

사용자 시점 기반 360 도 영상 스트리밍을 위한 다중 타일 추출기 구현

정종범, 이순빈, 김인애, 류은석

성균관대학교 컴퓨터교육과

uof4949@skku.edu, soonbinlee@skku.edu, inaelk@skku.edu, esryu@skku.edu

Implementing Multiple-tile Extractor for Viewport-dependent 360 Video Streaming

Jong-Beom Jeong, Soonbin Lee, Inae Kim, Eun-Seok Ryu

Department of Computer Education, Sungkyunkwan University

요 약

몰입감 있는 가상 현실 영상을 제공하기 위한 360 도 영상 부호화 및 전송 기술이 활발히 연구되고 있으나, 현재 가상현실 장비가 사용가능한 연산 능력 및 대역폭으로는 몰입감 있는 영상을 전송 및 재생하기에 한계가 있다. 따라서 본 논문은 고화질 360 도 사용자 시점 영상 제공을 위해 사용자 시점 타일을 추출하는 움직임 제한 타일 셋 기반 타일 추출기를 구현한다. 기존의 high-efficiency video coding (HEVC) 에서 구현되었던 타일 추출기와 달리 제안하는 추출기는 360 도 영상에 대한 비트스트림에서 여러 개의 타일을 추출한다. 이후 추출된 타일들은 전체 360 도 영상에 대한 저화질 비트스트림과 동시 전송되어 예상치 못한 사용자 시점 변경에 대응한다.

1. 서론

가상 현실 기술의 발전에 따라, 관련 시장이 급격히 증가하고 있다. 머리장착형 영상장치인 head-mounted display (HMD)의 보급이 증가하고 있고, 고품질의 가상 현실 영상을 제공하기 위한 효율적인 360 도 영상 스트리밍 기술이 요구되고 있다. 하지만, 현재 사용가능한 가상 현실 장비의 연산능력과 대역폭은 한계가 존재한다. 가상 현실에서 사용자가 어지러움을 느끼지 않을 정도의 영상을 제공하려면 12K 이상, 90 fps 이상의 초고화질 360 도 영상과 낮은 전송 지연이 필요하다[1]. 상기 요구 사항을 극복하기 위해 다양한 360 도 영상 스트리밍 방법이 제안되었다. 다시점 360 도 영상을 중요도에 따라 다운샘플링하는 연구[2], 360 도 영상을 타일로 분할하여 인코딩한 후 압축된 비트스트림 레벨에서 사용자 시점 타일만 추출 및 전송하는 연구[3-5], 복수 영상 전송 시 360 도 영상 간 중복성을 제거하는 연구[6] 등이 제안되었다.

HMD 를 통한 360 도 영상 렌더링 시 영상의 일부만 화면에 표현되기 때문에, 대역폭을 절약하고 병렬 복호화를 가능하게 하기 위해 high-efficiency video coding (HEVC) 에서는 타일 추출 및 스트리밍 기법이 제안되었다. HEVC test model (HM) 에 타일 추출기가 포함되었으나, 상기 추출기는 개별 타일 당 하나의 비트스트림을 생성하여 사용자 시점 타일 개수가 많아지면 오버헤드 (overhead)가 증가하는 단점이 있다.

본 논문은 HEVC 에서 동작하는 사용자 시점 기반 타일 스트리밍 시스템을 위한 다중 타일 추출기를 제안한다. 그림 1 은 제안하는 다중 타일 추출기 기반 스트리밍 시스템의 동작 흐름을 나타낸다. 제안하는 추출기는 개별 타일의 추출 및 복호화가 가능하도록 하는 motion-constrained tile set (MCTS) 을 사용해 부호화된 비트스트림에서 적어도 하나 이상의 사용자 시점 타일들을 추출하고, 단일 비트스트림을 생성 후 전송한다.

본 논문의 구성은 다음과 같다. 2 절에서는 배경 연구로 HEVC 에서의 타일을 소개한다. 3 절에서는 제안하는 다중 타일 추출기 설계 및 구현 내용을 설명한다. 4 절에서는 제안하는 타일 추출기의 구현 결과를 소개하고 마지막으로 5 절에서는 본 논문의 결론을 맺는다.

