



특집논문 (Special Paper)

방송공학회논문지 제29권 제6호, 2024년 11월 (JBE Vol.29, No.6, November 2024)

<https://doi.org/10.5909/JBE.2024.29.6.819>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

RAHT 노드 보간 기법을 통한 G-PCC 화면 간 속성값 예측 활성화와 압축 성능 향상

권도완^{a)}, 이성배^{a)}, 김규현^{a)*}

RAHT Node Interpolation Method for Enabling G-PCC Inter-Frame Attribute Prediction and Improving Compression Performance

Dowan Kwon^{a)}, Seongbae Rhee^{a)}, and Kyuheon Kim^{a)*}

요약

LiDAR(Light Detection and Ranging) 기술은 넓은 범위의 주변 환경 요소를 높은 정밀도로 촬영할 수 있는 기술이다. LiDAR 기술을 통하여 생성된 데이터는 다수의 점으로 구성된 포인트 클라우드로 제공되는데, 방대한 점의 수로 인하여 큰 용량을 지니고 있다. 이에 국제 표준화 단체 MPEG(Moving Picture Experts Group)은 포인트 클라우드의 처리, 저장 및 전송을 용이하게 하고자 포인트 클라우드 압축(Point Cloud Compression, PCC) 표준화를 진행하고 있다. 이 중 대표적인 PCC 표준 기술인 G-PCC(Geometry-based Point Cloud Compression)는 LiDAR 콘텐츠 압축을 하나의 목표로써 포함하고 있다. 현재 G-PCC에서는 RAHT(Region Adaptive Hierarchical Transform)를 활용한 속성값 화면 간 예측 기술 표준화가 활발히 진행되고 있으며 채택된 현재 RAHT 화면 간 예측 방식은 두 프레임에 공간상 같은 위치의 속성값이 존재할 때만 발생한다. 하지만 현재 RAHT 화면 간 예측 방식은 시간 축에 따라 동적으로 움직이는 사물이 많고 같은 물체를 표현하는 점들의 위치가 프레임마다 동일하지 않은 LiDAR 콘텐츠의 경우 실질적으로 낮은 예측 수행률을 갖는다. 따라서 본 논문은 LiDAR 콘텐츠의 RAHT 화면 간 예측률을 높여 최적의 화면 간 중복성 제거를 수행하기 위한 참조 가능 RAHT 노드 보간 방안을 제안한다.

Abstract

Light detection and ranging (LiDAR) is a technology that can capture a wide range of environmental elements with high precision. The data generated by LiDAR technology is presented as a point cloud consisting of a large number of points, which has a huge capacity due to the excessive number of points. In response, the international standardization organization MPEG (Moving Picture Experts Group) is standardizing Point Cloud Compression (PCC) to facilitate the processing, storage, and transmission of point clouds. Among them, G-PCC (Geometry-based Point Cloud Compression), a representative PCC standard technology, includes LiDAR content compression as one of its goals. Currently, G-PCC is actively standardizing the inter-prediction technique between the different frames using the Region Adaptive Hierarchical Transform (RAHT). RAHT inter-prediction method currently adopted by G-PCC only occurs when an attribute value exists at the same position in both frames simultaneously. However, the current RAHT inter-prediction method has a low prediction ratio for LiDAR content, where there are many dynamically moving objects along the time axis and the positions of points representing the same object are not the same from frame to frame. Therefore, this paper proposes a reference RAHT node interpolation scheme for optimal redundancy removal between frames with higher RAHT inter-prediction ratio of LiDAR content.

Keyword : Point Cloud Compression, G-PCC, RAHT, Inter Prediction

Copyright © 2024 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

I. 서론

LiDAR(Light Detection and Ranging)는 레이저 펄스(Pulse)를 활용하여 센서로부터 물체의 거리를 계산하고 반사된 펄스의 강도를 통하여 물체의 표면 정보를 수집하는 기술이다. LiDAR 기술을 활용하여 수집된 데이터는 사물의 위치를 표현하기 위해 펄스 반사 지점에 3차원 점을 생성하고 반사 펄스의 강도로부터 계산된 반사율(Reflectance)을 점의 속성으로 표현한다. LiDAR 센서는 수집된 다수의 점을 하나의 포인트 클라우드 콘텐츠로 제공하며, 광범위한 주변 환경을 높은 해상도로 제공하고 저조도 환경에서도 정밀한 촬영이 가능하여 자율주행, 디지털 트윈, 스마트 시티 등의 다양한 분야에서 주목받아 왔다.

현재 LiDAR 기술은 초당 최대 4.6백만 개의 포인트를 생성하는 만큼 더욱 정교해진 3D 포인트 클라우드를 제공한다. 하지만 이에 따라 증가한 데이터의 크기는 처리, 저장 및 전송에 어려움을 수반하며 상용화에 큰 제약 사항을 남겼다. 이에 국제 표준화 기구인 MPEG(Moving Picture Experts Group) 3DGH(Coding for 3D Graphics and Haptics)는 2017년에 포인트 클라우드 압축(Point Cloud Compression)^[1]을 위한 표준화를 시작하였으며, 이중 대표적인 포인트 클라우드 표준 기술인 G-PCC(Geometry-based Point Cloud Compression)^{[2][3]}에서는 LiDAR 센서로 수집한 포인트 클라우드 콘텐츠 압축을 목표로 한다. 현재 단일 프레임 내에서 화면 내 예측 기술이 적용된 G-PCC 1st

Edition은 표준화가 완료되었으며 참조 프레임을 활용한 화면 간 예측 기술이 포함된 G-PCC 2nd Edition이 지금까지 표준화 및 연구 중이다.

2022년 10월에 개최한 140차 MPEG 국제표준화 회의부터 24년 7월에 개최한 147차 회의까지 G-PCC 그룹에서는 RAHT(Region Adaptive Hierarchical Transform)를 적용한 화면 간 압축 연구가 활발하다. RAHT^[4]는 3차원 Haar 변환을 적용한 G-PCC 속성 압축 기법 중 하나로, RAHT를 활용한 화면 간 예측 기술은 기존 2D 영상 압축^{[5][6][7]}에서 사용했던 방법과는 다른 방식을 지니고 있다. 기존 2D 영상 압축에서는 움직임 정보(Motion Vector)를 찾기 위한 과정으로 움직임 추정 단계를 거쳤다면, RAHT를 활용한 화면 간 예측 기술은 움직임 추정 과정 없이 동일한 위치에 참조 가능한 속성값이 존재할 시 참조해오는 움직임 없음(Zero Motion) 방식을 채택하였다.

하지만 현 예측 방식은 LiDAR 데이터셋과 같이 점들의 분포가 불균일하고 불규칙한 콘텐츠에 대해서 낮은 수행률을 갖기 때문에 예측이 수행되지 않는 속성값들에 대한 화면 간 중복성이 여전히 존재한다. 따라서 본 논문에서는 RAHT 화면 간 예측 수행률을 활성화하여 최적의 화면 간 중복도 제거를 달성하기 위해 화면 간 예측이 수행될 수 없는 상황에서 참조 가능한 속성값을 보간하는 방안을 제안한다.

II. 배경 기술

본 논문에서 제안하는 RAHT 노드 보간 기법을 통한 G-PCC 화면 간 속성값 예측 활성화 기술은 G-PCC의 RAHT 및 이를 활용한 표준 기술들을 포함한다. G-PCC는 포인트 클라우드의 기하 정보를 우선 압축하고 복원된 기하 정보를 바탕으로 속성 정보를 압축하는 구조를 갖는다. 대표적인 G-PCC 기하 정보 압축 방법으로 옥트리 부호화(Octree Coding) 기법^{[8][9]}, 예측 트리(Predictive Geometry Tree Coding) 부호화 기법^{[10][11]}, 트라이슈프(Trisoup Coding) 부호화 기법^[12]이 있으며 속성 정보 압축 방법으로는 RAHT와 Predicting/Lifting 변환 부호화 방법^[13]이 있다. 각 부호화 기법에 따른 압축 성능은 [14][19]에서 확인할 수 있다.

a) 경희대학교(Kyung Hee University)

‡ Corresponding Author : 김규현(Kyuheon Kim)

E-mail: kyuheonkim@khu.ac.kr

Tel: +82-31-201-2442

ORCID: <https://orcid.org/0000-0003-1553-936X>

※ 이 논문의 연구 결과 중 일부는 한국방송·미디어공학회 2024년 하계 학술대회에서 발표한 바 있음.

