MPEG G-PCC 국제표준 기술

□ 변주형, 최한솔, 심동규 / 광운대학교

요약

본 고는 ISO/IEC JTC 1/SC 29/WG 7 MPEG(Moving Picture Experts Group) 3DG(3D Graphics coding) 그룹에서 진행되고 있는 포인트 클라우드 데이터 압축 표준 기술 중 하나인 G-PCC(Geometry-based Point Cloud Compression) 표준에 대하여 설명하고자 한다. G-PCC는 포인트 클라우드의 기하 정보와 속성 정보를 3차원 공간에서 서로 다른 기술을 이용하여 압축하는 표준으로, 무손실 압축 방법의 경우 10:1의 압축율을 제공하고 손실 압축의 경우 35:1 정도의 압축율을 보인다. 본 고에서는 G-PCC의 기하 정보와 속성 정보의 압축 방법을 상세히 설명하고 같은 기능을 수행하는 압축 기술 간의 성능을 비교하고자 한다.

I. 서 론

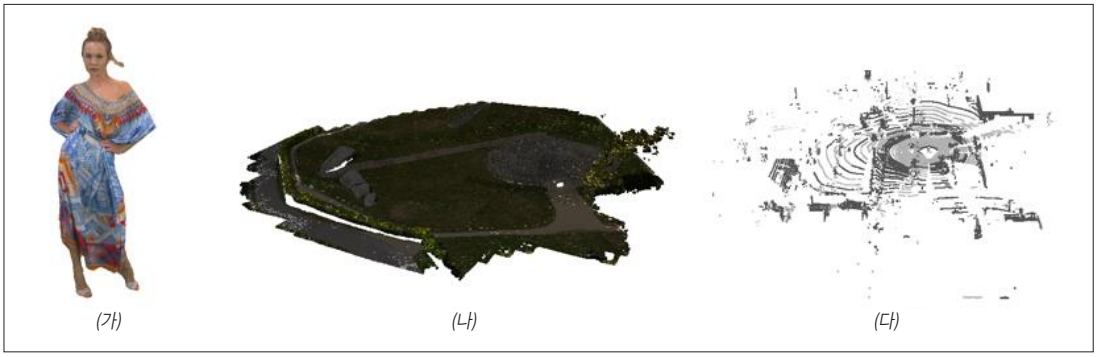
항공, 우주, 건설 및 컴퓨터 그래픽 등의 특수 분야에서 한정적으로 사용되던 3차원 영상은 최근 3D 영상 취득 장치의 발달과 더불어 기존 2D 영상의 한계를 뛰어넘어 3차원 공간의 현실 세계를 몰입감 있게 표현하고자 하는 요구에 의해 관심이 급증하고 있다. 특히 확장 현실(XR, Extended Reality) 기술 및 자율 주행 모빌리티 시스템을 위한 공간 데이터의 중요성 확대로 인해 각 3차원 데이터에 대한 취득 및 처리 기술이 각 분야에서 활발히 연구되고 있으며 최근에는 휴대폰과 태블릿 PC 등의 모바일 기기에도 3차원 공간 데이터를 취득하는 LiDAR(Light Detection And Ranging) 센서 및 뎁스 카메라 등이 탑재되어 3차원 데이터가 소비자들에 의해 서비스단에서의 생성 및 사용이 활발해지고 있다. 3차원 데이터는 점 구름 또는 포인트 클라우드라 불리는 데이터 형태로 표현할 수 있다. 포인트 클라우드 데이터는 3차원 공간상의 포인트들의 집합으로 각 포인트별 x, y, z의 좌표 값을 나타내는 기하 정보와 스칼라 또는 벡터 형태의 속성값을 가질 수 있다. 대표적인 속성값으로는 R, G, B 등의 색상 정보와 반사율 정보가 있으며 법선 벡터, 시간 정보 등 다양한 속성 정보를 포함할

수 있다. 일반적으로 하나의 3D 프레임 당 수십에서 수백만 개의 포인트로 구성되며 70만 개의 점을 갖는 포인트 클라우드의 원시 데이터를 30fps로 전송하기 위해서는 500MB/s 정도의 네트워크 대역폭이 요구된다[1]. 따라서 포인트 클라우드 데이터의 저장 및 전송을 위해서는 압축이 필수적이다. 이에 따라 국제표준화기구 ISO/IEC JTC1/SC29/WG 7 MPEG 3DG 그룹은 2017년의 CfP(Call for Proposals)[2]를 시작으로 실시간 3D 화상회의, VR/AR, 3D 자유 시점 스포츠 방송, 3D 지리 정보 시스템, 문화 유산 스캐닝, 3D 동적지도 기반 자율주행 내비게이션 등의 서비스 시나리오를 타겟으로 포인트 클라우드 압축 표준화를 시작하였다[3].

포인트 클라우드 압축 표준은 CfP를 기반으로 LiDAR 등의 장비를 통해 동적으로 취득된 포인트 클라우드를 타겟으로 하는 L-PCC(LiDAR Point Cloud Compression), 3D 그래픽 등과 같이 밀도가 높은 정적 포인트 클라우드를 위한 S-PCC(Surface Point Cloud Compression)와 동적 포인트 클라우드를 위한 V-PCC(Video-based Point Cloud Compression)의 세 가지 표준이 시작되었으며, 이후 L-PCC와 S-PCC 기술의 유사성으로 인해 하나의 표준인 G-PCC로 합쳐졌다.

MPEG에서는 포인트 클라우드 데이터의 종류 및 취득 방법에 따라 정적 포인트 클라우드 데이터의 카테고리 1, 동적 포인트 클라우드 데이터의 카테고리 2, 동적으로 움직이면서 취득한 카테고리 3으로 포인트 클라우드 데이터를 분류하고 있다[4]. G-PCC는 카테고리 1, 3의 데이터를 테스트 시퀀스로 하여 압축 표준을 진행하고 있다. 카테고리 1 데이터는 객체나 공간에 대해 점들의 밀도가 높은 단일 프레임의 포인트 클라우드로 구성되며 부호화 복잡도에 따라 ClassA와 ClassB로 구분된다. ClassB의 경우 ClassA 대비 부호화 복잡도가 큰 대용량 포인트 클라우드가 포함된다. 카테고리 3 데이터의 경우 동적으로 이동하면서 취득된 다수 개의 프레임을 갖는 frame-based 데이터와 대규모 공간에 대해 라이다 센서를 통해 취득된 포인트 클라우드와 2D 영상으로 취득된 칼라 영상이 정합된 단일 프레임의 fused 데이터로 구분된다. 카테고리 1, 3 데이터의 예시는 〈그림 1〉과 같다.

G-PCC는 현재 FDIS(Final Draft International Standard) 단계로 2021년에 국제표준으로 승인될 예정으로, 무손실 압축 방법의 경우 10:1의 압축율을 제공하고 손실 압축의 경우 허용 가능한 품질에서 최대 35:1의 압축율을 제공하고 있다.



