Overview of State-of-the-Art 3D Video Coding Technologies

Yo-Sung Ho and Yuseok Song
Gwangju Institute of Science and Technology (GIST), Gwangju 500-712, Korea
E-mail: {hoys, ysong}@gist.ac.kr

Keywords: 3D video coding, 3D-HEVC

Abstract

3D video coding enables compression of multi-view color video as well as depth videos. Texture-depth and inter-view similarities can be exploited which are not present in normal 2D video coding. In this paper, we introduce the state-of-the-art 3D video coding technologies that are included in 3D video coding extension of high efficiency video coding (3D-HEVC). This coding standard is currently being developed by the joint collaborative team on 3D video coding extension (JCT-3V) and is expected to be finalized in 2015.

1. Introduction

In March 2011, at the 96th meeting of moving picture experts group (MPEG), the 3D video coding (3DV) group issued a call for proposals (CFP) on 3D video coding technology in two categories, advanced video coding (AVC) based and high efficiency video coding (HEVC) based [1]. Numerous proposals were evaluated in the 98th meeting. The most effective design at the time was determined as the test model for development. In the following meetings, many techniques have been assessed and adopted while keeping the main framework consistent. In July 2012, joint collaborative team on 3D video coding extension (JCT-3V) continued the activities of MPEG 3DV.

3D-HEVC aims compression of multi-view texture and depth videos [2]. While 2D video can exploit only spatial and temporal redundancies, additional aspects are considered in 3D video coding such as inter-view and texture-depth redundancies. Numerous 3D-HEVC-specific coding tools have been developed by taking such redundancies into account [3].

The basic coding structure of 3D-HEVC is depicted in Fig. 1. The base view of texture data is coded by HEVC. Temporal-wise, always the base view is coded before the dependent views. In the default configuration, for each view, texture data are coded first and then followed by depth coding. Thus, in regards to texture coding, any modification to depth coding only affects the texture coding of dependent views while the coded base view remains the same. In this paper, we present dependent view coding technologies in Section 2 and depth map coding technologies in Section 3.

2. Dependent View Coding Technologies

2.1. Disparity-compensated prediction

Disparity-compensated prediction (DCP) is an additional prediction, similar to the concept of motion-compensated prediction (MCP) [4]. While MCP refers to coded pictures in the same view, DCP refers to dependent views in the same access unit. Already coded pictures are referred in the prediction process. Fig. 2 illustrates how DCP is constructed and used.

2.2. Advanced residual prediction

Advanced residual prediction (ARP) exploits residual correlation between views. A residual predictor is generated by aligning the motion between views. Three weighting factors are introduced to compensate the quality differences between views. Fig. 3 represents the prediction structure of ARP. This technique is also extended for inter-view residual as well.

Fig. 1. Basic coding structure of 3D-HEVC.
Fig. 2. Disparity-compensated prediction.
Fig. 3. Advanced residual prediction.
2.3. Depth-based block partitioning
Depth-based block partitioning (DBBP) is a block partitioning mode for texture block. Based on binary segmentation in the collocated depth block, the texture block is partitioned by two regions. Each partition is motion compensated and then merged later. Fig. 4 exhibits derivation of DBBP.

3. Depth Map Coding Technologies

3.1. Depth modeling modes
In addition to 35 intra modes of HEVC, three depth modeling modes (DMM) are used for depth coding. Such modes are based on either wedgelet or contour partition. These are specifically designed to represent sharp edges efficiently, which are characteristics of depth maps. Fig. 5 represents the partitions of DMM. Two wedgelet modes and one contour mode are defined; they differ in partition derivation and signaling procedures.

2.2. Segment-wise DC coding
Segment-wise DC coding (SDC) is a residual coding tool. Residual values are derived and coded, without transform and quantization. SDC is applied to all depth intra modes including HEVC intra prediction modes.

3.3. Motion parameter inheritance
Motion parameter inheritance (MPI) is based on similar motions of texture and depth videos at the same viewpoint. Texture video motion vector have quarter-sample accuracy while depth video uses full-sample accuracy motion vectors. Thus, motion vectors are quantized to the nearest full-sample positions. Fig. 6 displays the derivation of MPI.

4. Summary
We have introduced some of the core technologies in the upcoming 3D video coding standard, 3D-HEVC. Various coding methods that take inter-view and depth map characteristics into account have been added on top of HEVC. For dependent view coding technologies, we presented disparity-compensated prediction, advanced residual prediction and depth-based block partitioning. In addition, depth map coding technologies including depth modeling modes, motion parameter inheritance were covered.

References