

# 360 VR 영상의 프로젝션 포맷 및 성능 평가 방식

\*박성환 \*\*김규현

경희대학교

\* [gocheenee@khu.ac.kr](mailto:gocheenee@khu.ac.kr) \*\* [kyuheonkim@khu.ac.kr](mailto:kyuheonkim@khu.ac.kr)

## Projection format and quality metrics of 360 video

\*Park, Seong-Hwan \*\*Kim, Kyu-Heon

KyungHee University

### 요약

최근 사용자에게 더욱 몰입감 있는 콘텐츠를 제공하기 위한 기술에 대한 관심이 증가하고 있으며 그 중 가장 대표적인 것이 360 VR 영상이라고 할 수 있다. 미디어 표준화 단체인 MPEG(Moving Picture Experts Group)에서는 MPEG-I(Immersive) 차세대 프로젝트 그룹을 이용하여 이러한 움직임에 대응하고 있다. MPEG-I는 2021년 말 6DoF VR 영상을 목표로 8개의 파트가 표준화를 진행 중이다. 360 VR 영상의 경우 획득시 영상의 픽셀들이 3D 공간 상에 존재하게 되는데, 이를 처리 및 출력하기 위해서는 2D 영상으로 전환이 필요하며 이 때 사용되는 것이 Projection format이다. 현재 JVET(Joint Video Exploration Team)에서는 3D에서 2D로 전환이 이루어질 때 손실을 최소화 하기 위한 Projection format들에 대한 연구가 이루어지고 있다. 본 논문에서는 현재까지 제안된 다양한 Projection format들에 대하여 소개하고 이에 대한 성능 측정 방식에 대하여 소개한다.

### 1. 서론

미디어 기술이 발달함에 따라 사용자에게 더욱 몰입감 있는 콘텐츠를 제공하기 위한 기술들이 개발되고 있다. 그 중 가장 대표적인 것이 360 VR 영상이라고 할 수 있다. 이러한 요구에 따라 국제 표준화 단체인 MPEG(Moving Picture Experts Group)에서는 몰입형 미디어에 대한 차세대 프로젝트인 MPEG-I[1] 그룹을 제정하여 해당 기술에 대한 표준화를 진행 중이다. 또한 MPEG 과 VCEG(Video Coding Experts Group)은 JVET(Joint Video Exploration Team)이라는 그룹을 통해 8K영상, 360 VR 영상 등의 차세대 비디오에 대한 코딩 표준을 준비하고 있다. 현재의 인코더와 디코더는 2D 이미지 및 영상을 대상으로 하기 때문에 3D로 구성된 360 영상의 경우 처리가 불가능하다. 따라서 3D 영상과 2D 영상 간의 전환이 요구되며 이 때 프로젝션 포맷이 사용되게 된다. 프로젝션 포맷의 변환으로도 코딩 성능에 영향을 미칠 수 있기 때문에 해당 분야에 대한 연구가 활발히 진행 중이다.[2]

본 논문은 2장에서 MPEG-I에 대하여 설명하고, 3장에서는 Projection format에 대해 설명한다. 4장에서는 3장에서 설명한 Projection format의 성능 측정 방식에 대하여 설명하고 5장에서는 결론을 맺는다.

### 2. MPEG-I

MPEG-I는 국제 표준화 단체인 MPEG에서 진행 중인 몰입형 미디어에 대한 표준화 프로젝트이다. 이 프로젝트는 차세대 몰입형 비디

오 및 오디오 시스템 전체를 대상으로 하는 표준 기술들로 구성되어 있으며 6축 자유도를 갖는 몰입형 미디어를 목표로 표준화를 진행 중이다. MPEG-I는 그림 1에서 보여지는 것과 같이 2단계로 나뉘어 있으며 8개의 파트로 구성되어 있으며 각 파트는 몰입형 비디오, 오디오, 파일 포맷, 압축, 전송 등으로 구분되어 표준화가 진행되고 있다.

