INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11
MPEG2018/M41992
January 2018, Gwangju, Korea

Source	**Poznań University of Technology**
Status	**Input**
Title	**Free-viewpoint television demonstration for sports events**
Author	**Marek Domański, Adrian Dziembowski, Tomasz Grajek, Adam Grzelka, Krzysztof Klimaszewski, Dawid Mieloch, Robert Ratajczak, Olgierd Stankiewicz, Jakub Siast, Jakub Stankowski, Krzysztof Wegner**

Abstract

The document is related to the research done recently on Free- Viewpoint Television (FTV) at Poznań University of Technology, Poznań, Poland. The abovementioned research is aimed at development of a simple free-viewpoint television system that provides the virtual horizontal navigation. The system is featured by video acquisition using sparsely distributed pairs of video cameras, the depth estimation from video, and the virtual video rendering on a server. The system is developed for sports broadcasts (like volleyball) and cultural events (e.g. theater performances), as well as for interactive courses and manuals. The service is intended to be provided as video on demand, and technically available through a webpage thus insuring compatibility with the vast majority of the internet-connected devices on the market, like smartphones, tablets and laptops. The document is related to a demonstration of the user FTV terminal and a simple rendering server.

1 Introduction

The virtual navigation is a functionality of future interactive video services where a user is able to navigate freely around a scene, also stepping towards and outwards the scene. The communications systems that provide such a functionality are often called free-viewpoint television (FTV) [1,2]. Here, we restrict our considerations to the virtual navigation on a horizontal plane. The work is aimed at interactive internet-based systems operating as video-on-demand services that are shared by many viewers who are able navigate independently from each other [2].

In an FTV system, a viewer watches a scene from virtual viewpoints on an arbitrary navigation trajectory. At each virtual viewpoint, the corresponding view has to be synthesized and made available at the receiver. In the system demonstrated, we use the centralized model of view synthesis, where the views requested by all viewers are synthesized in the rendering servers of the service provider [2] – [5]. Such a centralized model is sensitive to delays in the bidirectional server-to-terminal communications, similarly to networked gaming. Therefore, the server-to-terminal distance has to be limited in order to limit the network delay. The experience of the authors suggests that such systems are expected to work smoothly within agglomerations.

The system is developed for broadcasts of sport events like basketball, volleyball, fencing, judo, karate etc. The is also aimed at coverages of cultural performance like amateur or professional theater performances. Other potential applications include interactive courses and manuals (medical, cosmetics, dancing, technical etc. FTV may be also beneficial for professionals, e.g. sports team coaches who can watch replays from different points of view in order to analyze the performance of contestants.

The goal is to develop relatively simple and inexpensive system that may work with a limited number of cameras and with somewhat irregular camera locations.

2 Video acquisition

In the sports hall of Poznan University of Technology, a new multi-camera system was built for experiments with FTV and other video processing techologies. This multi-camera system is mounted on a special rig installed on the high of ca. 4.5 meter above the floor (Fig. 2).

In [2] and [6], it was shown both theoretically and experimentally that for highly occluded scenes, pairing the cameras yields better quality of the estimated depth maps than when using the same number of cameras but uniformly distributed around a scene. It was shown that the quality of the virtual views is better for paired cameras when the percentage of occlusions exceeds about 25% of the frame area. Therefore, in the multi-camera system, the cameras grouped in pairs (Figs. 1-3).

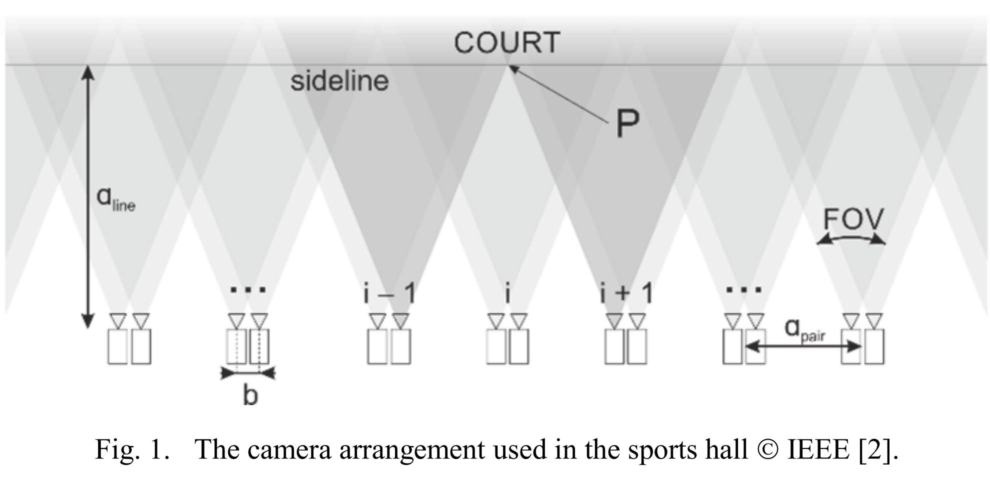

Fig. 1. The camera arrangement used in the sports hall © IEEE [2].

Fig. 2. Installation of the cameras in the sports hall.


)

Fig. 3. Multi-camera rig with a pair of cameras.

