Performance Evaluation of the Codec Agnostic Approach in MPEG-I Video-based Point Cloud Compression

Tianyu Dong, Kyutae Kim, and Euee S. Jang, Senior Member, IEEE

Hanyang University Seoul 04763 South Korea

Corresponding author: Euee S. Jang (e-mail: esjang@hanyang.ac.kr).

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) funded by the Korean government (MSIT) under Grant 2020-0-00452 (Development of Adaptive Viewer-centric Point Cloud AR/VR (AVPA) Streaming Platform).

ABSTRACT In this study, we evaluated the codec agnostic approach of video-based point cloud compression (V-PCC) by applying several video codecs to V-PCC. The main concept of V-PCC is to use a video codec to compress the 2D patch images generated from the 3D point cloud. As a new immersive media standard of the Moving Picture Experts Group (MPEG), V-PCC is designed to support the codec agnostic approach, which can be employed to compress point cloud data by using any video codec. At present, the V-PCC reference software is designed by using MPEG High-Efficiency Video Coding. We extended the evaluation of the video codec applicability for PCC with well-known MPEG video coding standards, such as Advanced Video Coding, Essential Video Coding, and Versatile Video Coding. During the evaluation, we identified several key strategies for applying a video codec to V-PCC to maximize compression efficiency or computational complexity. Furthermore, the coding efficiency and time complexity of each codec were tested. The evaluation tests revealed that V-PCC supports the codec agnostic approach and that the performance of the video codec has a positive correlation with the V-PCC final coding efficiency. Reviewing these key strategies would help to develop V-PCC with different video codecs based on their profiles and levels.

INDEX TERMS Computer graphics, point cloud compression, video codecs.

I. INTRODUCTION

As an immersive medium, point cloud data have been used in virtual reality (VR), augmented reality (AR), and mixed reality (MR). Consumer electronics or autonomous vehicles implemented in VR, AR, and MR applications use point cloud data as an immersive 3D media representation. Some of these electronic devices with time-of-flight (TOF) cameras can produce a large number of points. For a TOF camera with a resolution of 640 × 480 [1], the device generates 1.09 GB colorized points per second. These large amounts of data require an efficient compression method for storage and transmission. The Moving Picture Experts Group (MPEG) started standardizing point cloud data compression in 2017. This led to the development of the ISO/IEC 23090-part 5 video-based point cloud compression (V-PCC) [2], which can efficiently compress dynamic object point clouds.

In V-PCC, a patch generation process is adopted to generate three types of 2D images from 3D point cloud data: occupancy map (OMAP), geometry, and attribute [3]. V-PCC is designed to work with 2D video compression tools such that 2D images (after 3D to 2D transformation) can be compressed by using legacy video codecs, which are already implemented in billions of digital devices. As an important application of V-PCC, the codec is codec agnostic to increase the usability of V-PCC combined with different legacy codecs whenever possible.

There are few implementation examples showing that advanced video coding (AVC), versatile video coding (VVC), essential video coding (EVC), and other open-source video codecs can be used in V-PCC [4] – [7]. In a previous study, VP9, x264, and x265 contained in the FFmpeg package were shown to work together with V-PCC, and different codecs were used to compress different 2D components [4]. In another AVC-related study on V-PCC, the performance of the anchor and all-intra coding structures was evaluated [5]. An experiment based on the new VVC codec showed that VVC-based V-PCC achieved a coding...
gain similar to that for 2D videos [6]. Studies based on EVC indicate that the lossy compression video codec would significantly change the quality of V-PCC, causing all loss frames [7]. However, problems related to implementation have not been systematically discussed. In our previous research based on legacy image codec, we showed that JPEG can be used to compress geometry and attribute images with the all-intra coding structure [8]. We also observed that the video codec performance had a distinct impact on the V-PCC codec based on the research on the EVC baseline profile [8]. However, in these previous studies, the codec performance and computational complexity were not compared. Design-related strategies in which different video codec profiles and levels could be a limitation also need to be further discussed.

In this study, we extended our evaluation to three codecs: EVC, AVC, and VVC. Furthermore, we tested all-intra, low-delay, and random-access video coding structures. During the evaluation, we determined how to appropriately deal with pixel formats, coding structure accessibility, and bitrate assignment on geometry and attributes. We also discussed other problems, such as lossy occupancy settings.