<그림 1> (가) 카테고리 1 ClassA, "longdress_vox10_1300", (나) 카테고리 1 ClassB, "landscape_00014_vox14", (다) 카테고리 3 Frame-based, "ford_01_vox1mm-0100"

본 고의 구성은 다음과 같다. II장에서 G-PCC의 속성 정보와 기하 정보의 핵심 압축 기술에 대해 소개한다. III장에서는 G-PCC 참조 소프트웨어 TMC13v12.0[5]의 압축 기술의 성능을 비교하고, IV장에서 결론을 맺는다.

II. G-PCC 압축 기술

본 절에서는 G-PCC의 핵심 압축 기술에 대해 살펴본다. 〈그림 2〉는 G-PCC의 참조 소프트웨어인 TMC13 부호화기의 전체 구조도를 나타낸다.

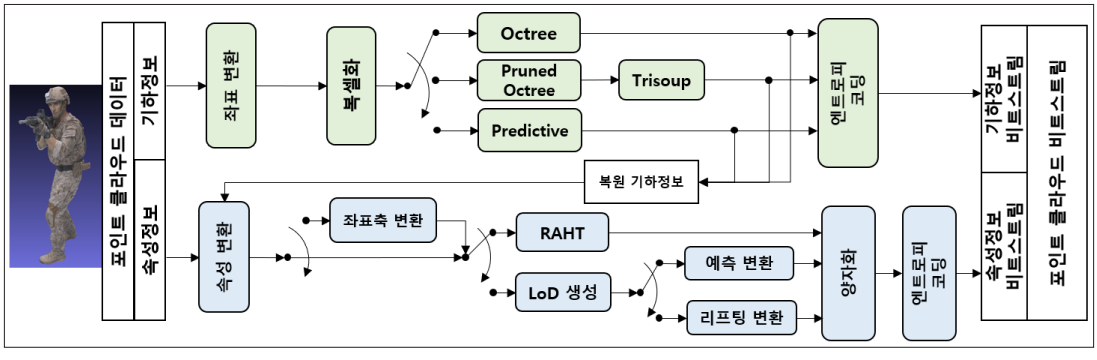
G-PCC는 기하 정보와 속성 정보가 서로 다른 압축 기술을 통해 압축이 수행되며 기하 정보의 경우 octree[6], trisoup, predictive[7]의 세 가지 압축 기술 중 선택적으로 하나의 기술을 통해 3차원 공간상의 기하 정보 압축을 수행한다. 속성 정보의 경우 부호화 된 기하 정보를 복원하여 생성된 복원 기하 정보를 사용하여 압축을 수행한다. 속성 정보 압축 기술로는 예측 변환(Predicting Transform), 리프팅 변환(Lifting Transform)[8] 및 지역 적응적 계층 변환(Region-Adaptive Hierarchical Transform)[9]이 있으며 세 기술 중 선택적으로 하나의 기술을 통해 압축을 수행할 수 있다.

1. 기하 정보 압축 기술

G-PCC 부호화기의 입력으로 들어오는 기하 정보는 월드 좌표계의 부동소수점의 형태일 수 있다. G-PCC 참조 소프트웨어 TMC13의 부호화기는 기하 정보를 압축하기에 앞서 해당 좌표를 압축을 위한 내부 좌표로 변환해주는 좌표 변환 과정을 수행한다. 좌표 변환은 식 (1)에 의해 수행되며, 입력 기하 정보 X_n^{orig}를 평행이동 파라미터 T와 스케일 파라미터 s를 통해 공간좌표가 $[0, 2^{d_x}) \times [0, 2^{d_y}) \times [0, 2^{d_z})$ 범위의 내부 좌표계로 변환을 수행한다.

$$X_n = (X_n^{orig} - T)/s, \quad T = (t_x, t_y, t_z) \qquad (1)$$

내부 좌표계로 변환된 기하 정보에 라운딩 연산을 수행하여 정수화하고, 동일한 정수 좌표로 맵핑되는 다수 개의 포인트의 속성 정보를 평균내어 복셀의 속성 정보를 지정하는 과정을 복셀화라고 한다. 복셀이란 $[0, 2^{d_x}) \times [0, 2^{d_y}) \times [0, 2^{d_z})$ 범위의 내부 좌표 공간에서 $1 \times 1 \times 1$ 크기의 공간을 복셀이라고 하고 손실 압축의 경우 복셀 당 최대 하나의 점만 존재하고 무손실 압축의 경우 복셀당 중복된 점의 개수가 전송된다. 포인트가 존

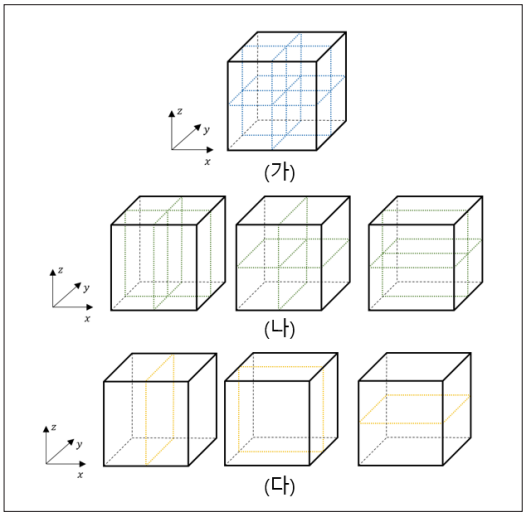


<그림 2> G-PCC 부호화기 구조도

재하는 복셀에는 포인트의 기하 정보와 속성 정보가 존재한다. 복셀화를 거친 내부 좌표 공간의 포인트 클라우드 기하 정보는 다음의 octree, trisoup, predictive 압축 기술 중 하나의 압축 기술을 통해 압축된다.

1) Octree 기반 기하 정보 압축 기술

Octree(OT) 기반의 기하 정보 코딩 방법은 복셀화된 기하 정보를 octree 구조 기반으로 재귀적으로 분할된 서브 노드의 occupancy 정보로 표현하여 부호화하는 방법이다. 최초 노드인 바운딩 박스는 $2^{d_x} \times 2^{d_y} \times 2^{d_z}$의 크기를 가질 수 있고 따라서 정육면체가 아닐 수 있다. 바운딩 박스를 시작으로 모든 노드가 $2^M \times 2^M \times 2^M$ 크기의 정육면체일 때까지 암시적으로 quad-tree(QT) 또는 binary-tree(BT) 구조로 분할되고 이를 암시적 QTBT 방법이라고 한다. 〈그림 3〉은 OT, BT, TT의 분할 방법을 나타낸다. 바운딩 박스를 〈표 1〉의 기준에 따라서 K번 암시적으로 QT 또는 BT 분할한 뒤, 〈표 2〉의 기준에 따라서 노드의 크기가 $2^M \times 2^M \times 2^M$일 때까지 암시적으로 QT 또는 BT 분할을 수행한다. 두 축의 길이가 같고 나머지 하나의 축의 길이가 작은 경우는 긴 두 축으로



<그림 3> (가) octree 분할(x-y-z 축 분할), (나) quad-tree 분할(좌측부터 x-y, x-z, y-z 축 분할), (다) binary-tree 분할(좌측부터 x, y, z 축 분할)

QT 분할을 수행하고, 한 축의 길이가 나머지 두 축의 길이보다 긴 경우는 긴 축으로 BT 분할을 수행한다. K번의 암시적 분할이 완료된 이후에 노드의 크기가 〈표 2〉의 조건을 만족하지 않을 경우, 〈표 2〉의 조건을 만족할 때까지 OT로 분할한 뒤 〈표 2〉의 조건에 따라서 노드의 크기가 $2^M \times 2^M \times 2^M$일 때까지 분할을 수행한다.

