LETTER Efficient Methods of Inactive Regions Padding for Segmented Sphere Projection (SSP) of 360 Video

Yong-Uk YOON[†], Yong-Jo AHN^{††}, Donggyu SIM^{††}, Nonmembers, and Jae-Gon KIM^{†a)}, Member

SUMMARY In this letter, methods of inactive regions padding for Segmented Sphere Projection (SSP) of 360 videos are proposed. A 360 video is projected onto a 2D plane to be coded with diverse projection formats. Some projection formats have inactive regions in the converted 2D plane such as SSP. The inactive regions may cause visual artifacts as well as coding efficiency decrease due to discontinuous boundaries between active and inactive regions. In this letter, to improve coding efficiency and reduce visual artifacts, the inactive regions are padded by using two types of adjacent pixels in either rectangular-face or circle-face boundaries. By padding the inactive regions with the highly correlated adjacent pixels, the discontinuities between active and inactive regions are reduced. The experimental results show that, in terms of end-to-end Weighted to Spherically uniform PSNR (WS-PSNR), the proposed methods achieve 0.3% BD-rate reduction over the existing padding method for SSP. In addition, the visual artifacts along the borders between discontinuous faces are noticeably reduced. key words: JVET, 360 videos, projection, visual artifacts, padding

1. Introduction

In recent years, 360 videos have been attracting increasing attention as a new media type that provides an immersive experience. In Joint Video Exploration Team (JVET) that investigates new video coding techniques with capability beyond High Efficiency Video Coding (HEVC), 360 video is dealt with as one of content categories in the scope of standardization along with Standard Dynamic Range (SDR) and High Dynamic Range (HDR) videos [1]. The JVET has established a software package denoted as 360Lib for 360 video coding, format conversions, and evaluation.

In general, existing video codecs are designed considering conventional 2D videos. Therefore, in the workflow of 360 video coding [2], firstly a 360 video is projected onto the 2D plane with a projection format, such as Equirectangular Projection (ERP), Cubemap Projection (CMP), Sphere Segmented Projection (SSP), etc. A sphere is generally projected onto multiple faces to be packed into a 2D frame. For example, CMP projects the sphere onto a cube of six faces, or SSP projects the sphere onto three faces of one rectangular and two circles. Then, the faces are packed onto a 2D rectangular image to be coded with a frame packing method. In some projection methods, a converted 2D image has inac-

Manuscript received April 4, 2018.

Manuscript publicized August 20, 2018.

- ^{††}The authors are with Digital Insights Inc., Nowon-gu, Seoul, 01890 Korea.
 - a) E-mail: jgkim@kau.ac.kr (Corresponding author) DOI: 10.1587/transinf.2018EDL8069

tive regions which are simply padded with gray color as well as discontinuities between faces. Such inactive regions and discontinuities may degrade coding efficiency and cause visual artifacts in the generated viewport which includes face boundaries. To overcome these drawbacks, various methods of frame packing and padding have been proposed [3]–[5].

In this letter, to improve coding efficiency and reduce visual artifacts in SSP, efficient methods of inactive regions padding by using two types of adjacent pixels in either rectangular-face or circle-face boundaries are presented. By padding the inactive regions with the highly correlated adjacent pixels, the discontinuities between active and inactive regions are reduced.

2. Segmented Sphere Projection (SSP)

The basic SSP [6] is a typical projection method in which inactive regions exist. As shown in Fig. 1, the SSP segments a sphere into three tiles: north pole, equator and south pole.



Fig. 1 Segmented sphere projection [6].



Fig.2 A padding method for SSP [3]. (a) Padding areas, (b) the north and south pole segments, (c) the south pole and equatorial segments

Manuscript revised June 29, 2018.

[†]The authors are with Korea Aerospace University, Deogyanggu, Goyang-si, Gyeonggi-do, 10540 Korea.

The boundaries of the three segments are 45° N and 45° S latitude. The four corners of each pole segments have inactive samples and are filled with the default gray color.

The inactive regions may cause visual artifacts as well as coding efficiency decrease due to discontinuities between active and inactive regions. To efficiently overcome these drawbacks, a padding method [3] has been proposed in JVET and adopted in the last 360Lib software [2]. In [3], as shown in Fig. 2, padding is applied around boundaries of two pole segments and the boundary between the south pole segment and the equatorial segment. The pixels in the padding region are obtained by blending the pixels in the circle boundary and the inactive region boundary. The weights of two pixels to be blended are determined according to the distance between the position to be padded in the inactive region to each boundary pixel. This padding method noticeably reduces the visual artifact and also enhances the BDrate performance. However, since there are still inactive regions after padding, further improvement may be possible.

