Performance Analysis on View Synthesis of 360 Video for Omnidirectional 6DoF

Hyun-Ho Kim, Ye-Jin Lee, and Jae-Gon Kim
Korea Aerospace University
{hhkim, yejin.lee}@kau.kr, jgkim@kau.ac.kr

Abstract

MPEG-I Visual group is actively working on enhancing immersive experiences with up to six degree of freedom (6DoF). In virtual space of omnidirectional 6DoF, which is defined as a case of degree of freedom providing 6DoF in a restricted area, looking at the scene from another viewpoint (another position in space) requires rendering additional viewpoints called virtual omnidirectional viewpoints. This paper presents the performance analysis on view synthesis, which is done as the exploration experiment (EE) in MPEG-I, from a set of 360 videos providing omnidirectional 6DoF in various ways with different distances, directions, and number of input views. In addition, we compared the subjective quality between synthesized images with one input view and two input views.

1. Introduction

Recently, with the increased commercial interests in deploying Virtual Reality (VR) applications, 360 video has become popular as a new media type giving immersive experiences. In order to enhance immersive experiences with up to six degrees of freedom (6DoF), MPEG-I Visual Group is actively working on it. In VR space of omnidirectional 6DoF which is considered in MPEG-I, looking at the scene from another viewpoint (another position in space) requires additional omnidirectional viewpoint rendering. Such additional rendered viewpoints are called virtual omnidirectional viewpoints. In this paper, we analyze the performance on view synthesis from a set of 360 videos in omnidirectional 6DoF in various ways with different distances, directions, and number of input views by using VSRS360 [1] and RVS [2]. In the experiments, the dataset of ClassroomVideo [3] was used, and the objective quality was evaluated by ERP WS-PSNR software provided by Zhejiang University [4].

2. Test Sequence

The details on the test sequence of ClassroomVideo provided by Philips are described in [3]. Figure 1 shows the frame 1 and 120 of the center view (texture image and depth map), and Figure 2 shows the arrangement of the viewpoints.

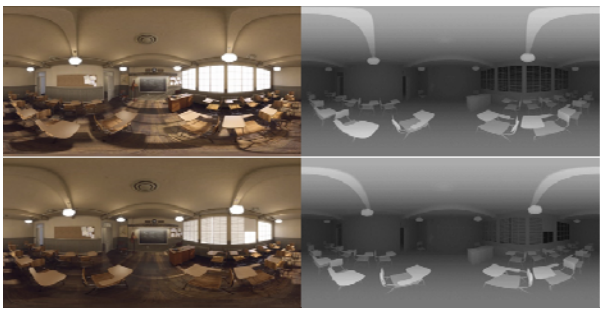

Figure 1. Frame 1 and 120 of the center view (texture image and depth map) [3]

3. Methodology

The goal of the experiments in this paper is to check the performance of the VSRS360 according to various synthesis conditions such the distance and direction between the synthesized view and the input view and the number of input views. For this, we performed three experiments as follows:

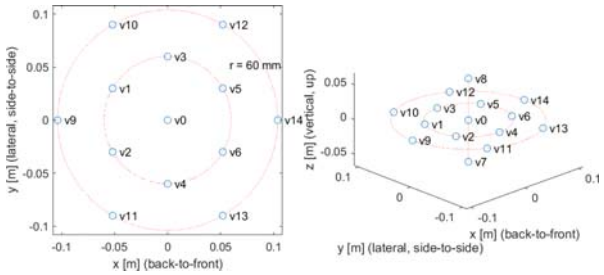

Figure 2. Viewpoint arrangement in ClassroomVideo sequence

- Calculate the average WS-PSNRs of the synthesized views at several positions with the same distance (60mm, 120mm) from the input view
- Calculate the average WS-PSNRs of the synthesized views in 6 directions (left, right, front-left, front-right, back-left, and back-right) at the distance of 60mm from the input view
- Calculate the average WS-PSNRs of synthesized views using two input views with diverse position combinations

The synthesis performance of each experiment measured by the ERP WS-PSNR software [4] is given in following sections.

Additionally, we compared the subjective quality of a synthesized view with a single input view and a synthesized view with two input views.

