

몰입형 미디어 포맷 표준화 동향

□ 이장원 / LG전자

요약

VR(Virtual Reality), AR(Augmented Reality) 컨텐츠의 획득과 소비가 가능한 사용자 디바이스들이 널리 보급되고 있는 가운데, MPEG(Moving Picture Experts Group)에서는 몰입형(immersive) 미디어의 압축과 포맷, 전송에 대한 표준 제정 작업이 활발히 진행 중이다. 본 논문에서는 몰입형 미디어 표준 프로젝트인 MPEG-I와 그 부속 표준의 하나이며 전방향 미디어 포맷에 대한 표준인 OMAF 표준의 기술 전반과 표준 동향에 대해 소개하고자 한다.

I. 서 론

최근 HMD(Head Mount Display)와 360도 카메라 등 몰입형 미디어를 처리하는 디바이스의 사용이 보편화 되면서, 몰입형 미디어 서비스 기술에 대한 관심이 높아지고 있다. CG(Computer Graphics) 기반의 컴퓨터 게임 등에서 주로 사용되던 VR 컨텐

츠 처리 기술들이 실제 촬영, 획득 가능한 실사 영상에까지 확대되어 사용되기 시작하면서, 사용자들은 360도 VR 비디오를 비롯한 몰입형 컨텐츠를 손쉽게 획득, 소비할 수 있게 되었다.

이러한 미디어 시장과 기술의 움직임 속에 MPEG은 다양한 디바이스와 플랫폼 간에 몰입형 미디어의 압축, 공유, 유통을 원활히 하기 위해, 몰입형 미디어 표준 프로젝트인 MPEG-I(Coded Representation of Immersive Media)[1]와 360도 미디어 저장 포맷 기술인 OMAF(Omnidirectional MediA Format)[2] 표준화 작업을 진행하고 있다.

본 고에서는 MPEG-I와 그 부속 표준인 OMAF의 기술 전반과 표준화 동향에 대해 소개하고자 한다. 먼저 II장에서는 MPEG-I의 표준화 현황과 기술 개요에 대해 소개하고, III장에서는 OMAF의 표준화 동향 및 요소기술들에 대해 설명하며, 마지막 IV장에서 향후 전망과 함께 결론을 맺는다.

II. MPEG-I 표준화 현황 및 기술 개요

2016년 10월 중국 청두 회의에서 MPEG은 몰입형 미디어를 위한 차세대 MPEG 프로젝트인 MPEG-I(Corded Representation of Immersive Media) 표준화를 위한 공식 활동을 시작하였다. 본 표준화에서는 새로운 미디어 서비스로 대두되고 있는 VR과 AR 서비스에 활용될 수 있는 몰입형 오디오, 비디오, 그래픽스 및 시스템 기술에 대한 표준화를 주요 목표로 삼고 있다.

MPEG-I는 비디오, 오디오에서 시스템 요소에 이르기까지 전체 컨텐츠 에코시스템을 대상으로 하는 표준 기술로 구성되며, 360 VR 비디오와 MR(Mixture Reality) 및 6-Dof(6 Degrees of Freedom) 관련 컨텐츠를 포함한 몰입형 미디어 전체를 고려 대상으로 한다.

MPEG-I는 총 2단계의 Phase와 8개의 파트로 정의되어 있고, 각 파트마다 서로 다른 표준으로 구성되어 있다. Phase1에서는 360도 VR 영상 시청 시나리오를 목표로 진행하며 Phase2는 6축 자유도를 가질 수 있는 6-Dof VR 영상도 포함을 한다. 이를 위해 각 파트 별로 표준화를 차수 중에 있거나 일부는 이미 진행하고 있다.

MPEG-I 파트 1(Immersive Media Architectures)은 MPEG-I 요구사항이 요약된 몰입형 미디어의 주요 사용 시나리오 및 아키텍처를 정의하며 MPEG-I의 전체 아키텍처를 포함하는 기술 보고서로 구성된다. 기술 보고서에는 또한 MPEG-I의 파트들에서 사용되는 몰입형 미디어와 관련된 주요 용어의 정의가 포함되어 있다.

MPEG-I 파트 2 표준인 OMAF(Omnidirectional MediA Format)는 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, 현재 표준화가 진행 중인 버전 2에는 3-Dof+미디어 지원 및 복수시점 360 비디오, 오버레이, 서브피쳐 트랙 등 버전 1 기능의 확장이 진행 중이다. OMAF의 표준화 현황 및 기술 개요에 대해서는 본 논문의 III장에서 자세히 소개한다.

