Overview of the Video-based Dynamic Mesh Coding (V-DMC) Standard Work

YiHyun Choi, Jong-Beom Jeong, Soonbin Lee, Eun-Seok Ryu
Department of Computer Science Education, Sungkyunkwan University, Seoul, South Korea
Emails: {ceh9141, uof4949, soonbinlee, esryu}@skku.edu

***Abstract*— The moving picture experts group (MPEG) has conducted various studies for standardizing the dynamic mesh coding (DMC) technology. In April 2022, unlike previous mesh compression standards with constant connectivity information, a new one was proposed with time-varying connectivity information, and five institutions responded to that call for proposal (CfP). One of these institutions' proposals was selected as the base model for video-based dynamic mesh coding (V-DMC). Currently, a new V-DMC test model (TM) is being developed based on the proposal. Since the release of the source of the proposal, experiments and feedback on many technical proposals and sources have emerged, and a new V-DMC TM has finally been released. This paper introduces the technical process leading up to the advent of V-DMC TM.**

Keywords— Video-based dynamic mesh coding, Virtual reality, Video-based point cloud compression

I. INTRODUCTION

Recently, with the expansion of the base of XR technology, technologies for XR content including immersive images are developing. Accordingly, several technologies for improving the quality of 3D content such as computer graphics and immersive video have been developed [1]-[4]. However, creating sophisticated XR content requires a model that can optimize large volumes of data. Consequently, technologies that can optimize large amounts of data have been widely researched.

A mesh is the most basic element in a 3D computer graphics model . A mesh is a set of polygons arranged in the shape of a polyhedron; 3D computer graphics models usually use a triangulation mesh as shown in Fig. 1. A mesh contains information about connectivity, geometry, mapping, vertex attributes, and attribute maps. If any of these five components varies over time, the mesh is called as a dynamic mesh. Because dynamic meshes contain much larger amounts of data compared with normal meshes, efficient compression, storage, and transmission of data become important. In light of this, moving picture experts group (MPEG) has been standardized dynamic mesh compression.

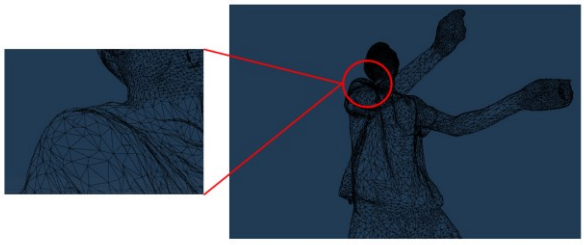

Fig. 1: Structure of a triangulation mesh.

In October 2021, a call for proposal (CfP) for dynamic mesh has been proposed by the MPEG. Unlike previous MPEG mesh compression standards such as frame-based animated mesh compression (FAMC) [5]. which has constant connectivity information with time-varying attribute maps and vertex attributes, a new standardization was proposed for dynamic meshes with time-varying connectivity information.

Apple, Interdigital, Nokia, Tencent, and Sony responded to this proposal [6], out of which proposal of Apple was adopted as a dynamic mesh coding (DMC) model because it performed best. On May 6, 2022, Apple released a source called P11. As the source of the adopted institution was disclosed, new research, feedback, and improvements to the source were actively proposed and discussed. Finally, in July 2022, a test model of a new video-based dynamic mesh coding (V-DMC) was released by referring to P11.

This paper introduces the process and trends of V-DMC, and the test model of V-DMC. First, the CfP, an anchor released by the MPEG and the encoding technology of proposal of Apple is introduced. Next, this paper introduces the P11 source of May 6, 2022, and the architecture of the V-DMC test model (TM) which was developed based on the P11. Finally, the proposals and feedbacks to P11 and V-DMC TM since P11 was released to the present is introduced.

II. BACKGROUND

This section describes the CfP released by the MPEG and the anchor provided for the performance evaluation. In addition, the technology of proposal of Apple, which was recognized for its outstanding performance compared with other responses to the CfP, is explained.

A. Call for Proposal of MPEG WG 7

According to the CfP of DMC released by the MPEG, the proposals for the CfP should be based on the V3C framework and specifications and required that the developed software reference the high efficiency video coding (HEVC) standard and that HEVC test model (HM) 16.21-screen content coding test model (SCM) 8.8 be used for video encoding. Futhermore, the MPEG provided an anchor for dynamic mesh compression for performance evaluation. The anchor encoded mesh using Draco and texture maps using HDRTools in a video format, using HM 16.21 with screen content coding extension (SCC) 8.8. Fig. 2 shows the anchor.

