

3D DCT 를 이용한 포인트 클라우드의 움직임 예측/보상 기법

이민석, 김보연, 윤상은, 황용해, 김준식, 김규현

경희대학교

qsibmini123@gmail.com, qhdus0901@gmail.com, dbstkd8980@gmail.com,
hyh717@khu.ac.kr, junsik@khu.ac.kr, kyuheonkim@khu.ac.kr

3D Motion Estimation and Compensation method for Point cloud video codec by 3D DCT

Minseok Lee, Boyeun Kim, Sangeun Yoon, Yonghae Hwang, Junsik Kim,
Khuheon Kim
Kyunghee University

요 약

포인트 클라우드는 3 차원 물체를 표현하기 위한 점들의 집합으로, 동적인 3 차원 데이터를 정밀하게 획득할 수 있기
이의 효율적인 압축의 필요성이 대두되고 있다. 기존 3D DCT(3D Discrete Cosine Transform)를 이용한 동적 객체의
포인트 클라우드 압축 방식은 Inter 프레임 압축을 고려하지 않아 압축시의 데이터 압축률에 한계가 있다. 따라서 본
논문은 이러한 문제점을 개선하기 위해 3D DCT 를 이용한 움직임 예측을 통하여 포인트 클라우드 영상의 I 프레임 및 P
프레임을 압축하는 방식을 제안한다.

1. 서론

포인트 클라우드는 3차원 물체를 표현하기 위한 점들의 집합으로, 동적인 객체의 데이터를 정밀하게 획득하여 표현할 수 있다는 장점이 있다. 하지만 포인트 클라우드는 기존의 2차원 데이터와는 달리 객체의 위치와 색상을 모두 표현하여야 하므로, 방대한 양의 데이터가 필요하다. 따라서, 이러한 대용량을 갖는 3D 포인트 클라우드의 효율적인 압축 방법에 대한 필요성이 대두되고 있다.

3D 동적 포인트 클라우드 객체를 압축하기 위해 2D 이미지 코덱인 JPEG에서 2D DCT(Discrete Cosine Transform) 모듈을 3D DCT 모듈로 대체하여 3D 포인트 클라우드의 압축을 진행한 방식이 연구된 바 있다. JPEG에서의 2D DCT 모듈은 8*8 매크로블럭 단위를 사용한다면, 3D DCT 모듈은 8*8*8 매크로큐브 단위로 DCT를 진행한다. 이때, 3D 포인트 클라우드는 빈공간이 존재하므로, 매크로큐브에 존재하는 빈 공간을 인접한 점들의 평균값으로 채우는 패딩(Padding) 과정을 거친다. [1][2][3]

이러한 방식은 3D DCT를 이용하여 포인트 클라우드 압축

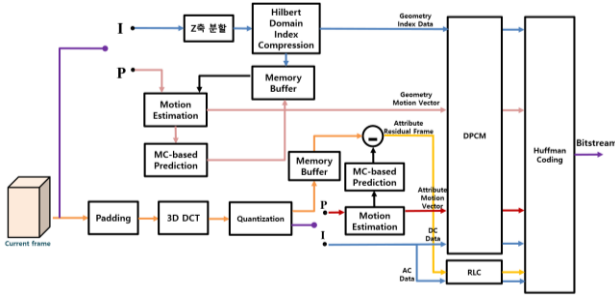
을 시도했다는 점에서 의미가 있지만, 여러 장으로 이루어진 동적 데이터의 압축을 고려하지 않는 Intra 프레임 기반의 압축 방식이라는 한계가 있다.

이러한 한계를 극복하고자 본 논문은 프레임 간의 움직임 예측을 통해 Inter 프레임을 생성하여, 여러 장의 영상 데이터를 보다 효율적으로 압축하는 3D 포인트 클라우드 동영상 압축 방안을 제안한다. [4]

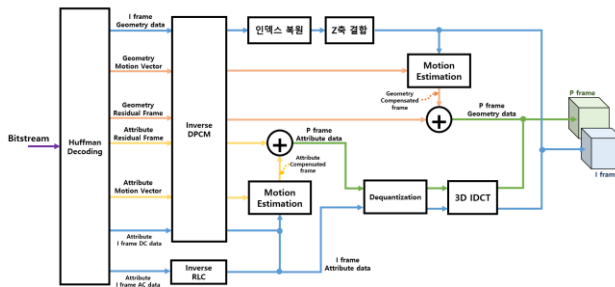
본 논문의 구성은 다음과 같다. 2장에서는 본 논문이 제안하는 3D DCT를 사용한 포인트 클라우드 영상 압축 방식을 자세히 정리한다. 3장에서는 이러한 압축 방식의 효율을 실험을 통해 확인하고, 4장에서 본 논문에 대한 결론을 맺는다.

