

## MPEG 몰입형 비디오를 위한 클러스터 기반 MV-HEVC 부호화 모드 결정

한창희<sup>1</sup>, 정종범<sup>2</sup>, 류은석<sup>2</sup>성균관대학교 교육학과<sup>1</sup>, 성균관대학교 컴퓨터교육과<sup>2</sup>

chillycorn@skku.edu, uof4949@skku.edu, esryu@skku.edu

Cluster-based MV-HEVC Coding Mode Decision for  
MPEG Immersive VideoChang-Hee Han<sup>1</sup>, Jong-Beom Jeong<sup>2</sup>, Eun-Seok Ryu<sup>2</sup>Department of Education, Sungkyunkwan University (SKKU)<sup>1</sup>Department of Computer Education, Sungkyunkwan University (SKKU)<sup>2</sup>

## 요 약

three degree of freedom (3DoF), three degree of freedom plus (3DoF+), six degree of freedom (6DoF) 등 몰입형 비디오의 높은 몰입감을 제공하기 위해 다중 비디오 영상을 효율적으로 처리하는 기법이 활발히 연구되고 있다. 이를 위해 원본의 몰입형 비디오가 입력되면 기본 시점 영상과 추가 시점 영상에서의 중복을 제거하고 기본 시점(basic view)에서는 보이지 않지만 추가 시점(additional view)에서는 보이는 영역을 추출하는 프루닝 과정이 이뤄지는 부호기에서의 부호화 모드 결정은 매우 중요하다. 본 논문은 test model for immersive video (TMIV)의 모드 중 하나인 MPEG immersive video (MIV) view mode 를 통해 만들어진 프루닝 (pruning) 그래프에서 선택된 시점들을 활용하여 뷰 간 중복성을 제거할 수 있는 효율적인 부호화 구조로 클러스터를 기반으로 병렬적으로 부호화하는 클러스터 기반 정렬 기법을 제안한다. 선택된 시점들을 인덱스 순서에 따라 부호화하는 기존 방법에 비해 제안하는 방법은 peak signal-to-noise ratio (Y-PSNR)에서 평균 3.9%의 BD-rate 절감을 보여주었다. 본 연구는 또한 더 객관적인 품질 측정을 위해 immersive video peak signal-to-noise ratio (IV-PSNR)에 의한 비교 결과도 함께 제공하며, 참조 순서에 맞게 정렬한 프루닝 기반 정렬 기법과의 비교도 함께 제공한다.

## 1. 서론

최근 코로나 19 로 인한 사회적 거리두기가 장기화됨에 따라 이를 극복할 수 있는 가상 현실 (virtual reality; VR) 기술을 활용한 콘텐츠와 3 차원 영상 디스플레이를 사용자의 머리에 장착하여 시청할 수 있는 장치(head mount display; HMD) 시장이 성장하고 있다. 하지만 현재의 VR 영상 서비스는 대부분 모바일 환경에서 작동되기 때문에 그래픽 처리가 된 영상에 비해 저품질로 제공되고 제한된 자유도 문제로 인해 몰입감이 떨어지게 된다.[1] 이에 따라 높은 화질을 지닌 막대한 양의 영상 콘텐츠를 실시간으로 전송하기 위한 VR 기술 개발이 활발하게 진행되고 있다.

몰입형 미디어에 대한 표준화 작업을 위해 등장한 MPEG

immersive (MPEG-I)는 몰입형 비디오의 몰입도를 사용자 시점 자유도(degree of freedom; DOF)에 따라 3DoF, 3DoF+, 6DoF 3 가지 단계로 구분하였고, MPEG immersive video (MIV) 기반의 3DoF+ 비디오 압축 표준을 개발하기 위한 참조 소프트웨어인 test model for immersive video (TMIV)를 제공하였다. 카메라 뷰가 많으면 그것을 가지고 모두 부호화하는 것은 효율성 측면에서 어려움이 있다. 따라서 TMIV 부호기는 픽셀들을 패치 형태로 추출한 후 아틀라스라고 불리는 영상 포맷으로 변환하게 되는데, 이때 중복되는 영역이 가장 많은 영상인 기본 시점 (basic view)와 추가 시점 (additional view) 로 구성된 아틀라스를 전송하는 MIV mode 와 기본 시점에서는 보이지 않지만 추가 시점에서는 보이는 영역을 추출하는 프루닝 과정 없이 특정 시점의 영상들을 아틀라스로 구성하여 보내 중간

시점을 생성하는 MIV view mode 2 가지 모드를 사용할 수 있다.[2]-[5]

본 논문은 그림 1 과 같이 MIV view mode 에서 View Optimizer 와 Pruning Cluster 를 통해 만들어진 프루닝 그래프를 활용하여 뷰 간 중복성을 고려한 코딩 구조를 통해 효율적인 압축을 이끌어낼 수 있는 방안인 클러스터 기반 부호화 기법을 제안한다.

