

MV-HEVC 기반 TMIV 에서의 성능 개선

이광순, 정준영, 오관정, 서정일

한국전자통신연구원

gslee@etri.re.kr, jyj0120@etri.re.kr, kjoh@etri.re.kr, seoji@etri.re.kr

Performance Improvement of TMIV based on MV-HEVC

Gwangsoon Lee, Kwan-Jung Oh, Jun Young Jeong, Jeongil Seo

Electronics and Telecommunications Research Institute (ETRI)

요 약

본문에서는 TMIV 에서 비디오 코덱으로 사용되고 있는 HEVC 를 MV(Multiview)-HEVC 로 대체한 실험결과를 소개하고 부호화 효율을 향상시키기 위한 아틀라스 생성 구조에 대해 제안하고자 한다. 이를 위해 본 논문에서는 푸루닝(pruning)된 패치를 패킹한 아틀라스 영상으로 구성된 MIV 앵커(anchor)에 MV-HEVC 를 적용하는 구조, 소스 시점영상으로부터 선택된 기본 시점영상으로만 패킹한 아틀라스 영상으로 구성된 MIV view 앵커에 적용하는 구조를 실험한다. 이와 더불어 부호화 효율을 향상시키기 위해 선택된 기본시점 영상을 패킹함에 있어서 2 개의 아틀라스영상에 걸쳐 가장 인접한 시점을 배치하는 아틀라스 영상 구조를 제안한다. 실험결과, 기존의 MIV 앵커에 MV-HEVC 를 적용하면 성능 개선이 거의 없으며, MIV view 앵커와 제안한 MIV view 앵커의 아틀라스영상 구조에 MV-HEVC 를 적용하면 객관적인 성능이 최대로 얻을 수 있음을 알 수 있었다.

1. 서론

MPEG-I 비디오 그룹에서의 이머시브 비디오(immersive video)는 몰입감을 극대화하기 위해 기존의 평면 비디오뿐만 아니라 전방위 360 비디오에 운동시차(motion parallax)의 부여가 가능한 것으로서, 현재 부호화 표준화의 완료단계에 있으며 그 표준의 명칭을 MIV(MPEG Immersive Video)라고 정하였다 [1]. 그림 1 은 MIV 표준 검증을 위해 개발된 테스트모델인 TMIV(Test Model for Immersive Video) [2] 인코더 및 디코더를 이용한 MIV 부복호화기 개념도를 설명하고 있다. TMIV 인코더는 다중의 텍스처와 깊이정보로 구성되는 이머시브 비디오로부터 공간적인 중복성을 제거한 후 렌더링에 필요한 패치들만으로 패킹한 텍스처 및 기하 아틀라스(atlas) 영상을 출력하고 이를 각각 기존의 HEVC 또는 VVC(Versatile Video Coding)들로 부복호화한다. TMIV 디코더는 디코딩된 텍스처 및 기하 아틀라스 영상으로부터 시청자의 시청위치에 부합되는 뷰포트(viewport)영상을 렌더링함으로써

운동시차를 재현하게 된다. 본문에서는 TMIV 에서 비디오 코덱으로 사용되고 있는 HEVC 를 MV-HEVC 로 대체하는 실험과 이때 부호화 효율을 향상시키기 위한 아틀라스 생성 구조에 대해 제안하고자 한다.

본 논문의 구성은 다음과 같다. 2 절에서는 MV-HEVC 기반의 TMIV 부복호화 실험을 위한 아틀라스 구조에 대해 설명하고 3 절에서 각 구조를 실험결과를 확인하고 마지막으로 본 논문에 대한 결론을 맺는다.