MPEG-I 파트 3은 MPEG-I 몰입형 미디어 서비스에 중요한 비디오 규격을 정의한다. 이를 위해 HEVC(High Efficiency Video Coding) 표준 이후로 Joint Video Exploration Team(JVET)을 구성하여 차기 영상 압축 기술을 위한 차기 코덱 표준화 작업을 시작하고 있다. 차기 코덱 표준은 VVC(Versatile Video Coding)로 명명되어 표준화가 진행 중이다.

MPEG-I 파트 4는 몰입형 오디오를 위한 규격을 정의한다. 이 부분에는 특히 6-Dof 컨텐츠 표현을 위해 필요한 특정 오디오 압축 및 재생 관련 표준이 포함된다. 차기 오디오 압축 표준의 공식 명칭은 향후 결정될 예정이다.

MPEG-I 파트 5(Point Cloud Coding)는 향후 몰입형 컨텐츠를 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 6-Dof 및 Mixture VR를 가능하게 하기 위한 기술로 알려져 있다.

MPEG-I 파트 6(Immersive Media Metrics)에는 몰입형 미디어 생태계 전반에 걸친 사용자 경험의 품질을 정량화하고 평가할 수 있는 미디어 메트릭 목록이 포함되어 있다.

MPEG-I 파트 7(Immersive Media Metadata)은 MPEG-I의 다른 파트에서 새롭게 정의된 메타데이터를 총괄하는 규격 파트이다. 특히 비디오, 오디오 및 시스템 등 여러 분야의 기술에 걸쳐 각기 다른 표준에서 공통으로 사용될 수 있는 메타데이터의 정의를 주요한 목표로 하고 있다.

MPEG-I 파트 8(Network Based Media Processing)에서는 몰입형 미디어의 네트워크 연동을 위한 인터페이스를 정의한다. 특히 6-DoF 컨텐츠의 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.

MPEG-I 타임 라인은 3 단계로 진행된다. Phase 1a 요구 사항은 360도 전방향 비디오를 가능하게 하는 기술을 지원하는 반면 Phase 1b 및 Phase 2 요구 사항은 각각 3DoF+ 및 6-DoF 미디어 컨텐츠를 지원하는 기술을 목표로 한다. Phase 1a의 표준화는 지난 2017년 말 완료 되었고, Phase 1b 시스템 및 비디오 분야에서 기술 표준화가 진행 중이며, Phase 2는 사용 시나리오 및 요구사항 논의 중에 있고 PCC 등 일부 시나리오 및 기술 분야에 있어서는 기술 표준화가 진행 중에 있다.

이미 시장에 360도 미디어 서비스 및 디바이스가 보급되어 있었기 때문에 이들을 위한 공통의 어플리케이션 포맷 제정을 위한 MPEG-A(Media Application Format)의 부속 표준으로 OMAF를 표준화하는 것으로 참여사들의 의견이 모아졌다. 이후 3-DoF 기반의 360도 미디어 저장 및 전송 규격에 대한 요구사항과 기술규격 설계를 위한 논의들이 본격적으로 이어졌고, 2016년 10월 중국 청두에서 개최된 116차 MPEG 회의에서는 3-DoF 기반의 저장, 전송 규격 표준뿐만 아니라, 6-DoF를 몰입형 미디어를 위한 비디오, 오디오, 그래픽스 및 시스템 기술들을 포괄하는 MPEG-I 표준 제정의 장기적 계획이 수립되었다.

OMAF는 이후 MPEG-I의 부속표준으로 이동하게 되었고, 2017년 말까지 3-DoF 표준화 완료라는 MPEG-I의 계획에 따라 표준화 작업의 속도를 높여, 2017년 1월에 Committee Draft(CD), 4월에 DIS(Draft International Standard), 7월에는 Study of DIS를 발간하였으며, 10월 마카오에서 개최된 120차 MPEG 회의에서 FDIS(Final Draft International Standard)를 발간함으로써 버전 1 표준화를 마무리 하였다. 현재 3DoF+미디어 지원 및 복수시점 360 비디오, 오버레이, 서브피쳐 트랙 등 확장 기능을 포함하는 버전 2 표준화를 2020년 초 완료를 목표로 진행하고 있다.

