

Przeegląd



MIESIĘCZNIK
STOWARZYSZENIA
ELEKTRYKÓW
POLSKICH



TELEKOMUNIKACYJNY

WYDAWNICTWO SIGMA NOT



WIADOMOŚCI
TELE
KOMUNIKACYJNE

8-9
2014

ISSN 1230-3496

Cena: 54,60 zł (w tym 5% VAT)

KSTiT 2014



Centrum Wykładowo-Konferencyjne Politechniki Poznańskiej

**XXX Krajowe Sympozjum
Telekomunikacji i Teleinformatyki
Poznań, 3-5 września 2014 r.**

- [14] Ling F.: *Optimal reception, performance bound and cut-off rate analysis of reference assisted coherent CDMA communications with applications*, IEEE Transactions on Comm. vol. COM-47 Oct. 1999
- [15] Ling F.: *Digital Rate Conversion with a Non-Rational Ratio for High Speed Echo-Cancellation Modems*, IEEE Proc. ICASSP'93, Vol. III, Apr. 1993, Minneapolis, MN

- [16] Proakis J. G., et. al: *Advanced Digital Signal Processing*, Macmillan Publishing Co., 1992
- [17] Rabiner L. R., Crochiere R. E.: *Multirate Digital Signal Processing*, Prentice Hall PTR, Englewood Cliffs, New Jersey, 1983
- [18] Ling F.: *Timing synchronization*, <http://www.slideshare.net/FuyunLing/timing-synchronization-f-lingv1>, Slideshare, 2013

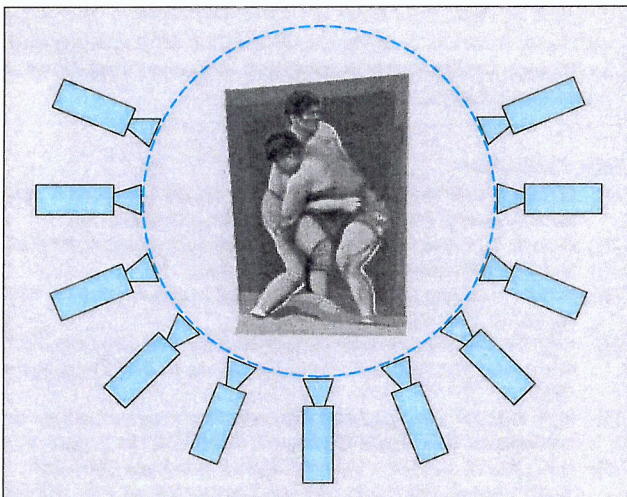
Marek DOMAŃSKI*, Adrian DZIEMBOWSKI*,
Agnieszka KUEHN*, Dawid MIELOCH*

KST:T 2014

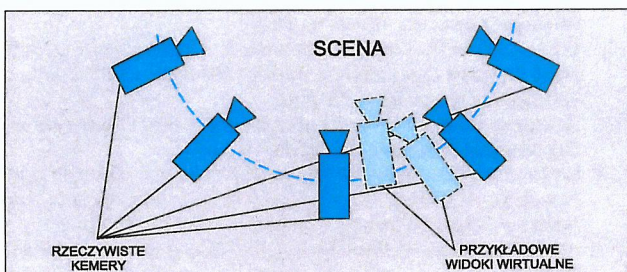
Telewizja swobodnego punktu widzenia – nowa usługa czy futurystyczna wizja?

TELEWIZJA SWOBODNEGO PUNKTU WIDZENIA

Systemy telewizji swobodnego punktu widzenia (*Free-Viewpoint Television – FTV*) umożliwiają pokazywanie sceny z punktów widzenia i z kierunków widzenia płynnie zmienianych przez widza podczas oglądania programu [1–3]. Widz samodzielnie wybiera punkt widzenia i w każdym kolejnym wybranym punkcie ogląda wirtualny widok wyliczony z odpowiedniej



■ Rys. 1. Zasada akwizycji obrazu w systemie wielokamerowym, pracującym dla potrzeb telewizji swobodnego punktu widzenia



■ Rys. 2. Synteza widoków wirtualnych ze zmianą położenia wirtualnego punktu widzenia – zmiana kierunku patrzenia na scenę i zmiana odległości od sceny

