

몰입형 비디오 부호화를 위한 점유맵 보정을 사용한 깊이의 동적 범위 확장

임성균, 황현종, *오관정, *정준영, *이광순, 김재곤
한국항공대학교, *한국전자통신연구원
{sglim, 200hhj}@kau.kr, *{kjoh, jyj0120, gslee}@etri.re.kr, jgkim@kau.ac.kr

Wider Depth Dynamic Range Using Occupancy Map Correction for Immersive Video Coding

Sung-Gyun Lim, Hyeon-Jong Hwang, *Kwan-Jung Oh, *Jun Young Jeong, *Gwangsoon
Lee, and Jae-Gon Kim
Korea Aerospace University, *ETRI

요 약

몰입형 비디오 부호화를 위한 MIV(MPEG Immersive Video) 표준은 제한된 3D 공간의 다양한 위치의 뷰(view)들을 효율적으로 압축하여 사용자에게 임의의 위치 및 방향에 대한 6 자유도(6DoF)의 몰입감을 제공한다. MIV의 참조 소프트웨어인 TMIV(Test Model for Immersive Video)에서는 복수의 뷰 간 중복되는 영역을 제거하여 전송할 화소수를 줄이기 때문에 복호화기에서 렌더링(rendering)을 위해서 각 화소의 점유(occupancy) 정보도 전송되어야 한다. TMIV는 점유맵을 깊이(depth) 아틀라스(atlas)에 포함하여 압축 전송하고, 부호화 오류로 인한 점유 정보 손실을 방지하기 위해 깊이값 표현을 위한 동적 범위의 일부를 보호대역(guard band)으로 할당한다. 이 보호대역을 줄여서 더 넓은 깊이값의 동적 범위를 사용하면 렌더링 화질을 개선시킬 수 있다. 따라서, 본 논문에서는 현재 TMIV의 점유 정보 오류 분석을 바탕으로 이를 보정하는 기법을 제시하고, 깊이 동적 범위 확장에 따른 부호화 성능을 분석한다. 제안기법은 기존의 TMIV와 비교하여 평균 1.3%의 BD-rate 성능 향상을 보여준다.

1. 서론

사용자의 움직임에 따른 시점의 변화까지 지원하는 6 자유도의 몰입형 비디오 부호화를 위한 MIV(MPEG Immersive Video) 표준은 버전 1이 완료되고 버전 2 표준 개발이 진행되고 있다[1]-[4]. MIV의 참조 소프트웨어인 TMIV[5]는 움직임 시차(motion parallax)를 제공하기 위해 다시점 영상과 함께 그 깊이(depth) 정보를 제공하고 이를 기반으로 원하는 위치의 가상 시점뷰를 합성한다. 따라서, 대용량의 다시점 영상들을 전송하기 위해서 시점 간의 중복성을 제거하는 방법을 사용하며, 복호화기에서는 제거된 영역과 그렇지 않은 영역을 구분할 수 있는 점유(occupancy) 정보가 필요하다. 각 화소별 점유 정보는 깊이 영상에 포함되어 전송되며, 깊이 영상 부호화로 인한 점유 정보 손실을 방지하고자 깊이값 동적 범위의 일부는 보호대역(guard band)으로 할당되어 있다. 본 논문에서는 압축 과정에서 점유 정보에 발생하는 오류를 복호화기에서 정정하는 기법을 제안하고 이를 활용해서 보호대역의 범위를 줄여 깊이 동적 영역을 더 많이 사용하고 렌더링 품질을 개선하고자 한다.

2. TMIV

TMIV의 몰입형 비디오 압축을 위한 주요 개념은 입력 시점 비디오들을 부호화 하기 전의 전처리를 통해 높은 상관관계를 가지는 뷰들 간의 중복성을 제거하여 압축할 데이터의 양을 최대한 줄이는 것이다. 그림 1은 TMIV 부호화기의 전체적인 구조이다. TMIV 부호화기에서는 입력된 서로 다른 시점의 비디오들을 몇 개의 그룹으로 나누고 각 그룹 내에서 기본 시점과 추가 시점으로 분류한다. 그룹별로 기본 시점과의 중복성이 있는 추가 시점의 영역들을 제거하는 프루닝(pruning) 과정이 수행된다. 프루닝 후에 남은 영역들은 직사각형 형태의 틀에 맞춰서 각각 개별적인 패치(patch)로 생성된다. 하나의 패치에는 전송되어야 할 잔여 화소들을 포함하는 유효 영역(valid region)과 프루닝에 의해 제거된 부분인 무효 영역(invalid region)이 모두 포함된다.

