

2D super resolution network를 이용한 Point Cloud 데이터 개선

*박성환 **김규현

경희대학교

* gocheenee@khu.ac.kr ** kyuheonkim@khu.ac.kr

Improvement of point cloud data using 2D super resolution network

*Park, Seong-Hwan **Kim, Kyu-Heon

KyungHee University

요약

미디어 기술은 사용자가 더욱 몰입감을 느낄 수 있는 방향으로 개발되어 왔다. 이러한 흐름에 따라 기존의 2D 이미지에 비해 깊이감을 느낄 수 있는 증강 현실, 가상 현실 등 3D 공간 데이터를 활용하는 미디어가 주목을 받고 있다. 포인트 클라우드의 수많은 3차원 좌표를 가진 여러 개의 점들로 구성된 데이터 형식이므로 각각의 점들에 대한 좌표 및 색상 정보를 사용하여 3D 미디어를 표현한다. 고정된 크기의 해상도를 갖는 2D 이미지와 다르게 포인트 클라우드는 포인트의 개수에 따라 용량이 유동적이며, 이를 기존의 비디오 코덱을 사용하여 압축하기 위해 국제 표준기구인 MPEG(Moving Picture Experts Group)에서는 Video-based Point Cloud Compression (V-PCC)을 제정하였다. V-PCC는 3D 포인트 클라우드 데이터를 직교 평면 벡터를 이용하여 2D 패치로 분해하고 이러한 패치를 2D 이미지에 배치한 다음 기존의 2D 비디오 코덱을 사용하여 압축한다. 본 논문에서는 앞서 설명한 2D 패치 이미지에 super resolution network를 적용함으로써 3D 포인트 클라우드의 성능 향상하는 방안을 제안한다.

1. 서론

그동안의 미디어 기술은 사용자가 더욱 몰입할 수 있는 콘텐츠를 제공하기 위해 개발되어 왔다. 또한 콘텐츠의 소비 범위를 확장하기 위하여 전송 기술이 개발되었으며 전송을 위한 압축 기술도 함께 발전해 왔다. 또한 최근에는 딥러닝 네트워크를 이용하여 미디어 분석 및 재생에 대한 연구들이 활발히 진행되고 있다.

몰입형 미디어는 증강 현실 (AR) 및 가상 현실 (VR) [1] [2]과 같이 3D 형식으로 개발되었으며 Voxel, Mesh 및 포인트 클라우드 [3]와 같은 3D 데이터 형식을 사용하여 표현된다. 이러한 3D 데이터 형식 중 포인트 클라우드는 객체를 색상 및 좌표 정보를 가진 여러개의 점들로 표현한다. 해상도에 따라 크기가 고정된 2D 이미지와 달리 포인트 클라우드 데이터의 크기는 포인트 수에 따라 가변적이다. 포인트 클라우드의 크기로 인해 전송을 통해 포인트 클라우드 데이터를 소비하는 데 어려움이 있어 포인트 클라우드 데이터의 압축에 대한 필요성이 제기되었다. 국제 표준화 그룹인 MPEG (Moving Picture Experts Group)은 3D 포인트 클라우드 데이터 압축 표준 기술 개발을 위하여 MPEG-I (Immersive) 차세대 프로젝트 그룹을 구성했다. MPEG-I에서는 비디오 기반 포인트 클라우드 압축을 (V-PCC)[4] 제안하였으며 해당 문서의 Part 5는 기존 2D 비디오 코덱을 사용하여 포인트 클라우드를 압축하는 방안에 대하여 설명하고 있다. V-PCC에서 3D 포인트 클라우드 데이터는 2D 이미지 패치[5]로 나뉘어 직교 방향을 따라 2D 이미지 [6][7]상에 배치되어 2D 비디오 코덱을 사용하여 압축된다 [8]. V-PCC는 3D 콘텐

츠를 2D 이미지 시퀀스 및 손실 비디오 코덱으로 변환하기 때문에 정보 손실이 불가피하며 재구성된 3D 포인트 클라우드의 품질 저하를 유발할 수 있다. 따라서 3D 포인트 클라우드의 시각적 품질을 향상시키기 위해 본 논문에서는 3D 포인트 클라우드의 재구성 과정에서 생성되는 2D 이미지를 super resolution network를 사용하여 향상하는 방안을 제안한다. 딥러닝 기반의 super resolution network는 기존의 기존 방식에 비해 주관적인 시각적 품질에서 더 높은 성능을 보여주고 있으며 다양한 네트워크가 여전히 연구되고 있다. 본 논문에서는 2장에서는 포인트 클라우드 데이터 향상에 사용되는 기본 기술을 설명하고, 3장에서는 본 논문에서 제안한 super resolution network를 이용한 포인트 클라우드 향상 기술의 과정을 설명하며 실험 결과를 주관적이고 객관적인 품질 측면에서 제시 및 분석하며, 마지막으로 향후 계획 및 결론을 4장에서 제안한다.