* Katedra Telekomunikacji Multimedialnej i Mikroelektroniki, Politechnika Poznańska, e-mail: domanski@et.put.poznan.pl

reprezentacji dynamicznej sceny. Ta reprezentacja jest przedtem wyznaczona z obrazów uzyskanych z pewnej liczby kamer rozmieszczonych wokół sceny (rys. 1). Należy jednak podkreślić, że w wyborze kolejnych punktów patrzenia na scenę widz nie jest ograniczony do widoków z rzeczywistych kamer, ale zazwyczaj ogląda widoki wirtualne (rys. 2), uzyskane w procesie syntezy. Syntezy wirtualnych widoków dokonuje się na podstawie wspomnianej odpowiedniej reprezentacji sceny, którą można nazwać obrazem swobodnego punktu widzenia. W badanych obecnie systemach najczęściej wykorzystuje się reprezentację w formacie „obraz wielowidokowy plus głębia” (*Multiview plus Depth – MVD*), składającą się z pewnej liczby widoków oraz map głębi, czyli tablic przedstawiających odległości poszczególnych punktów obrazu od kamery.

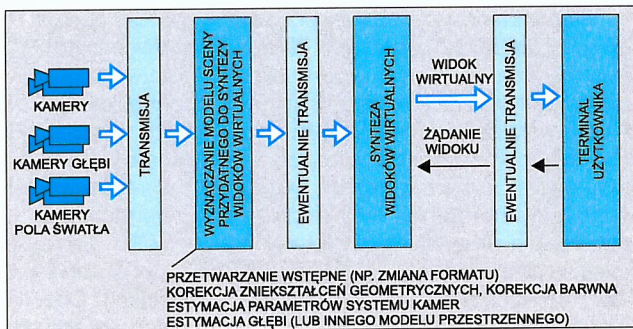
Telewizja swobodnego punktu widzenia może być stosowana do transmisji zawodów sportowych, na przykład zapasów, judo, walk sumo lub bokserskich. Podczas transmisji będzie ona dawała każdemu miłośnikowi tych sportów możliwość obejrzenia poszczególnych akcji z takiego kierunku, który go najbardziej interesuje. Podobne zastosowania można przewidywać na przykład dla konkursów tanecznych, sztuk teatralnych czy rejestrowanych za pomocą kamer przedstawień dziecięcych, oglądanych przez poszczególne rodziny najchętniej z kierunku, z którego aktualnie najlepiej widać dziecko z tej rodziny. Potencjalne zastosowania obejmują także różne instrukcje i materiały szkoleniowe, a nawet filmy fabularne i dokumentalne.

Obecnie trwają badania, mające na celu opracowanie systemów telewizji swobodnego punktu widzenia, które będą mogły być zastosowane do świadczenia usług realizowanych z wykorzystaniem Internetu. Szczególnie duży nacisk na takie badania kładzie się w Japonii, ale także w innych krajach Dalekiego Wschodu.

TELEWIZJA SWOBODNEGO PUNKTU WIDZENIA W SIECIACH TELEINFORMATYCZNYCH

Podstawowe etapy przetwarzania obrazów w systemie telewizji swobodnego punktu widzenia pokazano na rys. 3. W takim systemie obrazy ruchome są zbierane z kamer rozmieszczonych wokół sceny i przekazywane do dalszego przetwarzania. Liczba użytych kamer jest kompromisem pomiędzy kosztem systemu a jakością usługi. Wydaje się, że w wielu zastosowaniach wykorzystanie tylko dziesięciu – kilkunastu kamer umożliwi uzyskiwanie obrazu wirtualnego o wystarczająco dobrej jakości.