2. 3D DCT를 이용한 포인트 클라우드의 움직임 예측/보상 기법

본 논문은 포인트 클라우드의 보다 효율적인 압축을 위하여 포인트 클라우드의 색상 정보와 위치 정보를 압축하는 방법들을 각각 제시한다.



<Figure 1> 제안 기술의 부호화 과정 구조도



<Figure 2> 제안 기술의 복호화 과정 구조도

<Figure 1>, <Figure 2>은 위치 정보와 색상 정보를 부/복호화하는 과정을 나타낸 그림이며, 전체적인 압축 과정은 다음과 같다. 본 논문의 제안 기술은 I 프레임 압축부와 P 프레임 압축부로 나눌 수 있다. 입력 프레임이 I 프레임인 경우에는 색상 정보를 패딩과 3D DCT, 양자화(Quantization)를 거쳐 저장하며, 이는 서론에서 언급한 기존의 3D DCT 기반의 압축 방식과 동일하다. P 프레임의 압축은 3D DCT 매크로 큐브 기반의 움직임 예측을 통해 수행된다. 위치 정보는 Z 좌표마다 XY 평면과 평행하게 자르는 Z 축 분할을 수행한 후 공간 채움 곡선 알고리즘을 통해 획득된 인덱스를 통해 부호화된다. 입력 프레임이 P 프레임이면 이전 프레임을 참조하여 색상 정보와 위치 정보에 대한 움직임 예측 및 보상을 하여 데이터를 부호화한다. 복호화 과정은 부호화 과정의 역순이며, 제안 기술의 상세한 압축 방식은 다음 절을 통해 설명한다.

가. 색상 정보 압축 및 복원 과정

본 절에서는 본 논문에서 제안하는 3D DCT 매크로 큐브 기반의 움직임 예측 방식을 통해 P 프레임의 색상 정보를 압축 및 복원하는 방안에 대해 설명한다.

앞서 밝힌 바와 같이, 2D 영상과 달리 3D 영상은 빈 공간이 존재한다. 이러한 포인트 클라우드의 특성은 레퍼런스 프레임(이전 프레임)에서 타겟 프레임(현재 프레임)과의 움직임 예측 시 프레임 사이의 유사도를 평가하는데 어려움을 주며, 움직임 예측의 정확도를 크게 떨어뜨린다. 하지만, 패딩과 3D DCT를 거친 후에 움직임 예측을 진행하면 이미 큐브 내의 모든 점이 채워져 있기 때문에 이러한

문제를 해결할 수 있다. 따라서 움직임 예측 후에 DCT를 수행하는 기존의 2D 동영상 압축과는 달리, 패딩과 3D DCT를 적용한 뒤 움직임을 예측하는 3D DCT 매크로 큐브 기반 움직임 예측 방안을 제안한다.

본 논문에서 제안하는 3D DCT 매크로 큐브 기반 움직임 예측 방안은 I 프레임 압축 시 활용되는 패딩 후 3D DCT가 적용된 매크로 큐브를 활용하여 수행된다. 이때, 매크로 큐브 사이에서 효과적으로 움직임을 예측하기 위해 레퍼런스 프레임과 타겟 프레임의 매크로 큐브 사이의 MAD(Mean Absolute Difference)를 계산하여 3D 움직임 벡터를 탐색한다. 특히, 본 논문에서는 보다 정확한 움직임 탐색을 위해 WMAD(Weighted Mean Absolute Difference)를 활용하며, WMAD는 수식 (1)과 (2)를 통해 획득된다.

$$W(x, y, z) = \begin{cases} 1, & x + y + z = 0 \\ 0.75, & x + y + z = 1 \\ 0.5, & x + y + z > 1 \end{cases} \quad (1)$$

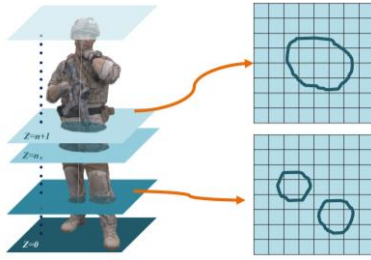
$$WMAD(A, B) = \sum_{x=0}^7 \sum_{y=0}^7 \sum_{z=0}^7 W(x, y, z) |A(x, y, z) - B(x, y, z)| \quad (2)$$