2. 배경 기술

Video-based point cloud compression

V-PCC는 기존 2D 비디오 코덱을 사용하여 3D 포인트 클라우드를 압축하는 방법이다. V-PCC의 인코딩 과정은 그림 1과 같으며 포인트 클라우드 데이터를 2D 이미지로 변환하기 위해 patch를 생성하는 patch generation으로 시작된다. Patch generation의 경우 비슷한 법선 벡터를 가진 포인트들을 XY, XZ 및 YZ 평면에 평행한 6개의 참조

평면을 기반으로 그룹화한다. Patch Packing 단계에서는 각 투영 평면에 투영되는 패치를 2D 그리드에 배치하며 각 패치가 2D 그리드에서 점유 여부를 나타내는 Occupancy map을 생성하며 각 patch가 어떤 평면에 배치 되었는지를 알 수 있는 정보인 Patch auxiliary information을 생성한다. 2D 그리드에 투영된 패치는 좌표 기반 Geometry와 색상 기반 Attribute 등 두 가지 정보로 구성되며 각각의 Geometry, Attribute 정보는 2D 이미지로 표현된다. 생성된 Occupancy, Geometry, Attribute map은 각각 2D 비디오 코덱을 사용하여 압축되어 Compressed bitstream을 획득할 수 있다.

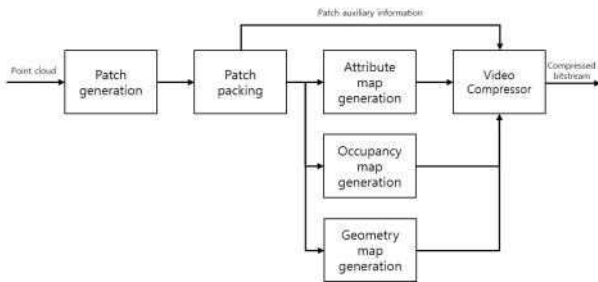


그림 1. V-PCC의 인코딩 과정

V-PCC의 디코딩 프로세스는 그림 2에서 나타난 것처럼 앞서 설명한 인코딩 과정의 역순으로 진행된다. Occupancy map, Geometry map, Attribute map은 2D 비디오 코덱을 사용하여 디코딩되고 디코딩된 map들과 Patch auxiliary information을 이용하여 3D 포인트 클라우드가 재구성된다.

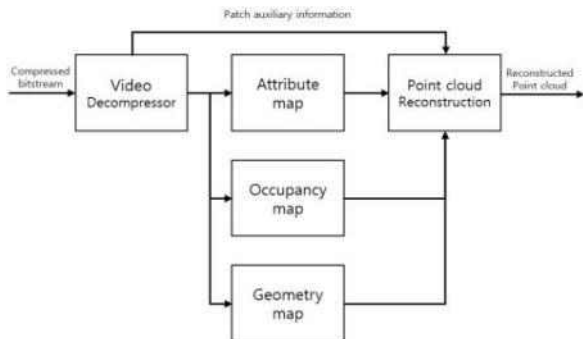


그림 2. V-PCC의 디코딩 과정

2D super resolution networks

최초의 신경망 기반의 super resolution network는 Image Super-Resolution using deep Convolutional Networks (SRCNN)[9]이며 SRCNN은 전통적인 super resolution 기술들에 비해 객관적인 성능을 향상시키기 위해 제안되었다. 이후 super resolution network에 대한 연구가 활발히 진행되었으며 이후 발표된 accurate image Super Resolution using Very Deep convolutional networks(VDSR)에서는[10] residual learning을 사용하여 깊은 layer를 쌓았고 높은 learning rate를 이용하여 SRCNN의 문제를 개선하였다. 또한, VDSR은 깊은 레이어를 이용하여 넓은 범위의 contextual 정보를 학습할 수 있다는 장점이 있다. 이러한 분석을 바탕으로 VDSR이 포인트 클라우드의 성능 향상에 사용될 super resolution network로써 가장 적절할 것으로 판단하여 VDSR을 적용하여 포인트 클라우드 품질을 개선하였다.

