

# Viewport-Adaptive Immersive Video Streaming System

정종범 / 성균관대학교 Multimedia Computing Systems Lab.

몰입형 콘텐츠를 다루는 차세대 미디어 환경에서는 가상현실을 지원하는 head-mounted display (HMD)를 쓰고 임의의 위치나 각도에서 자유롭게 영상을 시청할 수 있는 경험을 제공해야 한다. 이 때 고품질 실사 영상을 제공하기 위해 다수의 위치에서 동시에 취득된 영상을 사용할 수 있다. 그러나 다수의 카메라가 촬영한 고품질 색상 및 깊이 정보를 포함하는 몰입형 영상은 매우 큰 저장 공간을 요구해, 스트리밍 구현 시 부호화, 전송, 복호화, 렌더링 전반에 대한 많은 기술적 제약이 뒤따른다.

영상 압축 표준화 단체인 moving picture experts group (MPEG)은 high-efficiency video coding (HEVC/H.265), versatile video coding (VVC/H.266) 등의 기준 2차원 영상 압축 표준을 최대한 활용하면서, 추가 메타데이터를 통해 임의 시점 합성 (novel view synthesis)을 지원하는 MPEG immersive video (MIV, ISO/IEC 23090-12) 표준을 개발하였다. 이를 통해 다수의 영상을 3차원으로 재구성하여 영상 간 중복성을 제거하고, 잔차 정보를 2차원으로 재구성하여 2차원 영상 압축 도구를 통해 압축 및 3차원 공간 복원에 필요한 메타데

이터와 함께 전송하여 비트율을 종래 시스템 대비 크게 절감한다. 또한 전송 및 복호화해야 할 영상 개수를 크게 줄여 복호기 단에서의 복잡도를 줄이고, 이후 복호화된 영상을 기반으로 3차원 공간을 복원하여 사용자가 요청한 임의 시점을 생성 및 렌더링할 수 있다.

MIV 기반의 몰입형 영상 스트리밍 시스템 구현 시 원본 영상의 해상도가 크고 개수가 많을 경우 전송 및 복호화 측면에서의 부담이 크다. 따라서 사용자가 요청한 임의 시점을 검출하고, 해당 영역만 효율적으로 추출 및 전송/복호화하여 계산 복잡도를 줄이는 연구가 활발하게 진행 중이다. 본 연구는 이러한 배경에서 사용자 시점 적응적 몰입형 영상 스트리밍 시스템을 제안하고, 크게 세 가지 핵심 연구 주제를 중심으로 고품질, 저지연 스트리밍을 구현하는 방법을 소개한다.

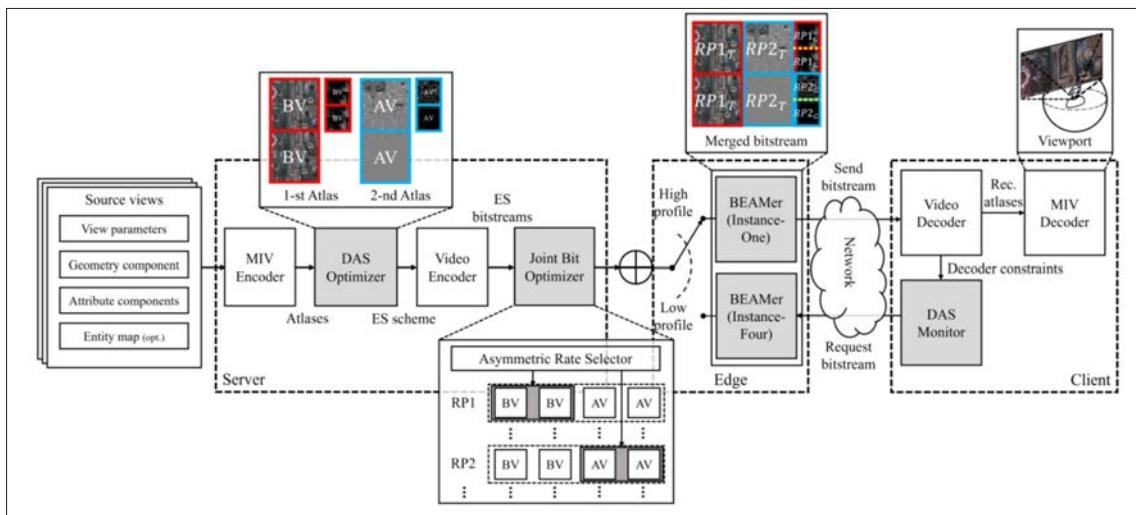
첫 번째 연구 주제는 종단간 몰입형 영상 스트리밍 시스템 구현을 다룬다. 2차원 형태로 촬영된 영상이더라도 거리를 나타내는 깊이 정보가 존재하면 영상을 3차원 공간으로 투영할 수 있는데, 이를 통해 영상 간 겹치는 영

