

MPEG-I 비디오 표준화 동향

□ 최기호, 방건* / 삼성전자, *ETRI

I. 서론

MPEG에서는 2016년 10월 중국 청두 회의에서 MPEG-I(Immersive) 표준화를 위한 공식 활동을 시작하였으며, 본 표준화에서는 새로운 미디어 서비스로 대두되고 있는 VR/AR 서비스에 활용될 수 있는 몰입형 오디오/비디오에 대한 표준화를 주요 목표로 삼았다[1].

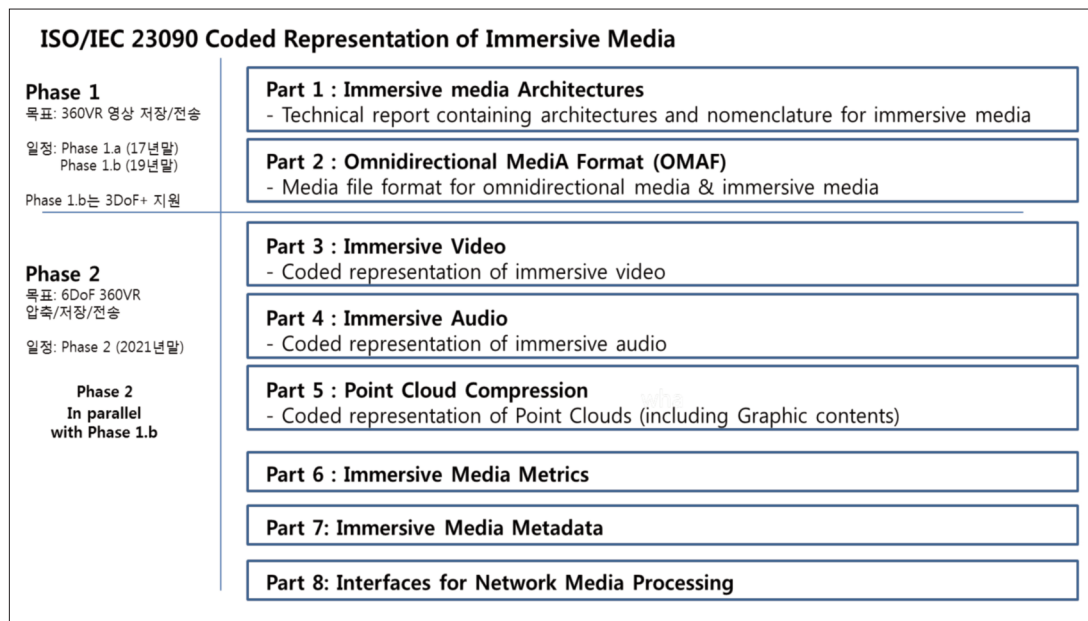
몰입형 비디오 표준은 현재 두 개의 큰 줄기로 진행되고 있다. 먼저는 고화질 Standard Dynamic Range(SDR)/High Dynamic Range(HDR)/360도 영상 압축 솔루션을 제공하기 위한 압축 표준 제정이고, 다른 하나는 몰입형 미디어의 효과적 기능을 제공하기 위한 기능성 기반 전송 표준 제정이다. 압축 표준은 MPEG과 ITU-T에서 함께 만든 Joint Video Expert Teams(JVET)에서 Versatile Video Coding Standard(VVC)라는 프로젝트 명으로 진행

중이며, 기능성 기반 전송 표준은 MPEG-I에서 3 Degrees of Freedom(Dof)(+)/6Dof으로 진행되고 있다.

본 논문에서는 이와 관련된 MPEG-I의 개요, 차세대 표준 비디오 코딩 표준인 VVC에서 논의되고 있는 360영상 관련 표준화 내용, 그리고 MPEG-I 6DoF에 대한 표준 기술 동향과 진행 상황에 대해 각각 소개하고자 한다.

1. MPEG-I 개요

MPEG-I (ISO/IEC 23090 Coded Representation of Immersive Media)는 몰입형 미디어에 대한 MPEG 차세대 프로젝트를 말한다. 비디오, 오디오에서 시스템 요소에 이르기까지 전체 콘텐츠 에코시스템을 대상으로 하는 표준 기술로 구성된 프로젝트이며, 360 VR(Virtual Reality) 비디오와



〈그림 1〉 MPEG-I 단계별 Part 구성

MR(Mixture Reality) 및 6DoF 관련 콘텐츠를 포함한 몰입형 미디어 전체를 고려 대상으로 한다.

1) MPEG-I 표준화 기술 동향

MPEG-I는 총 2단계의 Phase와 8개의 Part로 정의되어 있고, 각 Part마다 서로 다른 표준으로 구성되어 있다. Phase 1에서는 360도 VR 영상 시청 시나리오를 목표로 진행하며, Phase 2는 6축 자유도를 갖을 수 있는 6DoF VR 영상도 포함을 한다. 이를 위해 각 Part 별로 표준화를 착수 중에 있거나 일부는 이미 진행하고 있다.

■ Part 1: Immersive Media Architectures

MPEG-I Part 1은 MPEG-I 요구사항이 요약된 몰입형 미디어의 주요 사용 시나리오 및 아키텍처를 정의하며 MPEG-I의 전체 아키텍처를 포함

하는 기술 보고서로 구성된다. 기술 보고서에는 또한 MPEG-I의 Part에서 사용되는 몰입형 미디어와 관련된 주요 용어의 정의가 포함되어 있다.

■ Part 2: Omnidirectional Media Format(OMAF)

MPEG-I Part 2 표준인 OMAF는 MPEG-I의 시스템 표준 규격이며 몰입형 미디어를 위한 ISO Base Media File Format(ISOBMFF)의 확장 미디어 파일 포맷 규격이다. OMAF 버전 1은 전방향 360도 비디오의 저장 및 전송을 목표로 하였으며, 향후 버전에는 3DoF+ 및 6DoF 시나리오를 지원하기 위해 표준이 확장될 예정이다.

■ Part 3: Immersive Video

MPEG-I Part 3는 MPEG-I 몰입형 미디어 서비스에 중요한 비디오 규격을 정의한다. 이를 위해

High Efficiency Video Coding(HEVC) 표준 이후로 JVET을 구성하여 차기 영상 압축 기술을 위한 차기 코덱 표준 착수를 추진하고 있다.

■ Part 4: Immersive Audio

MPEG-I Part 4는 몰입형 오디오를 위한 규격을 정의한다. 이 부분에는 특히 6DoF 콘텐츠 표현을 위해 필요한 특정 오디오 압축 및 재생 관련 표준이 포함된다.

■ Part 5: Point Cloud Compression

MPEG-I Part 5는 향후 몰입형 콘텐츠를 구현하기 위한 핵심 기술로 예상되는 포인트 클라우드의 압축 관련 표준을 정의한다. 포인트 클라우드는 영상의 색을 담는 정보와 공간상의 좌표를 담는 정보를 가진 3차원 영상 표현 방법으로서 차기 6DoF 및 Mixture VR를 가능하게 하기 위한 기술로 알려져 있다.

■ Part 6: Immersive Media Metrics

MPEG-I Part 6에는 몰입형 미디어 생태계 전반에 걸친 사용자 경험의 품질을 정량화하고 평가할 수 있는 미디어 메트릭 목록이 포함되어 있다.

■ Part 7: Immersive Media Metadata

MPEG-I Part 7은 MPEG-I의 다른 Part에서 새롭게 정의된 메타 데이터를 총괄하는 규격 파트이다.