수식 (1)은 WMAD 계산 시 사용되는 가중치 행렬 $W(x, y, z)$ 의 수식이다. WMAD는 수식 (2)와 같이, 레퍼런스 프레임의 매크로 큐브 $A(x, y, z)$ 와 타겟 프레임의 매크로 큐브 $B(x, y, z)$ 의 차이에 가중치 행렬 $W(x, y, z)$ 를 적용하여 획득된다. WMAD 값이 0인 매크로 큐브가 레퍼런스 프레임에서 탐색되면, 해당 큐브의 좌표와 타겟 프레임의 큐브 좌표 사이의 차이 값을 움직임 벡터로서 저장한다. 만약 완전히 일치하는 큐브가 발견되지 않는다면 가장 유사한 큐브의 위치를 움직임 벡터로 저장하며 색상 값에 대한 차이 값은 잔차 프레임으로 저장된다. 이와 같은 과정은 타겟 프레임에 있는 모든 매크로 큐브에 대해 적용되고, 최종적으로 타겟 프레임의 모든 큐브에 대한 움직임 벡터가 저장된다. 그 후 저장된 움직임 벡터들과 잔차 프레임은 DPCM과 허프만 코딩을 거쳐 비트 스트림으로 압축된다. [5] 3D DCT를 이용한 색상 정보 복원 과정은 압축의 역순이며, <Figure 2>에 도시된 바와 같다.

나. 위치 정보 압축 및 복원

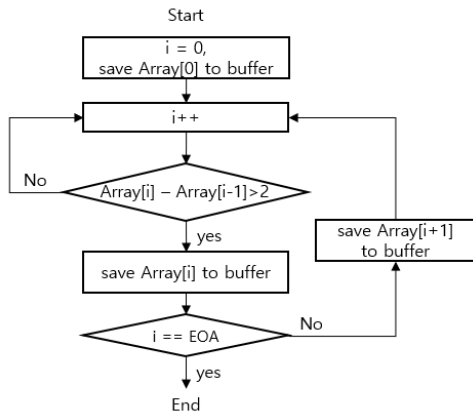
위치 정보의 I 프레임 압축 시에 본 논문이 주목한 점은, 입력 포인트 클라우드는 3차원 데이터이지만, 표면에만 데이터가 존재하며 내부는 비어 있는 형태를 띠고 있는 것이다. 따라서 이를 XY 평면과 평행하게 자른 단면은 대부분 원의 형태를 나타낸다. 본 논문에서는 이러한 포인트 클라우드의 특성을 기반으로 효율적인 위치 정보 압축을 위해, 포인트 클라우드를 각 Z 좌표마다 XY 평면과 평행

하게 자른 후 그 단면의 위치 정보를 압축하는 방안을 제안한다.



<Figure 3> 포인트 클라우드의 Z 축 분할

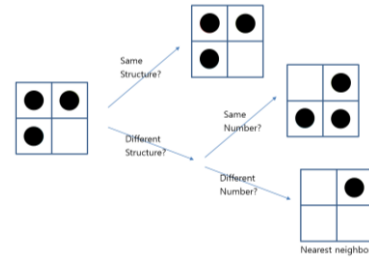
포인트 클라우드의 점의 유무를 0과 1로 모두 표현하게 되면 과도한 데이터양이 필요하므로, 점이 존재하는 위치에서의 인덱스 정보만을 저장한다. 이와 같은 인덱스 정보는 고차원 데이터를 1차원 인덱스화 할 수 있는 공간 채움 곡선 알고리즘을 통해 획득된다. 본 논문에서는 공간 채움 곡선 알고리즘 중 힐버트 커브[6]를 활용하여 인덱스 정보를 획득하였다.



<Figure 4> 인덱스 기반 위치 정보 압축 흐름도

힐버트 커브를 통해 인덱스로 변환된 각 점은 <Figure 4>의 흐름도와 같이 압축된다. 각 점의 인덱스는 순서대로 배열에 저장되며, 현재 배열과 이전 배열의 값의 차이가 2 이상이라면 현재 배열에 저장 되어있는 인덱스 값과 다음 배열의 인덱스 값을 버퍼에 저장한다. 이러한 작업을 연속적으로 진행하면, 연속되는 값들을 그룹 지어 그 첫 인덱스 값과 마지막 인덱스 값 만을 압축하는 효과를 획득할 수 있다. 이와 같은 부호화 방식을 거친 인덱스들은 DPCM과 허프만 코딩 후 비트스트림으로 저장된다.