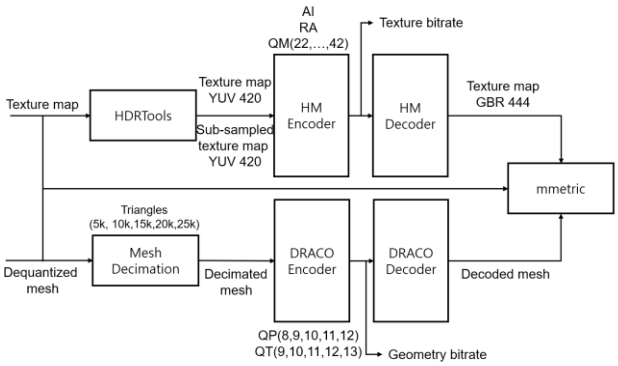

Fig. 2: V-DMC lossy anchor.

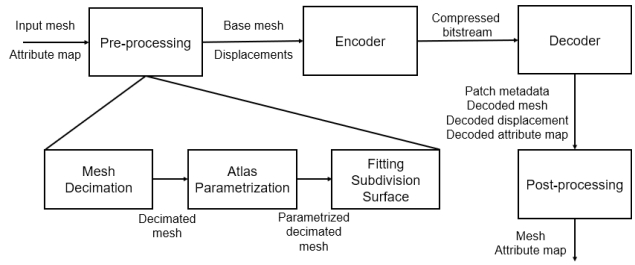

Fig. 3: Block diagram of the proposal by Apple.

B. *Response by Apple*

A high level diagram of proposal of Apple is shown in Fig. 3. Through the pre-processing process, the displacement information is used to create a decimated mesh that drastically decreases the amount of vertices. Subsequently, the mesh information is encoded using Draco, and the displacement information is converted into a wavelet transform. The wavelet transform is then converted to a set of wavelet coefficients, which are then quantized, and packed into a 2D image to enable video encoding. Proposal of Apple differs significantly from that of other institutions in that according to it, the decimated mesh is converted to the original mesh using the displacement vector values. In other words, pre-processing significantly contributes to compression efficiency.

The pre-processing method proposed by Apple is shown in Fig. 4. First, the original mesh data are decimated, and the number of vertices of the mesh is greatly reduced. Due to that, the data size of the mesh is also greatly reduced. The decimated mesh is then subdivided to generate iterations between vertices using a subdivision scheme. Proposal of Apple adopts the mid-point subdivision scheme, which places iterations in the middle of the two vertex coordinates. These iterations are displaced such that the decimated mesh is similar to the original mesh, and the displacement information is stored. Through this pre-processing, the proposal of Apple gets a high compression efficiency for DMC.

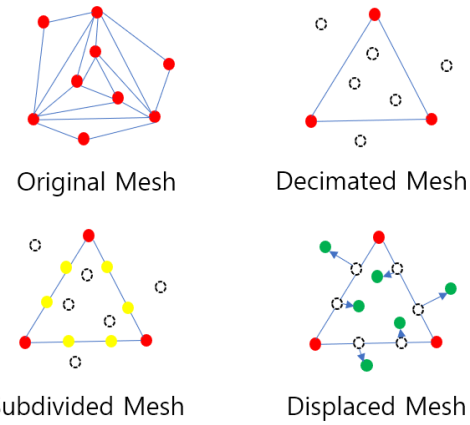

Fig. 4: Pre-processing method of Apple.

During decoding, the decoder can subdivide the decimated mesh without subdivision information because the subdivision scheme is set to the mid-point subdivision scheme. Thus, the decoder can generate the original mesh using only the decimated mesh and the displacement information. As a result, the decimated mesh has a much smaller data size than the original mesh. The displacement information is also stored in the form of a wavelet transform, which exhibits a high compression efficiency during the encoding process. As a result, the proposed method by Apple can achieve high compression efficiency because the original mesh can be reconstructed using only the reduced mesh data and displacement information.

Compression method of Apple achieved -15.6% better geometry peak signal-to-noise ratio (PSNR) and -55.5% better YPSNR compared with the anchor [6]. The evaluation figures are based on interframe encoding of random-access image-based rendering.

III. V-DMC TM AND P11

Five institutions responded to the V-DMC CfP, of which Apple and InterDigital showed higher compression performance than the anchor proposed by the MPEG. As a result, at the 138th meeting of the MPEG, proposal by Apple was chosen as a base for the new test model of V-DMC. The source code by Apple was then submitted and released on May 6, 2022. This section introduces P11 and V-MDC test models.

