

Video-based Dynamic Mesh Coding 표준화 동향 소개

□ 이종석 / 디지털인사이드

요약

본고에서는 현재 표준화 중인 MPEG Video-based Dynamic Mesh Coding(V-DMC)에 대한 표준화 동향과 V-DMC의 동적 메시 압축 기술을 소개하고자 한다. 전통적으로 3차원 데이터는 대부분 정적 메시로 고려되었기 때문에 기존의 메시 압축 표준 기술은 정적 메시 압축을 목적으로 하고 있다. 하지만, 최근 XR 콘텐츠, 게임, 메타버스와 같은 3차원 응용 서비스에서는 동적 메시가 주로 사용되고 있다. V-DMC는 이에 대응하기 위하여 동적 메시의 효율적인 압축을 위한 전처리 및 압축 기술을 포함하고 있다. 본고에서는 기존 메시 압축 기술과 다르게 V-DMC에서 새롭게 제안된 전처리 및 압축 기술에 대하여 살펴보고자 한다.

I. 서론

최근 3차원 데이터는 XR, 메타버스, 의료 영상, 게임, 내비게이션 뿐만 아니라 모바일 디바이스에서도 객체를 스캔하여 사용하는 등 3차원 데이터에 대한 사용량이 기하급수적으로 증가하는 추세이다. 일반적으로 3차원 데이터는 포인트 클라우드(point cloud)와 메시(mesh)를 의미하며, 취득에 사용된 센서의 정밀도에 따라서 획득되는 정점의 수가 증가하여 2D 비디오 데이

터 대비 높은 데이터양을 가질 수 있다. 따라서 3차원 데이터의 효율적인 압축 기술에 대한 필요성이 점점 증가하고 있다.

Google은 2016년 12월 포인트 클라우드 및 메시의 효율적인 압축을 위한 오픈소스 프로젝트인 DRACO[1] 개발을 시작하여 2017년 7월에 버전 1을 발표하였다. 현재 DRACO는 Google Maps, Google Earth, Chrome 등 위치 기반 서비스에서 주요하게 사용되고 있다. 하지만 DRACO는 메시 텍스처 맵에 대한 압축은 지원하지 않고,

정적 메시 압축을 목적으로 개발되었기 때문에 동적 메시 압축에는 적절하지 않다.

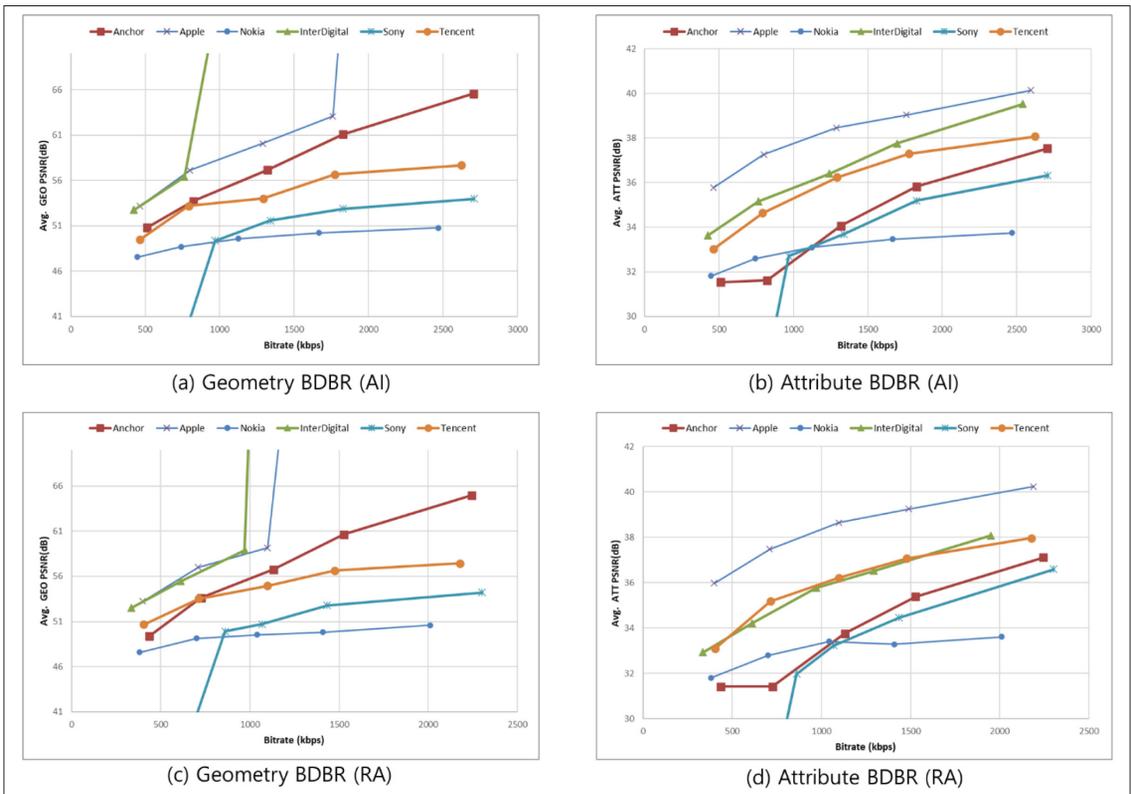
오랜 기간 3차원 데이터 압축 표준 기술을 개발해 온 국제 표준 기구 ISO/IEC JTC1/SC29/WG07의 MPEG-3DGH 그룹은 포인트 클라우드 압축 기술인 V-PCC(Video-based Point Cloud Compression), G-PCC(Geometry-based Point Cloud Compression)에 대한 기술 개발을 2018년부터 2년 간 수행하였고, V-PCC와 G-PCC 표준은 2021년에 국제 표준으로 제정 공표되었다. 그해 뒤이어 MPEG-3DGH 그룹은 2021년 10월 동적 메시 압축에 대한 CFP(Call for Proposals)[2]를 발행하여 <표 1>과 같은 표준화 일정으로 표준 기술 개발을 시작하였다. 2022년 4월 회의 이전에 Apple, Nokia, InterDigital, Sony, Tencent 등에서 동적 메시 압축 기술을 제출하였고 2022년 4월 회의에서 제출된 기술들에 대한 주관적, 객관적 압축 성능을 평가하고 비교하였다. <그림 1>은 CFP에 제출된 기술들에 대한 객관적 화질을 비교한 그래프이다.