※ This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program (Grant number: IITP-2024-2021-0-02046) and (No. RS-2023-00227431, Development of 3D space digital media standard technology) supervised by the Institute of Information & communications Technology Planning & evaluation (IITP) and some of the findings of this paper were presented at the 2024 Summer Conference of KIBME (Korea Institute of Broadcast and Media Engineers).

• Manuscript September 3, 2024; Revised October 23, 2024; Accepted October 23, 2024.

본 논문의 제안 기술은 속성 정보 압축 방법의 하나인 RAHT를 기반으로 한 기술이기에 본 절에서는 G-PCC의 RAHT와 RAHT 화면 간 예측 방법을 소개한다.

1. Region Adaptive Hierarchical Transform(RAHT)

RAHT는 Haar 변환을 바탕으로 3차원 포인트 클라운드를 계층화하면서 하나의 저주파 계수(Direct Current, DC)와 다수의 고주파 계수(Alternating Current, AC)로 변환하는 기술이다. RAHT 계층화 과정은 옥트리 분할과 유사한 방법으로 공간을 분할함으로써 진행된다. 하지만 옥트리 부호화에서는 각 노드가 8개의 서브 공간을 1비트의 점유 정보로 표현하여 총 8비트의 정보로 표현한다면, RAHT 트리의 노드는 8개의 서브 공간의 DC 계수와 각 공간에 속한 점들의 수로 표현된다. 자세히 설명하자면, 포인트 클라우드 공간 자체를 표현하는 루트 노드가 $2^M \times 2^M \times 2^M$ 차원으로 정의될 때, 깊이 d 에 속한 임의의 노드가 표현하는 공간은 $2^{M-(d+1)} \times 2^{M-(d+1)} \times 2^{M-(d+1)}$ 으로 정의된 8개의 하위 공간이 결합하여 $2^{M-d} \times 2^{M-d} \times 2^{M-d}$ 의 규칙으로 정의된다. 이때 각 하위 공간으로 정의되는 서브 노드들이 지닌 DC 계수와 각 공간 내부의 점들의 수가 RAHT 노드의 정보를 표현하게 되고 이들을 바탕으로 수식 (1)을 통하여 3D Haar 변환을 진행한다^[4].

$$\begin{bmatrix} DC \\ AC \end{bmatrix} = \frac{1}{\sqrt{w_i + w_j}} \begin{bmatrix} \sqrt{w_i} & \sqrt{w_j} \\ -\sqrt{w_j} & \sqrt{w_i} \end{bmatrix} \begin{bmatrix} DC_i \\ DC_j \end{bmatrix} \quad (1)$$

수식 (1)에서는 임의의 RAHT 노드에서 8개의 서브 노드 중 임의의 축 방향(x, y, z 축)으로 인접한 i, j 번째 노드들로 계산된 DC 계수와 AC 계수를 정의한다. w_i, w_j 는 i, j 번째 서브 노드의 공간에 속한 점들의 수이며 가중치(Weight)로 정의된다. 수식 (1)의 w 를 통하여 RAHT는 2D 이미지와 같이 제한된 격자 내에 모든 데이터가 존재하는 경우뿐만 아니라 점유/비점유 공간이 동시에 존재하는 포인트 클라우드 격자에서도 적용될 수 있다.

마지막으로 수식 (1)의 연산을 x, y, z 축 방향에 대해서 모두 적용한 임의의 RAHT 노드는 하나의 DC 계수와 최대

7개의 AC 계수를 갖게 된다. 이후 DC 계수는 다음 상위 계층의 변환을 위하여 부모 노드로 계승되며 AC 계수들은 엔트로피 부호화(Entropy Coding)^[15]된다.

2. RAHT 화면 간 예측

RAHT의 결과물 중 압축 대상은 하나의 DC 계수와 다수의 AC 계수들이다. 따라서 RAHT 화면 간 예측은 대상 포인트 클라우드의 RAHT 트리에 속한 임의의 노드로부터 출력되는 AC 계수들과 최종 DC 계수의 값을 참조 RAHT 트리로부터 예측하여 잔차(Residual)를 줄이는 것을 목표로 한다. 현재 G-PCC에서 RAHT를 활용한 AC 계수의 화면 간 예측은 움직임 추정 없이 서로 다른 포인트 클라우드 프레임의 RAHT 트리 구조상 같은 깊이, 같은 기하 정보를 갖는 RAHT 노드가 존재할 때만 예측을 진행한다. 그림 1은 기존 G-PCC의 RAHT 화면 간 예측 수행의 이해를 돕기 위한 자료이다.

그림 1은 부호화 대상 포인트 클라우드의 RAHT 트리와 참조 RAHT 트리가 주어졌을 때, DC 계수와 AC 계수의 예측 과정을 묘사하고 있다. 잔차 DC 계수 DC_{res} 는 참조 트리의 루트 노드가 가진 DC 계수 DC_{ref} 와 대상 트리의 루트 노드가 가진 DC 계수 DC_{tar} 과의 차이로 정의된다. 잔차 AC 계수 AC_{res} 의 경우, 같은 깊이 l 에 속한 RAHT 노드 중 모턴 코드(Morton Code)^[16]가 같은 경우 참조 AC 계수 AC_{ref} 와 대상 AC 계수 AC_{tar} 과의 차이로 정의된다.

하지만 LiDAR 콘텐츠는 프레임마다 RAHT 트리의 형태가 시간 축에 따라 크게 달라진다. 왜냐하면 LiDAR 센서 기반 자율주행 콘텐츠는 물체와 센서와의 거리, 차량의 속도, 센서의 프레임률(Frame rate)에 따라 연속된 프레임 내 존재하는 같은 물체일지라도 표현하는 점들의 수와 위치적 특성이 일정하지 않기 때문이다. 이러한 LiDAR 콘텐츠의 특성으로 인하여 프레임 간 두 노드가 동일한 위치를 공유할 때만 수행되는 기존 화면 간 예측 방법은 낮은 예측 수행률을 가지며 화면 간 중복성을 온전히 제거하기 어렵다. 이를 극복하는 방안으로 기존 2D 영상 압축에서 사용해왔던 움직임 벡터(Motion Vector)를 활용한 화면 간 예측 방법을 적용할 수도 있지만, 앞서

