

360 영상의 타일 기반 적응적 스트리밍을 위한 인트라 프레임 공유(IFS) 알고리즘

이시원, 하호진
한양여자대학교

lsiwon922@gmail.com, hjha@hywoman.ac.kr

Tile-based 360 Video Adaptive Streaming using IFS(Intra Frame Sharing) method

Siwon Lee, Hojin Ha
Hanyang Women's University

요약

최근 360 미디어 콘텐츠의 수요가 높아지면서 360 영상을 효율적으로 전송하기 위한 다양한 압축 및 전송 기법들이 연구되고 있다. 360 영상의 스트리밍 기법 중 하나인 타일 기반 적응적 스트리밍 기법은 360 비디오를 타일 그리드로 나누어 부호화하고, 뷰포트(Viewport)에 해당하는 타일만 스트리밍 할 수 있도록 한다. 기존의 타일 기반 스트리밍 기법은 뷰포트가 전환될 때 서버에 새로운 인트라 프레임을 요구하고, 지연 시간을 줄이기 위해 빈번한 RAP(Random Access Point)가 요구된다는 단점이 있다. 본 논문은 기존의 타일 기반 스트리밍 방법에서 인트라 프레임의 공유를 통해 비트율의 효율을 높이는 IFS(Intra Frame Sharing) 알고리즘을 제안한다. 제안된 IFS 방법은 낮은 화질의 인트라 프레임을 높은 화질의 인트라 프레임 부호화 시 공유하여, 뷰포트 스위칭 시점에서 다른 화질 간의 뷰포트 전환 작업이 필요 없어 비트율의 증가를 최소화한다. 실험에서, IFS 방법은 기존의 MCTS 기반의 적응적 스트리밍 방식과 비교했을 때 BD-rate(%)가 약 13% 절감되었다.

1. 서론

최근 virtual reality(VR)을 지원하는 head-mounted display(HMD) 기기들의 활발한 보급으로, 360 영상을 이용한 몰입형 실감미디어의 이용이 증가되고 있다. 사용자에게 몰입감 있는 360 영상을 제공하기 위해 최소 ultra high definition(UHD)이상의 고해상도가 요구되며, 이는 높은 대역폭을 필요로 한다. 이를 위해 비대칭 샘플링 기법, 타일 기반 motion-constrained tile set(MCTS)[1],[2] 등 고 비트율을 가진 360 영상의 효율적인 데이터 전송을 위해 다양한 압축 및 전송 기법들이 연구되고 있다.

360 영상을 전송하는 기존의 방법은 360 전체 영상을 클라이언트에 전송하고, 사용자가 보는 방향에 해당하는 영역(Viewport, 뷰포트)을 HMD와 같은 수신 장치를 이용해 렌더링 해서 보여준다. 하지만, 뷰포트는 360 영상의 일부분만 차지하기 때문에 360 영상 전체를 높은 품질로 전송하는 것은 불필요한 네트워크의 대역폭 낭비가 발생한다. 이런 불필요한 대역폭 사용을 줄이기 위해서 뷰포트 기반 적응적 스트리밍(Viewport-adaptive streaming, VAS)[3]-[5] 방법이 제안되고 있다. 그 중 타일 기반의 MCTS가 대표적이다. MCTS는 압축 과정에서 타일의 공간적, 시간적 독립성을 적용하여, 각 타일을 독립적으로 디코딩 할 수 있다. 기존 방식과는 달리 뷰포트에 기반해 뷰포트에 해당하는 영역을 더 높은 품질의 타일로 독립적으로 전송할 수 있어 대역폭의 활용에서 이점을 가진다. 그러나 MCTS 기반의 VAS 방식은 영상을 볼 때 사용자가 고개를 돌릴 경우, 새로운 뷰포트에 해당하는

인트라 프레임을 서버에 빈번하게 요청하게 되고, 이로 인해 비트율이 증가된다.