Obecnie brakuje efektywnej techniki kompresji obrazu, która wykorzystywałaby podobieństwa widoków tej samej sceny, ale pochodzących z różnych kamer rozmieszczonych wokół owej



■ Rys. 3. Podstawowe etapy przetwarzania obrazów w systemie telewizji swobodnego punktu widzenia

sceny. Wyniki doświadczeń wykonanych w Katedrze Telekomunikacji Multimedialnej pokazują, że dla kamer rozmieszczonych wokół sceny w przybliżeniu na okręgu co około 10 stopni zastosowanie nowej techniki kompresji obrazów wielowidokowych MV-HEVC [4] prowadzi do redukcji łącznej prędkości bitowej wszystkich strumieni zaledwie o kilka procent w stosunku do niezależnego kodowania poszczególnych widoków techniką HEVC [5]. Ponieważ widoki używane do estymacji głębi powinny odznaczać się dobrą jakością, można przewidywać, że przynajmniej w początkowym okresie rozwoju telewizji swobodnego punktu widzenia prawdopodobnie wszystkie widoki będą kodowane niezależnie z wykorzystaniem techniki HEVC skonfigurowanej dla zapewnienia niewielkiej stratności.

Przed właściwą transmisją obrazów z wielu kamer mogą być także przesyłane dodatkowe obrazy sceny ze specjalnymi znacznikami, wykorzystywane podczas wyznaczania parametrów i pozycji kamer, a także podczas korekcji zniekształceń geometrycznych obrazu powstałych wskutek niedoskonałości obiektywów kamer.

W ramach przetwarzania wstępnego dokonuje się także korekcji barwnej widoków, która kompensuje różnice występujące pomiędzy widmowymi charakterystykami poszczególnych kamer.

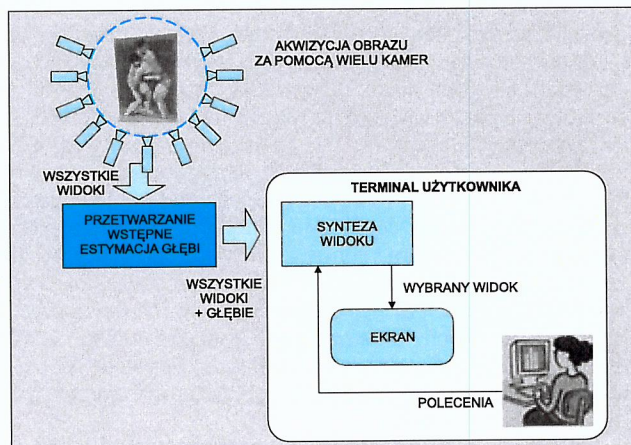
Następnie wyznacza się mapy głębi, co wymaga zastosowania znacznej mocy obliczeniowej. Otrzymane z rzeczywistych kamer widoki i wyznaczone odpowiadające im mapy głębi tworzą reprezentację „obraz wielowidokowy plus głębia” (MVD). Tę reprezentację należy dostarczyć do bloku syntezy widoków wirtualnych. Jeżeli estymacja głębi i synteza widoków wirtualnych są wykonywane w urządzeniach oddalonych od siebie, to reprezentacja MVD także musi zostać poddana kompresji. W tym celu będzie można wykorzystywać technikę 3D-HEVC [6], która obecnie jest w końcowym etapie opracowywania. Niestety, również ta technika jest opracowywana dla przesyłania wielu obrazów z kamer gęsto rozmieszczonych na łuku technika 3D-HEVC umożliwia tylko stosunkowo niewielkie, kilkuprocentowe lub kilkunastoprocentowe zmniejszenie prędkości bitowej w stosunku do prędkości uzyskiwanej z wykorzystaniem znacznie prostszego niezależnego kodowania poszczególnych widoków za pomocą już wzmiankowanej techniki HEVC [5].

W poszczególnych momentach czasu do bloku syntezy widoków wirtualnych do jednego ekranu użytkownika dostarcza się tylko jeden widok żądany przez widza. Taki widok może być poddany kompresji z wykorzystaniem typowej techniki kompresji obrazu HEVC [5]. Wybrany przez widza wirtualny widok nie jest już dalej wykorzystywany w przetwarzaniu obrazów i dlatego może być poddany stosunkowo silniejszej kompresji, niż w przypadku obrazów wielowidokowych przeznaczonych do dalszego przetwarzania.

