

대역폭 제한된 환경에서 다중시점 360VR 콘텐츠의 시점변경에 따른 전송지연에 대한 연구

*서봉석 황수빈 **김동호

서울과학기술대학교

*sbs91, hsb0710, **dongho.kim@seoultech.ac.kr

Viewpoint-Switching Delay of Multi-viewpoint 360VR in a Bandwidth-limited Environment

Bong-Seok Seo *Subin Hwang **Ye hoon Lee ***Dong Ho Kim

Seoul National University of Science and Technology

요약

시청자의 시점 위치를 선택하여 360VR 콘텐츠를 경험하는 다중시점 또는 자유시점 360VR 콘텐츠의 전송은 매우 높은 대역폭을 소비하게 된다. 특히 다중 사용자가 존재하는 셀룰러 환경에서는 특정 사용자에게 한정된 대역폭만을 할당할 수 있으므로 불가피하게 전송 품질 악화 또는 전송 지연을 겪게 된다. 다중시점(multi-viewpoint) 360VR 콘텐츠의 경우 시청자가 시점의 위치를 변경하면 해당 위치의 콘텐츠를 전송하므로, 한정된 대역폭이 주어지는 통신환경에서 시점을 이동할 때 마다 큰 전송 지연을 겪게 되어 사용자 경험이 매우 떨어지는 것을 예상할 수 있다. 본 논문에서는 다시점 360 비디오에서 시점 변경을 시도할 때 전송 지연을 줄이기 위하여 주변 위치의 영상을 적절히 Pre-Loading 하는 전송방식을 고려하고, 실질적인 5G 네트워크 환경에서 시점 변환에 따른 지연시간 성능을 분석하였다. 특정 수준 이하로 지연시간을 낮추기 위해서는 사용자 이동패턴에 대한 적절한 예측 기법이 요구되며, 정확도에 따른 지연시간 감소를 실험을 통하여 확인하였다.

1. 서론

360 비디오 서비스가 본격화된 이후로 시청자에게 강화된 실감 경험을 제공하기 위한 연구가 진행되고 있으며 시청자의 시점 위치를 변화하며 360 비디오 서비스를 경험할 수 있는 6DoF (Degree of Freedom)에 대한 연구 및 표준화가 MPEG-I(Moving Picture Expert Group - Immersive) 표준에서 논의되고 있다[1]. 3DoF + 또는 6DoF VR 서비스에서 다른 시점 (다른 위치 공간)에서 영상을 시청하려면 다른 위치에서 동시에 촬영한 여러 비디오를 사용하여 추가 시점의 영상을 렌더링해야 한다. 3DoF + 또는 6DoF 비디오를 재구성하는 여러 비디오를 전송하는 것은 4K 또는 8K 보다 고해상도의 영상을 지원하므로 매우 큰 부피를 가지고 있다[2].

모든 위치에서 360 비디오를 제공하는 완전한 6DoF는 현실적으로 제약이 있으므로 다중시점(multi-viewpoint)으로 양자화된 위치에서 3DoF 360VR(Virtual Reality) 콘텐츠를 구성하는 방식을 고려한다. 이때 요구되는 전송 데이터율은 3DoF 360 비디오 영상과 비교하여 양자화된 시점 위치의 개수에 비례하여 증가하는 것을 알 수 있다. 하지만, 한정된 대역폭을 갖는 통신시스템, 특히 여러 사용자가 존재하는 셀룰러 환경에서는 특정 사용자에게 매우 제한된 대역폭만을 할당할 수 있으므로 다중시점 360 비디오 서비스를 제공하는 것은 매우 제약적이다. 따라서, 불가피하게 전송 품질 악화 또는 전송 지연을 겪게 된다.