<그림 1>에서 Anchor는 DRACO+HEVC로 메시의 구조적 정보는 DRACO로 압축하고 텍스처 맵은 비디오 압축 표준 기술인 HEVC를 사용한 방법이다. <그림 1>에서 볼 수 있는 것 같이 Apple과 InterDigital의 기술이 Anchor보다 높은 압축 효율을 보였으며, 특히 Apple에서 제안한 기술이 attribute에 대한 압축 효율이 상대적으로 높은 것을 확인할 수 있다. 또한, 주관적 평가에서도 Apple의 기술이 가장 좋은 것을 확인하여, 최종적으로 그룹 내의 논의를 통해 Apple의 기술[3]을 기반으로 V-DMC의 참조 소프트웨어 모델인 TMM(Test Model Mesh)[4] 버전 1을 정하였다. MPEG-3DGH 그룹은 TMM을 기초로 V-DMC[5] 표준 기술 개발을 시작하여 2024년 10월 최종 표준 제정을 목표로 하고 있다.

본고에서는 앞서 설명한 V-DMC 표준화 동향을 더욱 쉽게 파악하기 위하여 V-DMC 표준화에서 사용되는 3차원 메시 데이터의 구조를 살펴보고 TMM의 기본적인 압축 구조에 대하여 설명하고자 한다.

<표 1> V-DMC 표준화 주요 일정[2]

Year	Month	MPEG meeting	Stage
2021	10	136	Release of the CFP
2022	01	137	Nominate the best proposal(s)
	04	138	
	07	139	Establishment of the first test model based on the best proposal(s) including the combined proposal
	10	140	Approval of WD 1.0
2023	01	141	Approval of WD 2.0
	04	142	Approval of WD 3.0
	07	143	Approval of CD
	10	144	
2024	01	145	Approval of DIS (prefer stable text)
	04	146	
	07	147	
	10	148	Approval of FDIS

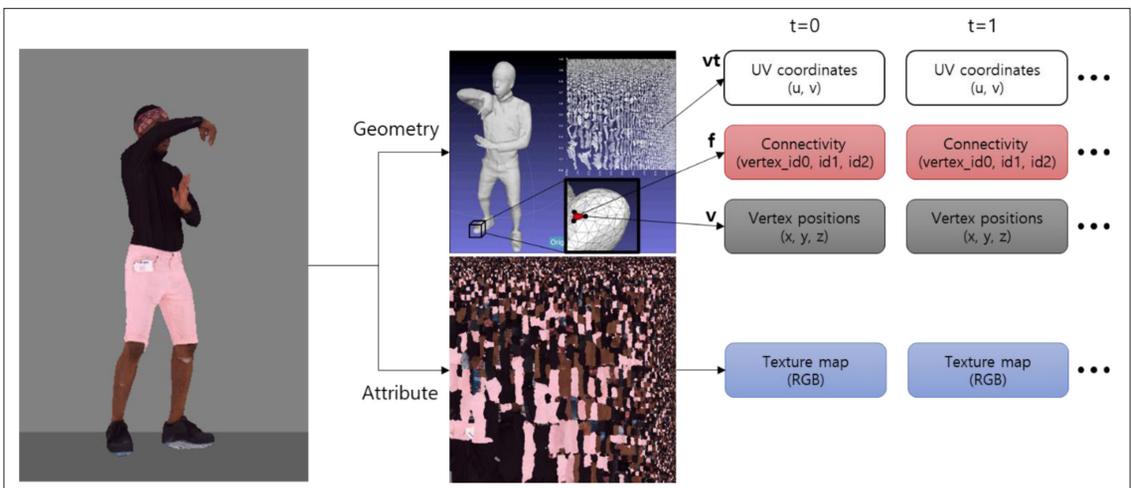


<그림 1> V-DMC CfP에 제출된 기술들의 객관적인 압축 성능 비교 그래프[6]

