TMIV를 위한 콘텐츠 적응형 패치 레벨 기하정보 스케일링

무하마드 우머 카클리, * 정준영, * 이광순, * 서정일

한국전자통신연구원

umar.isb@etri.re.kr, * jyj0120@etri.re.kr, * gslee@etri.re.kr, * seoji@etri.re.kr

Content Aware Patch Level Geometry Scaling for TMIV

Muhammad Umer Kakli, * Jun Young Jeong, * Gwangsoon Lee, * Jeongil Seo

Electronics and Telecommunications Research Institute, Daejeon, South Korea

Abstract

This paper presents a scheme to downscale the patches in geometry atlas based upon its contents for MPEG Immersive Video (MIV). In proposed scheme, the geometric contents of a patch is analyzed across the intra period, and a decision is made, whether to downscale a patch or not. The patches containing homogeneous information is downscaled by half, while the patches having salient information about the object such as edges are retained. Experimental results show that the bitrate saving is achieved by incorporating the patch level downscaling.

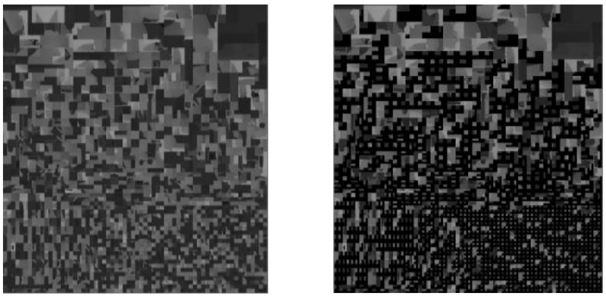
1. Introduction

Moving Pictures Expert Group (MPEG) is putting efforts to standardize the immersive video, which can be used in virtual reality (VR), augmented reality (AR), and mixed reality (MR). The undergoing standard is termed as MPEG Immersive Media (MIV). Broadly, MIV can be divided into three major modules, encoder, decoder and renderer. The MIV encoder packs the input views along with the geometry data into atlases consist of patches. It also generates the metadata to describe those atlases. These atlases are then encoded by the video codecs such as HEVC or VVenC. At the decoder side, first the video decoding is performed, after that, views are reconstructed by reversing the atlas packing process using transmitted metadata. Finally, views are rendered according to the viewport parameters.

During encoding process, once all the patches are written in geometry atlas, the atlas is downscaled by half of its original resolution to achieve the bitrate savings. We observe that some patches in geometric atlas consists of homogenous information, which can be downscaled before the downscaling of atlas. This patch level downscaling can help to further improve the bitrate savings. Therefore, in this paper, we present a scheme which determine the presence of a strong edge in a patch across the intra period. By utilizing this information, a decision is made, whether a patch can be downscaled or not. Additional metadata is also transmitted to ensure the upscaling of such patches at decoder side. The rest of the paper is organized as follows, Section 2 describe the proposed scheme in detail. The experimental details and results are discussed in Section 3, followed by conclusion is Section 4.

2. Proposed Scheme

A patch in a geometric atlas could have a horizontal or vertical edge that could be sharp or noisy. In order to determine the edge information in a patch, we first check difference of every pixel in a patch with its next five pixels along horizontal and vertical axis. If the difference of the current pixel with the next five pixels along horizontal or vertical axis is greater than 40% of current pixel value, we declare that as an edge pixel. We count the number of edge pixels across horizontal and vertical axis. The horizontal or vertical count greater than (1/3) of patch width or height, respectively, indicate the presence of a strong edge in patch. This procedure is repeated for all frames (in intra period) for the same patch. The patch is only downscaled, if more than (1/3) frames of the intra period for same patch indicate the absence of a strong edge. The patch is downscaled by half, if it satisfies the aforementioned conditions, and the metadata indicating the downscaling of a patch is also updated. The downscaling is performed using 2x2 max-pooling operation to preserve the thin objects and avoid the artifacts caused by in-between depth levels. Once the procedure is completed for all the patches, the whole atlas is downscaled by half of its original resolution. The atlas is then encoded with VVenC encoder. A fragment of atlas from Painter (SD) sequence with and without patch level scaling is shown in Fig. 1.



(a) Without patch level scaling (b) With patch level scaling

Fig. 1. Geometry atlas with and without patch level scaling

At the decoder side, once video decoding is performed, firstly, the geometric atlas is up-scaled using 2x2 nearest upscaling filter to its original resolution. Due to the max-pooling operation applied at the encoder side, some pixels are up-scaled as part of foreground, that actually belongs to the background. A color adaptive erosion filter is applied to suppress the artifacts caused by above-mentioned phenomenon. The foreground object shrinks in size due to the erosion, while color adaptivity ensure the correct spatial position of the depth edge. However, these edges could be noisy due to the non-linear characteristic of the erosion filter. Therefore, a contour smoothness filter is applied to smooth the edges in depth map. The same up-scaling procedure is again applied to the patches downscaled at encoder side. The rest of the decoding/rendering procedure is applied without any modification.

3. Experimentation and Results

The proposed scheme is implemented and tested with the source code provided by MPEG called TMIV (Test Model for MPEG Immersive Media) [1], under Common Testing Conditions (CTC) [2]. The intra period is set to 32, and the sequences consists of 97 frames. Table 1 show the result of proposed scheme for SD (Painter) and SJ (Kitchen) sequence. For SD sequence ~49% of the patches are downscaled, while for SJ sequence ~25% patches downscaled. This result in average bitrate savings of -2.9% and -2.2% for SD and SP sequence, respectively.

Sequence	Total No. of Patches	No. of Downscaled Patches	Bitrate Saving
Painter (SD)	3877	1903 (49.08 %)	-2.9 %
Kitchen (SJ)	5453	1398 (25.64 %)	-2.2 %

Table 1. Results of Proposed Scheme for SD & SJ sequence

4. Conclusion

A content aware patch level downscaling scheme is proposed for MIV. The scheme computes the strong edge information across the intra period for every patch. The patches consist of homogeneous region are downscaled, that helps to achieve the bitrate saving as shown in results.

Acknowledgments

This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2018-0-00207, Immersive Media Research Laboratory).

References

[1] B. Salahieh, C. Bachhuber, J. Jung, A. Dziembowski (Eds.), *Test Model 8 for MPEG Immersive Video*, ISO/IEC JTC 1/SC 29/WG 04 N 0050, January 2021, Online.

[2] J. Jung, B. Kroon (Eds.), *Common Test Conditions for MPEG Immersive Video*, ISO/IEC JTC 1/SC 29/WG 04 N 0051, January 2021, Online.