■ Part 8: Interfaces for Network Media

Processing

MPEG-I Part 8에서는 몰입형 미디어의 네트워크 연동을 위한 인터페이스를 정의한다. 특히

6DoF 콘텐츠의 높은 데이터량을 필요로 하는 몰입형 미디어 서비스를 위해 클라우드와 모바일 단말간에 복잡한 계산을 나눠 처리할 수 있도록 네트워크 인터페이스를 정의하여 전송 데이터 감소 및 단말의 배터리 런타임 향상을 가능하게 한다.

2) MPEG-I 표준화 진행 상황

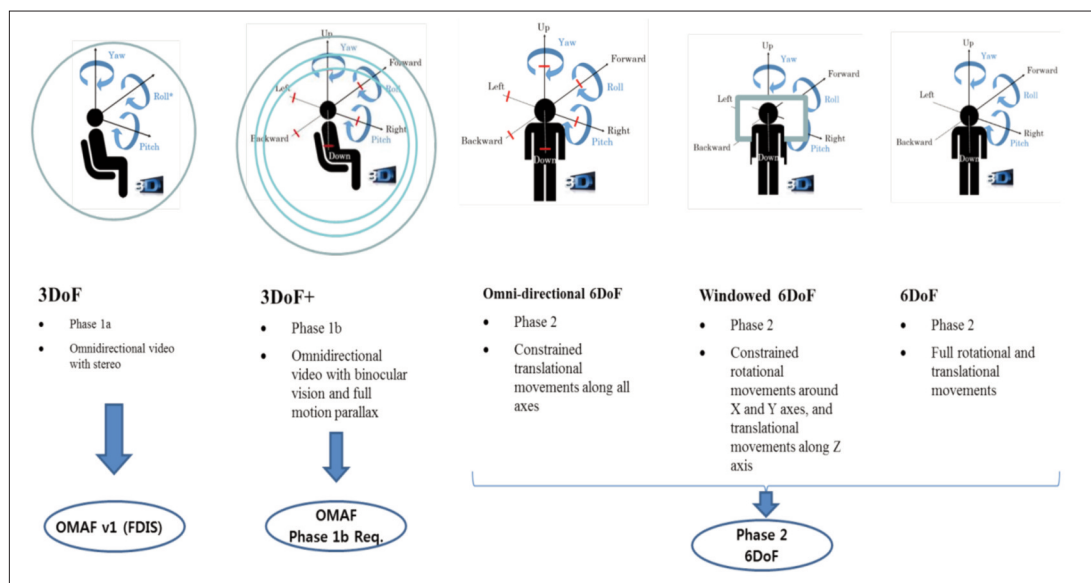
MPEG-I 타임 라인은 3단계로 진행된다. Phase 1a 요구사항은 360 전방향 비디오를 가능하게 하는 기술을 지원하는 반면 Phase 1b 및 Phase 2 요구사항은 각각 3DoF+ 및 6DoF 미디어 콘텐츠를 지원하는 기술을 목표로 한다. 현재 진행 상황은 Phase 1a는 완료되었고, Phase 1b는 요구사항 정의 중이며, Phase 2는 사용 시나리오 및 요구사항 논의 중에 있다.

〈표 1〉 MPEG-I 진행 상황

단계	목표 및 현황	완료 일정
Phase 1a	OMAF v1 (360도 VR)	17년말 (완료)
Phase 1b	OMAF v2 (3DoF+) 요구사항 논의 중	19년말
Phase 2	6DoF 사용 시나리오 및 요구사항 논의 중	21년말

3) MPEG-I 향후 전망 및 일정

MPEG-I에서 정의하는 사용 시나리오는 〈그림 2〉와 같이 다양한 몰입 환경을 포함한다. 3DoF의 360도 전방향 비디오, 3DoF+ 전방향 비디오를 시작으로, 네이게이션이 가능한 콘텐츠, Social VR 및 6DoF 서비스를 위하여 표준화가 진행될 것이며 MPEG-I는 차세대 몰입형 미디어에 대한 산업계 표준이 될 전망이다.



〈그림 2〉 MPEG-I 몰입감 단계 정의

II. 본 론

본 섹션에서는 MPEG-I 중 Part 3에 해당하는 MPEG-I 몰입형 미디어 서비스에 중요한 비디오 규격에 관하여 구체적으로 표준화 내용을 소개하고자 한다. 먼저는 HEVC 표준 이후 차기 영상 압축 기술을 위한 차기 코덱 표준화인 VVC이며, 다른 하나는 6DoF 서비스를 위한 표준화 제정이다.

1. VVC 360도 영상 코덱 표준화

2015년에 ISO산하 SC29/WG11 MPEG과 ITU산하 Q6/SG16 VCEG은 차세대 비디오 코딩 표준을 위한 스터디 그룹을 만들기로 합의한다. HEVC 이후 새로운 비디오 코딩을 준비하는 그룹인 JVET은 전통적인 2D영상을 포함한 새롭게 등장한 영상들 또한 차세대 표준의 대상으로 삼고자 하였다.

HEVC에서 커버하지 못하는 8K급 영상, 최근 새롭게 등장한 High Dynamic Range(HDR)영상, 그리고 Virtual Reality(VR)를 지원하기 위해서 등장한 360도 비디오 영상이다[2].

기존 MPEG 및 VCEG은 2D 이미지 및 영상을 코덱 표준의 대상으로 삼는다. 코덱 표준은 디코더의 선택스 및 디코딩 과정을 표준의 의무조항(normative)으로 삼기에, 인코딩 과정 및 선처리 과정(pre-processing), 후처리 과정(post-processing), 렌더링 등은 표준에서 제약하지 않고 있다. 하지만 360도 영상의 경우, 3D 영상을 2D 영상으로 혹은 그 반대의 절차가 필요하고, 그 과정에서 전-후처리는 압축 성능에 영향을 미치기 때문에, JVET에서는 선처리 및 후처리를 고려한 스터디가 진행되었으며, 특별히 프로젝션 변환과정이 코딩성능에 미치는 영향력에 대한 연구를 많이 진행하였다.

2017년 10월 미팅에서 JVET은 차세대 비디오 코덱 표준을 위한 CFP문서를 발간하였고, 2018년 4월 CFP 대응 결과를 바탕으로 차세대 비디오 코덱 표준의 향해를 시작하였다. 프로젝트 명은 Versatile Video Coding standard로 정하였고, 그 대상은 SDR, HDR, 360 Video이다. 현재 JVET은 2020년에 IS발간을 위해 표준화를 진행하고 있다[3].

본 섹션에서는 JVET그룹에서 표준화 진행중인 VVC 코덱 표준화 중 360도 영상을 중심으로 표준화를 설명하고자 한다.