II. 동적 메시 데이터

한 그림이다. 3차원 객체를 모델링한 메시 데이터는 일반적으로 geometry, attribute 데이터를 포함하고 있다. Geometry 데이터는 3차원 객체의 구조적인 정보를 의

<그림 2>는 동적 메시 데이터의 구조를 설명하기 위



<그림 2> 동적 메시 데이터 구조

미한다. 구조적인 정보는 정점 위치(vertex position), 텍스처 정점 위치(UV coordinates), 연결성(connectivity)을 포함한다. 정점 위치(vertex position)는 메시를 구성하는 각 삼각형의 꼭짓점인 정점의 3차원 좌표 정보이다. 텍스처 정점 위치는 정점에 대응하는 텍스처 맵상의 2차원 좌표를 의미하며 일반적으로 0과 1 사이의 값으로 정규화된 값이다. 마지막으로, 연결성은 메시를 이루는 삼각형의 3개의 정점의 인덱스이다. Attribute 데이터는 3차원 객체 표면의 색상 관련 정보를 가지고 있으며, 일반적으로 RGB 색 공간을 가지는 2차원 텍스처 맵인 이미지이다. 즉, 동적 메시는 <그림 1>과 같이 geometry, attribute 데이터를 묶어 하나의 프레임으로 하여 시간상으로 다수개의 프레임을 의미한다. 즉, 일반적인 비디오에 추가로 geometry 정보가 있기 때문에 비디오보다 많은 데이터양을 지닌다. 또한 geometry 데이터의 경우 정점이 sparse하게 존재하기 때문에 상당히 압축하기에 어려운 데이터 중 하나이며, 정점의 인덱스로 이루어진 연결성 정보 또한 정점의 인덱스이기 때문에 연속적인 연결성 정보의 중복성이 낮아 압축하기 어려운 데이터라고 볼 수 있다.

<표 2>는 V-DMC 표준화에 사용되는 테스트 시퀀스 정보를 정리한 표이다. 테스트 시퀀스는 원본 데이터의 크기에 따라서 A-C의 클래스로 분류되어 있으며, 크기가 작은 시퀀스들은 Class A로 분류되었고 크기가 큰 시퀀스들은 Class C로 분류되어 있다. 또한, 약 10초 길이의 30fps를

가지는 시퀀스들로 구성되어 있다. 테스트 시퀀스들은 약 2만 개의 정점을 가지고 있으며 약 4만 개의 삼각형 면을 가지고 있는 것을 확인할 수 있다. 텍스처 맵은 Class A, Class B는 2K×2K 해상도를 가지고 Class C는 4K×4K 해상도를 가진다. <표 2>의 정보에 따르면 텍스처 맵의 정보량이 상대적으로 많은 것을 확인할 수 있다.

III. V-DMC 압축 기술 소개

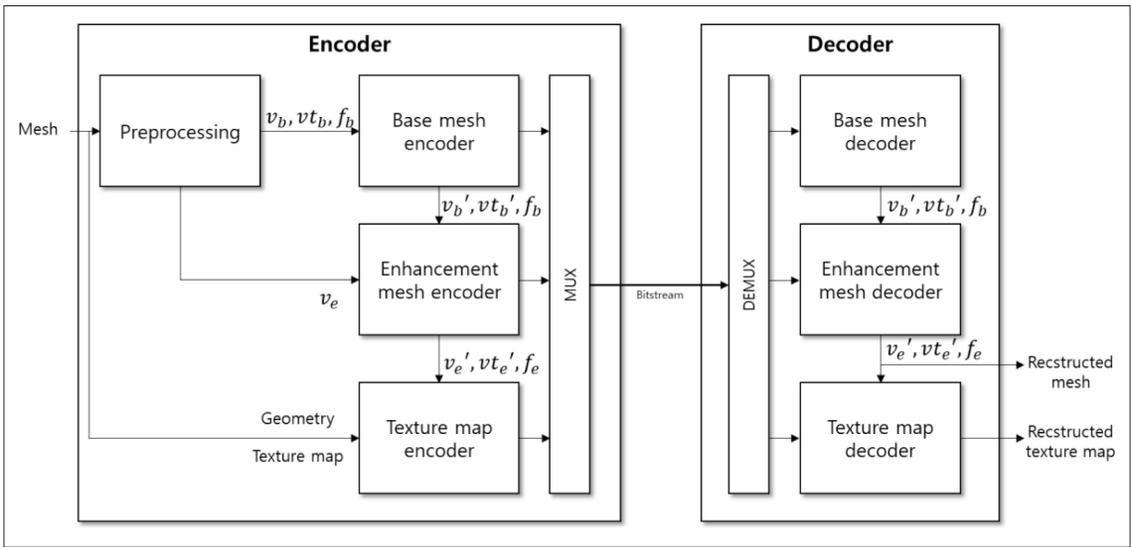
<그림 3>은 V-DMC의 참조 소프트웨어인 TMM의 부/복호화기 구조를 간략히 표현한 블록도이다. TMM의 부/복호화기에서 geometry 정보는 기본-방향 계층을 가지는 스케일러블 코딩 구조를 이용해 부호화되며, attribute 정보는 복원된 향상된 메시를 이용하여 텍스처 맵 부호화기를 통해 부호화된다. TMM의 복호화기는 부호화기의 역과정으로 기본 메시 복호화기, 향상 메시 복호화기, 텍스처 맵 복호화기 순서로 메시를 복원한다. 이후 절에서는 부호화기를 중심으로 각 블록을 조금 더 자세히 소개하고자 한다.