<표 1> K 번째 분할까지의 암시적 QT, BT 분할 조건

Condition	QT Partitioning	Condition	BT Partitioning
$d_z < d_x = d_y$	QT along x-y axes	$d_y < d_x \ and \ d_z < d_x$	BT along x axes
$d_y < d_x = d_z$	QT along x-z axes	$d_x < d_y \ and \ d_z < d_y$	BT along y axes
$d_x < d_y = d_z$	QT along y-z axes	$d_x < d_z \ and \ d_y < d_z$	BT along z axes

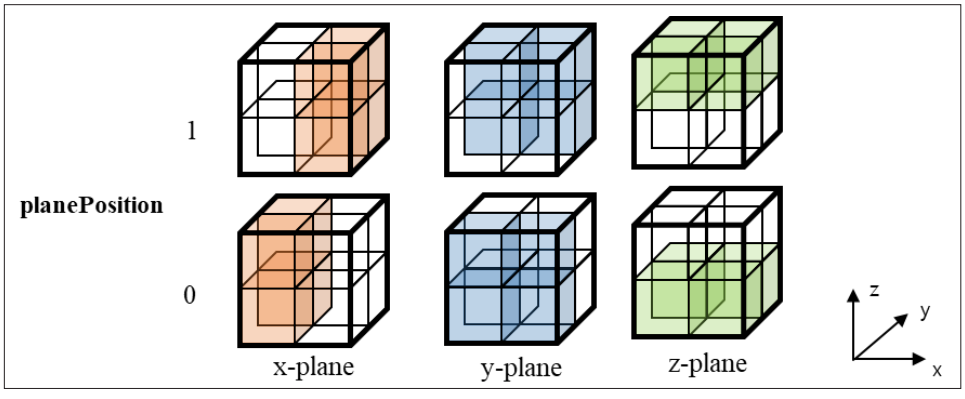
<표 2> K 번째 분할 이후의 암시적 QT, BT 분할 조건

Condition	QT Partitioning	Condition	BT Partitioning
$d_z = M < d_x = d_y$	QT along x-y axes	$d_y = M \le d_z < d_x \ or$ $d_z = M \le d_y < d_x$	BT along x axes
$d_y = M < d_x = d_z$	QT along x-z axes	$d_x = M \le d_z < d_y \ or$ $d_z = M \le d_x < d_y$	BT along y axes
$d_x = M < d_y = d_z$	QT along y-z axes	$d_x = M \le d_y < d_z \ or$ $d_y = M \le d_x < d_z$	BT along z axes

〈표 2〉에 따른 분할을 수행하여 노드의 크기가 $2^M \times 2^M \times 2^M$에 도달한 이후에는 각 노드에 대하여 최소 크기 노드에 도달할 때까지 OT 기반 분할을 수행한다. 노드는 *x*, *y*, *z* 축을 각각 절반으로 분할하여 8개의 서브 노드로 분할되고, 각 서브 노드에 포인트가 존재하는지의 유무가 총 8-bit의 occupancy로 표현된다. Occupancy는 현재 노드와 이웃하는 노드와의 관계를 이용하여 엔트로피 코딩을 수행한 후 전송된다. 포인트가 한 개 이상 존재하는 서브 노드의 occupancy는 1, 포인트가 존재하지 않는 서브 노드의 occupancy는 0의 값을 갖는다. 너비 우선 탐색 순서로 서브 노드의 크기가 $1 \times 1 \times 1$에 도달할 때까지 octree 분할을 수행한다.

Octree 구조에서 점유된 노드 정보를 보다 효율적으로 표현하기 위해 planar 코딩 모드가 도입되었다[10]. Planar 코딩 모드는 하나의 노드에 대하여 8-bit의 occupancy를 전송하는 대신, 노드 내에서 포인트가 존재하는 서브 노드들을 각 축방향의 평면 정보의 조합으로 표현하여 전송하는 모드이다. *x*, *y*, *z* 각 축에 수직인 여섯 개의 평면을 조합하여 포인트가 존재하는 다수개의 서브 노드의 위치를 표현한다. 이때 planar 모드의 사용 여부는 모든 노드에 전송되는 것이 아닌 과거 코딩된 노드가 평면일 가능성을 추적하여 확률에 기반하여 적격성을 판단하고, planar 모드에 적격하다고 판단된 노드에서 planar 플래그를 전송한다. Planar 모드가 사용되는 경우 평면의 방향 *x-plane*, *y-plane*, *z-plane* 인덱스와, 선택한 방향마다 평면의 위치 *planePosition*을 전송한다. 평면 정보만으로 서브 노드가 특정되지 않는 경우 추가 정보를 전송한다. 〈그림 4〉는 *x-plane*, *y-plane*, *z-plane*의 *planePosition* 값에 따른 평면을 나타낸다. *planePosition*의 엔트로피 코딩 과정에서 라이다로 취득된 기하 정보의 경우 라이다 센서의 취득 특성을 반영하여 *planePosition*의 엔트로피 문맥 정보를 선택하는 angular 모드[11]와 azimuthal 모드[12]가 추가되었다. Angular 모드의 경우 *z-plane*, azimuthal 모드의 경우 *y-plane*, *z-plane*의 *planePosition*을 엔트로피 코딩할 때 해당 문맥 정보를 통해 압축효율을 높인다.

Octree를 기반으로 기하 정보를 압축하는 데에 있어 주변 점과의 상관관계가 적은 isolated 점은 octree 기반 구조 대신 직접 해당 좌표를 전송해주는 것이 압축효율이 좋은 경우가 있다. 이를 DCM(Direct Coding Mode)이라 하며 G-PCC는 모든 점에 DCM 플래그를 전송하는 대신 주변 노드의 정보를 사용해 DCM의 적격성을 판단한 후 적격성을 만족하는 점들에 대해서만 DCM플래그를 전송하는 IDCM(Inferred Direct



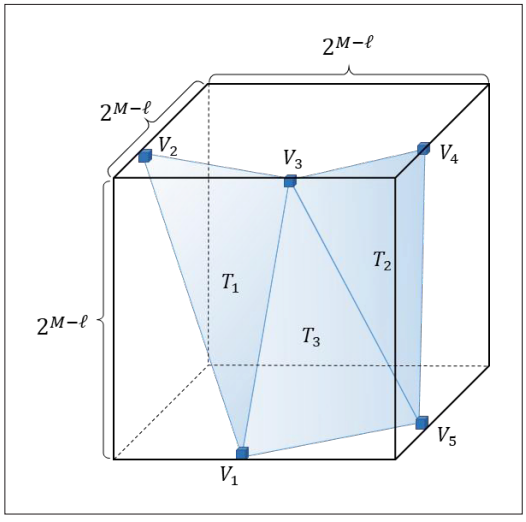
<그림 4> Planar 모드의 평면 방향과 *planePosition*

Coding Mode)[13]가 사용되고 있다. IDCM 모드도 라이다 센서를 통해 취득된 데이터의 경우 angular 모드와 azimuthal 모드의 문맥 정보를 사용하여 노드 내 포인트의 z축 좌표 및 x, y축 좌표의 엔트로피 코딩 효율을 높이는 기술이 사용되고 있다.

