

V3C 기반 3DoF+ 비디오 부호화 성능 분석

이예진, 윤용욱, 김재곤

한국항공대학교

{yejin.lee, yuyoon}@kau.kr, jgkim@kau.ac.kr

Performance Analysis of 3DoF+ Video Coding Using V3C

Ye-Jin Lee, Yong-Uk Yoon, and Jae-Gon Kim

Korea Aerospace University

요 약

MPEG 비디오 그룹은 MPEG-I 표준의 일부로 포인트 클라우드(Point Cloud) 압축을 위한 비디오 기반 포인트 클라우드 부호화(V-PCC)와 몰입형(immersive) 비디오 압축을 위한 MPEG Immersive Video(MIV) 표준을 개발하고 있다. 최근에는 포인트 클라우드 및 몰입형 비디오와 같은 체적형(volumetric) 비디오를 모두 압축할 수 있도록 V-PCC 와 MIV 를 통합한 V3C(Visual Volumetric Video-based Coding) 표준화를 진행하고 있다. 본 논문에서는 V3C 코덱을 사용한 3DoF+(3 Degree of Freedom plus) 비디오 부호화 방안을 분석한다. 또한 V3C 코덱의 2D 코덱으로 기존 HEVC 대신 VVC 를 사용할 경우의 부호화 성능 향상을 분석한다.

1. 서론

MPEG 비디오 그룹에서는 몰입형 미디어를 위한 MPEG-I(Immersive) 패키지 표준을 개발하고 있다. MPEG-I Part 5 에서는 포인트 클라우드(point cloud)와 몰입형(immersive) 비디오를 모두 압축할 수 있는 V3C(Visual Volumetric Video-based Coding)와 비디오 기반의 포인트 클라우드 부호화(V-PCC) 표준을, Part 12 에서는 3DoF+, 6DoF 의 몰입형 비디오 부호화를 위한 MIV(MPEG Immersive Video) 표준을 개발하고 있다[1].

체적형(volumetric) 비디오 중 하나인 포인트 클라우드(point cloud)는 색상, 반사율 등의 여러 속성을 갖는 3D 공간의 포인트 집합으로 구성되어 있다. 또 다른 체적형 비디오인 몰입형(immersive) 비디오는 3D 공간에서의 다양한 뷰(view) 포인트에 해당하는 가상의 합성 뷰나 3D 공간에 배치된 실제 카메라로 획득한 다수의 뷰로 구성된다. 이러한 체적형 비디오는 방대한 양의 데이터를 가지므로 저장 및 전송을 위해서는 보다 효율적인 압축 기술이 필수적이다.

이들 체적형 비디오 부호화 표준 기술은 기본적으로 3D 비디오를 2D 로 투영한 후 2D 부호화 코덱을 사용하는 것으로 공통적으로 아틀라스(Atlas), Geometry(depth), Attribute (texture), Occupancy Map 부호화로 구성된다.

본 논문에서는 V3C 코덱의 개요와 함께 V3C 코덱을 사용한 3DoF+ 비디오 부호화 성능을 제시한다. 또한, V3C 의 2D 코덱으로 HEVC 대신 VVC 를 사용했을 때의 부호화 성능 개선을 분석한다.

2. V3C 기반 3DoF+ 비디오 부호화

V3C 코덱은 V-PCC 를 기반으로 확장된 것으로 체적형 비디오(포인트 클라우드, 3DoF+ 비디오, 매쉬(mesh) 등)를 2D 영상으로 투영한 후 2D 코덱을 사용하여 압축을 수행한다. V3C 코덱은 포인트 클라우드와 MIV 에 필요한 압축 정보들을 수용할 수 있도록 확장 비트스트림 신택스를 제공하고 있다. V3C 의

신텍스는 포인트 클라우드와 MIV 의 공통된 압축 정보를 포함하고 V-PCC 비트스트림에서 사용되는 신텍스 일부가 MIV 를 수용할 수 있도록 확장되어 구성된다[2], [3].

