

공간 도메인 기반 캡슐화 방안

이상민 남귀중 이성배 *김규현

경희대학교

posing1606@khu.ac.kr nkj0427@khu.ac.kr rhee@khu.ac.kr kyuheonkim@khu.ac.kr

Spatial domain-based encapsulation scheme

Sangmin Lee, Kwijung Nam, Seongbae Rhee, Kyuheon Kim

Kyunghee University

요약

포인트 클라우드 데이터는 자율 주행 기술, 가상 현실 및 증강 현실에서 사용될 3차원 미디어 중 하나로 각광 받고 있다. 국제 표준화 기구인 MPEG(Moving Picture Expert Group)에서는 포인트 클라우드 데이터의 효율적인 압축을 위해 G-PCC(Geometry-based Point Cloud Compression) 및 V-PCC(Video-based Point Cloud Compression)의 표준화를 진행 중에 있다. 그 중, G-PCC는 본래 단일 프레임의 압축을 수행하는 정지 영상 압축 방식이지만, LiDAR(Light Detection And Ranging) 센서를 통해 획득된 동적 포인트 클라우드 프레임에 대한 압축의 필요성이 대두됨에 따라 G-PCC 그룹에서는 Inter-EM(Exploratory Model)을 신설하여 LiDAR 포인트 클라우드 프레임의 압축에 관한 연구를 시작하였다. Inter-EM의 압축 비트스트림은 G-PCC 비트스트림과 마찬가지로 효과적인 전송 및 소비를 위해 미디어 저장 포맷인 ISOBMFF(ISO-based Media File Format)으로 캡슐화될 수 있다. 이때, 포인트 클라우드 프레임들은 자율 주행 등의 서비스에 사용하기 위해 시간 도메인뿐만 아니라 공간 도메인을 기반으로도 소비될 수 있어야 하지만, 공간 도메인을 기반으로 콘텐츠를 임의 접근하여 소비하는 방식은 기존 2D 영상의 시간 도메인 기반 소비 방식과 차이로 인해 기존에 논의된 G-PCC 캡슐화 방안만으로는 지원이 제한된다. 이에, 본 논문에서는 G-PCC 콘텐츠를 공간 도메인에 따라 소비하기 위한 ISOBMFF 캡슐화 방안에 대한 파일 포맷을 제안하고자 한다.

1. 서론

포인트 클라우드는 3D 스캐너 또는 LiDAR(Light Detection And Ranging)를 통해 획득되는 다수의 점으로 구성되는 데이터로써, 3차원 형태인 좌표계의 x, y, z 축을 표현한 기하 정보와 반사율, RGB 값인 속성 정보로 구성된다. 이러한 포인트 클라우드 데이터는 수만에서 수억 개의 포인트로 구성되기 때문에 일반적인 2차원 미디어 데이터에 비해 더 큰 데이터 크기를 가진다. 이와 같이 거대한 용량의 포인트 클라우드 데이터를 효과적으로 활용하기 위해서는 포인트 클라우드의 압축이 필요하다. 이에 따라 국제 표준화 기구인 MPEG(Moving Picture Expert Group)에서는 G-PCC(Geometry-based Point Cloud Compression)[1] 및 V-PCC(Video-based Point Cloud Compression)[2]의 표준화를 진행 중에 있다. 그 중, G-PCC는 시간에 따라 변하지 않는 정적 개체 및 장면 포인트 클라우드를 압축하는 방안으로 제정되었다.

한편, 자율 주행[3] 서비스를 위해 LiDAR 센서 기반 포인트 클라우드가 주목받게 됨에 따라, G-PCC는 LiDAR 센서 기반의 데이터를 효율적으로 압축하기 위해 Inter-EM(Exploratory Model)[4]을 신설하여 LiDAR 센서 기반 포인트 클라우드 프레임의 압축에 관한 연구를 시작하였다. G-PCC의 비트스트림은 효율적인 전송을 위해 ISOBMFF(ISO-based Media File Format)[5]의 표준에 따라 캡슐화하는 방안을 사용하고 있지만, 기존의 미디어 콘텐츠 대부분은 2D 영상

으로써 시간 도메인에 따라 소비되었기에, ISOBMFF와 같은 미디어 포맷은 시간 도메인을 기반으로 한 방식만을 지원하고 있다. 그러나 G-PCC 콘텐츠는 사용자의 위치 정보를 바탕으로 한 서비스를 지원 가능해야 하므로, 시간 도메인뿐만 아니라 공간 도메인을 기반으로도 소비될 수 있어야 한다. 이에, 본 논문에서는 G-PCC 콘텐츠를 공간 도메인에 따라 소비하기 위해 ISOBMFF 캡슐화 방안을 제안하고자 한다.

