

## 볼류메트릭 비디오를 위한 MPEG Dynamic Mesh Coding (DMC) 표준화 동향

최이현, 정종범, 이순빈, 류은석

성균관대학교 컴퓨터교육학과

ceh9141@skku.edu, uof4949@skku.edu, soonbinlee@skku.edu, esryu@skku.edu

### Overview of MPEG Dynamic Mesh Coding (DMC) for Volumetric Video

YiHyun Choi, Jong-Beom Jeong, Soonbin Lee, Eun-Seok Ryu

Department of Computer Science Education, Sungkyunkwan University

#### 요 약

MPEG 에서는 frame-based animated mesh compression(FAMC)와 같은 동적 메쉬를 압축하는 기준을 만들어왔다. 그러나 이러한 동적 메쉬 압축 표준은 시변 연결성(time-varying) 정보가 아닌 일정한 연결성 정보를 가진 메쉬를 기준으로 개발되었기 때문에 실시간 스트리밍에서 사용하기 부적합하다. 따라서, moving picture experts group(MPEG)에서는 시변 연결성 정보를 사용하는 동적 메쉬 압축에 대한 새로운 표준을 제안하였는데, 본 논문에서는 MPEG 이 제의한 call for proposals (CfP)에 대해 여러 기관들이 제안한 메쉬 압축 기술들을 소개한다.

#### 1. 서론

가상 현실(Virtual Reality, VR)과 증강 현실(Augmented Reality, AR)을 포함한 확장현실(eXtended Reality, XR)은 다분야에서 융합기술로서의 활용성을 주목받아 급속도로 성장하고 있는 추세이며 이에 따라 몰입형 영상을 효율적으로 처리하기 위한 기술이 활발히 연구되고 있다[1-4]. 컴퓨터 그래픽스를 기본으로 실제 3D 환경을 구축하는 가상현실 콘텐츠는 물론이고 360 실사 영상을 Head Mounted Display(HMD)를 통해 제공하는 3 Degrees of Freedom(3DoF), 6 Degrees of Freedom(6DoF) 몰입형 영상 처리 과정에도 메쉬가 사용될 수 있다. 메쉬는 다면체의 형태를 구성하는 다각형들의 집합이며, 그래픽스를 비롯한 많은 3D 업계에서는 삼각형으로 이루어진 Triangulation Mesh 를 사용한다. 그림 1 에는 메쉬에 대한 예시가 이러한 메쉬는 컴퓨터 그래픽스를 기반으로 360 VR, 3D Capture, 3D Graphics 등 3D 부분에서 폭넓게 사용되고 있으며 컴퓨터 기술과 그래픽스 기술이 발전하며 더욱 정교한 모델을 표현할 수 있게 되면서 렌더링, 부호화, 복호화 과정에서 더 복잡한 연산을 요구하게 됐다.

메시 데이터는 Connectivity, geometry, mapping, 버텍스 attributes, attribute maps 정보로 구성되며 이 5 개의 정보 중 하나라도 시변 속성(time varying attribute)를 가지고 있다면 이를 동적 메쉬 (dynamic mesh) 로 규정한다. 점차 정교해지고 복잡해지는 메쉬 데이터를 효율적으로 저장, 실시간 전송, 부호화 및 복호화하기 위해서 MPEG 에서는 이에 대한 FAMC 같은 메쉬 압축 표준 기술을 연구해왔다. 그러나, 이러한 표준들은 실시간

스트리밍 시 연속적인 동적 메쉬 생성이 힘들고 충분한 압축률을 제공하지 못해 MPEG 에서는 새로운 표준 연구를 진행하고 있다. 이에 ISO/IEC SC29/WG7 에서는 시변 연관성 정보를 가지고 있는 동적 메쉬를 주제로 call for proposals (CfP) 를 제안하였고 그에 응해 5 개의 문서가 제출되었다. 본 논문에서는 이러한 연구 주제에 대해 여러 기관에서 제시한 동적 메쉬 압축의 기술에 대해 소개한다.

