IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 14, 2021, accepted December 10, 2021, date of publication December 20, 2021,
date of current version December 29, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3137036

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

TIANYU DONG, (Graduate Student Member, IEEE), KYUTAE KIM, AND EUEE S. JANG, (Senior Member, IEEE)
Department of Computer Science, Hanyang University, Seoul 04763, South Korea

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

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) funded by the Korean Government [Ministry of Science and ICT (MSIT)] (Development of Adaptive Viewer-Centric Point Cloud AR/VR (AVPA) Streaming Platform) under Grant 2020-0-00452.

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 a 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 using any video codec. The V-PCC reference software is currently designed using MPEG High-Efficiency Video Coding. We extended the evaluation of video codec applicability for PCC using well-known MPEG video coding standards, such as Advanced Video Coding, Essential Video Coding, and Versatile Video Coding. We identified several key strategies for applying a video codec to V-PCC to maximize the compression efficiency or computational complexity during the evaluation. Furthermore, the coding efficiency and time complexity of each codec are tested. The evaluation tests revealed that V-PCC supports the codec agnostic approach, and that the performance of the video codec is positively correlated 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

Point cloud data has been used in virtual reality (VR), augmented reality (AR), and mixed reality (MR) as immersive media. Consumer electronics or autonomous vehicles implemented in VR, AR, and MR applications use point cloud data as an immersive 3D media representation. Some 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 of colorized points per second. These large amounts of data require efficient compression methods for their storage and transmission. The Moving Picture Experts Group (MPEG) began standardizing point cloud data compression in 2017. This led to the development of 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 attributes [3]. V-PCC is designed to work with 2D video compression tools such that 2D images (after 3D to 2D transformation) can be compressed 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.

Some implementation examples show 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]. A previous study showed that VP9, x264, and x265 contained in the FFmpeg package could work

The associate editor coordinating the review of this manuscript and approving it for publication was Charalambos Poullis.

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 anchor and all-intra coding structures was evaluated [5]. An experiment based on the new VVC codec showed that VVC-based V-PCC achieves a coding gain similar to 2D videos [6]. Studies based on EVC indicate that the lossy compression video codec significantly changes the quality of V-PCC, causing all loss frames [7]. However, problems related to implementation have not been systematically discussed. Our previous research based on the legacy image codec showed that JPEG can be used to compress geometry and attribute images with an all-intra coding structure [8]. We also observed that the video codec performance had a distinct impact on the V-PCC codec based on research on the EVC baseline profile [8]. However, previous studies have not compared the codec performance and computational complexity. Design-related strategies in which different video codec profiles and levels could be limitations also need to be discussed further.

In this study, we extended our evaluation to include three codes: 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 discuss 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 strategies for different implementations. The implementation results are presented in Section IV. Finally, conclusions 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 an object can be projected onto 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 attributes. Next, V-PCC uses video encoders to compress images. Finally, the encoder merges all the video bitstreams and auxiliary information as a V-PCC bitstream.

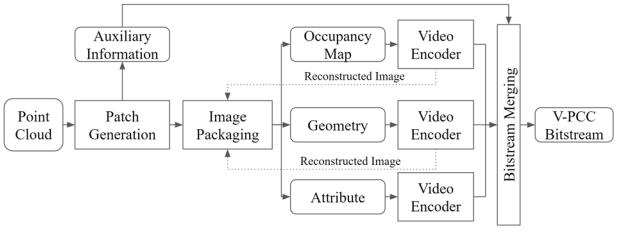

FIGURE 1. V-PCC encoder.

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

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

The size of the generated images, taken as examples, was 2560 × 1856. These images were stored in YUV color space. In OMAP, bit "1" represents an occupied signal and "0" represents an unoccupied signal. In the geometry images, the depths of these surfaces are represented in eight bits. The occupancy and geometry signals were stored in the Y-plane of the YUV color space. For the color attributes, the color data in the RGB color space were first converted into those in the RGB444-16bit intermedium 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 1. HEVC was used to evaluate the 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 MPEG AVC as a successful

TABLE 1. Codecs evaluated for V-PCC.

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

TABLE 2. Test dataset and specification.

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

standard because it is used in many digital media devices for Internet streaming and broadcasting services.

The VVC developed by the 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.

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

III. EVALUATION PROCESS

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

A. PROFILE AND LEVELS

For a video codec, a profile refers to a category of toolsets that regulate the coding efficiency and computational complexity or the limitation of the picture size and bitrate support. The profile specifies the picture resolution, pixel bit depth, and coding toolsets, and the levels of one codec define the maximum coding bitrate limits.

As shown in Table 2, 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, V-PCC generates an image size exceeding 1920×1080 pixels (also called high-definition resolution). Therefore, it has been supported only by higher profiles or levels in several codecs.

Second, the bit-depth support of video codecs is also related to it. In CTC, the main10 profile of the HEVC was used as the anchor configuration. The high-bit precision support represents color more precisely 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 efficiencies, 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 codec support are the pixel format and bitrate limitations. For example, if a lossless configuration is used in V-PCC encoding, the video codec needs to have YUV444 pixel support. Thus, the generated image of the color attribute can be archived under less distorted conditions.

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

B. INTER/INTRA FRAME ACCESSIBILITY

In a video codec, inter-frame accessibility refers to 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 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, including the C2-all-intra (AI) and C2-inter random-access (RA) configurations, is required for the lossy configuration of the CTC. C2-AI indicates that a single point cloud frame can be accessed without prediction from other frames. In contrast, the C2-inter random-access configuration uses inter-frame prediction to achieve a higher compression ratio. Simultaneously, the video codecs used in V-PCC should have a similar coding structure to maintain frame-level accessibility. HEVC and VVC have a 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 by the baseline profile [9]. This could cause problems during the design of the C2-AI condition [13].

TABLE 3. Codec encoding specification for V-PCC.

	Condition	Profile	Level (IDC)	Inter Prediction	Adaptive QP	MV Search Range
HEVC	AI	main	6.0	RPL	QP Offset for each frame (3~6)	64
	LD	main	6.0			
	RA	main10	6.0	RPL with B-frames		
AVC	AI	main	6.1	Internal POC	QP Offset for each prediction type (~3)	Increase to 64
	LD	main	6.1			
	RA	high10	6.1	Internal POC with B-frames		
EVC	AI	baseline	5.1	Internal POC	Internal delta QP Enable	64
	LD	baseline	5.1			
	RA	main	5.1	RPL with B-frames		Decrease to 64
VVC	AI	main10	auto	RPL	QP Offset for each frame (3~6)	64
	LD	main10	auto			
	RA	main10	auto	RPL with B-frames		Decrease to 64

Second, the partial access ability of V-PCC is required by region of interest (ROI)-based partitioning. Although an ROI is applied during the encoding process, different quality parameters can be used in different regions. The decoder can then selectively decode a specific region according to the demands 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 the coding efficiency [17], [18]. 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 can provide more coding gain. Furthermore, an increased QP for the 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 were mathematically identical to the original point clouds. However, lossy compression has a bitrate and quality that is 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, OMAP is important for reconstruction quality. Even under the CTC test lossy condition, the default configuration uses lossless coding to compress 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 values for less distortion [15].