A. *Source Code of Apple (P11)*

The architecture of the P11 consists of several applications, and the result of encoding and decoding can be obtained through several processes that utilize the obtained output as an input for other executable files.

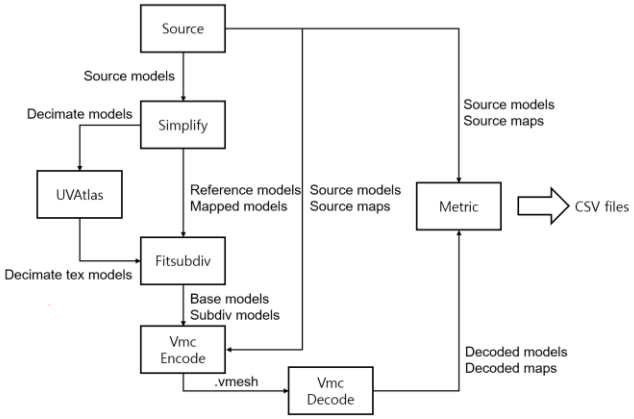

Fig. 5: Block diagram of intra encoding in P11.

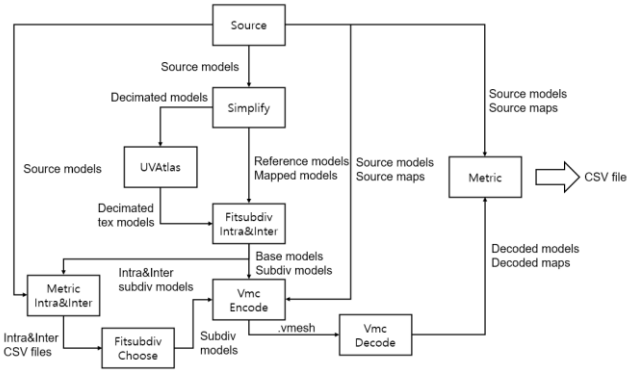

Fig. 6: Block diagram of low delay encoding in P11.

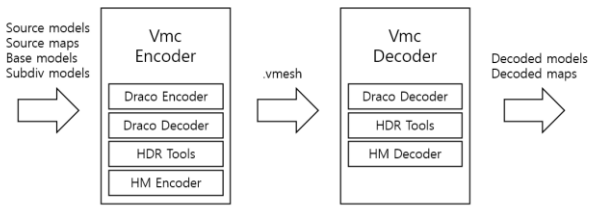

Fig. 7: Block diagram of Vmc encoding and decoding in P11.

Figs. 5 and 6 show the flow of the input and output between applications in intra and low-delay encoding. Simply, Geonof, UVAtlas, and Fitsubdiv performed proposed pre-processing process of Apple. In pre-processing, the base model data and subdiv model data are obtained and encoded. The output of encoding is in a .-vmesh format, which is decoded via the decoder to obtain the decoded model data and decoded map data. Encoding results can be obtained by comparing these data with the original model and map data.

Vmc encoder and decoder are encoding and decoding tools of P11. Vmc encoder and decoder are applications that perform encoding and decoding of P11. Diagrams of the Vmc encoder and decoder are shown in Fig. 7.

The Vmc encoder includes a Draco encoder, Draco decoder, HDRTools, and an HM encoder. The reason why the Draco decoder is included in the Vmc encoder is that the mesh data compressed by the Draco encoder is not only used as output, but is also reconstructed by the Draco decoder. The Draco decoder regenerates the original mesh through the reconstructed decimated mesh and reconstructed displacement information during the encoding process. The

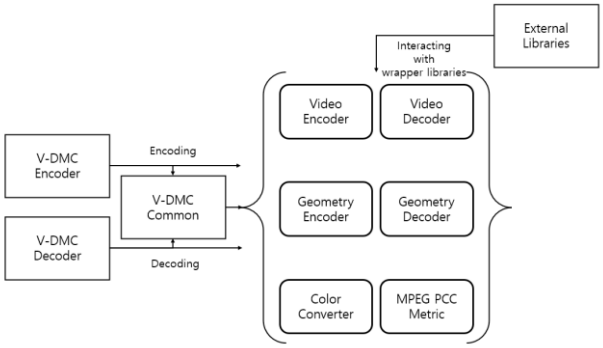

Fig. 8: V-DMC test model architecture.

reconstructed original mesh is used to compress the attribute map.Base models, which are decimated meshes obtained through pre-processing, and subdiv models with displacement information are used as inputs for encoding. A reconstructed mesh is used to update the displacement information and convert it into a wavelet transform.

