MPEG-DASH MPD for Tile-based Hybrid Stereoscopic 360-degree Video Streaming

Dongho You, Eunyoung Jeong and Dong Ho Kim
The Graduate School of Nano IT Design Fusion,
Seoul National University of Science and Technology (SeoulTech), Seoul 01811, South Korea
Email: {youdongho, jeunyoung, dongho.kim}@seoultech.ac.kr

***Abstract*—In this paper, we propose tile-based hybrid stereoscopic 360° video streaming service, in which low-quality 2D, high-quality 2D and 3D services can be provided by a single scalable HEVC (SHVC) en/decoder. In addition to this, its associated MPEG-DASH media presentation description (MPD) signaling is also presented as an example.**

I. Introduction

In the recent years, all markets related to virtual reality (VR) has grown tremendously, and global companies such as Facebook, Google, Samsung, LG and HTC compete to dominate this VR market. They commonly provide head-mounted display (HMD) devices that can display 360° video, as well as social media platforms providing the 360° video streaming service such as YouTube. In the platforms, it is common way to provide only monoscopic 360° video and display it on HMD device of users, but it can not provide 3D perception for the users. According to [1], stereoscopic 360° video makes the VR experience more immersive from the users rather than the monoscopic 360° video.

In this paper, therefore, we focus on an adaptive streaming service for the stereoscopic 360° video using HEVC tiling [2], and provide MPEG-DASH media presentation description (MPD) enabling hybrid video streaming in which low-quality 2D, high-quality 2D and 3D services can be provided in one scalable HEVC (SHVC) en/decoder [3]. In fact, studies of [4] and [5] also considered tile-based stereoscopic 360° video streaming using multi-view (MV) HEVC and SHVC, respectively, but these studies did not consider the proposed hybrid video streaming and its corresponding MPD signaling at all.

The remainder of this paper is organized as follows. Section II gives an overview of the proposed system considered for tile-based hybrid stereoscopic 360° video streaming. In section III, we provide examples of MPEG-DASH MPD enabling the proposed hybrid streaming service over HTTP. Finally, a short conclusion is presented in section IV.

II. Streaming System Overview

Fig. 1 shows an overview of the proposed tile-based hybrid stereoscopic 360° video streaming, in which both left and righ 360° videos are equirectangularly projected and separated into independent 16 tiles (4×4) for parallel processing. The 8K (7680×4320) right video is first down-sampled into low resolution video (e.g. QHD, 2560×1440), and encoded by a single SHVC encoder [3] with the original 8K left video. The

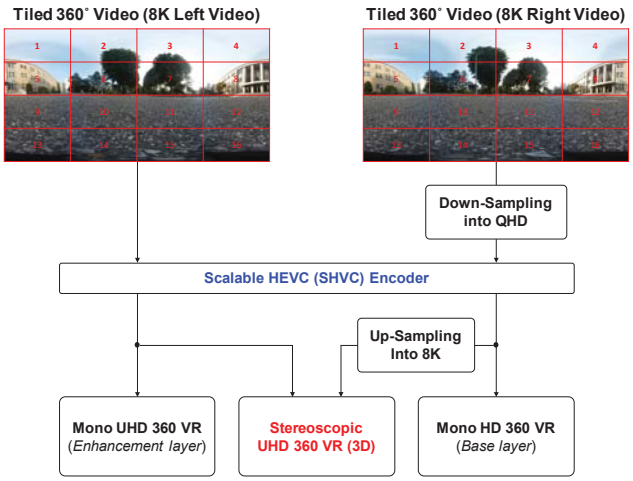

Fig. 1: Overview of tile-based hybrid stereoscopic 360-degree video streaming.

encoded low-quality (right) and high-quality (left) streams are called base and enhancement layers, respectively. The base layer is for a VR user within a bad network environment, whereas the enhancement layer is for a VR user within a good network environment. In addition to this, if a VR user is within a good network environment and capable of displaying stereoscopic 3D, both base and enhancement layers are combined to provide stereoscopic 360° videos for the user. Since the three services are provided by a single SHVC encoder, we call this method *"SHVC Tile-based Hybrid Stereoscopic 360° Video Streaming"*.

III. MPEG-DASH MPD Signaling

Fig. 2 shows an example of MPEG-DASH MPD [6] enabling the proposed hybrid video streaming service, in which some elements and attributes are omitted for its simplicity. Firstly, we use spatial relationship description (SRD), which describes the relationship between the tiles in the MPD. According to [7], the relationship shall be signaled using a `@schemeIdUri` with `urn:mpeg:dsh:srd:2014` and a `@value` attribute of a `SupplementalProperty` element at an `AdaptationSet`. Especially, the `@value` attribute must contain following parameters in order: `source_id`, `object_x`, `object_y`, `object_width`, `object_height`, `total_width`, `total_height` and

Fig. 2: Example of MPD signaling for SHVC tile-based bybrid stereoscopic 360-degree video streaming.