본 논문에서는 360 영상의 효율적인 전송을 위해 뷰포트 스위칭 시 비트율의 절감을 위해 인트라 프레임 공유(Intra Frame Sharing, IFS) 알고리즘을 제안한다. 뷰포트 스위칭 시 새로운 인트라 프레임에 대한 요구 없이 기존에 전송된 낮은 화질의 인트라 프레임을 지속적으로 사용할 수 있어 비트율 증가를 최소화할 수 있는 장점이 있다. 제안된 IFS 방법은 기존의 MCTS 기반의 적응적 스트리밍 방식과 비교하여 뷰포트 스위칭 시 비연속적인 화질의 발생을 최소화할 수 있을 것으로 기대된다.

본 논문의 구성은 다음과 같다. 2절에서는 관련 연구에 대해 살펴보고, 3절에서는 본 논문에서 제안하는 IFS 알고리즘에 대해 설명한다. 4절에서는 제안한 방법의 실험 결과를 설명하고, 마지막으로 5 절에서는 본 논문에 대한 결론을 서술한다.

2. 관련 연구

2.1. 뷰포트 기반 적응적 스트리밍(VAS)

360 영상을 전송하는 여러 가지 스트리밍 기법이 존재한다. 뷰포트에 독립적인 스트리밍 방식은 360 전체 스트림을 스트리밍 하는 가장 간단한 방법이다. 기존 비디오 스트리밍 시스템에서 수정이 필요 없는 장점이 있지만, 대역폭의 활용과 디코더 자원의 낭비가 된다는 단점이

있다. 이러한 단점을 보완하기 위해 뷰포트 기반 스트리밍 방법이 제안되었다.

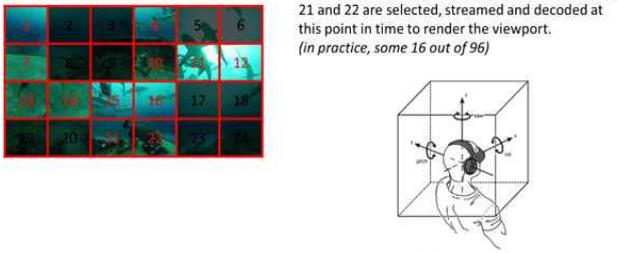


그림 1. 뷰포트 위치 기반 타일 선택
Fig. 1. Video selection based on viewport position[6]

VAS는 그림 1 과같이 뷰포트에 해당되는 부분을 선택해 독립적인 비디오 스트림으로 생성한다. 뷰포트에 해당하는 타일은 높은 화질로 렌더링하고, 나머지 영역도 커버하기 위해 뷰포트 외부의 영역은 더 낮은 화질의 스트림을 선택한다. 이렇게 하면 네트워크의 지연시간을 줄일 수 있다.

2.2. MCTS 기반 뷰포트 적응적 스트리밍(VAS)

앞서 말한 타일 기반 VAS 기법 중 최근 MCTS 기반 스트리밍 방법이 많이 연구되고 있다. MCTS는 H.265/HEVC(High efficiency video coding) 표준에 따라 프레임을 로우와 컬럼의 타일로 나누고, 각 타일은 독립적인 디코딩을 위하여 시간적 그리고 공간적 독립성을 가진다. 움직임 벡터 및 인트라 필터의 적용이 경계를 넘지 않도록 타일 내로 제한되고, 독립적인 부호화가 가능하다.

