가상 현실 비디오 프로젝션 처리를 위한 SEI 구문 분석기와 디코더 구현

정종범, 손장우, 장동민, 류은석
가천대학교
uof4949@gc.gachon.ac.kr, sjw6757@gc.gachon.ac.kr, dogzz9445@gc.gachon.ac.kr,
esryu@gachon.ac.kr

Implementation of SEI Parser and Decoder for Virtual Reality Video Projection Processing

JongBeom Jeong, Jang-Woo Son, Dongmin Jang, Eun-Seok Ryu
Gachon University

요 약

최근 360 도 가상현실을 지원하기 위한 비디오 시스템은 다양한 프로젝션에 대한 처리를 필요로 한다. 이를 위해 Moving Picture Experts Group (MPEG) 비디오 표준화 기술은 비디오에 대한 추가적인 정보들로 프로젝션을 처리하는 기술을 표준 채택하였다. 즉, 다양한 프로젝션의 비디오에 대응하는 비디오 메타데이터 처리를 H.265/HEVC(High Efficiency Video Coding)에서 제안된 Supplemental Enhancement Information (SEI) 메세지를 사용하여 지원한다. 본 논문은 비디오의 인코딩, 디코딩 시에 비디오 프로젝션 타입에 따라 다르게 처리하는 시스템의 구현 기술을 소개한다. 이를 위해 본 논문은 SEI 메시지 구문 분석기를 구현 시 HEVC Test Model(HM)을 이용하고, 디코더 구현 시 FFmpeg 라이브러리를 이용한다. 최종적으로 구현된 시스템은, 본 기관의 또 다른 구현 물인 실시간 360 비디오 플레이어에 통합되어 실시간 디코딩 및 다양한 프로젝션의 전/후처리를 문제 없이 지원하였다.

1. 서론

통신 기술의 급격한 발전에 힘입어, 음성 신호와 영상 신호를 포함하는 멀티미디어 데이터의 이용이 급속도로 증가하고 있다. 그러므로 동영상에 대한 수요도 지속적으로 증가하고 있으며 따라서 효과적인 동영상의 화질 향상 기법이 지속적으로 요구되고 있다. 그에 대한 예로 최근 들어 VR(Virtual Reality)를 지원하는 HMD(Head Mounted Display) 기기들의 보급이 활발하게 진행되고 있다. 삼성전자에서는 모바일 플랫폼을 기반으로 하는 삼성 갤럭시 기어 VR 을 출시하였고, Oculus 와 HTC 에서는 각각 pc platform 을 기반으로 하는 Oculus rift 와 HTC vive 를 출시하였다. 이에 MPEG 은 immersive media 의 지원을 위해 MPEG-I(Immersive) 표준화를 진행하고 있고, 2021 년까지 완료할 예정이다.

품질의 VR 영상을 HMD 기기를 통해 재생하기 위해서는 4K 정도가 되어야 하는데, 이 경우 HMD 기기가 처리해야 할 데이터의 양이 급격히 증가하게 된다. 최근에 상용화된 HMD 중 모바일 플랫폼을 기반으로 하는 기기의 비중이 높다는 것을 생각해 보면 이는 현실감 있는 VR 영상 시청에 있어 장애물이 된다. 때문에 2016 년 MPEG 에서 사용자가 VR 영상을 시청할 때 보는 영역인 뷰포트 영역을 처리하는 기술인 MCTS (Motion-Constrained Tile Set)가 논의되었고[1], 이를 구현한 기고문이 MPEG 에 제출되었다[2]. 사용자의 뷰포트를 추출하여 VR 기기로 스트리밍하는 기술이 제안되었다[3]. 또한 영상의 크기가 크므로 사용자의 관심 영역(ROI, Region of Interest)을 고화질로 인코딩하고, 나머지는 저화질로 인코딩하는 region-wise packing 이 제안되었다[4].

360 도 영상을 표현하려면 다양한 프로젝션 타입을 고려해야 하는데, 이 프로젝션에 대한 정보를 SEI message[5]에 담을 수 있다. SEI 메시지에 프로젝션에 대한 정보를 포함하여 인코딩하면 디코딩할 때 이를 사용하여 각 프로젝션에 대응하여 디코딩할 수 있다.

본 논문에서는 인코딩할 때 영상의 프로젝션에 대응하여 디코딩할 때 정보를 제공하는 SEI parser 를 구현하는 방법을 제안한다. HM 의 NAL parser 를 이용하여 SEI parser 를 구현하고, FFmpeg 라이브러리[6]를 이용하여 디코딩하며 이를 파이프라이닝하였다.