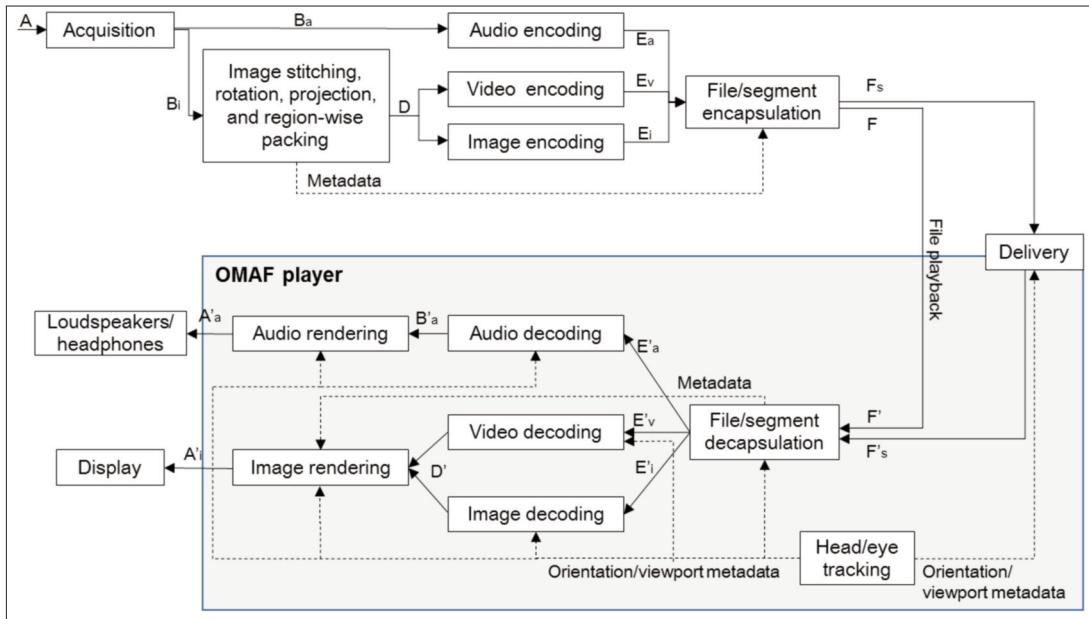
III. OMAF 표준화 현황 및 기술 개요

1. OMAF 표준화 현황

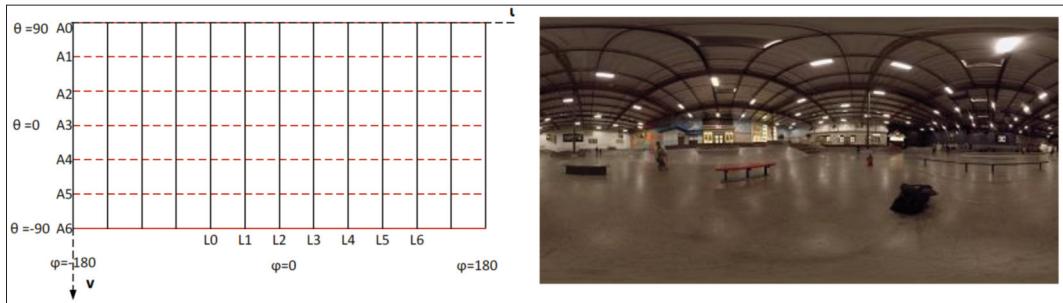
OMAF 표준화는 모바일과 PC 기반의 HMD가 대중화 되기 시작했던 2015년 10월, 스위스 제네바에서 열린 115차 MPEG 회의에서 시작되었다. 당시

2. OMAF 표준 기술 개요

OMAF는 360도 미디어 서비스를 위한 저장 포맷 표준으로, 기존의 2D(2-dimensional) 미디어 처리를 위한 획득, 압축, 저장, 전송을 위한 요소기술들을 기반으로 설계되고 있다. 또한, 360도 이미지, 비디오 처리를 위한 새로운 프로세싱 단계들과 여



(그림 1) OMAF 컨텐츠 동작 흐름도 [2]



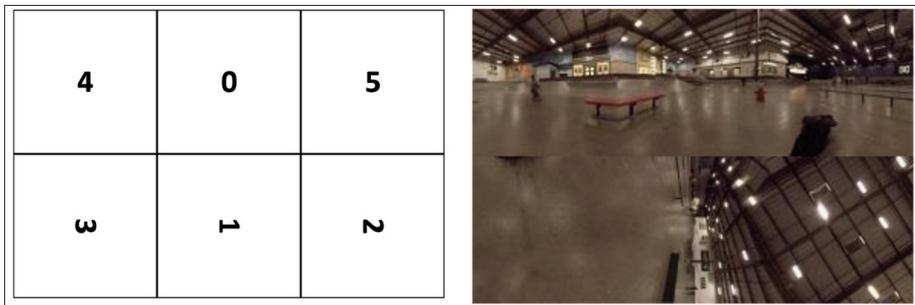
(그림 2) Equi-Rectangular Projection (ERP) [3]

기애 필요한 메타데이터들을 기술 범주로 포함한다. <그림 1>은 OMAF 컨텐츠의 동작흐름을 end-to-end 관점에서 도시하여 보여주고 있다.