1) 프로세싱 체인

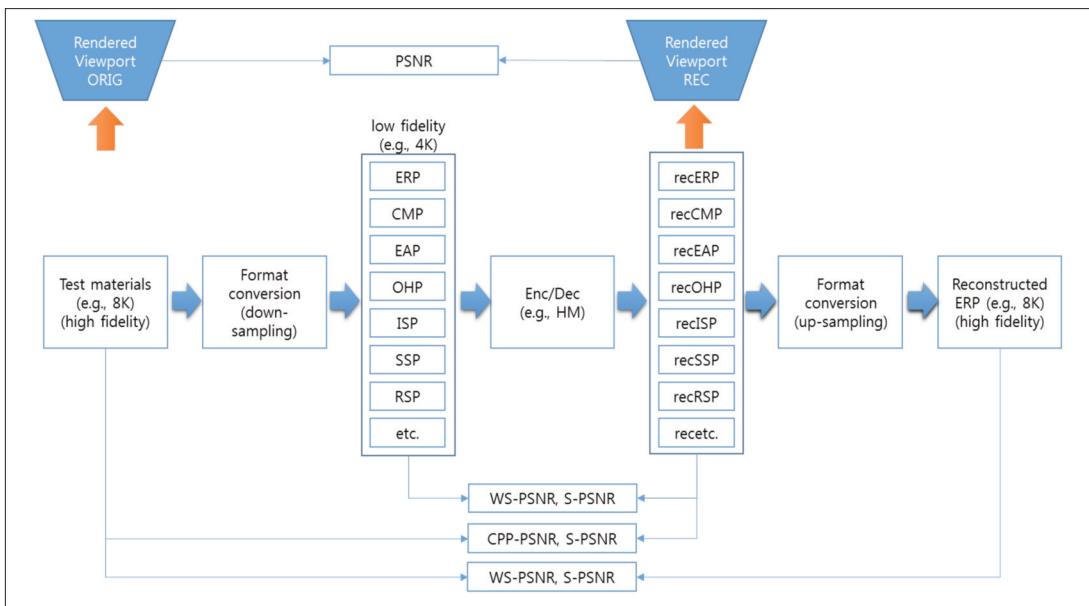
JVET에서 사용하는 360도 VR 압축을 위한 테스트 영상은 Equi-Rectangular Projection format (ERP) 프로젝션 방식으로 만들어진 2D기반 8K 해상도 영상을 사용한다. 현재 시장에서 사용되는 360 영상 카메라는 대부분 저장순간 2D기반 ERP

프로젝션 방식으로 저장되기에, 테스트 영상 또한 ERP 프로젝션 방식으로 만들어진 영상을 사용하였다.

〈그림 3〉은 JVET에서 정한 360도 영상 기본 프로세싱 체인을 나타낸다. 먼저 8K영상이 초기 입력으로 사용되고, 다운 샘플링 과정과 함께 사용하고 자 하는 포맷으로 변환한다. 이렇게 변환된 특정 프로젝션의 4K 2D영상이 인코더의 입력 값이 된다. 이후 인코딩/디코딩 과정을 거쳐 나온 특정 프로젝션의 4K 2D영상은 바로 Viewport renderer를 통해 3D 영상으로 변환 되거나, 또는 원본과 같은 ERP 8K 영상으로 업샘플링되어 아웃풋으로 출력되기도 한다.

2) 프로제션 방법

프로제션 방법은 3D를 구성하는 방식과 3D상에 있는 영상을 인코딩 하기 전에 2D영상으로 바꾸는



〈그림 3〉 360도 영상 기본 프로세싱 체인

과정을 말한다. 현재 JVET에서는 다양한 방식의 프로젝션 방식이 제안되어 있으며, 각 프로젝션 방식이 코딩 성능에 미치는 영향을 파악하기 위해, 성능적 혹은 기능적으로 이득이 있다고 판단되는 프로젝션들을 수집하여 SW에 구현하였다. 현재, 각 프로젝션 방식은 인/디코더와 함께 스테디되고 있다.

■ Equi - Rectangular Projection format(ERP)

산업계에서 가장 많이 쓰이고 있는 포맷이다. 3D를 구로 모델링하였으며, 3D영상을 2D영상으로 변환 시 경도를 x축으로 위도를 y축으로 바꾸어 매핑시킨다. 세계지도를 나타낼 때 가장 많이 쓰이는 방법과 동일하다[4].

■ CubeMap Projection format(CMP)

3D를 정육면체로 모델링하였으며, 3D영상을 2D영상으로 변환 시 6개의 사각형을 하나의 평면으로 매핑시킨다. 6개의 사각형을 하나의 평면으로 매핑 과정시 다양한 레이아웃이 가능하며, 현재는 3x2 레이아웃이 디폴트로 쓰이고 있다[5].

■ Equal - Area Projection format(EAP)

ERP 포맷과 유사하게 3D를 구로 모델링하였으며, 3D영상을 2D영상으로 변환 시 경도를 x축으로 위도를 y축으로 바꾸어 매핑시키지만, ERP와는 달리 경도를 y축 매핑 시 uniform 매핑이 아닌 non-linear 매핑을 쓴다. 해당 매핑을 쓰는 경우 2D상에서 나타내는 영역의 크기가 3D상과 같은 비율이 되는 장점이 있다[6].

■ OctaHedron Projection format(OHP)

3D를 정팔면체로 모델링하였으며, 3D영상을 2D

영상으로 변환 시 8개의 정삼각형을 하나의 평면으로 매핑시킨다. 8개의 정삼각형을 하나의 평면으로 매핑 과정 시 다양한 레이아웃이 가능하다[7].

■ IcoSahedron Projection format(ISP)

3D를 정이십면체로 모델링하였으며, 3D영상을 2D영상으로 변환 시 20개의 정삼각형을 하나의 평면으로 매핑시킨다. 20개의 정삼각형을 하나의 평면으로 매핑 과정 시 다양한 레이아웃이 가능하다[8].

■ Segmented Sphere Projection format(SSP)

ERP 포맷과 유사하게 3D를 구로 모델링하였으며, 3D영상을 2D영상으로 변환 시 3개의 부분으로 나누어 매핑시킨다. 위도 -45도부터 +45도까지는 ERP와 동일하게 2D상으로 매핑시키나, -45도 이하, +45도 이상인 부분은 원으로 2D상에 매핑시킨다. 각 부분은 6개의 정사각형으로 구성을 할 수 있다[9].

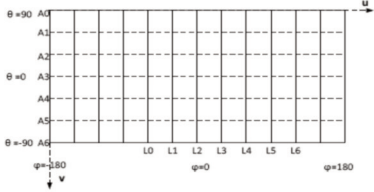
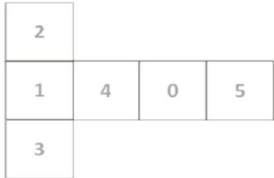
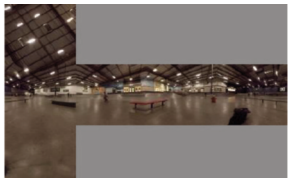
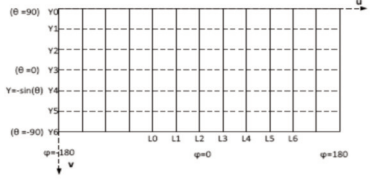
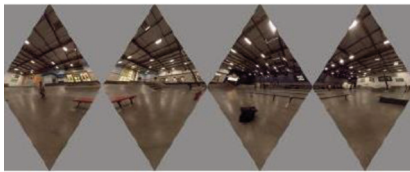
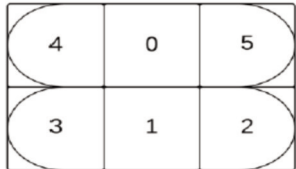
■ Rotated Sphere Projection(RSP)

ERP 포맷과 유사하게 3D를 구로 모델링하였으며, 3D영상을 2D영상으로 변환 시 2개의 페이스를 2D상으로 매핑시킨다. 해당 페이스 구성은 테니스 공을 구성하는 페이스와 동일하다[10].

■ 그 외의 프로젝션

위에서 언급된 프로젝션 외에 Truncated Square Pyramid (TSP)[11], Adjusted Cubemap Projection(ACP)[12], Equi - Angular Cubemap projection(EAC)[13], Hybrid Equi - angular Cubemap(HEC)[14] 등이 360Lib에서 지원한다. 대부분 CMP에서 파생된 프로젝션 방법들이다.

〈표 2〉 프로젝션 방식 및 예제

Projection	Frame packing	Example
Equi - rectangular		
Cubemap		
Equal - area		
Octahedron		
Icosahedron		
Segmented sphere		
Rotated Sphere		

3) 360도 영상 평가를 위한 객관적 평가

매트릭

비디오 코딩에서는 전통적으로 Peak Signal-to-Noise Ratio(PSNR)방식을 객관적 화질 평가 방식으로 사용하였다. 360도 영상의 경우 프로젝션을 고려한 수정된 PSNR 평가방법이 쓰이고 있으며, 객관적 화질 평가를 위해 프로세싱상에서 비교평가를 하기 위한 체인 상의 위치와 2D 이미지 상에서 중요도를 고려하는 방식으로 나누어 평가를 하고 있다. <그림 3>에서 구체적 확인이 가능하다.