다중시점 360VR 시스템에서는 클라이언트인 시청자가 시점의 위치를 변경하면 해당 위치의 콘텐츠를 서버로부터 전송하는 방식을 고려하며 한정된 대역폭이 주어지는 통신환경에서 시점을 이동할 때 마다 큰 전송 지연을 겪게 되어 사용자 경험이 매우 떨어지는 것을 예상할 수 있다. 실제로 시력이 좌/우 1.0인 사람의 경우 픽셀 분해능이 1°당 40픽셀이며, 이는 이상적인 360VR 영상인 12K 해상도와 동일하다. 현재 유튜브에서 제공하는 2K 30fps 영상의 대역이 3~6Mbps 라는 것을 고려했을 때 현재 영상 압축코덱을 기반으로 12K 영상을 보내기 위해서는 1개의 영상 당 200~300Mbps가 필요할 것으로 예측된다. 이처럼 큰 용량의 영상을 기반으로 다중시점 영상을 구현할 때, 사용자가 위치를 옮길 때 마다 큰 용량의 영상을 받아야 하기 때문에 높은 view-switching 지연이 필수적으로 발생한다. 이러한 시점 이동에 따른 지연시간을 줄이기 위해서는 사용자가 움직이기 전에 미리 영상을 다운로드 해야 하는데, 이를 위하여 기존보다 더운 많은 영상을 다운로드 받아야 하는 상황이 나타난다.

본 논문에서는 다시점 360 비디오에서 시점 변경을 시도할 때 전송 지연을 줄이기 위하여 주변 위치의 영상을 적절히 Pre-Loading 하는 전송방식을 고려하고, 실질적인 5G 네트워크 환경에서 시점 변환에 따른 지연시간 성능을 분석한다. 실제 5G 네트워크[3]에서 한사람에게 제공할 수 있는 네트워크 대역은 이상적인 환경과는 많이 다르기 때문에 2장에서 실제 네트워크 환경에서 제시하는 대역과, 콘텐츠의 필요 대역

에 대하여 논의하고, 이후 제한된 개수의 영상을 보낼 수 있는 네트워크에서 viewpoint-switching 지연을 줄여 사용자 경험을 높이기 위해 낮은 지연을 위한 예측 모델의 정확도를 가상으로 구상하고, 이를 적용했을 때 지연 확인을 통해 목표에 따른 위치 이동 예측 정확도를 확인할 수 있도록 한다. 최종적으로 현재 실험 결과와 이후 연구 방향에 대하여 논의하도록 한다.

2. 자유시점 360비디오의 전송 및 한계

자유시점 360비디오를 전송하기 위해서는 사용자가 움직일 수 있는 주변 영상을 미리 다운로드하여 디코딩 하고, 그 이미지를 기반으로 사용자 위치의 영상을 합성하는 과정이 필요하다. 이와 같은 과정을 위해 자유시점 콘텐츠를 구현했을 때 사용자는 영상 생성을 위해 적어도 2개에서 많으면 4개 이상의 영상을 전송받아야 한다.

자유시점 혹은 다시점을 구현하는 영상의 영상 대역을 확인할 경우 표 2와 같이 현재 HEVC(High Efficient Video Coding) 코덱을 활용하고, 이미지 손상을 고려하여 대역을 정할 경우 4K60p 영상은 20-40Mbps로 압축이 가능하며, 최근 발표 및 표준화가 진행된 VVC를 활용했을 경우 8K90p 영상은 90-130Mbps로, 12K3D120p 영상은 500-700Mbps로 압축이 가능하다. 또 360 영상의 경우 실제 사용자 시청 영역이 전체 영역이 될 수 없으므로, 사용자 시청영역만을 전송하는 것을 고려 했을 때 4K60p FoV는 10-20Mbps, 8K90pFoV는 30-50Mbps, 12K3D120pFoV는 200-300Mbps로 압축 가능하다. 따라서 현재 5G 네트워크의 이상적인 환경에서 이상적인 360VR 영상을 스트리밍 하려고 할 때 1명에게 3개 시점의 영상만을 제공할 수 있으며, 다중 사용자 환경에서는 12K가 아닌 더 작고 더 프레임레이트가 낮은 영상을 제공해야 한다. 이러한 문제는 360 영상 및 다시점 영상에서 가장 큰 문제로 여겨지며, 단일 시점 360 영상에서도 사용자가 보는 영역이 전체 이미지에 비해 작은 영역이기 때문에 생기는 비효율을 해결하기 위해 전체 이미지 중 일부만을 보내는 방법을 연구하여 전송방식을 다양화하고 있다[6-9].