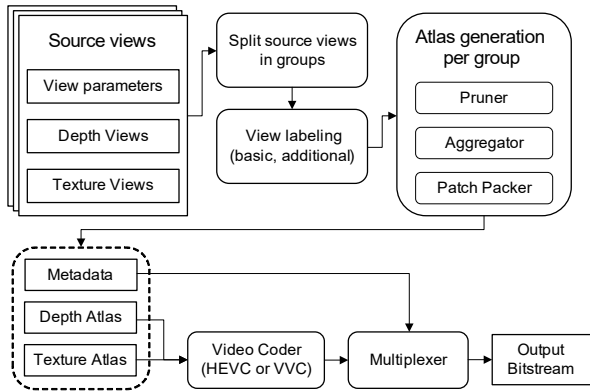


그림 1. TMIV 부호화기의 전체 구조도[5]

추가 시점에서 생성된 패치들은 기본 시점 영상과 함께 소수의 아틀라스에 패킹(packaging)된다. 패킹되는 패치들은 유효 영역이 서로 겹치지 않도록 하되 패치들이 차지하는 공간이 최소가 되도록 조밀하게 패킹된다. 패킹 공간을 최대한으로 줄이기 위해서 패치를 회전하여 패킹할 수도 있으며, 각 패치의 패킹 위치 및 회전 정보들은 메타데이터(metadata)로 구성되어 아틀라스 부호화 비트스트림(bitstream)과 함께 복호화기로 전송된다. 이런 메타데이터 정보는 TMIV 복호화기에서 아틀라스에 무작위로 패킹된 패치들이 추가 시점들의 원래 위치로 복원될 수 있도록 한다.

TMIV 에서 각 화소의 유효 여부를 나타내는 점유맵은 깊이 아틀라스에 포함되어 부호화기로 전송되며, 깊이값을 표현하기 위한 동적 범위의 일부를 점유맵 표현을 위해 사용한다. 복호화기에서는 압축 복원된 깊이 아틀라스의 화소값을 특정 임계값 T 와 비교하여 픽셀의 유효성을 판단하며, 부호화기에서는 깊이 아틀라스를 생성할 때는 부호화로 인해 점유맵에 오류가 발생하는 것을 방지하기 위해 일정 범위를 보호대역으로 둔다. 즉, 그림 2 와 같이 부호화기에서 깊이 아틀라스를 생성할 때, 비유효 화소의 깊이값은 0 으로, 유효 화소의 깊이값은 $2T$ 부터 최대값까지 선형적으로 스케일링 한다. 결과적으로 $[1, 2T - 1]$ 의 범위는 깊이값을 표현하기 위한 동적 범위로 사용되지 않는다.

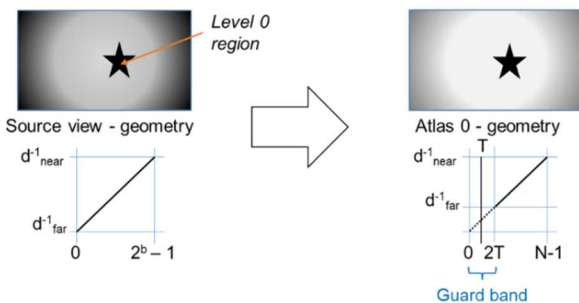


그림 2. TMIV 의 깊이값 보호대역을 이용한 점유 정보 표현[5]

3. 제안기법

TMIV 에서 생성된 아틀라스의 유효 여부는 정해진 블록 크기에 맞추어 블록 단위로 구분된다. 하지만 복호화기에서 점유 정보를 복원할 때는 압축 복원된 깊이 아틀라스로부터

화소 단위로 유효 여부를 결정한다. 부호화로 인해 깊이값에 오류가 발생한 경우에 점유 정보가 기존에 전송하고자 한 것과 달라지는 오류가 생길 수 있다. 임의의 한 블록 내에서 점유 정보에 오류가 발생한 화소는 그렇지 않은 화소보다 적을 가능성이 높다. 이 가정을 바탕으로 한 블록 내의 모든 화소의 유효 여부를 해당 블록에서 더 지배적인 경우로 통일시키는 보정 방법을 제안한다. 그림 3 은 아틀라스의 블록 크기가 8x8 일때 일부 화소에 발생한 점유 정보를 보정하는 제안기법의 적용 예를 보여준다[6].