JVET에서 평가방식을 다양하게 고려하는 이유는 다음과 같다. 먼저 전통적인 2D영상에서 사용하는 평가방식을 적용할 수 없는 신유형의 영상이기 때문이다. 2D영상을 위한 평가 방식은 많이 나와있으며 어느 정도 검증 절차도 마련되어 있지만, 360도 영상을 위한 표준 평가 방식은 현재 마련되어 있지 않다. 영상 평가 방식이 표준 제정에 쓰이는 기술을 선정하는 잣대라고 생각하여 보았을 때, 그룹에서 합의한 영상 평가 방식 마련은 굉장히 중요한 일이며 표준 시작 전 반드시 제정되어야 할 부분이다.

먼저 JVET에서는 프로세스 체인상 어떤 위치에서 평가하느냐에 따라 3가지 방식이 존재한다.

■ End-to-end 화질 열화 평가

입력영상인 8K ERP 포맷 이미지와 프로세싱체인 마지막 단계인 복원 8K ERP 포맷 이미지 사이의 PSNR을 측정하는 방식이다. End-to-end 화질 열화 평가방법을 사용하여 화질을 평가할 경우, 다운/업 샘플링에 따른 화질변화, 프로젝션 변환에 따른 화질 변화, 인/디코딩에 따른 화질 변화를 모두 포함하는 화질 평가가 가능하나, 각각의 개별적

성능은 측정하기 힘들다.

■ Cross-format 화질 열화 평가

입력영상인 8K ERP 포맷 이미지와 프로세싱체인 중 인/디코딩 직후 화질의 변화량을 측정하는 방식이다. 인/디코딩 이후 사용자가 사용하는 렌더러 입력영상 위치가 바로 인/디코딩 직후이기 때문에 해당 포맷의 사용이 고려되었다. 또한, 만약 인/디코딩에서 lossless 방식을 사용 할 경우, 해당 평가 방식은 오직 프로젝션 포맷팅 변화에 따른 화질 변화를 측정할 수 있기에, 프로젝션 포맷팅이 화질에 미치는 영향을 평가하기 위해서 사용한다.

■ Coding 화질 열화 평가

해당 화질 평가 방식은 전통적으로 쓰였던 인/디코딩 전-후단에서 화질을 평가하였던 방식과 동일하다. 프로젝션 방식에 상관 없이 오직 인/디코딩 변화에 따른 화질 변화를 측정한다.

두 번째로 포맷 변환과정을 고려한 4가지 평가 방식이 존재한다. 각 포맷이 변환과정에서 어떠한 영향을 받느냐에 따라 화질평가는 매우 달라질 수 있기 때문에, 포맷에 영향을 받지 않은 일반화된 매트릭이 JVET에서는 스터디 되었다.

■ PSNR

프레세싱 결과로 나온 영상이 원본과 얼마나 유사한지 측정하기 위해서 쓰이는 화질 평가 방식이다. 전통적인 비디오코딩에서 쓰는 객관적 화질 평가 방식이다. 원본 영상과 디코딩 영상과의 평균오차를 고려한 평가방식이다.

■ Spherically uniform PSNR(WS-PSNR)

기본적으로 PSNR 방식과 유사하다. 하지만 PSNR과 달리 2D상으로 매핑된 3D상에 있는 샘플의 의미를 해석하여 각 샘플에 가중치를 주는 방식으로 계산된다. 3D에서 2D로 프로젝션시 영상의 왜곡이 발생한다. 만약 해당 왜곡을 포함한 객관적 화질 평가가 이루어진다면, 코덱 성능을 정확히 파악하기 힘들다. 예를 들어 ERP포맷에서 폴부분과 적도 부분을 동일하게 화질 평가를 한다면 해당 평가 수치는 심하게 왜곡된 정보에 기반한 평가 수치가 될 것이다. 이를 극복하고자 JVET에서는 2D 프로젝션 평가시 포맷이 가지는 왜곡의 정도에 따라서 가중치를 달리 주는 방식의 PSNR를 제시하였다.

〈그림 4〉는 ERP포맷을 위한 WS-PSNR평가 방식의 예시이다[15]. (m, n) 도메인에서 가중치 $w(i, j)$ 는 다음과 같이 계산된다. 〈그림 4〉는 ERP포맷에서 $w(i, j)$ 수식으로 계산된 가중치를 그림으로 도식한 것이다. 밝은 색을 가지 것은 가중치가 크다는 의미이며, 어두운 영역은 가중치가 작음을 의미한다.

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



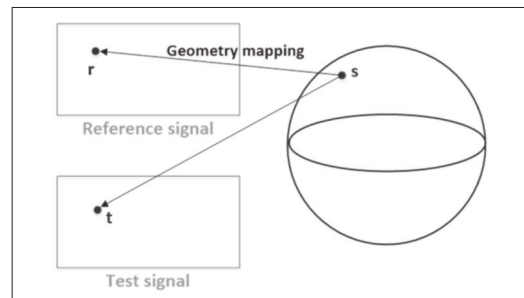
〈그림 4〉 ERP포맷에서 가중치 표시

■ Spherical PSNR(S-PSNR)

3D상에 있는 위치를 2D상으로 매핑하여 원본 영

상과 프로세싱 처리 영상의 PSNR을 측정하는 방식이다. 3D상에서의 모델링 방식으로 측정되기에 프로젝션 방식과 상관없이 PSNR 측정이 가능하다.

해당 객관 화질 평가 방식은 JVET 스터디 기간에 제안되었던 것으로, 두 개의 다른 프로젝션 포맷의 경우 객관적 화질 평가가 불가능하였던 단점을 극복하였다. 먼저 3D상의 특정 위치에 매핑하여 3D상에서 동일 위치를 파악한 뒤, 전통적인 MSE 기반의 PSNR을 측정하는 방식이다[16]. 특정프로젝션 포맷에 구애 받을 필요가 없기에 360도 영상의 화질을 평가하는 데에 좋은 매트릭이다.



〈그림 5〉 S-PSNR의 계산과정

■ Crasters Parabolic Projection(CPP-PSNR)

2D 영상을 Craster parabolic 프로젝션으로 바꾼 뒤 PSNR을 측정하는 방식이다[17]. Craster parabolic 프로젝션으로 변환을 시키는 이유는 앞서 소개한 WS-PSNR, S-PSNR과 동일하다. 각 포맷이 가지는 왜곡현상을 제거하고 의미있는 객관적 수치를 확보하기 위해서 정량화된 이미지 위에서 PSNR을 측정하는 것이다. 구체적으로 2D이미지를 Craster parabolic 2D영상으로 변환한 뒤 PSNR를 측정한다. 해당 Craster parabolic 2D영상은 의미상의 가치가 각 샘플에서 동일하기 때문에 왜곡 현상이 객관적 수치에 반영되는 부분이 적게 된다.



〈그림 6〉 CPP-PSNR을 위한 Craster parabolic 변환

4) 프로젝션 성능 평가

본 문단에서는 현재 2016년부터 제안된 프로젝션을 중심으로 객관적 화질 성능 결과를 제시한다. 평가 방식은 현재 JVET에서 공식적으로 사용하고 있는 End-to-end 화질 열화 기반에서 WS-PSNR을 측정한 수치를 나타낸다.

