

# 몰입형 비디오를 위한 MV-HEVC 적용 그룹 기반 TMIV 부호화 기법

김인에, 정종범, 이순빈, 류영일, 류은석  
 성균관대학교 컴퓨터교육과  
 inaelk@skku.edu, uof4949@skku.edu, soonbinlee@skku.edu,  
 yeongilryu@skku.edu, esryu@skku.edu

## Group-Based TMIV Encoding for Immersive Video using MV-HEVC

Inae Kim, Jong-Beom Jeong, Soonbin Lee,  
 Yeongil Ryu, Eun-Seok Ryu  
 Department of Computer Education, Sungkyunkwan University

### 요약

본 논문은 test model for immersive video (TMIV)의 metadata for immersive video (MIV) view mode 로 생성된 시점화면을 인접한 위치의 시점끼리 그룹으로 묶어 I-frame을 두 개 제공하는 그룹 방식을 적용 및 multi view-high efficient video coding (MV-HEVC)를 활용하여 부호화하는 기술에 대해 실험 비교하였다. 본 연구를 위하여 moving picture experts group (MPEG) 에서 공통 실험조건으로 제시된 TechnicolorMuseum 시퀀스에 대해 앞서 설명한 방식을 실험하였으며, 기본 MV-HEVC를 적용한 TMIV 부호화 방식 대비 MV-HEVC에 그룹 기반 방식을 적용한 부호화 방법은 Y-PSNR에서 평균 2.05%의 BD-rate를 절약하였다. 본 연구는 객관적인 평가를 위해 최신 몰입형 영상의 품질 평가 지표인 VIF, VMAF, IV-PSNR, MS-SSIM에 의한 비교 결과도 함께 제공한다.

### 1. 서론

미디어 콘텐츠 서비스 이용 주체가 가족에서 개인으로 변화하고, 개인이 선호하는 디바이스를 통해 각기 다른 디지털 콘텐츠를 소비하면서 미디어 소비시장이 나날이 발전되고 있다. 이와 동반하여 몰입감 있는 영상서비스를 위한 가상현실(virtual reality; VR) 기술개발이 활발하게 진행 중이다.

ISO 산하 국제표준 단체인 MPEG에선 몰입감 극대화를 위한 가상현실 기술의 표준화를 위해 MPEG-Immersive (MPEG-I)를 설립하여 high efficient video coding (HEVC) 이상의 영상 압축 기술개발을 위한 versatile video coding (VVC) 표준에 대해 활발히 논의 중이다.

MPEG-I(ISO/IEC 23090)에선 일반 비디오 대비 더욱 많은 시점을 전달해야 하는 몰입형 비디오의 특성을 살려 6 시점 자유도가 이용 가능한 압축표준을 수행할 수 있도록 2019년 3월 제네바 회의 시 CIP 대응을 통해 제안된 기술을 통합하여 실험 모델인 TMIV 소프트웨어를 정의하였고, 당시 결정된 5개의 CE를 진행한 결과를 바탕으로 10월 회의에서 TMIV 3.0으로 업그레이드 하였다[1]-[3].

TMIV는 부호화기 내부의 아틀라스 생성기(atlas constructor)에서 하나의 basic view와 시점 간 차이 정보 패치를 모은 아틀라스를 보내 나머지 추가 시점을 복원하는 MIV mode와 전체 시점 중 특정 시점의 영상들을 아틀라스로 구성하여 보내 중간 시점을 복원하는 MIV view mode 두 가지 방법을 지원한다. 본 연구는 MIV view mode로 생성하여 특정 시점의 화면을 모은 아틀라스를 인접한 위치의 시점끼리 두 그룹으로 분리하여 MV-HEVC를 적용한 그룹 기반 부호화 기술을 제안한다[2].