	codec	form	Mbps	fps
4K	HEVC	Sphere	20-40	60
		FoV	10-20	
8K	VVC	Sphere	90-130	90
		FoV	30-50	
12K3D	VVC	Sphere	500-700	120
		Fov	200-300	

표 1 360 Video의 전송 형태에 따른 대역 및 지연

5G 네트워크를 이용하여 다중시점 또는 자유시점 360 비디오를 전송한다고 가정할 때, 5G 네트워크에서 제공하고자 하는 최고 데이터 전송률은 2Gbps 이며[4], 실제 실내 환경 혹은 다중 사용자 환경에서는 3-400Mbps 로 감소된다[5]. 이를 확인하기 위해 NS-3의 mmwave를 통하여 28GHz, 73GHz를 활용하는 다중사용자 환경의 네트워크에서 사용자 당 데이터 처리율(throughput)을 확인하였다. 그림 1에 실험 환경을 제시하였다. 보행자 속도로 움직이는 사용자가 10명이고, 실내 환경에서 0.1초마다 움직이는 방향 및 속도를 바꾸는 경우 28GHz와 73GHz

대역을 carrier aggregation(CA) 했을 때 1명의 사용자 단말 당 데이터 처리율을 확인하는 실험을 진행하였으며, 100Mbps가 나오는 것을 확인할 수 있었다.

Mode	5G Terminal Configuration					Peak Data Rate
	direction	Bandwidth	Number of Stream	Modulation	Ratio	
1	UL	100Mhz	2	256QAM	3:7	230Mbps
	DL	100Mhz	4	256QAM	3:7	2Gbps
2	UL	100Mhz	2	64QAM	3:7	175Mbps
	DL	100Mhz	4	256QAM	3:7	2Gbps

표 2 5G 통신 모듈에서 최고 데이터레이트

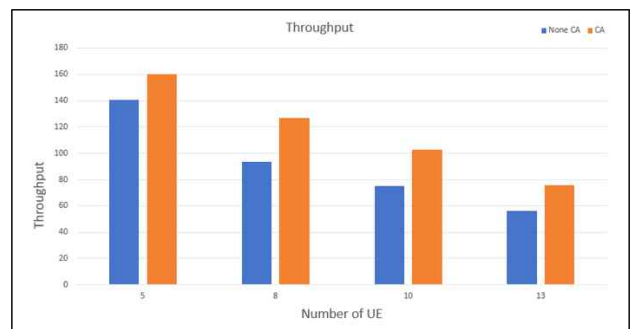
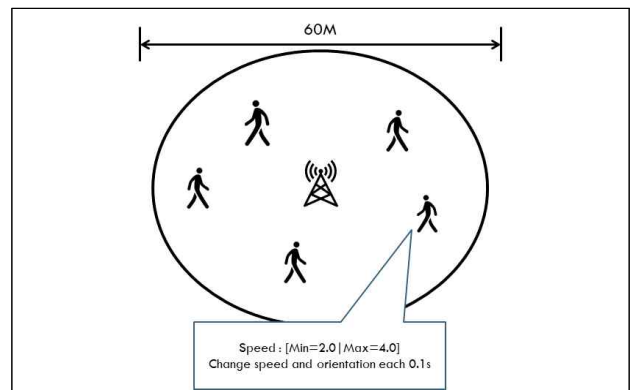


그림 1. 다중 사용자 환경에서 사용자(UE) 별 Throughput

실사를 기반으로한 6DoF 콘텐츠의 경우 시점 간 간격이 좁을수록 생성된 View 의 완성도가 높아지며, 사용자의 이동범위가 다운로드 받은 시점 이내로 제한되기 때문에, 사용자가 지연을 겪지 않게 하기 위하여 최대한 많은 위치의 영상을 다운로드해야 할 필요가 있다. 하지만 앞서 보인 것 과 같이 네트워크의 한계 때문에 실제 전송되는 영상의 수는 매우 한정적일 수 밖에 없다. 이때 영상의 종류에 따라 움직임의 범위가 늘어날 수 있고, 이때 영상이 로드 되지 않은 경우 로드가 완료될 때까지 콘텐츠를 즐길 수 없어 사용자 경험이 떨어지기 때문에 최대한 끊기지 않도록 유지하려고 해야 한다. 때문에 네트워크의 한계와 구현방법의 상충으로 실사를 활용한 자유시점 구현에 있어서 자연스럽게 위치 이동에 대한 예측을 고려해야만 한다.