기하 정보 양자화는 부동소수점으로 표현된 입력 기하 정보를 라운딩하여 복셀화 시키는 전처리 과정에서의 양자화와 더불어 바운딩 박스 내에서 지역적으로 비트와 화질을 조절하고자 octree 기반의 부호화 과정에서도 양자화가 수행된다. 양자화는 슬라이스 기반, octree 기반으로 수행될 수 있다. 슬라이스 기반의 양자화는 슬라이스 단위로 슬라이스 헤더에서 전송되는 슬라이스 base QP(Quantization Parameter)를 사용하여 양자화를 수행한다. Octree 기반의 delta QP를 사용할 특정 octree 깊이가 슬라이스 헤더에서 전송되고, 해당 깊이를 갖는 노드 단위로 delta QP가 선택적으로 전송된다. 슬라이스 단위로 전송된 base QP에 delta QP를 합산하여 최종 QP를 유도하고 양자화를 수행한다.

2) Trisoup 기반 기하 정보 압축 기술

Trisoup 기반 기하 정보 코딩 방법은 노드 내 기하 정보를 삼각형 메쉬로 근사화 하여 코딩하는 방법이다. $2^M \times 2^M \times 2^M$ 크기의 노드를 시작으로 octree 기반으로 분할을 수행하다가 $2^{M-l} \times 2^{M-l} \times 2^{M-l}$ 크기의 노드부터는 더 이상 분할하지 않고 trisoup 방법을 통해 평면을 근사화 한다. 이때 파라미터 l은 전송되며 l의 값이 0일 경우 octree 기반 기하 정보 압축 방법과 동일하게 수행된다. 노드 내 기하 정보를 근사화 할 삼각형 $T_{i,\ i=1,\cdots,n-2}$를 표



<그림 5> Trisoup 기반의 노드 기하 정보 근사화 예시

<표 3> 정렬된 n개의 꼭지점을 이용한 삼각형 구성 방법

n	Triangles
3	(1,2,3)
4	(1,2,3), (3,4,1)
5	(1,2,3), (3,4,5), (5,1,3)
6	(1,2,3), (3,4,5), (5,6,1), (1,3,5)
7	(1,2,3), (3,4,5), (5,6,7), (7,1,3), (3,5,7)
8	(1,2,3), (3,4,5), (5,6,7), (7,8,1), (1,3,5), (5,7,1)
9	(1,2,3), (3,4,5), (5,6,7), (7,8,9), (9,1,3), (3,5,7), (7,9,3)
10	(1,2,3), (3,4,5), (5,6,7), (7,8,9), (9,10,1), (1,3,5), (5,7,9), (9,1,5)
11	(1,2,3), (3,4,5), (5,6,7), (7,8,9), (9,10,11), (11,1,3), (3,5,7), (7,9,11), (11,3,7)
12	(1,2,3), (3,4,5), (5,6,7), (7,8,9), (9,10,11), (11,12,1), (1,3,5), (5,7,9), (9,11,1), (1,5,9)

현하기 위하여 삼각형의 꼭지점 $V_{j,\ j=1,\cdots,n}$의 개수와 좌표 정보가 전송된다. 복호화기에서는 전송받은 n개의 꼭지점 정보를 이용하여 〈표 3〉에 의하여 (n-2)개의 삼각형을 유도한다. 〈그림 5〉는 n=5일 때 세 개의 삼각형을 유도한 예시를 나타낸다. 부호화기는 해당 정보를 받아 삼각형을 유도한 뒤 $2^{M-l}\times 2^{M-l}\times 2^{M-l}$ 크기의 노드 내에서 x, y, z 축으로 d의 간격을 갖는 복셀들 중 유도된 삼각형의 평면과 교차하는 복셀을 복원된 기하 정보로 지정한다.

3) Predictive 기하 정보 압축 기술

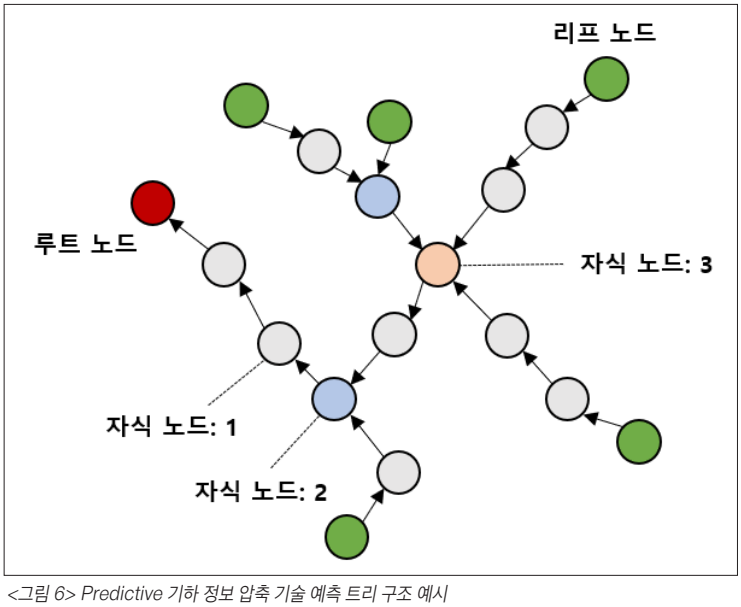
G-PCC 표준화 과정에서 octree 기반 기하 정보 압축 기술의 대안으로 predictive 기하 정보 압축 방법이 채택되었다. Predictive 기하 정보 압축 기술의 경우 octree 기반 압축 기술 대비 복호화 복잡도가 낮으며 응용 목적에 따라 부호화기의 지연 속도를 쉽게 조정할 수 있다는 장점이 있다. Predictive 기하 정보 압축 방법은 〈그림 6〉의 예측 트리 구조를 통해 기하 정보 압축을 수행한다.

Predictive 기하 정보 압축은 해당 예측 트리 구조를 기반으로 예측을 수행하며 각 노드들은 조상 노드를 통해서만 예측을 수행한다. Predictive 기하 정보 압축 기술은 총 4가지의 모드가 존재하며 모드는 〈표 4〉와 같다.