The remainder of this paper is organized as follows. Section II introduces the basic structure of the V-PCC codec and the four video codecs. Section III presents our strategies for different implementations. The implementation results are analyzed in Section IV. The conclusion and future research plans are summarized in Section V.

II. BACKGROUND

A. V-PCC

The V-PCC encoding process consists of patch generation, patch packaging, video compression, and bitstream merging, as shown in Fig. 1. According to the algorithm description of V-PCC [3], the 3D surface of one object can be projected into one of the 10 planes that are predefined in the 3D coordinate system. These patches are presented in the form of three images: OMAP, geometry, and attribute. Then, V-PCC uses video encoders to compress the images. Finally, the encoder merges all the video bitstreams and auxiliary information as a V-PCC bitstream.

The images produced by the patch generation are shown in Fig. 2. These images include an OMAP, a geometry image, and a color attribute image. The occupancy image shows the boundary of the 3D surface, and the geometry shows the depth of the patch surface. The attribute images store the corresponding color information. It should be noted that the geometry and color attribute images, unlike the OMAP, are in the formation of a near layer and a far layer of the patch images. The near and far layers are in odd and even ordered images, respectively.

FIGURE 2. Point cloud rendering image and its intermediate images.

The size of these generated images, taken as examples, is $2560 \times 1856$. These images are stored in the YUV color space. In the OMAP, bit “1” represents an occupied signal and “0” represents an unoccupied signal. In the geometry images, the depth of these surfaces is represented in eight bits. The occupancy and geometry signals are stored in the Y-plane of the YUV color space. For the color attributes, the color data in the RGB color space are first converted into those in the RGB444-16bit intermediate and then converted to those in YUV420 for video codec compression.

B. AVC, High-efficiency Video Coding (HEVC), EVC and VVC

In our implementation, we selected several video codecs for evaluation: HEVC, AVC, EVC, and VVC, as shown in Table I. HEVC was used for the evaluation of coding efficiency in the standardization of V-PCC. HEVC is well-known for its coding tree unit, which significantly improves compression efficiency. We also selected the MPEG AVC as a successful standard, because it is used in many digital media devices for internet streaming and broadcasting services.
The VVC developed by Joint Video Exploration Team (JVET) is a recently established standard [10] that aims to be the successor of HEVC. It exhibits a higher coding efficiency and can handle 360-degree videos and multiview videos. During the standardization of VVC, high-bit precision and high-dynamic-range coding were implemented together.

In contrast to AVC, HEVC, and VVC, EVC was developed by the MPEG separately [11]. EVC is considered to be a licensing-friendly video codec [12]. It provides royalty-free toolsets, only accepts technologies with transparent licensing terms, and exhibits an efficiency of coding similar to that of HEVC.

| Codec | Profile | Predecessor | Software |
|-------|---------|-------------|----------|
| AVC   | main    | H.263       | JM       |
|       | high10  | MPEG-2      | 19.0     |
| HEVC  | main    | H.264/AVC   | HM       |
|       | main10  |             | 16.22    |
| EVC   | baseline| H.264/AVC   | ETM      |
|       | main    | H.265/HEVC  | 7.1      |
| VVC   | main10  | H.265/HEVC  | VTM      |
|       |         |             | 9.0      |

### III. EVALUATION PROCESS

The evaluation of the video codecs includes codec specification analysis, implementation, and performance observation. In this section, the analysis of V-PCC and video codec is discussed, and a pre-implementation overview is provided. Several problems, such as selection of the video codec profile, frame level accessibility, and OMAP problems due to lossless coding, were noted while implementing other codecs.

#### A. PROFILE AND LEVELS

For a video codec, a profile refers to a category of tool set that regulates the coding efficiency and computational complexity or the limitation of the picture size and bitrate support. In detail, the profile specifies the picture resolution, pixel bit depth, coding tool sets, and the levels of one codec define the maximum coding bitrate limits.

