# An experimental Free-view Television System

Marek Domański, Krzysztof Klimaszewski, Jacek Konieczny, Maciej Kurc, Adam Łuczak, Olgierd Stankiewicz and Krzysztof Wegner

Chair of Multimedia Telecommunications and Microelectronics,
Poznań University of Technology,
Polanka 3, 60-965 Poznań, Poland,
[domanski,kklima]@et.put.poznan.pl
[jkonieczny,mkurc,aluczak,ostank,kwegner]@multimedia.edu.pl

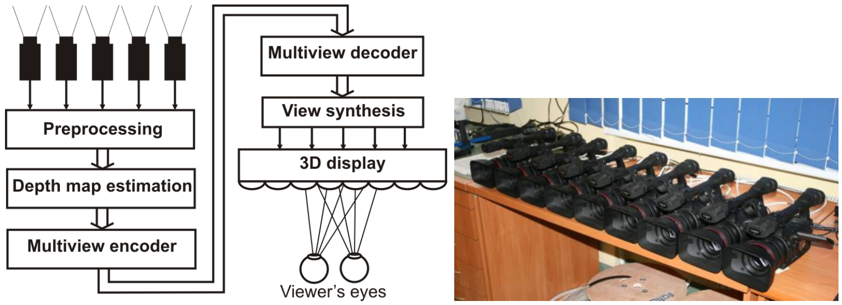
**Summary.** This paper presents an experimental prototype of Free-view television system, developed at Chair of Multimedia Telecommunications and Microelectronics at Poznań University of Technology. The general idea of the system, specific problems encountered during the implementation of its individual blocks, and also novel proposals have been highlighted. The results show that the quality achieved by proposed architecture is high both in objective and subjective manner.

## 1 Introduction

Nowadays, intensive research is undertaken to develop a new TV systems, which would offer not only higher quality of picture and sound, but also new functionality. One of desired features is ability to provide three dimensional content. An example of such a system is a Free-view Television (FTV) system, which enables the user to observe the scene from the view of his choice, or to experience illusion of depth on polarization or autostereoscopic displays. The results attained by other research laboratories around the world [8][9], show that still there are serious difficulties that disallow wide and common use of such systems. This work deals with this issue, by presenting an experimental prototype free-view television system, developed at the Chair of Multimedia Telecommunications and Microelectronics at Poznań University of Technology[1] (Fig 1a).

Our acquisition subsystem incorporates linearly positioned and equally spaced digital video cameras. The cameras deliver HD video content to preprocessing module, where specific views are rectified and color-balanced.

[1] This work was supported by the public founds as a research project in years 2007-2009. Prof. M. Domański is the recipeinet of Award MISTRZ form Foundation for Polish Science. K. Klimaszewski, J. Konieczny, M. Kurc, K. Wegner are the scholarship holders form Foundation for Polish Science.




a) b)

**Fig. 1.** a) Scheme of architecture b) Photography of acquisition subsystem

Color-balancing process reduces dissimilarities between characteristics of the cameras, while rectification process corrects positioning of the views. Then, multiview video is used to estimate depth of the scene. The resultant depth maps provide information about distances between the cameras and points in the scene. Together with the colour information, depth maps provide a 3D represenation of the scene. It is used in encoder, to increase compression efficiency. The decoder on the receiver side synthesizes new virtual views on demand. In spite of presentation technology (e.g. autostereoscopic or polarization display), two synthesized views are delivered to the user providing 3D experience.

## 2 System description

### 2.1 Acquisition

The main concept of the Free-view television system is to capture scene from multiple places simultaneously. Our acquisition system consists of nine high resolution cameras (Fig 1b), placed along straight line with constant distance between them. Linear arrangement of the cameras simplifies further processing. It is very important to synchronize all cameras in time in order to have frames captured at precisely the same moment. Among many cameras on the market we chose Canon XHG1 which is capable of recording video in Full-HD (1080p at 25fps), and what is more important, has dedicated line synchronization mechanism. Synchronization between all cameras is assured by the dedicated sync generator.

One of the major problem of capturing the video signal is the limited bandwidth to conventional hard drive. To overcome this problem we use special fast SSD hard drives arranged in RAID. We use 5 independent computers to capture and storage the video signal. Each captured video stream from the single camera is written directly to two SSD drives, connected in RAID for

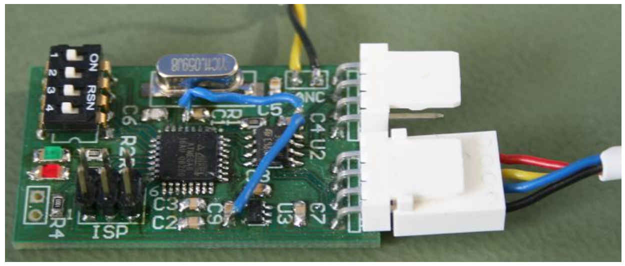


**Fig. 2.** Remote camera control unit

the best speed performance. Total capacity of our drives allows us to record up to 45 minutes of uncompressed multiview video sequences. The capture process can be controlled via internet which gives us great flexibility.