Wymienione części systemu mogą być w różny sposób rozdzielone pomiędzy zasoby usługodawcy i usługobiorcy. Kalibracja

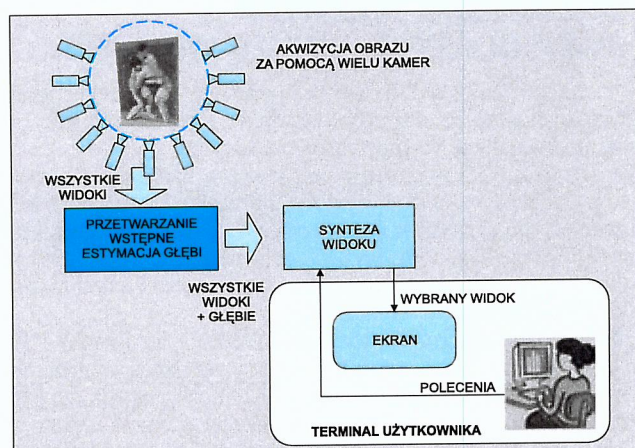
systemu, korekcja barwna oraz estymacja głębi są niezależne od wyboru widoku i dlatego powinny być wykonywane wspólnie dla wszystkich odbiorców usługi. Ze względu na duże wymagania dotyczące mocy obliczeniowej, zwłaszcza estymacja ruchu prawdopodobnie powinna być wykonywana centralnie, a więc z wykorzystaniem zasobów usługodawcy.

Jeśli chodzi o synteze widoków wirtualnych, to może ona być wykonywana zarówno w terminalu odbiorcy, jak i w komputerach usługodawcy. W pierwszym przypadku (rys. 4) do odbiorcy trzeba przesyłać całą reprezentację MVD ruchomej sceny, co przy wykorzystaniu obecnie dostępnych lub zaawansowanych w opracowywaniu technik opartych na technologii HEVC dla widoków o rozmiarach HD (1920 × 1080 punktów - *high definition*) i przy zachowaniu dobrej jakości odbieranego



■ Rys. 4. System telewizji swobodnego punktu widzenia z syntezą widoków prowadzoną lokalnie w terminalu widza

obrazu wymaga przepustowości łącza do odbiorcy rzędu 2 – 4 Mbit/s na jedną kamerę. Dla realistycznej liczby 10 kamer taka transmisja wymagałaby łącza do odbiorcy o zagwarantowanej przepustowości rzędu 20 – 40 Mbit/s, co jeszcze długo będzie przekraczało możliwości wielu potencjalnych usługobiorców. W tym wariantcie systemu telewizji swobodnego punktu widzenia wymagania stawiane terminalowi odbiorcy także są znaczne, gdyż właśnie w terminalu odbiorcy odbywa się synteza widoku wirtualnego. Jest to jednak zadanie możliwe do wykonywania na bieżąco w komputerze osobistym widza, gdyż obecnie produkowane dobre karty graficzne komputerów osobistych umożliwiają syntezę stereoskopowej pary w formacie HD z prędkością ponad 25 par obrazów na sekundę.



■ Rys. 5. System telewizji swobodnego punktu widzenia z syntezą widoków prowadzoną centralnie w serwerze

Wspomniane trudności nie pojawiają się w drugim wariancie konfiguracji systemu (rys. 5), w którym synteza widoków wirtualnych dla poszczególnych użytkowników odbywa się w zasobach usługodawcy. Do odbiorcy strumieniuje się tylko widoki wirtualne odpowiadające kolejno wybranym położeniom wirtualnej kamery sterowanej przez widza, co przy zastosowaniu nowej techniki HEVC wymaga dla obrazów o rozmiarach 1920×1080 punktów przepustowości rzędu 4 Mbit/s. W tym wariancie mogą pojawić się jednak trudności związane z opóźnieniem pomiędzy wydaniem polecenia wyznaczenia odpowiedniego widoku wirtualnego a rzeczywistym jego dostarczeniem do ekranu widza. To zagadnienie musi być jeszcze szczegółowo przebadane.