As shown in Table II, the dataset used in the common test condition (CTC) includes 10 bits and 11 bits geometry precision point cloud sequences for evaluation, and it generates different image sizes based on the point cloud size. For a class C sequence, the V-PCC generates an image size exceeding 1920 × 1080 (also called high-definition resolution). In several codecs, it will only be supported by higher profiles or levels.

Second, the bit depth support of video codecs is also related to it. In the CTC, the main10 profile of the HEVC is used as an anchor configuration. The high-bit precision support represents color in a more precise way inside the coding process but is also limited by the profile of each codec.

Third, some coding toolsets are regulated by the profiles. Tools that can improve coding efficiency, such as Transform 8 × 8 Mode in AVC, Delta QP for Cb, and Cr in EVC, can only be enabled in higher profiles. Other limitations that may be caused by codecs support are the pixel format and bitrate limitations. If a lossless configuration is used in V-PCC encoding, the video codec needs to have a YUV444 pixel support. In this way, the generated image of the color attribute can be archived in a less distorted condition.

Furthermore, patches generated by V-PCC can be placed in different locations of the image on different frames. Motion compensation has less efficiency when encoding these patches in sequence; thus, more bits are required to represent it. Finally, it might exceed the limits of the selected profiles or levels.

| Test Class | Dataset Name  | Frames | fps | Number of Points | Geometry Precision |
|------------|---------------|--------|-----|------------------|--------------------|
| A          | Queen         | 250    | 50  | ~1,000,000       | 10 bits            |
|            | Loot          |        |     | ~780,000         |                    |
|            | Red_and_Black | 300    | 30  | ~700,000         |                    |
| B          | Soldier       |        |     | ~1,500,000       |                    |
|            | Longdress     | 300    | 30  | ~800,000         | 10 bits            |
| C          | Basketball_player | 64    | 30  | ~2,900,000       | 11 bits            |
|            | Dancer        |        |     | ~2,600,000       |                    |

#### B. INTER/INTRA FRAME ACCESSIBILITY

In a video codec, inter-frame accessibility refers to an access behavior that requires reference frames inside one group of pictures (GOP), and intra-frame accessibility refers to individual access to a partial frame, which is a slice or tile in a picture. Similar to video codecs, access to a part of the frame or a specific frame of the point cloud sequence is also related to the video coding structure.

Access to a specific frame is required in the lossy configuration of the CTC, which includes the C2-intra and C2-inter random-access configurations. C2-intra means that a single point cloud frame can be accessed without prediction from other frames. In contrast, the C2-inter random-access configuration uses an inter-frame prediction to achieve a higher compression ratio. Simultaneously, the video codecs used in V-PCC should also have a similar coding structure to maintain frame-level accessibility. HEVC and VVC have their reference picture list (RPL), and they can design V-PCC-optimized picture order of coding (POC) structures. In the case of EVC, the configuration of the RPL is not supported in the baseline profile. This could cause problems during the design of the C2-intra condition [13].

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
Second, the partial access ability of V-PCC is required by the region of interest (ROI)-based partitioning. Although an ROI is applied in the encoding process, different quality parameters can be used in different regions. Then, the decoder can selectively decode a specific region based on the demand of the final viewer. This means that a video codec with a tile or slice-based intra-frame structure can satisfy this ability.

Additionally, adaptive QP options can improve coding efficiency. Because the near and far layers of the color attribute images are similar from a visual perspective, a QP offset applied to the prediction frame could have more coding gain. Furthermore, an increased QP for chroma components for color-attribute coding is recommended.

C. LOSSLESS OR NEAR LOSSLESS / LOSSY COMPRESSION
In the development of V-PCC, distortion control can be summarized as lossless or near-lossless/lossy. The losslessly compressed point clouds are mathematically identical to the original. On the other hand, lossy compression has a bitrate and quality controlled by a parameter [14]. Furthermore, near-lossless compression means that the number of points remains intact even if they share the same coordinates.

As one of the images generated from patch segmentation, the OMAP is important for reconstruction quality. Even in the CTC test lossy condition, the default configuration uses lossless coding to compress the OMAP. However, not all video codecs have lossless encoding options. Thus, lossy OMAP-related parameters in the encoder must be optimized. Parameters such as offsetLossyOM and thresholdLossyOM need to be set with codec specialized value for less distortion [15].

