

# 사용자 시점 기반 360 도 영상 타일 스트리밍을 위한 복호기 적응적인 단일 계층 타일 바인딩

장미, 정종범, 이순빈, 류은석  
성균관대학교 컴퓨터교육학과

smddl62@skku.edu, uof4949@skku.edu, soonbinlee@skku.edu, esryu@skku.edu

## Decoder-adaptive Single-layer Tile Binding for Viewport-dependent 360-degree Video Tiled Streaming

Mi Jang, Jong-Beom Jeong, Soonbin Lee, Eun-Seok Ryu

Department of Computer Science Education, Sungkyunkwan University

### 요 약

실감적인 가상 현실을 위해서는 고화질의 360 도 영상 스트리밍이 필요하다. 그러나 이는 높은 대역폭과 연산량을 요구하기 때문에 일반적인 가상 현실 기기로는 감당하기 힘들다. 이를 보완하기 위한 360 도 영상 부호화 및 전송 기술이 활발히 연구되고 있으며, 대표적으로 사용자 시점 기반 타일 스트리밍 기법 등이 있다. 본 논문은 기존의 CTU 기반 스트리밍과 타일 기반 스트리밍과 함께 복호기 적응적인 단일 계층 타일 바인딩을 활용한 타일 기반 스트리밍의 부호화 및 복호화 성능을 비교한다. 수행된 실험결과, 단일 계층 타일 바인딩을 활용한 타일 스트리밍 방법이 기존의 타일 스트리밍 기법에 비해 유사한 비트율 성능에 대비하여 복호화 시간에서 큰 이득을 볼 수 있음을 확인하였다.

### 1. 서론 및 배경 연구

최근 가상 현실 관련 기술의 발전에 따라, 다양한 종류의 head-mounted display(HMD) 디바이스, 360 도 카메라 등이 상용화되었다. 그러나 이러한 VR 기기로 360 도 영상을 볼 때 사용자가 멀미를 느끼지 않으려면 최소한 12K 의 고화질 영상이 필요하다. 고화질의 360 도 영상을 스트리밍하기 위해서는 고성능의 장비가 요구되고, 현재 널리 상용화된 디바이스 중 이를 만족하는 장비는 극히 일부이다 [1]. 이를 보완하기 위해, 사용자 시점에 위치한 타일들만 고화질로 전달하고 나머지 부분들은 저화질 영상으로 전달하여 이를 합쳐서 사용자에게 보여주는 타일 기반 스트리밍 등 효율적인 360 도 영상 스트리밍을 위한 연구가 진행되고 있다.

타일 (tile) 은 high-efficiency video coding (HEVC) 에서 제안된 개념으로, 슬라이스 (slice) 와 달리 motion-constrained tile set (MCTS) 기술로 인해 타일 간 시공간적 의존성이 제한되고, 각 타일은 독립적으로 추출 및 복호화가 가능하다. MCTS 기술은 전체 영상의 일부만을 사용하는 시스템의 경우 필요한 영역이 포함되는 타일만을 추출 및 전송하여 대역폭을 절약할 수 있게 한다[2-6]. 기존의 타일 스트리밍에서는 그림 1 과 같이 사용자 시점으로 선택된 타일들만 추출하여 전송한다.

그 후 전체 영상을 포함하고 있는 저화질 비트 스트림을 따로 전송한다. 복호기 측에서는 이 두가지 영상을 각각 복호화한 후, 저화질 및 고화질 영상을 사용하여 결과 영상을 출력한다.

