

몰입형 비디오에서의 비대칭 양자화 파라미터 성능 분석

이순빈, 정종범, 김인애, 류은석
 성균관대학교 컴퓨터교육과

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

Performance Analysis of Asymmetric Quantization Parameter On Immersive Video

Soonbin Lee, Jong-Beom Jeong, Inae Kim, Eun-Seok Ryu
 Department of Computer Education

요약

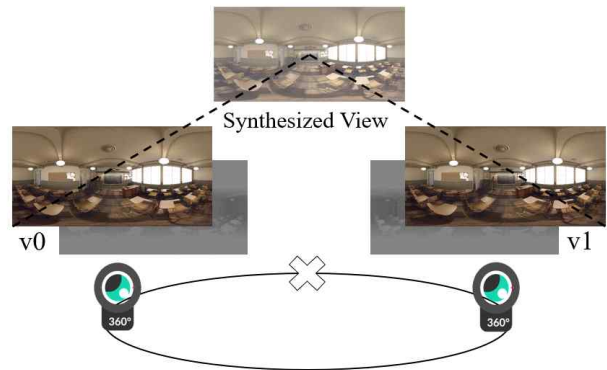
몰입형 비디오(Immersive Video)는 각 시점의 다수 영상과 깊이 지도 정보를 활용하여 사용자에게 보다 높은 자유도로 감상할 수 있게 하는 기술이다. 이러한 몰입형 비디오 서비스를 원활히 제공하기 위해서는 다수의 영상이 필요하므로 보다 높은 대역폭을 요구하게 된다. 본 논문에서는 MPEG-I (immersive) 그룹에서 표준화 진행중에 있는 기술인 MIV(model for immersive video)에서 채택하고 있는 바탕 시점(basic view) 모델에 대하여, 바탕 시점과 추가 시점에 대한 양자화 파라미터를 구분하여 선택적으로 영상의 비트레이트를 효과적으로 조절할 수 있는 방법을 제안한다. 추가 시점의 경우, 고주파의 텍스처 경계가 모인 특성상 비트레이트가 높게 나타나나 감상하는 사용자가 감상하는 영역에서 보다 적은 손실을 나타낸다. 제안하는 방법을 통해 상대적으로 큰 품질의 손실 없이 몰입형 비디오 영상의 비트레이트를 절감하는 것을 관찰하였다.

1. 서론

현재 표준화 단체 MPEG(moving picture experts group)에서는 여러 위치에 다양한 시점으로 존재하는 영상들을 취득하고 처리하여 사용자에게 화면 전환 이동에 따른 운동 시차(motion parallax)를 제공할 수 있는 몰입형 미디어 기술에 관한 표준화가 진행되고 있다. 이러한 표준화의 일환으로 2019년 3월 MPEG 126차 회의에서 MPEG-I (Immersive) 서브그룹은 6DoF(degrees of freedom)을 제공하는 몰입형 미디어를 지원하는 테스트 모델(Test Model for Immersive Video, TMIV)를 정의하였다[1,2].

6DoF 몰입형 미디어 기술 표준화는 몰입형 미디어를 위한 다수의 영상 처리를 목적으로 MIV(Model for Immersive Video) 기술에 대한 성능 탐색을 현재도 지속적으로 수행하고 있다. 6DoF기술은 자유로운 사용자의 시선과 움직임에 따른 시점을 지원하므로, 이를 위해 여러 위치에서 취득된 영상들을 동시에 압축 및 전송하여 사용자 시점의 영상을 취득하고 합성하는 과정에서 높은 대역폭과 연산량이 요구된다.

현재 MIV에서는 깊이 맵 기반 이미지 렌더링(DIBR) 기법을 바탕으로 하여 다수의 영상들을 깊이 맵과 함께 취득한 후, 깊이 맵을 바탕으로 각 영상의 시점과 거리를 계산하여 가상의 영상 시점을 합성하는 방식으로 진행된다. 따라서 사용자의 시선과 위치에 맞도록 시점을 합성해낼 수 있으며, 이로 인해 사용자는 보다 상호작용이 가능하고 몰입감 있는 미디어를 감상할 수 있게 된다.



(그림 1) 깊이 맵 기반 이미지 렌더링(DIBR)을 통한 중간 시점 생성의 예시

2. 비대칭 양자화 파라미터 할당

MIV에서는 바탕 시점(Basic View)과 각 시점의 고유한 영상 정보를 패치 단위로 모아둔 추가 시점(Additional View)으로 처리하는 모델을 채택하고 있다. 즉, 바탕 시점은 단일 원본 영상을 전송하고, 나머지 영상들은 바탕 시점과의 차이만을 계산한 잔차 영상 정보를 가지고 있게 된다[3].



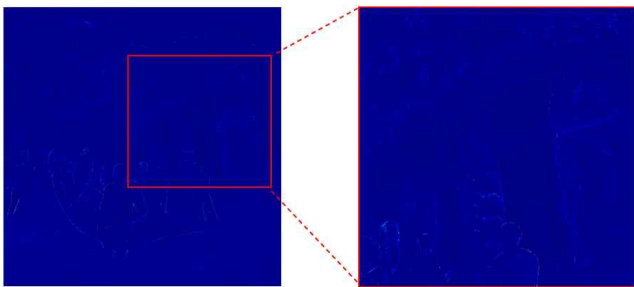
(그림 2) ClassroomVideo 6DoF 테스트 영상의 바탕 시점(좌), 추가 시점(우)

이 잔차 영상 정보 각각을 패치(patch)라고 하며, 이 패치들이 하나의 영상에 병합되어 아틀라스(atlas)가 된다. 바탕 시점은 단일 영상을 그대로 전송하지만, 추가 시점은 영상의 특성에 따라 최종 합성 영상 품질에 크게 영향을 달라지게 된다. 각 영상들의 중복성이 매우 클 경우 바탕 시점만으로도 대부분의 6차원 자유도 시점에 대응할 수 있으므로 추가 시점에 대한 중요성이 상대적으로 떨어지게 된다. 그러나 추가 시점의 경우 전체 영상이 아닌 각 차분 정보들을 부분적으로 병합한 영상으로 고주파 텍스처가 많이 나타내는 모습을 보여 영상 압축 과정에서 효율이 상대적으로 낮게 나타나게 된다.