Since release of P11, it has been widely researched. In fact, several researchers have released technical advancement proposals and feedback for P11. Many have highlighted that P11 creates a large number of input/output processes. Although input/output processes are helpful in checking the output of different applications, a large amount of input/output data when generated can consume a large amount of disk storage capacity and time. The MPEG also recognized these problems and claimed that they would be improved in test model of the V-DMC.

B. V-DMC TM Arcitecture

MPEG restructured the best proposal of the last CfP and released a test model of the new V-DMC on July 11, 2022. In the process of developing the test model, P11 was analyzed in detail to identify technological improvements, which were then studied them to ensure smooth DMC advancement in the future.

The V-DMC test model of the MPEG is shown in Fig. 8. First, they fixed the problems of P11 which structure is consist of many applications. The existing sub-application is integrated into one, and the application originally used in encoding and decoding is defined as a wrapper library and an external library. The wrapper library was newly developed with a video encoder and decoder consisting of HM, versatile video coding test model (VTM), FFMPEG libraries, geometry encoder, and decoder consisting of Draco and triangle fan-based compression (TFAN), and color converter with HDRTools contents. In addition, other applications of P11, as shown in Figs. 5 and 6, were also defined as external libraries and integrated into one application. The Gengof application was defined as the Sequence Information library, the Simplify application was defined as the Geometry Decision library. And the UVAtlas application was defined as the Textual Parametrization library, the Fitsubdiv application was defined as the Geometry Parameter library. Gengof is part of P11 and generates an initial group of frames generation. Consequently, the new V-DMC test model has three core codecs, namely, vdmcCommon, vdmcEncoder, and vdmcDecoder, which are easily accessible by referencing libraries.

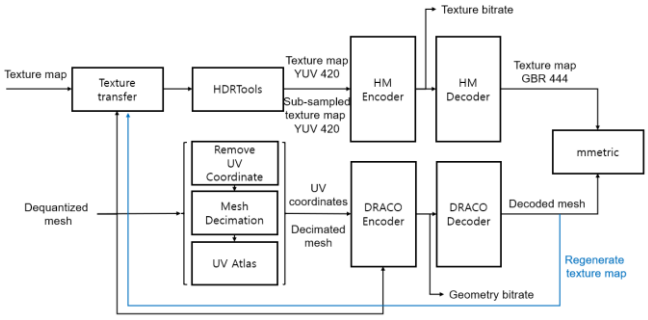

Fig. 9: Proposed new anchor.

IV. OTHER PROPOSALS

Many feedbacks on P11 were proposed from that the proposal was adopted to that the test model of V-DMC was released. This section introduces the technical improvements, feedback, and cross-checked contents of P11 proposed by other researchers.

A. New Anchor Proposal

As shown in Fig. 2, the anchor provided by MPEG at the time of CfP was encoded and decoded using Draco for geometry data compression and HEVC for texture data compression. Mesh decimation and down-sampling are included to reduce the bitrate. Because both processes operate independently and simultaneously, mesh decimation technology that preserved the textrue coordinates is required to be able to use the original texture map.

A further improved anchor has been proposed to address this issue and to efficiently evaluate the performance of DMC to be developed in the future. Proponent introduced that improved anchor uses texture map regeneration to generate texture maps more efficiently. They applied the pre-processing pipeline of P11, which obtained the best performance evaluation of the CfP responses to the new anchor.

The proposed anchor is shown in Fig. 9. In the previous anchor, decimation was not used in some sequences showing a higher bitrates than the others. In the case of a new anchor, all sequences can be decimated by deleting UV coordinates, regenerating texture maps with new UV coordinates obtained by the UV Atlas, and transferring textures using Meshlab. Proposers updated Draco to its latest version and quantized it using an external application as shown in P11. Finally, they modified the subsampling factor and texture QP to adjust it to be similarly to the total bitrate of P11.

B. Feedback for P11

In June 2022, Nokia, Sony, and Tencent released documents related to the P11 analysis and review. As other researchers have pointed out, too many applications require too much time and resources. They reported that there were multiple file read-write operation problems in P11, which can be a fatal drawback. In addition, certain compilers, such as Clang++-13 and complication instructions for Windows platforms, have been reported to be unavailable or have different results from those of P11.