4. Synthesis Explorations

4.1 View synthesis at the same distance

In this experiment, we synthesize views at the same distances (60mm, 120mm) from the input view. Figure 3 and Table 1 show the synthesizable viewpoints at the distance of 60mm and 120mm from each input view with the ClassroomVideo sequence. We calculated the average WS-PSNR of the views which are synthesized at several positions by each input view. The viewpoints of v7 and v8 are excluded in the experiment. They may not be suitable for the evaluation of performance since they have only one synthesizable view at the distance of 60mm (v0) and nothing at a distance of 120mm.

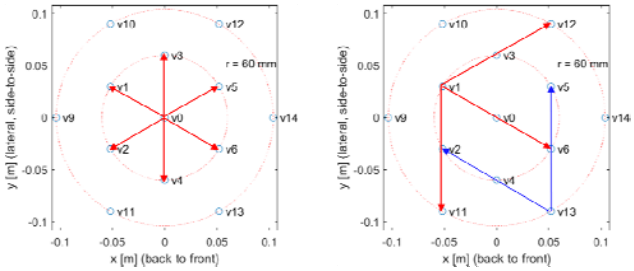

Figure 3. An example of viewpoints to be synthesized at the distance of 60mm and 120mm

Table 1. Viewpoint to be synthesized at the distance of 60mm and 120mm from each input view

input view	Synthesizable view at 60mm	Synthesizable view at 120mm
V0	V1,2,3,4,5,6	X
V1	V0,2,3,9,10	V6,11,12
V2	V1,4,9,11	V5,10,13
V3	V1,5,10,12	V4,9,14
V4	V0,2,6,11,13	V3,9,14
V5	V0,3,6,12,14	V2,10,13
V6	V0,4,5,13,14	V1,11,12
V9	V1,2	V3,4
V10	V1,3	V2,5
V11	V2,4	V1,6
V12	V3,5	V1,6
V13	V4,6	V2,5
V14	V5,6	V3,4

Table 2 shows the performance results of this experiment. There is no significant difference between the average WS-PSNR of the synthesized views from each input view at the same distance. However, there is a significant decrease in the average WS-PSNR when the distance between the input view and the synthesized view increases from 60mm to 120mm.

Table 2. Average WS-PSNR of the synthesized views at same distance from each input view

	WS-PSNR (VSRS360)		WS-PSNR (RVS 2.0)	
Input view	Distance of 60mm	Distance of 120mm	Distance of 60mm	Distance of 120mm
V0	22.66	X	31.44	x
V1	22.98	20.87	31.46	29
V2	22.94	20.92	31.51	29.56
V3	22.96	20.86	31.31	28.77
V4	22.88	20.89	31.5	29.85
V5	23	20.86	31.43	29.15
V6	22.95	20.89	31.51	29.77
V9	22.82	20.7	31.37	29
V10	23.04	20.94	31.15	28.72
V11	22.95	20.97	31.92	30.01
V12	23.03	20.96	31.12	28.88
V13	22.94	20.93	31.98	30.2
V14	22.84	20.74	31.57	29.44

4.2 View synthesis in various directions

In this experiment, we synthesize views in various directions using an input view at the 60mm distance. There are 6 synthesizable directions; left, right, front-left, front-right, back-left, and back-right. We calculated the average WS-PSNR of the views which are synthesized by the input views at several positions with the same direction.

Table 3 shows the performance results of this experiment. The average WS-PSNR of the synthesized views in left and right direction is a little bit higher than the other directions. This seems to be due to that the View Synthesis Reference Software (VSRS) [5], which is the base of the current VSRS360, is specialized in synthesizing left or right view from the input view rather than synthesizing front or rear view. On RVS results, the average WS-PSNR of the synthesized views in left direction is higher than the right direction.

Table 3. Average WS-PSNR of the synthesized views at each direction

Synthesis direction	Average WS-PSNR	
	VSRS360	RVS 2.0
Left	23.18	32.33
Right	23.09	30.98
Front_left	22.89	31.47
Front_right	22.86	30.99
Back_left	22.84	31.77
Back_right	22.81	31.26

4.3 View synthesis with two input views

In this experiment, we synthesize the view with two input views in various position combinations. Figure 4 shows the various way to synthesize the virtual view by using two input views with ClassroomVideo sequence. There are five combinations to synthesize the virtual view with two input views; left – right (yellow-colored line), front-left – front-right (red), back-left – back-right (blue), front-left – back-right (green), and back-left – front-right (purple).