During the missed-point refinement process, the points lost during the compression were recorded separately. These point values were added to the end of the generated images or a separate image to ensure that the information was correctly coded.

IV. RESULT

A. IMPLEMENTATION

In the implementation, our designs were categorized by an inter-frame access method under three conditions: all-intra (AI), random-access (RA), and low-delay (LD). The AI and RA conditions were the same as those of CTC, and they were designed to test the performance of the maximum frame-level access and maximum compression rate. The LD condition was chosen to evaluate the compression performance of the video codec with the most common configuration. The codec encoding specifications are presented in Table 3. The choice of each codec is discussed in the following sections.

1) COMMON CONDITION

Unlike CTC, we use lossy compressed OMAP encoding. The occupied signal was set to 255 because TMC version 7.2 does not handle the exception of low-quality OMAP compression. In addition, we changed the occupancy precision rate to 1:1 instead of 2 or 4.

The HEVC configuration guides the common video configurations. For example, in the all-intra condition, the GOP was set to 2, where the first picture was an I-frame, followed by a P-frame. However, the global 3D patch compensation tools are disabled in the random-access condition because they are only implemented in HEVC. Furthermore, the settings for different compression qualities were used as the

quantization parameter (QP) for each rating point of the predefined CTC values, as listed in Table 4.

TABLE 4. 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 CTC, we selected different profiles for different inter-frame access methods. We chose the main and main-10 profiles because they are the most commonly supported HEVC version 1 profiles. The other encoding configurations were the same as those used in CTC recommendations.

3) AVC

We chose AVC profiles as shown in Table 3. Based on the HEVC search range, we extended the search range from 32 to 64 for the RA configuration to improve inter-frame prediction. It should be noted that when the color attribute bitstream is encoded for class C data, the bitstream rate can 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 was coded into an intra-frame, and the far layer was coded into a p-frame.

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

5) VVC

VVC, a codec developed after HEVC, has a significant influence from HEVC. Most HEVC encoding configurations can be adapted for the VVC.

B. RESULT AND ANALYSIS

We used the test model version 7.2 in the experiment. The experiment was executed on computers with hardware and software configurations as shown in Table 5. The performance of each codec includes coding efficiency and time complexity. We also set the encoding parameter *nbThread* to eight for parallel encoding.

TABLE 5. 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

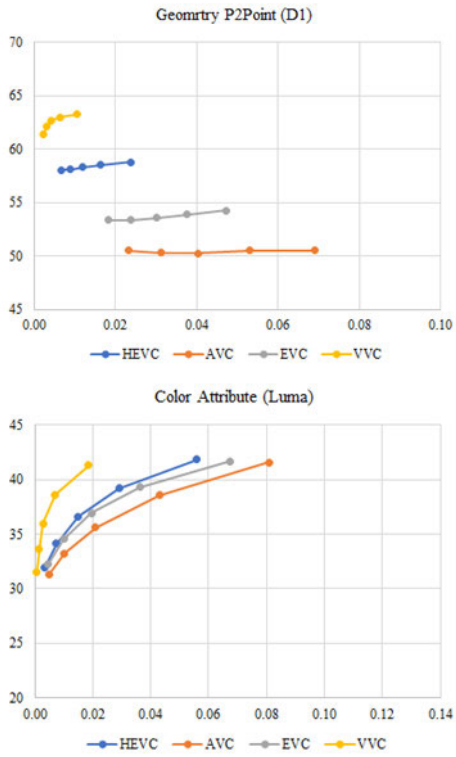

FIGURE 3. All-intra BDBR curve of "loot."

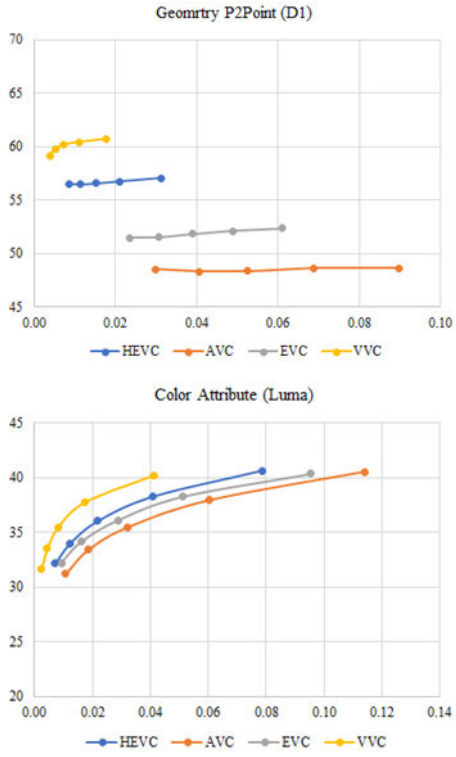

FIGURE 4. All-intra BDBR curve of "redandblack."

The coding efficiency is categorized into two parts: the individual performance of the geometry and color attributes

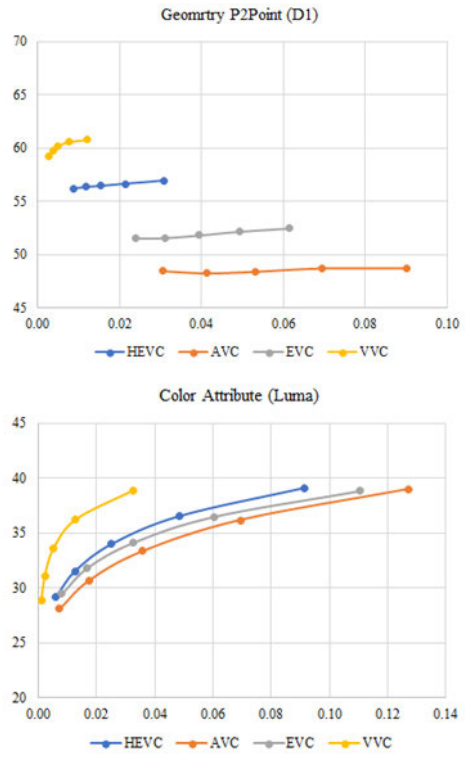

FIGURE 5. All-intra BDBR curve of "soldier."

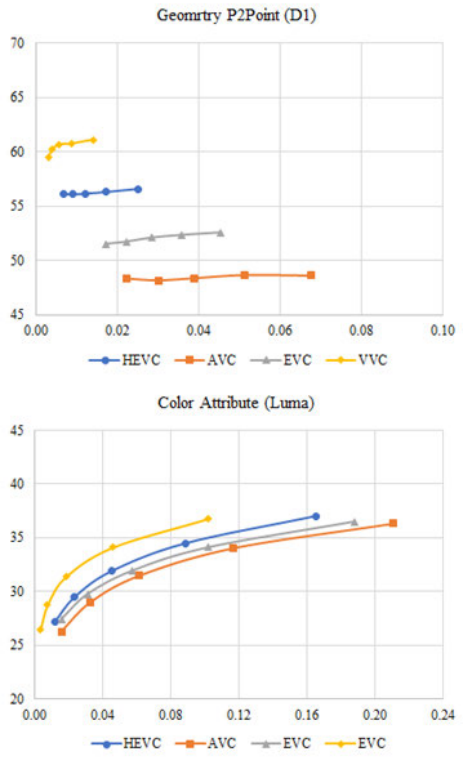

FIGURE 7. All-intra BDBR curve of "longdress."