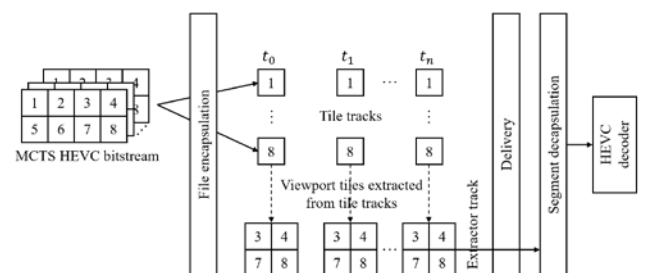


그림 1. 제안하는 다중 타일 추출기 구조도

2. 배경

Moving picture experts group (MPEG) 과 ITU-T video coding experts group (VCEG) 은 더 높은 영상 부호화 성능을 제공하기 위해 joint collaborative team on video coding (JCT-VC)을 설립 후, HEVC 표준화를 진행하였다[7]. 부호화해야 할 영상의 크기가 커짐에 따라 비트스트림의 복호화를 병렬적으로 진행해야 할 필요성이 제기되었고, 이에 타일이란 개념이 등장하였다. Advanced video coding (AVC) 에서의 슬라이스와 유사하게, 타일 또한 여러 개의 coding tree unit (CTU) 로 구된다. 또한, 서로 다른 슬라이스 및 타일 간에는 의존성이 존재하지 않아 독립적인 복호화가 가능하다. 하지만, 타일은 직사각형 형태로 구성된다는 차이가 있다. HEVC 에서 픽처는 적어도 하나 이상의 타일로 분할될 수 있으며, 움직임 벡터 (motion vector) 를 타일 내부로 제한하는 MCTS 를 통해 각 타일은 공간상에서 독립적이다[8]. 각 타일은 독립적으로 복호화될 수 있음과 동시에 원본 비트스트림에서 추출되어 독립된 비트스트림으로 출력될 수 있다. 단, 타일 추출 시 하나의 슬라이스가 하나의 타일을 포함하여야 한다.

타일은 병렬 복호화뿐만 아니라 선택적인 스트리밍 기법에도 사용될 수 있다. 예를 들어 HMD 를 통한 가상 현실 영상 재생 시 영상 자체가 360 도 전방위 영상이라 할지라도 화면에는 사용자의 field of view (FoV) 만큼만 영상이 재생된다. 따라서, 사용자 시점을 알고 있다면 영상의 일부만 전송하는 스트리밍 기법을 사용하여 대역폭을 절약할 수 있다. 사용자 시점 기반 타일 스트리밍의 경우 사용자 시점 예측 및 추출 지연 시간을 줄이는 것이 주요 과제인데, 영상에서 관심 영역을 특정하고 사용자 시점 예측에 사용하는 saliency 맵 (saliency map) 이 사용될 수 있다[9].

3. 360 도 영상 스트리밍을 위한 다중 타일 추출 기

본 절은 360 도 영상 스트리밍을 위한 다중 타일 추출기의 설계 및 동작 과정에 대해 기술한다. HEVC 비트스트림은 여러 개의 network abstraction layer (NAL) 유닛 (unit) 들로 구성되고, NAL 유닛은 video coding layer (VCL) NAL 유닛과 non-VCL NAL 유닛으로 나누어진다. VCL NAL 유닛은 실제 영상 데이터를 담고 있고, 슬라이스가 여기에 속한다. 2 절에 소개된 것처럼 슬라이스와 타일은 1 대 1 대응하므로, 만약 픽처가 8 개의 타일로 이루어졌다면 비트스트림의 각 픽처는 8 개의 슬라이스로 구성된다. Non-VCL NAL 유닛은 영상에 대한 메타데이터 등 영상 데이터 이외의 정보를 포함한다. HEVC 비트스트림에서의 non-VCL NAL 유닛으로는 video parameter set (VPS), sequence parameter set (SPS), 그리고 picture para-

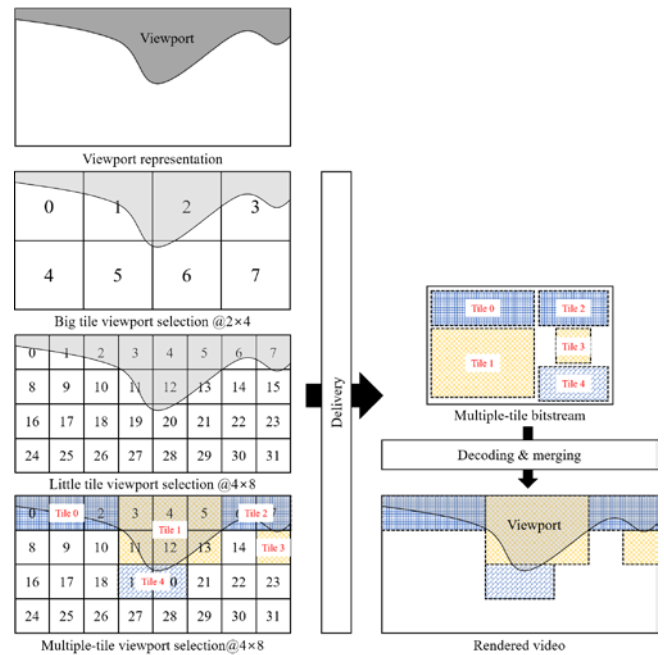


그림 2. 사용자 시점 기반 다중 타일 스트리밍 예시

-meter set (PPS) 등이 있다. 상기 파라미터 셋 (parameter set) 들은 비트스트림을 복호화할 때 필요한 정보들을 포함하나, 추가적인 정보가 필요한 경우 HEVC 부호기는 이 정보들을 supplemental enhancement information (SEI) message 의 형태로 비트스트림에 저장한다. MCTS 가 적용되어 부호화된 비트스트림의 경우 타일에 대한 파라미터 셋을 포함하는 extraction information sets (EIS) SEI message 를 포함한다.

