Chapter 5
Technical Challenges of 3D Video Coding

Cheon Lee
Gwangju Institute of Science and Technology, Republic of Korea

Yo-Sung Ho
Gwangju Institute of Science and Technology, Republic of Korea

ABSTRACT

Three-dimensional video (3DV) is expected to be the next multimedia technology that provides depth impression of observed scenery with multi-view videos. In fact, studies on 3D video have a long history, heading back two hundred years; but recently, it has risen as the hottest issue due to rapid progresses of IT technologies. Particularly, 3D video systems are the most promising technology in multimedia area. An extension of typical stereoscopic imaging, realistic and natural 3D video technologies are currently under development. In this chapter, the authors describe overall technologies of 3D video systems from capturing to display, including coding standards. Mainly, the chapter focuses on the recent standardization activities by MPEG (moving picture experts group) associated with 3D video coding.

INTRODUCTION

A new generation for three-dimensional video (3DV) has arrived. Countless 3D films have been produced since the record-breaking success of *Avatar*by *J. Cameron* in 2009; many 3D items such as 3D-TV are released continuously reflecting the market demands. Simultaneously, worldwide major broadcasting channels are testing 3D-TV services while 3D contents are under production.

DOI: 10.4018/978-1-4666-2660-7.ch005

In fact, researches on 3D imaging have a long history from studies of stereopsis. It is the process in human visual system leading to the perception of depth impression from two slightly different projections of the world onto retinas of the two eyes; it was first described by *Charles Wheatstone* in 1838 (Wheatstone, 1838). He invented a stereoscope to show that stereopsis can be realized by creating illusion of depth from flat pictures that includes differences in horizontal disparity. Stereopsis became popular during *Victorian* times with the invention of prism stereoscope by *David*

Brewster. It is used for stereoscopic photography to generate stereograms. In 1970s, stereoscopic films were produced by anaglyph images; depth cues can be seen with red and blue films. In recent years, research efforts have been strengthened due to the advanced IT technologies, e.g., multi-view camera arrays, 3D scene representation, coding and transmission applied to capturing, calibration, rendering and 3D display (Benoit, 2008), (Smolic, 2007), (Lang, 2008), (Pesquet-Popescu, 2009), (Milani, 2011), (Maitre, 2008), (Maitre, 2010), (Kubota, 2007).

In terms of content production, the 3D content provider considers how to implement depth cues to the video data and how to transmit them efficiently. Since the amount of 3D data is greater than a single viewpoint video, an efficient video coding for 3D video contents is highly required; it is a core technology in 3D video service. Recently, the moving picture experts group (MPEG), which is working group to set standards for audio and video compression and transmission, started activities on 3D video systems since 2001 (Smolic, 2006). As a first step, they explored various technologies related to 3D video in the name of 3D audio-visual (3DAV). After exploring related technologies, the multi-view video coding (MVC) was developed for compressing multiple viewpoint video data under the team named JVT (joint video team) in 2007; it was the first phase of FTV (free-viewpoint TV) work. Right after, as a second phase of FTV work, the standardization activity on 3D video coding (3DVC) started in 2008. The primary goal of 3DVC is to define a data format and associated compression technologies to enable high-quality reconstruction of synthesized views for various types of 3D displays(ISO/IEC JTC1/SC29/WG11, 2011).

Since MPEG 3DVC group considers future 3D video systems as well as coding standards, it is easy to understand what core technologies are related to 3D video systems. In this chapter, we describe the general framework for 3D video systems and their elemental technologies in detail. We also explain the related video coding methods on 3D video coding according to the history of standardization activities in MPEG.

Technologies of 3D Video Systems

The 3D video system employs multi-view video to render various viewpoint images of a scene to provide realistic and natural depth impression to users. Since it uses more than two viewpoint images for 3D effects differently from the single viewpoint video, there are many constraints for achieving a successful system such as capturing conditions, distortions between views, generating supplementary data for rendering, and efficient 3D rendering. Therefore, each part of technologies may determine the feasibility of 3D video systems. In this section, we describe the generic 3D video system and its related technologies.

FRAMEWORK FOR 3D VIDEO SYSTEMS

As we mentioned above, the future 3D video systems involve multi-view videos to support advanced stereoscopic displays and auto-stereoscopic displays. The most popular data format for these systems is the multi-view video plus depth data (MVD), where the depth information is used for rendering an arbitrary viewpoint video using view synthesis algorithms. This data format allows generating any virtual viewpoint image; hence a certain type of display can render 3D scenery by selecting proper viewpoint images. In detail, if a certain display is atypical 2D display, it can render only one viewpoint image; in general, 3D video systems support backward compatibility. For the stereoscopic display, a pair of viewpoint images are selected and rendered on the display. In the case of the auto-stereoscopic display, more than two viewpoint images, e.g., 9-views, are rendered by generating sufficient viewpoint images with the reconstructed data.