〈표 3〉은 각 프로젝션별 성능을 나타낸다. 성능 평가 환경은 360° video Common Test Condition (CTC)[18]를 따랐으며, 비교를 위한 앵커는 HM-16,16과 함께 PERP 프로젝션 방식이 사용되었다. 〈표 3〉의 성능을 확인해 보면, 코덱 내부의 툴의 변화 없이 프로젝션의 변화로 최대 10%에 가까운 성능이 향상됨을 확인 할 수 있다. VVC가 타겟으로

삼고 있는 HEVC 대비 성능 향상이 50%라는 것을 고려하여 볼 때, 코덱 내부의 변화 없이 프로젝션 변화만으로 약 10%로의 성능향상이 가능하다는 점에서 의미 있는 결과라고 판단된다.

최근 JVET 7월 미팅에서 HEC 프로젝션이 360Lib에 추가 채택되었다. 해당 프로젝션은 EAC, ECP, EAC와 동일한 접근법으로 만들어진 프로젝션임을 고려하여 볼 때 CMP에서 파생된 프로젝션 방법으로 간주된다. 성능은 EAC, ECP, EAC와 유사한 11.7%로 확인되었다[14].

2. MPEG-I 6DoF 소개

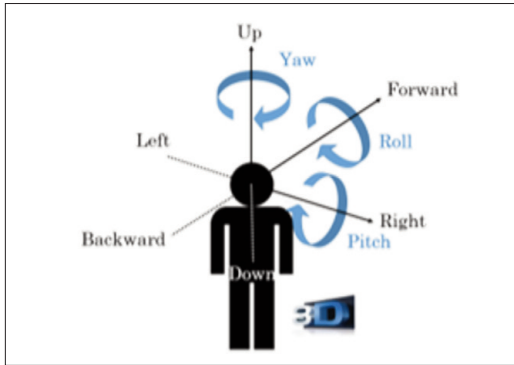
1) MPEG-I 6DoF 소개

MPEG-I에서 표준화로 진행될 예정인 MPEG-I Phase 2는 Phase 1.b에 이어 총 3개의 파트로 구성되어 있다. Phase 1에서는 주로 제한된 범위내에서의 사용자 움직임을 지원하는 3DoF 비디오에 대한 부호화 및 시스템 구조를 위한 표준화에 집중하고 있는 반면, Phase 2에서는 보다 넓은 범위에서 사용자의 자유로운 움직임을 지원하는 6DoF 비디오

〈표 3〉 각 프로젝션별 성능

Projection	E2E WS-PSNR for all sequences			E2E WS-PSNR for 8K sequences			E2E WS-PSNR for 6K sequences			E2E WS-PSNR for 4K sequences		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
CMP	-2.4%	-0.7%	-1.0%	-5.1%	-3.0%	-3.5%	0.1%	3.3%	2.8%	0.5%	-1.6%	-1.2%
AEP	-4.0%	-3.1%	-3.2%	-4.7%	-3.4%	-3.4%	-2.9%	-2.0%	-2.2%	-4.3%	-4.6%	-4.7%
COHP	-1.4%	1.5%	0.6%	-5.0%	-0.2%	-1.3%	3.5%	5.5%	4.5%	-0.4%	-1.1%	-1.3%
CISP	-3.3%	1.1%	0.2%	-7.5%	-1.3%	-2.7%	2.4%	6.2%	5.7%	-1.8%	-1.8%	-2.1%
SSP	-9.7%	-5.7%	-6.3%	-11.9%	-6.4%	-6.6%	-7.3%	-4.1%	-5.3%	-7.6%	-6.9%	-7.2%
ACP	-9.9%	-5.7%	-6.1%	-12.5%	-6.3%	-6.5%	-7.7%	-4.7%	-5.3%	-6.7%	-6.2%	-6.2%
RSP	-9.2%	-5.8%	-6.1%	-12.4%	-7.0%	-7.4%	-6.0%	-3.9%	-4.5%	-6.0%	-5.9%	-5.3%
ECP	-9.0%	-5.4%	-6.0%	-12.3%	-6.7%	-7.2%	-4.5%	-2.4%	-3.6%	-7.8%	-7.6%	-7.3%
EAC	-9.9%	-5.8%	-6.0%	-12.5%	-6.4%	-6.4%	-8.1%	-4.8%	-5.3%	-6.2%	-6.2%	-6.0%

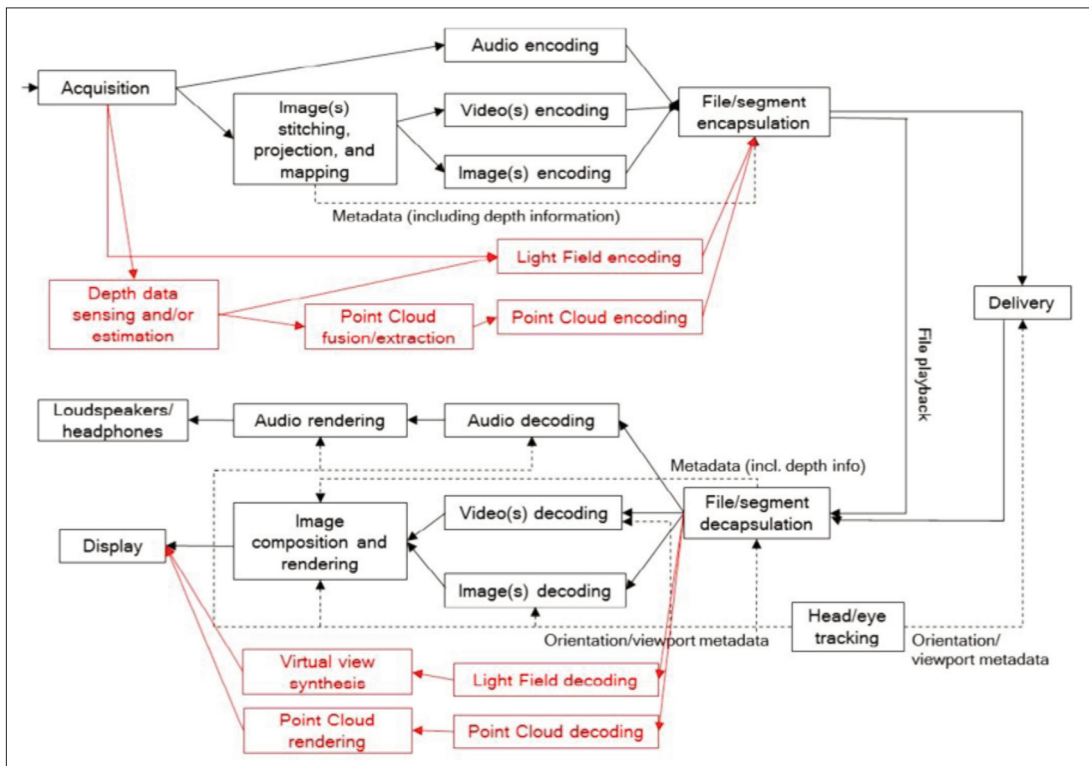
압축 기술 및 전송 기술을 표준화할 것으로 예상된다. <그림 7>은 6DoF 범위 내에서의 사용자 움직임을 예시로 보여주고 있으며, <그림 8>은 Phase 2의



<그림 7> 6DoF 사용자 움직임

6DoF를 지원하기 위한 MPEG-I 6DoF 시스템 구성도를 보여준다[20].