본 논문의 구성은 다음과 같다. 2 절에서는 배경연구로 그룹 기반 TMIV 부호화 기법을 소개한다. 3 절에서는 기존 기법인 시점 인덱스 기반 정렬과 프루닝 기반 정렬, 그리고 제안하는 부호화 방식인 클러스터 기반 정렬에 대해 소개한다. 4 절에서는 제안하는 부호화 방식과 기존 기법들 간의 품질 비교 실험과 복호화 시간 비교 실험을 소개하고 마지막으로 5 절에서는 본 논문의 결론을 맺는다.

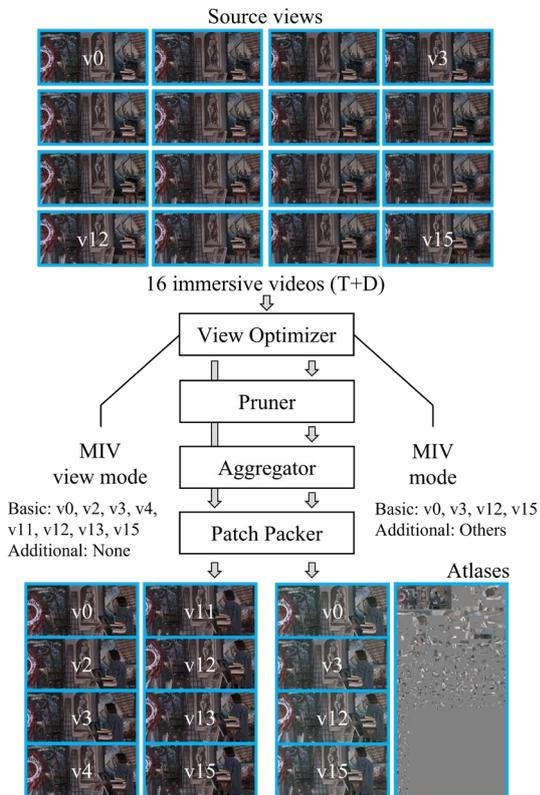


그림 1. 16 개의 예제를 사용한 TMIV 소프트웨어 MIV view mode 와 MIV mode (T: Texture, D: Depth)

## 2. 배경

본 절에서는 프루닝 기반 정렬과 클러스터 기반 정렬을 제안하게 된 배경인 그룹 기반 TMIV 부호화에 대해 소개한다. 그룹 기반 TMIV 부호화 기법은 뷰 간 순서를 고려하지 않은 채 모든 뷰를 한꺼번에 부호화 하는 방식으로 작동하는 기존의

TMIV MIV mode 를 개선하기 위해 제안되었다.[6] 여기서 그룹의 개념은 다수의 소스 뷰 비디오 정보와 카메라 매개변수를 기준으로 공간을 분할하고, 분할된 공간을 기준으로 전처리 과정을 수행하기 위한 것이다. 이러한 그룹 개념을 활용해 그룹 기반 TMIV 부호화는 MIV view mode 로 생성된 아틀라스를 인접한 시점과 인접한 시야각을 기준으로 그룹을 묶어 시점관 일관성을 유지한 채 MV-HEVC 부호화를 함으로써 부호화 처리 속도를 개선하고 렌더링의 품질을 향상시키는 결과를 가져왔다. 그러나 그룹 기반 TMIV 부호화 기법은 참조 순서에 따라 부호화 구조를 결정하는 것이 아니라 단순히 그룹을 기반으로 부호화를 한다는 한계를 지닌다. 또한 실험에 단일 테스트 시퀀스만 사용되어 추가적인 검증이 필요하다는 한계를 지닌다.[7]

## 3. 프루닝 기반 정렬과 클러스터 기반 정렬 기법

TMIV 8.0 소프트웨어에서 MIV view mode 를 선택하면 View Optimizer 와 Pruning Cluster 를 통해 주어지는 뷰 간 중복성을 고려해 기본 시점, 나머지 시점이 주어진다. 그림 2 는 이를 활용하여 시점 인덱스에 기반한 기존 MV-HEVC 부호화 방법인 (a) 시점 인덱스 기반 정렬, (b) 프루닝 기반 정렬, 그리고 본 논문에서 제안하는 (c) 클러스터 기반 정렬을 나타낸다. 이를 사용하여 서로 다른 3 가지 방법의 부호화 실험을 진행하였다.

