Efficient Image Size Selection for MPEG Video-based Point Cloud Compression

Qiong Jia, M. K. Lee, Tianyu Dong, Kyu Tae Kim and Euee S. Jang
Hanyang University

Abstract

In this paper, we propose an efficient image size selection method for video-based point cloud compression. The current MPEG video-based point cloud compression reference encoding process configures a threshold on the size of images while converting point cloud data into images. Because the converted image is compressed and restored by the legacy video codec, the size of the image is one of the main components in influencing the compression efficiency. If the image size can be made smaller than the image size determined by the threshold, compression efficiency can be improved. Here, we studied how to improve the compression efficiency by selecting the best-fit image size generated during video-based point cloud compression. Experimental results show that the proposed method can reduce the encoding time by 6 percent without loss of coding performance compared to the test model 15.0 version of video-based point cloud encoder.

1. Introduction

With the increasing development of 3D technology, point cloud data is in the spotlight for emerging 3D immersive services. Dynamic point clouds, comparable to point cloud representation of video frames, have been identified as an important media in AR/VR/MR applications. MPEG has recently standardized Video-based Point Cloud Compression (V-PCC) standardin 2021 for compressing dynamic point cloud data.

V-PCC has some unique processes in its design: 1) 3D to 2D patch conversion, 2) patch image generation, and 3) video compression. The first two processes are in fact to support the third process: video compression. In short, V-PCC tries to convert the 3D property of geometry and attributes (such as color) into 2D patch video for video compression. The major compression comes from the video compression.

V-PCC adopted the codec-agnostic approach, which allows any codec to encode patch video frames. Quite a few efforts are made to prove the codec-agnostic approach both inside and outside of MPEG [1].

The compression efficiency of V-PCC depends on many factors including the choice of codec, the patch packing strategy, and the coding mode (e.g., intra/inter frame coding, size of GOP, motion search range, etc.). Once a video codec is chosen for V-PCC, it is necessary to pack the patches to generate video frames such that the resulting compression efficiency is maximized.

There are many efficient patch packing approaches proposed at the encoder side in MPEG [2]. However, the problem is not simple because the shape of a patch is very arbitrary and varies in size. This results in a combined set of patches with large and small patches per each frame. In MPEG test conditions, a typical number of patches in each frame is around 100. Moreover, the number of patches may vary frame to frame, which makes the patch packing a NP-hard optimization problem.

In this paper, we tackle the group of pictures(GOP)-level image size determination process for further compression efficiency. Since most video codecs used in V-PCC operate in inter-frame coding mode, it is necessary to maintain the same image size over the video frames in the same GOP. This requirement makes the image size to be the maximum image size that can house the largest atlas (or image frame with packed patches) image frame in the given GOP.

The proposed method changes the current pre-determined default image size determination method in MPEG V-PCC to a minimal image size selection method by collecting all atlas image sizes in the GOP. In our method, the image size is not set to a default size, but to zero. Then this image size grows by sequentially evaluating the image atlas size of each frame and adopting the bigger image size from the comparison of the current minimal size and the current image atlas size. This simple algorithm makes it possible to achieve the smaller GOP image size selection than the reference software encoder in V-PCC. We found the benefits of the proposed method not only in compression efficiency, but also in computational complexity of both the encoder and the decoder due to the smaller image resolution.

This paper is organized as follows. In Section 2, we describe the details of the proposed method. Then, we discuss the experimental results of the proposed method in Section 3. Finally, we conclude the paper in Section 4.

2. Proposed Method

2.1 Existing method in MPEG V-PCC

In MPEG V-PCC, the patch packing is implemented in three processes: patch packing strategy, patch packing using precedence order, and patch flexible orientation [3].

