Global 3D Technology Forum

Session 1

3D Device & Processing

3D Video Processing and Coding Techniques

Yo-Sung Ho
(Prof, Gwangju Institute of Science
and Technology, Korea)

3D Video Processing and Coding Techniques

Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)
E-mail: hoyo@gist.ac.kr

Keywords: 3D Video, multi-view image, depth map, view synthesis, depth image-based rendering

Abstract

In this paper, we introduce video processing techniques for 3D video. We describe several challenging issues of 3D video processing, such as multi-view image capturing, preprocessing, multi-view video-plus-depth coding, and intermediate view synthesis.

1. Introduction

In recent years, various multimedia services have become available and the demand for three-dimensional (3D) video is growing rapidly. 3D video is considered as a main theme for the next-generation broadcasting system supporting natural viewing experiences in the true three dimensions.

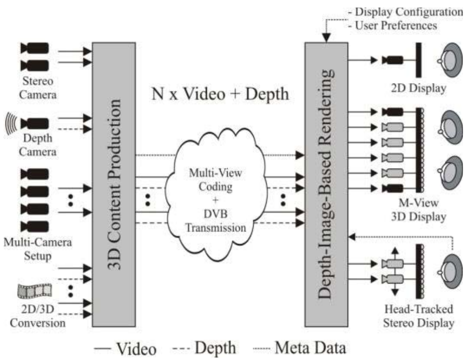

Fig.1. Three-dimensional video system

Fig. 1 shows the conceptual framework of the 3D video system [1]. The 3D video system includes several processes of 3D data acquisition, image processing/transmission, and rendering by means of 3D video data including multi-view color and depth images. We can generate the 3D video by utilizing cameras, such as stereoscopic cameras, depth cameras, or multi-view cameras.

3D video can be rendered by different types of displays, such as stereoscopic display, multi-view 3D display, head-tracked stereo display, or even 2D display. It can be compatible with the conventional 2D display by selecting one view by user preference.

In this paper, we introduce several challenging issues of 3D video processing and coding techniques. Specifically, 3D content creation, preprocessing, coding, and rendering techniques are explained in this paper.

2. 3D Video Generation

3D video, which contains color information and its corresponding depth map, can be obtained by means of several sensing devices. They are mainly classified into two categories: passive and active sensor-based method. The former calculates depth data indirectly by inferring the correlation of 2D images captured by two or more cameras. Stereo matching is a well-known process based on this method [2]. The latter, on the other hand, makes use of various types of sensors, such as laser, infrared ray (IR), or light pattern, to directly obtain depth information from the real 3D scene. Mostly, depth cameras, structured light, and 3D scanners are exploited in this method [3].

3. Preprocessing for 3D Video

3.1. Camera Calibration and Rectification

In order to extract intrinsic and extrinsic camera parameters, camera calibration has to be conducted. The intrinsic parameters describe the physical properties of the camera, and the extrinsic parameters are composed of rotation and translation of the camera. The approach using a grid pattern board is widely used in practice.

In many cases, we manually arrange the multiple cameras. This process induces geometric errors which are represented as the vertical mismatches and non-uniform disparities between adjacent views. Therefore, multi-view image rectification minimizing the geometric errors is an essential process to get high-quality multi-view images.

Transformation matrices for multi-view are calculated using the camera parameters and the virtual camera parameters [4]. The result of image rectification is shown in Fig. 2. The rectified multi-view image has the same vertical coordinates and uniform view interval among views.

Fig.2. Original and rectified multi-view images

3.2. Color Correction

Multi-view cameras cause a color inconsistency problem among views. This problem is unavoidable, although we use the same types of cameras and set them up equivalently.

Such a problem induces unnatural view change as well as performance degradation of further procedures including depth estimation and view interpolation.

Recently, various researches for color correction have been proposed. Among them, the correspondence-based algorithms take center stage, since they consider occlusion regions and be applied to any multi-view images [5]. Fig. 3 shows the result of color correction of multi-view images.