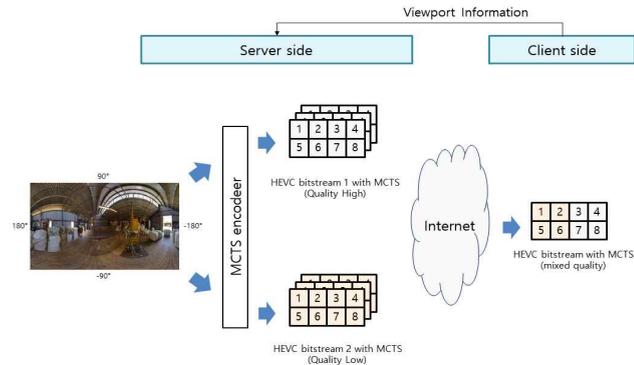


그림 2. MCTS를 이용한 타일 기반의 적응적 360 영상의 전송방법
Fig. 2. 360 video transmission using MCTS-based viewport-adaptive streaming(VAS)

그림 2는 MCTS를 이용한 360 영상의 전송 방법을 설명하고 있다. 높은 화질의 비트스트림과 낮은 화질의 비트스트림을 서버에 저장하고, 사용자의 관점에 따라 뷰포트 영역은 높은 화질의 타일로, 나머지 영역은 낮은 화질의 타일로 구성된 비트스트림을 전송한다. 기존의 H.265/HEVC를 이용해 전체 영상을 부호화하여 제공하는 스트리밍 방법에 비해, MCTS를 이용한 적응적 스트리밍 방법은 저장 공간과 네트워크의 대역폭을 절약할 수 있다는 장점이 있다. 그러나 사용자의 머리 움직임에 따른 뷰포트의 스위칭이 발생하면서 높은 화질과 낮은 화질의 타일셋을 변경하기 위해 빈번한 인트라 프레임들을 요구하고, 이는 비트율의 증가와 전송 지연(Latency)으로 이어진다.

3. 제안된 IFS(Intra Frame Sharing) 알고리즘

이 절에서는 뷰포트 스위칭 시에 MCTS 기반의 적응적 스트리밍 시스템의 성능 향상을 위한 IFS 알고리즘을 설명한다. 뷰포트의 변화에 따라 클라이언트는 새로운 뷰포트 정보를 전송하여, 뷰포트에 해당되는 높은 화질의 타일과, 나머지 낮은 화질의 타일들을 서버에 요청한다. 새로운 인트라 프레임을 다운받는 과정에서 순간적으로 비트량이 증가되고, 이는 원활한 뷰포트 변화에 따른 화질 변화를 어렵게 만든다. 본 논문에서는 스위칭 포인트에서의 원활한 뷰포트 전환을 위해 낮은 화질의 인트라 프레임을 공유하는 IFS 알고리즘을 제안한다.

제안된 IFS 알고리즘은 뷰포트 스위칭 시점에서 낮은 화질의 인트라 프레임을 공유 인트라 프레임(Shared I frame, SIF)으로 설정하고, 높은 화질의 인트라 프레임의 부호화에 SIF를 사용한다.

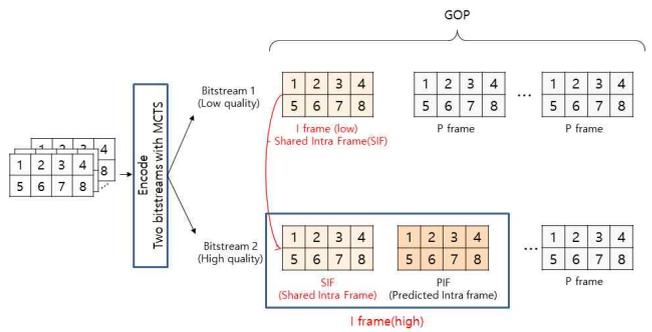


그림 3. 제안된 IFS를 이용한 부호화 과정
Fig. 3. Encoding process using proposed IFS method

그림 3은 IFS 방법의 부호화 과정을 나타낸다. 2개 화질의 비트스트림이 이용되며, 첫 번째는 낮은 화질, 두 번째는 높은 화질의 비트스트림을 나타낸다. 낮은 화질의 비트스트림은 기존의 타일 기반의 MCTS 부호화 과정을 동일하게 적용한다. 다른 점은 높은 화질의 비트스트림에서 인트라 프레임은 SIF와 예측된 인트라 프레임(Predicted intra frame, PIF)으로 구성된다. 여기서 PIF는 높은 화질의 인트라 프레임과 SIF 사이의 예측 과정으로 생성된 차분 프레임이다.

