

MIV의 효율적인 부호화를 위한 비대칭 임계값 기반 점유맵 보정

김동하, 임성균, 김정윤, 김재곤

한국항공대학교

{donghakim, sglim, jyun70}@kau.kr, jgkim@kau.ac.kr

Asymmetric Threshold-Based Occupancy Map Correction for Efficient Coding of MPEG Immersive Video

Dong-Ha Kim, Sung-Gyun Lim, Jeong-yoon Kim, and Jae-Gon Kim

Korea Aerospace University

요 약

MIV(MPEG Immersive Video)의 시험모델 TMIV는 다시점의 비디오와 깊이(depth) 비디오를 입력 받아 시점 사이의 중복성을 제거한 후 남은 텍스처(texture)와 깊이로 텍스처 아틀라스(atlas)와 깊이 아틀라스를 각각 생성하고 이를 압축한다. 각 화소별 점유(occupancy) 정보는 깊이 아틀라스에 포함되어 압축되는데 압축 손실로 인한 점유맵 오류를 방지하기 위하여 임계값 $T = 64$ 로 설정한 보호대역을 사용한다. 기존에 설정된 임계값을 낮추어 깊이 동적범위를 확대하면 보다 정확한 깊이값 표현으로 부호화 효율을 개선할 수 있지만 보호대역 축소로 점유맵 오류가 증가한다. 본 논문에서는 TMIV의 부호화기와 보호화기에 비대칭 임계값을 사용하여 보호대역 축소로 인한 점유맵 오류를 보정하면서 보다 정확한 깊이 값 표현을 통하여 부호화 효율을 개선하는 기법을 제안한다. 제안기법은 깊이 동적범위 확대와 비대칭 임계값 기반의 점유맵 오류 보정을 통하여 CG 시퀀스에서 2.2% BD-rate 이득과 주관적 화질 개선을 보인다.

1. 서론

움직임 시차(parallax)를 지원하는 6 자유도(DoF: degree of freedom) 몰입형 비디오 부호화를 위한 MIV(MPEG Immersive Video) 표준은 현재 버전 1이 완료됐고 버전 2 표준 개발을 진행 중이다[1-4]. MIV의 참조 소프트웨어인 TMIV(Test Model for Immersive Video)의 부호화기는 다시점 비디오와 깊이(depth) 비디오로 구성된 6DoF 비디오를 입력 받아 시점 사이의 중복성을 제거하는 푸루닝(pruning) 과정을 거친 후 제거되지 않은 텍스처(texture)와 깊이를 패킹(packaging)한 각각의 아틀라스(atlas)를 생성 압축한다. TMIV 복호화기는 복원한 아틀라스로부터 다시점의 비디오와 깊이를 재구성한다. 이때 아틀라스 생성시 제거되지 않은 화소를 나타내는 점유(occupancy) 정보가 필수적이다. 각 화소별 점유 정보는 깊이 아틀라스에 포함되어 전송되는데, 깊이 아틀라스의 압축

손실로 인한 점유 정보의 오류를 방지하기 위해서 깊이 동적범위에 보호대역(guard band)을 할당한다. 보호대역은 부호화기에서 설정된 깊이 점유맵 임계값으로 결정된다. 기존에 기본값으로 설정된 임계값을 낮추어 깊이 동적범위를 확대하면 보다 정확한 깊이값 표현으로 부호화 효율을 개선할 수 있다. 하지만 보호대역 축소로 점유맵 오류가 증가하고 이로 인한 시각적 아티팩트(artifact)가 발생할 수 있다. 본 논문에서는 TMIV의 부호화기와 복호화기에 비대칭 임계값을 사용하여 보호대역 축소로 인한 점유맵 오류를 보정하면서 보다 정확한 깊이 값 표현을 통하여 부호화 효율을 개선하는 기법을 제안한다.

2. 비대칭 임계값 기반 점유맵 보정

TMIV 복호화기는 점유 정보가 포함된 깊이 아틀라스를 입력

받고, 깊이 점유맵 임계값 T 를 사용하여 픽셀의 유효성을 판단한다. 이때, TMIV 부호화기는 깊이 아틀라스 생성시 점유맵의 오류가 발생하는 것을 방지하기 위해 일정 범위를 보호대역으로 설정한다. 그림 1은 TMIV의 보호대역을 적용할 때 깊이 정보의 동적범위를 묘사한다. 보호대역은 TMIV 부호화기에서 설정한 깊이 점유맵 임계값 T 에 따라 보호대역의 범위는 $[1, 2T - 1]$ 로 정의되어 이 범위의 깊이값은 동적범위로 사용되지 않는다. 유효한 깊이 정보 화소는 $2T$ 부터 최대값까지 선형적으로 스케일링 한다. 예를들어, 유효한 깊이정보 화소의 값이 0 이면 $2T$ 의 값으로 스케일링이 된다. 따라서, 깊이 동적범위는 $2T$ 부터 최대값이다.