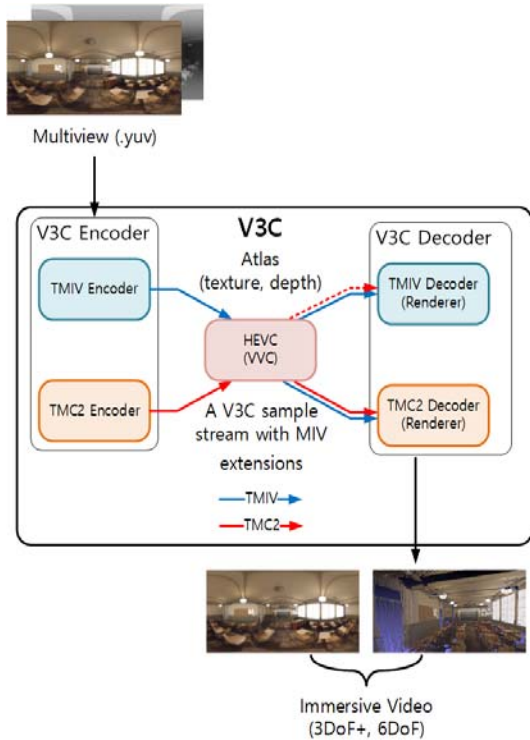


그림 1. V3C 코덱 구성도 및 3DoF+ 비디오 부호화

그림 1 은 V3C 코덱의 구성도이다. V3C 인코더는 V-PCC 의 시험모델인 TMC2(Test Model Category 2) 인코더와 3DoF+ 비디오 인코더인 TMIV(Test Model for Immersive Video) 인코더로 구성되어 있다. V3C 디코더는 마찬가지로 TMC2 디코더와 TMIV 디코더로 구성된다[3], [4]. 본 논문에서는 V3C 코덱의 TMC2 인코더와 디코더를 사용한 3DoF+ 비디오 부호화를 다룬다.

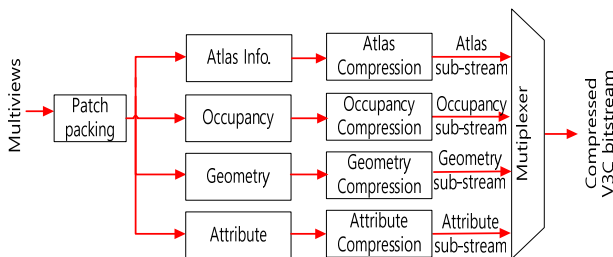


그림 2 V3C 인코더 구성도[1], [2]

그림 2 는 V3C 인코더의 구성도이다. 본 논문에서는 3DoF+ 비디오 압축을 위하여 V3C 인코더에 포함된 TMC2 를 사용한다.

V3C 의 인코더에서는 다수의 소스 뷰 비디오(텍스처 및 깊이)와 각 뷰의 카메라 파라미터를 입력으로 받아 전처리를 통하여 Geometry, Attribute 에 대한 패치(patch) 패킹 (packing)한 아틀라스와 Occupancy Map 을 생성하고 이들을 각각 압축한다.

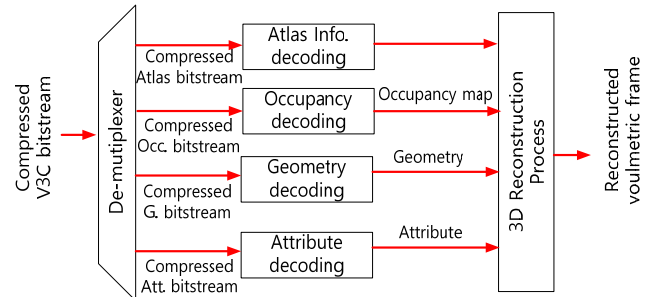


그림 3. V3C 디코더 구성도[1],[2]

그림 3 은 V3C 디코더 구성도이다. V3C 디코더에서는 V3C 비트스트림을 입력하여 복호화 과정을 거쳐 포인트 클라우드(ply 파일)를 복원한다. V3C 비트스트림에는 Occupancy, Geometry, Attribute 의 압축된 정보와 Atlas 정보가 포함된다. 여기서 2D 표현에서 3D 표현으로 역재구성(reconstruction)을 가능하게 하는 추가 아틀라스 정보도 비트 스트림에 포함되는데, 아틀라스의 각 패치는 2D 구성 요소에서 영역을 식별하고 이를 3D 공간으로 다시 적절하게 역투영하는데 필요한 정보를 포함하고 있다. 본 논문에서는 V3C 코덱을 사용하여 3DoF+ 비디오를 압축 전송하여 포인트 클라우드로 복원하고 그 부호화 성능을 확인한다.

3DoF+ MIV 테스트 시퀀스는 카메라 공간에서 가상 혹은 실제의 원근(perspective) 카메라 또는 ERP 카메라로 획득되며, 포인트 클라우드는 월드 좌표계에서 깊이를 가진 직교(orthographic) 카메라 형태로 복원된다[6]. 그림 4 는 원근 카메라를 사용하여 획득한 MIV 의 Classroom 시퀀스의 원본 뷰와 V3C 코덱으로 부호화하여 복원된 포인트 클라우드이다.

3. 실험결과

본 논문의 실험에서는 3DoF+ 비디오 부호화를 위하여 V3C 코덱을 사용하였으며, V3C 의 2D 코덱은 HEVC 의 시험모델 HM 16.20 를 사용하였다. 2D 코덱으로 VVC 를 사용한 경우의 성능비교에서는 VVC 시험모델 VTM 8.2 를 사용하였다. 실험조건은 RA(Random Access) 부호화 모드에서 MIV 의 CTC(Common Test Condition)과 V-PCC 의 CTC 를 따랐다[6], [7].