만약 뷰포트 스위칭이 발생해 높은 화질에서 낮은 화질의 비트스트림으로 이동할 경우, 높은 화질에서는 낮은 화질의 인트라 프레임(SIF)을 그대로 이용할 수 있기 때문에, 서버에 추가적인 인트라 프레임에 대한 전송을 요구하지 않아도 된다. 반대로, 낮은 화질에서 높은 화질의 비트스트림으로 스위칭이 발생될 경우, 낮은 화질의 SIF가 이미 클라이언트 측에 있기 때문에 높은 화질의 PIF만 다운로드하여 높은 화질의 인트라 프레임(SIF + PIF)을 생성할 수 있다. 이로 인해 스위칭 포인트에서 인트라 프레임의 빈번한 요청 없이 뷰포트 전환이 가능하다. 제안된 IFS 알고리즘은 높은 화질에서 인트라 프레임 복호화 시 기존의 MCTS 방법보다 약간의 복잡도를 증가시키지만, 비트율 감소 면에서 더 탁월하다.

4. 실험 결과

본 절에서는 제안된 IFS 기반의 viewport-adaptive streaming (IFS-VAS)의 성능을 분석한다. 제안된 IFS-VAS의 성능 비교를 위

해 MCTS 기반의 VAS(MCTS-VAS)와 VAS를 사용하지 않고 높은 화질의 ERP 영상을 모두 전송하는 방식(Non-VAS)을 이용하였다.

실험에 사용된 테스트 시퀀스는 equirectangular format(ERP) 포맷으로 4K 영상(3840×1920)으로 구성된 ‘AerialCity’를 사용하였다. 테스트 시퀀스의 부호화 및 복호화를 위해서 HEVC test model (HM) 16.20 [7] 버전이 사용되었으며, 타일링은 4×2로 설정하였다. 총 160 프레임을 사용하고, GOP 크기는 16으로 하여 random access point(RAP) 간격을 설정하였다. IFS-VAS에서는 3절에서 제시된 10개의 PIF를 높은 비트스트림 구성에 추가하였다.

제안된 IFS-VAS에서 사용되는 높은 화질의 비트스트림은 QP를 22, 26, 30, 37로 설정하고, 낮은 화질의 비트스트림은 높은 화질의 비트스트림 구성에 사용된 QP보다 7 작은 값을 설정하여 비트스트림을 구성하였다. 성능 평가를 위해 테스트 영상의 Luma 영상에 대한 BD-rate(%) [8]을 사용하였다.

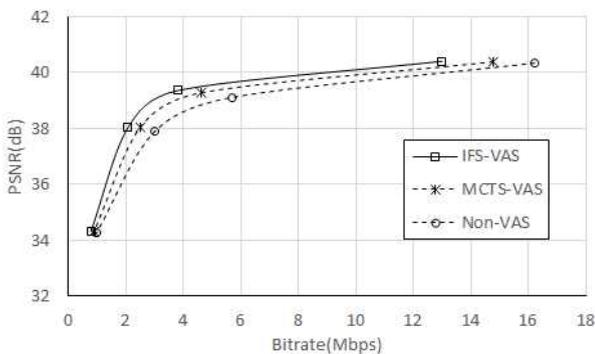


그림 4. IFS-VAS, MCTS-VAS 그리고 Non-VAS에 대한 RD-curves
Fig. 4. RD-curves of ‘AerialCity’ sequence for IFS-VAS, MCTS-VAS and Non-VAS

그림 3은 제안된 IFS-VAS와 MCTS-VAS, Non-VAS에 대해 높은 화질의 비트스트림에 설정된 4개의 QP에 대한 rate-distortion(RD) curve를 나타낸다. 모든 QP에서, 제안된 IFS-VAS는 MCTS-VAS와 Non-VAS보다 비트율이 절감되었다. 특히 높은 비트율에서 IFS-VAS는 MCTS-VAS에 비해 10% 이상의 대역폭을 줄일 수 있었다. 이는 뷰포트 스위칭 시 새로운 인트라 프레임에 대한 요구 없이 기존에 전송된 낮은 화질의 인트라 프레임(SIF)을 지속적으로 사용할 수 있어, 새로운 인트라 프레임 요구로 인한 비트율 증가를 최소화할 수 있기 때문이다.