현재 대부분의 연구들이 6DoF 형태에서의 고려가 아닌 단일 시점에서의 사용자 눈 이동 및 방향 이동 등을 고려하는 연구를 진행 중이다. 이상적인 목표로 표 1에서와 같이 영상의 해상도 및 초당 프레임 수의

목표치가 가파르게 올라가기 때문에 이를 해결하기 위해 다중시점 콘텐츠를 구현하여, 시점을 바꾸는 경우 생기는 지연을 측정하고자 한다.

3. 다중시점 360 비디오의 시점이동 지연 측정 실험

웹을 기반으로 한 시점이동 콘텐츠를 구현하기 위하여, A-FRAME 과 DASH(Dynamic Adaptive Streaming over HTTP) 플레이어를 통하여 사용하였다. A-FRAME은 Web에서 360환경을 구성할 수 있게 해주는 API 이며[20], 기본적인 DASH 기본적인 DASH 플레이어 구조에서는 하나의 Decoder가 여러개의 영상 버퍼를 가지지 못하여[21], 중간 뷰 생성을 위한 구조에서는 각 뷰마다 개별적인 디코더가 필요로 하므로, 1개의 View 마다 1개의 플레이어를 구현하여, 위치를 변경할 때 마다 새로운 플레이어가 플레이 될 수 있도록 하였다. 영상은 실제 8개의 시점을 동시 녹화하여 준비 하였으며, 총 3열로써 아래 그림과 같이 1열 3 포인트, 2열 2포인트, 3열 3포인트로 움직일 수 있는 위치가 존재한다. 6DoF를 구현하기 위한 방법으로 고려중인 RGBD의 경우 각 위치의 영상을 Decoding 한 뒤 Synthesis View를 합성하여 제공하는 과정을 거치므로, Synthesis View 합성 시간을 고려하지 않고 단순히 주변 위치로의 이동만을 고려했을 때 지연을 측정 및 평가한다.

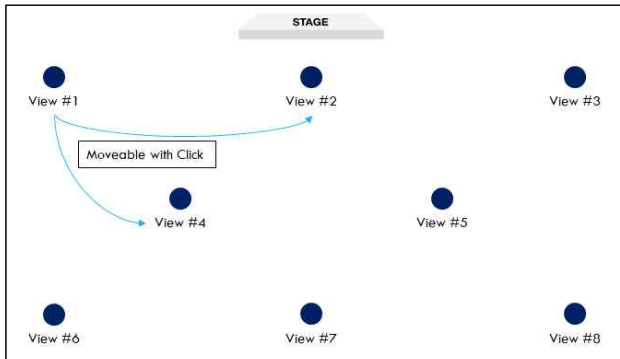


그림 2. Quantized 6DoF 시뮬레이션의 구성 도

실험은 현재 위치의 영상만을 받는 경우(1), 현재 위치와 랜덤으로 몇 개의 위치를 더 받는 경우(2), 이동을 예상하여 일정수준의 정확도로 다음 위치를 받는 경우(3)로 나누어 반복한다. 이때 지연시간은 위치 변경 명령(마우스 클릭 등) 으로 부터 다음 위치 영상 재생(일정량 이상의 버퍼를 채운 이후)순간 까지로 정한다. 이때 정의된 지연시간은 크게 2가지로 나뉘는데 한가지는 이미 버퍼가 있는 위치로 이동했을 경우 생기는 지연시간, 다른 한가지는 버퍼가 충분히 차 있지 않거나 없는 곳으로 이동했을 경우 생기는 지연시간이다. 지연시간은

$$\text{Viewpoint switching delay} = \text{processing delay} + \text{roundtrip delay} + \text{buffer delay} + \text{decoding delay}$$

로 나타낼 수 있으며 그림3과 같이 상황에 따라 전체 지연시간을 구분 할 수 있다.