1. 전처리

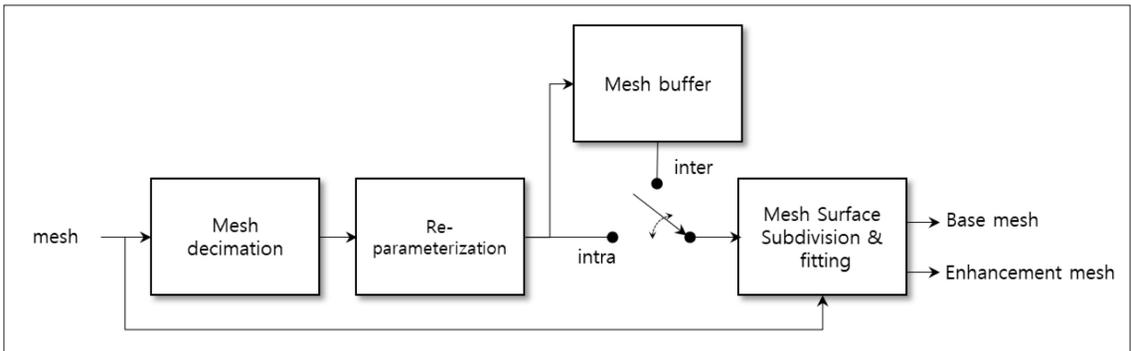
<그림 4>는 V-DMC 부호화기의 전처리 블록도이다. 전

<표 2> V-DMC 테스트 시퀀스 정보[2]

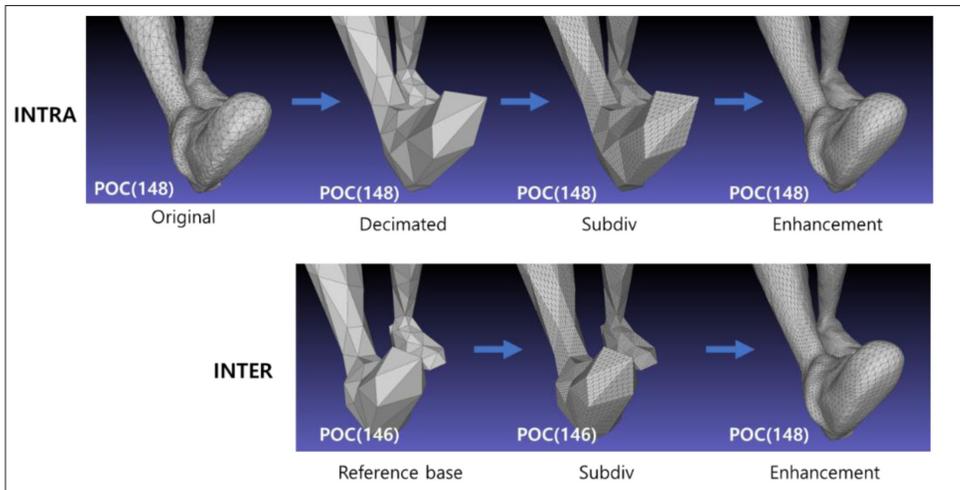
Test Class	Test material dataset filename	# Frames	# Vertices	# Faces	Geometry Precision	Texture Coord. Precision	Texture Map Size
A	longdress	300	22k	40k	10 bits	12 bits	2k x 2k
	soldier	300	22k	40k	10 bits	12 bits	2k x 2k
B	basketball_player	300	20k	40k	12 bits	12 bits	2k x 2k
	dancer	300	20k	40k	12 bits	12 bits	2k x 2k
C	mitch	300	16k	30k	12 bits	13 bits	4k x 4k
	thomas	300	16k	30k	12 bits	13 bits	4k x 4k
	football	300	25k	40k	12 bits	13 bits	4k x 4k
	levi	150	20k	40k	12 bits	13 bits	4k x 4k