Name	MCTS-VAS(%)	IFS-VAS(%)
AerialCity	-17.83	-31.32

표 1. Non-VAS 대비 IFS-VAS와 MCTS-VAS의 BD-Rate(%) 비교
Table 1. BD-Rate(%) comparison for 4x2 tiling of MCTS-VAS and IFS-VAS method relative to Non-VAS

표 1은 Non-VAS 대비 IFS-VAS와 MCTS-VAS의 BD-Rate(%)을 나타낸다. 음수 값은 기준이 되는 Non-VAS와 같은 peak signal to noise ration(PSNR)을 가질 때, MCTS-VAS와 IFS-VAS가 얼마나 많은 양의 비트율을 줄일 수 있는지 나타낸다. IFS-VAS와 MCTS-VAS는 Non-VAS와 비교하여 15% 이상의 비트율 감소를 나타내었다. 특히 IFS-VAS는 뷰포트 스위칭 시에, 높은 화질의 PIF만

을 서버로부터 다운로드하여 높은 화질의 인트라 프레임을 생성할 수 있어 비트율이 감소되었다. 반면, MCTS-VAS는 뷰포트 스위칭 시에 높은 비트율을 갖는 인트라 프레임을 서버에 빈번하게 요청하기 때문에, IFS-VAS에 비해 비트율이 증가되었다.

5. 결론

본 논문에서는 타일 기반 360 비디오의 적응적 스트리밍에서 효율적인 비트율 사용을 위해 IFS 알고리즘을 제안한다. IFS 알고리즘은 뷰포트 스위칭 시점에서 인트라 프레임을 공유함으로써 인트라 프레임 요청을 최소화하여 비트율을 감소할 수 있다. MCTS-VAS와 비교하여 실험한 결과에서 IFS-VAS가 비트율 감소 면에서 더 좋은 결과를 보였다.

Acknowledgement

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1A2C1010476).

참고 문헌

- [1] S. Lederer, “Today’s and future challenges with new forms of content like 360 AR and VR,” in MPEG workshop Global Media Technology Standards for an Immersive Age, 2017.
- [2] D. Podborski, E. Thomas, M. Hannuksela, S. Oh, T. Stockhammer, and S. Pham, Virtual reality and DASH, in International Broadcasting Convention, 2017.
- [3] A. Zare, A. Aminlou, M. M. Hannuksela, and M. Gabbouj, “HEVCcompliant tile-based streaming of panoramic video for virtual reality applications,” in Proceedings of the 2016 ACM on Multimedia Conference, pp. 601–605, 2016.
- [4] Y. Sanchez de la Fuente, R. Skupin, and T. Schierl, “Video processing for panoramic streaming using HEVC and its scalable extensions,” Multimedia Tools and Applications, vol. 76, no. 4, pp. 5631–5659, 2017.
- [5] R. Ghaznavi-Youvalari, M. M. Hannuksela, A. Aminlou, and M. Gabbouj, “Viewport-dependent delivery schemes for stereoscopic panoramic video,” in 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2017, pp. 1–4, 2017.
- [6] ISO/IEC JTC 1/SC 29/WG 11 MPEG2019/N18357 “Requirements for Immersive Media Access and Delivery”, March, 2019, Geneva, CH.
- [7] HM reference software 16.20. [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.20.
- [8] G. Bjontegaard, “Calculation of average PSNR differences between RDCurves,” VCEG-M33, 2001.