<표 4> Predictive 기하 정보 압축 예측 모드

모드 번호	예측 모드	예측자
0	*None*	·
1	*Delta 예측*	*p0*
2	*선형 예측 1*	*2*p0-p1*
3	*선형 예측 2*	*p0+p1-p2*

이때 p0, p1, p2는 각각 현재 포인트를 기준으로 parent, grandparent, grand-grandparent 노드를 의미한다. Predictive 기하 정보 압축 방법은 부호화기에서 참조 샘플로 사용될 점을 계산하여 조상 노드를 결정하며 사용자의 응용에 따라 참조 샘플을 계산할 범위를



<그림 6> Predictive 기하 정보 압축 기술 예측 트리 구조 예시

제한함으로써 부호화기의 계산 복잡도를 제어할 수 있다. 부호화기는 최적의 조상 노드를 계산하고 최적 모드를 결정하여 루트 노드부터 리프 노드까지 순차적으로 기하 정보의 압축을 수행하며 복호화기는 이를 순서대로 파싱받아 복원을 수행할 수 있으며 부호화기에서 각 노드 별 자식 노드의 수를 전송하고 복호화기는 이를 받아 참조 구조를 유도할 수 있다.

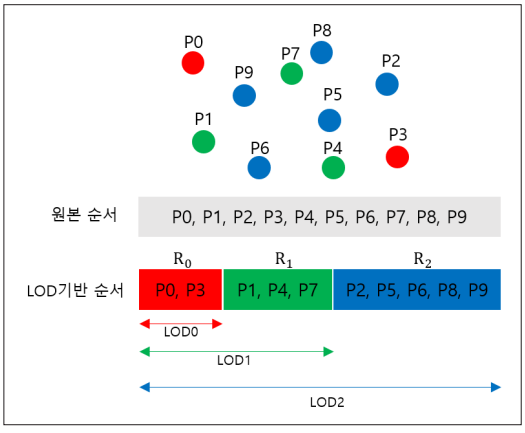
2. 속성 정보 압축 기술

속성 정보는 압축되기에 앞서 손실 기하 정보 압축으로 변경된 기하 정보에 적합한 속성값을 매핑하기 위하여 원본 포인트 클라우드 정보를 통해 부호화기에서 복원된 점들에 대한 속성값을 할당해주는 속성 변환 과정을 수행한다[14]. 이후 라이다 센서로 취득된 비균등한 샘플 거리를 갖는 원통 좌표계의 점들을 라이다 센서의 취득 특성을 통해 균등한 샘플 거리를 갖는 좌표 축으로 변환하는 기술이 선택적으로 적용될 수 있다[15]. 부호화기에서 좌표계 변환이 적용된 경우 복호화기에서는 원본 포인트 클라우드의 원통 좌표계로 복원하기 위해 좌표계 역변환 과정이 수행된다. 이후 속성 정보는 포인트를 정해진 규칙에 의해 세분화 레벨로 나누는 LoD(Level of Detail) 기술 기반의 예측 변환(Predicting Transform) 및 리프팅 변환(Lifting Transform)과 Haar 웨이블릿 변환 기반의 RAHT(Region-Adaptive Hierarchical Transform)가 있다. 예측 변환의 경우 무손실 속성 정보 압축 기술이며 리프팅 변환과 RAHT의 경우 손실 압축 기술로 사용자의 선택에 의해 세 가지 속성 정보 압축 기술이 선택적으로 사용될 수 있다. 각 압축 기술을 통해 예측 및 변환을 수행하여 생성된 잔차 신호/계수는 QP를 통해 유도된 양자화율로 양자화가 수행한 뒤 런-길이 코딩[16]을 수행하여 비트스트림이 생성된다.

1) 예측/리프팅 변환

예측 변환은 G-PCC의 무손실 압축 기술로 주변 포인트의 속성값을 사용하여 현재 포인트의 속성값을 예측하는 방법이며, 리프팅 변환은 손실 압축 기술로 예측 변환을 기반으로 추가적으로 업데이트 연산 과정을 수행한다. 두 기술은 복원된 기하 정보를 기반으로 부/복호화기에서 동일한 방식으로 생성된 LoD 구조를 기반으로 예측이 수행된다. LoD 생성은 원본 데이터의 순서로 저장되어 있는 포인트들을 포인트들 간의 L1 거리를 기반으로 세분화 레벨 $(R_l)_{l=0\cdots L-1}$로 재구성하는 과정이다. 세분화 레벨은 모든 포인트를 순회하며 각 세분화 레벨 별 포인트 간의 거리 $(d_l)_{l=0\cdots L-1}$, $(d_l < d_{l-1}, d_{L-1}=0)$ 이상의 포인트 집합을 R_0부터 R_{L-1} 순서로 순차적으로 각각의 세분화 레벨로 구성한다. 이때 n번째 LoD에 해당하는 포인트는 다음과 같이 정의된다.

$$LoD_n = \bigcup_{i=0}^{n} R_i = LoD_{n-1} \cup R_n \qquad (2)$$



<그림 7> LoD 생성 예시

〈그림 7〉은 LoD 생성 예시로 낮은 LoD는 점들 간의 간격이 큰, 낮은 밀도의 데이터 특성을 가지며 LoD 레벨이 높아질수록 점들의 밀도가 높아져 마지막 LoD는

포인트 클라우드의 모든 점을 포함한다. G-PCC는 해당 LoD 구조를 통해 특정 LoD 레벨의 점들까지만 복원함으로써 공간적 스케일러블 기능을 지원하고 있다. 예측 변환과 리프팅 변환의 기본 아이디어는 LoD_n의 점들의 속성값을 통해 세분화 레벨 R_{n+1}의 점의 속성값을 예측하는 것이다. 예측 변환의 경우 현재 포인트를 기준으로 복원된 기하 정보를 통해 이전의 부/복호화된 포인트 중 k개의 가까운 이웃점들을 계산한다. 이 때 k는 부호화기에서 사용자가 설정할 수 있는 파라미터로 k=3인 경우의 예측 모드는 〈표 5〉와 같다.

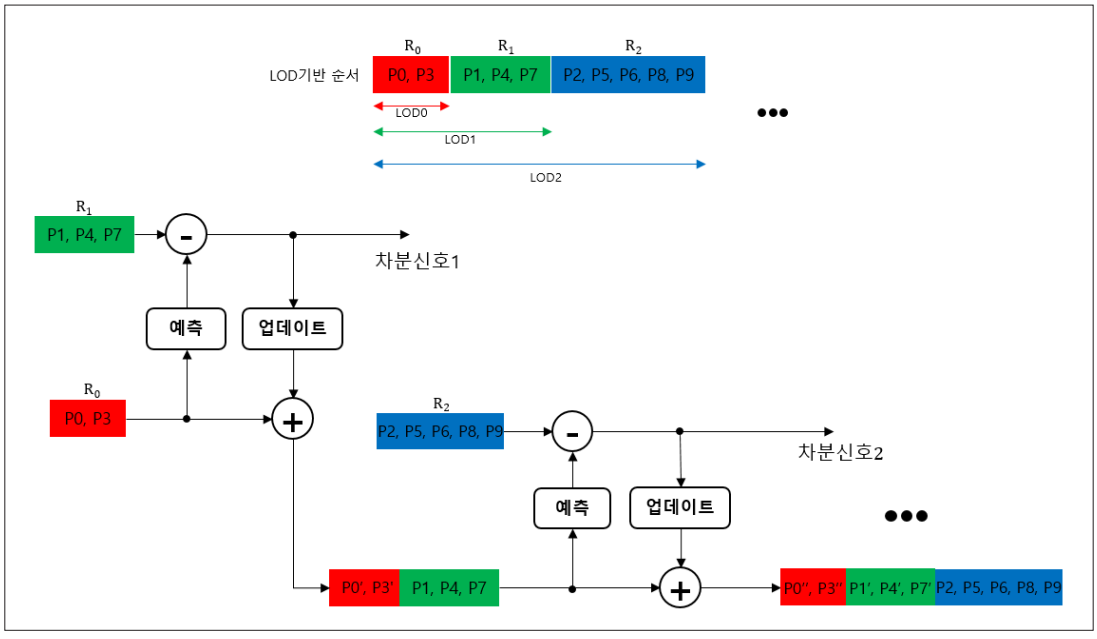
<표 5> 예측 변환 예측 모드 (k=3)