In the missed-point refinement process, the points lost during compression are recorded separately. Then, these point values are added to the end of the generated images or a separate image to ensure that the information is coded correctly.

VI. RESULT
A. IMPLEMENTATION
In the implementation, our designs are categorized by an inter-frame access method under three conditions: all-intra, random-access, and low-delay. The all-intra and random-access conditions are the same as the CTC, and they are designed to test the performance of the maximum frame level access and the maximum compression rate. The low-delay condition is chosen to evaluate the compression performance of the video codec with the most common configuration. The codec encoding specifications are listed in Table III. The choice for each codec is discussed in the following sections.

| Condition         | Profile | Level (IDC) | Inter Prediction | Adaptive QP                  | MV Search Range |
|-------------------|---------|-------------|------------------|-----------------------------|-----------------|
| HEVC All-Intra    | main    | 6.0         | RPL              | QP Offset for each frame (3–6) | 64              |
| HEVC Low-Delay    | main    | 6.0         | RPL with B-frames|                            |
| HEVC Random-Access| main10  | 6.0         | RPL              | QP Offset for each frame (3–6) | 64              |
| AVC All-Intra     | main    | 6.1         | Internal POC     | QP Offset for each prediction type (3–6) | Increase to 64 |
| AVC Low-Delay     | main    | 6.1         | Internal POC with B-frames | |
| AVC Random-Access | high10  | 6.1         | RPL              | QP Offset for each frame (3–6) | Decrease to 64 |
| EVC All-Intra     | baseline| 5.1         | Internal POC     | QP Offset for each frame (3–6) | 64              |
| EVC Low-Delay     | baseline| 5.1         | RPL              | QP Offset for each frame (3–6) | 64              |
| EVC Random-Access | main    | 5.1         | RPL              | QP Offset for each frame (3–6) | 64              |
| VVC All-Intra     | main10  | auto        | RPL              | QP Offset for each frame (3–6) | 64              |
| VVC Low-Delay     | main10  | auto        | RPL              | QP Offset for each frame (3–6) | 64              |

1) COMMON CONDITION
Unlike the CTC, we use lossy compressed OMAP encoding methods. The occupied signal is set to 255 because TMC version 7.2 does not handle the exception of low-quality OMAP compression. We changed the occupancy precision rate to 1:1, instead of 2 or 4.
different compression quality are used as the quantization parameter (QP) for each rating point of CTC predefined values, as shown in Table IV.

| Table IV | RATE POINT AND QP |
|---|---|
| Rate Point | R1 | R2 | R3 | R4 | R5 |
| Color QP | 42 | 37 | 32 | 27 | 22 |
| Geometry QP | 32 | 28 | 24 | 20 | 16 |

2) HEVC
In contrast to the CTC, we selected different profiles for different inter-frame access methods. We chose the Main and Main-10 profiles as they are the most commonly supported HEVC version 1 profiles. The other encoding configurations were the same as the CTC recommendations.

3) AVC
We chose profiles for AVC, as shown in Table III. Based on the HEVC search range, we extended the search range from 32 to 64 for the random-access configuration for better inter-frame prediction. It should be noted that when the color attribute bitstream is encoded for class C data, the bitstream rate could exceed the limits of level 5.2.

4) EVC
For the all-intra condition in EVC, we designed a picture coding order fixed for the baseline profile in the encoder, where the near layer is coded into an intra-frame, and the far layer is coded into a p-frame.

We also enabled delta QP to encode color attributes, and changed the default motion vector search range for random-access from 384 to 64.

5) VVC
VVC as a codec developed after the HEVC, which has a significant influence from HEVC. Most of the HEVC encoding configurations can be adapted to VVC.

