WHITE PAPER

The MPEG Dynamic Mesh Coding Standard

By Chao Cao
Senior Video Compression Engineer

Abstract

In recent years, 3D content has become increasingly ubiquitous, enabling immersive experiences across several platforms and devices. 3D meshes, which offer a versatile and efficient representation to accurately represent immersive 3D content, are widely used across a wide range of industries. A new Dynamic Mesh coding standard is currently under development as part 29 of the MPEG-I standard. In this paper, we give an overview of the ongoing MPEG standard and hope to bring the readers some insight into the future of immersive media.

Introduction

Background

With the advancements in 3D capture, modeling, and rendering technologies, 3D content has become increasingly ubiquitous, enabling immersive experiences across several platforms and devices. 3D meshes are one of these immersive contents that are widely used in the commercial market. As the complexity of models continues to increase to achieve the desired level of realism, a significant amount of data is associated with their creation and consumption. Achieving efficient compression of such content is of utmost importance to satisfy various requirements of applications.

Many mesh compression standards such as IC, MESHGRID, FAMC [1] were previously developed by MPEG to address dynamic meshes with constant connectivity and time varying geometry and vertex attributes. However, these standards do not consider time varying attribute maps and connectivity information. In recent years, an increasing amount of dynamic meshes are generated by using DCC (Digital Content Creation) tools and volumetric acquisition techniques such as [2]. This type of content is not supported by the existing standards.

In October 2021, a call for proposals was initiated by MPEG for Dynamic Mesh coding technology, targeting an efficient representation of dynamic objects and real-time acquisition environments and to directly handle dynamic meshes with time varying connectivity information and optionally time varying attribute maps. This standard targets lossy, and lossless compression for various applications, such as real-time communications, storage, free viewpoint video, AR and VR. Functionalities such as random access and scalable/progressive coding are also considered [3].

In April 2022, one of the proposals using video-based technology was adopted and became the new test model of the Video-based Dynamic Mesh Coding standard (V-DMC).

❖ Static and Dynamic Mesh

A mesh is composed of several polygons that describe the boundary surface of a volumetric object. Each polygon is defined by its vertices in 3D space and the information on how the vertices are connected, referred to as connectivity information. Optionally, various attributes, such as colors, normals, etc., could be associated with the mesh vertices or with the surface of the mesh. A static mesh represents a mesh that does not change in time while a dynamic mesh sequence represents a collection of static meshes that are changing in time. The change can be on its geometrical property or on its attributes such as color. For example, a mesh can be animated, morphed, or manipulated to simulate movement. Meanwhile, the color on a mesh surface can be different with various light conditions or shadings.

V-DMC workflow

As shown in Figure 1, a simplified version of the V-DMC workflow consists of pre-processing, encoding, decoding and post-processing stages. The input mesh can be a single static mesh or a sequence of dynamic mesh. The pre-processing stage simplifies the input mesh into a sparser version of the original mesh with associated information. After being encoded and decoded, a post-processing stage is used to improve the quality of the reconstructed mesh as the final output mesh.

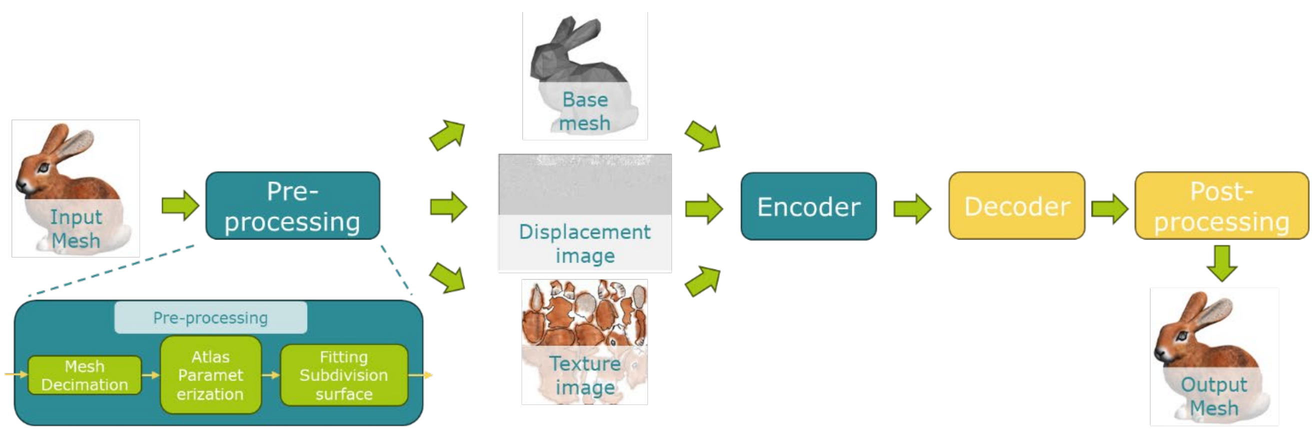


Figure 1. V-DMC workflow.