First, in the patch packing strategy, patches are placed to a defined minimum size block and specify the minimum distance between different blocks placed on a 2D grid. However, when there is not enough empty space to place the next patch, the height value of the grid will initially be doubled. Until the insertion of all patches is successful, then trim the height to the minimum required if it is not lower than the value originally specified in the encoder. In the process of patch packing using precedence order, the patches are rearranged by the priority indicated by a binary flag in the bitstream and allow larger patches to be decoded immediately [4]. In the patch flexible orientation process, there are 8 different orientation models employed: patches rotated by 0, 90, 180 and 270 degrees, and mirror images of all of these patches, to flexibly and compactly patch-pack the patches on 2D surface [5].

2.2 Proposed method

The existing method in V-PCC, although all the patches are placed in the 2D grid, not all the 2D grids are fully utilized. The actual space required by some patches is even less than the defined minimum size. Therefore, some space is wasted. We propose to initially set the minimum image height to 0, and then determine the height of the 2D grid according to the size of the large patch to be filled in first.

If the remaining space of the current 2D grid is not enough to fill the next patch, the height value of the grid will initially be doubled. If the height value is not changed until the insertion of all patches is successful, we will get a 2D grid smaller than the original defined minimum size.

The proposed method always guarantees the minimal size selection, whereas the existing method in V-PCC cannot cope with the case of the atlas image size smaller than the default size.

3. Results and Discussion

We tested HEVC-based V-PCC using the proposed method, and compared the coding efficiency under four conditions as follows: C2-all-intra (AI), CW-all-intra (AI), CW-low-delay (LD) and C2-inter random-access (RA). In the experiments, we used class a, b, and c sequences as test data. We used TMC2 version 15.0 in our Windows 10 development environment. The test PC has an i7-10700K CPU and a 32GB memory.

Under the C2-all-intra condition, Table 1 shows the geometry and color attribute encoding gains compared to the reference method. The results show a gain of 0.2 percent for geometry encoding and 0.5 percent for color attribute encoding. Table II shows the geometry and color attribute performance changes in the total bitstreams. The geometry coding gains in the total bitstream size are 0.4 percent and 0.3 percent for color attributes. Table 3 shows a 6 percent reduction in the coding process time.

Table 1 Geometry and color attribute ratio comparison of C2_ai.

	lossy geometry, lossy attributes [all intra]				
C2_ai	Geom. BD-TotGeomRate [%]		End-to-End BD-AttrRate [%]		
	D1	D2	Luma	Cb	Cr
Cat2-A	0.3%	0.3%	0.8%	0.8%	0.9%
Cat2-B	0.0%	0.0%	0.0%	0.0%	0.0%
Cat2-C	0.0%	0.0%	0.0%	0.0%	0.0%
Avg.	0.2%	0.2%	0.5%	0.5%	0.5%

Table 2 Geometry and color attribute to total rate comparison of C2_ai.

	lossy geometry, lossy attributes [all intra]				
C2_ai	Geom. BD-TotGeomRate [%]		End-to-End BD-AttrRate [%]		
	D1	D2	Luma	Cb	Cr
Cat2-A	0.7%	0.7%	0.6%	0.6%	0.6%
Cat2-B	0.0%	0.0%	0.0%	0.0%	0.0%
Cat2-C	0.0%	0.0%	0.0%	0.0%	0.0%
Avg.	0.4%	0.4%	0.3%	0.3%	0.3%

Table 3 Average geometric time of C2_ai.

C2_ai	Encoder Runtime		Decoder Runtime	
	Self	Child	Self	Child

Avg.	94%	88%	98%	90%

Under the CW-all-intra condition, Table 4 shows that there is no loss in geometry and color attribute encoding compared to the reference method. Table 5 shows a 2percent reduction in the encoding process time.

Table 4 Geometry and color attribute to total rate comparison of CW_ai.

CW_ai	lossless geometry, lossless attributes [all intra],bpip ratio [%]		
	Tot.Geom	Color	Total
Cat2-A	100%	100%	100%
Cat2-B	100%	100%	100%
Cat2-C	100%	100%	100%
Avg.	100%	100%	100%

Table 5 Average geometric time of CW_ai.