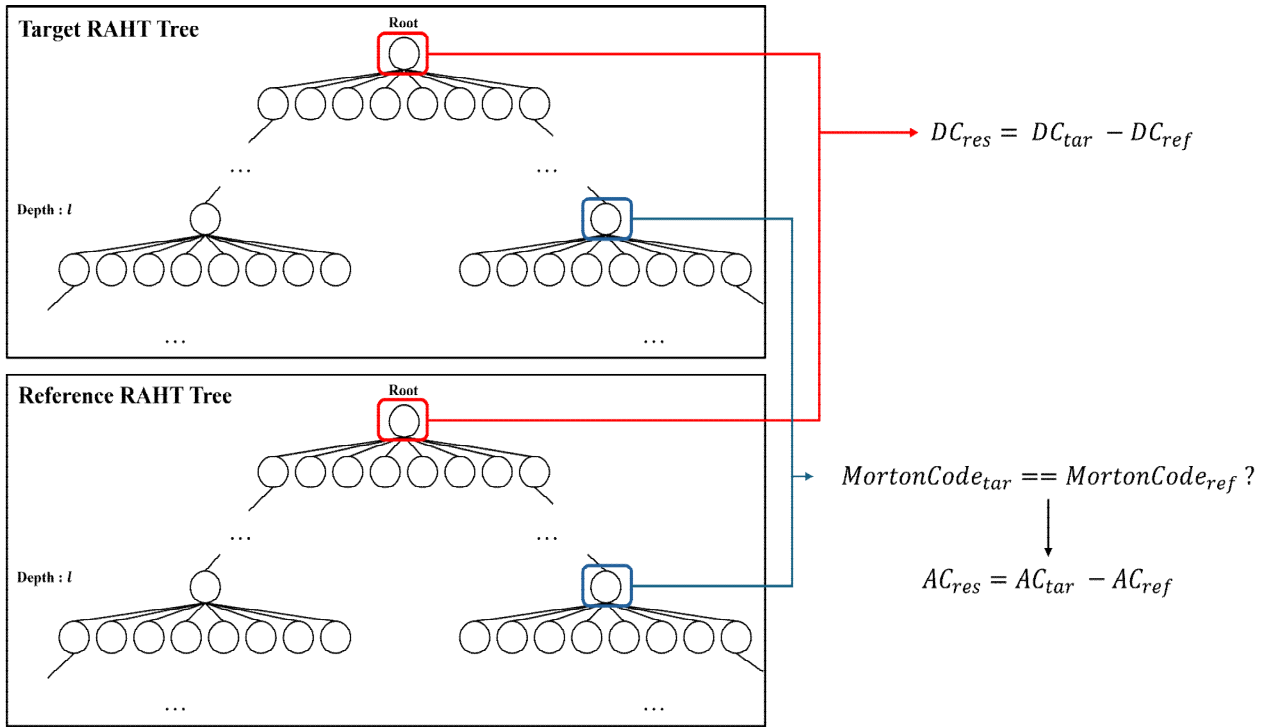


그림 1. RAHT 화면 간 예측 예시

Fig. 1. Example of RAHT inter frame prediction

언급한 LiDAR 콘텐츠의 특성에 의하여 움직임 벡터들의 분포가 다소 불규칙하고 그 수가 방대하여 효과적인 압축을 기대하기 어렵다. 이에 본 논문에서는 움직임 벡터와 같은 부가적인 정보 사용 없이 현 화면 간 예측 방법에 부합하지 않았던 노드들에 대한 예측 수행을 가능하게 하면서 최적의 화면 간 중복도를 소거하기 위한 RAHT 참조 노드 보간 방안을 소개한다.

III. RAHT 노드 보간 기법을 통한 G-PCC 화면 간 속성값 예측 활성화 기술

기존 RAHT 화면 간 예측 방법은 단순한 동작 및 연산으로 설계되어 코덱의 복잡도를 증가시키지 않으면서 압축 성능을 향상하였다. 하지만 II절에서 설명하였듯이, LiDAR 콘텐츠의 특성으로 인하여 기존 RAHT 화면 간 예측 기술만으로는 연속된 LiDAR 프레임의 속성값 중복

성을 효과적으로 제거하기 어렵다. 따라서 기존 기술을 통하여 화면 간 예측을 수행할 수 없던 RAHT 노드들에 대해서도 화면 간 예측을 수행하게 하며 보다 높은 화면 간 중복성 제거를 달성하고자 본 논문에서는 부호화 대상인 RAHT 노드와 같은 위치의 참조 RAHT 노드가 존재하지 않을 때, 해당 위치와 인접한 참조 프레임의 RAHT 노드들을 활용하여 참조 가능한 노드를 보간하는 기술인 RAHT 노드 보간 기법을 제안한다.

그림 2에서는 G-PCC의 RAHT 수행 과정을 간략히 나타낸 구조도이다. 그림에서 볼 수 있듯이, 모턴 코드로 1차원 정렬된 속성값이 입력되며 모턴 순서를 바탕으로 RAHT 트리를 구축한다. 이때 RAHT 트리는 II-1에서 언급한 구조로 생성되며 구축된 형상의 간단한 모습이 RAHT 트리 구축 모듈로부터 점선으로 표시되어 있다. 이후 RAHT 트리를 구성하는 각 노드는 화면 간 예측, 화면 내 예측^[21], 예측 없음 방법으로 부호화되며, 최종적으로 각 예측 방법으로부터 출력된 잔차들이 CABAC(Context-

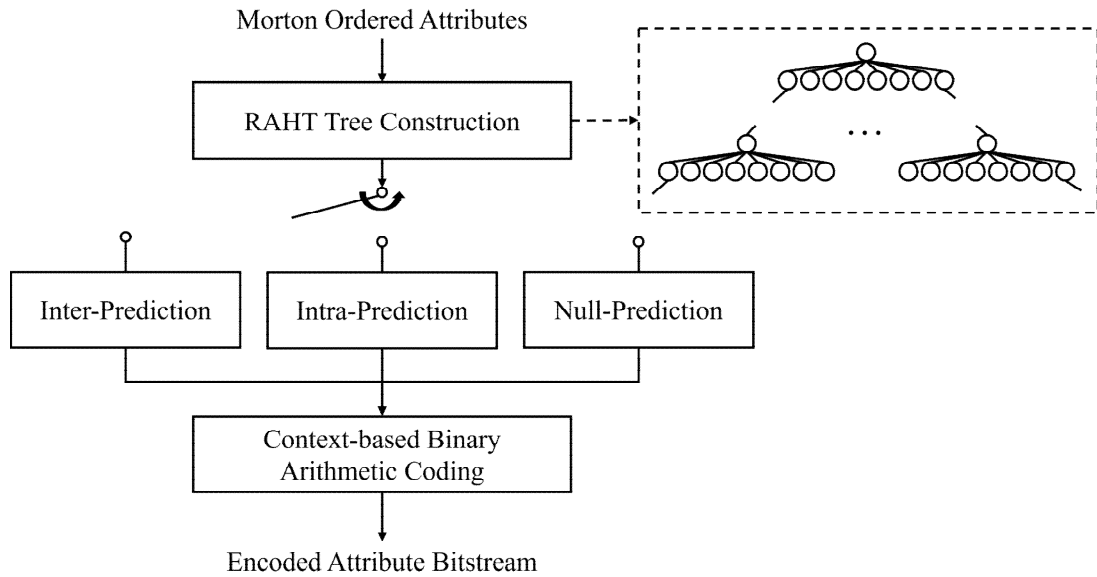


그림 2. 간략 RAHT 수행 구조도
 Fig. 2. Brief Structure of RAHT Procedure

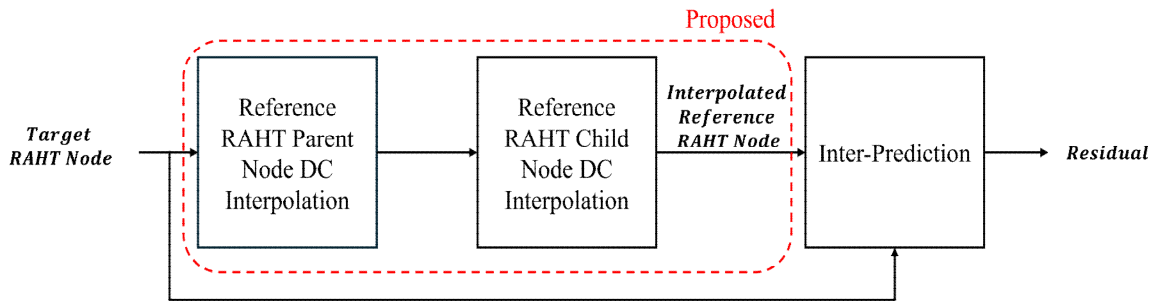


그림 3. 제안 기술이 적용된 RAHT 화면 간 예측 구조
 Fig. 3. Structure of proposed method in inter-prediction scheme

based Binary Arithmetic Coding)으로 엔트로피 부호화되어 종료된다. 그림 3은 제안 기술과 결합된 그림 2의 화면 간 예측 모듈을 묘사한다. II-2에서 언급하였듯이 RAHT 화면 간 예측 적격 여부는 부호화 대상 노드와 참조 노드의 위치가 같을 때만 시행된다. 따라서 그림 3은 적격 여부에 부합하지 않았던 부호화 대상 노드들에만 수행되며 보여지는 구조와 같이 참조 RAHT 부모 노드 보간 단계, 참조 RAHT 자식 노드 보간 단계를 통하여 참조 노드를 보간한 뒤 그림 2의 화면 간 예측 모듈과 동일한 과정을 거쳐 잔차를 계산한다.

