

Chapter
02MIV(MPEG Immersive Video)
표준기술 개발 동향

이광순_한국전자통신연구원 책임연구원
서정일_한국전자통신연구원 실장

I. 서론

이머시브 비디오는 시청자가 공간상에서 완전한 6DoF(Degree of Freedom) 자유도의 움직임 때 보여져야 할 영상을 재현함으로써 사실감과 몰입감을 동시에 제공할 수 있는 비디오로 정의할 수 있으며 최근 관련 기술 개발과 표준화가 활발히 진행되고 있다[1],[2]. MPEG에서는 2016년 10월부터 이머시브 비디오 관련 표준 제정을 위해 비주얼(Visual), 시스템 그룹 등에서 MPEG-I(Immersive) 프로젝트명으로 활발히 표준화를 진행해 왔다[3],[5]. MPEG-I 프로젝트는 고정된 중심을 기준으로 3축 회전 운동을 지원하는 Phase 1a(3DoF) 단계, 전방위 장면을 포함한 복수 개의 영상과 이를 기반으로 합성한 다수의 가상 영상을 기반으로 약간의 머리 움직임과 같이 매우 제한된 운동을 지원하는 Phase 1b(3DoF+) 단계 그리고 실제 생활하는 것과 같이 미디어 환경 내에서 자유롭게 이동이 가능한 Phase 2(6DoF) 단계의 총 세 단계로 구성된 로드맵 기반으로 진행되고 있다. 회전 운동에 해당하는 영상의 재현만 지원하는 3DoF의 경우, 즉 운동시차가 지원되지 않을 경우, 시청자는 불편함을 느끼면서 사실감과 몰입감을 느끼는데 한계가 있다고 알려져 왔으므로, 6DoF 기반의 이머시브 비디오에서는 완전한 운동시차의 지원을 기본적인 요구사항으로 다루어왔던 것이다. 특히, 6DoF 기반의 이머시브 비디오에서는 제한된 개수의 카메라로 촬영한 실사 다시점 영상을 처리하여 시청자의 움직임에 따라 연속적이고 실제 공간을 표현할 수 있는 실감 영상을 재현해야 하므로, 중간 시점영상 합성을 고려한 비디오 부호화 기술을 필요로 했다.

* 본 내용은 이광순 책임연구원(☎ 042-860-1676, gslee@etri.re.kr)에게 문의하시기 바랍니다.

** 본 내용은 필자의 주관적인 의견이며 IITP의 공식적인 입장이 아님을 밝힙니다.

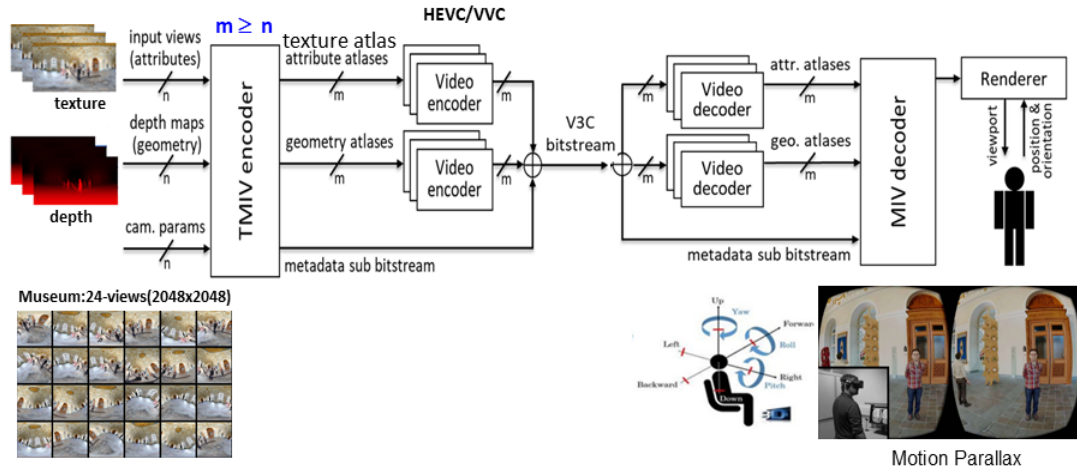
** 본고는 정부(과학기술정보통신부)의 재원으로 정보통신기술평가원의 지원을 받아 수행된 연구임(No.2018-0-00207, 이머시브전문연구실)

6DoF 제공을 위해 필수적인 공간영상 획득의 경우, 구글은 46개 반구형 카메라 어레이 기반 라이트필드(light field) 영상 획득과 딥러닝 기반 3차원 공간 표현 기법인 이머시브 라이트 필드 기술을 공개하였다[4].

