyższych metod, zmniejsza złożoność całego procesu estymacji.

2.3. Metoda zrównoleglenia obliczeń

W zaproponowanej metodzie zrównoleglenia obliczeń w estymacji map głębi każdy z n stworzonych wątków estymuje głębię o n razy mniejszej liczbie poziomów głębi (poprzez poziom głębi rozumie się odległość od płaszczyzny równoległej do płaszczyzny przetwornika kamery, reprezentującą wartości głębi możliwe do wyznaczenia). Pojedynczy proces estymacji przeprowadza optymalizację funkcji kosztu dla różnych zbiorów poziomów głębi. Z tego powodu, jeżeli obliczenia są przeprowadzane na komputerze z wielordzeniowym procesorem, to całkowity czas przetwarzania dla estymacji głębi może zostać zmniejszony. Dzięki zaproponowanemu podziałowi obliczeń zachowane są zarówno spójność międzywidokowa map głębi i ich spójność czasowa, które są kluczowe dla jakości syntezy widoków wirtualnych.

Mapy głębi o zredukowanej liczbie poziomów głębi wyznaczone przez różne procesy estymacji muszą być połączone w jedną końcową mapę głębi. Proces łączenia jest przeprowadzany poprzez optymalizację funkcji kosztu, jednak dla każdego segmentu podczas jednego cyklu łączenia rozpatrywane są tylko dwa poziomy głębi – wyznaczone przez wątki t oraz $t + 1$. Zatem, podczas jednego cyklu łączenia, łączone są ze sobą dwie mapy głębi. Aby uzyskać końcową mapę głębi dla zrównoleglenia na n wątków, wymagane jest $\lceil \log_2(n) \rceil$ cykli łączenia.

Inne uznane metody zrównoleglenia estymacji map głębi bazujące na optymalizacji grafu skupiają się na zrównolegleniu algorytmu rozwiązywania grafu. Przedstawiane metody często opisują implementację algorytmu graph cut na procesorach graficznych. Czas obliczeń jest zmniejszony wtedy nawet na 40-krotne, jednakże, takie implementacje wprowadzają ograniczenia w liczbie widoków możliwych do użycia podczas estymacji (w związku z relatywnie małym rozmiarem pamięci przeznaczanej dla procesorów graficznych) i zakładają regularną siatkę wierzchołków, niemożliwą do wykorzystania w estymacji opartej na segmentacji.

Proponowana przez autora metoda nie wprowadza żadnych ograniczeń dotyczących optymalizowanego grafu, przez co zrównoleglenie międzywidokowo spójnej estymacji głębi opartej na segmentach jest możliwe.

3. Wyniki eksperymentalne

3.1. Porównanie z metodą DERS

Zaproponowana przez autora metoda estymacji głębi została porównana z metodą zaimplementowaną w oprogramowaniu Depth Estimation Reference Software przygotowywanym i rozwijanym przez grono ekspertów MPEG ISO/IEC [13]. Oprogramowanie DERS wykorzystuje punktową metodę estymacji opartą na optymalizacji grafu powszechnie używaną w środowisku naukowym w badaniach dotyczących telewizji swobodnego punktu widzenia. Metoda zaimplementowana w DERS, tak jak metoda autora, również wykorzystuje algorytm graph cut, ale jest oparta na optymalizacji przeprowadzanej dla punktów a nie segmentów widoków.

Jakość map głębi w przeprowadzonych badaniach jest mierzona pośrednio, poprzez jakość widoków wirtualnych syntezy z użyciem estymowanych map głębi. Dla końcowego użytkownika systemu telewizji swobodnego punktu widzenia to właśnie jakość widoku wirtualnego określa całościową jakość systemu. Do celów syntezy widoków wirtualnych użyto metody zaimplementowanej w oprogramowaniu Virtual View Synthesis Software (VSRS) [14].