1. 참조 RAHT 부모 노드 보간 단계

본 기술은 부호화 대상 포인트 클라우드의 공간에 속한 대상 RAHT 노드와 동일한 위치의 참조 RAHT 노드의 서브 노드들의 DC 계수 보간을 목표로 한다. 제안 기술의 첫 번째 단계로서, RAHT 서브 노드들을 포함하는 부모 노드의 DC 계수를 보간하는 예시를 그림 4에서 보여주고 있다. 그림 4의 $2 \times 2 \times 2$ 규격의 RAHT 노드는 현재 G-PCC에서 부호화 중인 계층보다 한 단계 상위에 존재하는 부모 RAHT 노드들이며, 이웃 부모 RAHT 노드들의 DC 계수를

$$N = \{n | n \in RT_{ref}, c_{morton} - \epsilon < n_{morton} < c_{morton} + \epsilon\} \quad (2)$$

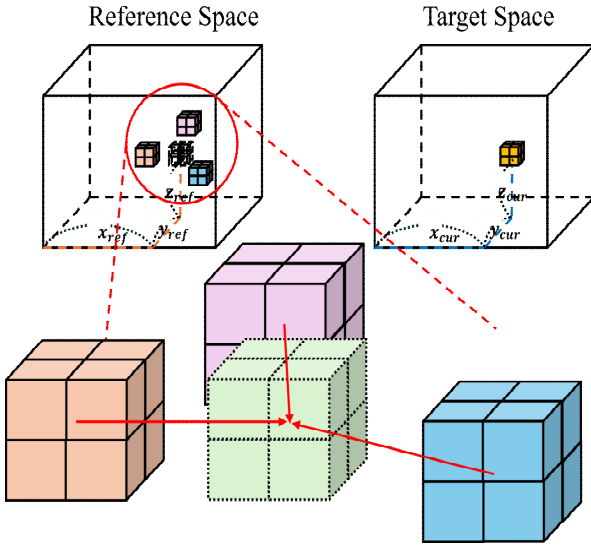


그림 4. RAHT 노드 보간 기법을 통한 G-PCC 화면 간 속성값 예측 활성화 기술 첫 단계 예시

Fig. 4. Example of first stage in RAHT Node Interpolation Method for Enabling G-PCC Inter-Frame Attribute Prediction

활용하여 현재 보간될 위치에 DC 계수를 생성하는 모습을 묘사하고 있다. 이때 수식 (2)는 보간 대상 노드 c 와 그 위치 c_{morton} 에서 해당 DC 계수를 보간하기 위해 사용되는 이웃 RAHT 노드들의 정의를 나타낸다. 참고로 c_{morton} 은 현재 부호화 대상 RAHT 노드의 부모 노드 위치와 같다.

보간 시 사용되는 참조 노드 집합 N 은 참조 RAHT 트리 RT_{ref} 에 속한 참조 노드 n 으로 구성되며, n 은 수식 (2)에서 보여지는 바와 같이 c_{morton} 와 모턴 거리로 상수 ϵ 만큼 떨어진 범위 내에 정의된다. 이때 ϵ 는 사전에 지정되는 임의의 상수이며 모턴 거리란 모턴 코드 값의 차이를 말한다. 수식 (2)에 따라 N 은 RAHT 트리 내 임의의 계층에서 RAHT 노드가 하나만 존재하는 경우를 제외한 모든 상황에서 적어도 하나 이상의 원소를 갖는다. 이에 위의 수식을 활용하여 LiDAR 콘텐츠와 같이 점들의 분포가 드문 경우에도 낮은 복잡도로 보간 시 사용될 이웃 노드들을 찾아낼 수 있다.

이어 N 에 속한 모든 참조 노드 n 을 활용하여 보간 대상 RAHT 노드 c 의 DC 계수를 수식 (3)과 같이 계산한다.

$$c_{DC} = \sum_{n \in N} \zeta(n) n_{DC} \quad (3)$$

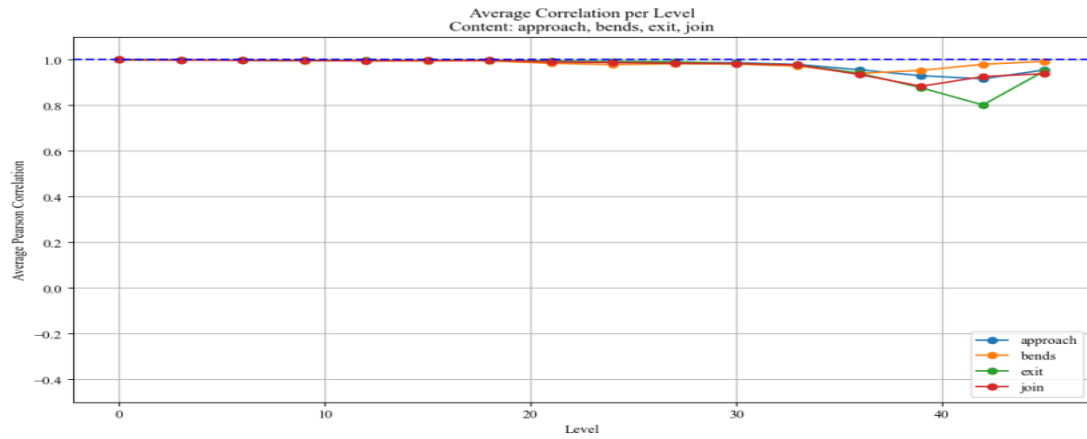
$$\text{where } \zeta(a) = \frac{\frac{1}{D(a_{morton}, c_{morton})}}{\sum_{n \in N} \frac{1}{D(n_{morton}, c_{morton})}}$$

수식 (3)에서 확인할 수 있듯이, c_{DC} 는 수식 (2)에서 정의한 참조 노드 n 들의 DC 계수의 가중합으로 정의되며 D 는 두 개의 매개변수 RAHT 노드의 모턴 코드를 입력받아 이를 3차원 좌표로 변환하여 유클리드 거리를 계산하는 함수다. 따라서 가중치 $\zeta(a)$ 은 a 에 가변적이며 N 에 속한 모든 참조 노드들과 c 와의 거리의 역수합과 a 와 c 의 거리의 역수의 비율로 정의된다. 이는 본 논문의 제안 기술이 LiDAR 콘텐츠의 RAHT 트리에 속한 임의의 노드와 주변 노드의 반사율에 대한 상관관계가 거리에 반비례한다는 가정을 바탕으로 설계되었음을 의미한다.

