

G-PCC RAHT 화면 간 속성 정보 압축 : 기하 정보 기반 지역 움직임 적용 방안 연구

권도완, 이성배, 김규현*

경희대학교, 경희대학교, *경희대학교

kdwys97@khu.ac.kr, rhee@khu.ac.kr, *kyuheonkim@khu.ac.kr

RAHT inter frame attribute coding in G-PCC : Study of local motion applications using geometry-based information

Kwon Dowan, Rhee Seongbae, Kim Kyuheon*

Kyunghee Univ., Kyunghee Univ., *Kyunghee Univ.

요약

최근 AR/VR 산업 활성화에 따라 3D 데이터 활용도가 높아지고 있다. 이에 MPEG (Moving Picture Expert Group) 표준화 단체에서는 3D 데이터의 한 종류인 포인트 클라우드 데이터 압축을 위한 표준화 진행이 활발하다. 포인트 클라우드는 기하 정보와 속성 정보를 가지는 점 데이터들의 집합으로 이를 위한 대표적인 압축 표준화 기술로 포인트 클라우드의 3차원 공간적 특성을 활용하여 새로운 압축 방법을 적용한 Geometry-based Point Cloud Compression (G-PCC) 기술이 존재한다. G-PCC 기술에서 포인트 클라우드는 기하 정보 압축과 속성 정보 압축으로 나누어 부/복호화를 진행하며, 현재 G-PCC 기술은 화면 간 속성 압축 기술 고도화에 집중하고 있다. G-PCC의 속성 압축 기술 중 Region Adaptive Hierarchical Transform (RAHT)의 화면 간 예측 기술은 현재 움직임 정보가 없이 진행된다. 따라서 본 논문에서는 G-PCC 기술 중 속성값 압축을 위해 사용되는 RAHT 기술의 고도화를 위하여 RAHT 화면 간 예측 기술에 움직임 정보를 적용하는 연구를 진행하였다. 괄목할만한 BD-Rate (Bjontegaard Delta Rate)은 관측되지 않았지만, 몇 콘텐츠에만 기존 RAHT 대비 좋은 성능을 보이는 결과가 관측되었으며, RAHT 화면 간 예측의 새로운 방법을 확보하였다.

I. 서론

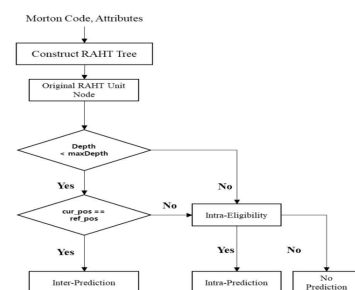
포인트 클라우드는 3차원 공간에 속한 데이터점들의 집합으로 3차원 공간상의 위치를 표현하는 기하 정보와 RGB나 반사율과 같은 고유한 속성 정보를 가지고 있다. 포인트 클라우드는 데이터점들을 통하여 형상을 나타낼 수 있다는 점에서 자율주행, 로봇 주행, 3D 콘텐츠 개발 등 다양한 분야에서 활용도가 높아지고 있다. 하지만 수집된 포인트 클라우드 콘텐츠는 2D 데이터에 비해 크기가 방대하여 저장 및 전송에 큰 어려움을 겪고 있다. 이에 국제 표준화 기구인 ISO/IEC JTC1 산하의 MPEG에서는 포인트 클라우드 압축을 진행하기 위하여 3차원 공간의 특성을 활용하여 압축을 진행하는 Geometry-based Point Cloud Compression (G-PCC) 기술을 개발해왔다.

G-PCC 압축 기술[1]은 기하 정보를 압축한 후, 복원된 기하 정보를 바탕으로 속성 정보를 압축하며, 속성 정보 압축 방법으로 예측/리프팅 변환과 Region Adaptive Hierarchical Transform (RAHT)[2]를 표준으로 구성하고 있다. 이 중 RAHT는 2차원 Haar Wavelet 변환을 3차원 점에 적용하는 속성 압축 방법으로, 포인트 클라우드 프레임 자체를 루트 (Root)로 설정하고 이후 재귀적으로 8개의 정육면체 형태의 공간으로 나누는 팔진 트리 (Octree) 구조를 활용하여 각 트리 레벨 별로 Haar Transform을 적용하는 기술이다. 최근 G-PCC 그룹 내에서 RAHT 화면 간 예측 기술[1]이 주요 쟁점으로 떠오르고 있다. 하지만 현재 RAHT 화면 간 예측에서는 기존 2D 영상의 화면 간 예측 기술과 다르게 움직임 정보를 적용하지 않고 같은 기하 정보를 갖는 점이 존재한다면 해당 점의 속성 정보를 참조한다. 이에 본 논문에서는 RAHT 화면 간 예측에