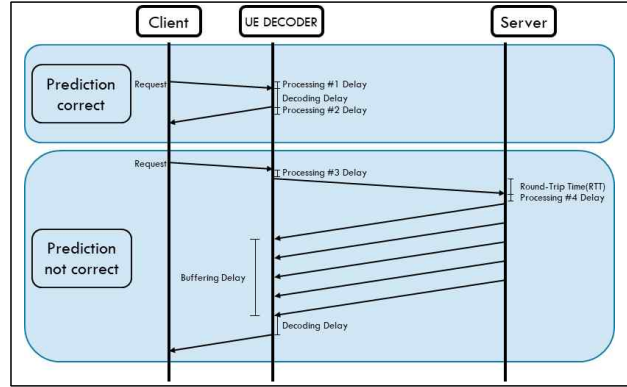


그림 3. 시점이동 시 생기는 Delay 구조 도 또한 일정 수준의 정확도를 가지는 알고리즘은 간단하게 Preload된 시점 한 개당 40%의 정확도로 독립시행을 수행하는 형태로 가정하였다.

$$n < 7, P = 1 - (1 - P_c)^n \quad \text{else } n = 7, P = 1.$$

P: 시점 이동 예측 정확도

P_c: Preload된 영상 당 시점 이동 예측 정확도

이때 현재 데모에서 사용된 영상의 개수가 8개이므로, 현재 위치를 제외하고 Preload된 위치가 7개가 되면 모든 위치를 Preload 하게 되는 것이기 때문에 100%확률이 된다.

그림3에서 보이는 것 중 Processing Delay 및 Decoding Delay는 컴퓨터 사양 및 코드 자체 지연에 의해 결정되며, RTT와 Buffering Delay는 네트워크 환경에 따라 달라진다. 현재 실험을 위한 시스템 구성은 Ryzen 1800X, RTX2080ti, 1Gbps LAN로 구성되어있으며, RTT와 Buffering Delay 조절을 위해 브라우저 개발자 기능을 통하여 네트워크 지연 40ms와 대역폭을 특정 수준으로 고정하여 유선네트워크가 연결되어 있지만, 무선의 환경을 구현했다.

실험에서 사용되는 영상은 각각 1080P 해상도에 6Mbps의 대역폭을 필요로하며, 브라우저의 네트워크 대역은 영상 대역의 정수배 만큼 할당하여 진행하였다. 각 상황별 시점이동은 600회 진행 하였고, 최저값, 최고값 각 5개 씩을 제외한 500개의 시점이동 지연시간의 평균을 기록하였음.

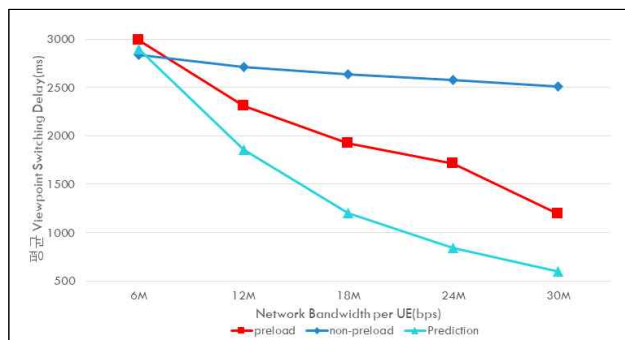


그림 4. 상황 별 평균 시점이동 지연시간

4. 실험 결과 및 논의

그림4와 같이 1번 환경에서는 네트워크 대역이 증가함에 따라 약한

우하향을 보임을 알 수 있다. 그리고 2번 환경에서는 상대적으로 가파르게 선형감소함을 알 수 있으며, 마지막 3번에서는 지수함수와 같은 형태로 지연시간이 감소한다. 평균 지연만을 놓고 보았을 때 1초 이하의 지연을 보이기 위해서는 예측없는 단순 Preload의 경우 4개이상, 예측을 했을 경우 Preload의 경우 2개 이상의 추가 시점을 같이 다운로드 해야 한다.

본 실험은 8개라는 적은 수의 시점을 가지지고 있고, Synthesis View 합성도 이루어지지 않아 실제 RGBD를 기반으로 한 자유시점 콘텐츠의 지연시간과는 차이가 있을 것으로 예상 가능하다. 또한 그래프 내에는 전체 평균 시점이동 지연이 기록되었지만, 실제 Preload된 위치로 이동할 경우 200ms 대 지연을 보여 Preload 되지 않았을 경우 대비 10배 이상의 Latency 감소 효과를 볼 수 있음을 확인 했다.