### 3.1 시점 인덱스 기반 정렬

시점 인덱스 기반 정렬 기법은 프루닝 그래프에서 선택된 시점들의 순서에 따라 부호화를 수행한다. 따라서 이 방식은 기본 시점과 나머지 시점 간의 중복성을 고려하지 않은 채 부호화가 수행된다.

### 3.2 프루닝 기반 정렬

프루닝 기반 정렬 기법은 프루닝 그래프에서 주어진 기본 시점과 나머지 시점을 참조 순서에 맞게 정렬하여 부호화를 수행한다. 이 때 정렬되는 방식은 프루닝 방식에 기초한 것으로 뷰 간 중복성이 어느 정도 고려된 상태에서 부호화된다. 이를 통해 나머지 시점 간 중복성을 어느정도 제거할 수 있다.

### 3.3 클러스터 기반 정렬

본 논문에서 제안하는 두 번째 부호화 방식인 클러스터 기반 정렬을 소개한다. 클러스터 기반 정렬 기법은 기본 시점과 나머지 시점 간의 관계를 고려해 클러스터를 기반으로 병렬적으로 부호화를 수행한다. 이를 통해 뷰 간 중복성을 제거함과 동시에 병렬적으로 부호화를 수행할 수 있게 됨으로써 더 높은 부호화 효율을 가져올 수 있다.

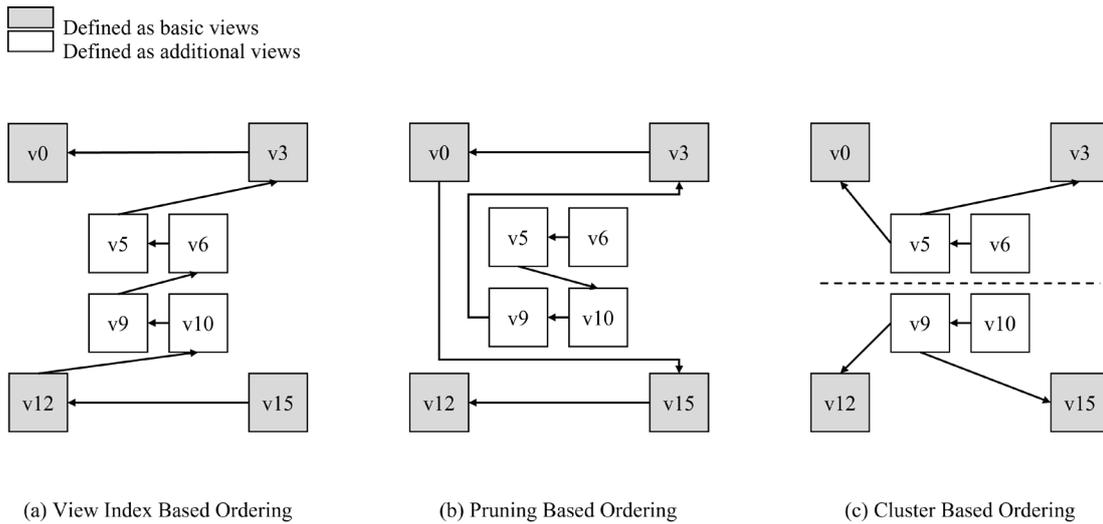


그림 2. 프루닝 그래프를 활용해 제안하는 코딩 구조도

#### 4. 구현 및 실험 결과

본 논문에서는 실험을 위한 mode 결정은 TMIV 8.0 을 사용하였으나 실제 압축 실험은 TMIV 8.0 부터는 atlas 가 4 개로 제한되어 일부 영상이 누락되는 경우가 있고 해상도 제한도 있기 때문에 다시점 영상 압축에 더 적합한 TMIV 3.0 을 사용하였다. 또한 MV-HEVC 참조 소프트웨어인 MV-HEVC test model (HTM) 13.0 을 사용하였으며, MIV 에 대한 실험조건인 common test conditions (CTC)를 준수하였다.[8] 프루닝 기반 정렬과 클러스터 기반 정렬을 사용한 방법과 기존 부호화 방식인 시점 인덱스 기반 정렬을 비교하였으며 테스트 시퀀스는 *Painter* (2048 \* 1088), *Frog* (1920 \* 1080) 2 가지가 사용되었다. 품질 평가 기법은 immersive video peak signal-to-noise ratio (IV-PSNR), (Y-PSNR) 2 가지를 사용하였다.