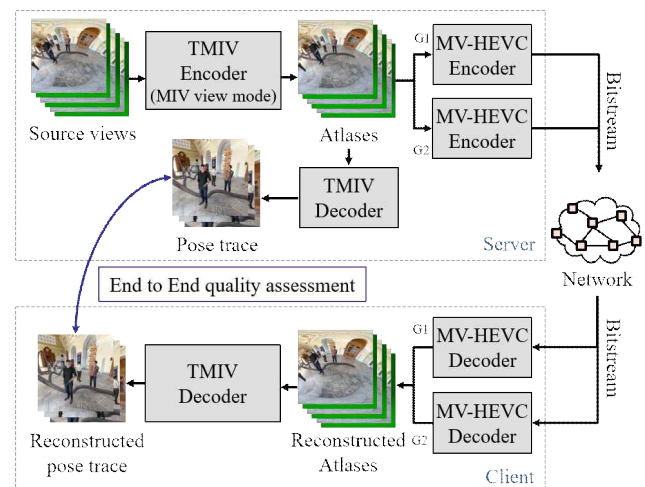
### 2. 관련연구

본 절에서는 Group-based TMIV에 대해 소개한다. 기존 TMIV의 몰입형 부호화기는 복호화기 측의 렌더러 품질향상

을 고려하지 않고 최저 데이터를 부호화하여, 위치에 따른 연관성이 무시된 채로 배경 패치가 렌더링 되기 때문에 다중 패스 렌더러가 실패한다. Group-based TMIV의 특징은 기존 TMIV의 MIV mode와 같이 모든 뷰를 한꺼번에 인코딩 하는 것이 아니라 중요도가 높은 패치(전경 물체, 가장자리 영역 등)를 각각 인접한 시점의 여러 그룹으로 나누어 각 그룹을 별도로 부호화 및 렌더링하는 것이며, HEVC 압축 기술의 테스트 모델인 HM을 통해 부호화하였다. 이 기술은 부호화 처리 속도를 개선하고 렌더링 품질을 향상한다. 그러나, 해당 기술은 MIV mode에서 성능향상을 나타내었으나, MIV view mode에서는 성능이 감소한다[4].

### 3. 그룹 기반 TMIV 부호화 기법

본 절에서는 MV-HEVC 에서의 그룹 기반 TMIV 부호화 기법 적용의 설계 및 참조 구조에 대하여 기술한다.



(그림 1) MV-HEVC 활용 그룹 기반 TMIV 부호화 블록도

TMIV 부호화기에서 MIV view mode를 선택하면 원본 시점 영상 중 특정 시점의 영상을 아틀라스로 생성한다. 이와 동시에 품질평가를 위한 포즈 추정 영상을 생성한다. 부호화된 아틀라스는 인접한 시점 위치에 따라 G1과 G2 두 개 그룹으로 묶어 MV-HEVC를 적용하여 부호화한다. MV-HEVC 부호화기를 통해 생성된 비트스트림은 네트워크를 통해 클라이언트로 전송되고 MV-HEVC 복호화기를 통해 재구성된 아틀라스로 복호화한다. 그림 1은 제안하는 그룹 기반 TMIV 부호화 방식의 진행순서를 블록도로 나타낸 것이다. 기본 MV-HEVC로 부호화한 TMIV 아틀라스는 1개의 I-frame과 나머지 시점 영상은 P-frame으로 부호화하며, 이 두 frame을 참조하여 B-frame을 생성하는 구조이다. 본 논문에서 제안하는 그룹 기반 방식은 2개의 I-frame과 각 그룹에 배치된 시점 수에 따라 P-frame으로 부호화하고, 그룹별 각각 I-frame과 P-frame을 참조하여 B-frame을 생성한다. 이때 참조구조는 시간의 흐름에 따라 순차적인 구조를 지니게 된다.

4. 실험결과

본 논문에서는 그룹 기반 부호화 기술을 MV-HEVC에 적용함과 동시에 다시점 영상처리의 부호화 성능향상의 효율을 측정하기 위해 MPEG-I에서 몰입형 영상에 대한 실험조건인 common test condition (CTC)를 준수하였다.

(표 1) TechnicolorMuseum 시퀀스 정보

Sequence name	Resolution	No. of Source views	MIV View Anchor	View Per Group
Technicolor Museum	2048x2048	24	v0, v1, v4, v8, v11, v12, v13, v17	G1: v0, ..., v11 G2: v12, ..., v23

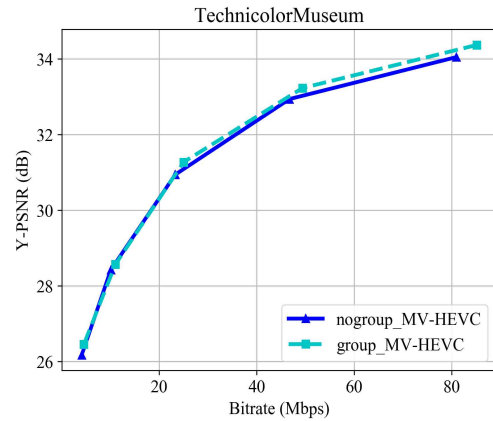
실험에는 CTC에 정의된 TechnicolorMuseum 시퀀스를 사용하였으며, 표 1은 실험에 활용된 시퀀스의 시점정보가 어느 그룹에 해당하는지 보여준다. 본 실험은 MV-HEVC test model (HTM) 13.0 에 기반하여 구현하였다. 그룹을 적용한 아틀라스에 MV-HEVC를 적용하여 생성된 원본 포즈 추정 영상과 클라이언트에서 재구성된 포즈 추정 영상의 품질을 비교하여 성능을 평가 및 기존 부호화 기술과 비교한다. 본 연구의 결과는 최신 몰입형 영상의 품질평가 기준인 weighted to spherically uniform peak signal to noise ratio(WS-PSNR), visual information fidelity(VIF), video multimethod assessment fusion (VMAF), multi-scale structural similarity(MS-SSIM), immersive video PSNR (IV-PSNR)을 지원하는 소프트웨어를 사용하여 측정했다[5].