모드 인덱스	예측 모드
0	거리 기반 가중 평균
1	첫 번째 가까운 점
2	두 번째 가까운 점
3	세 번째 가까운 점

예측 모드는 k개의 예측 포인트의 속성값 중 하나를 예측자로 사용하는 모드와 k개의 이웃점들의 속성값을 포인트 간의 거리에 반비례한 가중치를 통해 가중 평균을 계산하는 모드 중 최적의 모드를 결정하여 해당 모드를 통해 예측을 수행한다. 이때 예측 변환은 주변 이웃점들의 속성값의 차이 값을 확인하여 모드 정보를 생략하는 기술이 채택되어 부/복호화기에서 동일한 방법으로 k개의 이웃 포인트의 속성값 중 가장 큰 값과 가장 작은 값의 차이가 특정 임계값 이하인 경우에는 예측 모드의 전송/파싱 없이 거리 기반 평균 예측을 사용한다. 결정된 예측 모드를 통해 예측 신호를 생성하여 원본 신호와의 차이 값인 차분 신호를 생성한다.

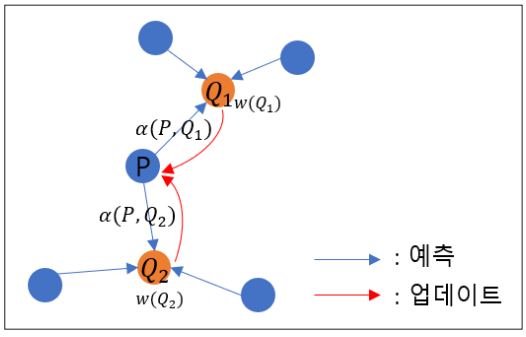
리프팅 변환의 경우 예측 변환에 업데이트 연산이 추가된 구조로 리프팅 변환 과정은 〈그림 8〉과 같다.

리프팅 변환은 예측 변환과 동일하게 LoD_n의 점들의 속성값을 통해 n+1번째 세분화 레벨 R_{n+1}의 점들의 속



<그림 8> 리프팅 변환 예시

성값을 예측하며 이때 예측 변환과 달리 항상 거리 기반 가중치 평균을 사용한다. 업데이트 연산은 예측을 통해 생성된 잔차 신호를 통해 예측에 사용된 예측자를 업데이트해주는 과정이다. 점 P에 대한 업데이트 연산 과정의 예시는 〈그림 9〉와 같으며 점 P를 사용하여 예측을 수행한 점들의 집합이 $\Delta(P)$일 때 점 P에 대한 업데이트 값은 식 (3)을 통해 계산된다.



<그림 9> 점 P에 대한 업데이트 연산 예시

$$Update(P) = \frac{\sum_{Q\in\Delta(P)}[\alpha(P,\ Q) \times w(Q) \times D(Q)]}{\sum_{Q\in\Delta(P)}[\alpha(P,\ Q) \times w(Q)]} \quad (3)$$

α는 두 점 사이 거리에 반비례한 가중치를 의미하며 D는 차분 신호, w는 해당 포인트에 대한 영향 가중치를 의미한다. 리프팅 변환은 항상 k개의 가까운 이웃점들의 속성 값을 사용하여 예측을 수행하여 복원된 기하 정보를 통해 점들 간의 예측 참조 횟수를 계산할 수 있다. 이를 통해 영향 가중치는 복원된 기하 정보를 통해 해당 점이 다른 점을 예측하는데 참조 점으로 사용될수록 큰 값을 갖도록 설계되어 업데이트 연산에 사용되고 추가적으로 각 잔차 속성값에 영향 가중치 값의 제곱근 값을 스케일링해주어 각 점들의 예측 참조 횟수에 따른 적응적인 양자화 방법을 채택하여 사용하고 있다. 또한 예측변환과 리프팅 변환은 참조 구조상 낮은 LoD의 점들이 다른 점들을 예측하는데 많이 참조 점으로 사용되는 점을 고려하여 LoD 레벨 별로 델타 양자화 파라미터를 전송하여 양자화율을 제어하는 방법이 채택되었다[17].

2) 지역 적응적 계층 변환(RAHT)

지역 적응적 계층 변환, RAHT는 Haar 웨이블릿 변환의 변형으로 2×2×2 큐브 단위로 3차원 공간상의 데이터를 x, y, z축으로의 변환을 수행한다. RAHT의 변환 과정은 다음의 식을 통해 정의된다.

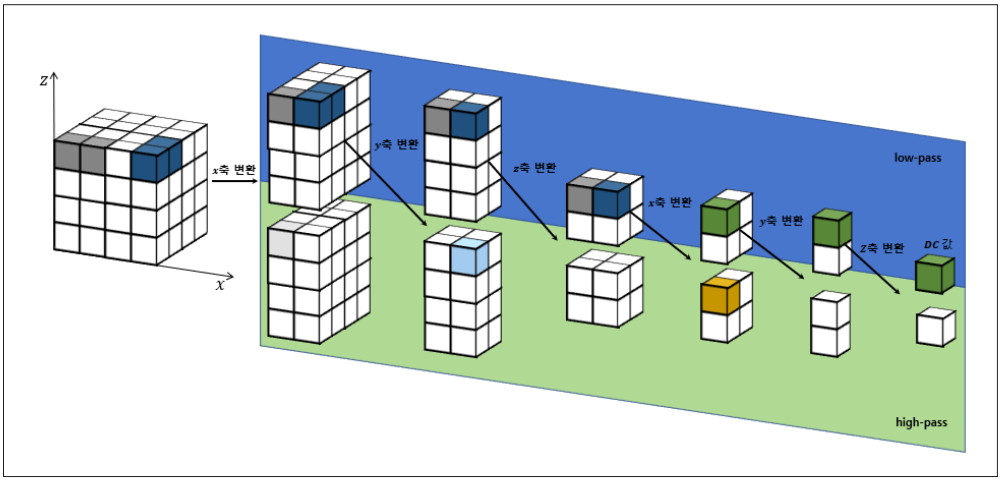
$$\begin{bmatrix} g_n^l \\ h_n^l \end{bmatrix} = T_{w1w2} \begin{bmatrix} g_{2n}^{l+1} \\ g_{2n+1}^{l+1} \end{bmatrix}, \quad (4)$$

$$T = \frac{1}{\sqrt{w_{2n}^{l+1} + w_{2n+1}^{l+1}}} \begin{bmatrix} \sqrt{w_{2n}^{l+1}} & \sqrt{w_{2n+1}^{l+1}} \\ -\sqrt{w_{2n+1}^{l+1}} & \sqrt{w_{2n}^{l+1}} \end{bmatrix}$$

$$w_n^l = w_{2n}^{l+1} + w_{2n+1}^{l+1} \quad (5)$$

각 축방향으로 인접한 두 포인트에서 저주파 성분과 고주파 성분으로 변환되며 이때 변환에 사용되는 가중치 값은 식 (5)와 같이 계산되어 밀도가 높은 공간에 대해서 높은 가중치가 할당된다. 각 축으로의 변환은 저주파수 계수가 하나의 DC값이 될 때까지 재귀적으로 변환이 수행되며 변환되는 축방향으로 인접한 포인트가 없을 시 해당 계수는 저주파 영역으로 복제된다. 〈그림 10〉은 4×4×4 영역에 대한 변환 예시이다.