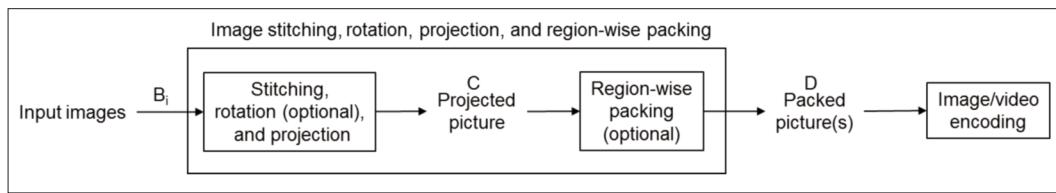
360도 미디어는 오디오 센서와 카메라들로부터 획득된다. 일반적으로 360도 비디오, 이미지는 여러 대의 카메라로부터 획득되며, 획득된 이미지를 하나로 연결하는 스티칭(stitching), 3D 공간상의 전 방향 이미지를 정해진 포맷의 사각형 픽처에 투영하는 프로젝션(projection)과정을 거치게 된다.

다양한 프로젝션 포맷들이 연구/개발되고 있으며, OMAF는 현재 시장에서 보편적으로 사용되는 Equi-Rectangular Projection(ERP), Cube Map Projection (CMP)을 지원하고 있다. <그림 2>, <그림 3>은 각각 ERP, CMP에서의 프로젝션 방법을 도시하여 보여주고 있다.

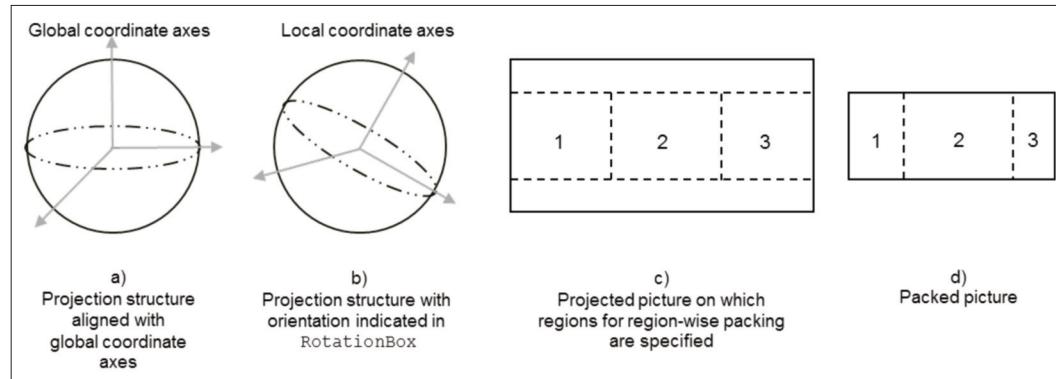
<그림 4>에서와 같이 프로젝션의 결과물인 프로젝티드 픽처(projected picture)는 선택적으로 리전 와이즈 패킹(region-wise packing) 과정을 거칠



〈그림 3〉 Cube Map Projection (CMP) [3]



〈그림 4〉 리전와이즈 패킹 (region - wise packing) [2]



〈그림 5〉 팩트피쳐 생성을 위한 프로세싱 단계 [2]

수 있다. 이 과정에서 프로젝티드 피쳐를 구성하는 영역들의 위치, 크기 및 방향을 변형시킬 수 있으며, 전체적인 압축 효율을 높이거나, 특정 뷰포트 (viewport)에 해당하는 영역의 해상도를 다른 영역에 비해 높게 처리함으로써 사용자 시점 기반 프로세싱에 사용될 수 있도록 피쳐의 구성을 가공하여 팩트 피쳐(packed picture)를 생성한다.

이외에도 OMAF는 〈그림 5〉와 같이 스티칭, 프로

젝션, 리전와이즈 패킹에 걸친 팩트피쳐 생성 과정에 있어서, 3차원 공간의 회전(rotation) 정도와 360도 컨텐츠의 시야범위(coverage)를 시그널링 할 수 있는 메타데이터를 제공한다. 이러한 메타데이터를 이용하여, 중심점(center point) 위치 이동을 통한 압축 효율 향상, 360도 이하의 비디오 컨텐츠 지원 등 다양한 컨텐츠 형태와 목적을 갖는 몰입형 미디어 어플리케이션 서비스가 가능하도록 하고 있다.