V. CONLUSION

Five institutions responded to the CfP for DMC by MPEG and proposal of Apple had Y-PSNR, 55.5 percent higher than anchor, and Geometry PSNR, 26.3 percent higher in Inter, random access rendering. And proposal of Apple was adopted as the base model of the new V-DMC test model. Subsequently, the initial V-DMC test model was released in July 2022, referring to the open source model by Apple, P11. Accordingly, proposals and experiments for technical improvement of the P11 and V-DMC test models are being actively researched, and these proposals will be reviewed and researched in the future to upgrade the V-DMC TM and DMC technology.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2022R1F1A1074935).

REFERENCES

[1] JongBeom Jeong, Dongmin Jang, Jang-Woo Son, Eun-Seok Ryu, “Bitrate efficient 3DoF+ 360 video view synthesis for immersive VR video streaming”, In Proceedings of the 2018 International Conference on Information and Communication Technology Convergence, pp. 581-586, 2018.

[2] Jong-Beom Jeong, Soonbin Lee, Dongmin Jang, Eun-Seok Ryu, “Towards 3DoF+ 360 Video Streaming System for Immersive Media”, IEEE Access, vol. 7, pp. 136399-136408. 2019.

[3] Soonbin Lee, Jong-Beom Jeong, Inae Kim, Sangsoon Lee, Eun-Seok Ryu, “Screen Content Coding Analysis to Improve Coding Efficiency for Immersive Video”, Journal of Broadcast Engineering, vol. 25, no. 6, pp. 911-921, 2020.

[4] Soonbin Lee, Jong-Beom Jeong, Eun-Seok Ryu, “Efficient Group-Based Packing Strategy for 6DoF Immersive Video Streaming”, In Proceedings of the 2022 International Conference on Information Networking, pp. 310-314, 2022.

[5] Khaled Mamou, Detlev Marpe, Karsten Müller, Titus Zaharia, Françoise Prêteux, “Frame-based Animated Mesh Compression: integration of the CABAC arithmetic encoder”, ISO/IEC JTC 1/SC 29/WG 11, MPEG/m14493, 2007

[6] YiHyun Choi, Jong-Beom Jeong, Soonbin Lee, Eun-Seok Ryu, “Overview of MPEG Dynamic Mesh Coding (DMC) for Volumetric Video”, The Korean Institute of Broadcast and Media Engineers (KIBME) Summer Conference, Jun. 20-22, 2022.

[7] WG 07 MPEG 3D Graphics coding, “CfP for Dynamic Mesh Coding”, ISO/IEC JTC 1/SC 29/WG 7,MPEG/N231, 2021.

[8] Khaled Mamou, Jungsun Kim, Alexis Tourapis, Dimitri Podborski, Krasimir Kolarov, “Apple’s Dynamic Mesh Coding CfP Response”, ISO/IEC JTC1/SC29/WG7, MPEG/m59281, 2022.

[9] Danillo B. Graziosi, Ali Tabatabai, Alexandre Zaghetto, Satoru Kuma, Kao Hayashi, Rajan Joshi, Madhukar Budagavi, “[V-DMC] On anchor improvement inversigation”, ISO/IEC JTC 1/SC 29/WG 7, MPEG/m60119, 2022.

[10] Sebastian Schwarz, Patrice Rondao Alface, Aleksei Martemianov, Danillo Graziosi, Ali Tabatabai, Satoru Kuma, Kao Hayashi, Jun Tian, Xiang Zhang, “Input contribution to EE 4.2 Investigation on V-DMC source code implementation”, ISO/IEC JTC 1/SC 29/WG 7, MPEG/m59982, 2022.

[11] Vladyslav Zakharchenko, “[V-DMC][EE4.1] Report on suggested tool testing for V-DMC”, ISO/IEC JTC 1/SC 29/WG 7, MPEG/m60148, 2022.

[12] François-Louis Tariolle, Maja Krivokuća, Julien Ricard, Olivier Mocquard, “[V-DMC] InterDigital cross-check of Apple CfP proposal results using the provided mpeg-vmesh-tm code”, ISO/IEC JTC 1/SC 29/WG 7, MPEG/60254, 2022.

[13] David Flynn, Khaled Mammou, “uvatlas convergence and P11 comparison”, ISO/IEC JTC1/SC29/WG11, MPEG/60562, 2022.

[14] David Flynn, Khaled Mammou, “V-DMC: removing per-frame mode decision pessimism”, ISO/IEC JTC1/SC29/WG11, MPEG