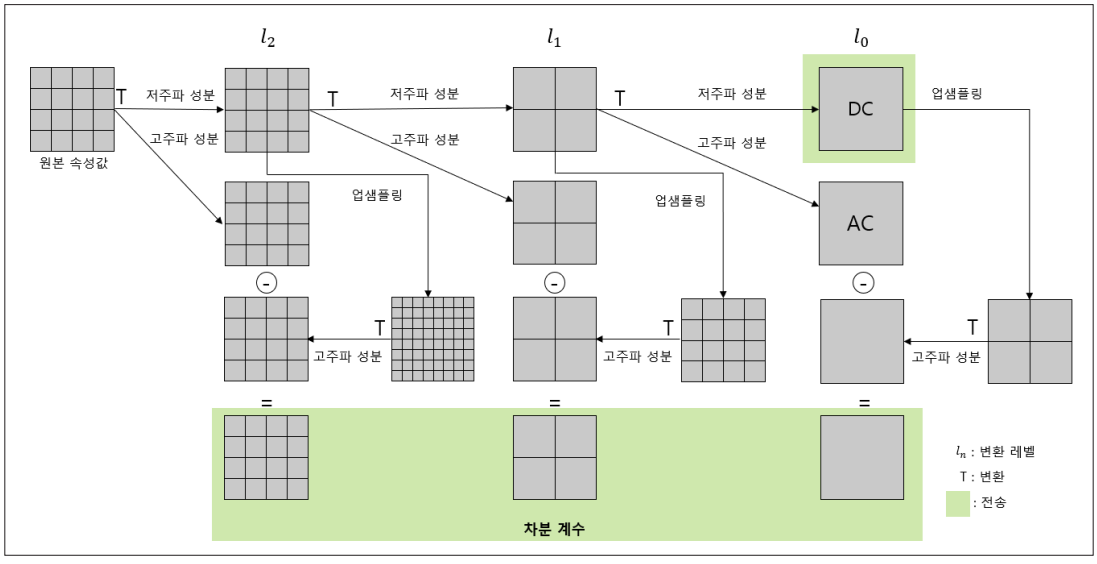
변환을 통해 생성된 저주파 계수와 고주파 계수 중 DC값과 고주파 계수가 양자화된 후 엔트로피 코딩을 통해 부호화되며, 복호화기에서는 역변환 과정을 통해 저주파수 계수를 생성하여 복원을 수행한다. 또한 현재 G-PCC는 RAHT를 통해 생성된 고주파 계수에 대한 변환 도메인 예측 기술이 채택되어 RAHT의 성능을 크게 향상시켰다[18]. 〈그림 11〉은 변환 도메인 고주파 계수



<그림 10> RAHT 4×4×4 변환 예시

예측의 부호화 과정을 이차원 영역에서 표현한 것이다.

변환 도메인 예측은 변환을 통해 생성된 l_n의 저주파 계수를 l_{n+1}과 동일한 개수를 갖도록 업 샘플링한다. 이후 업 샘플링된 계수를 변환을 통해 생성된 고주파 계수를 l_n의 고주파 계수의 예측자로 사용하여 예측을 수행한다. 변환 도메인 예측이 수행됨에 따라 기존의 고주파 계수를 전송하는 것이 아닌, 차분 고주파 계수를 전송함에 따라 전송 비트량이 크게 줄어 RAHT의 성능이 향상되었다.



<그림 11> RAHT 변환 도메인 예측의 2D 예시

III. G-PCC 성능평가

G-PCC의 기하 정보와 속성 정보 압축 기술의 성능을 비교하기 위해 참조 소프트웨어 TMC13v12.0을 통해 CTC(Common Test Condition)[4]에 따라서 성능 평가를 진행하였다. CTC의 실험 조건 중 CW(lossless geometry, lossless attribute) 조건을 통해 octree와 predictive 기술의 압축 성능을 비교했으며 C1 (lossless geometry, lossy attribute) 조건을 통해 리프팅 변환과 RAHT의 압축 성능을 비교하였다.

〈표 6〉은 octree 기반의 기하 정보 무손실 압축 대비 predictive 기하 정보 무손실 압축의 bpip(bit per input point) ratio와 부/복호화 시간을 비교한 결과이다. 실험 결과를 보면 octree 기반의 기하 정보 압축 기술이 카테고리 1 데이터에 대한 압축 성능이 더 좋으며 predictive 압축 기술의 경우 카테고리 3에서 octree 기반 압축 대비 약 14%의 비트 감소를 확인할 수 있다. 부/복호화 시간의 경우 predictive 기술이 octree보다 2배 이상의 부호화 시간이 증가하지만 복호화 시간은 20% 감소하는 것을 확인할 수 있다.

〈표 7〉은 C1조건에서 기하 정보가 무손실 octree 기반으로 압축된 상황에서의 리프팅 변환과 RAHT의 손실 압축 성능을 비교한 결과이다. 전체 성능을 비교해 보면 RAHT가 크로마 성분을 제외한 나머지 속성 정보에 대해 리프팅 변환 대비 BD-rate[19] 향상을 보였으며, 특히 밀도가 낮은 카테고리 3 데이터에서 모든 속성 값에 대해 BD-rate 향상을 보였다. 하지만 부/복호화 시간은 모두 RAHT가 리프팅 변환보다 많은 시간이 소요됨을 확인할 수 있다.