Pomimo potencjalnych trudności z opóźnieniami, ostatni wariant systemu telewizji swobodnego punktu widzenia wydaje się bardziej interesujący. Ten wariant jest interesujący pod względem technicznym, gdyż unika się strumieniowania do odbiorcy danych z bardzo dużą prędkością bitową. Ponadto może być też bardziej interesujący dla operatora sieci telekomunikacyjnej pod względem ekonomicznym, gdyż w takiej konfiguracji systemu sprzedaje on swoim klientom nie tylko usługi przesyłania danych, ale także usługi ich przetwarzania. Dlatego ten wariant wydaje się też łatwiejszy do szybkiego wdrożenia do użytkowania w sieciach teleinformatycznych. Prawdopodobny scenariusz wprowadzenia usługi telewizji swobodnego punktu widzenia polega na sprzedaży tej usługi użytkownikom Internetu, którym usługodawca będzie także dostarczał swoje oprogramowanie służące do komunikacji z serwerem i prezentacji wybieranych widoków monoskopowych lub stereoskopowych. Taki scenariusz rozwoju nie wymaga ustalania nowych norm, ale może spowodować pojawienie się na rynku różnych wzajemnie niekompatybilnych systemów oferowanych przez różnych operatorów.

Wdrożenie systemu telewizji swobodnego punktu widzenia nawet w omawianym uproszczonym wariancie wymaga jednak rozwiązania różnych problemów badawczych zwłaszcza związanych z poprawą estymacji głębi, z syntezą widoków wirtualnych z niedokładnych map głębi, a także z organizacją systemu. Natomiast dla tego wariantu systemów FTV już tylko stosunkowo niewielki wysiłek badawczy jest potrzebny w zakresie kompresji.

Ponadto dla telewizji swobodnego punktu widzenia należy zbudować odpowiednie systemy przesyłania dźwięku. W prostych systemach, w których syntezy widoku wirtualnego dokonuje się w serwerze, również synteza dźwięku może być wykonywana w tym serwerze. Takiej syntezy można dokonywać korzystając z dźwięków zarejestrowanych przez mikrofony poszczególnych kamer oraz z informacji o wyznaczonej lokalizacji poszczególnych kamer.

NORMALIZACJA W ZAKRESIE TELEWIZJI SWOBODNEGO PUNKTU WIDZENIA

Dalszy rozwój telewizji swobodnego punktu widzenia będzie wymagał ustalenia międzynarodowych norm dotyczących:

- modeli naturalnych scen trójwymiarowych,
- kompresji modeli scen dynamicznych – w szczególności może to być technika kompresji obrazów i dźwięków z wielu dowolnie rozmieszczonych kamer (obrazy przesyłane wraz z odpowiednimi danymi umożliwiającymi wyznaczenie parametrów systemu wielokamerowego oraz ewentualnie z odpowiednimi mapami głębi lub analogicznymi informacjami),
- komunikacji pomiędzy terminalem widza a serwerem.

Obecnie jeszcze nie wiadomo, czy wszystkie wymienione obszary zostaną objęte normami. Być może pojawią się jeszcze inne obszary, w których normalizacja zostanie uznana za konieczną. Oczywiście obecnie nie wiadomo też, jakie wymagania zostaną postawione przyszłym normom, ani tym bardziej jak będą wyglądały unormowane techniki.

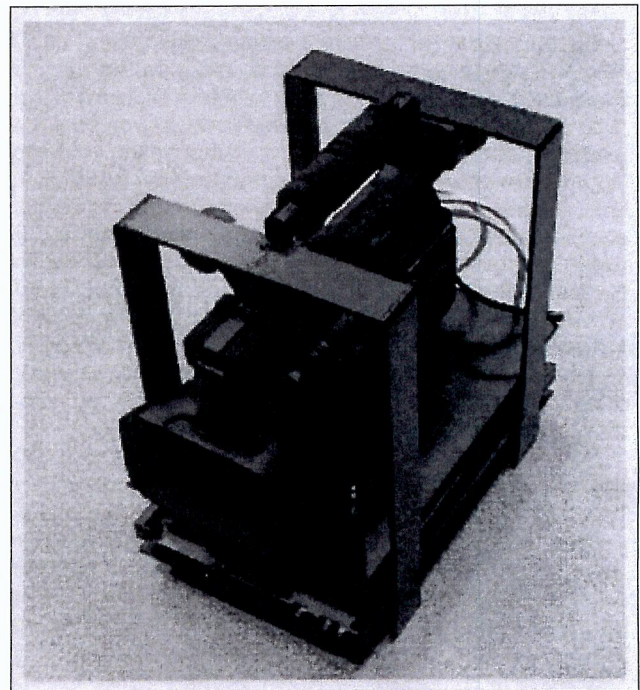
Wspomniane wyżej obszary przyszłej normalizacji obecnie są przedmiotem badań naukowych. Dlatego grupa ekspertów MPEG (*Moving Pictures Experts Group*), oficjalnie nazywana ISO/IEC JTC1/SC29/WG11, od zeszłego roku prowadzi przeglądy wyników badań w zakresie telewizji swobodnego punktu widzenia [7]. Podczas spotkań tej grupy prezentuje się wyniki prac badawczych dotyczących różnych zagadnień związanych z telewizją swobodnego punktu widzenia. Odpowiednie prace są systematycznie publikowane na stronach internetowych grupy MPEG, a w części także na stronach grupy JCT-3V (*ITU-T/ISO/IEC Joint Collaborative Team on 3D Video Coding Extension Development*). Obecnie przeprowadza się też skoordynowane międzynarodowe eksperymenty badawcze (*Exploration Experiments*) dotyczące zagadnień związanych z telewizją swobodnego punktu widzenia.