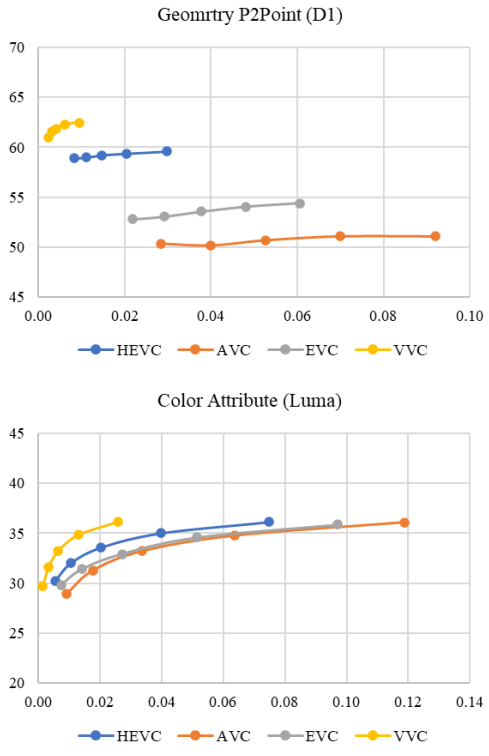

FIGURE 6. All-intra BDBR curve of "queen."

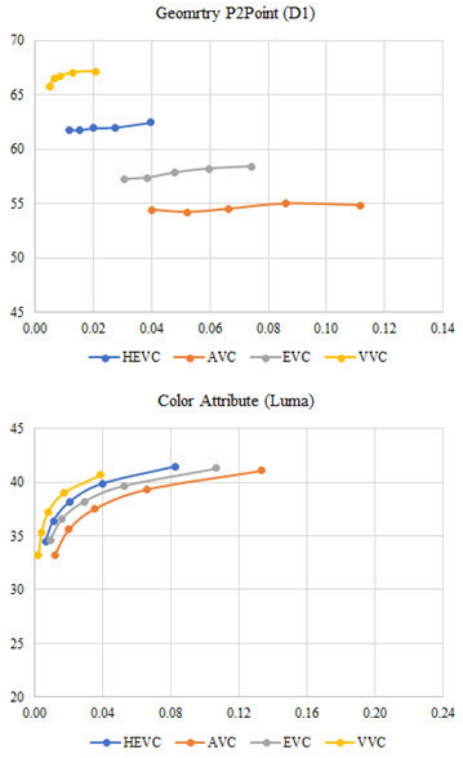

FIGURE 8. All-intra BDBR curve of "basketball_player."

and the performance of the total bitrate. The geometry performance is measured by the point cloud error (PC_error) in the PSNR with point-to-point (P2Point (D1)) and point-to-plane (P2Plane (D2)) errors [15]. PC_error measures the

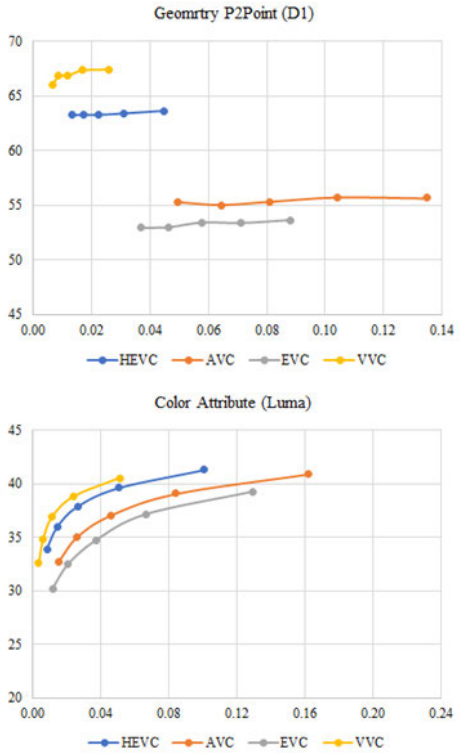

FIGURE 9. All-intra BDBR curve of "dancer."

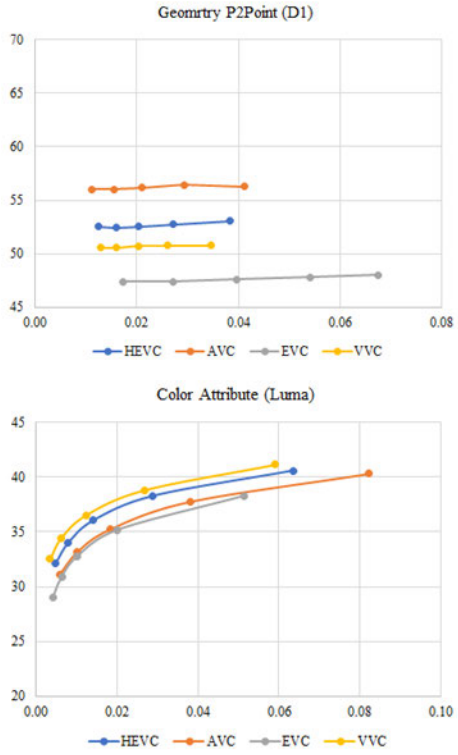

FIGURE 11. Low-delay BDBR curve of "redandblack."

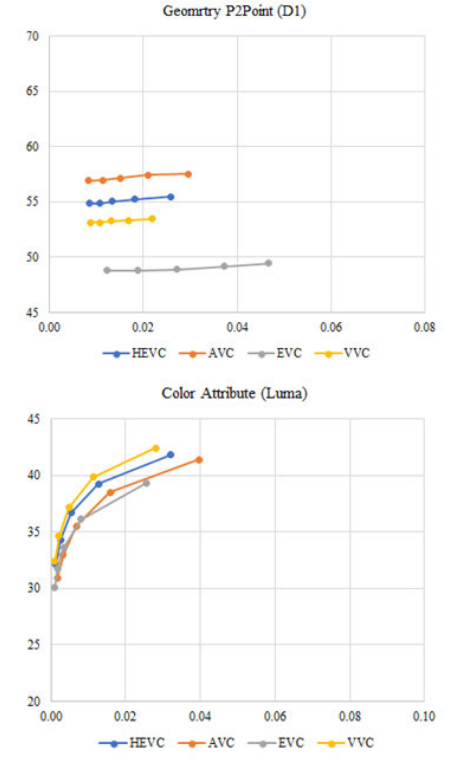

FIGURE 10. Low-delay BDBR curve of "loot."