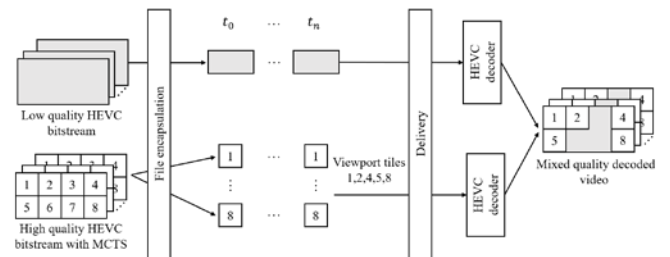


그림 1. 기존의 타일 스트리밍 구조도

이러한 방식의 타일 기반 스트리밍은 저화질과 고화질 시점 타일의 두 계층의 영상을 각각 전송해야 한다는 점에서 대역폭의 손해가 발생할 수 있다. 또한 복호기의 측면에서도 두 영상을 각각 복호화한 후 병합하는 과정이 필요하다는 단점이 있다. 따라서 본 논문은 고화질 사용자 시점 타일들과 함께 저화질 전체 타일들을 병합하여 하나의 계층으로 묶어서 전송하는 단일 계층 타일 바인딩 (binding) 을 사용한 타일 기반 스트리밍 기법을 제안한다. 본 논문의 구성은 다음과 같다. 2 절에서 기존

타일 기반 스트리밍 기법과 coding tree unit (CTU) 단위 품질 제어 기법, 그리고 제안하는 단일 계층 타일 기반 스트리밍 기법을 소개하고, 3 절에서 제안하는 방법과 기존 방법들의 품질 및 복호화 시간을 비교하는 실험을 및 결과를 제시하며, 마지막으로 4 절에서는 본 논문의 결론을 맺는다.

## 2. 복호기 적응적인 단일 계층 타일 바인딩

본 절은 복호기 적응적인 단일 계층 타일 바인딩을 이용한 타일 기반 스트리밍의 동작 과정에 관해 기술한다. 타일이 아닌 CTU 를 기반으로 하는 스트리밍에서는 그림 2 와 같이 원본 360 도 영상에서 사용자 시점 CTU 영역은 높은 품질, 그 외 영역은 낮은 품질로 부호화를 수행 및 전송한다. 이와 같은 방법에서는 대역폭 측면에서 이점을 가질 수 있으나, 사용자 시점이 바뀔 때마다 새로 부호화를 수행해야 하므로 부호화 시간이 지연 시간에 추가되어 종래의 타일 기반 스트리밍 방식 대비 지연 시간이 증가하는 단점이 있다.

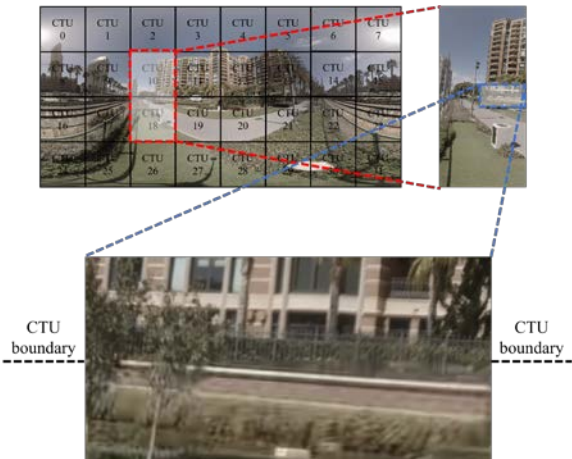


그림 2. CTU 기반 품질 제어 예시

그림 3 는 그림 1 에서의 타일 스트리밍과 달리 저화질의 비트스트림을 그대로 보내는 것이 아닌, MCTS 기술을 사용하여 부호화한 결과를 이용한다. 사용자 시점 타일에 포함된 타일들을 포함한 모든 타일들을 사용자 시점의 고화질 타일들과 함께 전송한다. 그 후, 복호기에서는 먼저 사용자 시점의 고화질 타일들과 나머지 부분들의 저화질 타일들을 병합하여 혼합 화질의 단일 비트스트림을 생성하고 복호화를 수행한다.

