

2D 보간법 기반의 DASH 포인트 클라우드 스트리밍 시스템

조하성¹, 고윤민², 송황준^{1*}

포항공과대학교 컴퓨터공학과¹, 한동대학교²

lifeofcho@postech.ac.kr, yunmin@handong.ac.kr, *hwangjun@postech.ac.kr

2D Interpolation-based Enhancement for MPEG-DASH Point Cloud Streaming System

Hasung Cho¹, Yunmin Go², Hwangjun Song^{1*}

Dept. of Computer Science, POSTECH¹, Handong Global Univ²

요약

본 논문은 MPEG-DASH 기반의 적응적 포인트 클라우드 스트리밍 환경에서 실시간성을 보장하고 품질을 향상시키는 시스템을 제안한다. 제안하는 시스템은 포인트 클라우드를 여러 품질로 스트리밍 함으로써 초기 지연시간을 줄인다. 또한 수신한 포인트 클라우드의 대해 최적의 포인트 사이즈 결정 및 2D 보간법 기반의 post-processing 을 통해 시각적 품질을 향상시킨다. 제안하는 시스템은 실제 스트리밍 환경 구축을 통해 실험하였고, 시각적 품질 향상시킴을 검증하였다.

I. 서론

최근 애플 비전 프로의 등장을 계기로 메타 버스 및 XR(Extended Reality)로 대표되는 실감형 콘텐츠 산업이 크게 성장하고 있다. 이러한 콘텐츠는 사용자에게 높은 몰입감과 현실과 같은 가상 경험을 제공하기 위해 3 차원 공간을 사실적으로 표현하는 기술을 필요로 하며, 포인트 클라우드는 그 핵심 기술 중 하나로 주목받고 있다.

그러나 고품질 포인트 클라우드는 수십만에서 수백만 개의 포인트로 구성되어 데이터 용량이 상당히 크다. 이로 인해 스트리밍 시 발생하는 높은 초기 지연시간은 사용자의 체감품질을 저하한다.

이러한 문제를 해결하기 위해 MPEG-DASH[1]와 같은 적응형 스트리밍 기술이 사용된다. MPEG-DASH는 네트워크 상태에 따라 다양한 품질의 포인트 클라우드를 인코딩하여 전송하지만, 대역폭이 낮은 환경에서는 저품질, 즉 밀도가 낮은 포인트 클라우드를 전송함으로써 사용자의 시각적 품질을 저하한다. 이를 보완하기 위해 클라이언트 측에서 업샘플링을 통해 품질을 높이는 기법들이 있으나, 대부분의 기법은 3 차원 공간에서 직접 연산을 수행함으로 실시간 스트리밍 환경에 적용하기에는 계산 복잡도가 높다는 한계가 있다[2].

따라서 제안하는 시스템은 MPEG-DASH 기반의 적응적 포인트 클라우드 스트리밍 환경에서 실시간성을 보장하며 시각적 품질을 향상시키는 기법을 제안한다. 제안하는 기법은 수신된 저품질 포인트 클라우드에 2D 보간법을 적용하여, 낮은 지연시간과 높은 체감 품질을 동시에 달성하는 것을 목표로 한다.

II. 본론

그림 1 은 제안하는 시스템의 구조를 나타낸다. DASH 서버는 원본 포인트 클라우드를 여러 LoD (Level of Details) 로 표현하기 위해, 먼저 Octree 구조를 활용하여 전체 공간을 복셀(Voxel) 단위로 분할한다. 그 후, 각 복셀 내에 포함된 포인트들을 균등하게

level 단위로 그룹화한다. 최종적으로 생성된 level 들을 수식 1 과 같이 총 10 개의 LoD 로 구성 후 저장한다.