3. 2D SR network를 이용한 Point cloud 품질 개선

본 논문에서는 앞서 설명한 2D 패치 이미지의 품질을 개선하기 위해 2장에서 소개한 VDSR을 적용하였다. Super resolution에서는 업스케일링 배수에 따라 출력 이미지의 해상도가 결정된다. 예를 들어 업스케일링 배수가 2이고 입력 이미지가 1280x1280 인 경우 출력 이미지는 2560x2560이다. 2 장에서 설명한 것처럼 3D 포인트 클라우드를 2D 패치 이미지로 변환하면 Attribute map, Geometry map, Occupancy map 등 3 개의 이미지가 생성되는데 3D 포인트 클라우드를 재구성 할 때 Attribute map과 Geometry map은 Occupancy map을 참조하므로 세 이미지의 해상도가 동일해야 한다. 결과적으로 업스케일링 배수는 Attribute map뿐만 아니라 Geometry map과 Occupancy map에도 동일하게 적용되어야 한다. 따라서 Geometry 및 Attribute map의 해상도를 동일하게 유지하기 때문에 업 스케일링 배수는 1로 설정하여 실험을 진행하였다.

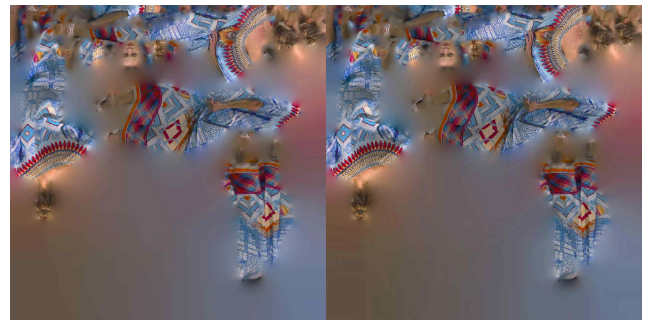


그림 3. 원본(좌)과 super resolution이 적용된 Attribute map(우)

2D 패치의 Attribute map에 super resolution network를 적용한 결과는 그림 3과 같으며, 원본 이미지에 비해 주관적인 시각적 품질이 향상되었음을 알 수 있다. 본 논문에서는 3D 포인트 클라우드 데이터를 2D 이미지로 변환하는 V-PCC 과정에서 발생하는 정보 손실을 보상하기 위해 2D super resolution network를 이용한 3D 포인트 클라우드 향상을 제안하였으며 제안된 방법을 검증하기 위해 객관적인 3D PSNR과 주관적인 시각 품질을 모두 측정하였다.

표 1. V-PCC와 제안 방식의 3D PSNR 비교

Sequence		3D PSNR(dB)		End to end PSNR[dB]		
		D1	D2	Y	Cb	Cr
dress	V-PCC	71.44	75.29	36.93	39.89	39.08
	제안방식	71.44	75.29	26.44	35.36	34.51
loot	V-PCC	70.69	74.67	40.07	43.24	39.31
	제안방식	70.69	74.67	28.27	34.74	33.82
soldier	V-PCC	71.24	75.14	41.95	49.50	49.59
	제안방식	71.24	75.14	28.61	45.41	45.06

표 1에서 볼 수 있듯이 PSNR은 3개의 포인트 클라우드 시퀀스에 대해 측정되었으며, 'V-PCC'와 '제안방식'은 각각 V-PCC 과정을 통해 재구성된 포인트 클라우드, 제안된 방식이 적용되어 재구성된 포인트 클라우드를 원본 포인트 클라우드와 비교하여 측정된 PSNR 값을 나타낸다. Super resolution network는 Geometry map에는 적용되지 않았으므로 Geometry PSNR은 변경되지 않았다. 3D PSNR을 제외하고 나머지 PSNR의 경우 super resolution network가 적용되지 않은 경우보다 대체적으로 낮은 것으로 나타났다.