움직임 정보를 적용하였을 때, 기존 RAHT의 성능이 향상될 수 있을지에 대한 가능성을 확인하기 위하여 RAHT 화면 간 예측 기술의 새로운 구조를 적용하여 실험을 진행하였다.

II. 본론

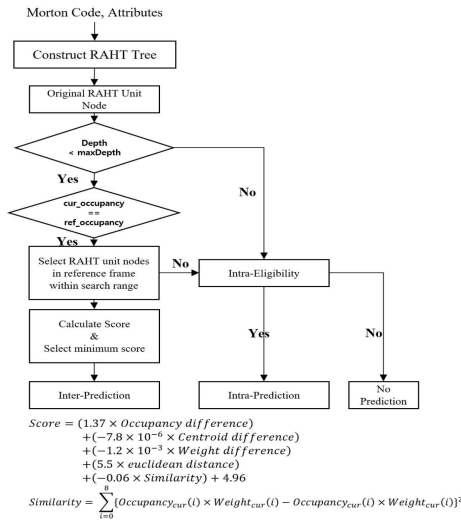
기존 RAHT의 순서도는 그림1[1]과 같다. 현재 프레임과 참조 프레임의 기하 정보를 나타내는 모턴 코드 (Morton code)와 속성 정보를 바탕으로 RAHT 팔진 트리를 구성하고 제한된 트리 깊이(maxDepth)까지 RAHT 화면 간 예측이 진행된다.



<그림1. 기존 RAHT 화면 간 예측 순서도>

하지만 깊이 정보 비교를 거친 후, 같은 모턴 코드를 가진 참조 노드가 존재할 때 RAHT 화면 간 예측이 진행되는 것을 그림1에서 확인할 수 있다. 이러한 방식은 실질적으로 프레임 간 움직임 정보가 없을 때, 즉 일정한 시간 동안 물체가 움직이지 않을 때 적합한 방법이다. 하지만 LIDAR 센서와 같이 이동형 물체에 탑재된 상태에서 데이터를 수집할

경우 많은 물체의 움직임이 감지될 것으로 예상할 수 있기에 해당 방식은 타당하지 않다. 이에 본 논문은 그림2에서 보여주고 있는 새로운 RAHT 화면 간 예측 구조도를 제안한다.



<그림2. 제안 RAHT 화면 간 예측 순서도와 Score>

앞서 설명하였다시피, RAHT는 팔진 트리로 구성되기에, 하나의 노드는 8개의 자식 노드를 가지며 자식 노드들은 부모 노드의 점유 공간 안에 존재하는 노드들이다. 그리고 기하 정보 압축이 끝나고 속성 정보 압축이 진행되기 때문에 각 노드의 위치 또한 알 수 있다. 따라서 두 노드 간의 점유 정보 차이 (Occupancy difference), 공간에 속한 점 개수의 차이 (Weight difference), 노드 간 거리 (Euclidean distance), 무게 중심 차이 (Centroid difference) 계산이 가능해진다. 그림2에서 볼 수 있듯이, 그림1과 달리 RAHT 트리 구조의 참조 노드와 현재 노드의 위치 비교 후, 참조 노드 주변의 노드 중 임의로 설정한 검색 범위 내의 노드들을 모두 참조 후보 노드들로 지정한다. 이는 물체의 지역적 움직임이 발생하였을 때 현재 노드와 비교하여 같은 위치의 참조 노드보다 더 유사한 참조 노드를 찾아오기 위함이다. 이후 참조 후보 노드들의 기하학적 정보를 바탕으로 점수를 계산하고, 점수를 바탕으로 현재 노드와 가장 유사한 형태를 가진 노드를 참조한다. 실제 속성값의 비교가 아닌 기하 정보를 바탕으로 지역 움직임을 추정하는 이유는 방대한 양의 포인트 수로 인해 부호화 대상 노드마다 지역 움직임 정보를 생성하여 압축하는 방법은 압축 성능에 악영향을 줄 수 있기 때문이다. 이에 앞서 설명한 기하 정보를 비교하여 적절한 참조 노드를 선택하는 방법을 적용하였다. 사용되는 점수는 다중 회귀 분석을 통하여 실제 속성값과 참조 노드의 속성값의 차이를 점유 정보 차이, 점 개수의 차이, 무게 중심 차이 등의 기하 정보로 설명한 수식으로 생성된 값이다. 추가로 유사도 (Similarity)는 본 논문에서 사용된 값으로 각 노드의 자식 노드들의 점유 정보와 포함하고 있는 점 개수의 곱의 MSE로 표현된다.