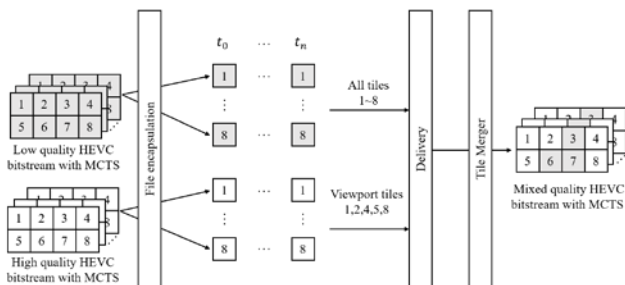


그림 3. 제안하는 단일 계층 타일 바인딩 구조도

## 3. 실험 및 결과

본 논문은 사용자 시점 기반 CTU 기반 스트리밍, 타일 기반 스트리밍, 그리고 제안하는 단일 계층 타일 바인딩을 이용한 타일 기반 스트리밍 방법을 비교한다. joint video experts group (JVET) 에서 정의한 360 도 영상에 대한 common test conditions (CTC)[7] 를 준수하여 실험을 진행하였다. CTU 단위 품질 제어를 사용하기 위해 본 실험에서는 부호화 시 HEVC 가속 소프트웨어인 Kvazaar 를[8], 복호화 시 HEVC 참조 소프트웨어인 HEVC test model (HM) 을 사용하였다[9]. 부호화 및 복호화 엔 32 프레임을 사용하였으며 사용자 시점 움직임 예시는 [10]의 0 번 옵션을 사용하였다. 각 방법마다 사용자 시점을 고려하지 않은 사용자 시점 독립적인 방법을 anchor 로 정의 및 포함하여 4 가지 방법으로 각각 4 가지 quantization parameter (QP) 값을 사용하여 부호화 및 복호화 과정을 실험하였다. 테스트 시퀀스로 Trolley 시퀀스가 사용되었다. 품질 평가 기법은 immersive video peak signal-to-noise ratio (IV-PSNR)[11], Y-PSNR 2 가 지를 사용하였다.

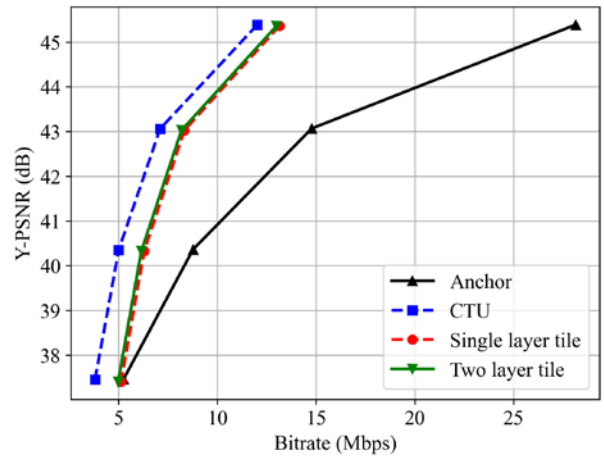


그림 4. RD-curves (Y-PSNR)

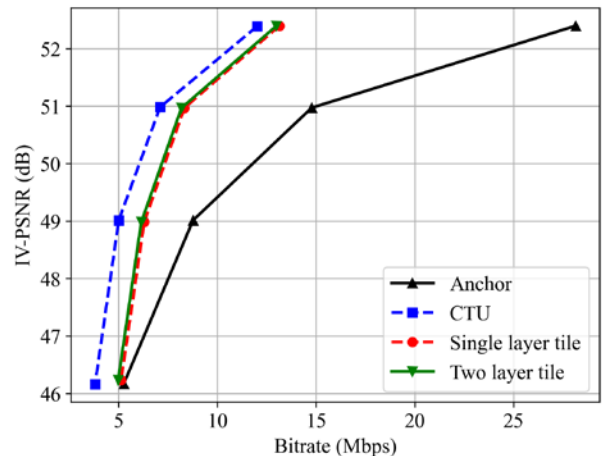


그림 5. RD-curves (IV-PSNR)

그림 4 와 그림 5 는 4 가지 부호화 방식의 rate-distortion (RD) curve 를 나타낸다. 단일 계층 타일 바인딩을 이용한 타일 스트리밍 방법(Single layer tile)의 경우, 두 가지 평가 기법에서 모두 기존 타일 스트리밍 기법에 비해 크게 떨어지지 않는 품질을 보였음을 알 수 있다. 그림 6 의 복호화 시간의 측면에서는 모든 QP 에 대해 기존 타일 스트리밍 기법 대비 복호화 시간을 크게 절약할 수 있음을 알 수 있다.