Porównanie jakości map głębi estymowanych przy użyciu proponowanej metody oraz DERS zaprezentowano w tabeli 1, razem z porównaniem czasu estymacji dla obu metod. Dla wszystkich sekwencji testowych proponowana metoda estymuje mapy głębi o lepszej jakości niż metoda odniesienia. Wzrost jakości wyrażonej przez PSNR lumeny wynosi ponad 2.5 dB. Jednocześnie, czas trwania estymacji dla proponowanej metody jest zawsze krótszy – średnio czterokrotnie względem metody odniesienia.

Tab. 1. Porównanie jakości widoków wirtualnych zsyntezowanych używając map głębi dla proponowanej metody i oprogramowania DERS.

Sekwencja testowa	PSNR lumeny widoku wirtualnego [dB]			Czas estymacji głębi dla jednej ramki mapy głębi [s]		
	DERS	Metoda autora	Zysk	DERS	Metoda autora	Stosunek
Ballet	27.93	28.69	0.76	882	499	57%
Breakdancers	31.13	32.19	1.06	949	255	27%
BBB Butterfly	29.97	33.20	3.23	593	279	47%
BBB Rabbit	22.59	27.21	4.62	744	92	12%
Poznań Blocks	21.97	27.20	5.23	1445	313	22%
Poznań Blocks2	25.67	28.12	2.45	1060	210	20%
Poznań Fencing2	26.74	28.60	1.86	2254	391	17%
Poznań Service2	23.69	25.51	1.82	2780	305	11%
	Średnia: 2.63			Średnia: 27%		

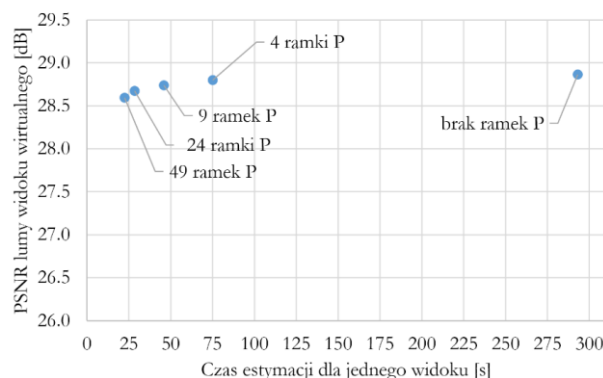
3.2. Wpływ poprawy spójności czasowej

Zbadano wpływ proponowanej metody poprawy spójności map głębi na ich jakość oraz czas trwania estymacji. Estymacja map głębi dla sekwencji testowych została powtórzona dla różnej liczby ramek głębi typu P pomiędzy ramkami głębi typu I. Jakość głębi wyrażoną poprzez jakość widoków wirtualnych zaprezentowano na rys. 3. Użycie ramek głębi typu P w pomijalnym stopniu zmniejszyło jakość map głębi (niecałe 0.25 dB). Jednakże, czas trwania estymacji został znacząco zmniejszony (prawie 15 krotnie).

W celu określenia poprawy spójności czasowej map głębi, zsyntezowane widoki wirtualne zostały zakodowane za pomocą kodera HEVC [3]. Poprawa spójności czasowej jest mierzona za pomocą zmniejszenia rozmiaru strumienia bitowego zakodowanych widoków wirtualnych w porównaniu do widoków wirtualnych zsyntezowanych z użyciem głębi wyestymowanych bez użycia proponowanej metody poprawy spójności.

Użycie proponowanej metody poprawy spójności map głębi powoduje znaczące zmniejszenie rozmiaru strumienia bitowego zadokowanych widoków wirtualnych, średnio o prawie 32%. Takie zmniejszenie rozmiaru

oznacza zwiększenie wydajności predykcji czasowej kodera, co oznacza, że kolejne ramki kodowanej sekwencji były do siebie bardziej podobne. Potwierdza to znaczące zwiększenie spójności czasowej widoków wirtualnych, a więc również estymowanych map głębi.