3. Proposed Inactive Regions Padding (IRP)

If the inactive regions are filled with the adjacent pixels in the given sphere video, the high correlation between active and inactive regions could be kept in the projected 2D video. As shown in Fig. 1, two circles of the SSP are in contact with the four rectangular faces at the latitude of $\pm 45^{\circ}$ in the given sphere video. Therefore, in this paper, the inactive regions are filled with the pixels obtained by two different ways of Test 1 and Test 2 as shown in Fig. 3.

Test 1 uses the average of the pixels at the boundary of the corresponding rectangular face adjacent with each inactive region. In this way, the inactive regions are padded with adjacent pixels of 3D sphere in which two circles are in con-



Fig. 3 IRP for SSP. (a) Test 1, (b) Test 2, (c) an example of IRP

tact with the four rectangular faces at the latitude of $\pm 45^{\circ}$. Test 2 uses the average of the pixels at boundaries of the corresponding quadrant circle faces adjacent with each inactive region in 2D. The correspondences between inactive regions and associated faces are indicated with the same colors and numbers in Fig. 3 (a) and (b), respectively.

Figure 3 (c) shows an example of SSP with the proposed padding method of Test 2. The pixels in the existing padding region are obtained by the same way of blending in [3]. In other words, the boundary pixels of the inactive region derived by the proposed padding method and the boundary pixels of the circle are blended according to the distance to each pixel to be blended at the position in the padded region.

As a result, the discontinuities appearing at the boundaries between inactive and active regions can be almost removed, which result in reducing visual artifacts and improving coding efficiency.

4. Experimental Results

The proposed padding methods were implemented on the 360Lib reference software [7] and the HM-16.16 [8] is used to evaluate the coding efficiency. The evaluation was done according to the JVET Common Test Conditions (CTC) and the evaluation procedures for 360 videos [9].

4.1 Objective Quality

The Test 1 and Test 2 were compared to the Padded ERP (PERP) [2] and the SSP [3] which have been adopted in the last 360Lib software. Table 1 and 2 show the average Bjontegaard-Delta rate (BD-rate) in the five test sequences

E2E S-PSNR E2E WS-PSNR Anchor: PERP Y U ν Y U v -6.5% 8K -11.8% -4.7% -6.6% -12.0% -4.7% SSP 6K -7.0% -2.8% -5.2% -7.0% -2.9% -5.3% [3] -10.9% -4.3% -4.3% Avg -6.3% -11.0% -6.3% -5.2% -6.8% 8K -12.1% -5.2% -6.8% -12.2% Test 1 6K -6.9% -2.8% -5.4% -6.9% -2.9% -5.5% -4.7% -11.1% -4.7% -6.6% -11.2% -6.6% Avg 8K -12.1% -5.2% -7.0% -12.2% -5.2% -7.0% -3.2% -5.4% -7.0% -3.3% -5.4% Test 2 6K -7.0% -11.1% -4.8% -4.8% -6.7% -6.7% -11.2% Avg

 Table 1
 Coding performance of the proposed methods (anchor: PERP)

 Table 2
 Coding performance of the proposed methods (anchor: SSP [3])

Anchor:	E2E S-PSNR			E2E WS-PSNR		
SSP	Y	U	V	Y	U	V
Test 1	-0.3%	-0.4%	-0.3%	-0.3%	-0.4%	-0.3%
Test 2	-0.3%	-0.5%	-0.5%	-0.3%	-0.5%	-0.5%

 Table 3
 Viewport PSNR of the proposed method in two dynamic viewports (anchor: SSP[3])

Anchor:	PSNR-DYN-VIEWPORT 1			PSNR-DYN-VIEWPORT 2		
SSP	Y	U	V	Y	U	V
8K	-0.3%	-0.5%	-0.15%	-0.3%	-0.25%	-0.45%
6K	0.0%	0.1%	-0.1%	0.0%	0.3%	-0.8%
Average	-0.2%	-0.4%	-0.1%	-0.3%	-0.1%	-0.5%



selected in JVET call for proposal [10] with various quality metrics. To provide different weights depending on the 2D position, the Weighted to Spherically uniform PSNR (WS-

Table 5The evil viewport of static viewport 2 centered at vertex 2



Test 2 of the proposed method



PSNR) and Spherical PSNR (S-PSNR) are used. In WS-PSNR, the distortion at each sample position is weighted by the area on the sphere covered by the given sample po-

sition. End-to-End (E2E) PSNR is calculated between the original ERP and the reconstructed ERP. Although the resolution is increased in the proposed method and [3] due to the padding, the resolution is kept in the calculation of E2E PSNR since the padded regions are cropped when the SSP is reconverted to ERP.