B. RESULT AND ANALYSIS
The test model version 7.2 was used in the experiment. The environment of this experiment was executed on computers with hardware and software configurations, as shown in Table V. The performance of each codec includes coding efficiency and time complexity. We also set the encoding parameter nbThread to eight for parallel encoding.

| Table V | TEST COMPUTER SPECIFICATION |
|---|---|
| Processors | Intel Core i7-10700K (16 Cores) |
| Memory | DDR4 3200MHz (Dual channel) |
| Project Develop Environment | Microsoft Visual Studio 2017 |
| Operating System | Microsoft Windows Server 2019 |

The coding efficiency is categorized into two parts: the individual performances on the geometry and color attributes, and the performance on the total bitrate. The geometry performance is measured by the point cloud error (PC error) in PSNR with point-to-point (P2Point (D1)) error and point-to-plane (P2Plane (D2)) error [15]. The color attribute performance was measured by the PC error on luminance, Chroma Cb, and Chroma Cr. According to the CTC document, both P2Point and P2Plane use a larger mean square error (MSE) of the original versus reconstructed point cloud alternately in the peak signal-to-noise ratio (PSNR) calculation, as shown in (1).

\[
PSNR = 10 \log_{10} \left( \frac{3p^2}{\max(\text{PC Error}_A, \text{PC Error}_B)} \right)
\]

We evaluate each coding structure in the following sections by graphical representation of the relation between PSNR and bit size instead of BD-rate comparison.

1) ALL-INTRA CASE
In this section, we consider every sequence for a more precise analysis. We plotted the bits per input point (BPIP) and PSNR measured from PC Error on the x and y axes, respectively. Figs. 3–9 show the rate-distortion (RD) of the geometry and luma component of the color attribute for all-intra coding structures. Each figure includes the HEVC, EVC, AVC, and VVC results.

The geometry RD curve indicates that the PSNR differences from the minimum to the maximum are 0.48 dB, 0.28 dB, 1.04 dB, and 1.56 dB for each codec. This means that the changes in geometry QP could not help to improve the total geometry quality under a lossy OMAP situation. For the color part, each codec had a relative performance as their claimed coding efficiency. The VVC performed best and, EVC showed coding gains compared to AVC with the baseline profile.

An uncommon result was obtained for the EVC encoding the dancer sequences, as shown in Fig. 9. It performs worse than AVC in terms of both geometry and color.
FIGURE 3. All-intra BDBR curve of “loot.”

FIGURE 4. All-intra BDBR curve of “redandblack.”

FIGURE 5. All-intra BDBR curve of “soldier.”

FIGURE 6. All-intra BDBR curve of “queen.”
2) LOW-DELAY CASE

Figs. 10–16 depict the RD of the geometry and luma component of the color attribute for the low-delay coding structure. As shown in these figures, AVC exhibits the best performance, whereas EVC only achieves less than 50 dB for geometry coding.

For the color attribute part, the VVC shows the best performance. EVC and AVC yield the same equivalent efficiency of most sequences.
FIGURE 10. Low-delay BDBR curve of “loot.”

FIGURE 11. Low-delay BDBR curve of “redandblack.”

FIGURE 12. Low-delay BDBR curve of “soldier.”

FIGURE 13. Low-delay BDBR curve of “queen.”
3) RANDOM-ACCESS CASE

The random-access performance of each codec is shown in Figs. 17–23. The geometry RD curve shows that the PSNR differs for each codec and is larger than the all-intra condition. It is noted that VCC with a lossy OMAP also achieves a good performance, with the D1 PSNR exceeding 65 dB.

The color part of VVC also reaches a high PSNR with a relatively low bitrate. However, AVC-based codecs show poor efficiency, where the OMAP-excluded geometry bitstream size is significantly larger than other codecs. On the color part of the bitstream, the EVC achieved a performance similar to that of HEVC. However, the AVC needs to double the BPIP to have a similar quality of VVC.
FIGURE 17. Random-access BDBR curve of “loot.”

FIGURE 18. Random-access BDBR curve of “redandblack.”

FIGURE 19. Random-access BDBR curve of “soldier.”

FIGURE 20. Random-access BDBR curve of “queen.”
4) COMPLEXITY EVALUATION

For the complexity metric, we take the geometric mean of each single runtime and sequence. The video processing and video-excluded times were measured separately. The video-excluded time includes the processing time of each thread used in the TMC. In Table VII, it is the all-intra runtime that compares three codecs with the HEVC-based codec.