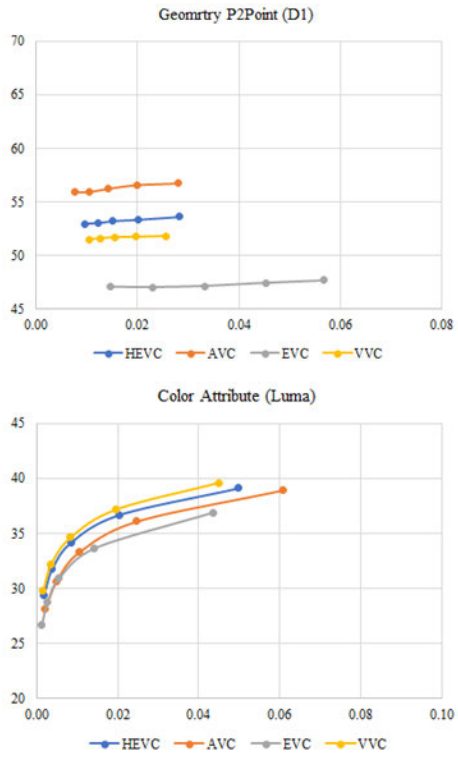

FIGURE 12. Low-delay BDBR curve of "soldier."

color attribute performance of luminance, chroma Cb, and chroma Cr. According to the CTC document, both P2Point and P2Plane alternately use a larger mean square error (MSE) of the original versus the reconstructed point cloud in the

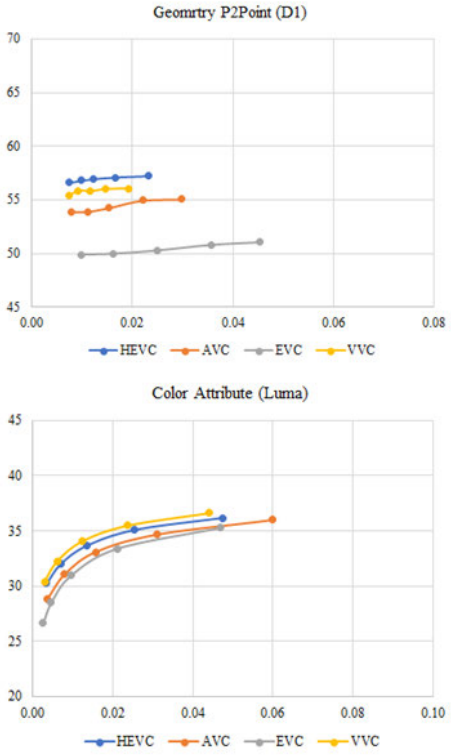

FIGURE 13. Low-delay BDBR curve of "queen."

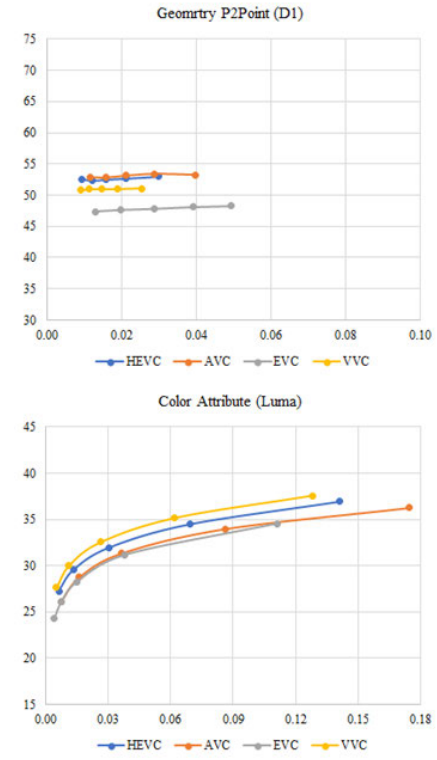

FIGURE 14. Low-delay BDBR curve of "longdress."

peak signal-to-noise ratio (PSNR) calculation, as shown in (1). In (1), $e_{\mathbf{B},\mathbf{A}}$ is the error of point cloud **B** relative to

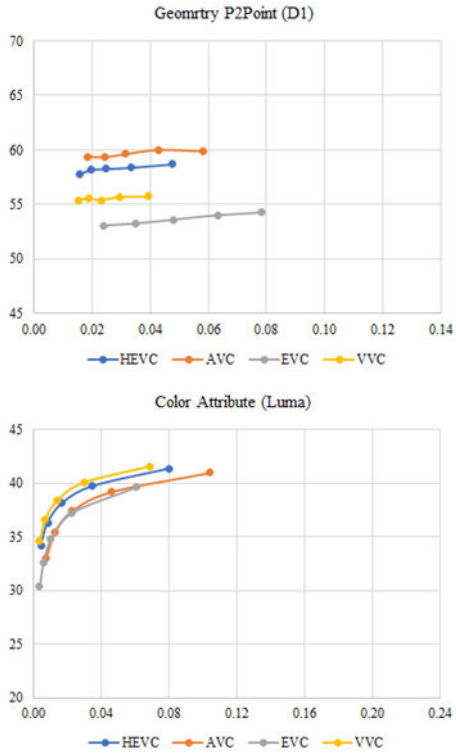

FIGURE 15. Low-delay BDBR curve of "basketball_player."

reference point cloud **A**., and p is the peak constant value of the geometry precision for each sequence in Table 2.

$$PSNR = 10\log_{10}\left(\frac{3p^2}{max(e^{DX}_{\mathbf{B},\mathbf{A}}, e^{DX}_{\mathbf{A},\mathbf{B}})}\right), \tag{1}$$

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

1) ALL-INTRA CASE

In this section, we consider every sequence for 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 compared with the original data for the AI coding structures. Each figure shows the HEVC, EVC, AVC, and VVC results, respectively.

The geometry RD curve indicates that the lowest to highest PSNR differences 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. Each codec had a relative performance as their claimed coding efficiency for the color part. VVC performed the best, and EVC showed coding gains compared to AVC with the baseline profile.

An uncommon result was obtained for EVC encoding of the dancer sequences, as shown in Fig. 9. It performs worse than AVC in terms of both the geometry and color.

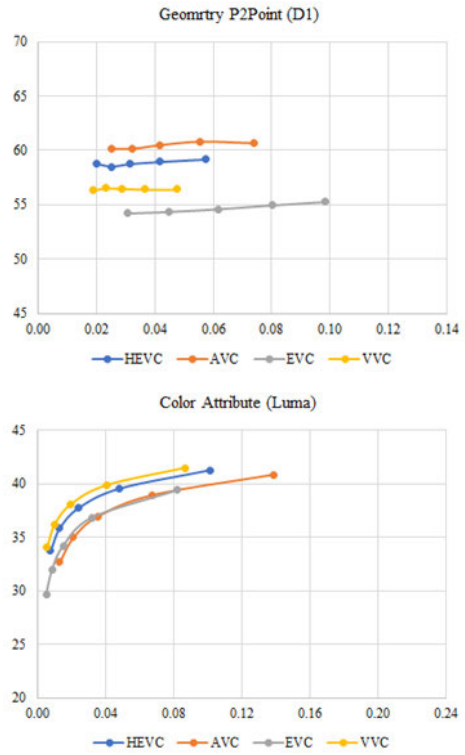

FIGURE 16. Low-delay BDBR curve of "dancer."