그림 4. 포인트 클라우드 원본 및 제안방식 적용 결과

그러나 재구성된 포인트 클라우드를 주관적 품질 측면에서 비교했을 때 본 논문에서 제안한 방식을 적용할 경우 텍스처의 품질 향상이 나타난 것을 그림 4를 통해 확인할 수 있었다. 이러한 객관적이고 주관적인 결과를 바탕으로 제안된 2D super resolution network를 이용한 포인트 클라우드 품질 개선 방법은 주관적인 시각적 품질을 향상시킬 수 있지만 원래의 포인트 클라우드에 가까운 속성 값을 얻는 데 어려움이 있음을 알 수 있었다. 이는 신경망이 유사한 패턴을 기반으로 값을 생성하는 속성을 가지고 있어 더 나은 시각적 품질을 제공하지만 원본과는 다른 색상 값을 출력하기 때문인 것으로 생각된다.

4. 결론

최근에는 사용자에게 몰입감 있는 콘텐츠를 제공하기 위하여 AR, VR 등 3D 공간 데이터를 이용한 미디어 데이터가 주목받고 있다. 포인트 클라우드는 각 포인트에 대한 3D 좌표와 색상 정보를 이용하여 3D 미디어 데이터를 표현하기 위한 데이터 형식이다. 포인트 클라우드는 2D 이미지보다 용량이 크기 때문에 이를 압축하는 기술이 요구되었으며 국제 표준기구인 MPEG은 V-PCC를 개발했다. V-PCC는 3D 포인트 클라우드 데이터를 패치로 분해하고 패치를 2D 이미지로 투영한 후 기존 2D 비디오 코덱을 사용하여 압축하는 기술이다. 이러한 과정으로 인해 3D 포인트 클라우드를 2D 이미지로 변환하고 기존 비디오 코덱을 사용하여 2D 이미지를 인코딩하는 동안 데이터 손실이 발생할 수 있으며 이러한 데이터 손실로 인해 재구성된 포인트 클라우드의 품질이 저하될 수 있다. 이에 본 논문에서는 포인트 클라우드 패치의 2D 영상에 super resolution network를 적용하여 포인트 클라우드 데이터를 향상시키는 방법을 제안 하였다. Attribute map에 super resolution network를 적용하면 PSNR의 객관적인 결과는 낮았지만 주관적인 시각적 품질은 향상되는 것을 확인할 수 있었다.

This work was supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No.2020-0-00452, Development of Adaptive Viewer-centric Point cloud AR/VR(AVPA) Streaming Platform)

참고 문헌

- [1] "Ubiquitous AR to dominate focused VR by 2022," <https://techcrunch.com/2018/01/25/ubiquitous-ar-to-dominatefocused-vr-by-2022/>, accessed: 2019.
- [2] "Creating 3D Content with Reality Composer," https://developer.apple.com/documentation/realitykit/creating_3d_content_with_reality_composer, accessed: 2019.
- [3] Emin Zerman, Cagri Ozcinar, Pan Gao, and Aljosa Smolic, "Textured Mesh vs Coloured Point Cloud: A Subjective Study for Volumetric Video Compression", 2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)
- [4] "Text of ISO/IEC CD 23090-5 Video-based Point Cloud Compression," ISO/IEC JTC1/SC29/WG11 MPEG2019/N18670, Gothenburg, Sweden, Oct. 2019.
- [5] H. Hoppe, T. DeRose, T. Duchamp, J. A. McDonald, W. Stuetzle, "Surface reconstruction from unorganized points," Proc. SIGGRAPH, pp. 71-78, 1992.
- [6] Schwarz Sebastian, Miska M. Hannuksela, FakourSevom Vida, Sheikhi-Pour Nahid, "2d video coding of volumetric video data", Picture Coding Symposium (PCS) 2018 - Proceedings, pp. 61-65, 2018.
- [7] S. Schwarz, et al., "Emerging mpeg standards for point cloud compression," in IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 9, no. 1, pp. 133-148, Mar. 2019.
- [8] "High efficiency video coding test model, H M 16 . 1 8 + S C M 8 . 7 , " https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.18+SCM-8.7/, accessed: 2019.
- [9] "Image Super-Resolution Using Deep Convolutional Networks", Chao Dong, Chen Change Loy, Kaiming He, Xiaoou Tang, ECCV, 2014
- [10] "Accurate Image Super-Resolution Using Very Deep Convolutional Networks", Jiwon Kim, Jung Kwon Lee and Kyoung Mu Lee, CVPR, 2016