<그림 3> V-DMC TMM 부/복호화기 블록도



<그림 4> TMM 전처리 블록도



<그림 5> 전처리 블록에서 화면 내 및 화면 간 예측을 위한 전처리 과정

처리 과정은 화면 내 예측을 위한 전처리 과정과 화면 간 예측을 위한 전처리 과정으로 나뉘며, 전처리 다음 과정인 기본 메시 부호화기에서 화면 내 예측 및 화면 간 예측을 모두 수행한 후 부호화 효율이 좋은 예측 방법을 선택한다. 따라서, 전처리 과정에서는 화면 내 예측 전처리와 화면 간 예측 전처리가 모두 수행되어야 하며, <그림 5>는 화면 내 전처리 및 화면 간 전처리 과정을 예시로 표현한 그림이다. 화면 내 전처리 과정은 먼저 메시 단순화 과정(Mesh decimation)이 수행된다. 3차원 객체의 모델을 취득할 때 센서 노이즈와 같은 다양한 요인으로 인해 원본 메시의 정점 및 연결성은 균일하지 않을 수 있다. V-DMC는 메시지를 전송받은 수신단에서 3차원 모델을 렌더링하기 위한 전처리 과정을 부호화 이전에 수행함으로써 수신단의 복잡도를 낮추는 방법을 채택하였다. 즉, 원본 메시는 단순화되고 다시 부호화기에서 업 샘플링되어 균일한 메시로 수정된다.

메시 단순화 과정 이후에는 단순화된 메시에 대하여 재매개변수화(re-parameterization)가 수행된다. 재매개변수화 과정은 단순화된 메시지를 기준으로 연속된 메시들을 다수 개의 패치로 분류하여 새롭게 메시지를 구성하는 단계이다. 즉, 3차원 공간에서는 연속된 삼각형이지만 텍스처 맵에서 그에 대응되는 삼각형들은 서로 다른 위치에 존재할 수 있다. 이러한 경우 불연속성이 커지기 때문에 텍스처 정점과 텍스처 맵을 부호화하는데 효율적이지 않다. 따라서, 재매개변수화를 통해 연속된 삼각형들의 텍스처 정점의 위치를 서로 이웃하도록 수정하여 전체적인 부호화 효율을 향상시킬 수 있다.

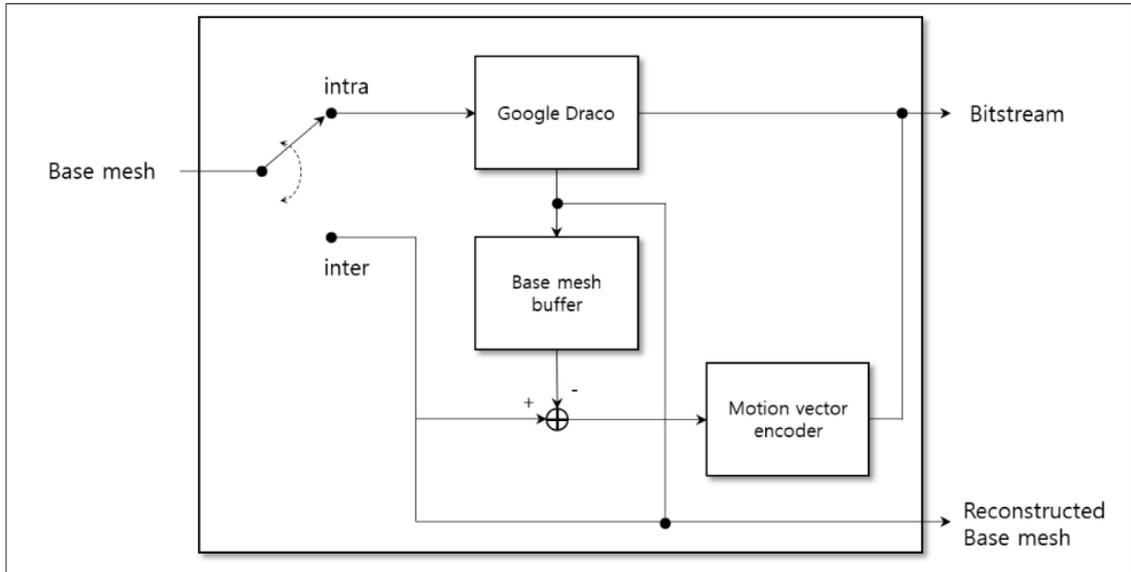
재매개변수화된 메시는 이후 화면 간 예측 전처리를 위하여 버퍼에 저장되며, 평면 분할 및 피팅(fitting) 과정이 수행된다. 평면 분할 및 피팅 과정은 앞서 설명한 균일한 메시지를 만들기 위한 업 샘플링 과정이다. 먼저 평면 분할은 각 삼각형을 4개의 작은 삼각형으로 분할하는 과정으로, 삼각형의 각 모서리 중점에 분할 정점을 생성하고 기존 정점과 새로 생성된 정점을 연결하여 4개의 삼각형을 생성할 수 있다. 이 평면 분할은 재귀적으로 수행될 수 있