(표 2) 기본 MV-HEVC 적용 대비 그룹 기반 MV-HEVC 부호화 적용 BD-rate 절약을

Bitrate	Y-PSNR	VIF	VMAF	MS-SSIM	IV-PSNR
High bitrate	-3.44%	-4.13%	-2.66%	-4.51%	-3.81%
Low bitrate	-0.67%	-3.23%	-1.47%	-2.70%	-1.60%
Average	<b>-2.05%</b>	<b>-3.68%</b>	<b>-2.06%</b>	<b>-3.61%</b>	<b>-2.71%</b>

본 연구는 5가지 특정한 quantization parameter (QP) 값으로 양자화하여 부호화한 영상을 기준으로 QP 1~4까지를 고대역폭 영상, QP 2~5까지를 저대역폭 영상으로 정의하여 고대역폭과 저대역폭 영상 두 가지에 대하여 평가하는 MPEG의

CTC 방식을 준수한다. 표 2는 제안하는 부호화 방식을 적용하여, 고대역폭 영상과 저대역폭 영상 모두에서 BD-rate 절감이 가능함을 보여준다. 제안하는 부호화 방식 사용 시 Y-PSNR, VIF, VMAF, MS-SSIM, IV-PSNR 에 대해 평균 2.06%, 3.68%, 2.06%, 3.61%, 2.71%의 BD-rate 절감이 가능하다. 그림 2는 기존 부호화 기술 및 제안하는 그룹 기반 부호화 기술의 rate-distortion (RD) curve를 보여준다. 그림 2에서 나타내듯이 고대역폭에서 기존 부호화 기술보다 더 좋은 품질의 영상을 제공할 수 있음을 알 수 있다.



(그림 2) 기존 부호화 기술 및 제안하는 부호화 기술 RD-curve

5. 결론 및 향후 연구

본 논문은 TMIV에서 MIV view mode로 생성된 아틀라스를 인접한 시점끼리 그룹으로 묶어 MV-HEVC에서 부호화하였을 때, Y-PSNR에서 평균 2.05%의 BD-rate 절감이 가능함을 확인하였다. 현재 수행한 연구 결과를 토대로, 추가 테스트 시퀀스에 대한 실험과 사용자의 시점 방향에 따라 선택된 시점 그룹만을 전송할 수 있는 그룹 선택기 개발을 후속 연구로 진행할 계획이다.

ACKNOWLEDGMENT

이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No.2020-0-00231, 5G 엡지클라우드 기반 VR·AR 저지연 스트리밍 기술 개발)과 2020년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1A2C1010476).

참고문헌

[1] 박현수, 박상호, 강제원, “6 자유도 전방위 몰입형 비디오의 압축 코덱 개발 및 성능분석”, 방송공학회논문지 제24권 제6호, 2019.

[2] B. Salahieh, B. Kroon, J. Jung, and M. Domański, “Test Model for Immersive Video,” 126th MPEG meeting of ISO/IEC JTC1/SC29/WG11, MPEG/w18470, 2019.

[3] 이광순, 정준영, 신홍창, 서정일, “이머시브미디어를 3DoF+ 비디오 부호화 표준 동향”, 전자통신동향분석, 34권 제6호, p.156-163, 2019

[4] B. Salahieh, S. Bhatia, J. Boyce, “Group-Based TMIV”, 127th MPEG meeting of ISO/IEC JTC1/SC29/WG11, MPEG/m49409, 2019.

[5] 정종범, 이순빈, 김인애, 류일웅, 김성빈, 장동민, 류은석, “3DoF+ 영상에 대한 객관적 영상 품질 평가 기법 성능 비교에 대한 연구”, 한국인터넷정보학회 추계학술대회발표논문집, 제 21권 1호, p.7-8, 2020.