1) Pre-processing stage

The pre-processing stage consists of three main processes: a decimation step, an atlas parameterization step, and a subdivision step. As shown in Figure 2, the "Mesh Decimation" simplifies the "Original" mesh into a sparser mesh noted as "Decimated"/"Base" mesh with fewer number of vertices and faces while maintaining its original geometric properties depending on the requirements of the application. Then, the "Atlas Parameterization" creates a parameterization of the input 3D mesh onto a 2D texture space for texture mapping. It helps to map the original 2D textures to the compressed representation.

The “Fitting Subdivision surface” step subdivides the “Decimated” mesh into a “Subdivided” mesh which is a denser mesh for predicting the original input mesh. The “Subdivided” mesh is then deformed to match the detailed geometric properties of the original input mesh as “Displaced” Mesh. A displacement map is computed from the “Deformed” mesh and the “Subdivided” mesh as the residual of the prediction and later represented as a 2D displacement image. Finally, the texture from the original input mesh is mapped onto a reparametrized 2D texture image. A 3D “Base” mesh, a 2D displacement image and a 2D texture image are compressed individually by the encoder.

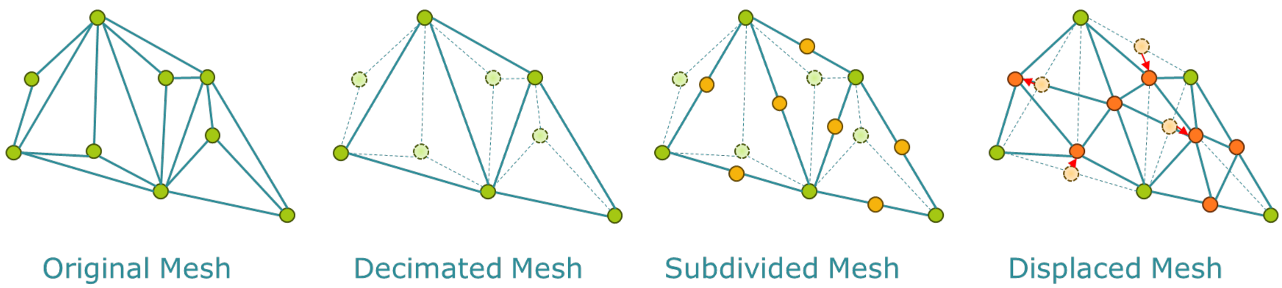


Figure 2. Pre-processing example of a mesh.

2) Encoding and decoding stage

In order to compress the mesh and its associated information efficiently, different codecs are used as shown in Figure 3. For example, a mesh codec called DRACO [4] from Google was used to code the mesh in the test model before version 4.0 and got replaced by an MPEG EdgeBreaker implementation [5]. Different video codecs such as the HM [6] and VTM [7] reference software can be used to code the displacement and texture images. Additionally, a motion codec can also be used to compress the motion data of the mesh when in inter coding mode.