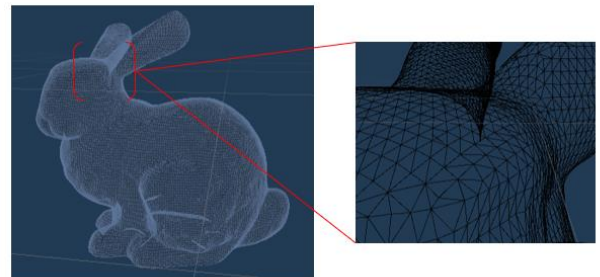


그림 1. Mesh WireFrame

#### 2. 동향

본 절에서는 MPEG의 동적 메쉬 압축 Cfp 에 응한 Apple, InterDigital, Sony, Nokia, Tencent 가 제시한 압축 기술에 대해 소개한다.

##### 2.1 Apple

Apple 은 메쉬 압축 과정에서 메쉬 데이터, attribute map 을 input 으로 요구한다.

전처리 (pre-processing) 과정은 메쉬를 인코딩하기 전에 압축 효율을 올리기 위해 메쉬 데이터를 조정하는 과정이다. Apple 에서는 전처리 과정에서 메쉬 데이터를 줄이고 (decimation), 버텍스 간에 정점을 두고 (subdivision) 그 정점들을 변형 (displace) 하여 원래 메쉬와 비슷한 구조를 재현한다. 데이터가 적어진 메쉬 데이터와 변형 벡터 (displacement vector)만 가지고도 원래의 메쉬를 유사하게 재현할 수 있기 때문에 압축효율을 크게 증가시킬 수 있었다.

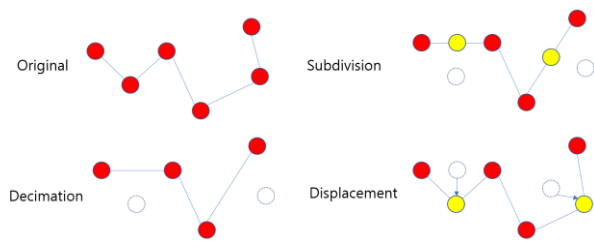


그림 2. Apple 의 Mesh Sampling

메쉬의 인코딩 코덱으로는 오픈소스인 Draco 메쉬 코덱을 선택했다. inter 메쉬 인코딩의 경우에 참조 메쉬와 관련있는 다른 프레임의 메쉬간의 버텍스 좌표 차이들의 집합인 모션필드 (motion field) 를 생성하였다. 이런 모션 필드를 인코딩 결과물에 포함시킴으로써 디코딩 과정에서 참조 메쉬와 모션필드를 사용해서 관련있는 프레임의 메쉬 데이터를 얻을 수 있게 되었다.

최종적으로 메쉬의 인코딩의 최종 산출 결과물은 압축된 메쉬 데이터/변형 (displacement) 비트스트림, 그리고 텍스처 정보를 담은 attribute 비트스트림을 합친 비트스트림이 되었다.

## 2.2 InterDigital

InterDigital 은 UV atlas 생성하고 압축하는 기술에 집중하여 제안하였다. 전처리과정에서 decimation 된 메쉬로 Boundary first flattening (BFF) 를 사용하여 비디오 압축에서 일반적으로 사용하는 group of pictures (GOP) 수준으로 UV 패치를 생성한다. 그 후, 생성된 UV 패치를 타일로 분할하며 그 각 타일은 원래의 텍스처를 재투영한 정보를 가지게 된다. 그 후, 첫번째 프레임의 타일을 인접한 다른 타일과의 차이를 최소화시키기 위해 재정렬하고 부호화 효율을 높이기 위해 화면 간 추적을 시행한다. 이를 통해 원래의 텍스처 atlas 보다 더 압축에 효율적인 타일 텍스처 atlas 를 사용하게 되었으며 부호화된 UV 좌표는 복호화된 메쉬 geometry 정보만을 이용해서 복호화하기 때문에 UV 좌표를 비트스트림으로 인코딩할 필요가 없어졌다.