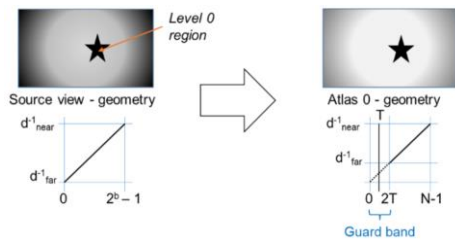


그림 1. TMIV의 깊이값 보호대역을 이용한 점유 정보 표현

TMIV에서는 기본값으로 깊이 점유맵 임계값 $T = 64$ 로 설정되어 있다. 임계값을 낮추면 보호대역의 크기가 감소하여 표현할 수 있는 깊이 동적범위가 확장되어 렌더링 품질이 증가할 수 있다[5]. 하지만, 동시에 깊이 점유맵 임계값 감소로 인해 기존의 보호대역으로 생성되지 않던 점유맵 오류가 생성된다.

그림 2는 기존의 부호화기의 점유맵 임계값 $T = 64$ 를 사용하지 않고, $T = 4$ 를 사용하여 Group(R) 시퀀스의 97 프레임 깊이 아틀라스를 압축 복원했을 때, 점유맵 오류 중 하나인 무효화소가 유효화소로 바뀌는 오류의 분포 히스토그램이다. 해당 아틀라스를 TMIV 부호화기로 렌더링을 진행한다면 임계값 $T = 4$ 이하의 오류들은 제거되고 렌더링을 진행하지만 임계값 이상의 오류들은 존재하게 되어 렌더링 화질에 시각적 아티팩트를 야기할 수 있다. 따라서, 복호화기 점유맵 임계값을 부호화기의 임계값 보다 높은 $T = 6$ 으로 사용하면 그림 2와 같이 파란색 면적만큼의 오류 보정을 할 수 있다.

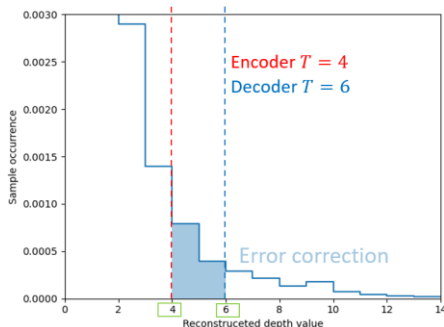


그림 2. 무효화소를 유효화소로 결정한 점유맵 오류 분포 히스토그램(R sequence, 97 frame)

이와 같이 본 논문은 부호화기의 임계값을 낮추고 복호화기의 점유맵 임계값을 비대칭으로 부호화기에서 사용한 값 보다 높은 적절한 임계값을 사용하는 점유맵 오류 보정을 기법을 제안한다[6].

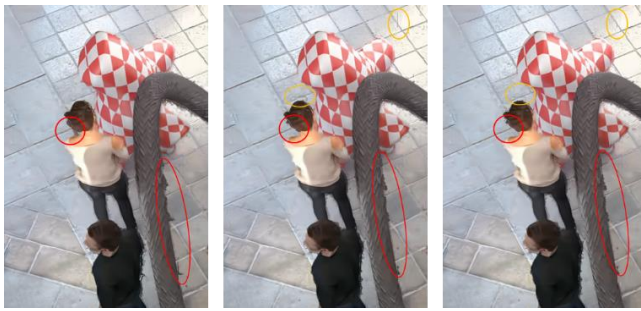
3. 실험결과

실험은 부호화기의 $T = 4$ 로 고정하고 복호화기의 $T = \{4, 6\}$ 으로 설정하여, 점유맵 임계값을 낮추어 단순히 깊이 동적 확대를 적용했을 때와 이와 함께 비대칭 임계값을 적용하여 오류 보정을 수행했을 때의 부호화 성능비교를 한다. TMIV CTC[7]에 따라 Bit-rate 대비 WS-PSNR 을 사용하여 BD-rate 성능을 측정했다. TMIV 버전 14.0[8]에 구현하였으며 VVenC[9]를 사용하여 생성된 97 프레임 아틀라스를 압축했다. 표 1은 각 복호화기 T 값의 다른 TMIV 앵커(anchor)[10]에 대비 성능향상을 나타낸 것이다. 표 1과 같이 T 의 값을 낮추어 깊이 동적 확대만을 적용했을 때와, 제안기법을 함께 적용했을 때의 각각 BD-rate 이득은 2.1%와 2.2%로 미미한 성능개선이 있다.