In order to select the camera type, one has to considered several requirements. In our case, we assumed full HD resolution and a high-sensitivity sensor with low noise as the illumination is moderate in the hall. Moreover, the cameras must be able to work in a synchronous mode, i.e. all cameras has to register a frame at the same time instant. On the other hand, the cameras have to be inexpensive in order to meet strict budget restrictions. Therefore popular consumer GoPro Hero4 Black cameras are used (cf. Fig. 3).

The cameras GoPro Hero4 Black are used with their standard wide-angle lenses, therefore strong geometrical correction is needed (cf. Fig. 4).

Fig. 4. One frame from a volleyball match sequence: before lens distortion removal (left) and after lens distortion removal (right).

3 Multiview test sequence “Poznan Volleyball2”

In the sports hall of Poznan University of Technology, a new multi-camera a new multiview test sequence “Poznan Volleyball2” was produced (cf. Figs 5 and 6).

Fig. 5. Selected original views obtained from the cameras.

Fig. 6. The same views after correction.

4 Rendering server and user terminal

The rendering server responds to the requests from a user and streams video for the requested viewpoint. This requires the video frames are synthesized according to the current viewpoint defined by a user. As the viewpoint may change rapidly for individual frames, the video is streamed with compression in the intraframe mode (MJPEG or AVC All Intra).

The user does not need to install special software on his/her device. The FTV service is available through a webpage. The content access is based on WebSocket API, which allows a web browser to initialize and maintain TCP connection with any server. WebSocket API provides general support for any type of connection and requires higher layer of transmission protocols for audio and video (Fig. 7 and 8).

The demonstration includes also the presentation of the service where a viewer uses his or her smartphone or tablet to watch the virtual video and to change the virtual observation point. For demonstration, the rendering server is implemented on a laptop that is capable for real-time view synthesis.

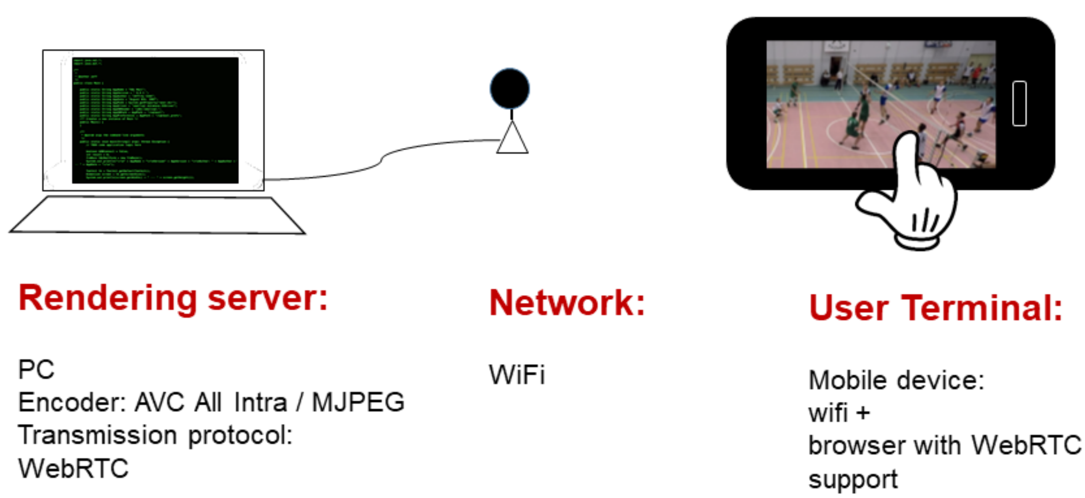

Fig. 5. Demonstration of the rendering sever and user terminals

Fig. 6. Selected views after virtual zooming as requested by a viewer.

5 Conclusion

All these results together with other results already published encourage us to believe that the development of usable FTV systems will be possible within the next very few years.

Acknowledgement

This work was supported by The National Centre for Research and Development, Poland under Project no. TANGO1/266710/NCBR/2015.

References

[1] M. Tanimoto, et al., "FTV for 3-D spatial communication", *Proc. IEEE*, vol. 100, pp. 905-917, 2012.

[2] O. Stankiewicz, M. Domański, A. Dziembowski, A. Grzelka, D. Mieloch, J. Samelak, "A free-viewpoint television system for horizontal virtual navigation," IEEE Trans. Multimedia, Early access, 2018.

[3] M. Domański, "Practicing free-viewpoint television: multiview video capture and processing," in: M. Tanimoto, T. Senoh, "FTV seminar report," *ISO/IEC JTC 1/SC 29/WG 11*, Doc. MPEG M34564, July 2014.

[4] M. Domański, A. Dziembowski, D. Mieloch, A. Łuczak, O. Stankiewicz, K. Wegner, „A practical approach to acquisition and processing of free viewpoint video," *Picture Coding Symposium PCS 2015*, Cairns, 2015, pp. 10-14, also in IEEExplore.

[5] J. Kim, J. Jang, D. Ho Kim, "Design of platform and packet structure for the free-viewpoint television", *18th IEEE International Symposium on Consumer Electronics*, Jeju Island, 2014.

[6] M. Domański, A. Dziembowski, A. Grzelka and D. Mieloch, "Optimization of camera positions for free-navigation applications," in *2016 International Conference on Signals and Electronic Systems* (ICSES), Kraków, 2016, pp. 118-123, also in IEEExplore.