Sequence name	View Index Based	Pruning Based	Cluster Based
Technicolor Painter	0.00%	-0.08%	<b>-4.48%</b>
Intel Frog	0.00%	-2.44%	1.66%
<b>Average</b>	<b>0.00%</b>	<b>-1.26%</b>	<b>-1.41%</b>

표 1. 기존 기법 대비 제안하는 기법의 IV-PSNR BD-rates

Sequence name	View Index Based	Pruning Based	Cluster Based
Technicolor Painter	0.00%	-0.78%	<b>-8.33%</b>
Intel Frog	0.00%	-3.36%	0.56%
<b>Average</b>	<b>0.00%</b>	<b>-2.07%</b>	<b>-3.89%</b>

표 2. 기존 기법 대비 제안하는 기법의 Y-PSNR BD rates

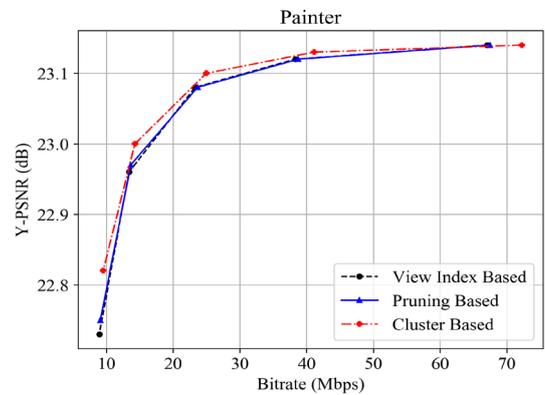


그림 3. 고대역폭에서의 RD-curves (*Painter*)

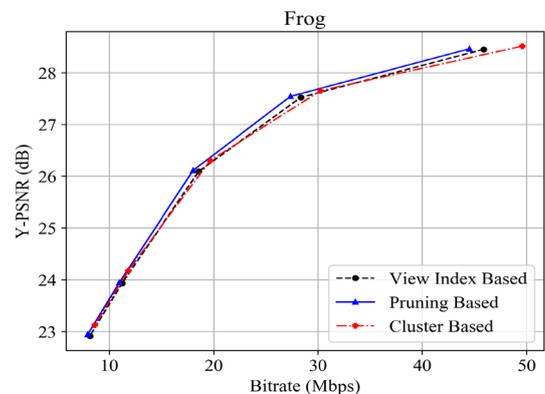


그림 4. 고대역폭에서의 RD-curves (*Frog*)

본 연구는 테스트 시퀀스마다 다른 5 가지 quantization parameter (QP) 값으로 영상을 부호화 하여 QP 1 에서 4 까지는 고대역폭 영상, QP 2 에서 5 까지는 저대역폭 영상으로 정의하여 두 대역폭에 대해 평가하는 MPEG-I 의 CTC 방식을 준수하여

진행되었다. 표 1 과 표 2 는 제안하는 클러스터 기반 정렬 방식이 기존의 부호화 방식인 시점 기반 정렬과 프루닝 기반 정렬보다 고대역폭, 저대역폭 영상 모두에서 평균적으로 BD-rate 를 절약할 수 있음을 보여준다. 제안하는 부호화 방식은 IV-PSNR 과 Y-PSNR 에서 평균 1.4%, 3.9%의 BD-rate 를 절약할 수 있다. 그림 3 과 그림 4 는 *Painter* 와 *Frog* 테스트 시퀀스에서의 기존 부호화 방식과 제안하는 부호화 방식의 rate-distortion (RD) curve 를 보여준다. 이를 통해 제안하는 부호화 방식이 기존의 부호화 방식들보다 평균적으로 더 좋은 품질의 영상을 제공할 수 있음을 알 수 있다. 또한 그림 5 와 그림 6 은 각 테스트 시퀀스에서 제안하는 방식이 기존 부호화 방식보다 복호화 시간을 평균 0.9% 절약할 수 있음을 보여준다. 이때 제안하는 방식이 병렬적으로 부호화, 복호화 될 수 있다는 점을 고려하면 기존 방식보다 복호화 시간도 더 절약할 수 있음을 알 수 있다.