표 1. 비대칭 임계값 적용 BD-rate 부호화 성능 (97 frames, 부호화기 $T = 4$)

CG Sequence	$T = 4$ in encoder			
	$T = 4$ in decoder		$T = 6$ in decoder	
	High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR	High-BR BD rate Y-PSNR	Low-BR BD rate Y-PSNR
Classroom	1.2%	1.8%	1.1%	1.8%
Museum	-0.3%	0.2%	-0.3%	0.2%
Fan	-0.4%	0.5%	-0.4%	0.5%
Kitchen	-2.9%	-1.2%	-2.9%	-1.2%
Chess	-4.1%	-1.6%	-4.1%	-1.6%
Group	-6.4%	-3.6%	-6.4%	-3.7%
Average	-2.1%	-0.7%	-2.2%	-0.7%

그림 3은 제안기법의 주관적 화질을 비교한 것이다. 붉은색 원은 부호화기의 $T = 4$ 로 하여 깊이 동적 확대에 의한 화질 개선을 TMIV 앵커와 비교한 것이고, 주황색 원은 제안기법 적용시 시각적 아티팩트가 감소된 부분을 나타낸다. 표 2와 그림 3를 통하여, 제안기법은 TMIV 14.0 대비 많은 객관적 부호화 성능은 유사하지만 보다 정확한 깊이 값의 표현과 점유맵 오류 보정을 통하여 시각적 아티팩트를 방지하면서 눈에 띄는 주관적 화질 개선을 확인하였다.



Anchor Decoder $T = 4$ Decoder $T = 6$
가) Group 시퀀스



Anchor Decoder $T = 4$ Decoder $T = 6$
나) Kitchen 시퀀스

그림 3. 제안기법의 주관적 화질 개선(encoder $T = 4$)

4. 결론

본 논문에서는 점유맵 임계값을 낮추어 점유 정보의 보호대역을 줄이고 깊이의 동적범위를 확장하여 보다 정확한 깊이값 표현으로 렌더링 화질을 개선한다. 이때, TMIV의 부호화기와 복호화기에 비대칭 점유맵 임계값을 사용하여 보호대역 축소로 인한 점유맵 오류를 보정하여 시각적 아티팩트를 줄이는 기법을 제시하였다.

제안기법은 깊이 동적범위 확대와 비대칭 임계값 기반의 점유맵 오류 보정을 통하여 CG 시퀀스에서 2.2% BD-rate 이득과 뚜렷한 주관적 화질 개선을 보였다. 제안기법은 6DoF 몰입형 비디오 부호화 효율 향상을 위한 기법을 활용될 것으로 기대된다.

Acknowledgement

본 논문은 2022 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. 2018-0-00207, 2017-0-00486).

참고 문헌(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] 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.

[6] D. Kim, S. Lim, J. Kim, K. Oh, J. Jeong, and G. Lee, "Wider Depth Dynamic Range with Occupancy Error Correction," ISO/IEC JTC1/SC29/WG04, m60846, Oct. 2022.

[7] J. Jung, B. Kroon, "Common Test Conditions for MPEG Immersive Video," ISO/IEC JTC1/SC29/WG4, N0203, Apr. 2022.

[8] B. Salahieh, J. Jung, A. Dziembowski (Eds.), "Test Model 14 for MPEG Immersive Video," ISO/IEC JTC1/SC 29/WG4, N0242, Jul. 2022

[9] "VVenc", V0.3.0.1, [online], Available: <https://github.com/fraunhoferhhi/vvenc/releases/tag/v0.3.1.0>

[10] A. Dziembowski, J. Jung, and B. Kroon, "Report of MPEG immersive video CTC anchor generation," ISO/IEC JTC 1/SC 29/WG4, N0234, Jul. 2022.