본 논문의 구성은 다음과 같다. 2 절에서는 관련 연구에 대해 살펴보고, 3 절에서는 본 논문에서 제안하는 설계 방법에 대해 설명한다. 마지막으로 4 절에서는 본 논문에 대한 결론을 서술한다.

2. 관련 연구

2.1. 프로젝션

VR 을 지원하는 HMD 에서 360 도 영상을 보려면 기존의 2 차원 영상을 입체적으로 매핑하는 과정이 필요한데, 이 매핑 방법을 프로젝션이라고 한다. 프로젝션 방식의 예로 ERP (Equirectangular Projection), CMP(Cubemap Projection),

OHP(Octahedron projection), SSP(Segmented sphere projection), EAC(Equi-Angular cubemap projection) 등이 있다.

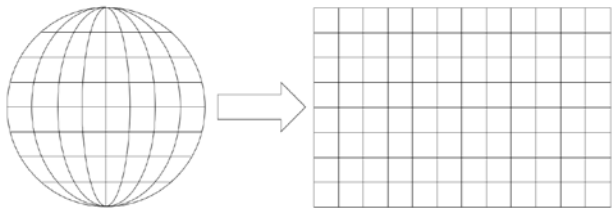

그림 1. 360 미디어의 Equirectangular Projection(ERP)
Fig. 1. Equirectangular Projection of 360 media

MPEG 에서 다루는 360 미디어를 지원하기 위한 OMAF(Omnidirectional Media Format)에서는 360 비디오를 처리하기 위한 핵심 표준 기술들에 대해 정의하고 있다[7]. 프로젝션 기법 중 ERP 방식은 360 비디오를 2 차원 영상에서 나타내는 데 가장 많이 쓰이는 방식이고, 그림 1 에서 알 수 있듯 구 형태의 영상을 2 차원 영상으로 매핑한다. 때문에 구의 위아래 끝점에 해당하는 부분이 2 차원으로 매핑될 때 중복되어 표현되는 문제가 있다.

CMP 방식은 최근 많이 쓰이기 시작한 방식으로, 그림 2 에서 보는 바와 같이, 구 형태로 표현되는 영상을 각각 육면체의 면들에 매핑시켜 표현한다.

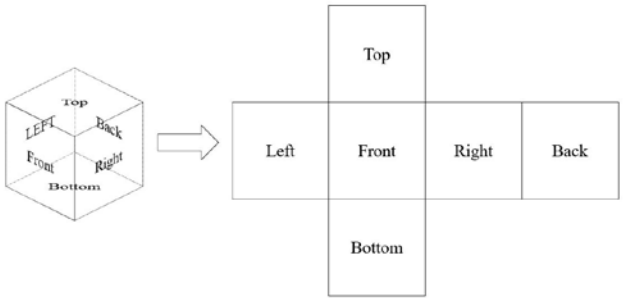

그림 2. 360 미디어의 Cubemap Projection(CMP)
Fig. 2. Cubemap Projection of 360 media

2.2. 영역 기반 패킹 (Region-wise packing)

360 미디어는 프로젝션 단계를 통해 2 차원의 이미지로 투영되어지는데, 영역 기반 패킹을 이용하면 이 이미지의 특정 영역의 위치, 크기를 변경할 수 있다. 그림 3 을 보면 알 수 있듯 영역 기반 패킹은 사용자의 관심 영역이 될 확률이 높은 중간 영역은 고화질로 인코딩하고, 상대적으로 관심 영역이 될 확률이 낮은 위, 아래 영역은 크기를 줄이고 저화질로 인코딩한다. 이러한 전략을 통해 360 미디어의 압축률을 높이거나 사용자 뷰포트에 종속적인 비트스트림 생성을 구현[8]할 수 있다.

영역 기반 패킹이 적용되어 인코딩된 영상의 경우 디코딩할 때 해당 영상에 대한 영역 기반 패킹 메타데이터가 전달될 수 있어야 한다.

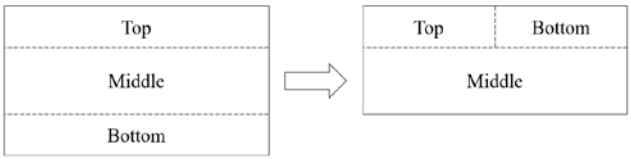

그림 3. 영역 기반 패킹의 실시 예
Fig. 3. A working example of region-wise packing

2.3. SEI 메시지

인코딩된 HEVC 비트스트림은 NAL(Network Abstraction Layer)로 불리는 패킷의 형태로 저장되거나 전송된다. 그림 4 에서 알 수 있듯 NAL 은 VPS(Video Parameter Set), SPS(Sequence Parameter Set), PPS(Picture Parameter Set), PREFIX_SEI, Slice 를 포함한다.