제안하는 다중 타일 추출기는 입력된 비트스트림을 NAL 유닛 단위로 분석하고, EIS SEI message 에서 VPS, SPS, PPS 를 가져온다. 이후 다중 타일 추출기는 추출할 타일 정보에 맞게 파라미터 셋들을 수정한다. SPS 에서는 출력 비트스트림의 픽처 너비와 높이가 수정된다. 만약 입력 비트스트림이 대칭적으로 분할되었다면, PPS 에서는 타일의 행 및 열 개수를 수정하고 루프 필터 (loop filter) 옵션을 비활성화한다. 만약 비대칭적으로 분할되었다면, 각 타일의 너비 및 높이 또한 수정된다. 추출할 타일을 포함한 슬라이스의 헤더에서는 출력 비트스트림의 픽처 너비 및 높이에 따라 슬라이스 세그먼트 주소가 수정된다. 이후 제안하는 타일 추출기는 입력 VCL NAL 유닛을 출력 VCL NAL 유닛으로 변환하고, Non-VCL NAL 유닛과 함께 출력 비트스트림에 저장한다.

그림 2 는 제안하는 다중 타일 추출기를 이용한 사용자 시점 기반 타일 스트리밍 예시를 나타낸다. 그림에서 알 수 있듯 픽처를 더 많은 개수의 타일로 분할할수록 추출 및 전송할 영역이 줄어들며, HM 에 포함된 기존의 타일 추출기는 타일 하나 당 하나의 비트스트림을 생성하여 타일 개수가 많으면 복호화 시 오버헤드가 늘어나는 단점이 있다. 그러나 제안하는 다중 타일

추출기는 단일 비트스트림의 여러 개의 타일을 포함할 수 있어 복호기 개수를 추가로 요구하지 않으면서 대역폭을 절약할 수 있어 상기 오버헤드를 극복할 수 있다. 또한, 초고해상도 영상에 대한 타일 비트스트림의 경우 적은 개수의 타일이 포함되어 있더라도 픽처의 크기가 변하지 않는다면 모바일 디바이스와 같은 제한된 자원을 제공하는 환경에서 복호화하는 데에 어려움이 있다. 그러나 그림 2 에 도시되어 있듯 제안하는 다중 타일 추출기는 출력 비트스트림의 픽처 크기를 조절할 수 있어 클라이언트 사양 또는 대역폭에 따른 유연한 타일 추출이 가능하다.

4. 구현 결과

본 논문은 제안하는 다중 타일 추출기를 검증하기 위해 15360×7680 크기를 가지는 2 개의 테스트 시퀀스, *Drone*, *LakePark* 를 6×6 배열로 분할하여 부호화하였다. 양자화 파라미터 (quantization parameter, QP) 로는 22 를 사용하였다. 사용자 시점 정보는 JVET 에서 권고하는 동적 사용자 시점을 사용하였다[10]. 그림 3 은 [10]에 기초하여 사용자 시점 타일만을 추출하고 복호화한 결과를 보여준다. 그림 4 는 타일 2 개만을 추출하여 픽처의 해상도를 변경하고 하나의 비트스트림을 생성 후 복호화한 결과를 나타낸다. 전체 타일 중 일부만 추출 및 병합해도 문제없이 복호화가 되는 것을 확인하였고, 타일 경계에서 타일 간 간섭 없이 복호화가 이루어짐을 확인하였다.