(a) 원본



(b) 복원된 포인트 클라우드

그림 4. V3C 코덱을 이용한 MIV 압축복원 예(Class A Classroom)

V3C 인코더로 입력된 3DoF+ 시퀀스를 무손실로 압축하고 RA 부호화 모드로 복호화하여 원본 포인트 클라우드에 가까운 무손실 압축 포인트 클라우드를 생성하였다. 4 개의 3DoF+ 시퀀스 대상으로 무손실 압축 포인트 클라우드와 QP 별 손실 압축한 포인트 클라우드의 평균 압축률을 비교하여 표 1 에 나타내었다.

3DoF+ 시퀀스로부터 V3C 를 통해 압축된 포인트 클라우드와 부호화 성능을 확인을 위해서는 3DoF+ 비디오의 Normal Vector 성분의 포인트 클라우드가 필수적이다. 현재 3DoF+ 비디오의 멀티 뷰 비디오로부터 생성된 포인트 클라우드에 대해서 Metrics Error 계산이 불가능하다[7]. 이를 대신하여 V3C 부호화 과정에서 입력된 3DoF+ 시퀀스로부터 생성된 Attribute 와 Geometry 아틀라스를 2D 코덱으로 VTM 을 사용하였을 때의 HM 대비 아틀라스 부호화 성능을 표 2 에 제시 하였다. 해당 Attribute 와 Geometry 아틀라스 비트스트림이 V3C 디코더 단에서 포인트 클라우드를 재구성하게 되므로 HM 대비 VTM 사용시의 포인트 클라우드의 부호화 개선 성능도 있을 것으로 예상된다.

4. 결론

MPEG 비디오 그룹에서는 MPEG-I 표준의 한 파트로 포인트 클라우드 및 몰입형 비디오 등의 체적형(volumetric) 비디오 부호화를 위한 V3C 표준화를 진행하고 있다. 본 논문에서는 V3C 코덱을 사용한 3DoF+ 비디오 부호화 방안과 부호화 성능을 분석하였다.

본 논문의 실험에서는 V3C 의 V-PCC 시험모델인 TMC2 를 이용하여 다수의 소스 뷰로 구성된 3DoF+ 비디오를 부호화하여 포인트 클라우드로 복원할 수 있음을 보였다. 또한, V3C 의 2D 코덱으로 VTM 을 사용하였을 때의 HM 대비 아틀라스의 부호화 성능 개선 정도를 확인하였다. 향후, 포인트 클라우드와 몰입형 비디오를 모두 수용하여 보다 효율적으로 부호화 할 수 있는 V3C 의 개선이 필요하다.

표 1. V3C 의 포인트 클라우드의 평균 압축률(RA, 17 frames)

Sequences		Compression Ratio
Computer Graphic	SA	257:1
	SB	40:1
Natural Content	SD	137:1
	SE	108:1
Total		135:1

표 2. HM 대비 VTM 의 아틀라스 부호화 성능(RA, 17 frames)

Sequences		High-BD rate Y-PSNR	Low-BD rate Y-PSNR
Computer Graphic	SA	-12.7%	-16.4%
	SB	-10.6%	-10.5%
Natural Content	SD	-11.6%	-12.1%
	SE	-11.0%	-13.9%
Total		-11.47%	-13.23%

Acknowledgement

이 논문은 과학기술정보통신부의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. 2017-0-00486).

- [1] "WD of V3C and V-PCC (2nd Edition)", ISO/IEC JTC1/SC29/WG11 Output Document N19523, Online Meeting, July, 2020.
- [2] "V-PCC Future Enhancements (V3C + V-PCC)," ISO/IEC JTC1/SC29/WG11 Output Document N19329, Online Meeting, Apr. 2020.
- [3] B. Kroon, and D. Graziosi, "Report on cross-test model testing of 3VC," ISO/IEC JTC1/SC29/WG11 Input Document m53266, Online Meeting, Apr. 2020.
- [4] D. Graziosi et al., "[VPCC][specification]On the integration between MIV and V-PCC," ISO/IEC JTC1/SC29 WG11 Input Document m51044, Geneva, CH, Oct. 2019.
- [5] "V-PCC Codec Description," ISO/IEC JTC1/SC29/WG11 Output Document N19332, Online Meeting, Apr. 2020.
- [6] "Common Test Conditions for Immersive Video," ISO/IEC JTC1/SC29/WG11 Output Document N19214, Online Meeting, Apr. 2020.
- [7] "Common Test Conditions for PCC," ISO/IEC JTC1/SC29/WG11 Output Document N19324, Online Meeting, Apr. 2020.