The results show that the overall video encoding excluded time did not change significantly. However, the video encoding times were different for each codec. AVC and EVC require less video processing time, whereas the VVC encoding time increased to 9.6 times that of HEVC. On the decoder side, owing to the points being reconstructed on the AVC and EVC codec, the decoding time is reduced by 14%–19%.

| Codec  | Encoder runtime | Decoder runtime |
|--------|-----------------|-----------------|
|        | Video excl. thread time | Video enc. thread time | Video excl. thread time | Video dec. |
| AVC    | 100%            | 16%             | 81%            | 76%        |
| EVC    | 99%             | 37%             | 86%            | 91%        |
| VVC    | 100%            | 962%            | 99%            | 112%       |
As presented in Table VII, the three codecs require the random-access condition runtime. The table shows that the overall video encoding excluded time was reduced by 7%–8% for AVC and EVC. The EVC and VVC have used more video processing time up to 4.4 times and 3.2 times, but AVC encoding time only takes 17% of the HEVC one. With fewer points reconstructed on the EVC decoder, the decoding time is reduced by 8%. The AVC codec used 27% less than the anchor.

| TABLE VIII | CODING COMPLEXITY OF LOW-DELAY |
|------------|---------------------------------|
| Encoder runtime | Decoder runtime |
| Video excl. thread time | Video enc. | Video excl. thread time | Video dec. |
| AVC | 83% | 924% | 85% | 66% |
| EVC | 84% | 13% | 94% | 79% |
| VVC | 88% | 292% | 102% | 111% |

| TABLE IX | CODING COMPLEXITY OF RANDOM-ACCESS |
|-----------|-----------------------------------|
| Encoder runtime | Decoder runtime |
| Video excl. thread time | Video enc. | Video excl. thread time | Video dec. |
| AVC | 93% | 17% | 106% | 73% |
| EVC | 92% | 448% | 92% | 101% |
| VVC | 99% | 325% | 100% | 101% |

5) SUBJECTIVE COMPARISON
The final result of the rendered point-cloud data was confirmed. As presented in Figs. 24–26, we capture the first frame of the reconstructed sequence along the z-axis. Each figure shows the captures of different codecs with the same coding conditions but with rate point three (R3). The rendering software is the PCC renderer [16] used by the MPEG.

Fig. 24 shows the image of each codec for all-intra conditions. As expected, on the edge of the rendered object, a lossy OMAP generates many noise points. VVC shows the best quality, whereas AVC shows the worst quality.

The low-delay condition results for each codec are shown in Fig. 25. VVC shows the best quality, which does not match the result with the lowest RD curve. AVC shows the worst quality with considerable noise when the RD curve is above the HEVC one.

The captured results of the random-access condition for each codec are presented in Fig. 26. VVC shows the best quality, and AVC shows the worst quality, as expected. EVC shows better quality than AVC in terms of color attributes.

FIGURE 24. Rendering result of all-intra case.

FIGURE 25. Rendering result of low-delay case.

FIGURE 26. Rendering result of random-access case.

It is known that the inter-predicted frame inside a GOP has a different quality owing to the delta QP. We captured the 16th frame of each codec. Fig. 27 shows the 16th frame under low-delay conditions. The head part of the EVC encoded object has a major color distortion compared to the other three. As depicted in Fig. 28, the VVC-encoded object has different subjective reviewing results compared to the
RD-curve result, which shows color distortion on the face part of the model. This indicates that the design of the coding structure has a direct quality impact on the final result.

**FIGURE 27.** Rendering result of 16th frame low-delay case.

**FIGURE 28.** Rendering result of 16th frame random-access case.

**VII. CONCLUSION**

In this study, we adopted three video codecs, namely, AVC, EVC, and VVC, into V-PCC and evaluated the agnostic design of V-PCC from several perspectives. The implementation and test results confirmed that the video codec performance interferes with the V-PCC codec system. In future works, the coding performance should be improved and optimized to help the development and spread of V-PCC technology.