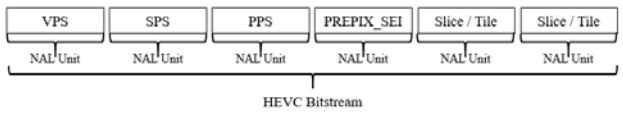

그림 4. HEVC NAL 의 구조와 구성요소
Fig. 4. Structure and syntax elements of HEVC NAL

NAL 유닛 중 PREFIX_SEI 유닛은 비트스트림에 대한 SEI 메시지를 포함한다. 그림 5 는 H.264 프레임의 SEI 메시지에 대한 구조[9]를 보여주고 있다. SEI 메시지를 구성하는 요소 중 payloadType 은 NAL 유닛의 종류를 구분할 때 사용된다[10].

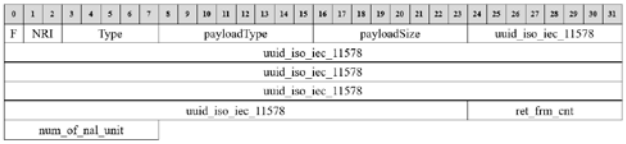

그림 5. H.264 프레임에 대한 SEI message 의 구조
Fig. 5. SEI message structure of H.264 frame

3. 설계

3.1. SEI 메시지 구문 분석기

다양한 360 미디어에 대응하여 인코딩하기 위해 본 논문에서는 HM(HEVC test model)[11]을 사용하였다. 그림 6 은 HM 의 SEI 클래스에서 PayloadType 에 추가한 요소들을 나타낸다. ERP, CMP, sphere_projection, 영역 기반 패킹, omni_viewport 를 처리하기 위해 추가하였고, 각각 150, 151, 152, 153, 154 의 값으로 나타내어진다.

PayloadType	Value
EQUIRECTANGULAR_PROJECTION	150
CUBEMAP_PROJECTION	151
SPHERE_PROJECTION	152
REGIONWISE_PACKING	153
OMNI_VIEWPORT	154

그림 6. SEI 클래스에서의 PayloadType 정의

Fig. 6. Definition of PayloadType in SEI class

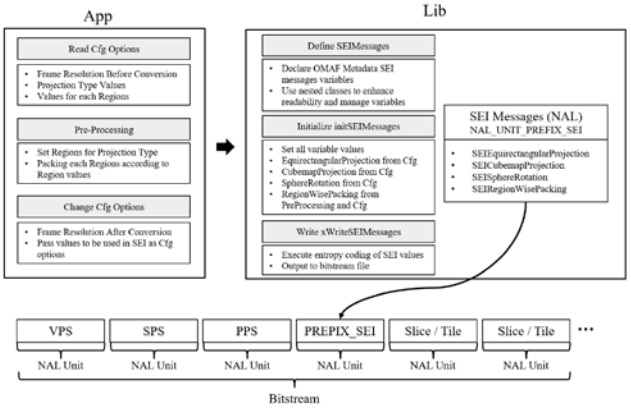

그림 7. HM 디코더에서의 SEI message 처리 과정
Fig. 7. SEI message processing of HM decoder

그림 7 은 HM 디코더에서 SEI 메시지를 처리하는 과정을 나타낸다. NAL 유닛을 처리할 때 PREFIX_SEI 를 찾고, 따로 저장한다. VPS, SPS, PPS, Slide 를 처리할 때가 되면 따로 저장해 둔 PREFIX_SEI 를 이용하여 영상이 ERP, CMP, 영역 기반 패킹일 때를 구분하고 각자 다르게 처리한다.

3.2. FFmpeg 디코더

HM 은 HEVC 표준을 따르는 코드들이 포함된 레퍼런스 소프트웨어인 관계로 실시간 처리에 한계가 있다. 때문에 본 논문에서는 실시간 디코딩을 위해 FFmpeg 라이브러리를 이용하여 디코더를 구현하였다. 그림 8 은 HM 의 NAL parser 와 FFmpeg 디코더를 활용하여 프레임 단위로 처리하는 디코더의 구조를 명시하였다. HM 의 TAppCommon, TLibCommon, TLibDecoder 프로젝트를 참조하고, FFmpeg 라이브러리를 포함시켰다. 처리 과정은 플레이어를 초기화하고, SEI 메시지를 파싱한 후 FFmpeg 을 이용하여 YUV 데이터를 얻어오고, 파싱한 SEI 메시지를 이용하여 후처리를 진행한 뒤 플레이하는 과정으로 이루어진다.