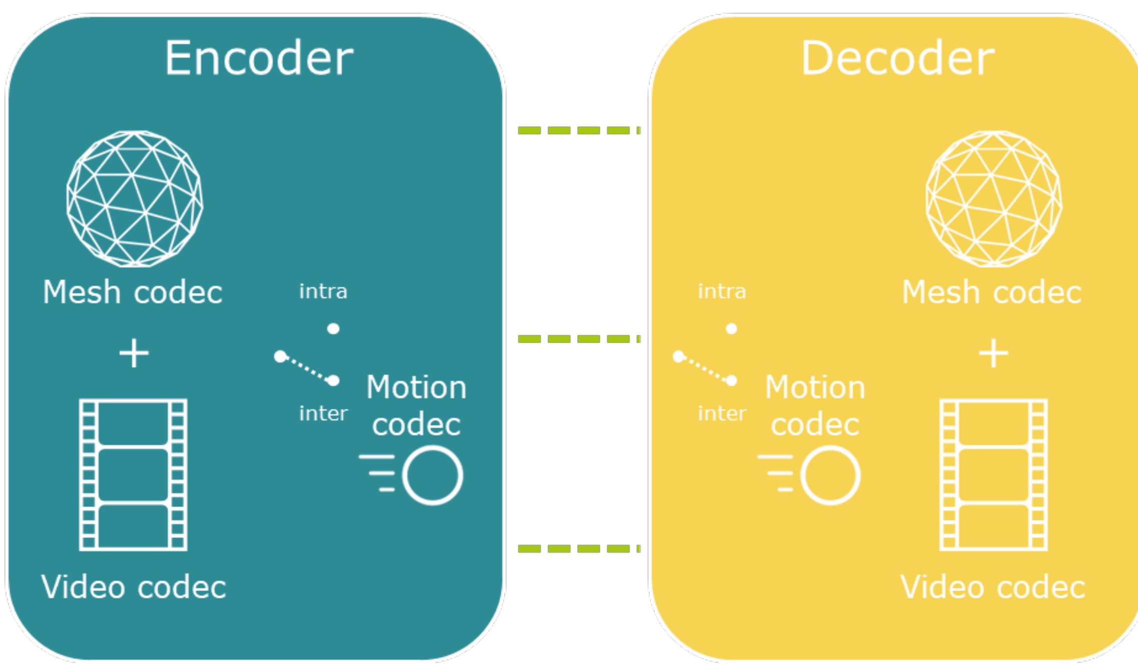


Figure 3. The encoder and decoder of the V-DMC software.

3) Post-processing stage

The reconstructed mesh is further refined at the post-processing stage for better visual quality or other desired properties. However, it is still an early stage of the V-DMC development, and not much effort has been focused on this process. In the future development, various technologies may be foreseen such as smoothing, texture remapping, connectivity optimization, etc.

Conclusion

In this paper, we have briefly reviewed the workflow of the V-DMC standard. The main advantage of the V-DMC is that it takes advantage of the existing Visual Volumetric Video-based Coding (V3C) standard [8] which can greatly shorten the development duration of the standard so that it can be available for the commercial market. Moreover, it will evolve with the advancement of existing video coding standards and mesh coding technologies.

It is exciting to see tremendous innovative techniques have opened new possibilities for immersive experiences by enabling efficient storage and transmission of 3D mesh data. We can anticipate even more astonishing immersive experiences that push the boundaries of what is possible, captivating users in increasingly realistic and engaging virtual and augmented environments.

References

[1] Information technology - Coding of audio-visual objects - Part 16: Animation Framework eXtension (AFX), ISO/IEC 14496-16.

[2] A. Collet, M. Chuang, P. Sweeney, D. Gillet, D. Evseev, D. Calabrese, H. Hoppe, A. Kirk and S. Sullivan, "High-quality streamable free-viewpoint video," in ACM Transaction on Graphics (SIGGRAPH), 2015.

[3] CfP for Dynamic Mesh Coding, ISO/IEC JTC 1/SC 29/WG 7/N00231, Online, October 2021.

[4] Google Draco, 01 2021, [online] Available: htps://google.github.io/draco/.

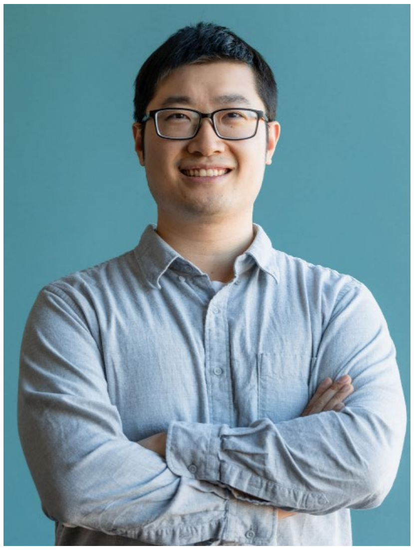
[5] Edgebreaker base mesh codec integration, ISO/IEC JTC 1/SC 29/WG 7/m62603, Antalya, Turkey, April 2023.

[6] HM reference software for HEVC, 06 2023, [online] Available: htps://vcgit.hhi.fraunhofer.de/jvet/HM.

[7] VTM reference software for VVC, 06 2023, [online] Available: htps://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM.

[8] Information Technology—Coded Representation of Immersive Media—Part 5: Visual Volumetric Video-Based Coding (V3C) and Video-Based Point Cloud Compression (V-PCC), Standard ISO/IEC 23090-5, N19329, May 2020.

About the Author



Chao Cao received a bachelor's degree in Telecommunication Engineering from Shanghai University, Shanghai, China, in 2013, a master's degree in multimedia networking from the Université Paris-Saclay, Paris, France, in 2018 and a Ph.D. degree in point cloud compression at Télécom SudParis, Institut Polytechnique de Paris, Évry, France, in 2022. He has been participating and contributing to the moving picture experts group (MPEG) since 2018. He is currently a Senior Video Compression Software Engineer in Ofinno focusing on the technologies for video-based dynamic mesh coding, a coming standard of MPEG-I part 29. His research interests include image, video, point cloud and mesh compression.

Chao Cao

Senior Video Compression Software Engineer, Ofinno