MPEG-5 EVC Encoder Improvement for V-PCC

Tianyu Dong   Euee S. Jang
Hanyang University

Abstract

In this paper, we proposed an improved method on the picture order of coding (POC) of MPEG-5 Essential video Coding (EVC) encoder to support a short intra period for Video-based Point Cloud Compression (V-PCC). As a codec-agnostically designed standard, V-PCC claimed to be able to work with a lot of codecs. Current EVC test model software shows that the baseline profile could not provide appropriate POC calculation. The proposed method offers a solution to this POC-related problem and provides up to 44.6% coding grains for EVC based V-PCC.

1. Introduction

Immersive media applications have not been more popular before in our everyday lives. As a new standard result of MPEG-Immersive standardization activities, V-PCC is taking a lot of focus. People take a lot of work on its development and applications.

The V-PCC is designed as a codec agnostic standard, where it can co-work with a lot of video codecs. Now the V-PCC test model (TMC2) can work with MPEG-Advanced Video Coding (AVC), High Efficiency Video Coding (HEVC), and newly finished Versatile Video Coding (VVC) [1]. Our early research shows that MPEG-EVC can also be implemented with V-PCC [2].

The V-PCC compresses the dynamic point cloud frames in sequence occupancy map, geometry, and color attribute videos [3]. The encoder segments the 3D surfaces into 2D patches with depth information then generates images containing both near and far parts. Then these images are compressed into coded video bitstreams. The decoder follows the opposite process, and it uses decompressed video to reconstruct the point clouds.

To evaluate the performance of V-PCC, the working group (WG07) of MPEG extinguished a common test condition (CTC) for the experts to follow [4]. One of the conditions in the CTC is called lossy all-intra (AI), which means the V-PCC decoder can have frame-level free access inside a coded group of frames (GOF). This condition required the video codec to produce independent video frames for corresponding point cloud frames.

However, the EVC encoder did not support such a coding structure when using the baseline profile with a low-delay (LD) configuration [5]. When we use the baseline profile, the encoder cannot set near images as intra frames (I-frame) or set far images as prediction frames (P-frame) [6].

We propose a method to resolve such (POC) related problems, and the EVC encoder could encode videos with a shorter intra period (I-period). It provides independent I-frame access for the near frame and enhances the compression efficiency for the EVC based V-PCC.

This paper is organized into the following chapters. Section 2 reviews the background technology used in V-PCC, followed by the introduction of the proposed method in Section 3. Section 4 shows the result of the proposed method. And we review the impact of our practice in the end.

2. Background

The V-PCC encoder segments point clouds into patches in different projection directions. Then these patches fit into three different images, as shown in Fig. 1, which are occupancy map (a), geometry (b, c), and color attribute (d, e). Additionally, the geometry and color attribute have two layers from the near side (b, d) and far side (c, e) of the projection. These two images are very similar to each other.

Figure 1. Patch images generated from one frame of point cloud.

The TMC2 uses HEVC as their default video codec. The HEVC encoding configuration for AI condition code the near layer into I-frame while the far layer into P-frame. This kind of coding structure will have a higher coding efficiency.

However, our research on the EVC test model (ETM) shows that by giving a shorter I-period in the LD conditions, the encoder will have the wrong POC [6]. Which could not generate intra-prediction (IP) coded frames.

3. Proposed method

We use the *poc_derivation* function to generate the correct POC for the IP structured condition to select reference pictures. And the calling of the *poc_derivation* function is only needed when it is not an instantaneous decoder refresh (IDR), a bi-prediction frame, or a random-access condition. The decoder uses the *poc_derivation* function to calculate the POC values. We also called the *set_nalu* function before initializing the POC values shown in Fig. 2.

```
if ( !reference picture list ) /*baseline profile*/
{
   set_nalu()
   if ( nalu!=IDR && !B_Slice && lowdelay)
      evc_poc_derivation()
   evc_picman_refp_init()
}
else{ /*main profile initialization process*/}
if ( /*reference picture list is used*/) set_nalu()
```

Figure 2. Proposed code modification.

We implemented the proposed method into ETM. It does not change behavior for encoding the EVC CTC test data, and ETM can be used in the V-PCC.

4. Result

We tested the EVC based V-PCC using our proposed method to compare the coding efficiency in the AI conditions. We use the video-level AI coding method as a reference, and the test data are class a and b sequences. In the experiment, we use ETM version 7.1 and TMC2 version 12.0, in the development environment Windows 10 with Visual Studio 2017. The hardware is an AMD 5800X 4.8GHz CPU with 32GB DDR4 3200MHz memory.

Table I shows the geometry and color attribute coding gain compared to the reference method. The result shows 32% gains in the geometry coding and 45% gains in the color attribute coding.

Table I. Geometry and color attribute ratio comparison.

	Geom. BD Total Geom.Rate [%]		End-to-End BD Attribute Rate [%]		
	D1	D2	Luma	Cb	Cr
Cat2-A	-32.1	-22.7	-44.6	-52.6	-52.2
Cat2-B	-31.6	-44.4	-46.5	-53.5	-53.5
Avg.	-32.0	-27.1	-45.0	-52.8	-52.4

Table II shows the geometry and color attribute performance changes in the total bitstreams. The geometry coding gains in the total bitstream size are 21% and 44.6% for color attribute. Table III shows that the proposed method has reduced 19% video encoding time.

Table II. Geometry and color attribute to total rate comparison.

	Geom. BD Total Rate [%]		End-to-End BD Total Rate [%]		
	D1	D2	Luma	Cb	Cr
Cat2-A	-21.4	-1.7	-44.3	-50.0	-49.6
Cat2-B	-19.5	-43.2	-45.6	-50.7	-50.7
Avg.	-21.0	-10.0	-44.6	-50.1	-49.8

Table III. Average geometric time

	Encoder Runtime		Decoder Runtime	
	Self	Child	Self	Child
Avg.	99%	81%	82%	74%

5. Conclusion

The result shows that the proposed method reduces the bit size of the video stream without significant quality affection. It also provides frame-level accessibility. It could enhance the development of EVC based V-PCC.

Acknowledgments

This work was partly 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, 50%) and National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1F1A1061737, Basic Research of Pre- and Post-Processing Technologies of Point Cloud Compression, 50%)

References

[1]. 3DG, "V-PCC Test model 2," [Online], http://mpegx.int-evry.fr/software/MPEG/PCC/TM/mpeg-pcc-tmc2.

[2]. T. Dong, E. Jang, "Review on codec-agnostic approach for MPEG V-PCC," 2020 The Korean Institute of Broadcast and Media Engineers Summer Conference, July 2020.

[3]. 3DG, "V-PCC Codec Description VPCC v12 Algorithm Description," ISO/IEC JTC 1/SC 29/WG 11, N00012, [Online], Oct. 2020.

[4]. 3DG, "Common Test Conditions for V-PCC," ISO/IEC JTC 1/SC 29 WG07, N00038, [Online], Oct. 2020.

[5]. EVC Ahg, "EVC test model," [Online], http://mpegx.int-evry.fr/software/MPEG/Video/EVC/ETM.

[6]. T. Dong, K. Kim, E. Jang, "EVC encoder improvement on baseline coding order refinement," ISO/IEC JTC 1/SC 29/WG04, m55968, [Online], Jan. 2021.