<그림 8>의 MPEG-I Phase 2 6DoF 시스템 구성도에서 보인 것처럼 6DoF 비디오로써 Light Field 비디오 및 포인트 클라우드 비디오를 그 대상으로 고려하고 있으며, 이를 중심으로 Omnidirectional 6DoF, Windowed 6DoF, Dense Light Field를 주제로 탐색 실험이 진행 중에 있다. 탐색 실험이라함은 본격적인 표준화를 진행하기 전에 관련 기술들의 성숙도를 점검하는 단계로서 1년 정도 진행하는 것이 보통이다. 다만, 6DoF에 대한 탐색 실험은 위에서 언급한 세가지를 동시에 모두 진행하기에는 무리가 있는 상황이라 단계적으로 진행될



<그림 8> MPEG-I Phase 2 6DoF 시스템 구성도

예정이며, Windowed 6DoF가 우선적으로 표준화 시작 직전 단계인 Call for Evidence(CfE)에서 보기 위한 공통 실험 조건을 정하는 단계인 CTC 단계로 접어들었으며, 122차 미국 샌디에고 MPEG 회의에서 그 문서 초안이 공개되었다. 다음장에서는 세가지 6DoF에 대해 간략한 특징과 함께 현재 진행되고 있는 탐색 실험에 대해 간략하게 소개한다.

2) MPEG-I 6DoF 탐색 실험

■ Omnidirectional 6DoF

Omnidirectional 6DoF는 360 비디오를 기본으로 사용자의 움직임에 따라 사용자 시점에 맞춰 제한된 공간에서 6DoF를 지원하는 특징을 가지고 있다[21]. 2016년까지 MPEG FTV에서는 3rd Phase FTV(Free Navigation Television)를 위한 탐색 실험과 이를 위한 Call for Evidence Response를 수행한 바 있다. 특히, 3rd Phase FTV에서는 Super Multiview Video Coding과 Free Navigation Video Coding에 대한 표준화를 진행하였다. 다만, Call for evidence 과정에서 특별한 evidence를 보지 못하여 위의 두 시나리오에 대해서는 표준화 단

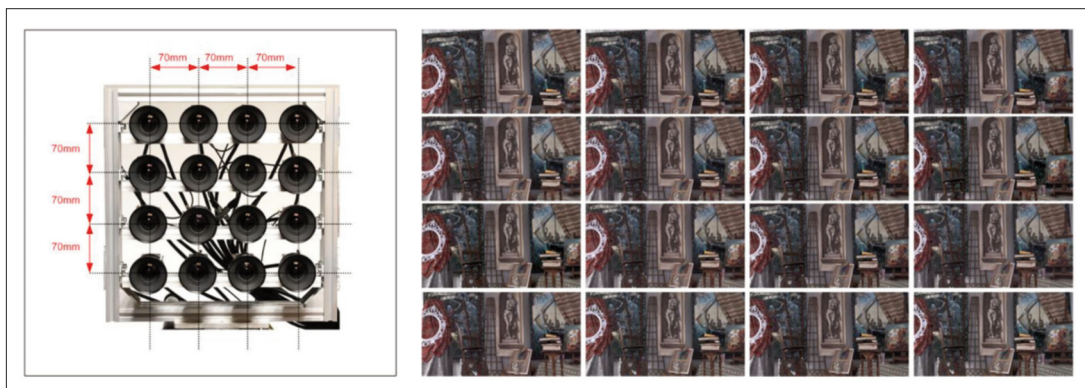
계까지 이르지 못하였다. 그리고 2015년 10월 MPEG FTV AHG에서는 360 3D video라는 포맷에 대한 탐색 실험이 새롭게 시작되었으며, 이는 2016년 10월 청두 회의에서 MPEG-I 프로젝트가 MPEG에서 시작되면서 Omnidirectional 6DoF video로 발전하였고 이에 대한 탐색 실험이 계속 진행되고 있다.

현재의 탐색 실험에서는 360도 비디오에 대한 깊이 영상을 추정하는 기술에 대한 EE-Depth와 추정된 깊이 영상과 360도 비디오를 사용하여 사용자 움직임에 따른 시점에 가상의 360도 비디오 영상을 생성하는 EE-Synth로 나누어 진행 중이다[22][23].

■ Windowed 6DoF

Windowed 6DoF를 위한 영상은 다시점 카메라 시스템으로 촬영된 영상들로부터 사용자 움직임에 따라 가상시점 영상을 생성하는 것이 특징이다. 특히 하나의 시점 영상으로부터 이웃한 시점영상들까지 사용자가 바라보는 창문 형태의 영역으로 제한하고 있다[24].

Windowed 6DoF의 탐색 실험에서는 <그림 9>에

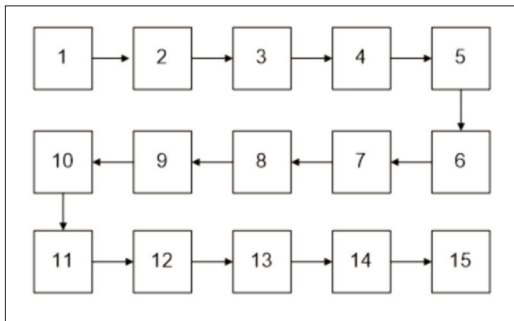


<그림 9> Windowed 6DoF 카메라 배열 형태(좌) 및 획득 영상(우)의 예

서와 같이 각 카메라로부터 자연스러운 시점 변화를 주기 위해 가상시점 영상 생성 실험과 가상시점 영상 생성을 위해 깊이 영상 추정 실험을 진행하고 있다.

또한, 카메라로부터 얻은 시점 영상과 추정된 깊이 영상에 대한 부호화 탐색 실험을 진행하고 있다. 현재 부호화 탐색 실험을 통해 시점간 예측 순서에 따라 압축 성능이 달라질 수 있다는 것을 알 수 있었다. 그 중 한가지 방법으로 순차적 시점간 예측 순서는 <그림 10>과 같다[25].

다양한 시점간 예측 순서가 있을 수 있는데, 이는 현재 부호화 탐색 실험이 계속 진행 중에 있다. 동



<그림 10> 시점간 예측을 위한 시점 스캔 방식의 예

시에 공통 실험 조건에 대한 논의가 진행 중에 있으며, 여기서 부호화 성능 평가는 MV-HEVC를 기준으로 사용하고 있으며, 아래의 조건으로 실험을 수행하는 것을 논의하고 있다[26].

- 시점간 예측 방식은 <그림 10>과 같은 스캔 방향을 사용
- 계층적 B구조는 GOP8사이즈 및 인트라 24주기로 사용
- 움직임 예측은 고속 다이아몬드 검색 방법을 이용하고 탐색 범위는 96으로 설정
- 실험에 사용되는 컬러영상과 깊이영상에 대한

양자화 파라미터(QP)스텝은 <표 4>와 같다.

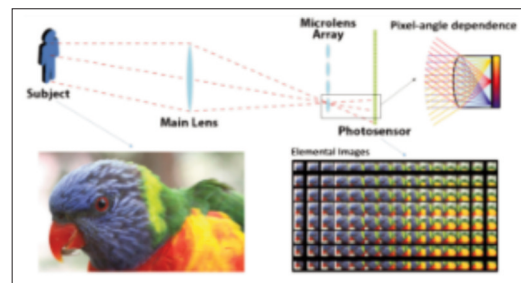
<표 4> 양자화 파라미터

Sequence	Medium bitrate	Low bitrate
Technicolor Painter and ULB Unicorn, planeA	T: 25 30 35 40 D: 34 39 42 45	T: 30 35 40 45 D: 39 42 45 48

부호화 및 복호화 성능을 평가하기 위해 복호화된 입력 영상을 사용해 가상시점 영상을 생성하여 원본 영상과 PSNR 비교를 수행한다. 이러한 평가를 통해 부/복호화에 사용된 비디오 코덱이 가상시점 영상의 품질에 미치는 영향을 파악할 수 있다.