Należy się spodziewać, że te wszystkie działania doprowadzą w roku 2016 do wezwania do zgłaszania propozycji technik (*Call for Proposals*), które następnie zostaną poddane dokładnej ocenie eksperymentalnej. W ten sposób ze zgłoszonych technik będą wybrane najlepsze i na ich podstawie rozpocznie się przygotowywanie norm, które prawdopodobnie zostaną przyjęte jako normy ISO/IEC (*International Standardization Organization / International Electrotechnical Committee*) oraz ewentualnie także jako zalecenia ITU (*International Telecommunication Union*).

BUDOWA EKSPERYMENTALNEGO SYSTEMU TELEWIZJI SWOBODNEGO PUNKTU WIDZENIA NA POLITECHNICE POZNAŃSKIEJ

Prowadzone w Katedrze Telekomunikacji Multimedialnej i Mikroelektroniki Politechniki Poznańskiej badania dotyczące telewizji swobodnego punktu widzenia koncentrują się na stosunkowo prostych systemach. W pracach przyjęto, że wdrażanie telewizji swobodnego punktu widzenia należy rozpocząć od systemów tanich w budowie i eksploatacji.

Kierując się wspomnianymi założeniami, zbudowano wstępną wersję prostego systemu wykorzystującego typowe kamery bez kosztownych zmechanizowanych obiektywów [3,8,9]. Dla



■ Rys. 6. Bezprzewodowy moduł kamerowy



■ Rys. 7. Nagrywanie sekwencji za pomocą bezprzewodowego systemu złożonego z 10 kamer rozmieszczonych na łuku

takich kamer w Katedrze opracowano i wyprodukowano moduły umożliwiające bezprzewodową synchronizację wszystkich kamer, a także bezprzewodowe sterowanie oraz zbieranie danych kontrolnych i kalibracyjnych z całego systemu kamer (rys. 6). W ten sposób powstał 10-kamerowy system, w którym kamery można rozmieszczać w dowolnych położeniach wokół sceny, a obraz można rejestrować bez konieczności łączenia modułów kamerowych kablami (rys. 7). Założono, że system rejestracji obrazu będzie obsługiwany przez jedną osobę, co zbliży koszty rejestracji programu w takim systemie do kosztów rejestracji klasycznych programów telewizyjnych za pomocą jednej lub dwóch kamer. Bardziej szczegółowe opisy modułów kamerowych można znaleźć na przykład w pracach [3] oraz [9]. Ponadto dostosowano algorytmy estymacji głębi oraz syntezy widoków do potrzeb systemów z dowolnie rozmieszczonymi kamerami, a następnie przygotowano odpowiednie oprogramowanie wzorcowe [10,11].

Wspomniany system wykorzystano do wytworzenia sekwencji testowych niezbędnych do dalszych badań nad telewizją swobodnego punktu widzenia. Podczas rejestracji kilku sekwencji, dokonywanej zarówno w pomieszczeniach, jak i na zewnątrz budynku, kamery ustawiano w przybliżeniu na łuku wokół sceny. Dla części wyprodukowanych sekwencji wyznaczono już także mapy głębi. Jedną z tych sekwencji, nazwaną *Poznan Blocks*, jest już używana przez grupę MPEG w planowych eksperymentach badawczych.