2. 본론

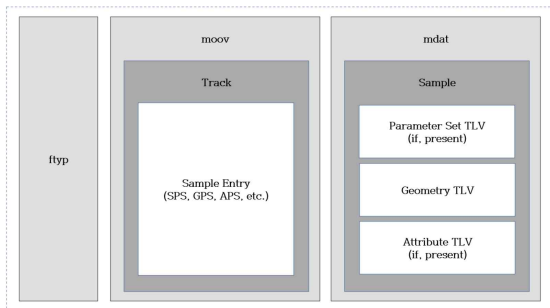
G-PCC 비트스트림[6]은 <그림 1>과 같이 TLV(Type Length Value) 단위로 구성되며, SPS(Sequence Parameter Set), GPS(Geometry Parameter Set), Geometry Data, APS(Attribute Parameter Set), Attribute Data, Tile Inventory, Boundary Marker, Constant Attribute, User Data 타입을 가진다.

tlv_type	Description
0	Sequence Parameter Set
1	Geometry Parameter Set
2	Geometry Data
3	Attribute Parameter Set
4	Attribute Data
5	Tile Inventory
6	Boundary Marker
7	Constant Attribute
8	User Data

tlv_encapsulation[]	Descriptor
tlv_type	u8[]
tlv_num_payload_bytes	u32
for(i = 0; i < tlv_num_payload_bytes; i++)	u8[]
tlv_payload_byte[i]	u8[]

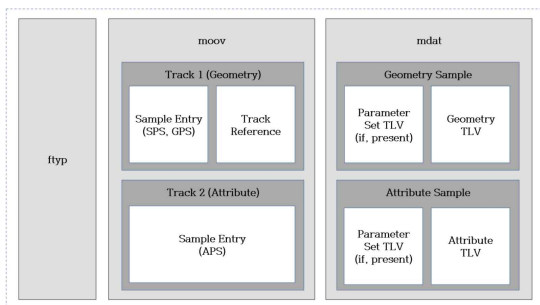
<그림 1> G-PCC 비트스트림 구조

G-PCC 비트스트림 캡슐화의 형태는 두 가지로 분류되는데 단일 트랙 캡슐화와 다중 트랙 캡슐화로 볼 수 있다. 단일 트랙 캡슐화는 <그림 2>와 같이 한 개의 트랙 단위에 기하 정보 및 속성 정보를 모두 저장하는 형식이며, 샘플은 G-PCC TLV 단위로 구성된다.



<그림 2> G-PCC 비트스트림의 단일 트랙 캡슐화 파일 구조

다중 트랙 캡슐화는 <그림 3>과 같이 기하 정보와 속성 정보를 각각 다른 트랙에 저장하는 형식이고, 압축 결과물 또한 Geometry TLV와 Attribute TLV를 구분하여 각각 다른 트랙으로 구성된다.