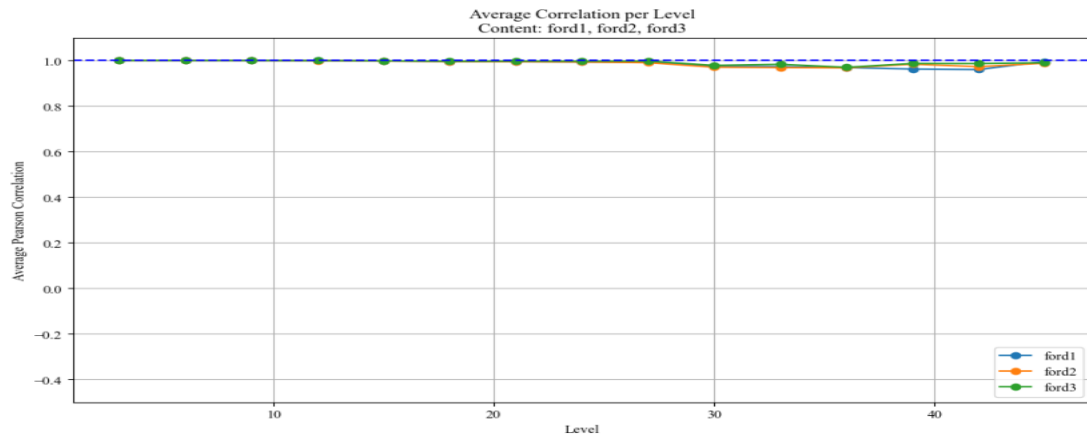
2. 참조 RAHT 자식 노드 보간 단계

제안 기술의 두 번째 단계로서, 참조 RAHT 자식 노드 보간 단계는 수식 (3)을 바탕으로 보간된 부모 노드의 DC 계수를 활용하여 부모 노드 내부의 $1 \times 1 \times 1$ 차원으로 정의되는 서브 노드(혹은 자식 노드)들의 DC 계수를 계산하기 위해서 각 서브 노드의 점유 정보와 가중치들을 정의하고 서브 노드들의 DC 계수를 보간한다. 본 제안 기술에서 각 서브 노드의 점유 정보는 부호화 대상 RAHT 노드의 서브 노드들의 점유 정보와 동일하게 사용한다. RAHT는 수식 (1)을 바탕으로 x, y, z 축에 인접한 서브 노드들을 바탕으로 DC 계수와 AC 계수를 순차적으로 계산하는데, 이 순서가 모든 RAHT 노드에 대하여 동일하다. 이에 점유 패턴이 같은 임의의 두 RAHT 노드의 AC 계수 발생 시점은 동일하기 때문에 예측 후 수행되는 Haar 도메인 상의 차분 연산이 두 노드 모두 계수가 존재하는 구간에서만 적용된다. 이러한 사실을 바탕으로 동일한 점유 패턴을 보간 노드에 적용하는 방안은 예측 과정의 차분 연산에서 불필요한 잔차 발생을 방지할 수 있음을 알 수 있다.

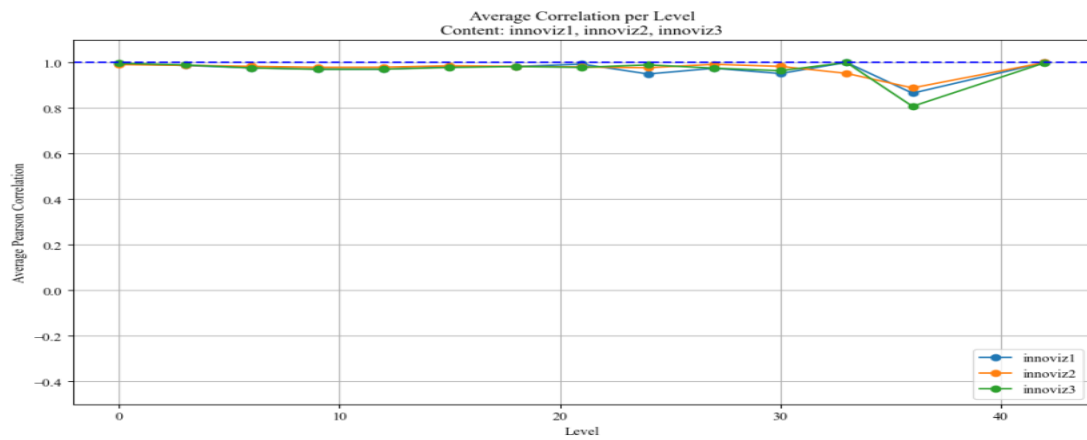
(Dowan Kwon et al.: RAHT Node Interpolation Method for Enabling G-PCC Inter-Frame Attribute Prediction and Improving Compression Performance)



(a)



(b)



(c)

그림 5. LiDAR 콘텐츠 그룹별 각 RAHT 계층에 속한 임의의 RAHT 노드의 가중치와 속성값의 평균 상관계수 (a) Qnxadas 콘텐츠 결과 (b) Ford 콘텐츠 결과 (c) InnovizQC 콘텐츠 결과

Fig. 5. The average correlation coefficient between the weights and attribute values of random RAHT node in each RAHT layer per LiDAR content group (a) Result of Qnxadas (b) Result of Ford (c) Result of InnovizQC

이어서 보간 대상 노드의 가중치들은 부호화 대상 노드의 가중치들과 동일하게 사용하였다. 이에 대한 이유는 그림 5를 통하여 설명할 수 있다. 그림 5에서는 LiDAR 콘텐츠별 임의의 RAHT 노드의 가중치들과 속성값들의 상관도를 RAHT 계층별로 분석한 결과를 보여주고 있으며, 이때 속성값들은 리프 노드일 경우 원본 반사율에 해당하며 그 외의 계층에 속한 노드일 경우 자식 노드들이 속한 계층에서 계산된 DC 계수들에 해당한다. $2 \times 2 \times 2$ 규격으로 정의된 RAHT 노드에 속한 $1 \times 1 \times 1$ 규격의 서브 노드들의 가중치와 속성값은 동일한 스캐닝(Scanning) 방식으로 1차원 정렬되었으며 각각 0~1 사이의 값들로 정규화가 된 후 피어슨 상관계수(Pearson Correlation) 측정이 진행되었다. 그림 5에서 확인할 수 있듯이, 콘텐츠들 모두 리프 노드에 가까워질수록(혹은 낮은 level일수록) 두 요소 간의 상관관계는 더욱 강해지는 것을 확인할 수 있으며, 루트 노드에 가까운 계층일수록(혹은 높은 level일수록) 가중치와 속성값 패턴의 상관도가 불규칙하지만, 여전히 높은 상관도가 존재하는 것을 알 수 있다. 따라서 본 제안 기술에서 보간된 RAHT 노드의 서브 노드의 DC 계수들은 그림 5에서 확인한 사실을 바탕으로 수식 (4)를 통하여 보간한다.

$$c_{DC}^{sub} = c_{DC} \times \frac{s_{weight}}{\sum_{a \in S} a_{weight}}, \text{ where } S = \{s | s \in RT_{tar}, s_{morton} \gg 3 = c_{morton}\} \quad (4)$$

수식 (4)에서는 보간한 RAHT 노드의 서브 노드 c_{DC}^{sub} 의 DC 계수 c_{DC}^{sub} 를 정의하며 기호 $\gg$ 는 뒤따르는 수치만큼 하위 비트를 시프팅(Shifting)하는 연산자로서 사용되었다. 노드 s 는 부호화 대상 RAHT 트리 RT_{tar} 에 속하며 자신의 모던 위치의 하위 3비트를 제외한 나머지 비트들은 보간 대상 모던 위치 c_{morton} 과 같다. 이는 기하적으로 s 보다 한 단계 상위 계층에 존재하는 c 가 정의하는 공간에 노드 s 가 포함되는 것을 의미한다. 이러한 사실을 바탕으로 수식 (4)는 앞선 수식 (3)을 통하여 보간한 c_{DC} 를 활용하여 서브 노드들의 DC 계수를 계산하는 과정을 보여주며, S 에 속한 임의의 s 의 가중치 s_{weight} 와 모든 s 의 가중치의 합

으로 나눈 비율을 활용하여 c_{DC}^{sub} 를 계산한다. 이와 같은 계산 과정은 아래 그림 6과 같다.

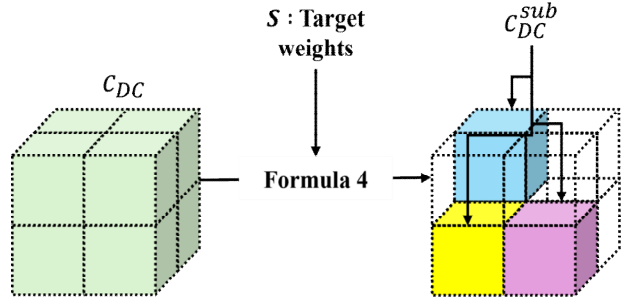


그림 6. 보간된 RAHT 부모 노드 DC 계수를 바탕으로 수식 (4)를 활용한 서브 노드 보간 예시

Fig. 6. Example of child node interpolation using Formula (4) based on interpolated RAHT parent node DC coefficients