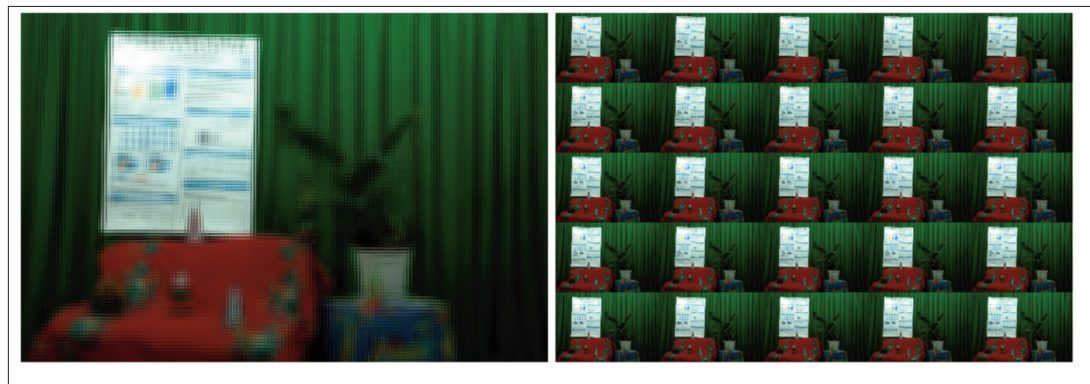
■ Dense Light Field 부호화 성능 탐색 실험

Dense Light Field 비디오는 micro lens를 사용하여 만든 lenslet을 2차원 배열 형태로 구성한 조밀한 영상을 의미한다. 2차원 배열의 조밀한 다시점 영상을 의미한다. 아래의 <그림 11>에서처럼 라이트필드 영상은 상 앞에 대물 렌즈를 두어 상의 초점을 맞추고 중간에 2차원 배열의 조밀하게 배열된 마이크로 렌즈로 다른 시점의 상을 감광기(Photosensor)에 맺게 함으로서 획득할 수 있다[27].



<그림 11> 라이트필드 영상 획득 개념

Dense Light Field 영상은 <그림 12>에서처럼 다시점 영상으로 변환이 가능하다.

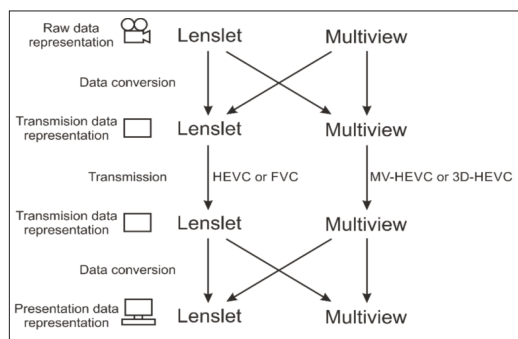


〈그림 12〉 Dense Light Field 영상(좌)와 변환된 다시점 영상(우)의 예

부호화 탐색 실험에서는 이러한 Dense Light Field 영상과 변환된 다시점 영상의 압축 효율에 대한 비교 실험을 진행하고 있다. 아래의 〈그림 13〉은 Dense Light Field 영상의 부호화 탐색 실험의 과정을 보여주고 있다.

Dense light Field 부호화 탐색 실험의 목적을 요약하면 아래와 같다.

- Lenslet 비디오와 다시점 비디오 부호화 기술에 대한 성능 탐색
- 서로 다른 포맷 간의 변환을 통해 품질 저하 비교 탐색
- 변환시킨 포맷에 대한 부호화 성능 탐색
- 현재 기술에 대한 한계점을 정의하기 위해 탐색



〈그림 13〉 Dense Light Field 부호화 탐색 실험

III. 결론

1. VVC 360 표준 향후 전망 및 일정

샌디에고에서 열린 10차 JVET미팅에서는 VVC와 관련된 CfP 결과가 발표되었다. SDR, HDR, 360도 영상 이렇게 3가지 항목으로 CfP를 응답받았으며 360도 영상의 경우 12개의 제출본을 차기 테스트 모델을 위해서 평가하였다. 각 기관에서 제출한 기술은 각자가 개발한 프로젝션 포맷에 코어 코덱을 적용하였으며, 제출된 문서에 따르면, 객관적 성능이 모두 우수한 것으로 나타났다[19].

CfP 응답을 평가하기 위해 MPEG 테스트 체어가 평가단을 구성하여 주관적 화질 평가를 진행하였다. 평가단은 Anchor(현재 기술)인 ERP 프로젝션 +HEVC 코덱과 비교하여 제출된 기술이 동일 비트레이트에서 얼마나 좋은 성능을 나타내는 지 점수를 매겼다. 1(원본과 다름) - 5(원본과 동일) 사이의 점수를 매겨 각 응답 기술의 주관적 성능향상을 확인하였는데, 평균적으로 동일화질이라 간주되는 구간에서 Anchor에 비해 절반의 비트레이트가 확인되었다. 다시 말해 동일 화질을 나타내기 위해서 응답한

기술은 2배의 압축 성능을 나타내는 것이 증명되었다.

해당 미팅에서는 360도 영상을 위해서 특정 포맷을 표준으로 제정하지 않았고, Core Experiment 단계를 통해 360도 영상 압축을 위해 적합한 포맷 및 코딩 툴을 만들어 나가고 있다. 특별히 360도 영상을 위한 압축 표준에서 아직 토론이 필요한 사항은 SDR을 위한 기술과의 연계성이다. 현재 360도 영상 압축 기술은 프로젝션 포맷을 중심으로 진행되어 왔기에, SDR 코덱을 포함하는 형태로 되어 있지만, 360 영상 특징은 SDR과 상이한 부분이 존재하기에, 360도 영상을 위한 새로운 코딩툴 혹은 기존 SDR툴의 변형이 필요한 상황이다.

CFP 응답 이후 2회의 추가 JVET 미팅을 거쳤지만, 해당 논의에 대한 결론은 아직 나지 않는 상황이며, 360도 영상을 위한 새로운 코딩툴 혹은 기존 SDR툴의 변형이 필요하다면, VVC의 새로운 프로파일로 만들지, 아니면 SDR을 위해 정해진 코덱의 변형을 허용해서 하나의 프로파일로 합칠 지의 논의가 차후 회의에서 결정될 예정이다. VVC의 남은 일정은 <표 5>와 같다.

<표 5> 표준화 진행 일정

일정	내용
2018/04	테스트 모델 선정 및 Working Draft 발간
2018/07	Working Draft2 발간
2018/10	Working Draft3 발간
2019/04	Committee Draft 발간
2019/10	Final Committee Draft 발간
2020/10	Final Draft International Draft 발간

2. 6DoF 표준 향후 전망 및 일정

MPEG에서는 6DoF 표준화 완료 시점을 2021년

~2022년을 목표로 잡고 있으며, MPEG-I에서 다루어지고 있는 표준 중에 가장 중심이 되는 몰입형 비디오는 라이트필드 비디오와 포인트클라우드 비디오라고 할 수 있다. 현재 포인트클라우드 이미 2017년 10월 마카오 MPEG미팅을 시작으로 표준화가 활발하게 진행되고 있으며, MPEG-I Visual AhG에서는 6DoF와 관련된 Omnidirectional 6DoF, Windowed 6DoF, Dense Light Field 부호화에 대한 탐색 실험을 통해 현재 기술 수준과 부족한 기술이 무엇인지를 파악해 나가고 있는 단계에 와 있다. 현재의 동향으로 봤을 때, 1~2년내로 6DoF 부호화에 대한 표준화가 본격적으로 시작될 것으로 예상된다.