In order to control all cameras simultaneously we developed custom camera control system (Fig 2) which allows us to manipulate each of camera settings remotely. Control system is based on Sony LANC interface which is supported by our cameras.

### 2.2 Preprocessing

A very important step in preparation of multiview data is rectification. In majority of contemporary systems cameras are placed in line, facing the same direction with parallel optical axes. As it is impossible to adjust cameras' positions mechanically, a processing step is needed in order to simulate such camera arrangement. This step is called rectification.

Rectification is a process of adjusting optical axes and ensuring that the line that connects all cameras' optical centers is parallel to all images' horizontal axes [13]. In general, accurate rectification is possible only for two cameras. Rectification for more then two cameras has to be an approximation, as no general solution is possible. Rectifying transform can be calculated by using cameras' intrinsic and extrinsic parameters. It is also necessary to remove distortions caused by camera lens.

The first step in rectification is camera system calibration. In our system it is performed by using a checkerboard pattern. Program for processing checkerboard images is written with the help of OpenCV [14] library functions and the rectifying transform for camera system is calculated basing on results of calibration of all cameras. Development of the software is constant and aims towards automatic calibration and rectification with satisfying quality and accuracy.

Additional problem with cameras is white balance mismatch - every camera can, and usually does have, a different white balance setting. In order to avoid annoying effects in synthesized views and problems with processing, color correction is conducted. In this step, images from all cameras are subject to color adjustment based on color calibration for all cameras. Color calibration is performed by using dedicated color calibration checkerboard.

### 2.3 Depth map estimation

The aim of depth estimation is to deliver a 3D model of an analyzed scene. In general, it employs three steps of the estimation: matching of the images, optimization of the solution and post-processing. Our research concerns all of these steps.

Matching of the scene elements in neighbouring views is the first step in the estimation process. Typical approach is to compare blocks of pixels with respect to sums of absolute differences (SAD) between intensities of matched images. Our analysis of alternative methods show that it is better to use more advanced similarity measures [4]. For example, use of RANK.AD similarity measure provides gain of about 1.2dB (PSNR) in comparison to basic SAD measure.

As for optimization step, we use belief-propagation algorithm (BP) [2][3] with some novel improvements. Sought disparity map is modelled as a 2-dimensional Markov-field. BP algorithm employs iterative analysis of similarity map in order to find optimal solution. The optimization goal function is influenced by similarity map and smoothness constraints.

Each point of the scene (for which disparity is sought) is represented by a single node of the field. Nodes of the mesh communicate with others by message passing mechanism. Each message contains information about beliefs of node, specifically, all possible disparities for considered point. After each iteration of message passing, the self-beliefs of each node are updated, with respect to incoming messages. We propose that instead of classical rectangular 4-way mesh, messages are passed in 8 directions. The process is continued until convergence. Then, optimal disparity value for each node is selected and corresponding depth value is computed.

The precision of depth map estimation depends on resolution of processing mesh. A straight-forward method of obtaining better precision (like half-pixel precision) would be using higher resolution, which would result in tremendous increase of computational complexity. Thus, we use an original algorithm, which first estimates pixel-precise depth map with use of BP technique and attains quarter-pixel precision afterwards, with help of post-processing MLH technique [5][6].

### 2.4 Compresion

The amuount of data in multiview system is tremendous. This gives a rise to a demand of an effective way of compressing the data, and produced a challenge for compression algorithms. In our research we use MVC codec (an extension of existing SVC scalable codec) that is a proposed standard for multiview compression. We also take part in research conducted in MPEG committee towards establishing MVC as a standard.