P 프레임의 위치 정보 압축 과정은 다음과 같다. 우선 전체 포인트 클라우드를 2*2*2 크기의 단위 매크로 큐브로 나눈다. 그 후 움직임 벡터를 찾고자 하는 타겟 프레임 매크로 큐브와 레퍼런스 프레임 내 큐브 사이의 구조를 비교한다.



<Figure 5> 위치 정보 움직임 예측 과정

매크로 큐브들을 비교할 때의 우선 순위는 <Figure 5>와 같다. 타겟 프레임 큐브와 완전히 똑같은 구조의 큐브가 레퍼런스 프레임에서 발견된다면 그 큐브의 좌표와 타겟 프레임 큐브 좌표 사이의 차이 값을 움직임 벡터로서 저장한다. 만약 그러한 큐브가 존재하지 않으면 큐브 내에 존재하는 점의 개수가 동일한 타겟 프레임 큐브를 저장한다. 만일 점의 개수가 타겟 프레임 큐브와 같은 큐브 또한 존재하지 않는다면, 가장 가까운 큐브를 임의로 움직임 벡터로 저장한다. 이와 같은 과정은 타겟 프레임에 있는 모든 매크로 큐브에 대해 적용되고, 최종적으로 타겟 프레임의 모든 큐브에 대한 움직임 벡터가 저장된다. 그 후 움직임 벡터는 DPCM과 허프만 코딩을 거쳐 비트 스트림으로 압축된다. 위치 정보를 복원할 때에는 부호화 과정에서 선정된 매크로 큐브들이 사용되며, 이 과정에서 원본과는 다른 위치값이 생성된다.

3. 실험 결과

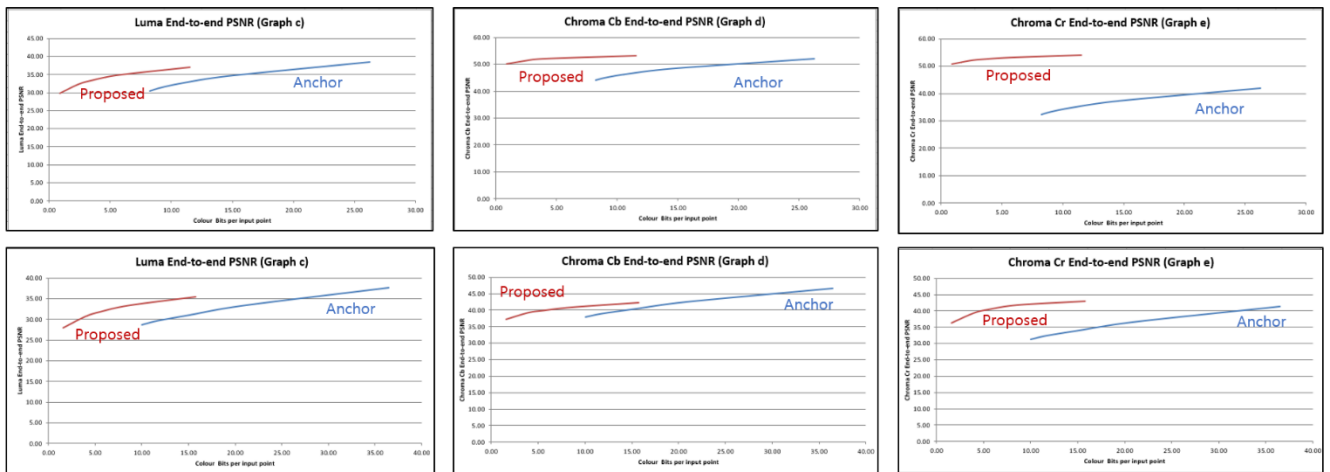
본 장에서는 제안 기술의 압축 효율을 확인하며, 이를 검증하기 위해 서론에서 언급한 기존의 압축 기술과 본 논문에서 제안하는 움직임 예측기반 압축 기술을 비교한다. 실험을 위한 코드는 Visual Studio 2019의 C++을 통해 제작되었으며, 8I의 'longdress'와 'soldier'라는 포인트 클라우드 영상 샘플이 사용되었다.