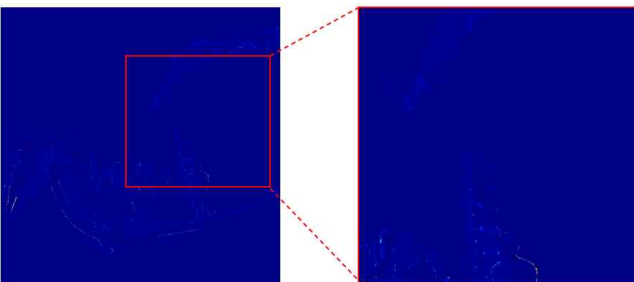
따라서 본 논문에서는 바탕 시점과 추가 시점의 중요도 여부에 따라 양자화 파라미터를 다르게 할당함으로써 사용자 시점의 큰 품질 손실 없이 비트레이트를 효율적으로 절감시킬 수 있는 방법을 제안한다.

3. 실험 결과 및 분석

본 논문에서는 제안하는 방법의 효율을 검증하기 위해 MPEG-I에서 정의되어 있는 공통 실험 조건(Common Test Conditions) 영상 3개를 선정하여 사용하였다. 영상 부호화에는 HEVC test model (HM) 16.20 버전을 사용하였고, 패치를 1개 이상 포함하는 영상을 추가 시점으로 정의하였다[4]. 양자화 파라미터 테이블은 [22,27,32,37,42], QP1에서 QP5까지로 구성되어 있으며 추가 시점의 경우 바탕 시점보다 양자화 파라미터의 값을 1단계 높여 부호화를 진행하였다.



(그림 3) QP1과 QP2 합성 영상의 차이 픽셀
원본 영상(좌), 확대 영상(우)



(그림 4) QP1과 Proposed(QP1+2)의 합성 영상의 차이 픽셀
원본 영상(좌), 확대 영상(우)

그림 3은 바탕 시점과 추가 시점 모두 각각 단일하게 QP1과 QP2로 압축하였을 때의 합성 영상의 차이를 나타낸다. 그림 4는 바탕 시점은 QP1, 추가 시점은 QP2로 품질을 낮추어 압축하였을 때의 합성 영상의 차이를 나타낸다. 그림 3과 달리 그림 4에서는 오직 합성에서 영향을 끼치는 부분에 대해서만 손실이 일어난 것을 확인할 수 있다. 이는 영상의 추가

시점의 정보량에 따라 압축에 있어서의 품질 손실이 기존의 영상들과는 다른 방식으로 일어날 수 있음을 시사한다.

실험 지표의 경우 공통 실험 조건을 따라 바탕 시점과 추가 시점을 부호화하고, 이를 다시 복호화한 뒤 바탕 시점과 추가 시점으로 다시 각 시점의 원본 영상들을 복원했을 때의 평균에 대하여 품질 평가를 계산하였다. 영상의 품질 평가에는 WS-PSNR, IV-PSNR, VMAF가 사용되었다. 표 1은 비대칭 양자화 파라미터 할당을 이용하여 압축을 진행하였을 때의 기존 대비 bjontegaard delta rate (BD-rate) 절감률을 보여준다.

(표 1) 비대칭 양자화 파라미터 할당 방법의 BD-rate 절감률

Sequence name	WS-PSNR	IV-PSNR	VMAF
ClassroomVideo	-9.76%	-9.99%	-4.76%
Museum	-17.97%	-6.88%	-11.53%
Average	-13.86%	-8.43%	-8.14%

4. 결론

본 논문은 몰입형 비디오 압축 전송 과정에서 비대칭 양자화 파라미터 할당 방법의 효율을 검증하였다. 추가 시점의 중요도에 따라 절감률이 크게 달라짐을 관찰하였으며, 비교적 추가 시점의 중요도가 떨어지는 영상의 경우 BD-rate 절감을 확인하였다. 향후 연구로 사용자 시점에 대한 시나리오를 가정하여 시점에 대한 영상(Viewport) 렌더링을 진행하여 감상하였을 때에 대한 품질 손실을 확인하여 연구를 확대할 계획이다.

ACKNOWLEDGMENT

이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1A2C1010476). 또한 이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2020-0-00231, 5G 엣지클라우드 기반 VR·AR 저지연 스트리밍 기술 개발)

참고문헌

- [1] Mary-Luc Champel, Thomas Stockhammer, Thierry Fautier, Emmanuel Thomas, Rob Koenen, "Quality Requirements for VR", 116th MPEG meeting of ISO/IEC JTC1/SC29/ WG11, MPEG 116/m39532, 2016.
- [2] Jong-Beom Jeong, Soonbin Lee, Dongmin Jang, EunSeok Ryu, "Towards 3DoF+ 360 Video Streaming System for Immersive Media", IEEE Access, 7, pp. 136399-136408, 2019.
- [3] "Working Draft 3 of Immersive Video", ISO/IEC JTC1/SC29/WG11 MPEG2020/ w18794, Geneva.
- [4] "Common Test Conditions for Immersive Video", ISO/IEC JTC1/SC29/WG11 MPEG2020/ w18789, Geneva.