3. 아웃라이어 처리

앞서 설명한 RAHT 노드 보간 기법을 통한 G-PCC 화면 간 속성값 예측 활성화 기술은 수식 (3)과 같이 RAHT 노드 간 거리에 따른 DC 계수의 상관도를 기반으로 설계되었기 때문에 이러한 관계성이 유의미한 범위에만 적용할 수 있다. 직관적으로, 리프 노드에 가까운 계층일수록 센서로부터 수집된 데이터에 근접하기 때문에 연속 신호 혹은 미분 가능한 신호로 수렴한다는 것을 알 수 있으며 이로부터 깊은 계층일수록 이웃 간 DC 계수의 유사도가 높을 것을 유추할 수 있다. 그림 7은 이에 대한 사실적 근거를 묘사하며 RAHT 트리의 깊이에 따라 이웃 노드 간의 DC 계수 MSE (Mean Square Error)가 달라짐을 보여준다. 이때 MSE는 정규화된 값으로 각 콘텐츠의 계층마다 계산된 MSE 중 최대 MSE를 1로 변환하고 나머지 값들을 정규화한 값이다. 얇은 깊이일수록 RAHT 노드와 이웃 노드 간의 DC 계수의 유사도가 떨어지는 추세를 확인할 수 있지만 계층이 깊어질수록, 즉 원본 포인트 클라우드의 차원과 유사해질수록, 앞서 언급한 직관과 같이 이웃 간 DC 계수 MSE가 크게 줄어드는 경향을 확인할 수 있다. 이러한 사실을 바탕으로, 본 제안 기술이 이웃 노드 간 DC MSE가 가장 적은 최하위 계층(리프 노드 계층)부터 특정 상위 계층까지만 적용될 수 있도록 기술 적용 시작 깊이 정보를 알리기 위한 상수 α 를

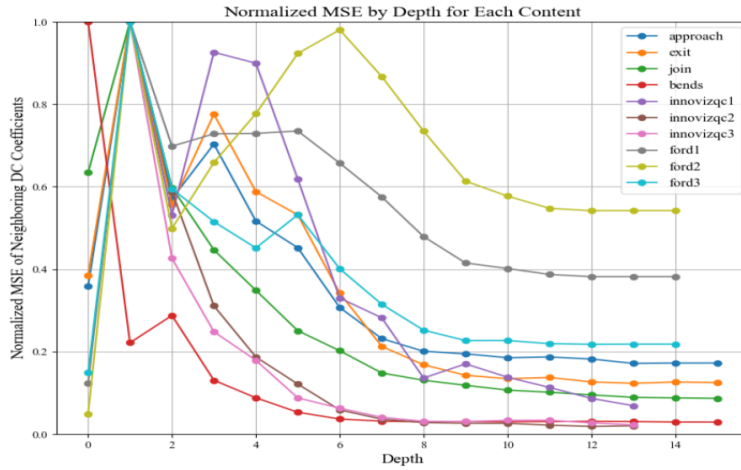


그림 7. RAHT 깊이별 임의의 노드와 이웃 노드들 간의 DC MSE
Fig. 7. DC MSE between an arbitrary node and its neighbors per RAHT depth

설정한다. 사용자 정의 상수이자 0~1 사이의 실수값을 갖는 α 는 시그널링(Signalling)을 통해 부/복호화기에서 같은 값으로 사용되며 부호화 대상 RAHT 트리의 최대 깊이에서 α 를 곱해 시작 깊이 정보를 계산하여 제안 기술의 적용 범위를 제한한다.

4. RAHT 노드 보간 기법을 통한 G-PCC 화면 간 속성값 예측 활성화 기술 최종 구조

그림 8은 G-PCC 내부에서 동작하는 제안 기술의 순서도를 묘사한다. 본 순서도는 기존 RAHT 화면 간 예측을 기반으로 제안 기술을 적용할 수 있도록 재설계되었으며 제안 기술과 관련된 부분은 모두 붉은 테두리로 표시하였다. 순서도에 등장하는 tar_pos 는 부호화 대상 RAHT 노드의 모던 코드, ref_pos 는 참조 대상 RAHT 노드의 모던 코드, tar_weights 는 부호화 대상 RAHT 노드의 가중치 정보들을 나타낸다. 또한 III-3절에서 진행한 실험 결과를 바탕으로, 이웃 노드 간의 DC 계수의 오차가 적은 계층에서만 본 기술이 적용될 수 있도록 α 를 통하여 깊이를 제한하며 window_size 는 수식 (2)의 ϵ 을 의미한다.

그 외로 대상 부호화 노드와 같은 위치에 있는 참조 노드가 존재할 시 기존 화면 간 예측이 진행되는 점은 G-PCC 기존 방법과 같으며, 존재하지 않는 경우 전체 깊이 (depth_total)에 α 를 곱한 값과 현재 RAHT 트리 깊이의 비

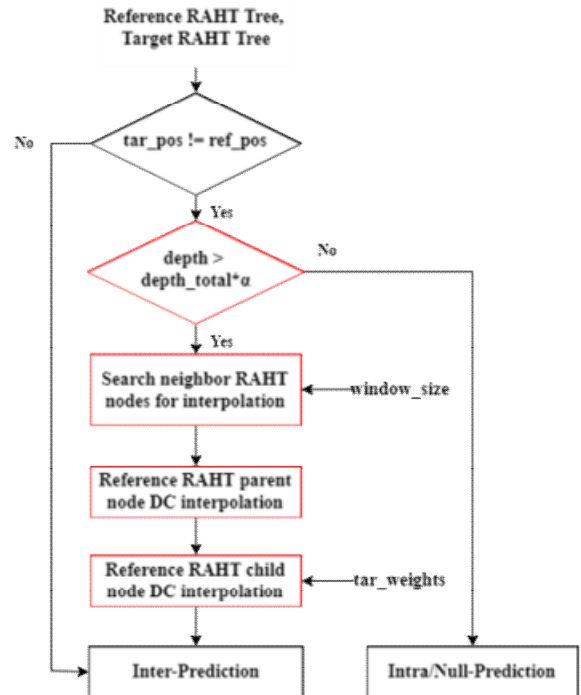


그림 8. 제안 기술 순서도
Fig. 8. Flowchart of proposed method

교 결과에 따라 제안 기술, 화면 내 예측(Intra-Prediction) 부호화 방법 혹은 어떠한 예측도 진행하지 않는 예측 없음(Null-Prediction) 부호화 방법의 수행 여부가 결정된다.

IV. 실험 결과

1. 실험 데이터셋 및 구성 환경

본 실험에 사용된 데이터셋은 G-PCC Common Test Condition(CTC)^[17]에 정의된 모든 동적 LiDAR 콘텐츠들이며, Ford_01_q1mm, Ford_02_q1mm, Ford_03_q1mm 각 1,500프레임과 BlackBerry Ltd와 QNX Software System Ltd에서 2018년도에 MPEG에 제공한 qnxadas-junction-approach 74프레임, qnxadas junction exit 74프레임, qnxadas motorway join 500프레임, qnxadas navigating bends 300프레임과 InnovizQC1, InnovizQC2, InnovizQC3^[18] 각 300프레임으로 구성된다. 더불어 G-PCC의 CTC 콘텐츠가 아닌 Cognata 회사의 무료 시뮬레이션 LiDAR 데이터셋인 LiDAR-velodyneHDL64E 콘텐츠 250프레임을 사용하여 추가적인 실험을 수행하였다. 실험에서 사용된 G-PCC

Test Model은 TMC13v26.0rc-2^{[19][20]}이고 기하 정보와 속성 정보 모두 손실 압축을 수행하는 C2 실험 환경 아래에서 수행되었으며, 기하 정보 압축은 옥트리 부호화 기법을 적용하였고 속성 정보 압축으로 RAHT를 적용하였다.