본 논문에서는 G-PCC Common Test Condition (CTC)[3]에 제시된 환경에서 진행하였으며, TMC13v24-rc1 테스트 모델을 사용하였다. 테스트 콘텐츠는 LiDAR spinning 콘텐츠 (BB-QNX)와 LiDAR non-spinning (InnovizQC) 콘텐츠에 대하여 32 개의 연속 프레임에 사용하였다. 또한 기하 정보 압축 방식은 옥트리 부호화 (Octree Coding)[1], 속성 정보 압축 방식은 RAHT로 진행하였다.

본 논문에서 제안한 순서도에 따른 결과는 표1과 같다. 속성 부/복호화

단에서 진행되는 실험이기에 표1에서 볼 수 있듯이 기하 정보에 관한 결과는 변화가 없으며 속성 정보 결과에 해당하는 속성 BD-Rate 결과값에 대하여 변화가 관측되었다. qnxadas junction approach와 qnxadas junction exit 콘텐츠의 경우 -0.1%의 이득이 발생하였으나, qnxadas motorway join의 경우 0.2%의 손실이 발생하였고, InnovizQC1 콘텐츠의 경우 압축 성능에서 변화가 관측되지 않았다.

Content (r01~r06, 32 sequence)	Geom. BD Total GeomRate [%]		End-to-End BD AttrRate [%]	Category
	D1	D2	Reflectance	
qnxadas junction approach	0	0	-0.1	Spinning LiDAR
qnxadas junction exit	0	0	-0.1	
qnxadas motorway join	0	0	0.2	
InnovizQC1	0	0	0.0	Non-Spinning LiDAR

<표1. 제안된 RAHT 적용 결과>

III. 결론

본 논문은 움직임 정보를 사용하지 않는 기존 RAHT 화면 간 예측 기술에 기하 정보를 바탕으로 지역 움직임을 적용하는 데에 의의가 있다. 하지만 압축 성능 면에서, 해당 방법으로 다양한 콘텐츠에 일정하고 좋은 성능을 보장하기에는 다소 어려움이 있다. 그 이유는 LiDAR 콘텐츠마다 특성이 상이하며 사용된 기하 정보들만으로는 다양한 LiDAR 콘텐츠의 속성값 특성을 아우르기에 부족하기 때문이다. 따라서 추후 속성 정보와 상관도가 높은 새로운 기하 정보를 생성하여 적용해보는 후속 연구를 진행할 예정이다.

ACKNOWLEDGMENT

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-2024-00227431, Development of 3D space digital media standard technology) supervised by the Institute of Information & communications Technology Planning & evaluation (IITP).

참고 문헌

- [1] G-PCC 2nd edition codec description, document N00648, ISO/IEC JTC1/SC29 WG7 3DG, July. 2023.
- [2] R. L. de Queiroz and P. A. Chou, "Compression of 3D Point Clouds Using a Region-Adaptive Hierarchical Transform," in IEEE Transactions on Image Processing, vol. 25, no. 8, pp. 3947-3956, Aug. 2016
- [3] Common test conditions for G-PCC, document N00722, ISO/IEC JTC1/SC29 WG7 3DG, November. 2023
- [4] G-PCC Test Model v24, document N00717, ISO/IEC JTC1/SC29 WG7 3DG, November. 2023