$$LoD_i = LoD_{i-1} + level_i \quad (1)$$

포인트 클라우드는 설정된 포인트 사이즈에 따라 렌더링 될 수 있는 데, 이때 포인트 사이즈는 사용자의 체감 품질에 큰 영향을 미칠 수 있다. 그림 2 에서는 같은 LoD 에서 서로 다른 포인트 사이즈가 가지는 시각적 품질을 나타낸다. 따라서 Optimization module 에서는 현 LoD 에서 수식 2 를 최소화하는 최적의 포인트 사이즈 결정 후 MPD 에 URL 과 함께 저장한다.

$$\alpha \cdot area_{hole} + (1-\alpha) \cdot area_{overlap} \quad (2)$$

α 는 0.5, $area_{hole}$ 과 $area_{overlap}$ 는 특정 포인트 사이즈로 렌더링 했을 시 포인트들 사이로 비어 보이는 area 와 여러 포인트가 동일한 픽셀을 덮어쓰는 area 의 픽셀 단위 면적을 각각 의미한다. Overlap area 와 hole area 는 모든 LoD 의 대해 서버에 사전에 계산되어 저장된다.

DASH 클라이언트는 network monitor, MPD parser, decoder, post-processing module 를 포함하고 back-point culling 과 포인트 사이즈 설정을 수행한다.

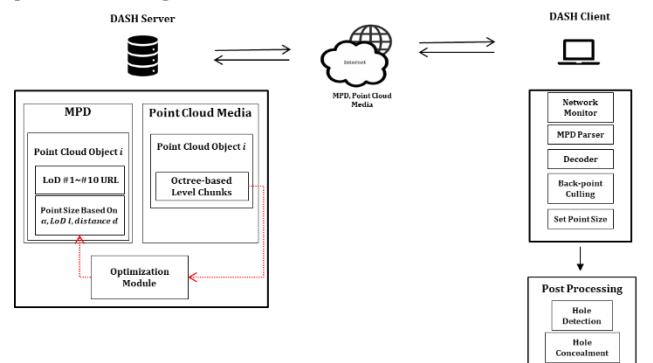


그림 1. 제안하는 시스템 구조



(a) 포인트 사이즈 3, (b) 포인트 사이즈 12
그림 2. 포인트 사이즈에 따른 품질 차이.

MPD parser는 수신한 MPD를 분석한다. Network monitor는 가용가능한 대역폭을 주기적으로 추정하고 현재 네트워크 상태를 고려하여 요청할 LoD를 결정하고 DASH 서버에 요청한다. 요청한 LoD가 도착하면 back-point culling과 MPD에서 제공한 최적 포인트 사이즈를 적용한다. Back-point culling은 viewport를 기준으로 포인트 클라우드의 후면 포인트들을 제거하는 과정이다. 낮은 LoD 즉 저품질 포인트 클라우드는 렌더링 되었을 때 후면 포인트들이 전면에서 보일 수 있다. 이러한 포인트들은 사용자의 체감품질을 저하시키기 때문에 디스플레이되어 지기 전에 제거한다.

후면 포인트들이 제거된 후 post-processing 모듈을 적용한다. 최적 포인트 사이즈를 적용했을 시, hole area 뿐만 아니라 overlap area 또한 고려하기 때문에 렌더링 된 포인트 클라우드에는 hole이 발생할 수 있고 이러한 hole들은 시각적 품질을 저하할 수 있다. Post-processing 모듈에서는 발생하는 hole 픽셀들을 검출하고, 검출한 hole에 대해서 hole이 아닌 이웃 픽셀들을 활용한 2D 보간법을 통해 hole concealment를 수행한다. 이 때 post-processing 모듈의 처리시간은 평균적으로 60ms가 소요되기 때문에 실시간성이 보장된다.

III. 실험 결과

실제 스트리밍 환경을 구축하여 제안하는 시스템 성능을 분석하였다. DASH 클라이언트는 WebGL 기반 Three.js 라이브러리를 통해 웹 브라우저에 렌더링한다. DASH 서버는 NGINX, Point Cloud Library (PCL) 기반으로 구축하였고 USB 무선 랜 카드(802.11ac)를 사용하여 와이파이와 연결한다. 실험 시 사용한 포인트 클라우드는 Long Dress[2]다.