<그림 3> G-PCC 비트스트림의 다중 트랙 캡슐화 파일 구조

G-PCC ISOBMFF에서 G-PCC 콘텐츠의 공간 도메인에 따른 소비를 위해서 본 논문에서는 공간 정보 박스를 제안하고자 한다. 프레임 공간 정보 박스 생성은 ISOBMFF 표준의 SampleEntry 박스 내부에 Frame Spatial Information 박스('gpfs')를 생성하여 3차원 공간 데이터 프레임의 위치 정보를 저장하여 공간적 임의 접근을 지원할 수 있다.

```

Definition
Box Types:          'gpfs'
Container:          GPCCSampleEntry ('gpe1', 'gpeg', 'gpc1', 'gpcg', 'gpeb', 'gpcb')
Mandatory:         No
Quantity:          Zero or one

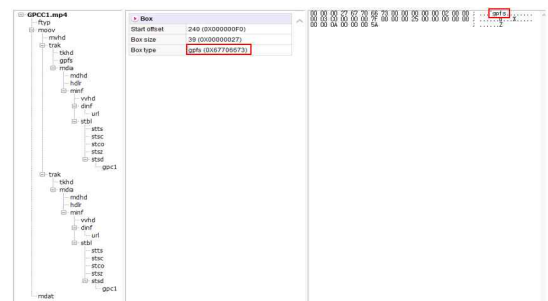
Syntax
aligned(8) class GPCCFrameSpatialInfoBox extends FullBox('gpfs', 0, 0) {
    unsigned int(15)  num_SpatialInfo;
    bool(1)          frame_altitude_skip_flag;
    bool(1)          frame_speed_skip_flag;
    bool(1)          frame_direction_skip_flag;
    unsigned char(6) reserved;
    for (int i=0; i < num_SpatialInfo; i++) {
        unsigned int(32) frame_latitude;
        unsigned int(32) frame_longitude;
        if(!frame_altitude_skip_flag) {
            unsigned int(32) frame_altitude;
        }
        if(!frame_speed_skip_flag) {
            unsigned int(32) frame_speed;
        }
        if(!frame_direction_skip_flag) {
            unsigned int(32) frame_direction;
        }
    }
}
    
```

<그림 4> Frame Spatial Information 박스('gpfs')의 구조

<그림 4>에서 보는 바와 같이, 만들어진 프레임 공간 정보 박스는 LiDAR 등의 센서를 통해 획득된 3차원 공간 데이터 프레임의 절대 좌표 (GPS) 또는, 3차원 공간 데이터 영상 그룹의 상대 좌표를 포함한다. Frame Spatial Information 박스의 num_SpatialInfo는 위치 정보를 갖는 3차원 공간 데이터 프레임의 수 정보를 나타낸다. frame altitude skip flag, frame speed skip flag, frame direction skip flag는 각각 frame altitude, frame speed, frame direction 파라미터의 스킵 여부에 대한 정보를 나타낸다. frame latitude와 frame longitude는 프레임의 위도 및 경도에 대한 정보를 나타내며, frame altitude skip flag, frame speed skip flag, frame direction skip flag는 프레임의 고도, 속도 및 방향에 대한 정보를 나타낸다. 생성된 박스에서 제공된 3차원 공간 데이터 프레임의 위치 정보를 통해 사용하는 위치를 기반으로 3차원 공간 데이터의 공간적 임의 접근을 수행하고 원하는 콘텐츠를 제공 받아 소비할 수 있다.

3. 실험 결과

본 논문은 G-PCC 콘텐츠의 파일 포맷의 spatial random access를 위한 ISOBMFF 캡슐화 방안 제안하였다. 제안한 방안이 타당성을 검증하기 위해 포인트 클라우드의 캡슐화 결과 확인과 공간 도메인을 기반으로 소비 여부, 캡슐화된 파일의 디코딩을 실험을 진행하였다.



