

Widz wybiera punkt widzenia

MAREK DOMAŃSKI

Wydział Elektroniki i Telekomunikacji, Politechnika Poznańska
domanski@et.put.poznan.pl

Prof. dr hab. inż. Marek Domański jest kierownikiem Katedry Telekomunikacji Multimedialnej i Mikroelektroniki na Politechnice Poznańskiej. Stworzył w niej zespół specjalizujący się w badaniach dotyczących przetwarzania i kompresji obrazów ruchomych oraz innych sygnałów multimedialnych.

Prace dotyczące ruchomych obrazów swobodnego punktu widzenia (*free-viewpoint video*) prowadzone są w Katedrze Telekomunikacji Multimedialnej i Mikroelektroniki Politechniki Poznańskiej

W coraz większym stopniu wykorzystujemy obrazy do przekazywania danych za pomocą systemów technicznych, takich jak telewizja, internet czy płyty DVD i Blu-ray. Już teraz większość informacji przesyłanych w sieciach telekomunikacyjnych reprezentują ruchome obrazy. Ich rosnący udział w światowym ruchu telekomunikacyjnym jest spowodowany m.in. rozwojem techniki, która pozwala na gromadzenie, przetwarzanie i przesyłanie obrazów doskonale oddających naturalne sceny. Pojawienie się powszechnie dostępnych systemów na to pozwalających jest wynikiem bardzo rozległych i intensywnych badań naukowych dotyczących techniki obrazów. Fascynujący rozwój tych badań jest kontynuowany m.in. w celu stworzenia systemów pozwalających na przekazywanie obrazów, które będą jeszcze bardziej realistycznie niż obecnie odtwarzać rzeczywistość.

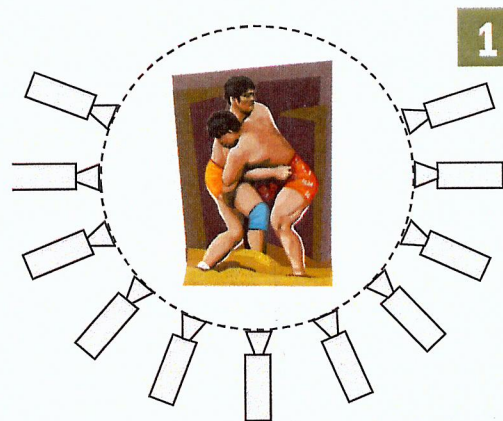
Badania prowadzone przez nasz zespół w Katedrze Telekomunikacji Multimedialnej i Mikroelektroniki dotyczą ruchomych obrazów swobodnego punktu widzenia (*free-viewpoint video*). Obrazy te charakteryzują się możliwością interakcji – widz może sam swobodnie obracać scenę, i to bez zatrzymywania ruchu postaci i innych obiektów. Namiastkę obrazu swobodnego punktu widzenia stanowi fragment filmu „Matrix”, w którym jedna z postaci została zatrzymana w trakcie skoku, a następnie była wirtualnie obracana wraz z całą sceną.

Celem naszych badań jest stworzenie praktycznych i możliwie tanich systemów, w których widz samodzielnie wybiera punkty widzenia i w każdym z nich ogląda wirtualny widok wyliczony z ruchomego obrazu swobodnego punktu widzenia, będącego odpowiednią reprezentacją 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

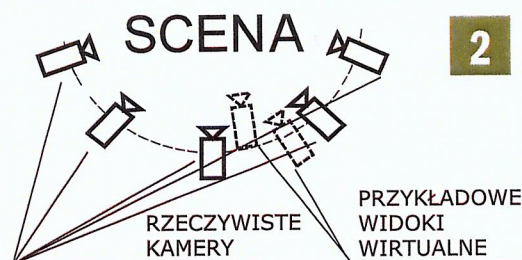
kolejnych punktów patrzenia na scenę widz nie jest ograniczony do widoków z rzeczywistych kamer, ale zazwyczaj ogląda widoki wirtualne (rys. 2) uzyskane w procesie syntezy. Dokonuje się jej na podstawie wspomnianej odpowiedniej reprezentacji sceny, która jest właśnie obrazem swobodnego punktu widzenia.

Ruchomy obraz swobodnego punktu widzenia i posługująca się nim telewizja swobodnego punktu widzenia mogą być stosowane do transmisji zawodów sportowych, na przykład zapasów, dżudo, walk sumo lub bokserskich. Sprawdzi się także w przypadku konkursów tanecznych, sztuk teatralnych czy rejestrowanych za pomocą kamer przedstawień dziecięcych, oglądanych przez rodziców najchętniej z kierunku, z którego najlepiej widać ich dziecko. Potencjalne zastosowania obejmują także różne instrukcje i materiały szkoleniowe, a nawet filmy fabularne i dokumentalne.