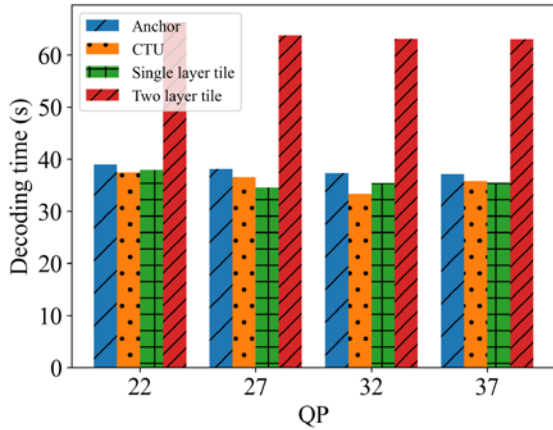


그림 6. 복호화 시간 비교 그래프

표 1 은 실험 결과 Bitrate 및 PSNR 값을 나타내며, 표 2 는 각 방법에서의 부호화 및 복호화에 걸린 시간(초)을 나타낸다. Two layer tile 방법에서의 복호화 시간은 base layer 와 tile layer 의 복호화 시간을 합산한 값을 의미한다.

표 1. 실험 결과 Bitrate 및 PSNR

QP	Anchor				
	Total kbps	Y-PSNR	U-PSNR	V-PSNR	IV-PSNR
22	28134.98	45.39	52.03	51.48	52.40
27	14779.04	43.07	50.57	50.41	50.97
32	8767.08	40.36	48.77	49.03	49.01
37	5255.83	37.46	46.38	46.94	46.17
QP	CTU				
	Total kbps	Y-PSNR	U-PSNR	V-PSNR	IV-PSNR
22	12016.58	45.38	52.02	51.47	52.39
27	7097.33	43.06	50.55	50.39	50.98
32	5004.11	40.35	48.80	49.01	49.01
37	3807.05	37.46	46.27	46.99	46.16
BD-rate Y-PSNR		-46.03%			
BD-rate U-PSNR		-45.00%			
BD-rate V-PSNR		-44.25%			
BD-rate IV-PSNR		-44.32%			
QP	Single layer tile				
	Total kbps	Y-PSNR	U-PSNR	V-PSNR	IV-PSNR
22	13138.23	45.36	52.02	51.48	52.40
27	8329.75	43.03	50.52	50.39	50.97
32	6295.63	40.32	48.76	49.02	48.98
37	5126.96	37.41	46.43	47.00	46.23
BD-rate Y-PSNR		-33.76%			

BD-rate U-PSNR		-32.33%			
BD-rate V-PSNR		-31.35%			
BD-rate IV-PSNR		-31.31%			
QP	Two layer tile				
	Total kbps	Y-PSNR	U-PSNR	V-PSNR	IV-PSNR
22	12988.40	45.36	52.02	51.48	52.40
27	8179.92	43.03	50.52	50.39	50.97
32	6145.80	40.32	48.76	49.02	48.98
37	4977.13	37.41	46.43	47.00	46.23
BD-rate Y-PSNR		-35.16%			
BD-rate U-PSNR		-33.80%			
BD-rate V-PSNR		-32.88%			
BD-rate IV-PSNR		-32.83%			

표 2. 실험 결과 encoding 및 decoding time

QP	Anchor	
	Enc T [s]	Dec T [s]
22	150.99	38.97
27	125.46	38.07
32	116.16	37.32
37	111.65	37.16
QP	CTU	
	Enc T [s]	Dec T [s]
22	131.50	37.47
27	117.60	36.52
32	113.61	33.33
37	111.34	35.77
QP	Single layer tile	
	Enc T [s]	Dec T [s]
22	94.67	37.97
27	97.54	34.60
32	91.69	35.37
37	87.02	35.43
QP	Two layer tile	
	Enc T [s]	Dec T [s]
22	94.67	66.24
27	97.54	63.74
32	91.69	63.06
37	87.02	63.00