실험은 LoD2 저품질 포인트 클라우드를 스트리밍 받은 경우 2D 보간법을 통한 post-processing을 적용한 것과 적용하지 않은 결과를 포인트 클라우드 전면 기준으로 비교한다.

그림 3은 SSIM값을 사용하여 비교한 결과를 나타낸다. LoD2인 경우는 밀도가 낮은 포인트 클라우드이기 때문에 포인트 사이즈가 비교적으로 큰 8일 때 가장 높은 SSIM값인 0.634를 기록했다. Post-processing module을 통해 hole detection과 hole concealment를 적용한 결과는 0.734이고

그림 4는 시각적 화질 비교를 나타낸다. Post-processing을 적용하지 않은 경우는 후면 포인트들이 전면에서 보이면서 시각적 품질을 낮춘다. 결론적으로 객관적 수치와 주관적 시각 품질을 모두 향상시키는 것을 확인할 수 있다.

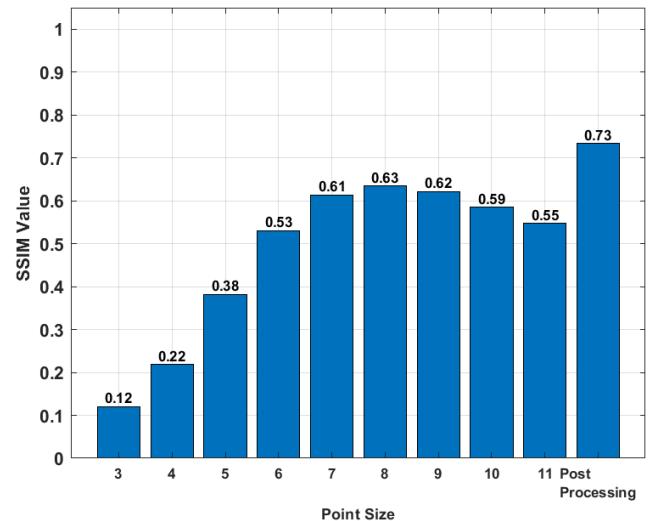
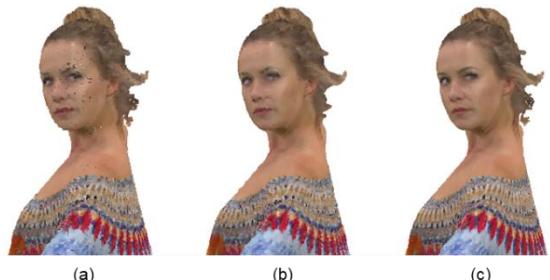


그림 3. LoD2 기준 SSIM 값 비교



(a) Post processing 적용 X, (b) Post-processing 적용, (c) 원본.
그림 4. 시각적 화질 비교

IV. 결론

본 논문에서는 MPEG-DASH 기반 적응형 스트리밍 기술을 사용하여 데이터 용량이 큰 원본 포인트 클라우드를 효과적으로 여러 LoD로 구성하였고, 이에 따라 최적의 포인트 사이즈 및 post-processing을 적용함으로써 저품질 포인트 클라우드일 때 시각적 품질을 높였다. 실제 스트리밍 환경을 구축하여 성능 검증을 완료하였고, 향후 포인트 클라우드 비디오에도 적용하는 연구를 진행할 것이다.

참고 문헌

- [1] T. Stockhammer, "Dynamic Adaptive Streaming over HTTP – Standards and Design Principles", in Proceedings of ACM MMSys, pp. 133-144, 2011
- [2] LI, Ruihui, et al. Pu-gan: a point cloud upsampling adversarial network. In: Proceedings of the IEEE/CVF international conference on computer vision. 2019. p. 7203-7212.
- [3] JPEG Pleno Database: 8i Voxelized Full Bodies (8iVFBv2) – A Dynamic Voxelized Point Cloud Dataset. (n.d.). <http://plenodb.jpeg.org/pc/8ilabs/>