OMAF는 미디어 데이터의 프로세싱과 관련된 메타데이터뿐만 아니라, 몰입형 어플리케이션 서비스에서의 사용자 경험 향상을 위한 타임드 메타데이터(timed - metadata)를 제공하고 있다. 초기 시점(initial viewpoint), 추천 뷰포트(recommend viewport)와 같이 360도 미디어의 재생 시간 흐름에 따라 변화하는 시점, 뷰포트 관련 데이터들을 제공하고 있다.

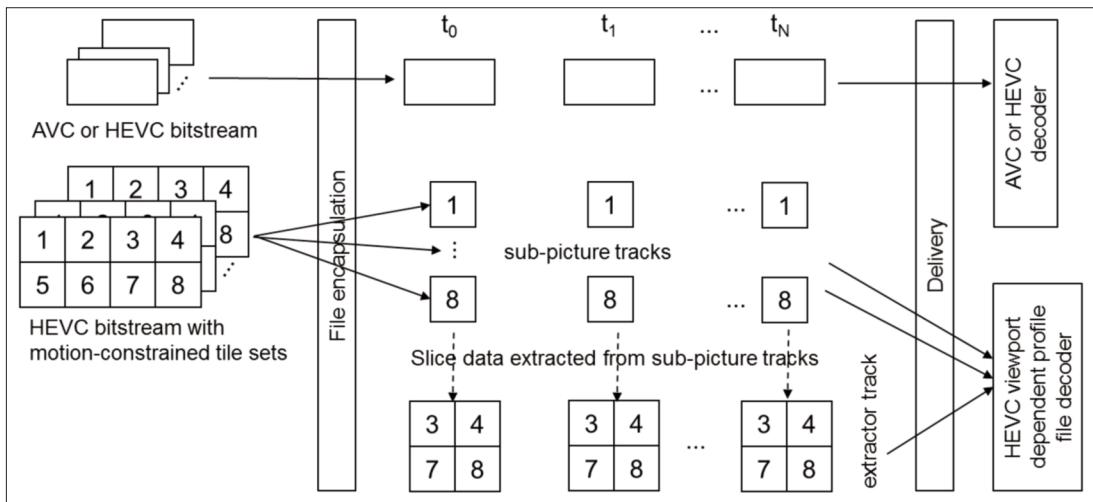
360도 미디어는 HEVC, MPEG-H Audio 등의 규격으로 부호화되며, 앞서 설명한 메타데이터들과 함께 ISO base media file format[4] 기반의 파일 포맷에 저장된다. 저장된 파일은 파일 기반 서비스, 본 장의 후반부에 설명될 DASH(Dynamic Adaptive Streaming over HTTP), MMT(MPEG Media Transport) 기반의 360도 미디어 전송 및 스트리밍을 통해 사용자에게 전달된다.

사용자 디바이스에서는 전달된 파일을 파싱(parsing)하여 미디어 데이터와 메타데이터를 추출하고, 미디어 데이터를 복호화한다. HMD 등의 출

력 장치를 통해 복호화된 미디어 데이터와 메타데이터를 통해 360도 미디어를 3D 공간에 복원하고, head/eye tracking을 통해 사용자 시점에 해당되는 360도 미디어의 부분을 몰입감 있는 형태로 사용자에게 제공하게 된다.

360도 비디오의 압축과 전송에 있어서 많은 관심을 받고 있는 또 하나의 기술요소는 사용자 시점 기반 프로세싱(viewport dependent processing)이다. 360도 비디오의 특성상 사용자는 항상 전체 비디오의 일부에 해당하는 뷰포트만을 시청하고 있으므로, 뷰포트에 해당하는 일부분에 해당되는, 대역폭 등의 자원을 집중하여 사용한다면, 프로세싱 효율을 향상 시킬 수 있을 것이다.

〈그림 6〉은 두 비트스트림과 MCTS(Motion-Constrained Tile Set)를 이용한 사용자 시점 기반 프로세싱의 예를 보여준다. 저화질로 전체 비디오를 부호화하여 하나의 비트스트림을 구성하고, 고화질의 비디오는 MCTS로 분할하여 부호화한다. 고화질의 비디오는 사용자의 뷰포트에 따라 해당



〈그림 6〉 두 비트스트림과 MCTS를 이용한 사용자 시점 기반 프로세싱의 예 [2]

MCTS만 전송, 복호화된다. 사용자 시점에 상관없이 항상 제공되는 저화질의 비디오는 사용자 시점 이동과 실제 뷰포트 이미지가 사용자에게 보여지기 까지의 지연시간(Motion - to - photon latency)에 사용자에게 제공될 수 있어, 사용자 경험 악화를 최소화 할 수 있다.