으며 부호화기에서 반복 횟수를 입력받아 수행하며, 복호화기에서도 동일한 횟수로 분할하기 위하여 해당 정보를 복호화기로 전송한다. 마지막으로 피팅 과정이 수행된다. 피팅 과정은 평면 분할된 메시지를 원본 메시와 같아지도록 반복적으로 정점의 위치를 수정하는 과정이다. 이때, 평면 분할된 메시와 원본 메시 사이 거리의 차이의 합이 작아지도록 최적화가 수행된다. 최종적으로 피팅된 메시는 향상 메시가 되고 향상 메시지를 평면 단순화하여 기본 메시지를 생성할 수 있다.

화면 내 예측 전처리 이후에는 화면 간 예측 전처리가 수행된다. 먼저 버퍼에 저장된 단순화된 메시 중 참조 메시로 하여 평면 분할 및 피팅 과정을 수행한다. 즉, 3차원 객체가 움직여 이동된 정점을 피팅 과정을 통해 움직임 보상해 준다. 움직임 보상된 메시는 향상 메시가 되고, 평면 단순화를 통해 기본 메시지를 생성한다. 화면 간 예측 전처리 과정에서 참조 메시지를 사용해 피팅을 함으로써 기본 메시 부호화기의 화면 간 예측에서 현재 프레임과 참조 프레임 정점이 일대일 대응이 되어 단순히 정점의 위치를 차분하여 정점의 움직임 벡터를 생성할 수 있다. 기존 비디오 압축이나 PCC의 부호화기에서 화면 간 예측을 위하여 수행되던 움직임 추정 과정이 V-DMC에서는 전처리로 빠지면서, V-DMC 부호화기는 기존 방법들 대비 상대적으로 계산 복잡도가 낮다. 또한 부호화기에서 화면 간 예측으로 결정되었다면, 텍스처 정점 및 연결성은 전송하지 않고 참조 메시와 동일한 텍스처 정점과 연결성을 사용한다.

2. 기본 메시 압축

<그림 6>은 기본 메시 부호화기 블록도이다. 기본 메시 부호화기는 기본 메시지를 부호화하여 비트스트림을 생성하고 동시에 향상 메시지를 복원한다. 화면 내 프레임인 경우 기본 메시의 정점, 텍스처 정점, 연결성 정보가 DRACO를 통해 부호화된다. 전통적으로 메시의 연결성 정보를 부/복호화하는 다양한 방법 중 Edgebreaker[7]