현재 실험에서 가상적으로 제시한 위치 이동 정확도 모델은 4개 이상의 시점이 Preload 될 때 약 88% 이상의 정확도를 가지고, 현재 실험 모델에서 평균 800ms 이내의 지연시간을 가지게 된다. RGBD의 경우 공간을 돌아다니기 위해 최소 2개 이상의 영상을 다운로드 하여야 하므로, 3개 이상의 영상이 프리로드 되는 구간에서 80% 이상의 정확도를 가지게 될 경우 시점이동 후 생기는 Hole의 복원이 신속하게 이뤄져 위화감이 빠르게 사라질 것으로 예상된다.

5. 결론

본 논문에서는 5G 네트워크 환경에서 자유시점을 보낼 경우 생기는 대역 문제와 이를 해결하기 위하여 이동패턴 예측을 통한 저지연의 필요성에 대하여 논의 했다. 실제로 자유시점에 대한 실험을 구현하거나 환경 구축을 하지 못했지만, Quantized 6DoF로 자유시점과 비슷한 환경을 구현했으며, 이때 시점이동 지연시간에 대한 실험을 진행했다. 실제 RGBD를 활용한 Synthesis View 합성을 통한 6DoF에서는 다중 영상 렌더링 및 합성으로 지금과 같은 환경에서 보다 지연 요소가 더 많아지지만, 반대로 새로운 View의 영상을 다운로드 받지 못하더라도 Hole이 존재하는 채로 생성된 이미지를 제공할 수 있기 때문에 지연에서 조금 더 자유로울 수 있을 것으로 예상된다. 앞으로 실제 RV[4]를 활용한 6DoF 환경 구현 및 단일 이미지로 생성된 다른 위치의 이미지가 사용자 경험에 있어서 어느정도의 영향을 끼치는지와 종합적인 판단을 통한 목표 시점이동 지연을 확인하고 이를 위한 예측 알고리즘을 연구 할 계획이다.

Acknowledgment

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2016-0-00144, 시청자 이동형 자유시점 360VR 실감미디어 제공을 위한 시스템 설계 및 기반기술 연구)

6. 참고 문헌

[1] "MPEG-I Use Cases for omnidirectional 6DoF, windowed 6DoF, and 6DoF," ISO/IEC JTC1/SC29/WG11, w16768, April 2017.

- [2] <http://www.immersiveshooter.com/2017/08/31/explaining-360-video-resolution-how-measure-quality-comparisons/>
- [3] Shafi, Mansoor, et al. "5G: A tutorial overview of standards, trials, challenges, deployment, and practice." IEEE journal on selected areas in communications 35.6 (2017): 1201-1221.
- [4] <https://www.gsmathrive.com/shenzhen/2020/08/12/cloud-ar-vr-whitepaper/>
- [5] Shafi, Mansoor, et al. "5G: A tutorial overview of standards, trials, challenges, deployment, and practice." IEEE journal on selected areas in communications 35.6 (2017): 1201-1221.
- [6] Hannuksela, Miska M., Ye-Kui Wang, and Ari Hourunranta. "An overview of the OMAF standard for 360 video." 2019 Data Compression Conference (DCC). IEEE, 2019.
- [7] Niamut, Omar A., et al. "MPEG DASH SRD: spatial relationship description." Proceedings of the 7th International Conference on Multimedia Systems. 2016.
- [8] Hosseini, Mohammad, and Viswanathan Swaminathan. "Adaptive 360 VR video streaming: Divide and conquer." 2016 IEEE International Symposium on Multimedia (ISM). IEEE, 2016.
- [9] Corbillon, Xavier, et al. "Viewport-adaptive navigable 360-degree video delivery." 2017 IEEE international conference on communications (ICC). IEEE, 2017.
- [10] Schwarz, Sebastian, et al. "Emerging MPEG standards for point cloud compression." IEEE Journal on Emerging and Selected Topics in Circuits and Systems 9.1 (2018): 133-148.
- [11] <https://support.google.com/youtube/answer/78358>
- [12] <https://aframe.io/>
- [13] <https://github.com/Dash-Industry-Forum/dash.js/wiki>