The most important feature of MVC codec influencing compression efficiency is introduction of inter view reference frames. A simplified scheme of reference derivation is presented in Fig 3. It can be seen that one of the

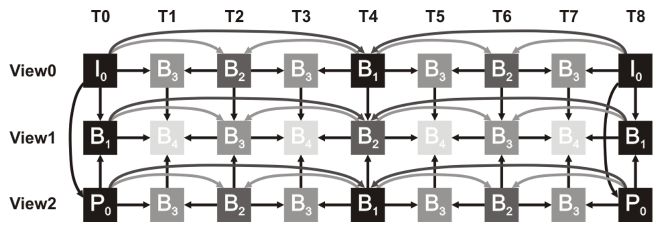



**Fig. 3.** MVC reference scheme

views (View 0) is encoded as in standard AVC coder, but the other ones can utilize some or all coded frames of a neighboring view in process of disparity compensation by including them to reference lists. It can be shown that MVC outperforms simulcast coding (i.e. coding of every view separately) by approximately 50% in bitstream size, providing the same quality in terms of PSNR [10]. As there are different coding structures possible to implement in MVC codec, it is important to adjust coding structure for optimal performance. During our research we have tested several improvements to MVC coding, like disparity prediction method [11], rotating reference macroblocks, then utilizing three dimensional vectors in motion/disparity compensation and recently joint compression of views and corresponding depth maps [12].

### 2.5 View synthesis

Ability of synthesizing views allows reduction of bandwidth required by the whole system, because only selected views have to be transmitted. It also provides additional functionality to synthesize new views in virtual camera positions. In general, the whole task can be described as a translation of the scene model into coordinates of desired camera location and then projection into 2D image (Fig 4).

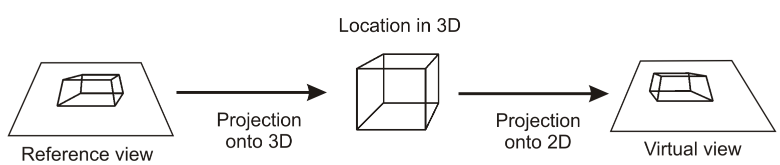



**Fig. 4.** View synthesis through rendering

In our algorithm, a view is synthesized basing on two neighboring views. Two separate virtual views are synthesized at the position of desired view and then these views are merged together to avoid problems with occlusions [7].

### 2.6 Presentation

Introduced system enables presentation of 3D video content which gives a viewer a unique 3D experience. During our research we have tested widely known presentation methods in order to choose the best one regarding the quality of the 3D illusion. We have tested three types of stereoscopic devices: autostereoscopic display, polarization display and polarized projectors. Despite the intensive development in recent years, autostereoscopic displays still produce weak 3D experience and offer very narrow viewing angle, and thus we turned to polarization based techniques. For small audience, we used polarization LCD display which offers the greatest quality to the spectators. In order to present the 3D content to a bigger audience, we developed a dedicated presentation technique using two HD projectors with circular polarizer and special screen preserving light polarization.

### 2.7 Quality evaluation

As described above, the 3D television system is composed of many novel elements. Evaluation of the quality of those elements is difficult because in most cases there are no references available. The most common method is to evaluate quality of the whole system in output-versus-input manner. The output of the system, i.e. synthesized views delivered to the user are compared with reference views acquired with real cameras, which are the input of the system (Fig 5).

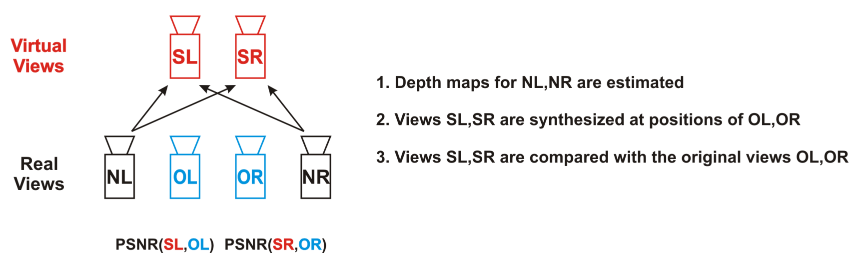



**Fig. 5.** Camera configuration and procedure for quality evaluation

## 3 Results

The experiments were conducted with the use of standard multiview MPEG test sequences and also with sequences acquired with our system. The results are presented for 'Book Arrival' (HHI [8]) sequence.

As can be seen from Fig 6, the results of view synthesis are of high quality (28÷35dB of PSNR), and were almost indistinguished by the subjects from the reference view during subjective evaluation. The analysis of influence of coding on overall quality of synthesized views also was performed with comparison to reference view. It can be concluded that depth map can be compressed more efficiently than a view and that for a wide range of QD parameter (quantization parameter for depth encoding) synthesized view quality is not influenced by changes in QD. Dashed horizontal line (Fig 6 right) corresponds to quality of view synthesied from uncompressed data. Curved dashed line shows quality of view synthesized with depth map compressed with QD resulting from the proposed formula (1).

$$QD = 49 - 0.005 \cdot (76 - QP)^2 \tag{1}$$

This formula allows to estimate a maximal QD parameter that does not cause degradation of synthesis quality below 0.5 dB compared to maximum synthesis quality for given QP. It therefore sets the limiting value of QD for a given QP.