저장 포맷인 OMAF 표준과 연계하여, 몰입형 미디어의 전송 및 스트리밍을 지원하기 위한 기존 전송기술과 OMAF의 접목 방안 또한 논의되어 OMAF 표준의 일부로 포함되었다.

DASH를 통한 OMAF의 전송 기술은 IP/TCP 환경에서 사용자 시점기반 화질/대역폭 적응적 360도 미디어 서비스를 제공하기 위한 방안들을 표준화 하였다. <그림 7>에 도시된 바와 같이, DASH를 통한 OMAF 전송을 위해 360도 미디어 전송

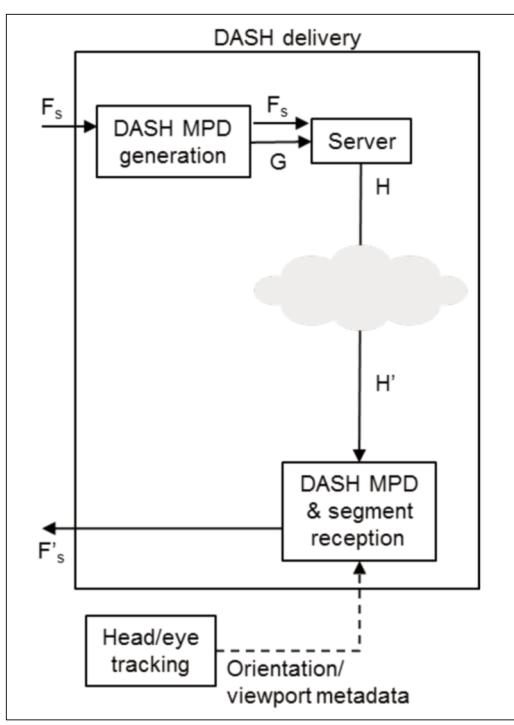
관련 메타데이터 지원을 위한 DASH MPD(Media Presentation Description) 확장 방안과, head/eye tracking을 통해 사용자 시점에 해당하는 DASH segment의 선택적 송수신 방안이 고려되어 있다.

MMT를 통한 OMAF 전송 기술은 IP/UDP 환경에서 360도 미디어를 사용자 시점, 제작자 시점 기반으로 효율적으로 전송하기 위한 방안들을 표준화하였다. <그림 8>에 도시된 바와 같이, MMT를 통한 OMAF 전송을 위해 360도 미디어 전송 관련 메타데이터 생성과, head/eye tracking을 통해 사용자 시점에 해당하는 MMTP Flow의 효율적 송수신 방안이 고려되어 있다.

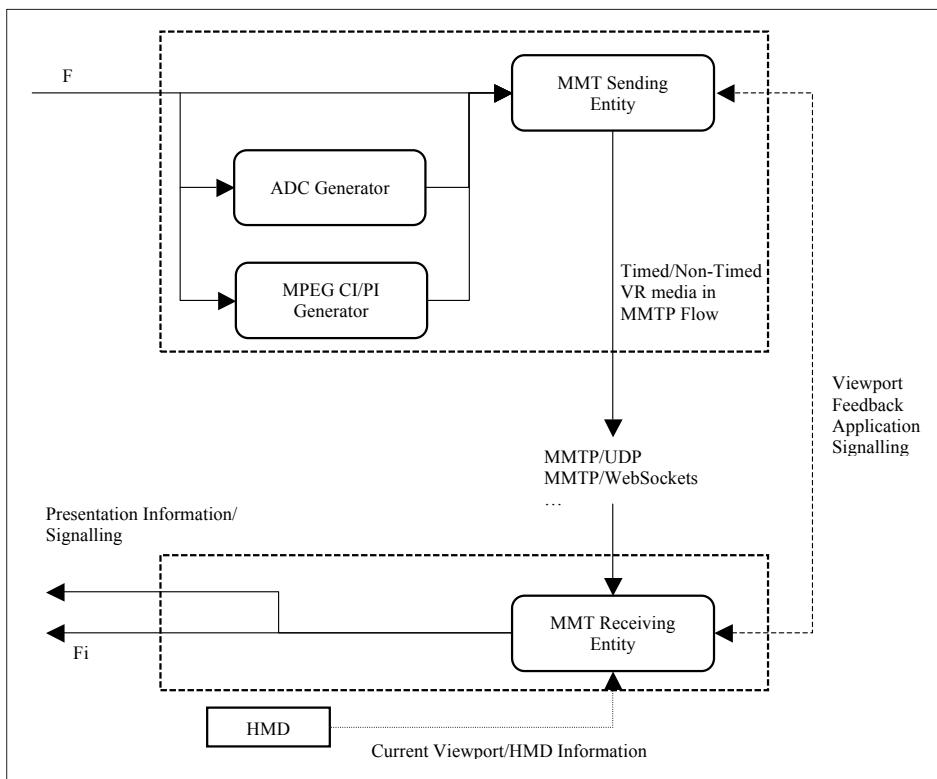
현재까지 소개한 360도 전방향 미디어 지원을 OMAF 버전 1의 기술 요소들을 기반으로, 기능 확장을 위한 OMAF 버전 2 표준화가 현재 진행 중에 있다. OMAF 버전 2는 3-Dof+ 미디어 지원을 기본 목표로 하며, 복수시점 360도 미디어, 오버레이, 360 비디오의 서브픽처(sub-picture) 등 다양한 기능을 제공하고 있다.