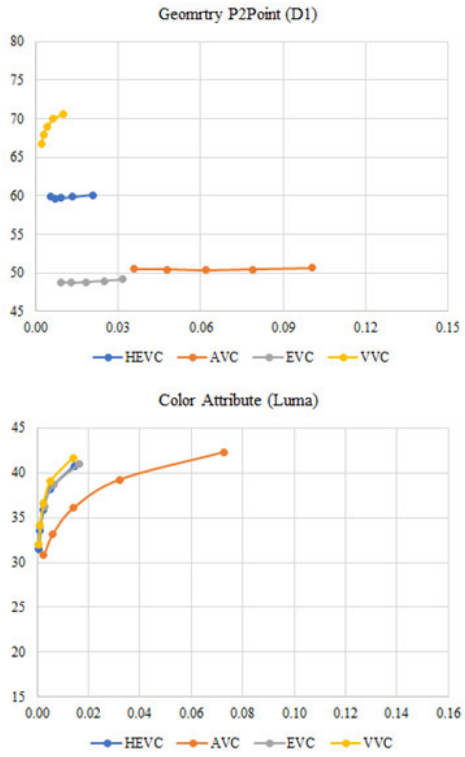

FIGURE 17. Random-access BDBR curve of "loot."

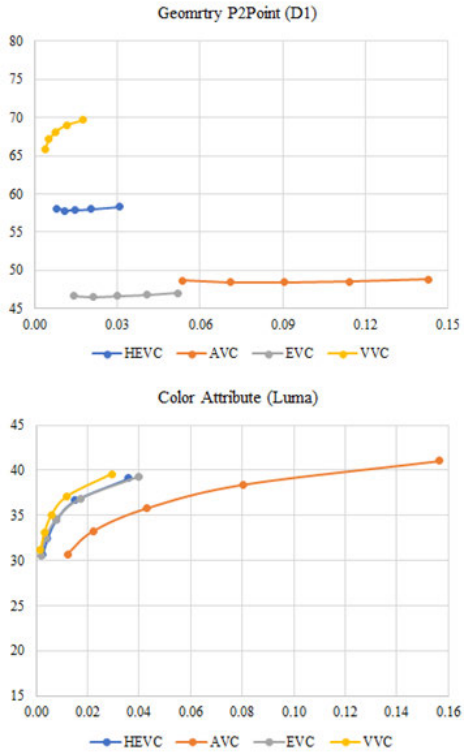

FIGURE 18. Random-access BDBR curve of "redandblack."

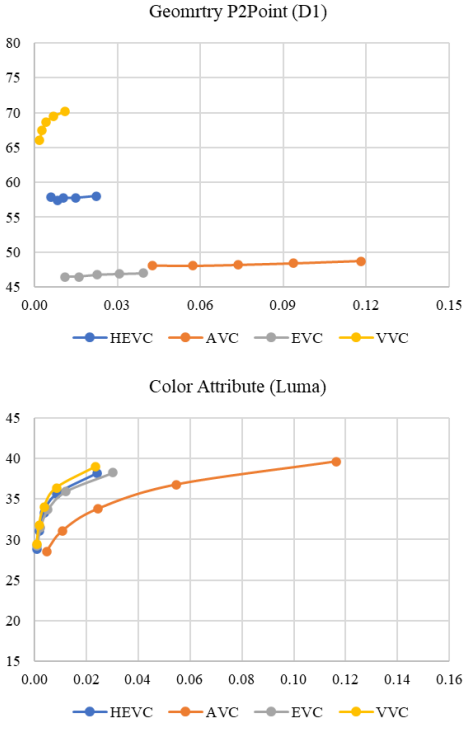

FIGURE 19. Random-access BDBR curve of "soldier."

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. These figures show that AVC exhibits the best performance, whereas EVC only achieves less than 50 dB for geometry coding.

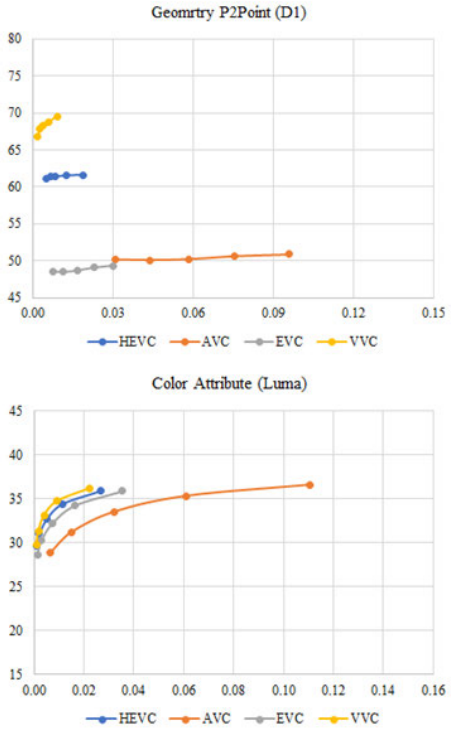

FIGURE 20. Random-access BDBR curve of "queen."

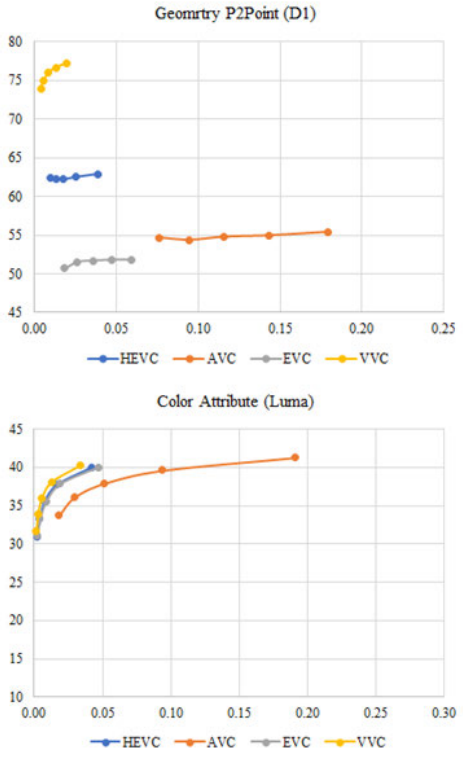

FIGURE 22. Random-access BDBR curve of "basketball_player."

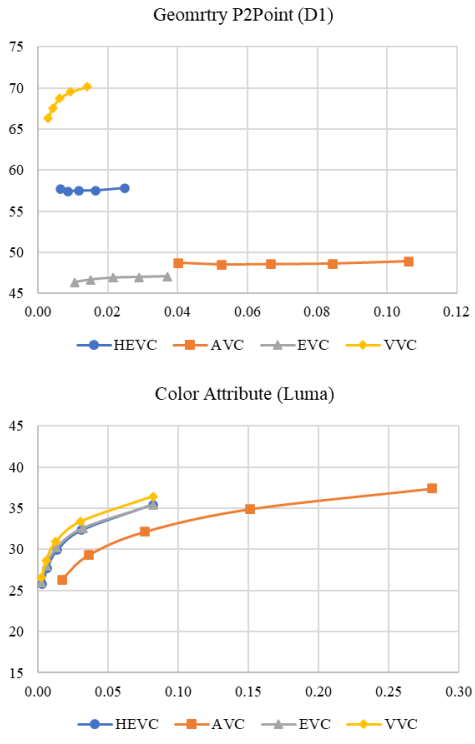

FIGURE 21. Random-access BDBR curve of "longdress."