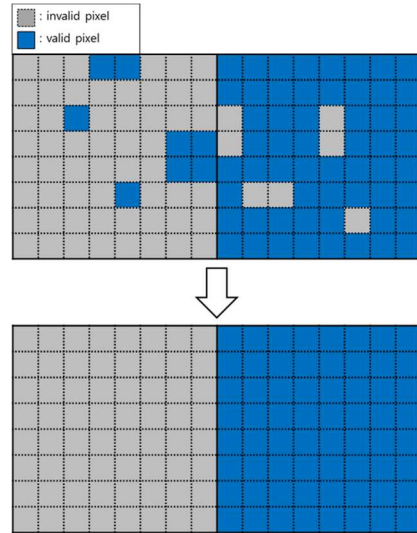


그림 3. 블록 크기가 8x8 일 때, 제안하는 보정 방법의 예[6]

제안하는 방법을 사용하여 점유맵의 오류를 보정하려면 임의의 블록 내에서 오류가 발생한 화소 수가 그렇지 않은 경우보다 적다는 가정이 성립해야 한다. 따라서, 표 1 과 같이 MIV 의 공통 테스트 조건(Common Test Conditions, CTC)[7]에서 정의된 가장 높은 QP 조건으로 아틀라스를 부호화할 때 점유 정보에 오류가 발생한 블록 수를 관찰하였다. 또한 제안기법으로 오류를 보정할 수 없는 블록 수도 확인하였다.

표 1 의 결과와 같이 T 의 값이 작을수록 점유 정보에 오류가 발생할 확률이 높기 때문에 오류 블록 수가 증가하고 제안기법으로 보정할 수 없는 블록 수도 증가한다. 그러나 T 가 기본값인 64 에서 16 으로 감소할 때까지는 제안기법으로 대부분의 블록을 바르게 보정할 수 있음을 확인했다.

표 1. T 에 따른 아틀라스 내 평균 오류 블록 수와 제안기법 적용 후의 평균 오류 블록 수 비교

Threshold value	Total blocks	Average number of error blocks	Average number of unable correction
$T = 4$	591600	23749.75	702.75
$T = 16$		2747.75	1.5
$T = 32$		280.5	0
$T = 64$ (default)		68	0

4. 실험결과

제안기법의 성능을 평가하기 위해서 입력 시점 비디오와 동일한 시점으로 합성된 렌더링 비디오를 비교하여 화질을 측정했다. MIV CTC 에 따라 WS-PSNR[8]과 IV-PSNR[9]을 사용했다. 표 2 는 $T = 32$ 로 생성된 아틀라스와 제안기법을 적용했을 때의 BD-rate 성능을 보여준다. 제안기법은 TMIV12.0 버전에 구현하였으며, VVenC[10]를 사용하여 CTC 에 따라 생성된 아틀라스를 압축했다. 기존 TMIV 방식에 비해서 의미 있는 성능 향상이 있는 것을 확인할 수 있으며, 휘도 성분 WS-PSNR 에 대하여 평균 1.3%의 이득을 보여준다.

표 2. $T = 32$ 로 조정된 후 제안기법을 적용한 몰입형 비디오 부호화 성능 실험결과(Anchor: TMIV12.0)

Sequences (resolution)	WS-PSNR (Y)	IV-PSNR
SA (4096x4096)	0.4%	0.4%
SB (2048x2048)	-0.2%	0.2%
SO (1920x1080)	-0.5%	-0.3%
SJ (1920x1080)	-6.7%	0.4%
SD (2048x1088)	-0.0%	-0.0%
SE (1920x1080)	-0.0%	-0.0%
SP (1920x1080)	-0.0%	-0.0%
SN (2048x2048)	-1.3%	0.6%
SR (1920x1080)	-3.8%	-1.7%
Average	-1.3%	-0.1%