**REFERENCES**

[1] LUCID Vision Labs Inc., “Time of flight gets precise: Enhanced 3D measurement with Sony DepthSense technology,” Richmond B.C., Canada, 2017

[2] Coded representation of immersive media Part 5: Visual volumetric video-based rendering (V3C) and video-based point cloud compression (V-PCC), ISO/IEC FDIS 23090-5, 2020

[3] 3DG, "V-PCC codec description VPCC TMv11 algorithm description," ISO/IEC JTC 1/SC 29/WG 11, w19526, [Online], Jun. 29-Jul. 3, 2020

[4] V. Zakharchenko, "[VPCC] [EE4FE 2.7 Report] Evaluation of the results of multiple video codec integration in V-PCC software,” ISO/IEC JTC 1/SC 29/WG 11, m53334, Alpbach, Apr. 20-24, 2020

[5] Y. -T. Tsai, C. -L. Lin, C. -L. Lin, Y. -H. Lee, J. -L. Lin, Y. -J. Chen, and C. -C. Ju, "[VPCC] [EE4FE 2.7 related] AVC codec integration in V-PCC test model,” ISO/IEC JTC 1/SC 29/WG 11, m53410, Alpbach, Apr. 20-24, 2020

[6] D. Mehlem and C. Rohlfing, "[VPCC] Versatile video coding for VPCC,” ISO/IEC JTC 1/SC 29/WG 11, m53510, Alpbach, Apr. 20-24, 2020

[7] E. Chang, J. Cha, T. Dong, K. Kim, and E. Jang, "On lossy occupancy map compression in V-PCC," ISO/IEC JTC 1/SC 29/WG 11, m34604, [Online], Jun. 29-Jul. 3, 2020

[8] T. Dong et al., “MPEG video-based point cloud coding based on JPEG,” 3DSA, IDW ’19, 2019

[9] Y. -K. Wang and J. Chen, [EVC] On MPEG-5 EVC profiles, tiers, and levels,” ISO/IEC JTC 1/SC 29/WG 11, m48981, Gothenburg, Jul. 8-12, 2019

[10] International Telecommunication Union (ITU-T), Infrastructure of audiovisual services — Coding of moving video (H.266), ISO/IEC 23090-3, 2020

[11] Information technology - General video coding - Part 1: Essential video coding, ISO/IEC 23094-1:2020

[12] K. Choi, J. Chen, D. Rusanovskyy, K. Choi, and E. S. Jang, "An overview of the MPEG-5 essential video coding standard [standards in a nutshell],” in IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 160-167, May 2020, doi: 10.1109/MSP.2020.2971765.

[13] “Common test conditions for essential video coding,” ISO/IEC JTC 1/SC 29/WG 11, N19003, Brussels, Jan. 13-17, 2020

[14] 3DG, “PCC requirements,” ISO/IEC JTC 1/SC 29/WG 11, N17353, Gwangju, Jan. 22-24, 2018

[15] 3DG, “Common test conditions for point cloud compression”, ISO/IEC JTC 1/SC 29/WG 11, N18883, Gothenburg, July 2018

[16] 3DG, “MPEG 3DG renderer,” Online, http://mpegx.intevry.fr/software/MPEG/PCC/mpeg-pcc-renderer

**TIANYU DONG** was born in Linyi, Shandong, China, in 1987. He received the B.S. degree in 2011 and the M.S. degree in electronics and computer engineering from Hanyang University, South Korea, in 2014. He is currently a Ph.D. candidate in the Department of Computer Science, Hanyang University. He started his research on video compression in 2011. His interests include multimedia-related compression technology and application. He has authored 10 articles on conference papers.

**KYUTAE KIM** was born in Anseong-si, Gyeonggi-do, Korea, in 1996. He received the B.S. degree in computer engineering from Gangnam University, South Korea, in 2020. He is currently a M.S. student in the Department of Computer Science, Hanyang University.
EUEE S. JANG (SM) received the B.S. degree in 1991 from Jeonbuk National University, South Korea, and the Ph.D. degree from SUNY at Buffalo, Buffalo, NY, USA, in 1996. He is currently a Professor at the Department of Computer Science and Engineering, Hanyang University, Seoul, South Korea. He has authored more than 325 articles on MPEG standardization, more than 90 journal articles and conference papers, more than 65 pending or accepted patents, and two book chapters. His research interests include image/video coding, reconfigurable video coding, and computer graphics objects. He has also received three ISO/IEC certificates of appreciation for contributions to MPEG-4 development. Finally, he received the Presidential Award from the Korean Government for his contribution to MPEG standardization.