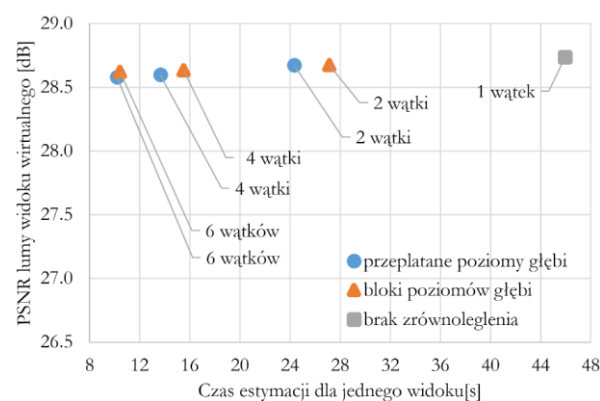


Rys. 3. Średnia jakość widoku wirtualnego zsyntezowanego z użyciem map głębi estymowanych dla różnych liczb ramek głębi typu P między ramkami głębi typu I oraz czas estymacji.

3.3. Wpływ zrównoleglenia obliczeń

W ostatnim zaprezentowanym eksperymencie zbadano wpływ proponowanej metody zrównoleglenia obliczeń na jakość map głębi i czas trwania estymacji map głębi. Estymacja została powtórzona dla różnych scenariuszy zrównoleglenia: bez użycia zrównoleglenia (estymacja używając jednego wątku) oraz zrównoleglenie na 2, 4 oraz 6 wątków. Ponadto, testy zostały przeprowadzone dla dwóch proponowanych schematów zrównoleglenia: dla poziomów głębi przeplecionych między wątkami i dla estymacji dla poziomów głębi podzielonych na bloki kolejnych poziomów.

Wyniki przeprowadzonych eksperymentów (jakość estymowanych map głębi wyrażona poprzez jakość widoku wirtualnego oraz czas trwania estymacji map głębi) zaprezentowano na rys. 4.



Rys. 4. Średnia jakość widoku wirtualnego zsyntezowanego z użyciem map głębi estymowanych dla różnej liczby użytych wątków i różnych typów zrównoleglenia oraz czas estymacji głębi.

Nawet dla zrównoleglenia na 6 wątków zmiana jakości estymowanych map głębi jest nieznaczna (jedynie 0.1 dB), jednak czas trwania estymacji został zmniejszony średnio więcej niż 4,5 krotnie. Zaprezentowane wyniki

wskazują, że zaproponowana metoda zrównoleglania obliczeń stanowi znaczące i przydatne rozszerzenie proponowanej metody estymacji map głębi.

4. PODSUMOWANIE

Prezentowano oryginalną metodę estymacji map głębi umożliwiającą wyznaczanie map głębi spójnych międzywidelkowo dla dowolnego rozstawienia kamer systemu FTV. Estymacja jest przeprowadzana dla segmentów, a nie dla poszczególnych punktów wejściowych widoków. Proponowana metoda estymuje głębie o jakości wyższej niż metoda odniesienia DERS powszechnie używana w środowisku naukowym w badaniach dotyczących telewizji swobodnego punktu widzenia. Ponadto, zaproponowana metoda zmniejsza czas estymacji w porównaniu z metodą odniesienia (średnio czterokrotne zmniejszenie czasu trwania estymacji).

Dodatkowo, przedstawiono oryginalną metodę poprawy spójności map głębi pozwalającą na nawet 15-krotne zmniejszenie czasu trwania estymacji głębi. Znacząca poprawa spójności map głębi została potwierdzona (używając autorskiej metody pomiaru spójności czasowej) poprzez zauważalne zmniejszenie prędkości bitowej zakodowanych widoków wirtualnych nawet o 32%. W metodzie estymacji zaimplementowano również oryginalną metodę zrównoleglania obliczeń w estymacji map głębi dla estymacji bazującej na optymalizacji z wykorzystaniem grafów. Zaproponowana metoda pozwoliła na 4,5-krotne zmniejszenie czasu trwania estymacji dla zrównoleglenia 6-wątkowego.