<그림 9>는 복수시점 360 비디오 서비스의 형상을 보여주고 있다. 하나의 360도 비디오 시청 중 사용자의 상호작용을 통해 동일한 컨텐츠의 다른 시점에서 획득된 360도 비디오로 이동할 수 있다. 경기장의 트랙을 따라 이동하는 카메라의 예와 같이 복수 시점의 360도 비디오 지원은 고정된 카메라 시점에서뿐만 아니라 가변적 위치의 카메라 시점 또한 지원한다.

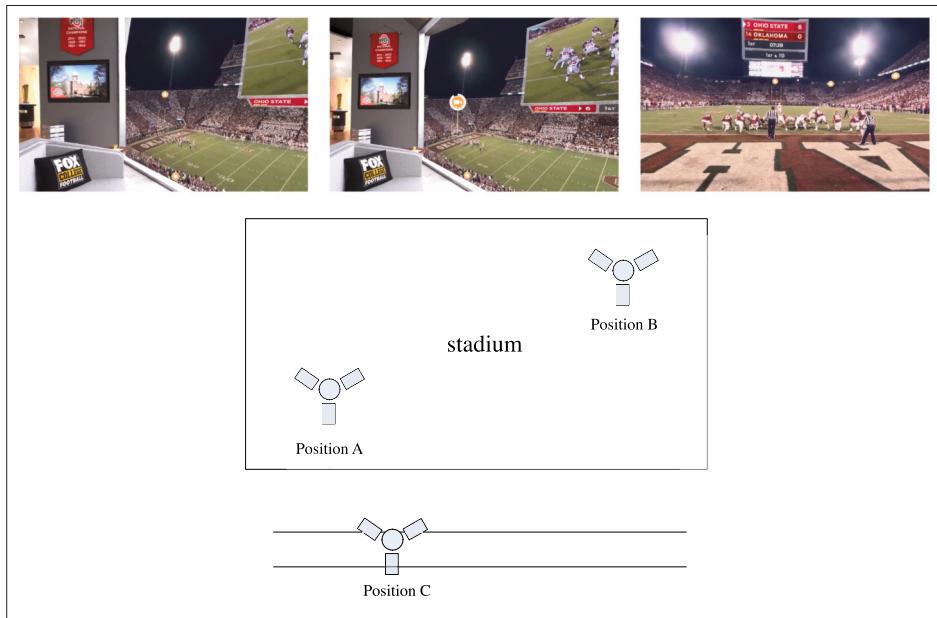
<그림 10>은 오버레이가 포함된 360도 비디오 서비스의 형상을 보여주고 있다. 360 비디오에 2D 또는 3D 비디오, 이미지가 올려진 형태의 영상을 사용자에게 제공할 수 있으며, 광고, 위치 기반 서비스 등 다양한 VR 어플리케이션에서 활용이 가능하



<그림 7> DASH를 통한 OMAF 컨텐츠 전송 [2]



〈그림 8〉 MMT를 통한 OMAF 컨텐츠 전송 [2]



〈그림 9〉 복수시점 360 비디오 지원 [5][6]



〈그림 10〉 오버레이가 포함되는 360 비디오 [7]