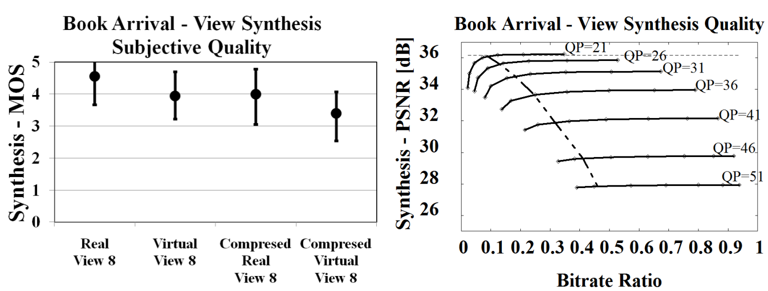



**Fig. 6.** System quality evaluation

## 4 Summary/Conclusions

A novel experimental Free-view Television system, developed at Chair of Multimedia Telecommunications and Microelectronics at Poznań University of Technology, has been presented. The architecture of the system and its components have been characterized, and novel proposals have been proposed. The presented results show that the presentation framework, which consists of whole system except coding module provides promising quality of rendered video. Coding experiments also show that proposed algorithm provides additional gains in comparison to standard MVC framework.

## References

1. "Description of Exploration Experiments in 3D Video", ISO/IEC JTC1/SC29/WG11, MPEG 2008/N9596, Antalya, Turket, January 2008
2. J.S. Yedidia, W.T. Freeman, Y. Weiss, "Undestanding Belief Propagation and Its Generalizations", Exploring Artificial Intelligence in the New Millennium, Chap. 8, pp. 239-236, January 2003 (Science Technology Books)
3. O. Stankiewicz, K. Wegner, "Depth Map Estimation Software version 3", ISO/IEC JTC1/SC29/WG11, MPEG 2008/M15540, Hannover, Germany, June 2008
4. K. Wegner, O. Stankiewicz, "Similiarity measures for depth estimation", 3DTV-Conference 2009 The True Vision Capture, Transmission and Display of 3D Video, Potsdam, Germany, 4-6 May 2009
5. O. Stankiewicz, K. Wegner, "Depth Map Estimation Software version 2", ISO/IEC JTC1/SC29/WG11 MPEG2008/M15338, Archamps, France, April 2008
6. O. Stankiewicz, K. Wegner, "Application of Middle Level Hypothesis algorithm for improvement of depth maps produced by Depth Estimation Reference Software", ISO/IEC JTC1/SC29/WG11 MPEG 2009/M16028, Lausanne, Switzerland, February 2009
7. M. Gotfryd, K. Wegner, M. Domaski,"View synthesis software and assessment of its performance", ISO/IEC JTC1/SC29/WG11, MPEG 2008/M15672, Hannover, Germany, June 2008
8. I. Feldmann, M. Mller, F. Zilly, R. Tanger, A. Smolic, P. Kauff, T. Wiegand, "HHI Test Material for 3D Video", ISO/IEC JTC1/SC29/WG11, MPEG 2008/M15413, Archamps, France, April 2008
9. Yo-Sung Ho, Eun-Kyung Lee, Cheon Lee "Multiview Video Test Sequence and Camera Parameters", ISO/IEC JTC1/SC29/WG11, MPEG 2008/M15419, Archamps, France, April 2008
10. K. Klimaszewski "Ocena efektywnoci kompresji sekwencji wielowidokowych", Krajowa Konferencja Radiokomunikacji Radiofonii i Telewizji 2008, Wrocaw, Poland, 9 - 11 April 2008
11. K. Klimaszewski "Disparity prediction in multiview coding", Pozna University of Technology Academic Journals - Electrical Engineering no 54, pp. 89-96. Pozna, Poland, 2007
12. K. Klimaszewski, K. Wegner, M. Domaski "Influence of distortions introduced by compression on quality of view synthesis in multiview systems", 3DTV-Conference 2009 The True Vision Capture, Transmission and Display of 3D Video, Potsdam, Germany, 4-6 May 2009
13. Yun-Siik Kang Cheon Lee Yo-Sung Ho, "An Efficient Rectification Algorithm for Multi-View Images in Parallel Camera Array", 3DTV Conference 2008 The True Vision Capture, Transmission and Display of 3D Video, 28-30 May 2008
14. OpenCV homepage `http://sourceforge.net/projects/opencvlibrary/`