Szczególna przydatność przedstawionej metody estymacji głębi została już potwierdzona przez jej wdrożenie w działającym systemie FTV, opracowanym przez Katedrę Multimedia Telekomunikacji i Mikroelektroniki Politechniki Poznańskiej [11]. Mapy głębi, które można z powodzeniem stosować do syntezy widoku wirtualnego, można estymować za pomocą proponowanej metody w czasie poniżej 5 sekund na jedną ramkę. W związku z tym czas od rejestracji sekwencji do prezentacji swobodnej nawigacji ulega znaczącemu skróceniu, co może przyspieszyć praktyczne wykorzystanie systemów FTV.

LITERATURA

[1] Ceulemans B., Lu S., Lafruit G., Munteanu A. 2018. "Robust Multiview Synthesis For Wide-Baseline Camera Arrays". *IEEE Transactions on Multimedia* 20(9): 2235-2248.

[2] Dziembowski A., Grzelka A., Mieloch D., Stankiewicz O., Wegner K., Domański M. 2016. "Multiview synthesis — Improved view synthesis for virtual navigation". *2016 Picture Coding Symposium (PCS)*: 1-5.

[3] HEVC reference codec. Dokumentacja i implementacja dostępna online: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/.

[4] Kolmogorov V., Zabini R. 2004. "What energy functions can be minimized via graph cuts?". *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(2): 147-159.

[5] Lafruit G., Domański M., Wegner K., Grajek T., Senoh T., Jung J., Kovács P., Goorts P., Jorissen L., Munteanu A., Ceulemans B., Carballeira P., García S., Tanimoto M. 2016. "New visual coding exploration in MPEG: Super-MultiView and Free Navigation in Free viewpoint TV". *2016 Proceedings of the Electronic Imaging Conference: Stereoscopic Displays and Application*, San Francisco: 1-9.

[6] Li L., Zhang S., Yu X., Zhang L. 2018. "PMSC: PatchMatch-Based Superpixel Cut for Accurate Stereo Matching". *IEEE Transactions on Circuits and Systems for Video Technology* 28(3): 679-692.

[7] Mieloch D. 2018. „Depth Estimation in Free-Viewpoint television”. Praca doktorska.

[8] Mieloch D., Dziembowski A., Grzelka A. 2016. "Segmentacja obrazu w estymacji map głębi". *Przegląd Telekomunikacyjny* 88(6): 241-244.

[9] Nur G., Dogan S., Arachchi H., Kondoz A. 2010. "Impact of depth map spatial resolution on 3D video quality and depth perception". *3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video*: 1-4.

[10] Seo D., Jo K. 2017. "Multi-layer superpixel-based MeshStereo for accurate stereo matching". *10th International Conference on Human System Interactions (HSI)*: 242-245.

[11] Stankiewicz O., Domański M., Dziembowski A., Grzelka A., Mieloch D., Samelak J. 2018. "A Free-viewpoint Television System for Horizontal Virtual Navigation". *IEEE Transactions on Multimedia* 20(8): 2182-2195.

[12] Stankiewicz O., Domański M., Wegner K. 2015. "Estimation of temporally-consistent depth maps from video with reduced noise". *3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*: 1-4.

[13] Stankiewicz O., Wegner K., Tanimoto M., Domański M. 2013. "Enhanced Depth Estimation Reference Software (DERS) for Free-viewpoint Television". Dokument ISO/IEC JTC1/SC29/WG11 MPEG M31518.

[14] Stankiewicz O., Wegner K., Tanimoto M., Domański M. 2013. "Enhanced view synthesis reference software (VSRS) for Free-viewpoint Television". Dokument ISO/IEC JTC 1/SC 29/WG 11, MPEG M31520.

[15] Tanimoto M. 2014. "FTV standardization in MPEG.". *2014 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*: 1-4.

[16] Yao C., Tillo T., Zhao Y., Xiao J., Bai H., Lin C. 2014. "Depth Map Driven Hole Filling Algorithm Exploiting Temporal Correlation Information". *IEEE Transactions on Broadcasting* 60(2): 394-404.

[17] Zhu H., Wang Q., Yu J. 2017 "Occlusion-Model Guided Antiocclusion Depth Estimation in Light Field". *IEEE Journal of Selected Topics in Signal Processing* 11(7): 965-978.