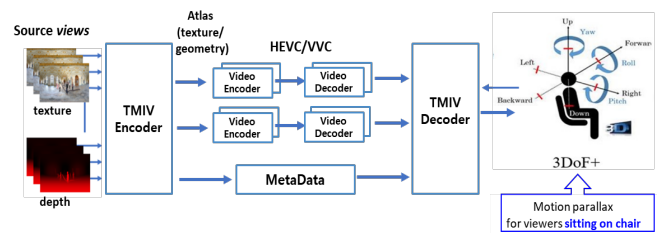


그림 1. MIV 를 위한 TMIV 부복호화 개념

2. MV-HEVC 기반의 TMIV 부복호화 구조

본 논문에서 MV-HEVC 를 TMIV 에 적용하여 부복호화 성능을 실험하기 위한 구조는 그림 2 에서와 같다. 현재 TMIV 부복호화에서 텍스처 및 기하 아틀라스 영상을 부복호화하기 위한 비디오 코덱을 MV-HEVC 로 대체한다. TMIV 앵커는 아틀라스가 2 개의 쌍(텍스처 및 기하)으로 제한되었으므로 시점간 예측을 위한 계층의 개수를 2 로 설정한다.

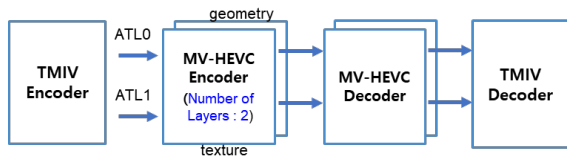
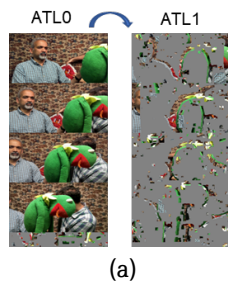


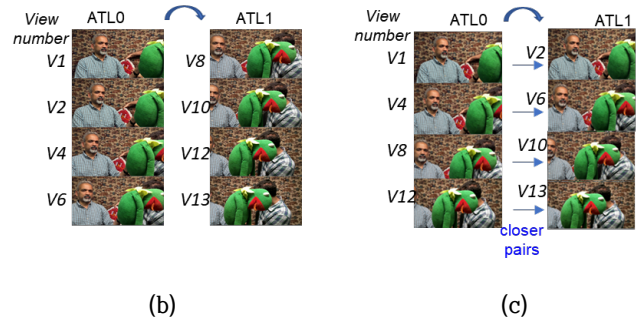
그림 2. MV-HEVC 적용 실험을 위한 TMIV 부복호화 구조

MV-HEVC 를 적용하기 위한 아틀라스 영상의 구조는 그림 3 에서와 같다. 첫번째 실험에서는 TMIV MIV 앵커에서 HEVC 를 MV-HEVC 를 대체하고 기본 시점영상과 푸루닝(pruning)된 패치들로 아틀라스 영상을 구성하였으며, 두번째 실험에서는 TMIV MIV View 앵커에서 HEVC 를 MV-HEVC 로 대체한다. 즉, 모든 소스 시점영상에서 일부 선택된 기본시점 영상만을 패킹한 아틀라스영상을 부복호화한다. 한편, TMIV MIV View 앵커는 푸루닝되지 않는 기본시점 영상만으로 아틀라스를 구성하지만, 선택되어 패킹되는 시점영상의 순서는 아틀라스내에서 순차적으로 배치된다. 제안된 아틀라스 영상 구조에서는 기본 시점영상을 구성할 때 MV-HEVC 에서 시점간 예측(interview prediction) 시 가장 인접한 시점이 되도록 한다. 즉, 선택된 기본시점 영상을 패킹함에 있어서 2 개의 아틀라스영상에 걸쳐 가장 인접한 시점이 배치되도록 하여 시점간부호화 효율을 높이도록 한다.



(a)

그림 3. MV-HEVC 를 적용하기 위한 아틀라스 영상 구조: (a) MIV 앵커, (b) MIV view 앵커, (c) 제안한 MIV view 앵커



(b)

(c)

그림 3. (계속)

3. 실험 결과

TMIV7.0 CTC [3]조건하에서 3 개의 실사 MIV 콘텐츠를 이용하여 그림 3 에서의 아틀라스 구조로 실험을 수행하여 객관적 성능을 분석하였다. 그림 4(a)는 MIV 앵커 대한 실험 결과로서, BD 율은 평균 -0.4~-0.6%, IV-PSNR 은 -2.0~-3.2%의 이득을 보였다. 그림 4(b)는 MIV View 앵커에 대한 실험 결과로서, BD 율은 평균 -0.6~-0.7%, IV-PSNR 은 -3.8~-4.2%의 이득을 보였다. 반면에 그림 4(c)는 제안한 MIV View 앵커에 대한 실험결과로서, 기본 시점영상을 아틀라스에 걸쳐 인접하게 재배치함으로써 기존 MIV View 앵커에 비해 WS-PSNR 에서 약 -7%, IV-PSNR 에서 -9.2~-10.7%의 추가 이득이 있음을 알 수 있다.