2. 실험 결과 및 분석

표 1은 본 제안 기술의 ϵ 에 따른 부/복호화기의 시간 복잡도 분석을 위하여 각 콘텐츠에 대해서 $\epsilon = 1$ 일 때의 부/복호화기 복잡도 대비 $\epsilon = 2, 3, 4$ 일 때의 복잡도 비율의 평균 백분율을 나타낸다. 표 1에서 확인할 수 있듯이 부호화기에서는 ϵ 이 1 증가할수록 약 20%의 시간 복잡도 증가율을 보였으며 복호화기에서는 약 10%의 시간 복잡도 증가율을 보였다. 또한 표 1에서 나타나지 않았지만, 압축 성능에 영향을 줄 만한 복원 PSNR 향상 혹은 비트스트림 크기 감소는 관측되지 않았다. 따라서 본 논문에서는 가장 낮은

표 1. ϵ 에 따른 제안 기술의 시간 복잡도 비교

Table 1. Time complexity comparison of proposed method for ϵ

	Encoder Time Complexity Ratio (%)			Decoder Time Complexity Ratio (%)		
	$\epsilon=2$	$\epsilon=3$	$\epsilon=4$	$\epsilon=2$	$\epsilon=3$	$\epsilon=4$
Average Time Complexity Ratio	114	137	151	110	120	132

표 2. 제안 기술 성능 평가표

Table 2. Performance evaluation of proposed method

		Geom. BD TotGeomRate [%]		End-to-End BD AttrRate [%]	Geom. BD TotalRate [%]		End-to-End BD TotalRate [%]
		D1	D2	Reflectance	D1	D2	Reflectance
am-frame spinning	ford_01_q1mm	0.0	0.0	0.0	0.0	0.0	0.0
	ford_02_q1mm	0.0	0.0	0.0	0.0	0.0	0.0
	ford_03_q1mm	0.0	0.0	0.0	0.0	0.0	0.0
	qnxadas-junction-approach	0.0	0.0	-0.3	0.0	0.0	-0.1
	qnxadas-junction-exit	0.0	0.0	-0.1	0.0	0.0	0.0
	qnxadas-motorway-join	0.0	0.0	-0.2	0.0	0.0	-0.1
	qnxadas-navigating-bends	0.0	0.0	-0.2	0.0	0.0	0.0
am-frame non-spinning	InnovizQC1	0.0	-	-0.1	0.0	-	0.0
	InnovizQC2	0.0	-	-0.1	0.0	-	0.0
	InnovizQC3	0.0	-	0.0	0.0	-	0.0
Time Complexity	Encoder	126%					
	Decoder	108%					

표 3. 시뮬레이션을 통하여 획득한 Cognata 사의 LiDAR-velodyneHDL64E 데이터셋에 대한 성능 결과

Table 3. Performance results on Cognata's LiDAR-velodyneHDL64E dataset obtained through simulation

		Geom. BD	TotGeomRate [%]	End-to-End BD	AttrRate [%]	Geom. BD	TotalRate [%]	End-to-End BD	TotalRate [%]
		D1	D2	Reflectance		D1	D2	Reflectance	
am-frame spinning	LiDAR-velodyneHDL64E	0.0	-	-0.5		0.0	-	-0.2	
Time Complexity	Encoder	134%							
	Decoder	120%							

복잡도에서 제안 기술이 수행될 수 있도록 ϵ 을 1로 설정하였다.

III-3절의 계층별 이웃 DC 계수들의 MSE 특성과 α 를 0.25, 0.5, 0.75로 나누어 실험한 결과를 바탕으로 qnxadas-junction-exit 콘텐츠에 대하여 0.75, LiDAR-velodyne HDL64E 콘텐츠에 대해서는 0.25, 그 외의 다른 콘텐츠들은 0.5로 적용하였다.

성능 평가 방법으로 압축 분야 성능 평가 방법으로 사용되는 Bjontegaard Delta-Rate^[22]를 적용하였으며, 표 2에서 볼 수 있듯이, 다수의 콘텐츠에 대해서 이득이 발생하였다. qnxadas junction approach, qnxadas junction exit, qnxadas motorway join, qnxadas navigating bends, InnovizQC1, InnovizQC2 콘텐츠에 대해서는 속성 BD-Rate(End-to-

End BD-AttrRate[%]) 0.1% 이상의 이득이 발생하였고, 그 중 2개의 콘텐츠에 대해서 전체 BD-Rate(End-to-End BD-TotalRate[%])에서도 이득이 관측되었다. Ford 1, 2, 3의 경우 적은 수치의 영향이 관측되었으며, 구체적으로는 Ford 1, 3 콘텐츠에서 각각 0.0003%, 0.0004%의 속성 BD-Rate 손실이 발견되었고, Ford 2 콘텐츠의 경우 0.0002%의 이득이 관측되었다. 또한 InnovizQC3는 각각 0.0002%, 0.0001%의 속성 BD-Rate 이득이 발생하였고 최종 평균 속성 BD-Rate로 0.1%의 비약적인 이득이 관측되었다.

표 3은 시뮬레이션을 통하여 획득한 LiDAR-velodyne HDL64E 콘텐츠에 관한 결과를 보여준다. 표 3에서 나타나듯이, 속성 BD-Rate와 전체 BD-Rate에서 각각 0.5%, 0.2%의 압축 이득이 발생한 것을 확인할 수 있으며 실제 LiDAR

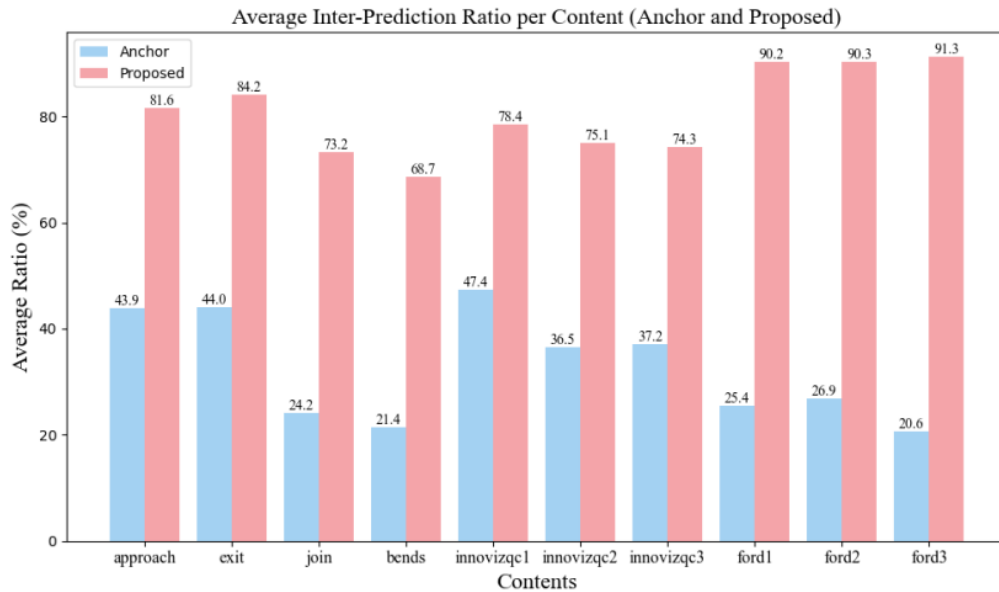


그림 9. 제안 기술과 기존 기술의 화면 간 예측률 비교

Fig. 9. Comparison of inter-prediction ratio between anchor and proposed

센서로 촬영한 콘텐츠보다 더 높은 압축 이득이 관측되었다. 이는 이웃 RAHT 노드의 DC 계수 간 거리에 따른 상관도를 바탕으로 보간되는 본 제안 기술의 특성상 LiDAR 콘텐츠가 지닌 노이즈(Noise)에 직접적인 영향을 받기 때문에 노이즈의 영향이 적은 시뮬레이션 데이터셋에서 더 큰 압축 성능 향상이 발생한 것으로 추측된다.

추가로 본 제안 기술의 화면 간 예측률 활성화 정도를 파악하기 위하여, 그림 9에서는 제안 기술과 기존 기술의 화면 간 예측률 비교를 보여준다. 그림을 통하여 확인할 수 있듯이, 제안 기술은 기존 기술보다 평균 47.96% 더 많은 예측률을 보이며 Ford 3 콘텐츠에서 70.67%의 최대 예측률 증가를 관측하였다.