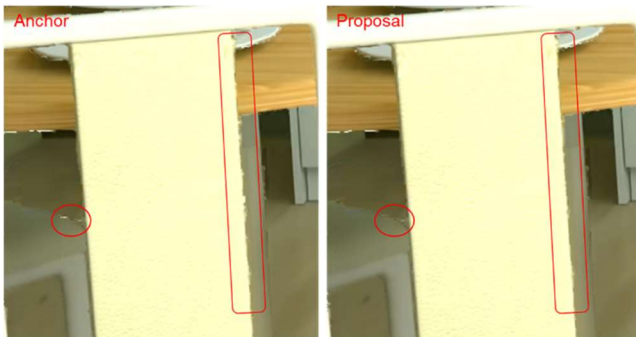


그림 4. ‘Kitchen’ 시퀀스에서 렌더링 된 시점 영상의 예 (좌: 기존 TMIV, 우: 제안기법)

그림4는 Kitchen(SJ) 시퀀스를 사용하여 기존 TMIV12.0 과 제안기법의 주관적 화질을 비교한 것이다. TMIV 로 생성된 아틀라스로 렌더링한 시점 영상에서 보이는 시각적 아티팩트(visual artifact)가 제안기법을 적용한 결과에서는 제거된 것을 확인할 수 있다. 이를 통해 제안기법은 표 1 에서 확인한 객관적 화질뿐만 아니라 주관적 화질 측면에서도 눈에 띄는 개선을 보이는 것을 알 수 있다.

5. 결론

본 논문에서는 몰입형 비디오 부호화(MIV)의 깊이 아틀라스 압축 과정에 발생한 오류로 인한 점유맵 오류의 보정 방법과 이를 함께 사용하여 더 넓은 깊이 동적 범위의 사용을 통해 향상된 부호화 성능을 제시한다. 즉, 제안하는 보정 방법을 통해 점유맵 오류를 방지하기 위해 깊이 동적 범위에

할당된 보호 대역을 줄이고 깊이 동적 범위를 확장함으로써 보다 정확하게 깊이값을 표현하고, 렌더링 화질을 개선시킬 수 있다. 실험결과에서 제안기법으로 휘도성분의 WS-PSNR 에서 BD-rate 1.3%의 비트율 절감의 부호화 효율 개선과 두드러진 주관적인 화질 개선을 얻을 수 있음을 확인할 수 있었다.

Acknowledgement

이 논문은 2022 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. 2018-0-00207, 이머시브 미디어 전문연구실).

References

- [1] “MPEG-I Use Cases for Omnidirectional 6DoF, Windowed 6DoF, and 6DoF,” ISO/IEC JTC1/SC29/WG11, N16768, Apr. 2017.
- [2] M. Wien, J. M. Boyce, T. Stockhammer, and W.-H. Peng, “Standardization Status of Immersive Video Coding,” IEEE J. Emerg. Select. Topics Circuits Syst., vol. 9, no. 1, pp. 5-17, Mar. 2019.
- [3] “Text of ISO/IEC FDIS 23090-12 MPEG Immersive Video,” ISO/IEC JTC1/SC29/WG04, N00111, Jul. 2021.
- [4] “Preliminary WD1 of ISO/IEC 23090-12 MPEG Immersive video Ed. 2,” ISO/IEC JTC1/SC29/WG04, N00176, Jan. 2022.
- [5] B. Salahieh, J. Jung, A. Dziembowski (Eds.), “Test Model 11 for Immersive Video,” ISO/IEC JTC1/SC29/WG04, N00142, Oct. 2021.
- [6] S. Lim, H. Hwang, J. Kim, K. Oh, J. Jeong, and G. Lee, “Wider Depth Dynamic Range Using Occupancy Map Correction,” ISO/IEC JTC1/SC29/WG04, m59549, Apr. 2022.
- [7] J. Jung, B. Kroon, “Common Test Conditions for MPEG Immersive Video,” ISO/IEC JTC1/SC29/WG04, N00169, Jan. 2022.
- [8] “WS-PSNR Software Manual,” ISO/IEC JTC1/SC29/WG11, N18069, Oct. 2019.
- [9] A. Dziembowski, “Software manual of IV-PSNR for Immersive Video,” ISO/IEC JTC1/SC29/WG04, N00013, Oct. 2020.
- [10] VVenC software, [Online]. Available at: <https://github.com/fraunhoferhhi/vvenc/tree/v0.2.0.0>