메쉬 부호화 과정에서 quadratic error metrics(QSlim)를 사용하여 decimation 수행하고 구글의 오픈소스인 Draco 를 사용하여 양자화 및 메쉬 압축을 진행하였다.

## 2.3 Nokia

Nokia 는 메쉬 압축 과정에서 메쉬 정보, 재질 (material), 텍스처를 input 으로 요구한다. 전처리에서 중복된 버텍스들을 합치고, 같은 좌표를 가지지만 같은 법선 벡터 또는 텍스처를 가지지 않는 버텍스를 제거한다. 연관성 있는 면들을 군집화하고 그 군집들을 병합한다. 일정한 샘플링 비율에 맞게 크기를 조정하고 연속적인 프레임의 패치가 하나의 참조 패치를 기준으로 낮은 차이를 가지도록 재정렬한다. 생성된 패치를 기준으로 atlas 를 생성하고 geometry, occupancy, attribute 정보를 가지고 있는 비트스트림을 생성한다.

최종적으로, atlas, geometry, occupancy 비트스트림, attribute 비트스트림을 인코딩의 최종 산출 결과물로 얻는다.

## 2.4 Tencent

Tencent 는 메쉬의 중복되고 독립된 버텍스와 비정상적인 삼각형을 제거하고 스무싱 필터(smoothing filters)를 통해 메쉬의 밀집도를 조절하여 효율적으로 UV atlas 를 생성한다. UV atlas 를 샘플링하여 버텍스의 위치와 UV 정보를 가지고 있는 UV 차트를 만든다. 그 차트로부터 경계선 추출 및 분류 (chart boundary extraction & classification)하여 적절한 샘플링과 패킹을 거쳐 Geometry atlas 를 생성하게 되는데, 이 Geometry image 는 video codecs 을 통해 압축되게 된다. 그 후, 경계선 버텍스의 위치와 UV 정보를 가지고 있는 경계선 버텍스를 양자화 및 엔트로피 코딩하여 메쉬 데이터에 대한 인코딩을 마무리한다.

최종적으로는 최종산출 결과물은 엔트로피 인코딩된 차트 경계선 정보와 비디오 인코딩된 Geometry image, attribute image 정보가 된다.

## 2.5 Sony

Sony 는 기존의 V3C 표준에서 개발한 도구들 중 사용할 적절한 도구를 선택하고 수치와 함수를 조정했다. Sony 가 그 전의 CfP 에서 제안했던 메쉬 추적과 패치에 관련된 패치 타입을 MIV 에서 수행한 것과 유사하게 동적 메시에 적합하게 확장하였다. 패치 생성 과정에 깊이 필터링 블록을 추가하여 2D 공간에서의 연결성을 강화하고 동시에 원래의 UV 좌표는 유지할

수 있도록 확인 과정을 추가했다. 연결성 패치 과정에서 메쉬의 삼각형의 수에 따라 효율이 다르므로 MPEG 의 SC3DMC 와 구글의 Draco 를 통합하여 둘 다 사용하였다. 메쉬 압축 효율을 증가시키기 위해 메쉬 단순화(Simplification)를 시행한다. 이 때, 단순화로 인해 메쉬의 질이 떨어지지 않도록 패치 과정에는 단순화된 메쉬를 사용하고 깊이 맵을 생성할 때는 원래의 메쉬의 깊이 정보와 단순화된 메쉬의 깊이 정보를 이용해서 최종적으로 산출해낼 깊이 정보를 패치하였다. 또한, 디코딩 과정에서 면 세분화(face subdivision)를 적용하여 재구성된 면들이 원래와 근접하도록 하였다.

### 3. 평가

Image-based Rendering 와 InterFrame Encoding 을 기준으로 기관별 성능평가는 다음과 같다.