각 영상 샘플에 대하여, 배경 기술을 적용한 시퀀스는 16장의 I 프레임, 제안 기술을 적용한 시퀀스는 4장의 I 프레임과 12장의 P 프레임으로 구성하였다. 제안 기술에서 사용한 양자화 계수는 <Table 1>과 같고, 배경 기술에서 사용한 양자화 계수는 {4,8,12,16,20}이다.

제안 기술의 양자화 계수에 따른 PSNR과 총 비트스트림 크기에 대한 결과는 <Table 1>과 같다.

Contents	QP	PSNR					Bitstream Size
		D1	D2	Y	Cb	Cr	Total (bits)
Long Dress	2	84.39878	92.91551	35.44516	42.31009	42.95737	269794208
	4			33.37238	40.89961	41.81994	178622120
	6			31.87329	39.89334	40.52046	140383768
	8			30.73888	39.23861	39.41578	119872000
	14			28.53204	37.63458	37.00039	93680440
Soldier	2	86.29909	94.70299	37.056793	53.1375937	53.9395187	287216720
	4			34.861043	52.3561875	53.0612687	188062712
	6			33.320475	51.9513375	52.4376312	150347296
	8			32.201187	51.3106562	51.9248625	131601216
	14			30.1656	50.466625	50.866125	109661968

<Table 1> PSNR & 비트스트림 크기 결과



< Figure 6> 상 : Soldier RD - Curves, 하 : Longdress RD - Curves

배경 기술의 위치 정보 압축기술은 무손실 압축인 데 반해 제안 기술은 손실 압축을 쓰는 관계로 전체 데이터에 대한 정확한 비교가 어렵다. 따라서 두 기술 간의 색상 정보인 YCbCr에 대한 압축 효율만을 비교했으며, <Table 2>의 Bjontegaard-Delta BitRates (BDBR)에서 확인할 수 있듯이 색상 값에 대해 최소 58.3%에 달하는 이득이 있었다. [7]

Contents	End-to-End BD Attribute Rate [%]		
	Y	Cb	Cr
Long Dress	-69.9%	-91.4%	-100%
Soldier	-66.0%	-58.3%	-87.7%

<Table 2> BDBR 결과 : Anchor vs Proposal

<Figure 6>와 <Figure 7>은 배경 기술과 제안 기술 사이 색상 값의 X축인 비트 레이트(bit-rate) 대비 Y축인 PSNR을 비교한 결과이다. PSNR이 비슷할 경우 제안 기술의 비트 스트림의 크기가 상대적으로 작고, 비트 스트림의 크기가 비슷할 때 제안 기술의 PSNR이 상대적으로 높다는 것을 확인할 수 있다.

4. 결론

본 논문은 3D DCT를 및 움직임 보상을 통해 포인트 클라우드 영상의 I 프레임 및 P 프레임의 압축을 구현하였다. 제안 기술은 여러 장의 프레임에 대한 효율적 압축이 불가능하였던 기존의 방식의 한계를 극복하고, 향상된 압축률을 보였음에 의의가 있다. 그러나 색상 정보의 움직임 예측 과정에 있어서 기존 방식에 비해 PSNR이 다소 떨어진다는 단점이 있다. 따라서, 움직임 예측의 정확도를 개선하는 등의 후속 연구를 통해 색상 정보 측면에서의 성능 개선이 필요하다.

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)

참고문헌

- [1] R. K. W. Chan and M. C. Lee, "3D DCT quantization as a compression technique for video sequences," Proceedings. International Conference on Virtual Systems and Multi Media VSMM '97 (Cat. No.97TB100182), Geneva, Switzerland, pp. 188-196, 1997.
- [2] G. K. Wallace, "The JPEG still picture compression standard," in IEEE Transactions on Consumer Electronics, vol. 38, no. 1, pp. xviii-xxxiv, Feb. 1992.
- [3] J.B. O'Neil, "Entropy coding in speech and television differential PCM systems," IEEE Trans. Inform. Theory IT-17, 758-761, 1971.
- [4] G. Morrison, "Video coding standards for multimedia: JPEG, H.261, MPEG," IEE Colloquium on Technology Support of Multimedia, pp. 2/1-2/4, 1992.
- [5] Huffman, D.A. A method for the construction of minimum redundancy codes. In Proceedings IRE, vol. 40, pp. 1098-1101, 1962.
- [6] HILBERT, "D. Uber die stetige Abbildung einer Linie auf ein Flächenstück," Mathematische Annalen 38, 3, 459-460, 1891.
- [7] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," Document VCEG-M33, Austin, Texas, USA, Apr. 2001.