* * *

Ten krótki artykuł ma przekonywać, że telewizja swobodnego widzenia nie jest futurystyczną mrzonką, ale może być rzeczywistością w ciągu zaledwie paru lat. Przed praktycznym uruchomieniem systemów telewizji swobodnego punktu widzenia trzeba jeszcze rozwiązać pewne problemy badawcze. Ponieważ istnieją oparte

na doświadczeniach koncepcje rozwiązania tych problemów, można się spodziewać, że proste, a więc stosunkowo tanie w budowie i eksploatacji, systemy telewizji swobodnego punktu widzenia będą mogły być eksploatowane już za parę lat. Ponieważ dla specyficznych typów rejestrowanych zdarzeń również przygotowanie treści w takich systemach prawdopodobnie nie będzie znacznie droższe od klasycznych transmisji, to takie systemy mogą rzeczywiście stosunkowo szybko wejść na rynek. Z pewnością jednak rynek telewizji swobodnego punktu widzenia pozostanie tylko małą częścią całego rynku telewizyjnego.

Przygotowanie pracy było wsparte ze środków publicznych na badania naukowe.

LITERATURA

- [1] Tanimoto M., Tehrani M. P., Fujii T., Yendo T.: *Free-viewpoint TV*. IEEE Signal Processing Magazine, vol. 28, January 2011
- [2] Tanimoto M.: *FTV: Free-viewpoint Television*, Signal Processing: Image Communication, vol. 27, July 2012
- [3] Domański M., Łuczak A., Klimaszewski K.: *Telewizja swobodnego punktu widzenia*. Przegląd Telekomunikacyjny i Wiadomości Telekomunikacyjne, nr 2-3/2013, vol. LXXXII
- [4] Tech G., Wegner K., Chen Y., Hannuksela M., Boyce J.: *MV-HEVC Draft Text 8*, JCT-3V of ITU-T and ISO/IEC Doc. JTC3V-H1004, March 2014, Valencia, Spain
- [5] Domański M.: *Nowe technologie kompresji obrazu ruchomego dla nowych usług multimedialnych*, Przegląd Telekomunikacyjny i Wiadomości Telekomunikacyjne, nr 8-9/2012, vol. LXXXI
- [6] Tech G., Wegner K., Chen Y., Yea S.: *3D-HEVC Draft Text 4*, JCT-3V of ITU-T and ISO/IEC Doc. JTC3V-H1001, March 2014, Valencia, Spain
- [7] Tanimoto M., Senoh T., Naito S., Shimizu S., Horimai H., Domański M., Vetro A., Preda M., Mueller K.: *Proposal on a New Activity for the Third Phase of FTV*, Doc. ISO/IEC JTC 1/SC 29/WG 11 no. MPEG2012/M30229/M30232, Vienna, Austria, July – August 2013
- [8] Doc. ISO/IEC JTC 1/SC 29/WG 11 no. MPEG/N14552, FTV Seminar Report, July 2014
- [9] Domański M., Dziembowski A., Kuehn A., Kurc M., Łuczak A., Mieloch D., Siast J., Stankiewicz O., Wegner K.: *Experiments on Acquisition and Processing of Video for Free-Viewpoint Television*, 3DTV CON, Budapest, July 2014, to appear in IEEEExplore
- [10] Stankiewicz O., Wegner K., Tanimoto M., Domański M.: *Enhanced Depth Estimation Reference Software (DERS) for Free-viewpoint Television*, ISO/IEC JTC 1/SC 29/WG 11, Doc. MPEG M31518, Oct. 2013
- [11] Stankiewicz O., Wegner K., Tanimoto M., Domański M.: *Enhanced View Synthesis Reference Software (VSRS) for Free-viewpoint Television*, ISO/IEC JTC 1/SC 29/WG 11, Doc. M31520, Oct. 2013

PORTAL INFORMACJI TECHNICZNEJ

www.sigma-not.pl

największa baza publikacji on-line