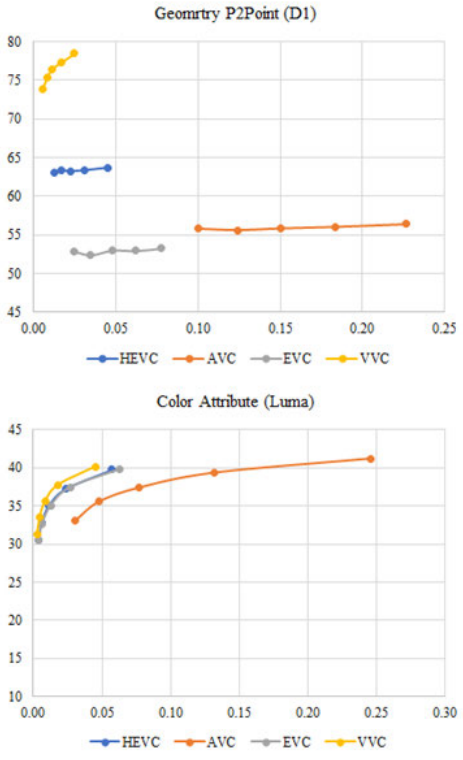

FIGURE 23. Random-access BDBR curve of "dancer."

For the color attribute part, VVC shows the best performance. EVC and AVC yielded the same equivalent efficiencies for most sequences.

3) RANDOM-ACCESS CASE

The RA performance for each codec is shown in Figs. 17–23. The geometry RD curve shows that the PSNR

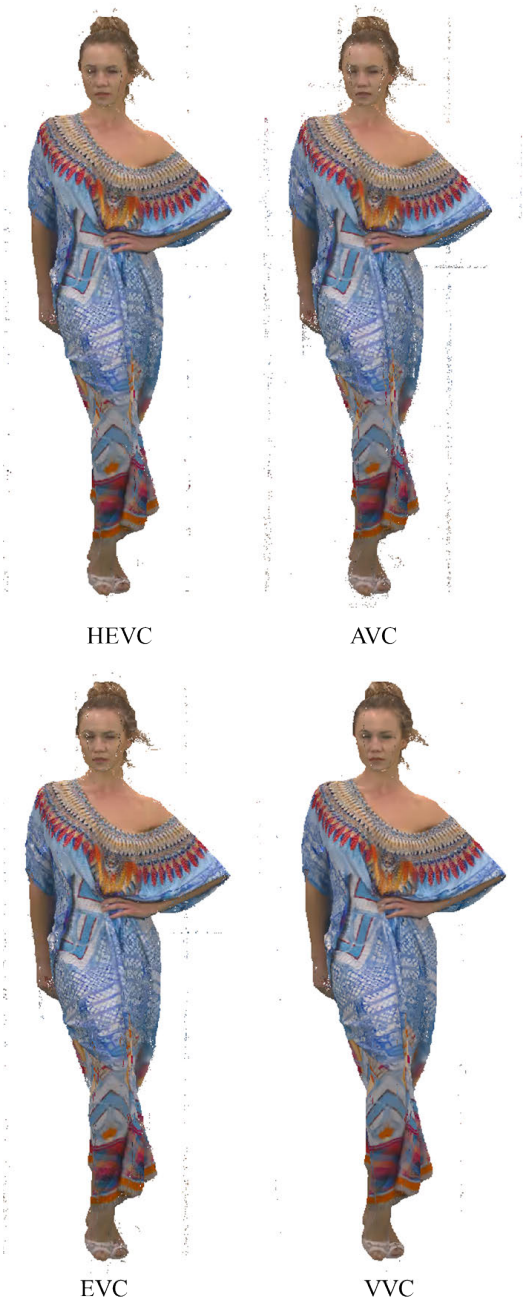

FIGURE 24. Rendering result of all-intra case.

differs for each codec and is larger than that under the AI condition. It is noted that VCC with a lossy OMAP also achieves good performance, with D1 PSNR exceeding 65 dB.

The color part of the 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 the other codecs. On the color part of the bitstream, the EVC achieved a performance similar to the HEVC. However, the AVC needs to double the BPIP to have a similar quality to VVC.

TABLE 6. Coding complexity of all-intra.

	Encoder runtime		Decoder runtime	
	Video excl. thread time	Video enc.	Video excl. thread time	Video dec.
AVC	100.5%	16.4%	81.6%	75.3%
EVC	98.8%	36.6%	85.7%	89.9%
VVC	100.8%	965.1%	98.8%	111.5%

TABLE 7. Coding complexity of low-delay.

	Encoder runtime		Decoder runtime	
	Video excl. thread time	Video enc.	Video excl. thread time	Video dec.
AVC	91.5%	921.1%	84.2%	65.2%
EVC	91.9%	12.9%	93.6%	78.3%
VVC	94.4%	292.0%	100.1%	110.0%

TABLE 8. Coding complexity of random-access.

	Encoder runtime		Decoder runtime	
	Video excl. thread time	Video enc.	Video excl. thread time	Video dec.
AVC	96.6%	17.0%	106.5%	73.3%
EVC	94.9%	456.4%	92.4%	100.5%
VVC	98.8%	325.1%	99.8%	100.7%

4) COMPLEXITY EVALUATION

We take the geometric mean of every runtime and sequence as the complexity metric. The video processing and video-excluded times were measured separately. We conducted this test twice and measured the running times of both tests. The video-excluded time includes the processing time of each thread used in the TMC. Table 6 shows the AI runtime that compares the three codecs with the HEVC-based codec.

The result shows that the overall video-excluded encoding thread time did not change significantly. However, the video encoding times differ for each codec. For example, 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

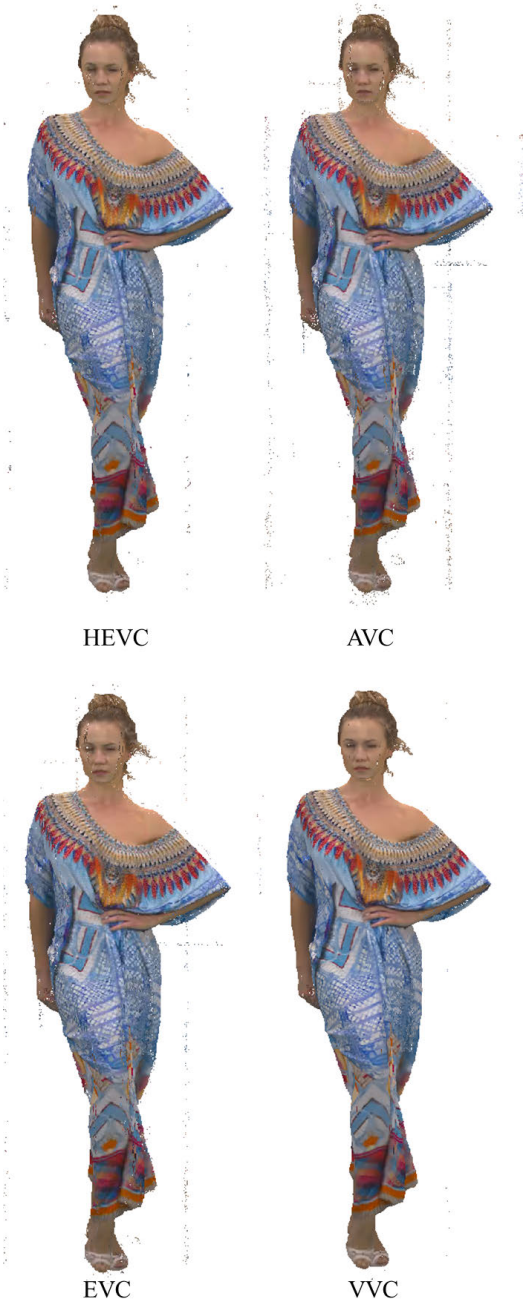

FIGURE 25. Rendering result of low-delay case.