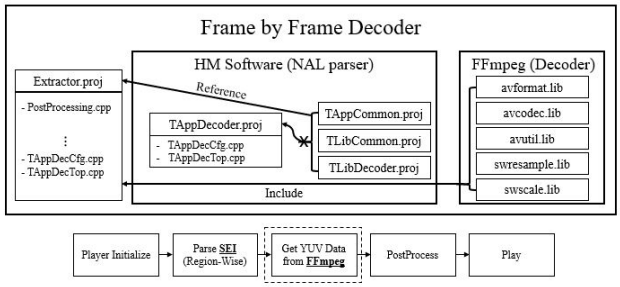

그림 8. HM 디코더에서의 SEI message 처리 과정
Fig. 8. SEI message processing of HM decoder

4. 실행 결과

그림 9 는 영역 기반 패킹이 적용된 영상의 프레임을 보여주고 있다. 왼쪽 그림은 영역 기반 패킹 전처리만 적용된 비트스트림을 HM 의 기본 디코더로 디코딩한 프레임을 나타낸다. 오른쪽 그림은 영역 기반 패킹 전처리가 적용된 비트스트림을 본 논문에서 제안하는 디코더로 영역 기반 패킹 후처리를 적용하여 디코딩한 프레임을 나타낸다.

그림 9. 영역 기반 패킹 전처리, 후처리 된 프레임
Fig. 9. Region-wise packing preprocessing, postprocessing applied frame

5. 결론

본 연구는 가상 현실 미디어에 대한 다양한 프로젝션에 대응하는 SEI 구문 분석기와 실시간 디코더 구현 기술을 설명한다. HM 을 이용하여 HEVC 표준을 따르면서도 실시간 처리를 위해 FFmpeg 라이브러리를 사용하였고, 영역 기반 패킹에 대한 메타데이터를 SEI 메시지를 통해 전달하였다. 본 논문에서는 HM 의 SEI 클래스에서 PayloadType 요소들을 추가하여 사용하였으나, 표준화 미팅에서 SEI 메시지에 대한 기고안이 여전히 제출되고 있으므로[12], 향후 이를 지원하는 연구 개발을 계속 진행할 계획이다.

참고문헌

[1] Ye-Kui Wang, Hendry, Marta Karczewicz. (2016). Viewport dependent processing in VR: partial video decoding. 116th MPEG meeting of ISO/IEC JTC1/SC29/WG11, MPEG 116/ m38559.

[2] Robert Skupin, Yago Sanchez, Karsten Sühring, Thomas Schierl, Eun-Seok Ryu, Jangwoo Son. (2017). 120th MPEG meeting of ISO/IEC JTC1/SC29/WG11, MPEG 120/ m41626.

[3] Jang-Woo Son, Dongmin Jang, Eun-Seok Ryu. (2018). Implementing Motion-Constrained Tile and Viewport Extraction for VR Streaming. ACM Network and Operating System Support for Digital Audio and Video 2018 (NOSSDAV2018).

[4] Sejin Oh, Soojin Hwang. 2017. OMAF: Generalized signaling of region-wise packing for omnidirectional video. 118th MPEG meeting of ISO/IEC JTC1/SC29/ WG11, MPEG2017/ m40423.

[5] Gary J. Sullivan, Haoping Yu, Shun-ichi Sekiguchi, Huifang Sun, Thomas Wedi, Steffen Wittmann, Yung-Lyul Lee, Andrew Segall, and Teruhiko Suzuki. (2007). New Standardized Extensions of MPEG4-AVC/H.264 for Professional-Quality Video Applications. ICIP 2007, I-13 - 16.

[6] FFmpeg Team. (2013). FFmpeg. URL http://ffmpeg.org/

[7] 오세진. (2017). 360 미디어를 위한 MPEG Omnidirectional Media Format (OMAF) 표준 기술. 방송공학회논문지 제 22 권 제 5 호, 2017.9, 600-607.

[8] Jang-Woo Son, Dongmin Jang, Jong Beom Chung, Eun-Seok Ryu. (2018). Implementing 360 Video Tiled Streaming System. ACM Multimedia Systems 2018.

[9] ISO/IEC. (2010). 14496-10:2010 Information technology—Coding of audiovisual objects—part 10: Advanced Video Coding.

[10] Jill Boyce, Hyun-Mook Oh, Gary J. Sullivan, Alexis Tourapis, Ye-Kui Wang. (2018). Working draft of Additional Supplemental Enhancement Information for HEVC (Draft 1). Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 N17417.

[11] Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute. (2017) High Efficiency Video Coding (HEVC) reference software HM. Accessed 13 March, https://hevc.hhi.fraunhofer.de/.

[12] Chad Fogg, Walt Husak, Gary J. Sullivan, Alexis Tourapis, Ye-Kui Wang. (2018). Working draft of Additional Supplemental Enhancement Information for AVC (Draft 1). 121th MPEG meeting of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, MPEG121/n17425.