`spatial_set_id`, and these denote an identifier for the original video, the horizontal and vertical positions of the top-left corner of the associated tile, the width and height of the associated tile, the width and height of the original video, and an identifier for a group of tiles, respectively. All parameters are non-negative integer in decimal representation, and the last `spatial_set_id` is optional. Let us see the line 19 of Fig. 2 as an example. Since we consider a original 8K left 360° video separated into 16 tiles (i.e. the resolution of each tile is 1920×1080), the `@value` of the first tile of left video is described as (2, 0, 0, 1920, 1080, 7680, 4320), in which assuming that the `source_id` of left video is "2". Similarly, the `@value` of the second tile of left video can be also described as (2, 1920, 0, 1920, 1080, 7680, 4320).

In addition to the spatial relationship of tiles via SRD, we should describe the stereoscopic relationship between the left and right videos. According to [8], the stereoscopic relationship of the left and right videos shall be signaled using a `@schemeIdUri` with `http://dashif.org/guidelines/dash-atsc-videoposition` and a `@value` attribute of a `EssentialProperty` element at the `AdaptationSet` for "left video". Especially, the `@value` attribute must be the same as an `@id` attribute of the `AdaptationSet` for "right video". Let us see the line 20 of Fig. 2 as an example. Since the first tile of left video is a pair with the first tile of right video, the `@value` attribute of `EssentialProperty` element for the first tile of left video is described as "right_tile_1". It is the same as the `@id` of `AdaptationSet` for the first tile of right video (line 4). Similarly, the `@value` attribute of `EssentialProperty` element for the second tile of left video can be also described as "right_tile_2", which is the same as the `@id` of `AdaptationSet` for the second tile of right video (line 10).

Finally, the dependency between the base and enhancement layers is also described in the MPD. According to [8] and [9], if the original video is encoded scalably by H.264/SVC or H.265/SHVC, the dependency shall be signaled using a `@dependencyID` attribute at the `Representation` for "enhancement layer". Especially the `@dependencyID` attribute must be the same as an `@id` attribute of the `Representation` for "base layer". Let us see the line 21 of Fig. 2 as an example. Since the first tile of enhancement layer (left) cannot be decoded without the first tile of base layer (right), the `@dependencyID` attribute of `Representation` for the first tile of enhancement layer is described as "right_nHD_1". It is the same as the `@id` of `Representation` for the first tile of base layer (line 6). Similarly, the `@dependencyID` attribute of `Representation` for the second tile of enhancement layer can be also described as "right_nHD_2", which is the same as the `@id` of `Representation` for the second tile of base layer (line 12).

IV. Conclusion

In this paper, the tile-based hybrid stereoscopic 360° video streaming service is proposed to provide three services (i.e. low-quality 2D, high-quality 2D and 3D) simultaneously with one SHVC en/decoder. Furthermore, the associated MPD signaling is described in three categories: spatial relationship, stereoscopic pairing and dependency mapping. Finally, we expect that it provides not only efficient and adaptive service in accordance with users' network environments, but also virtual reality (VR) experience more immersive.

Acknowledgment

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. 2016-0-00144, Moving Free-viewpoint 360VR Immersive Media System Design and Component Technologies).

References

[1] M. Rowell, "Stereo vs Mono 360 video for VR,", Sept. 2015 [Online]. Available: http://360labs.net/blog/stereo-vs-mono-360-video-vr.

[2] K. Misra, A. Segall, M. Horowitz, S. Xu, A. Fuldseth and M. Zhou, "An overview of tiles in HEVC," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 6, pp. 969–977, Dec. 2013.

[3] J. M. Boyce, Y. Ye, J. Chen and A. K. Ramasubramonian, "Overview of SHVC: Scalable extensions of the high efficiency video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 20–34, Jan. 2016.

[4] R. Skupin, Y. Snchez and T. Schierl, "Compressed domain processing for stereoscopic tile based panorama streaming using MV-HEVC," in *Proc. IEEE Internaitonal Conference on Consumer Electronics - Berlin (ICCE-Berlin)*, Berlin, Germany, 06–09 Sept. 2015.

[5] Y. Snchez, R. Skupin and T. Schierl, "Video processing for panoramic streaming using HEVC and its scalable extensions," *Multimedia Tools and Applications*, vol. 76, no. 4, pp. 5631–5659, Feb. 2017.

[6] I. SSodagar, "The MPEG-DASH standard for multimedia streaming over the internet," *IEEE MultiMedia*, vol. 18, no. 4, pp. 62–67, Apr. 2011.

[7] O. A. Niamut, E. Thomas, L. D'Acunto, C. Concolato, F. Denoua and S. Y. Lim, "MPEG DASH SRD: spatial relationship description," in *Proc. International Conference on Multimedia Systems (MMSys)*, Klagenfurt, Austria, 10–13 May 2016.

[8] *Guidelines for Implementation: DASH-IF Interoperability Point for ATSC 3.0 (Version 1.0)*, DASH Industry Forum, Jan. 31, 2016. [Online]. Available: http://dashif.org/wp-content/uploads/2017/02/DASH-IF-IOP-for-ATSC3-0-v1.0.pdf.

[9] ISO/IEC JTC1/SC29/WG11, "Information technology - Dynamic adaptive streaming over HTTP (DASH) – Part 1: Media presentation description and segment formats", ISO/IEC 23009-1:2012, 2012.