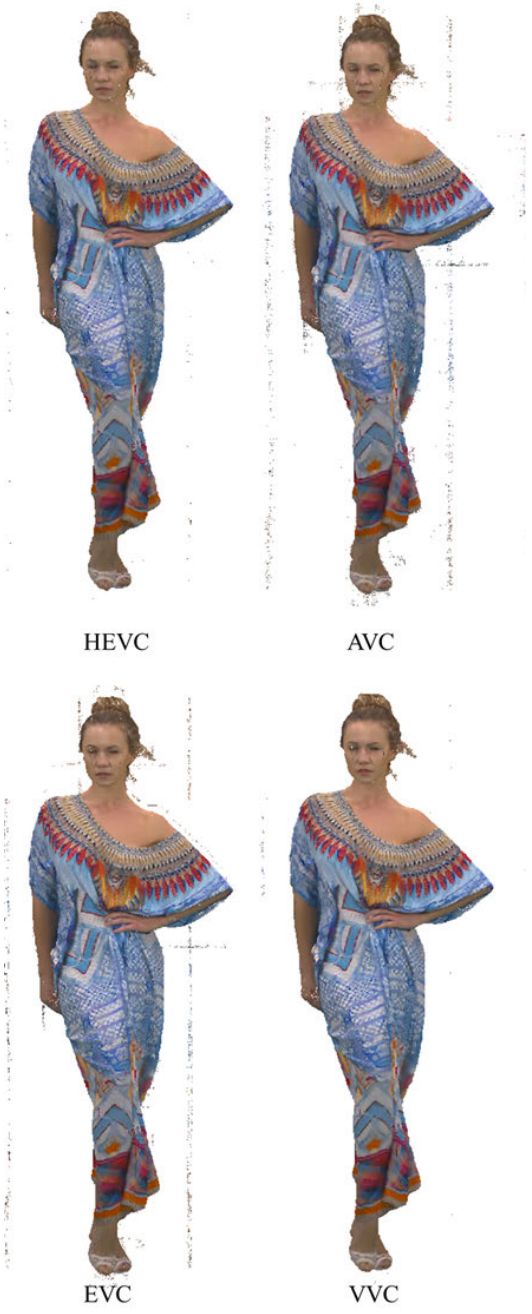

FIGURE 26. Rendering result of random-access case.

the AVC and EVC codec, the decoding time was reduced by 14%–19%.

Table 7 shows that the three codecs used less time for the video-excluded process under the LD condition. Again, EVC showed the shortest video encoding time and AVC showed the longest video encoding time.

Table 8 shows that the three codecs required runtime for the RA condition. 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%. As a result, the AVC codec used 27% less time than the anchor.

5) SUBJECTIVE COMPARISON

The results of the rendered point cloud data were confirmed. As shown in Figs. 24–26, we captured 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 used by MPEG is PCC renderer [16].

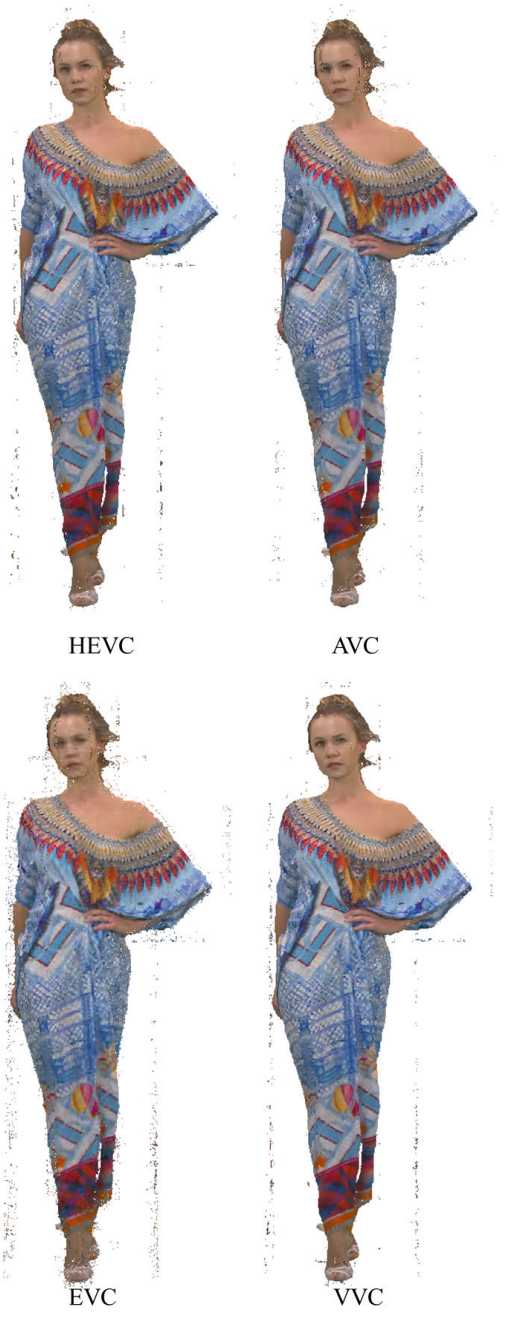

FIGURE 27. Rendering result of 16th frame low-delay case.