<표 6> Octree 기반 압축 대비 predictive 압축의 무손실 압축 성능 비교

CW	Geometry bpip ratio [%]
Cat1-A average	137.3%
Cat1-B average	103.7%
Cat3-fused average	85.5%
Cat3-frame average	86.9%
Overall average	104.2%
Avg. Enc Time [%]	222%
Avg. Dec Time [%]	80%

IV. 결 론

본 고에서는 MPEG에서 진행중인 포인트 클라우드 압축 표준인 G-PCC에 대해 소개하고 G-PCC에 포함

<표 7> 리프팅 변환 대비 RAHT의 손실 압축 성능 비교

C1	Attributes			
	BD rate [%]			
	Luma	Chroma Cb	Chroma Cr	Reflectance
1) Cat1-A average	1.4%	15.4%	42.3%	
1) Cat1-B average	-6.9%	10.8%	35.7%	
Cat3-fused average	-11.3%	-4.1%	-3.6%	-6.9%
2) Cat3-frame average				-6.3%
Overall average	-3.5%	12.0%	36.3%	-6.5%
Avg. Enc Time [%]	134%			
Avg. Dec Time [%]	158%			

된 핵심 압축 기술에 대해 살펴보았다. G-PCC는 초기 표준화 시작단계 대비 다양한 압축 기술들이 추가됨으로써 압축 성능 향상을 이루었다. 3차원 공간상의 기하 정보 압축 기술로는 octree 구조를 기반으로 암시적 QTBT 분할 기술, planar 모드, IDCM 모드가 채택되었으며, 주변 노드의 정보 및 라이다 취득 데이터인 경우 라이다 센서의 취득 특성을 고려하여 엔트로피 코딩의 문맥 정보를 선택하는 기술이 추가되어 압축 성능을 높였다. 또한 predictive 압축 기술이 추가되어 라이다 센서로 취득된 데이터에 대하여 octree 기반 압축 기술 대비 압축 성능을 향상시켰다. 속성 정보를 압축하는 방법으로는 예측/리프팅 변환과 RAHT가 선택적으로 사용된다. RAHT의 변환 도메인 예측이 추가됨으로써 압축 성능 향상을 이뤘으며 속성 정보 압축을 수행하기 전 라이다 센서를 통해 취득된 데이터의 특성을 고려한 좌표축으로의 변환을 적용하여 성능을 향상시키는 기술이 적용되었다.

하지만 현재 G-PCC 표준은 AI(All Intra)환경에서의 압축만을 수행하고 있어 동적 포인트 클라우드 시퀀스를 압축하는 성능에 한계가 존재한다. 이에 따라 MPEG은 2023년 완료를 목표로 G-PCC 표준 기반의 G-PCC 버전 2의 표준화를 시작할 예정이다. G-PCC 버전 2에서는 포인트 클라우드 간 시간적 중복성을 제거하여 동적 포인트 클라우드 시퀀스의 압축 성능을 향상시키는 기술들이 추가될 예정이다.

참 고 문 헌

[1] C. Cao, M. Preda, and T. Zaharia, “3D Point Cloud Compression: A Survey,” Web3D '19: The 24th International Conference on 3D Web Technology, Los Angeles, CA, 1-9 Jul. 2019

[2] MPEG 3DG, “Call for proposals for point cloud compression v2,” ISO/IEC JTC1/SC29/WG11 N16763, Hobart, AU, Apr. 2017

[3] C. Tulvan, R. Mekuria, Z. Li, and S. Laserre, “Use Cases for Point Cloud Compression,” ISO/IEC JTC1/SC29/WG11 N16331, Geneva, CH, Jun. 2016

[4] MPEG 3DG, “Common Test Conditions for PCC,” ISO/IEC JTC1/SC29/WG11 N18474, Geneva, CH, Mar. 2019

[5] MPEG 3DG, “G-PCC Test Model v12,” ISO/IEC JTC1/SC 29/WG 7 N00005, Online, Oct. 2020

[6] D. Meagher, “Geometric modeling using octree encoding,” Computer Graphics and Image Processing, vol. 19, no. 2, pp. 129-147, Jun. 1982

[7] Z. Gao, D. Flynn, A. Tourapis, and K. Mammou, “[G-PCC][New proposal] Predictive Geometry Coding,” ISO/IEC JTC1/SC29 WG11 m51012, Geneva, CH, Oct. 2019

[8] K. Mammou, A. Tourapis, J. Kim, F. Robinet, V. Valentin, Y. Su, “Proposal for improved lossy compression in TMC1,” ISO/IEC JTC1/SC29/WG11 m42640, San Diego, US, Apr. 2018

[9] G. P. Sandri, P. A. Chou, M. Krivokuća, R.L. de Queiroz, “Integer alternative for the region-adaptive hierarchical transform,” IEEE Signal Processing Letters, vol. 26, no. 9, pp. 1369-1372, Jul. 2019

[10] S. Lasserre, S. Flynn, “[GPCC][new input] Planar mode in octree-based geometry coding,” ISO/IEC JTC1/SC29/WG11 m48906, Gothenburg, SE, Jul. 2019

[11] S. Lasserre, J. Taquet, “[GPCC] [CE 13.22] related] An improvement of the planar coding mode,” ISO/IEC JTC1/SC29/WG11 m50642, Geneva, CH, Oct. 2019

[12] S. Lasserre, J. Taquet, "[GPCC] [CE13.22 related] The azimuthal coding mode," ISO/IEC JTC1/SC29/WG11 m51596, Brussels, BE, Jan. 2020
[13] S. Lasserre, D. Flynn, "[PCC] Inference of a mode using point location direct coding in TMC3," ISO/IEC JTC1/SC29/WG11 m42239, Gwangju, KR, Jan. 2018
[14] A. Vosoughi, S. Yea, S. Wenger, S. Liu, "[G-PCC][New proposal] Improved color transfer," ISO/IEC JTC1/SC29/WG11 m49407, Gothenburg, SE, Jul. 2019
[15] Hyun-Mook Oh, Sejin Oh, "[G-PCC] [EE13.43] Report on coordinate conversion," ISO/IEC JTC1/SC29/WG11 input document m54622, Online, Jul. 2020
[16] T. Sugio, "[G-PCC] Attribute residual coding in TMC13," ISO/IEC JTC1/SC29 WG11 m46108, Marrakech, MA, Jan. 2019
[17] C. D. Han, "[G-PCC] CE13.19 Report on Attribute Layer Quantization Control," ISO/IEC JTC1/SC29/WG11 m47834, Gothenburg, SE, Jul. 2019
[18] D. Flynn, S. Lasserre, "Report on up-sampled transform domain prediction in RAHT," ISO/IEC JTC1/SC29/WG11 m49380, Gothenburg, SE, Jul. 2019
[19] G. Bjøntegaard, "Calculation of Average PSNR Differences Between RD Curves," document VCEG-M33, ITU-T SG 16/Q 6, Austin, TX, 2001

필 자 소 개

변 주 형

- 2019년 2월 : 광운대학교 컴퓨터공학과 학사
- 2021년 2월 : 광운대학교 컴퓨터공학과 석사
- 2021년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전

최 한 솔

- 2018년 2월 : 광운대학교 컴퓨터공학과 학사
- 2020년 2월 : 광운대학교 컴퓨터공학과 석사
- 2020년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전

필 자 소 개

심 동 규

- 1993년 2월 : 서강대학교 전자공학과 공학사
- 1995년 2월 : 서강대학교 전자공학과 공학석사
- 1999년 2월 : 서강대학교 전자공학과 공학박사
- 1999년 3월 ~ 2000년 8월 : 현대전자 선임연구원
- 2000년 9월 ~ 2002년 3월 : 바로비젼 선임연구원
- 2002년 4월 ~ 2005년 2월 : University of Washington Senior research engineer
- 2005년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 교수
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전