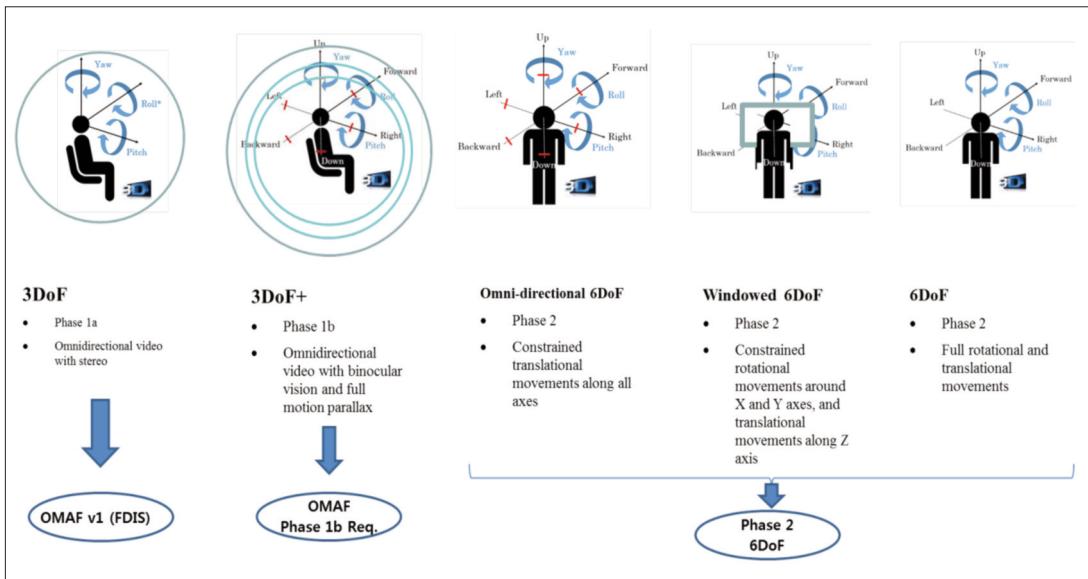
다.

IV. 결 론

지금까지 몰입형 미디어 표준인 MPEG-I와 그 부속 표준으로서 몰입형 미디어 포맷에 대한 시스템 표준인 OMAF의 표준화 동향 및 기술 전반에 대

해 소개하였다.

MPEG-I에서 정의하는 사용 시나리오는 〈그림 11〉과 같이 다양한 몰입 환경을 포함한다. 표준화가 완료된 3-Dof의 360도 전방향 비디오와 현재 표준화가 진행중인 3-Dof+ 전방향 비디오에서 나아가, 앞으로의 네이게이션이 가능한 컨텐츠, Social



〈그림 11〉 MPEG-I의 몰입감 단계 정의 [2]

VR 및 6-DoF 서비스를 위하여 표준화가 진행될 것이며 MPEG-I는 차세대 몰입형 미디어에 대한 산업계 표준이 될 전망이다.

2015년 시작된 OMAF 표준화 작업은 3-DoF 지원을 위한 버전 1 표준화 작업을 마무리하고, 3-DoF+ 및 다양한 사용자 경험 지원을 위한 버전 2 표준화 작업을 활발히 진행하고 있다. 현재까지 MPEG-I와 OMAF는 시장의 요구에 따라 혼존하는 몰입형 미디어 서비스 기술을 표준에 담아내고,

다양한 어플리케이션 서비스 시나리오 지원, 전방 향 미디어 프로세싱 효율성 증대 등 향후 몰입형 미디어 서비스 기술의 방향에 대한 전망도 제시하고 있다. MPEG-I와 OMAF는 표준화 계획에 발맞춰 6-DoF 기반 몰입형 미디어 서비스 기술 제공에 이르기까지 지속적으로 표준화가 진행될 것으로 전망되며, 앞으로도 관련 시장의 요구와 최신의 몰입형 미디어 기술 동향을 충실히 반영하는 표준 기술이 될 것을 기대한다.

참고 문헌

- [1] ISO/IEC JTC1/SC29/WG11/N17685, "Technical Report on Architectures for Immersive Media", MPEG 122, Apr. 2018
- [2] ISO/IEC JTC1/SC29/WG11/N17827, "WD 2 of ISO/IEC 23090-2 OMAF 2nd edition", MPEG 123, July 2018.
- [3] Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, "360Lib Software Manual", MPEG 117, Jan. 2017
- [4] ISO/IEC JTC1/SC29/WG11/N15177, "Text of ISO/IEC 14496-12 5th edition", MPEG 111, Feb. 2015.
- [5] ISO/IEC JTC1/SC29/WG11/M42483, "Support of switching between content from different viewpoints", MPEG 122, Apr. 2018
- [6] ISO/IEC JTC1/SC29/WG11/M42457, "Support of multiple viewpoints", MPEG 122, Apr. 2018
- [7] ISO/IEC JTC1/SC29/WG11/M42466, "Spatial alignment of OMAF visual content", MPEG 122, Apr. 2018

필자 소개

이장원



- 2007년 : 경희대학교 전자공학과 학사
- 2013년 : 경희대학교 전자전파공학과 박사
- 2013년 ~ 현재 : LG전자 CTO부문 선임연구원
- 주관심분야 : image processing, media streaming, broadcasting