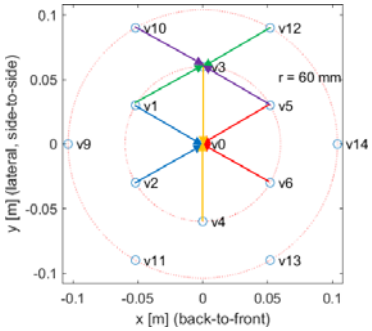

Figure 4. Various way to synthesize view with two input views

Table 4 shows the number of synthesizable view and the average WS-PSNR of the synthesized view at each combination of the position. It shows that the average WS-PSNR of the synthesized views which are placed in between the two input views are a little bit higher than others (yellow-/green-/purple-colored case in Figure 4).

Table 4. Number of synthesizable view and the average WS-PSNR of the views

Input views	Number of synthesizable view	Average WS-PSNR (Y)	
		VSRS360	RVS 2.0
2 front view (red)	6 (v0 from v5, v6 v1 from v0, v3 v2 from v0, v4 v3 from v5, v12 v4 from v6, v13 v9 from v1, v2)	23.53	32.77
2 side view (yellow)	5	23.89	34.32
2 rear view (blue)	6	23.42	32.33
Front_left + Back_right (green)	5	23.71	34.07
Front_right + Back_left (purple)	5	23.80	34.06

Also, compared to the results of previous experiments, using two input views gives better performance than using one input view when synthesizing virtual view.

4.4 Subjective quality

In this experiment, we compared the subjective quality of a synthesized view with a single input view and a synthesized view with two side input views. Figure 5 shows the first frame of the viewpoint v0 synthesized from the input view of v3, and Figure 6 shows the first frame of the viewpoint v0 synthesized from two input views of v3 and v4. Although both images have significant noise and the synthesis results are not satisfactory, if we compare the subjective quality of two synthesized images, the view synthesized with a single input view (Figure 5) looks better. However, when comparing the objective image quality (WS-PSNR), the synthesized view with two input views shows better performance.

Figure 5. v0 synthesized from v3
(WS-PSNR (Y): 23.10)

Figure 6.v0 synthesized from v3 and v4
(WS-PSNR (Y): 23.87)

5. Conclusions

This paper presents the experimental results on virtual view synthesis using VSRS360 in various ways with different distances, directions, and number of input views.

Based on the experimental results, the VSRS360 gives somewhat expected performances in the view synthesis below.

- Using a closer input view gives better performance than using a farther view input view
- Using a view at the same distance gives quite similar performance results regardless of the position of the input view
- Synthesizing left or right views gives better results than synthesizing front or back view
- Using two input views gives better results than synthesizing with one input view

In addition, the current version of VSRS360 can be improved with the following extensions when we consider diverse use cases of omnidirectional 6DoF.

- Extension to support other projection formats in addition to ERP in the view synthesis. Conversion of the ERP input view into other projection formats which have less distortion in the 2D projection image than ERP format (e.g. CMP).
- Improvement of synthesis performance when two views are used. In addition, use more input views to synthesize virtual view which are located in neighbor positions.

Acknowledgement

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by Korea Government (MSIT) (2017-0-00486).

References

[1] M. Domański et al., “Extended VSRS for 360degree video,” ISO/IEC JTC1/SC29 WG11, m41990, Jan. 2018.

[2] B. Kroon, “Proposed update of the RVS manual”, ISO/IEC JTC1/SC29 WG11, m44757, Oct. 2018.

[3] B. Kroon, “3DoF+ test sequence ClassroomVideo”, ISO/IEC JTC1/SC29 WG11, m42415, Apr. 2018.

[4] Y. Sun et al., “ERP-WSPSNR software manual,” ISO/IEC JTC1/SC29 WG11, m43816, July 2018.

[5] T. Senoh et al., “VSRS improvement,” ISO/IEC JTC1/SC29 WG11 m38979, Oct 2016.