역을 제거 가능하다. 상기 과정은 프루닝(pruning)이라 정의되며, 시점 간 중복성 제거의 기준이 되는 기본 시점과 중복성이 제거되는 추가 시점을 분류하여 프루닝을 시행한다. 중복성은 3차원 공간에서 판별하더라도, 이를 2차원 영상에 투여하여 여전히 비디오 압축 기술을 사용 가능하다. 한편, 프루닝만으로는 영상의 개수를 줄일 수 없기 때문에, 잔차 정보들만을 추출 및 병합하는 패킹(packing) 과정을 통해 전송할 영상의 개수를 크게 줄일 수 있다. 이때 각 잔차 정보를 다시 원본 영상으로 복원하기 위한 메타데이터를 별도 저장하며, 이를 통해 원본 영상 복원 및 임의 시점 생성이 가능하다. 본 기술을 통해 기존의 HEVC simulcast 시스템 대비 부호화 시간을 83.81%, 픽셀 수를 82.73%, 비트율을 27.45% 절감할 수 있음을 확인하였다.

두 번째 연구 주제는 사용자 시점 적응적 타일 기반 스트리밍 기술에 대해 다룬다. 기존의 단일 360도 영상에서의 사용자 시점을 검출하는 기술을 확장하여, 다수의 영

상으로 구성된 몰입형 영상의 깊이 정보를 이용해 각 시점별 사용자 시점을 검출하고, 이를 통해 사용자 관심 영역에 해당하는 타일을 검출한다. 상기 타일들을 고품질로 전송하되, 갑작스러운 사용자 시점 변경에 대비하여 저화질 전체 영역 영상은 사용자 시점과 관계없이 전송한다. 또한 개별 타일별로 비트스트림을 생성하여 전송 및 복호화도 가능하지만, 이로 인해 발생하는 부하를 줄이기 위해 본 연구는 다수의 사용자 시점 타일을 추출하여 단일 비트스트림을 생성하는 기술을 추가하였다. 상기 연구를 MIV 기반 스트리밍 시스템에 적용하면, 사용자 시점은 고품질로 감상 가능하면서도 그 외 시점은 저품질의 영상이 표현하기 때문에 비트율을 크게 절약 가능하다. 또한 단일 비트스트림 생성을 통해 타일 개수와 상관없이 복호화 시간 및 메모리 사용량을 유지 가능하다. 본 연구를 통해 기존 기술 대비 12.04%의 비트율 절감, 55.51%의 복호화 시간 절감을 확인하였다.

세 번째 연구 주제는 서브피쳐 기반 복호기 제어 및



<그림 1> 제안하는 복호기 적응적 타일 추출/병합 및 비대칭 품질 할당 시스템

## 졸업논문 소개

혼합 품질 할당 기법에 대해 다룬다. <그림 1>은 제안하는 시스템을 도시한다. MIV에서 사용하는 기본 시점과 추가 시점의 특성을 구분하기 위해 영상을 서브피쳐(subpicture)로 분할하여 상기 시점들을 구분한다. 분할 시 서브피쳐를 다시 한 개 내지는 두 개의 피쳐로 병합할 수 있도록 하며, 서브피쳐 내 시점 정보에 따라 비트율을 차등 할당한다. 상기 서브피쳐를 담은 비트스트림을 전송 시, 단말의 복호기가 제한된 수의 인스턴스를 지원할 경우 서브피쳐를 병합 및 전송한다. 실험 결과 서브피쳐를 도입하여 부호화 시간은 3.45% 증가하나, 복호화 시간은 26.83% 감소함을 확인하였다. 또한 상기 서브피쳐 단

위 비트율 할당을 통해 비트율을 19.52% 절감 가능함을 확인하였다.

제안한 연구들은 MIV와 같은 몰입형 영상 압축 표준, HEVC, VVC 등의 영상 압축 표준과 정합되어 사용될 수 있으며, 표준 기술과 호환되도록 설계하여 기존 하드웨어를 그대로 사용하면서도 성능 향상이 가능하다. 또한 사용자 중심의 설계를 통해 동일 비트율에서 품질 향상을 이루면서도 처리 시간을 줄여 경량화된 단말에서의 적용 역시 가능하다. 따라서 제안한 연구들은 확장 연구를 통해 몰입형 영상 스트리밍 시스템에 도입 가능할 것으로 기대된다.

### 정종범



- 2018년 8월 : 가천대학교 컴퓨터공학과 학사
- 2020년 1월 ~ 2020년 3월 : University of California, Santa Barbara 방문연구원
- 2021년 8월 ~ 2022년 1월 : Purdue University 방문연구원
- 2022년 9월 ~ 2023년 8월 : 성균관대학교 글로벌융합학부 강사
- 2023년 9월 ~ 2024년 8월 : 성균관대학교 실감미디어공학과 강사
- 2025년 2월 : 성균관대학교 컴퓨터교육학과 박사
- 2025년 2월 ~ 현재 : 한국전자통신연구원 미디어부호화연구실 박사후연구원
- ORCID : <https://orcid.org/0000-0002-7356-5753>
- 주관심분야 : Immersive video processing, learning-based computer vision, multimedia applications