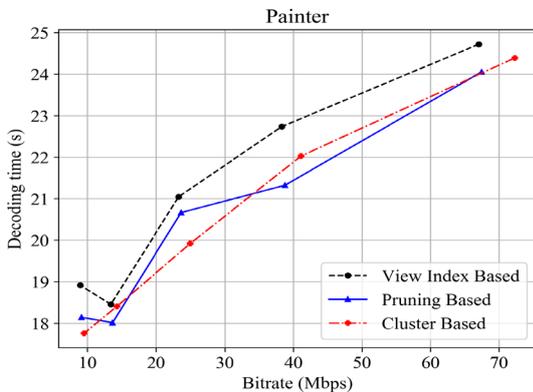


그림 5. 복호화 시간 비교 (*Painter*)

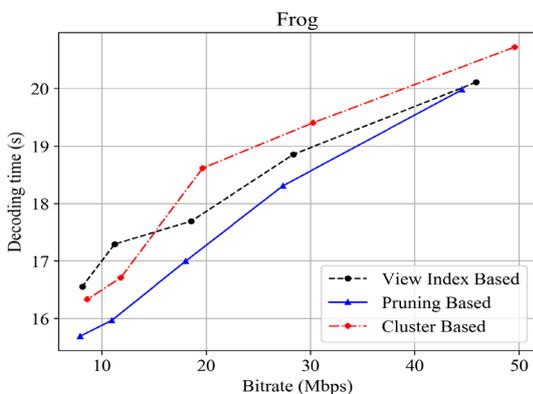


그림 6. 복호화 시간 비교 (*Frog*)

## 5. 결론

본 논문은 3Dof, 3Dof+, 6Dof 등 몰입형 비디오를 효율적으로 처리하기 위해 TMIV 에서 압축 효율을 높이고 복호화 시간을 줄일 수 있는 최적의 코드 구조로 프루닝 기반 정렬과 클러스터 기반 정렬을 제안한다. 제안하는 방식은 MIV view

mode 를 통한 기존의 부호화 방식인 시점 기반 정렬과 비교했을 때 IV-PSNR 에서 평균 1.4%, Y-PSNR 에서 평균 3.9%의 BD-rate 를 절약할 수 있음을 확인하였다. 또한 복호화 시간도 기존 방식보다 0.9%를 절약할 수 있음을 확인하였다. 향후 연구에서는 다양한 시야각을 가진 고해상도 테스트 시퀀스를 가지고 부호화 실험을 진행하여 더 많은 상황을 고려할 예정이다.

## Acknowledgement

이 논문은 2021 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2020-0-00231-002, 5G 엣지클라우드 기반 VR·AR 저지연 스트리밍 기술 개발).

## 참고문헌

- [1] 호요성, "MPEG-I 표준과 360 도 비디오 콘텐츠 생성", 전자공학회지 제 44 권 제 8 호, 2017 년, pp.52-57
- [2] 박현수, 박상효, 강제원, "6 자유도 전방위 몰입형 비디오의 압축 코덱 개발 및 성능분석", 방송공학회논문지 제 24 권 제 6 호, 2019.Ja
- [3] 이광순, 서정일, "MPEG 이머시브 비디오 표준기술 동향과 전망", 정보통신기획평가원 주간기술동향 1969 호, 2020 년
- [4] 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.
- [5] 김현호, 김재곤, 이광순, "MPEG-I 3DOF+ 비디오 부호화 표준 기술", 한국정보통신기술협회 TTA 저널 185 호, 2019년, pp.17-25
- [6] Inae Kim, Jong-Beom Jeong, Soonbin Lee, Yeongil Ryu, Eun-Seok Ryu, "Group-Based TMIV Encoding for Immersive Video using MV-HEVC", Korean Society for Internet Information (KSII) Fall Conference, Oct. 30-31, 2020.
- [7] B. Salahieh, S. Bhatia, J. Boyce, "Group-Based TMIV", 127th MPEG meeting of ISO/IEC JTC1/SC29/WG11, MPEG/m49409, 2019.
- [8] Joel Jung, Bart Kroon, Jill Boyce, "Common Test Conditions for Immersive Video", ISO/IEC JTC1/SC29/WG11, MPEG/w18789, 2019.