CW_ai	Encoder Runtime	Decoder Runtime	
	Self	Self	Child
Avg.	98%	99%	96%

Under the CW-low-delay condition, Table 6 shows that there is no loss in encoding geometry and color properties compared to the reference method. Table 7 shows a 2 percent reduction in the encoding process time.

Table 6 Geometry and color attribute to total rate comparison of CW_ld.

CW_ld	lossless geometry, lossless attributes [inter, low delay],bpip ratio [%]		
	Tot.Geom	Color	Total
Cat2-A	100%	100%	100%
Cat2-B	100%	100%	100%
Cat2-C	100%	100%	100%
Avg.	100%	100%	100%

Table 7 Average geometric time of CW_ld.

CW_ld	Encoder Runtime	Decoder Runtime	
	Self	Self	Child
Avg.	98%	99%	98%

Under the C2-all-intra condition, Table 8 shows the geometry and color attribute coding changes compared to the reference method. The results show a 3.8 percent loss in geometry encoding and a 6.2 percent loss in color attribute encoding. Table 9 shows the geometry and color attribute performance changes in the total bitstreams. The geometry coding losses in the total bitstream size are 4.3 percent and 4.3 percent for color attributes. Table 10 shows a 5 percent reduction in the coding process time.

Table 8 Geometry and color attribute ratio comparison of C2_ra.

C2_ra	lossy geometry, lossy attributes [inter, random access]				
	Geom. BD-TotGeomRate [%]		End-to-End BD-AttrRate [%]		
	D1	D2	Luma	Cb	Cr
Cat2-A	6.4%	5.9%	10.6%	11.5%	12.3%
Cat2-B	1.1%	0.9%	0.9%	1.6%	1.4%
Cat2-C	0.0%	0.0%	0.0%	0.0%	0.0%
Avg.	3.8%	3.5%	6.2%	6.8%	7.3%

Table 9 Geometry and color attribute to total rate comparison of C2_ra.

lossy geometry, lossy attributes [inter, random access]					
C2_ra	Geom. BD-TotGeomRate [%]		End-to-End BD-AttrRate [%]		
	D1	D2	Luma	Cb	Cr
Cat2-A	7.3%	6.7%	7.4%	8.0%	8.4%
Cat2-B	1.1%	0.9%	0.9%	1.3%	1.2%
Cat2-C	0.0%	0.0%	0.0%	0.0%	0.0%
Avg.	4.3%	4.0%	4.3%	4.8%	5.0%

Table 10 Average geometric time of C2_ra.

C2_ra	Encoder Runtime		Decoder Runtime	
	Self	Child	Self	Child
Avg.	95%	91%	98%	91%

4. Conclusion

The experimental results show that our proposed method reduces the encoding time by 6 percent under the C2-all-intra condition without loss of encoding performance. The encoding time was also reduced to varying degrees under the other three conditions.

ACKNOWLEDGMENT

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2020-0-00452, Development of Adaptive Viewer-centric Point cloud AR/VR(AVPA) Streaming Platform).

Reference

[1]. Tianyu, Dong, and Euee S. Jang. "Review on codec-agnostic approach for MPEG V-PCC." Proceedings of the Korean Society of Broadcast Engineers Conference. The Korean Institute of Broadcast and Media Engineers, 2020.

[2]. Graziosi, D., et al. "An overview of ongoing point cloud compression standardization activities: Video-based (V-PCC) and geometry-based (G-PCC)." APSIPA Transactions on Signal and Information Processing, 9, 2020.

[3]. 3DG, "V-PCC codec description," ISO/IEC JTC 1/ SC 29 WG07, N00100, [Online], Apr. 2021.

[4]. Danillo Graziosi, et al. "[V-PCC] (CE2.24-related New) Patch packing using precedence order," ISO/IEC JTC 1/SC 29 WG07, m48919, [Gothenburg], Jul. 2019.

[5]. Danillo Graziosi, "[PCC] TMC2 Patch Flexible Orientation," ISO/IEC JTC 1/SC 29 WG07, m43680, [Ljubljana], Jul. 2018.