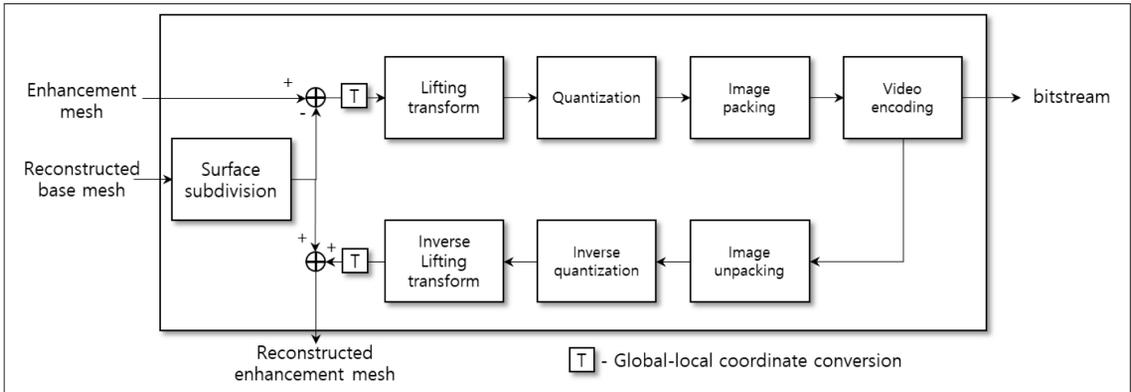


<그림 6> 기본 메시 부호화기 블록도

가 압축 효율 및 계산 복잡도 측면에서 탁월한 성능을 보여 V-DMC에서도 DRACO에 포함된 Edgebreaker를 사용한다고 볼 수 있다. 복원된 기본 메시는 화면 간 예측을 위하여 버퍼에 저장되고 동시에 향상 메시 부호화기로 전달된다. 화면 간 프레임인 경우 메시 버퍼에 저장된 참조 메시와 현재 메시의 정점 위치를 서로 차분하여 움직임 벡터를 생성하여 움직임 벡터 부호화기를 통해 부호화를 수행한다. 앞서 전처리 과정에서 설명한 것과 같이 화면 간 예측 전처리 과정에서 이미 참조 메시와 현재 메시가 일대일 대응되었기 때문에 단순히 정점끼리 차분하여 움직임 벡터를 얻을 수 있다. 마지막으로 움직임 벡터 부호화기는 움직임 벡터에 대하여 압축을 수행하며 DRACO를 통해 복원된 메시의 연결성을 이용하여 현재 부호화하고자 하는 정점과 연결된 정점들의 복원된 움직임 벡터를 이용하여 현재 움직임 벡터를 예측하고 현재 움직임 벡터와 차분값을 압축하여 복호화기로 전송한다. 여기서, 기본 정점은 향상 메시의 기본 골격이 되기 때문에 정확도를 위하여 움직임 벡터는 무손실 압축을 수행한다.

3. 향상 메시 압축

<그림 7>은 향상 메시 부호화기 블록도이다. 원본 향상 메시와 복원된 기본 메시지를 부호화하여 비트스트림을 생성하는 과정이다. 먼저, 복원된 기본 메시를 평면 분할(Surface subdivision)하여 원본 향상 메시와 차분하여 정점별 차이 벡터를 생성한다. 향상 메시와 복원된 기본 메시 역시 일대일 대응되어 있기 때문에 단순히 차분 연산이 가능하다. 차이 벡터는 normal-tangential-bitangential로 구성된 지역 좌표계로 변환된다. 이는 전처리 과정의 피팅 과정에서 평면의 normal 방향으로 차이가 최소화되도록 반복적으로 피팅되기 때문에 전역 좌표계보다 지역 좌표계의 에너지 밀집도가 높아진다. 이에 따라 전역 좌표계 대비 높은 부호화 효율을 얻을 수 있다. 이후 차이 벡터의 각 축에 대하여 웨이블릿 변환의 일종인 리프팅 변환(lifting transform)[8]을 수행한다. 리프팅 변환은 G-PCC 표준에서도 사용되는 변환 방법으로서 입력 데이터를 재귀적인 예측 및 갱신 과정을 통해 주파수 도메인으로 변환하는 방법이다. 주파수 도메인으로 변환된 차이 벡터



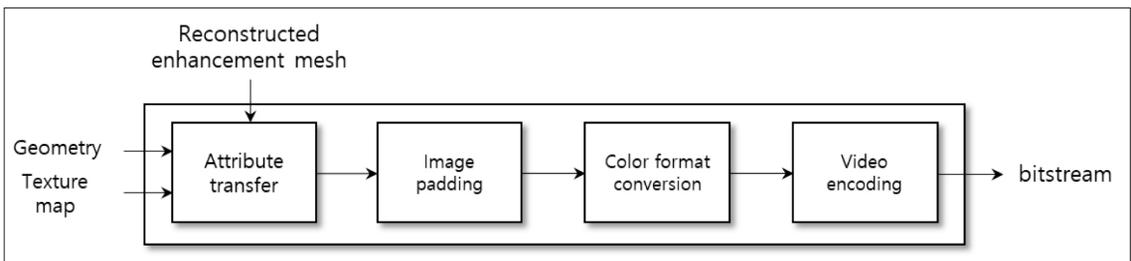
<그림 7> 향상 메시 부호화기 블록도

는 이후 양자화되고 축별로 2차원으로 패킹되어 YUV444 포맷의 이미지로 변환된다. 이때, 차이 벡터는 저주파 성분부터 고주파 성분 순서대로 이미지의 우하단에서 좌상단 방향으로 16x16 블록을 최대 크기로 하는 2x2 inverse z-scan 순서대로 패킹된다. 마지막으로 패킹된 이미지는 비디오 압축 표준인 HEVC로 부호화된다. 이후 과정은 앞선 과정의 역과정이 수행되며 최종적으로 향상 메시가 복원된다.