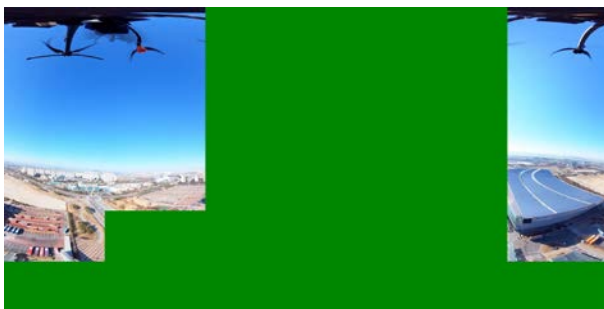


그림 3. 추출된 사용자 시점 다중 타일 영상 (Drone)

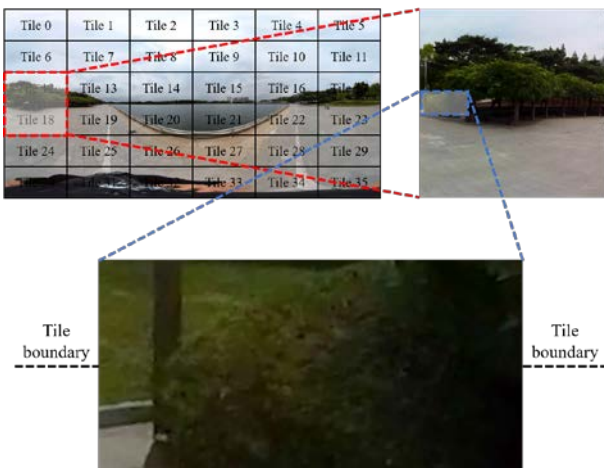


그림 4. 추출된 다중 타일 및 경계 (LakePark)

5. 결론

본 논문은 HEVC 에서의 사용자 시점 기반 360 도 영상 스트리밍을 위한 다중 타일 추출기를 제안한다. 제안하는 타일 추출기는 여러 개의 목표 타일을 추출하고 non-VCL NAL 유닛 및 슬라이스의 일부 정보를 수정하여 단일 비트스트림을 생성하고, 필요할 경우 픽처의 크기를 변경할 수 있다. 제안하는 추출기에 의해 추출된 비트스트림은 HEVC 복호기에 의해 문제 없이 복호화됨을 확인하였다. 향후 연구에서는 유연한 스트리밍을 위해 추출된 타일을 병합하는 기법을 연구할 계획이다.

Acknowledgement

이 논문은 2020 년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.2018-0-00765, 6DoF 지원 초고화질 몰입형 비디오의 압축 및 전송 핵심 기술 개발). 이 논문은 또한 2020 년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1A2C1010476).

참고문헌

- [1] Mary-Luc Champel, Thomas Stockhammer, Thierry Fautier, Emmanuel Thomas, Rob Koenen. 2016. Quality Requirements for VR. 116th MPEG meeting of ISO/IEC JTC1/SC29/ WG11, MPEG 116/m39532.
- [2] Jong-Beom Jeong, Dongmin Jang, Jangwoo Son, Eun-Seok Ryu. 2018. 3DoF+ 360 Video Location-Based Asymmetric Down-Sampling for View Synthesis to Immersive VR Video Streaming. *Sensors* 18 ,9(2018), 3148.
- [3] Jang-Woo Son, Dongmin Jang, Eun-Seok Ryu. 2018. Implementing Motion-Constrained Tile and Viewport Extraction for VR Streaming. In *Proceedings of the 28th ACM Network and Operating System Support for Digital Audio and Video*, pp. 61-66.
- [4] Jangwoo Son, Eun-Seok Ryu. 2018. Tile-Based 360-Degree Video Streaming for Mobile Virtual Reality in Cyber Physical System. *Computers & Electrical Engineering*, 72, 361-368.
- [5] Jeong-Beom Jeong, Soonbin Lee, Il-Woong Ryu, Tuan Thanh Le, Eun-Seok Ryu. 2020. Towards Viewport-dependent 6DoF 360 Video Tiled Streaming for Virtual Reality Systems. In *Proceedings of the 28th ACM*

International Conference on Multimedia, pp. 3687-3695.

- [6] Jong-Beom Jeong, Soonbin Lee, Dongmin Jang, Eun-Seok Ryu. 2019. Towards 3DoF+ 360 Video Streaming System for Immersive Media. *IEEE Access*, 7, pp. 136399-136408.
- [7] Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, Thomas Wiegand. 2012. Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Transactions on circuits and systems for video technology*, 22(12), pp. 1649-1668.
- [8] Alireza Zare, Alireza Aminlou, Miska Hannuksela, Moncef Gabbouj. 2016. HEVC-compliant tile-based streaming of panoramic video for virtual reality applications. In *Proceedings of the 24th ACM international conference on Multimedia*, pp. 601-605.
- [9] Soonbin Lee, Dongmin Jang, Jong-Beom Jeong, Eun-Seok Ryu. 2019. Motion-constrained tile set based 360-degree video streaming using saliency map prediction. In *Proceedings of the 29th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*, pp. 20-24.
- [10] Tomohiro Ikai, Yukinobu Yasugi, Tomoko Aono. 2017. Ahg8: Dynamic viewport generation for 360° video evaluation. 127th MPEG meeting of ISO/IEC JTC1/SC29/WG11, MPEG2017/m39669.