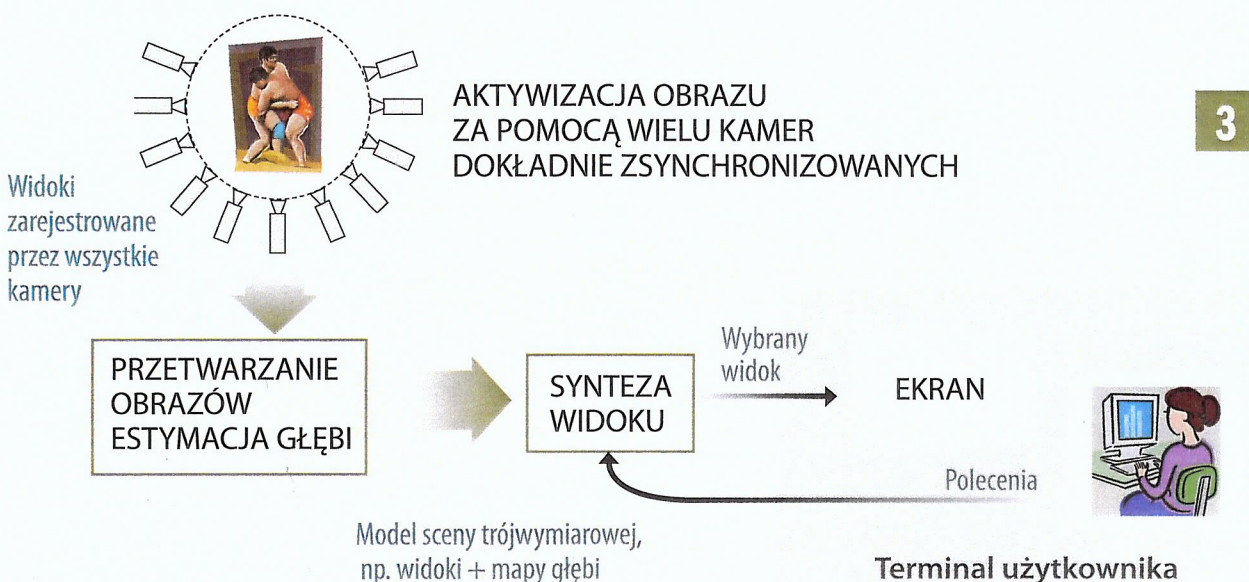
W tej chwili opracowujemy systemy telewizji swobodnego punktu widzenia, które będą mogły być realizowane z wykorzystaniem internetu. Użytkownik będzie oglądał obraz na ekranie komputera, tabletu lub inteligentnego telefonu, a wybór punktu widzenia przekazywał, poruszając palcami po płycie dodatkowej lub po ekranie dotykowym.



Kamery rozmieszczone wokół sceny rejestrują ruchomy obraz z różnych punktów widzenia



Widoki rzeczywiste rejestrowane rzeczywistymi kamerami oraz widoki wirtualne syntezowane programem komputerowym



Uproszczony schemat systemu telewizji swobodnego punktu widzenia

Ogólny schemat systemu telewizji swobodnego punktu widzenia pokazano na rysunku 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, jego złożonością a jakością obrazów wirtualnych wytwarzanych w terminalach użytkownika. Wydaje się, że w wielu zastosowaniach wykorzystanie tylko dziesięciu-kilkunastu kamer pozwoli na uzyskanie obrazu wirtualnego o wystarczająco dobrej jakości. Użycie tylko dziesięciu kamer wymaga rozłożenia wokół sceny około 40 kabli, co w wielu przypadkach zawodów sportowych i występów artystycznych nie jest możliwe ze względu na wygodę i bezpieczeństwo publiczności. Dlatego nasz zespół opracował nowatorskie moduły kamerowe umieszczone na statywie (rys. 4). Takie moduły pozwalają na bezprzewodowe sterowanie kamerami z jednego punktu, a także na ich precyzyjną synchronizację z dokładnością do około 2 mikrosekund. W ten sposób powstał praktyczny system precyzyjnej synchronizacji rejestracji obrazu za pomocą wielu kamer.

Najważniejsze prace badawcze są związane z algorytmami równoczesnego przetwarzania wielu obrazów ruchomych. Należy tu wymienić: metody dokładnego (z dokładnością do mikrometrów) wyznaczania położenia przetworników kamer umieszczonych na statywach, metody dokładnej kompensacji błędów układów optycznych obiektywów, metody kompensacji różnic charakterystyk czułości barwowej poszczególnych przetworników (nawet poszczególne kamery tego samego typu mają różne charakterystyki) oraz metody kompensacji różnic oświetlenia w polach widzenia poszczególnych kamer. Szczególnie trudne jest opracowanie efektywnej metody estymacji głębi, czyli odległości pomiędzy obiektami odpowiadającymi poszczególnym punktom obrazu a płaszczyzną obrazu. Wspomniane algorytmy zostały opisane w literaturze dla prostszych przypadków, a opracowanie metod właściwych dla kamer rozmieszczonych dowolnie wokół sceny jest niebanalnym problemem, nad którym pracujemy. System wielu kamer generuje bardzo dużą ilość danych, co wyma-

ga zastosowania nowych metod kompresji. W 2011 roku udało nam się w tym zakresie uzyskać duży sukces, gdyż opracowaliśmy oryginalną metodę jednoczesnej kompresji obrazów ruchomych pochodzących z wielu kamer rozmieszczonych stosunkowo gęsto. W wyniku badań w zagranicznych laboratoriach uznano, że nasza metoda kompresji jest jedną z dwóch najlepszych na świecie technik kompresji i jest znacznie lepsza od metod opracowanych przez wiele zespołów badawczych, w tym przez ośrodki badawcze wielkich koncernów (np. Sony, LG, Samsung, Ericsson). Dlatego nasza metoda została wykorzystana w procesie tworzenia nowej normy kompresji zwanej 3D-HEVC. Niestety, zarówno nasza metoda, jak i metody uzyskane przez inne zespoły nie są efektywne w kompresji obrazów ruchomych uzyskiwanych z wielu kamer rzadko rozmieszczonych wokół sceny. Dlatego obecnie pracujemy także nad metodami kompresji właściwymi dla takich zastosowań. Jest to duże wyzwanie, gdyż efektywne metody kompresji są bardzo skomplikowane – na przykład wspomniana już metoda poprzednio opracowana w naszym zespole była realizowana za pomocą oprogramowania liczącego około 100 000 linii kodu.

Liczymy, że będziemy mogli kontynuować wspomniane prace, gdyż otwierają one nowe obszary fascynujących zastosowań systemów technicznych dla udostępniania interaktywnych treści wizualnych.