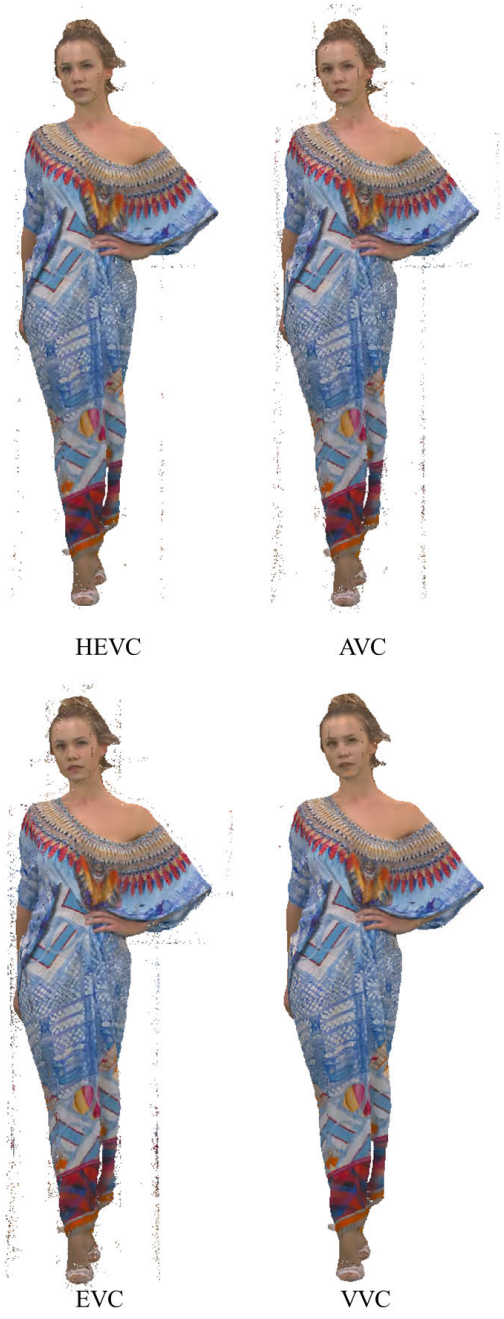

FIGURE 28. Rendering result of 16th frame random-access case.

Fig. 24 shows an image of each codec for the AI conditions. As expected, the lossy OMAP generated many noise points on the edge of the rendered object. VVC showed the best quality, whereas AVC showed the worst.

The LD condition results for each codec are shown in Fig. 25. VVC showed the best quality, which did not match the result of the lowest RD curve, as shown in Fig. 14. However, AVC showed the worst quality with considerable noise when the RD performance of AVC was better than HEVC.

The captured results for the RA condition for each codec are shown in Fig. 26. As expected, VVC showed the best quality and AVC showed the worst quality. However, EVC showed better quality than AVC in terms of the color attributes.

It is known that the interpredicted frame inside a GOP has a different quality, owing to the delta QP. Therefore, we captured the 16th frame for 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 from the RD curve result in Fig. 21, which shows color distortion on the face part of the model. This indicated that the design of the coding structure had a direct impact on the final result.

V. CONCLUSION

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

REFERENCES

[1] *Time of Flight Gets Precise: Enhanced 3D Measurement With Sony DepthSense Technology*, [1] LUCID Vision Labs, Richmond, BC, Canada, 2017.

[2] *Coded Representation of Immersive Media Part 5: Visual Volumetric Video-Based Voding (V3C) and Video-Based Point Cloud Compression (V-PCC)*, Standard ISO/IEC FDIS 23090-5, 2020.

[3] *V-PCC Codec Description VPCC TMv11 Algorithm Description*, Standard ISO/IEC JTC 1/SC29/WG, 3DG, Jun. 2020.

[4] V. Zakharchenko, *[VPCC] [EE4FE 2.7 Report] Evaluation of the Results of Multiple Video Codec Integration in V-PCC Software*, Standard ISO/IEC JTC 1/SC 29/WG, Alpbach, Austria, Apr. 2020.

[5] Y.-T. Tsai, C.-C. 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*, Standard ISO/IEC JTC 1/SC 29/WG, Alpbach, Austria, Apr. 2020.

[6] D. Mehlem and C. Rohlfing, *[VPCC] Versatile Video Coding for VPCC*, Standard ISO/IEC JTC 1/SC 29/WG, Alpbach, Austria, Apr. 2020.

[7] E. Chang, J. Cha, T. Dong, K. Kim, and E. Jang, *On Lossy Occupancy Map Compression in V-PCC*, Standard ISO/IEC JTC 1/SC 29/WG, Jul. 2020.

[8] T. Dong, *MPEG Video-Based Point Cloud Coding Based on JPEG*, document IDW 19, 3DSA, 2019.

[9] Y.-K. Wang and J. Chen, *[EVC] On MPEG-5 EVC Profiles, Tiers, and Levels*, Standard ISO/IEC JTC 1/SC 29/WG, Gothenburg, Sweden, Jul. 2019.

[10] *Infrastructure of Audiovisual Services—Coding of Moving Video (H.266)*, Standard ISO/IEC 23090-3, International Telecommunication Union (ITU-T), 2020.

[11] *Information Technology—General Video Coding—Part 1: Essential Video Coding*, Standard 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 Process. Mag.*, vol. 37, no. 3, pp. 160–167, May 2020, doi: 10.1109/MSP.2020.2971765.

[13] *Common Test Conditions for Essential Video Coding*, Standard ISO/IEC JTC 1/SC 29/WG, Brussels, Belgium, Jan. 2020.

[14] *PCC Requirements*, Standard ISO/IEC JTC 1/SC 29/WG, 3DG, Gwangju, South Korea, Jan. 2018.

[15] *Common Test Conditions for Point Cloud Compression*, Standard ISO/IEC JTC 1/SC 29/WG, 3DG, Gothenburg, Sweden, Jul. 2018.

[16] 3DG. *MPEG 3DG Renderer*. Accessed: Dec. 19, 2021. [Online]. Available: http://mpegx.int-evry.fr/software/MPEG/PCC/mpeg-pcc-renderer

[17] C. Fu, E. Alshina, A. Alshin, and Y. Huang, "Sample adaptive offset in the HEVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1755–1764, Dec. 2012, doi: 10.1109/TCSVT.2012.2221529.

[18] H. Yuan, R. Hamzaoui, F. Neri, and S. Yang, "Model-based rate-distortion optimized video-based point cloud compression with differential evolution," in *Proc. ICIG*, 2021, pp. 735–747, doi: 10.1007/978-3-030-87355-4_61.

TIANYU DONG (Graduate Student Member, IEEE) was born in Linyi, Shandong, China, in 1987. He received the B.S. and M.S. degrees in electronics and computer engineering from Hanyang University, South Korea, in 2011 and 2014, respectively, where he is currently pursuing the Ph.D. degree with the Department of Computer Science.

He started his research on video compression, in 2011. He has authored ten articles on MPEG standardization and five other journal articles and conference papers. His research interests include multimedia-related compression technology and application.

KYUTAE KIM was born in Anseong-si, Gyeonggi-do, South Korea, in 1996. He received the B.S. degree in computer engineering from Gangnam University, South Korea, in 2020. He is currently pursuing the M.S. degree with the Department of Computer Science, Hanyang University.

EUEE S. JANG (Senior Member, IEEE) received the B.S. degree from Jeonbuk National University, South Korea, in 1991, and the Ph.D. degree from SUNY at Buffalo, Buffalo, NY, USA, in 1996.

He is currently a Professor with the Department of Computer Science and Engineering, Hanyang University, Seoul, South Korea. He has authored more than 325 articles on MPEG standardization, 90 journal articles and conference papers, 65 pending or accepted patents, and two book chapters. His research interests include image and video coding, reconfigurable video coding, and computer graphics.

Dr. Jang received three ISO/IEC certificates of appreciation for his contributions to MPEG-4 development. Finally, he received the Presidential Award from the Korean Government for his contribution to MPEG standardization.