As a result, the proposed methods obtain, in terms of E2E S-PSNR and WS-PSNR, on average 0.3% BD-rate gain on luma component and up to on average 0.5% gain on chroma component compared to the existing SSP. Especially, Test 2 can achieve BD-rate gain without loss in 6K sequences and more chroma gain compared to Test 1. Test 2 could be more efficient because it uses more correlated pixels in the 2D frame to be encoded.

4.2 Subjective Quality

The evil viewports which include the boundaries between faces are provided for the quality comparison in terms of visual artifacts [11], [12]. The center of a decoded evil viewport is indicated by the predefined vertex. In CTC [9], two vertexes are predefined for each projection format. For SSP, as shown in Fig. 3(c), the vertexes are (yaw, pitch) = $(-180^\circ, 45^\circ)$ and $(90^\circ, -45^\circ)$: vertex 1 is along the picture border at discontinuous edge; vertex 2 is within the picture at the position with a maximum number of discontinuous edges. In addition, for each test sequence, two viewports (viewport 1 and viewport 2) are defined to indicate the center of 2D image to be coded [9]. Accordingly, for each sequence, there are four decoded evil viewports for subjective quality comparison: (1) static viewport 1 centered at vertex 1, (2) static viewport 1 centered at vertex 2, (3) static viewport 2 centered at vertex 1, and (4) static viewport 2 centered at vertex 2.

In the visual quality comparisons, we use the results of Test 2 in the viewports centered at vertex 2 since Test 2 has better gain than Test 1 and vertex 2 has more discontinuous edges between faces than vertex 1. In Table 4 and 5, the decoded evil viewports of static viewport 1 and 2 centered at vertex 2 for the sequence "*ChairliftRide*", which are coded with the constant QP of 37, are shown. In the basic SSP's result, visual artifacts of round and vertical lines are clearly visible at face edges. On the other hand, in the results of JVET-G0097 [3] and the proposed method, it is shown that these visual artifacts are significantly reduced. In particular, the proposed method visually slightly better in terms of texture details and color although it is not apparent. In addition, the BD-rate comparison with SSP [3] in two dynamic viewports [9] is shown in Table 3.

5. Conclusions

In this letter, efficient methods of inactive regions padding

for SSP of 360 videos are proposed to improve coding efficiency and reduce visual artifacts. The experimental results show on average 0.3%, 0.5%, and 0.5% BD-rate reduction on each component over the existing SSP, respectively. In addition, the visual artifacts reduction can be noticeable by inactive regions padding with the proposed methods.

Acknowledgments

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by Korean Government (MSIT) (2017-0-00352).

References

- J. Ohm and M. Wien, "Future video coding tools and developments beyond HEVC," Proc. ICIP 2017, Oct. 2017.
- [2] Y. Ye, E. Alshina, and J. Boyce, "Algorithm descriptions of projection format conversion and video quality metrics in 360Lib (Version 5)," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-H1004, Oct. 2017.
- [3] Y.-H. Lee, H.-C. Lin, J.-L. Lin, S.-K. Chang, and C.-C. Ju, "EE4: ERP/EAP-based segmented sphere projection with different padding sizes," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-G0097, July 2017.
- [4] A. Abbas, "AHG8: An update on RSP projection," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-H0056, Oct. 2017.
- [5] Y.-U. Yoon, D.-H. Park, J.-G. Kim, Y. Ahn, and D. Sim, "Methods of padding inactive regions for segmented sphere projection," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-I0023, Jan. 2018.
- [6] C. Zhang, Y. Lu, J. Li, and Z. Wen, "AHG8: Segmented sphere projection for 360-degree video," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-E0025, Jan. 2017.
- [7] 360Lib, [Online]. Available at: https://jvet.hhi.fraunhofer.de/svn/ svn_360Lib/
- [8] HM reference software, [Online]. Available at: https://hevc.hhi. fraunhofer.de/svn/svn_HEVCSoftware/
- [9] J. Boyce, E. Alshina, A. Abbas, and Y. Ye, "JVET common test conditions and evaluation procedures for 360° video," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-H1030, Oct. 2017.
- [10] A. Segall, V. Baroncini, J. Boyce, J. Chen, and T. Suzuki, "Joint call for proposals on video compression with capability beyond HEVC," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-H1002, Oct. 2017.
- [11] Z. Deng, L. Xu, and J. Boyce, "AHG8: Subjective test pilot study of 360° video projection/packing formats," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-F0083, March 2017.
- [12] J. Boyce, E. Alshina, and D. Zeng, "Subjective testing method for comparison of 360 video projection formats using HEVC," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-F1004, March 2017.