#### 4. 결론

본 논문은 고화질 가상 현실 영상에서 사용자 시점 기반 360 도 영상 스트리밍을 위한 타일 기반 스트리밍 방식에서 이용될 수 있는 단일 계층 타일 바인딩을 제안한다. 제안하는 타일 바인딩 방법은 기존의 두 계층으로 사용자 시점 기반 타일과 그 외의 영상을 보내는 방법에 비해 큰 품질 하락 없이 복호화 시간을 크게 절약할 수 있음을 확인하였다. 향후 연구에서는 더욱 다양한 고화질 테스트 시퀀스에 대해 다양한 조건들로 부호화 실험을 진행할 계획이다.

### Acknowledgement

이 논문은 과학기술정보통신부에서 시행한 한국전자통신연구원의 연구개발지원사업의 지원을 받아 수행된 연구임 (No.2022-0-00022-001, 초실감 메타버스 서비스를 위한 실시간 입체영상 공간컴퓨팅 기술 개발)

### 참고문헌

- [1] J. -W. Son, D. Jang, E. -S. 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 (NOSSDAV), pp. 61-66.
- [2] J. -B. Jeong, S. Lee, I. Kim, E. -S. Ryu. 2021. Performance Analysis of Viewport-dependent Tiled Streaming on 16K Ultra High-quality 360-degree Video. KSII Journal of Internet Computing and Service (JICS), Vol. 22, No. 3.
- [3] S. Lee, D. Jang, J. -B. Jeong, S. Lee, E. -S. Ryu. 2019. Tile-Based 360 Degree Video Streaming System with User's gaze Prediction. Journal of Broadcast Engineering (JBE), Vol. 24, No. 6.
- [4] J. -B. Jeong, S. Lee, I. Kim, S. Lee, E. -S. Ryu. 2020. Implementing VVC Tile Extractor for 360-degree Video Streaming Using Motion-Constrained Tile Set. Journal of Broadcast Engineering (JBE), Vol. 25, No. 7.
- [5] J. -B. Jeong, S. Lee, I. Kim, E. -S. Ryu. 2021. Implementing Viewport Tile Extractor for Viewport-Adaptive 360-Degree Video Tiled Streaming. 2021 International Conference on Information Networking (ICOIN), pp. 8-12.
- [6] T. T. Le, J. -B. Jeong, S. S. Lee, J. Kim, E. -S. Ryu. 2021. An Efficient Viewport-Dependent 360 VR System Based on Adaptive Tiled Streaming. CMC-Computers, Materials & Continua, Vol. 66, No. 3, pp. 2627-2643.
- [7] J. Boyce, E. Alshina, A. Abbas, Y. Ye. 2017. JVET common test conditions and evaluation procedures for 360° video. Standard ISO/IEC JTC1/SC29/WG11, MPEG118/ n16891.
- [8] M. Viitanen, A. Koivula, A. Lemmetti, A. Ylä-Outinen, J. Vanne, T. D. Hämäläinen. 2016. Kvazaar: open-source HEVC/H. 265 encoder. In Proceedings of the 24th ACM international conference on Multimedia, pp. 1179-1182.
- [9] HM reference software, [Online]. Available at: [https://hevc.hhi.fraunhofer.de/svn/svn\\_HEVCSoftware/](https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/).
- [10] T. Ikai, Y. Yasugi, and T. Aono. 2017. Ahg8: Dynamic viewport generation for 360° video evaluation. Standard ISO/IEC JTC1/SC29/WG11, MPEG2017/m39669.
- [11] A. Dziembowski. 2019. Software manual of IV-PSNR for Immersive Video. Standard ISO/IEC JTC1/SC29/ WG11, MPEG127/n18709.