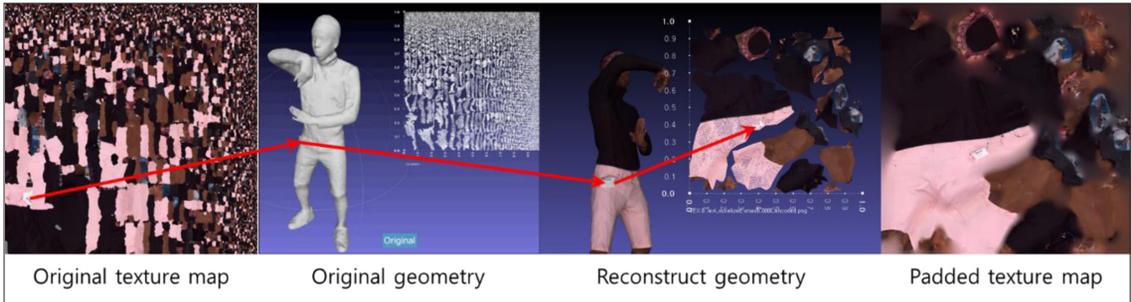
4. 텍스처 맵 압축

<그림 8>은 텍스처 맵 부호화기 블록도이다. 텍스처 맵 부호화기는 텍스처 맵을 비디오 압축 표준을 통해 압

축하는 과정으로 그 전처리 과정인 속성 전이(attribute transfer)와 패딩이 수행된다. 속성 전이 과정은 메시의 geometry 데이터가 전처리, 압축 및 복원 과정을 거치면서 원본과 다르게 수정되는데, 속성 전이 과정을 통해 원본 텍스처 맵의 색상 정보를 수정된 메시에 매핑하는 과정이라고 할 수 있다. <그림 9>의 예시와 같이 원본 메시의 geometry 정보와 복원된 메시의 geometry 정보를 비교하여 대응하는 위치를 찾고 원본 텍스처 맵에서 색 정보를 가져와 현재 geometry에 매핑하여 새로운 텍스처 맵을 생성한다. 새롭게 생성된 텍스처 맵은 빈 공간을 채워주는 패딩이 수행된다. 일반적으로 텍스처 맵은 <그림 9>의 3번째 그림과 같이 빈 공간이 존재한다. 이런 빈 공간은 이후 수행될 비디오 압축에서 날카로운 고주파 성분을 생



<그림 8> 텍스처 맵 부호화기 블록도



<그림 9> 텍스처 맵 부호화기 블록도

성하여 부호화 효율을 감소시키게 된다. 이를 막기 위하여 V-DMC에서는 이미지를 반복적으로 다운 샘플링 및 업 샘플링하면서 빈 공간을 채워 나가는 푸시-풀(push-pull) 패딩 방법이 사용된다. 푸시-풀 패딩은 V-PCC 표준에서도 사용되는 패딩 기술로 원래의 텍스처 영역과 빈 공간이 자연스럽게 연결되어 높은 압축 효율성을 보인다. 패딩된 이미지는 YUV420 포맷으로 변환되어 차이 벡터와 동일하게 HEVC를 통해 압축된다. 텍스처 맵 부호화기에서는 전송받은 비트스트림만 복호화하여 바로 3차원 렌더링에 사용될 수 있다.

IV. 결론

본고에서는 현재 표준화 진행 중인 V-DMC에 대한 표준화 동향과 동적 메시 압축 기술을 살펴보았다. 기존의 정적 메시 압축 기술들과 다르게 V-DMC는 전처리 과정을 통해 움직임 추정을 수행하여 부호화기의 계산 복잡도를 낮추면서 기본-향상 다계층의 압축 구조를 통해 효율적으로 메시의 연결성 정보를 부호화하는 특징이 있었다. 또한, V-DMC 표준은 Apple의 주도로 표준화가 진행되고 있으며 회의 과정에서 계산 복잡도, 수신단 병렬성 및 렌더링 효율성 등을 고려하면서 기술 개발을 진행하고 있기 때문에 앞으로 모바일 장치에서 V-DMC가 3차원 데이터 저장 및 전송에 활용될 수도 있을 것으로 예상된다.

참 고 문 헌

- [1] Google, Draco: 3D Data Compression, <https://github.com/google/draco>
- [2] ISO/IEC JTC 1/SC 29/WG 02, "CfP for Dynamic Mesh Coding," N145 Oct, 2021
- [3] ISO/IEC JTC 1/SC 29/WG 07, "Apple's Dynamic Mesh Coding CfP Response," m59281, Apr, 2022.
- [4] ISO/IEC JTC 1/SC 29/WG 07, "V-DMV-DMC TMM4.0," N595, Apr, 2023.
- [5] ISO/IEC JTC 1/SC 29/WG 07, "WD 3.0 of V-DMC," N615, May 2023.
- [6] ISO/IEC JTC 1/SC 29/WG 07, "Consolidated answers to the Dynamic Mesh CfP," m59299, Apr, 2022.
- [7] J. Rossignac, "Edgebreaker: Connectivity compression for triangle meshes," *IEEE Trans. Vis. Comput. Graphics*, vol. 5, no. 1, pp. 47-61, Jan, 1999.
- [8] W. Sweldens, "The lifting scheme: A construction of second generation wavelets," *SIAM journal on mathematical analysis*, vol. 29, no. 2 pp 511-546, 1998.

저 자 소 개



이 종 석

- 2016년 : 광운대학교 전자공학과 학사
- 2018년 : 광운대학교 전자공학과 석사
- 2018년 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- 2020년 ~ 현재 : 디지털인사이트 선임연구원
- 주관심분야 : 비디오압축, 3차원데이터처리, 스파이킹뉴럴네트워크, 위성영상처리, 딥러닝