3. 마무리

2K, 4K 영상의 등장은 소비자에게 고화질 영상을 일상생활에서 즐길 수 있게 하였다. 그러나 소비자는 이제 고화질 영상을 넘어 현장감 있는 영상을 몰입감 있게 느끼길 원한다. 이에 발맞추어 새롭게 등장한 360도 영상/VR/AR 기술은 소비자에게 생동감 넘치는 실감성을 제공하는 기술로 빠르게 성장하고 있으며, 시장 또한 급격하게 확대되고 있다. 그러나 현존하는 표준 기술로는 360도 영상/VR/AR 표현에 소요되는 막대한 데이터의 용량을 처리하기에는 충분하지 않은 수준이다. 또한, 기능성 측면에서도 현존하는 사용자 머리의 움직임 기반 YAW, Pitch, Roll만을 정의하는 3DoF만으로는 소비자가 원하는 몰입형 경험을 제공하기에는 부족한 부분이 존재한다. 이러한 요구사항을 바탕으로 새롭게 진행 중인 VVC 360도 영상 표준 코덱과 6DoF 표준 재정이, 시장에서 요구하고 있는 저장 공간 및 전송대역폭의 문제를 해결과 더불어 소비

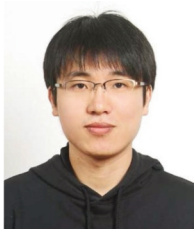
자가 원하는 움직임 자유도를 제공하여 몰입감 및 사실감을 극대화 할 수 있는 기술을 표준에서 제공할 수 있을지 지속적으로 지켜봐야 할 것이다.

참고 문헌

- [1] "Planned Phases for Standardization of Immersive Media", ISO/IEC JCT1/SC29/WG11/N16543, Chengdu, CN, October 2017.
- [2] "Requirements for a Future Video Coding Standard v5", ISO/IEC JCT1/SC29/WG11/N17074, Torino, IT, July 2017.
- [3] A. Segall, V. Baroncini, J. Boyce, J. Chen, T. Suzuki, "Joint Call for Proposals on Video Compression with Capability beyond HEVC", JVET-H1002, Macao, CN, Oct. 2017.
- [4] Smolic, Aljoscha, and David McCutchen. "3DAV exploration of video-based rendering technology in MPEG." IEEE Transactions on Circuits and Systems for Video Technology 14.3 (2004): 348-356.
- [5] Corbillon, Xavier, et al. "Viewport-adaptive navigable 360-degree video delivery." Communications (ICC), 2017 IEEE International Conference on. IEEE, 2017.
- [6] M. Zhou, "AHG8: A study on quality impact of line re-sampling rate in EAP", JVET-G0051, Torino, IT, July 2017.
- [7] H.-C. Lin, C.-Y. Li, J.-L. Lin, S.-K. Chang, C.-C. Ju, "AHG8: An efficient compact layout for octahedron format", JVET-D0142, Chengdu, CN, Oct. 2016.
- [8] V. Zakharchenko, E. Alshina, K. P. Choi, A. Singh, A. Dsouza, "AhG8: Icosahedral projection for 360-degree video content", JVET-D0028, Chengdu, CN, Oct. 2016.
- [9] C. Zhang, Y. Lu, J. Li, Z. Wen, "AhG8: segmented sphere projection (SSP) for 360-degree video content", JVET-D0030, Chengdu, CN, Oct. 2016.
- [10] Adeel Abbas, David Newman, "AHG8: Rotated Sphere Projection for 360 Video", JVET-F0036, Hobart, AU, 31 Mar. 2017.
- [11] G. Van der Auwera, M. Coban, Hendry, M. Karczewicz, "AHG8: Truncated Square Pyramid Projection (TSP) For 360 Video", JVET-D0071, Chengdu, CN, Oct. 2016.
- [12] M. Coban, G. Van der Auwera, M. Karczewicz, "AHG8: Reference picture extension of ACP format 360-degree video", JVET-G0058, Torino, IT, July 2017.
- [13] M. Zhou, "AHG8: A study on Equi-Angular Cubemap projection (EAC)", JVET-G0056, Torino, IT, July 2017.
- [14] Y.-H. Lee, J.-L. Lin, S.-K. Chang, C.-C. Ju, "CE13: Modified Cubemap Projection in JVET-J0019 (Test 5)," JVET-K0131, Ljubljana, SI, Jul. 2018.
- [15] Y. Sun, A. Lu, L. Yu, "AHG8: WS-PSNR for 360 video objective quality evaluation", JVET-D0040, Chengdu, CN, Oct. 2016.
- [16] B. Vishwanath, Y. He, Y. Ye, "AHG8: Area Weighted Spherical PSNR for 360 video quality evaluation," JVET-D0072, Chengdu, CN, Oct. 2016.
- [17] V. Zakharchenko, E. Alshina, A. Singh, A. Dsouza, "AhG8: Suggested testing procedure for 360-degree video", JVET-D0027, Chengdu, CN, Oct. 2016.
- [18] E. Alshina, J. Boyce, A. Abbas, Y. Ye, "JVET common test conditions and evaluation procedures for 360° video", JVET-H1030, Macao, CN, Oct. 2017.
- [19] V. Baroncini, "Results of Subjective Testing of Responses to the Joint CfP on Video Compression Technology with Capability beyond HEVC", JVET-J0080, San Diego, US, Apr. 2018.
- [20] "New Work Item Proposal on Coded Representation of Immersive Media," ISO/IEC JTC1/SC29/WG11, N16541, Jan. 2017.
- [21] Gun Bang, Gauthier Lafruit, "Introduction to 360 3D video application and requirement for FTV discussion", ISO/IEC JTC1/SC29/WG11, w37351, Oct. 2015.

- [22] Krzysztof Wegner, Olgierd Stankiewicz, Adrian Dziembowski, Dawid Mieloch, Marek Domański, Omnidirectional 6-DoF/3-DoF+ rendering, ISO/IEC JTC1/SC29/WG11, M40806, July 2017.
- [23] Gun Bang, Bart Kroon, "Exploration Experiments for MPEG-I: Omnidirectional 6DoF", ISO/IEC JTC1/SC29/WG11, N17722, July 2018.
- [24] J. Jung, and G. Lafruit, "Exploration Experiments for MPEG-I: Windowed-6DoF," ISO/IEC JTC1/SC29/WG11, N17721, July. 2018.
- [25] J. Jung, "Limitation of Multi-View Extensions of HEVC and Fixes for MPEG-I Phase 2," ISO/IEC JTC1/SC29/WG11, N17133, Oct. 2017.
- [26] J. Jung, and G. Lafruit, "Common Test Conditions on 3DoF+ and Windowed 6DoF.", ISO/IEC JTC1/SC29/WG11, N17726, July 2018.
- [27] Mehrdad Teratani (Panahpourtehrani), Kazuhiro Hara, "Exploration Experiments for MPEG-I: Compression of Dense Representation of Light Fields", ISO/IEC JTC1/SC29/WG11, N17723, July 2018.

필자 소개



최기호

- 2008년 : 한양대학교 정보통신대학 졸업 (Summa Cum Laude)
- 2012년 : 한양대학교 일반대학원 전자컴퓨터통신공학과 박사졸업
- 2014년 : 한양대학교 전자컴퓨터통신공학과 Post Doc.
- 2014년 ~ 현재 : 삼성전자 Samsung Research 책임연구원
- 주관심분야 : 비디오 코딩 표준화, 멀티미디어 관련 연구



방건

- 2014년 : 고려대학교 컴퓨터전파통신공학과 박사졸업
- 2006년 : ATSC T3/S2 ACAP 표준화 Editor
- 2012년 : MIT RLE ATSP group 방문 연구원
- 2000년 ~ 현재 : 한국전자통신연구원 방송미디어연구소 책임연구원
- 주관심분야 : 2D/3D 비디오 부호화, 영상 처리, 컴퓨터 비전, 인공지능