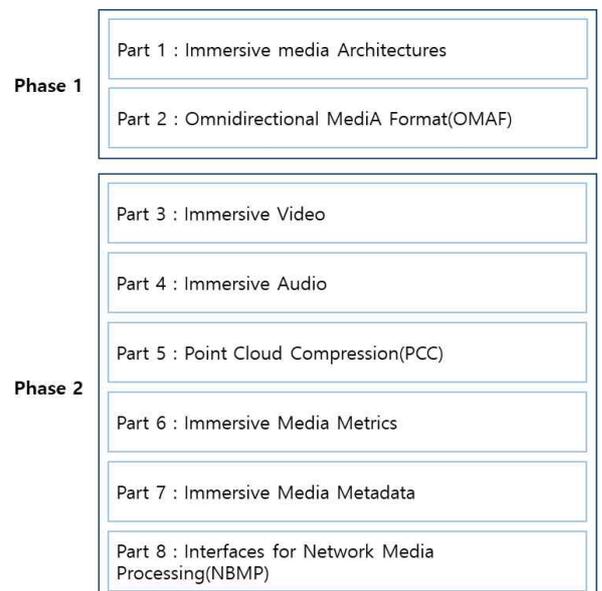


그림 1 MPEG-I 단계 및 파트 구성

### 3. Projection format

360 VR 비디오는 그림 2와 같은 절차를 거쳐 소비되게 된다. 360 VR 전용 카메라를 통해 습득된 영상은 해당 카메라에 지정된 프로젝션 포맷으로 저장되게 된다. 이후 다운샘플링 및 프로젝션 포맷 변경이 이루어지게 되고 이 결과물을 이용하여 인코딩 / 디코딩 작업을 수행한다. 디코딩 된 비디오는 다시 프로젝션 포맷 변경 및 업샘플링 과정을 거친 뒤 원래의 해상도로 복원되어 출력되게 된다. 360 VR 비디오는 전방향에 대한 정보를 제공하지만, 사용자가 실제 소비하는 뷰포트는 하나의 방향만을 가지기 때문에 이를 이용하여 객관적 성능평가가 가능해진다.

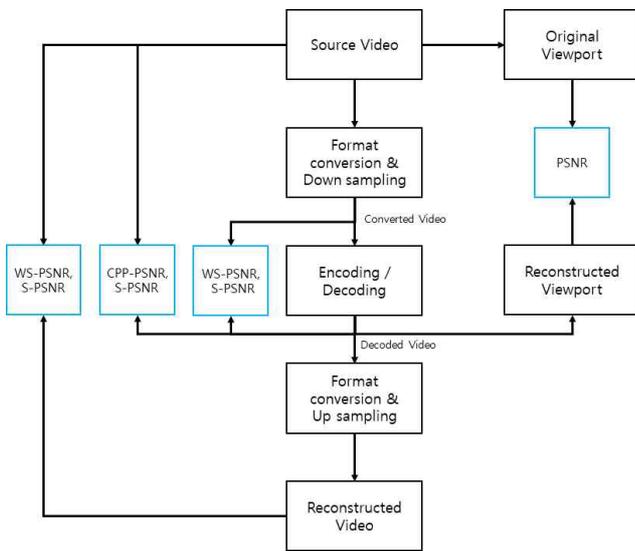


그림 2 360 VR 비디오의 프로세싱 체인

포맷 변경시 사용되는 프로젝션 포맷에 따라서 3D에서 2D로 변환될 때 왜곡되는 정도의 차이가 있기 때문에 인코딩/디코딩을 포함한 후처리 작업에 왜곡에 의한 성능 차이가 발생하게 된다. 따라서 프로젝션 포맷들에 대한 다양한 연구가 이루어지고 있으며 다음과 같은 프로젝션 포맷들이 논의되고 있다.

CubeMap Projection (CMP)		ERP와 더불어 가장 많이 쓰이는 방식. 정육면체를 모델링하여 6개의 사각형을 하나의 평면에 매칭시키는 방식
Adjusted CMP (ACP), Equi-angular CMP(EAC), Hybrid equi-angular(CMHEC)		CMP 방식을 기반으로 forward / backward 조정함수를 추가한 것
OctaHedron Projection (OHP)		정팔면체를 모델링하여 8개의 삼각형을 하나의 평면에 매칭시키는 방식
IcoSahedron Projection(ISP)		정 20면체를 모델링하여 20개의 삼각형을 하나의 평면에 매칭시킨다.
Segmented Shpere Projection(SSP)		위도 -45~45는 ERP와 동일하게 매칭하고, 극점부분에 해당하는 위도 -45 이하, 45 이상 부분은 원 형태로 매칭시켜 6개의 정사각형으로 만드는 방식
Rotated Sphere Projection(RSP)		테니스 공의 구성 방식과 유사하게 구를 2개의 면으로 맺이시킨 방식

표 1 Projection fomats

Projection Format	Packing	설명
EquiRectangular Projection (ERP)		가장 대중적으로 쓰이는 방식. 구를 모델링 하여 경도를 x축으로, 위도를 y도로 매칭시키는 방식
Adjusted equal-area Projection(AEP)		압축 효율과 극점 왜곡을 향상시키기 위해 위도 및 경도의 계산식을 수정한 방식

### 4. 360 VR 비디오의 성능 평가 방식

비디오 코딩의 객관적 성능 평가 방식으로는 PSNR(Peak Signal to Noise Ratio) 방식이 대중적으로 사용되고 있다. 하지만 360 VR 비디오의 경우 3D로 구성되어 있기 때문에 2D를 대상으로 하는 PSNR 방식은 적합하지 않다. 따라서 JVET에서는 다음과 같은 성능 평가 방식들을 사용하고 있다.