Institution	Geometry_PSNR	Y-PSNR
Apple	-15.6%	-55.5%
InterDigital	-26.3%	-18.9%
Nokia	78.5%	101.6%
Tencent	61.0%	-8.3%
Sony	135.6%	93.7%

표 1. Inter, Random Access Image-based Rendering

표시되는 퍼센트는 같은 메쉬를 대상으로 각 기관이 제시한 압축 결과물과 비교 기준으로 제시한 압축 과정의 결과물을 비교한 것이다. 퍼센트가 마이너스를 기록하면 기준 압축보다 더 좋은 압축 결과를 얻은 것으로 볼 수 있다.

기준이 되는 압축 과정은 Draco 와 SC3DMC 같은 정적 메쉬 압축을 위한 최첨단 코덱을 사용했고 텍스처 맵은 비디오 포맷으로 변경되어 HEVC HM 16.21 을 사용하였다.

### 4. 결론

MPEG 의 동적 메쉬 압축에 대한 cfp 에는 5 개의 기관이 응했으며 Apple 과 InterDigital 이 MPEG 에서 기준으로 제시한 압축보다 각각 두가지 평가기준에서 성능이 우수했다. InterDigital 의 동적 메쉬 압축 결과 Geometry PSNR 이 기준보다 26.3% 더 높았고, Apple 의 동적 메쉬 압축 결과 Y-PSNR 이 무려 55.5% 상승한 결과를 보였다. 여러 기관들이

제시한 제안을 참고하여 향후 7 월에 진행될 미팅에서 DMC 의 초기 tool 이 결정될 것이다

### 참고문헌

- [1] JongBeom Jeong, Dongmin Jang, Jang-Woo Son, Eun-Seok Ryu. 2018. 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.
- [2] Jong-Beom Jeong, Soonbin Lee, Dongmin Jang, Eun-Seok Ryu. 2019. Towards 3DoF+ 360 Video Streaming System for Immersive Media. IEEE Access, vol. 7, pp. 136399-136408.
- [3] Soonbin Lee, Jong-Beom Jeong, Inae Kim, Sangsoon Lee, Eun-Seok Ryu. 2020. Screen Content Coding Analysis to Improve Coding Efficiency for Immersive Video. Journal of Broadcast Engineering, vol. 25, no. 6, pp. 911-921.
- [4] Soonbin Lee, Jong-Beom Jeong, Eun-Seok Ryu. 2022. Efficient Group-Based Packing Strategy for 6DoF Immersive Video Streaming. In Proceedings of the 2022 International Conference on Information Networking, pp. 310-314.
- [5] Patrice Rondao Alface, Aleksei Martemianov, Sebastian Schwarz, Lauri Ilola, Lukasz Kondrad, Jozsef Szabo, Christoph Bachhuber. 2022. Nokia's response to Cfp for Dynamic Mesh compression. Standard ISO/IEC JTC1/SC29/WG7, MPEG/m59274.
- [6] Khaled Mammou, Jungsun Kim, Alexis Tourapis, Dimitri Podborski, Krasimir Kolarov. 2022. Apple's Dynamic Mesh Coding Cfp Response. Standard ISO/IEC JTC1/SC29/WG7, MPEG/m59281.
- [7] Jean-Eudes Marvie, Celine Guede, Julien Ricard, Olivier Mocquard, Maja Krivokuća, François-Louis Tariolle. 2022. InterDigital's Response to Dynamic Mesh Coding Cfp. Standard ISO/IEC JTC1/SC29/WG7, MPEG/m59285.
- [8] Xiang Zhang, Chao Huang, Jun Tian, Xiaozhong Xu, Shan Liu. 2022. Tencent's Dynamic Mesh Coding Cfp response. Standard ISO/IEC JTC1/SC29/WG7, MPEG/m59295.

- [9] Danillo B. Graziosi, Satoru Kuma, Kao Hayashi, Ohji Nakagami, Ali Tabatabai. 2022. Sony's Dynamic Mesh Coding Call for Proposal Response. Standard ISO/IEC JTC1/SC29/WG7, MPEG/m59284.
- [10] CfP for Dynamic Mesh Coding. 2021. ISO/IEC JTC 1/SC 29/WG 7. 38