<그림 5> G-PCC 비트스트림 'gpfs' 캡슐화 결과

제안한 방안에 따라 G-PCC 비트스트림을 ISOBMFF에 위치 정보를 담는 박스를 적용하여 캡슐화했다. <그림 5>는 ISOBMFF 표준의

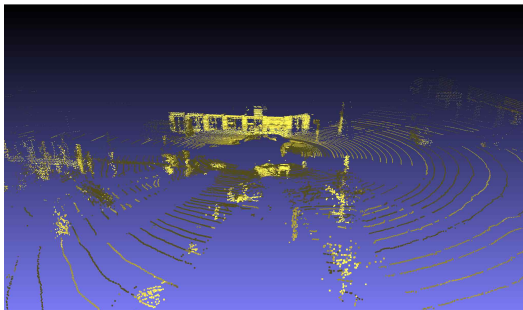
SampleEntry 박스 내부에 'gpsfs' 박스를 보여주고 박스의 내부 정보로 공간 정보를 비트 화하여 나타내주고 있다.

```

Microsoft Visual Studio 디버그 콘솔
MPEG-G-PCC
SampleCount : 20
1. longitude : 127.081192, latitude : 37.243145
2. longitude : 127.081207, latitude : 37.243138
3. longitude : 127.081026, latitude : 37.243152
4. longitude : 127.080799, latitude : 37.243152
5. longitude : 127.080577, latitude : 37.243150
6. longitude : 127.080744, latitude : 37.243303
7. longitude : 127.080677, latitude : 37.243434
8. longitude : 127.080555, latitude : 37.243532
9. longitude : 127.080319, latitude : 37.243510
10. longitude : 127.079580, latitude : 37.243506
11. longitude : 127.080637, latitude : 37.243150
12. longitude : 127.080974, latitude : 37.243505
13. longitude : 127.080677, latitude : 37.243434
14. longitude : 127.080395, latitude : 37.243502
15. longitude : 127.080019, latitude : 37.243510
16. longitude : 127.079580, latitude : 37.243506
17. longitude : 127.079590, latitude : 37.243506
18. longitude : 127.079580, latitude : 37.243506
19. longitude : 127.079580, latitude : 37.243506
20. longitude : 127.079580, latitude : 37.243506
-----
size : 8 sample
Decoding : 3
>rec_3 bin
  
```

<그림 6> 공간 도메인 샘플 결과

또한 생성된 ISOBMFF 파일을 통해 위치 정보를 확인할 수 있다. <그림 6>은 공간 도메인을 기반한 소비의 검증이다. 샘플들은 위치 정보로 확인할 수 있고 원하는 GPS 값을 가지는 샘플을 G-PCC 비트스트림으로 출력하여 복원된 비트스트림을 소비할 수 있음을 확인하였다.



<그림 7> ISOBMFF 기반의 G-PCC decoding 결과

생성된 ISOBMFF 파일로 G-PCC의 디코딩이 가능함을 <그림 7>에서 검증하였다. 따라서 본 논문에서 제안한 G-PCC 콘텐츠를 공간 도메인에 따라 소비하기 위한 ISOBMFF 캡슐화 방안인 파일 포맷을 통해 G-PCC 데이터의 공간 정보에 따른 소비가 가능함을 확인하였다.

4. 결론

시간 도메인 기반으로 G-PCC의 콘텐츠는 사용자가 원하는 시점의 데이터를 시간적 임의 접근을 통해 활용할 수 있지만, 위치의 선택적인 콘텐츠 활용을 고려하지 못한다는 한계가 존재한다. 본 논문은 이러한 한계를 극복하고 G-PCC 콘텐츠를 공간 도메인에 따라 소비하여 포인트 클라우드의 원하는 위치를 선택하여 활용할 수 있음을 확인하였다. 또한 ISOBMFF 캡슐화 방안인 파일 포맷을 샘플 화하여 사용자가 선택적으로 소비할 수 있음을 검증하였다. 이러한 G-PCC 콘텐츠로 시간뿐만 아니라 공간 도메인으로 소비를 제공할 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

This work was supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No.2020-0-00452, Development of Adaptive Viewer-centric Point cloud AR/VR(AVPA) Streaming Platform)

5. 참고문헌

- [1] ISO/IEC JTC 1/SC 29/WG 07/N0004, "Text of ISO/IEC FDIS 23090-9 Geometry-based Point Cloud Compression," Oct. 2020.
- [2] "Video-based Point Cloud Compression," ISO/IEC 23090-5, Jan. 2021
- [3] P. Sun, X. Zhao, Z. Xu, R. Wang and H. Min, "A 3D LiDAR Data-Based Dedicated Road Boundary Detection Algorithm for Autonomous Vehicles," in IEEE Access, vol. 7, pp. 29623-29638, 2019.
- [4] ISO/IEC JTC 1/SC 29/WG 07/N00189, "Inter-prediction Exploration Model (Inter-EM) v3.0"
- [5] ISO/IEC 14496-12, "Coding of audio-visual objects - Part 12: ISO base media file format", February 2015