Fig.3. Original and color-corrected images

4. Compression of 3D Video

As an extended version of multi-view video coding, the 3D video coding compresses the multi-view video as well as their corresponding depth data. Since the depth data provides the geometrical information of a scene, a virtual viewpoint image can be generated by viewpoint shifting. However, the consequent problem is depth map coding.

Since the depth data has different characteristics from the color image, conventional coding techniques are not proper for depth coding. Therefore, we reduce the bitrate of depth coding by sharing the coding information, such as motion vector and block size [6]. In addition, by using the inter-view redundancies between views, the multi-view depth data can be compressed with the view synthesis prediction method [7].

One important objective of depth map in the 3D video system is supplementary data for virtual view generation. Therefore, the distortion due to the depth coding is relatively not important if the rendering quality is not changed. In this sense, a depth reconstruction filter is designed to rebuild the depth edges after decoding [8].

5. Virtual View Synthesis

Virtual view synthesis is a key technology to realize 3D video, and various view synthesis methods have been studied. The 3D warping technique is an appropriate method for view synthesis because it generates an arbitrary viewpoint image by using two forward and backward projection processes. Using depth information for reference images, we change the viewpoint of the reference image into the target viewpoint. During this process, holes can be generated due to the occlusion problem that is one of the most difficult issue in the view synthesis process.

To deal with the hole problem, we proposed a boundary noise removal method [9]. The boundary noise is generated by inaccurate depth estimation around depth discontinuity. It induces quite annoying boundary noises near the object boundary. We can remove the annoying noises by replacing them with texture information from the reference image. Using the method, we can generate high quality multi-view images for the 3D video system. Fig. 4 shows the result of the view synthesis without/with boundary noise removal.

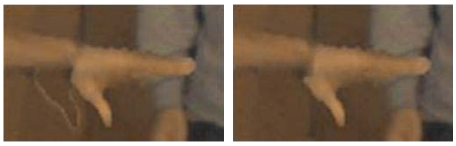

Fig.4. Results without/with boundary noise removal

6. Summary

In this paper, we have explained several challenging issues of 3D video processing techniques: 3D data creation, preprocessing, compression, and virtual view synthesis. Basically, we have used depth information for natural 3D view generation. The techniques that are introduced in this paper can be useful for various 3D multimedia applications.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2011-0030822).

References

1. ISO/IEC JTC1/SC29/WG11, "Applications and Requirements on FTV," N9466, Oct. 2007.
2. S. Lee, K. Oh, and Y. Ho, "Segment-based Multi-view Depth Map Estimation using Belief Propagation from Dense Multi-view Video," *3DTV Conference*, p. 193-196 (2008)
3. S. Lee, Y. Ho, "Real-time Stereo View Generation using Kinect Depth Camera," *APSIPA ASC*, p. SP3.12(1-4) (2011).
4. Y. Kang and Y. Ho, "An Efficient Image Rectification Method for Parallel Multi-Camera Arrangement," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 3, p. 1041-1048 (2011).
5. J. Jung and Y. Ho, "Color Correction Algorithm Based on Camera Characteristics for Multi-View Video Coding," *Signal, Image and Video Processing*, p. 1-12 (2012).
6. H. Oh and Y. Ho, "H.264-based Depth Map Sequence Coding using Motion Information of Corresponding Texture Video," *Lecture Notes in Computer Science*, vol. 4319, p. 898-907 (2006).
7. C. Lee, B. H. Choi, and Y. Ho, "Efficient Multi-view Depth Video Coding using Depth Synthesis Prediction," *Optical Engineering*, vol. 20, p. 1-14 (2011).
8. K. Oh, S. Yea, A. Vetro, and Y. Ho, "Depth Reconstruction Filter for Depth Coding," *Electronics Letters*, vol. 45, p. 305-306 (2009).
9. C. Lee, Y. Ho, "View Synthesis using Depth Map for 3D Video," *APSIPA ASC*, pp. 1-8 (2009).