Sequence		High-BR	Low-BR	High-BR	Low-BR	High-BR	Low-BR
		BD rate	BD rate	BD rate	BD rate	BD rate	BD rate
		Y-PSNR	Y-PSNR	VMAF	VMAF	IV-PSNR	IV-PSNR
Painter	SD	-8.2%	-8.6%	-8.8%	-9.8%	-9.9%	-11.3%
Frog	SE	3.3%	4.7%	1.7%	3.7%	0.8%	2.5%
Carpark	SP	3.8%	2.1%	2.8%	1.4%	3.1%	-0.9%
MIV		-0.4%	-0.6%	-1.4%	-1.6%	-2.0%	-3.2%

(a)

Sequence		High-BR	Low-BR	High-BR	Low-BR	High-BR	Low-BR
		BD rate	BD rate	BD rate	BD rate	BD rate	BD rate
		Y-PSNR	Y-PSNR	VMAF	VMAF	IV-PSNR	IV-PSNR
Painter	SD	-2.2%	-2.3%	-3.9%	-4.3%	-3.3%	-3.5%
Frog	SE	-3.3%	-3.2%			-4.1%	-4.6%
Carpark	SP	3.3%	3.6%	-4.3%	-4.5%	-4.1%	-4.6%
MIV		-0.7%	-0.6%	-4.1%	-4.4%	-3.8%	-4.2%

(b)

Sequence		High-BR	Low-BR	High-BR	Low-BR	High-BR	Low-BR
		BD rate	BD rate	BD rate	BD rate	BD rate	BD rate
		Y-PSNR	Y-PSNR	VMAF	VMAF	IV-PSNR	IV-PSNR
Painter	SD	-9.5%	-9.6%	-9.1%	-10.3%	-11.6%	-13.0%
Frog	SE	-14.0%	-13.1%	-17.9%	-16.9%	-17.1%	-16.1%
Carpark	SP	2.3%	0.5%	0.4%	0.3%	1.2%	-3.0%
MIV		-7.1%	-7.4%	-8.9%	-9.0%	-9.2%	-10.7%

(c)

그림 4. 객관적인 실험결과: (a) MIV 앵커, (b) MIV view 앵커, (c) 제안한 MIV view 앵커

4. 결론

본 논문에서는 시점간 예측기술을 MIV 에 적용하기 위하여 HEVC 를 MV-HEVC 로 대체하여 객관적인 성능을 비교하였다. 또한 아틀라스 영상간에 간단하게 시점간 예측기술을 적용할 수 있는 구조를 제안하고 객관적인 성능이 향상됨을 검증하였다. VVC 에서도 시점간 예측기술이 적용 가능하도록 표준이 확장된다면, 확장된 VVC 기반의 TMIV 부호화에 응용하여 부호화 효율을 향상시킬 수 있을 것이다.

감사의 글

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임(No. 2018-0-00207, 이머시브전문연구실)

참 고 문 헌

- [1]. J. Boyce, R. Doré, and V. K. M. Vadakital, Committee Draft for Immersive Video, Standard ISO/IEC JTC1/SC29/WG11MPEG/N19482, Jul. 2020. [5] B.
- [2]. Salahieh, B. Kroon, J. Jung, and A. Dziembowski, Test Model 6 for Immersive Video, Standard ISO/IEC JTC1/SC29/WG11MPEG/N19483, Jul. 2020
- [3]. J. Jung, B. Kroon, Common Test Conditions for MPEG Immersive Video, ISO/IEC JTC1/SC29/WG04 N0051, Jan. 2021, Online.