V. 결 론

본 논문에서는 G-PCC의 RAHT 트리의 노드 중 화면 간 예측이 수행될 수 없는 노드들에 대해서 참조 가능한 RAHT 노드를 보간하여 포인트 클라우드 속성값의 화면 간 중복률을 더욱 감소시키고 손실 압축 환경에서의 비약적인 압축 성능 향상을 보였다. 현재 G-PCC의 RAHT 화면 간 예측은 모든 계층에 대해서 화면 내 예측을 적용한 결과와 화면 간 예측을 적용한 결과 중 RDO(Rate-Distortion Optimization) 수행 과정에서 계산된 비용이 더 적은 방법을 택하는 하이브리드(Hybrid) 방식으로 구성된다. 따라서 본 논문의 제안 기술을 적용하여 RAHT 화면 간 예측률을 활성화함으로써 두 가지 예측 방법에 대한 더욱 공정한 RDO 평가를 수행할 수 있을 것으로 기대된다. 하지만 현재 제안 기술은 세 가지 문제가 남아있다. 첫 번째로 제안 기술에서 사용되는 깊이 정보에 대한 임계값 설정 방법은 그림 7의 실험 결과와 α 가 0.25, 0.5, 0.75일 때에 대한 대략적인 실험을 바탕으로 수행되었기에 완전한 최적화라고 볼 수 없으며, RAHT 트리의 각 계층의 DC 계수와 노드들의 위치적 분포의 관계에 기반하여 효율적으로 제안 기술의 범위를 설정할 수 있는 임계값 유도 방안이 필요하다. 마지막으로 본 제안 기술은 수식 (4), (5)의 보간 알고리즘으로 인하여 높은 복잡도를 갖고 있으며 거리에 따른 DC 계수의 상관관계를 바탕으로 설계되었기에 LiDAR 센서의 하드웨

어 특성에 따라 성능의 편차가 발생할 수 있다. 이에 보간 알고리즘의 개선 및 복잡도 완화가 향후 연구로 남아있다.

참 고 문 헌 (References)

- [1] S. Schwarz et al. "Emerging MPEG Standards for Point Cloud Compression", IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol.9, No.1, pp.133-148, Mar. 2019.
doi: <https://doi.org/10.1109/JETCAS.2018.2885981>
- [2] W. Zhang, F. Yang, Y. Xu, M. Preda. "Standardization Status of MPEG Geometry-Based Point Cloud Compression (G-PCC) Edition 2", 2024 Picture Coding Symposium (PCS), Taichung, Taiwan, pp.1-5, June, 2024.
doi: <https://doi.org/10.1109/PCS60826.2024.10566443>
- [3] H. Hur. "MPEG Geometry-based Point Cloud Compression (G-PCC) Standardization Technology", Broadcasting and Media Magazine, vol.28, No.3, pp.12-22, 2023, Apr.
- [4] R. L. de Queiroz, P. A. Chou. "Compression of 3D Point Clouds Using a Region-Adaptive Hierarchical Transform", IEEE Transactions on Image Processing, vol.25, No.8, pp.3947-3956, Aug. 2016.
doi: <https://doi.org/10.1109/TIP.2016.2575005>
- [5] T. Wiegand, G. J. Sullivan, G. Bjontegaard, A. Luthra, "Overview of the H.264/AVC video coding standard", IEEE Transactions on Circuits and Systems for Video Technology, vol.13, No.7, pp.560-576, July, 2003.
doi: <https://doi.org/10.1109/TCSVT.2003.815165>
- [6] G. J. Sullivan, et al. "Overview of the high efficiency video coding (HEVC) standard", IEEE Transactions on Circuits and Systems for Video Technology, vol.22, No.12, pp.1649 - 1668, Dec. 2012.
doi: <https://doi.org/10.1109/TCSVT.2012.2221191>
- [7] B. Bross, et al. "Overview of the versatile video coding (VVC) standard and its applications", IEEE Transactions on Circuits and Systems for Video Technology, vol.31, No.10, pp.3736 - 3764, Oct. 2021.
doi: <https://doi.org/10.1109/TCSVT.2021.3101953>
- [8] Ruwen Schnabel, Reinhard Klein. "Octree-based Point-Cloud Compression", Spbg, 6:111-120, 2006.
- [9] Y. Huang, et al. "A Generic Scheme for Progressive Point Cloud Coding", IEEE Transactions on Visualization and Computer Graphics, vol.14, No.2, pp.440-453, Mar. 2008.
doi: <https://doi.org/10.1109/TVCG.2007.70441>
- [10] D. Flynn, et al. "Predictive Geometry Coding", ISO/IEC JTC1/SC29 WG7 3DG m51012, Oct. 2019.
- [11] D. Flynn and K. Mammou, "G-PCC EE13.8 report on low-latency coding for automotive/mapping applications", ISO/IEC JTC1/SC29 WG7 3DG m51012, Apr. 2020.
- [12] S. Lasserre, "Part 1 improving trisoup summary, results and perspective", ISO/IEC JTC1/SC29 WG7 3DG m59288, Mar. 2022.
- [13] K. Mammou, et al. "Proposal for Improved Lossy Compression in TMC1", ISO/IEC JTC1/SC29 WG7 3DG m42640, Apr. 2018.
- [14] "G-PCC performance evaluation and anchor results", ISO/IEC

- JTC1/SC29 WG7 3DG N00941, July, 2024.
- [15] T. Sugio, "Attribute Residual Coding in TMC13", ISO/IEC JTC1/SC29 WG7 3DG m46108, Jan. 2019.
- [16] Guy M. Morton. "A computer oriented geodetic data base and a new technique in file sequencing" IBM Germany Scientific Symposium Series, 1966.
- [17] "Common test conditions for G-PCC", ISO/IEC JTC1/SC29 WG7 3DG N00944, July, 2024.
- [18] Geert Van der Auwera, et al. "New non-spinning LiDAR sensor datasets", ISO/IEC JTC1/SC29 WG7 3DG m60706, Oct. 2022.
- [19] "G-PCC test model TM13 v26", ISO/IEC JTC1/SC29 WG7 3DG N00863, Apr. 2024.
- [20] "G-PCC 2nd edition codec description", ISO/IEC JTC1/SC29 WG7 3DG N00942, July, 2024.
- [21] D. Flynn and S. Lasserre, "CE13.18 report on upsampled transform domain prediction in RAHT", ISO/IEC JTC1/SC29 WG7 3DG m49380, July, 2019.
- [22] G. Bjontegaard, "Calculation of average PSNR differences between RDcurves," 13th VCEG-M33 Meeting ITU-T Q.6/SG16 VCEG-M33, Austin, TX, USA, Apr. 2001.
- [23] J. Kim, S. Rhee, H. Kwon, K. Kim, "LiDAR Point Cloud Compression by Vertically Placed Objects Based on Global Motion Prediction" IEEE Access, vol.10, pp. 15298-15310, Dec. 2022, doi: <https://doi.org/10.1109/ACCESS.2022.3148252>
- [24] Y. Hwang, J. Kim, K. Kim, "2D interpolation of 3D Points using Video-based Point Cloud Compression", Journal of Broadcast Engineering, vol.26, No.6, pp.692-703, 2021, Nov. doi: <https://doi.org/10.5909/JBE.2021.26.6.692>

저 자 소 개

권 도 완



- 2024년 2월 : 경희대학교 전자공학과 공학사
- 2024년 3월 ~ 현재 : 경희대학교 전자공학과 석사과정
- ORCID : <https://orcid.org/0009-0006-2150-4488>
- 주관심분야 : 영상처리, 멀티미디어 시스템, 포인트 클라우드 압축

이 성 배



- 2019년 2월 : 경희대학교 전자공학과 공학사
- 2021년 2월 : 경희대학교 전자정보융합공학과 공학석사
- 2021년 3월 ~ 현재 : 경희대학교 전자정보융합공학과 박사과정
- ORCID : <https://orcid.org/0000-0002-7392-6870>
- 주관심분야 : 딥러닝, 영상처리, 멀티미디어 시스템

김 규 현



- 1989년 2월 : 한양대학교 전자공학과 공학사
- 1992년 9월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학석사
- 1996년 7월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학박사
- 1996년 ~ 1997년 : 영국 University of Sheffield, Research Fellow
- 1997년 ~ 2006년 : 한국전자통신연구원 대화형미디어연구팀장
- 2006년 ~ 현재 : 경희대학교 전자정보대학 교수
- ORCID : <http://orcid.org/0000-0003-1553-936X>
- 주관심분야 : 디지털 방송, 영상처리, 멀티미디어 통신, 디지털 대화형 방송