#### 1) WS-PSNR(Weighted to Spherically uniform PSNR)

WS-PSNR[3]은 원본 영상과 테스트 영상과의 평균 오차를 고려한다는 점에서 PSNR과 유사하지만 3D 상에서 샘플링을 할 때 발생하는 오차를 고려하여 가중치를 주는 방식으로 계산된다. 3D 상의 정보를 2D 상의 프로젝션 포맷으로 변환 할때에 발생하는 왜곡을 가중치

를 통해 완화시켜주는 방식이며 이때 사용되는 수식은 다음과 같다.

$$w(i, j)_{ERP} = \cos \frac{(j + 0.5 - N/2)\pi}{N}$$

### 2) S-PSNR(Spherical PSNR)

S-PSNR[4]은 원본 영상과 테스트 영상으로부터 각각 생성된 3D 상의 위치를 대상으로 PSNR을 측정하는 방식이다. 이 방식의 경우 두 개의 영상이 서로 다른 해상도나 프로젝션 포맷을 가지더라도 비교가 가능하다는 장점이 있다. 3D 상에서 해당 이미지의 인접위치를 계산할 때 가장 가까운 이웃점을 사용하는 방식과 interpolation을 사용하는 방식에 따라 각각 S-PSNR-NN, S-PSNR-I로 나뉜다.

### 3) CPP-PSNR(Crasters Parabolic Projection PSNR)

CPP-PSNR[5] 방식은 원본 영상 및 테스트 영상을 Craster Parabolic 프로젝션 포맷으로 변환한 뒤 PSNR을 측정하는 방식이다. Craster Parabolic의 경우 구를 2D 상으로 펼쳤을 때 왜곡을 최소화하는 방식으로, S-PSNR에서 사용되는 방식과 비슷하기 때문에 두 개의 영상이 서로 다른 해상도나 프로젝션 포맷을 가지더라도 비교가 가능하다는 장점을 갖는다.

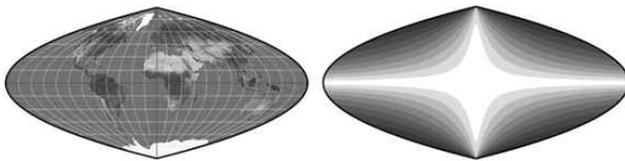


그림 9 Craster Parabolic

## 5. 결론

미디어 기술이 발달함에 따라 사용자에게 더욱 몰입감 있는 콘텐츠를 제공하기 위한 기술 표준화가 활발히 진행 중이다. 현재까지 완료된 360 VR 비디오의 프로젝션 포맷 및 성능 평가 방식들에 대해 알아 보았고 이를 기반으로 프로젝션 포맷 기반 영상 왜곡 최소화에 대한 연구가 진행될 예정이다.

\* 본 논문은 과학기술정보통신부 및 정보통신기술진흥센터의 출연금 등으로 수행하고 있는 한국전자통신연구원의 360 영상 왜곡 최소화를 위한 프로젝션 포맷 연구의 위탁연구과제의 연구결과입니다.

### 참고 문헌

- [1] "ISO/IEC 23090 Coded Representation of Immersive Media"
- [2] "ISO/IEC Algorithm descriptions of projection format conversion and video quality metrics in 360Lib version 9", 2019
- [3] "AHG8: WS-PSNR for 360 video objectivequality evaluation", Joint Video Exploration Team of ITU-T SG16WP3 and ISO/IEC JTCl/SC29/WG11, JVET-D0040, 2016
- [4] "A Framework to Evaluate Omnidirectional Video Coding Schemes," 2015 IEEE International Symposium on Mixed and

Augmented Reality, 2015

[5] "AHG8: Suggested testing procedure for 360-degree video," JointVideo Exploration Team of ITU-T SG16 WP3 and ISO/IECJTCl/SC29/WG11, 2016