최근에는 CG 기반의 게임이나 비현실적 공간으로 표현되는 메타버스에 현실 공간을 표현 하기 위한 기술 개발의 필요성이 대두됨에 따라 실사 기반으로 현실감과 몰입감을 동시에 지원할 수 있는 이머시브 비디오 기술의 활용도가 지속적으로 증대되고 있다. 본 고에서는 국제 표준화 기구인 SC29 WG4에서 표준화한 MIV(MPEG Immersive Video) 표준[5] 기술에 대해 소개하고자 한다.

II. SC29 WG4에서의 이머시브 비디오 표준화

JTC 1(Joint Technical Committee 1)의 SC(Sub Committee)29 산하에는 7개의 작업 반(WG)과 3개의 자문반(AG)이 있으며, 이 중에서 WG4는 MPEG Video coding을 위한 작업반으로, MIV(MPEG Immersive Video), EVC(Essential Video Coding), LCEVC (Low Complexity Enhancement Video Coding), NNC(Neural Network for Multimedia Content Description), LVC(Lenslet Video Coding) 등의 표준화를 담당하고 있다. 현재 MIV 관련 에드혹 그룹(Ad-hoc group)에서 몰입감과 사실감을 부여하는데 필수적인 운동 시차의 지원이 가능한 이머시브 비디오 부호화 표준화를 진행 중에 있다. 표준화 초기에는 3DoF+, windowed 6DoF 비디오 등으로 불려지다가 표준 발간 시점에 명칭을 MIV라고 정하였다. 전방위로 확장된 시청 공간을 구성하는데 필요한 다수 시점영상의 부호화를 위해서는 시스템이 복잡해지고 비디오 코덱의 개수가 증가할 수밖에 없다. 따라서, MIV 표준은 전통적인 비디오 코딩에서의 인코딩 효율 향상과 더불어 입출력 인터페이스, 압축 처리를 위한 코덱의 개수 등을 나타내는 픽셀율(pixel rate) 증가를 억제하고 비교적 적은 수의 2D 비디오 코덱을 통해 사용자의 인터렉션에 따라 최적의 주관적 화질을 재현하는 것에 초점을 맞추고 있다. 즉, [그림 1]과 같이 MIV에서는 전방위 카메라를 통해 획득되는 다수의 텍스처(texture) 및 기하정보(geometry)를 이용하여 공간상의 중복성을 제거한 후 아틀라스(atlas)라고 불리워지는 새로운 인코딩 포맷을 생성하는 TMIV 인코더를 추가함으로써, 적은 수의 비디오 코덱을 통해 압축 처리할 수 있다. 디코더단에서는 다시 후처리와 중간시



(자료) Boyce, Jill M., et al., "MPEG Immersive Video Coding Standard," Proceedings of the IEEE, 2021.

[그림 1] MIV 기반 이머시브 비디오 부복호화 아키텍처

점영상 합성을 통해 시청자의 움직임에 따라 동적으로 끊어짐 없이 뷰포트(viewport) 영상을 재생하도록 한다. 이 모든 과정에 관한 제어정보를 부호화하기 위해서는 메타데이터가 필요하며 이러한 메타데이터와 비트스트림 구조가 주로 표준화되었다. [그림 1]의 예에서는 24개의 반 ERP(Half equi-rectangular projection) 포맷으로 생성된 테스트 영상(Museum) [6]이 TMIV 인코더에 입력되었을 경우이며, 소스 영상이 24개의 텍스처와 깊이영상으로 대용량이었지만 아틀라스는 2쌍의 텍스처와 깊이맵으로 구성될 수 있으므로, 상당히 작은 수의 HEVC 또는 VVC 비디오 코덱을 통해 압축 전송이 가능함을 보여주고 있다.

MIV 표준은 "ISO/IEC 23090-12:2021(E), Information technology - Coded representation of immersive media - Part 12: MPEG Immersive video"라는 표준명으로 제정되었다. MIV 표준의 참조 규격은 ISO/IEC 23090-5, Information technology - Coded representation of immersive media - part 5: V3C(Visual Volumetric Video-based Coding) and V-PCC(Video-based Point Cloud Compression)이다[7]. 즉, 패치방식의 아틀라스 포맷은 포인트클라우드 표준 중의 하나인 V-PCC와 공통적인 규격으로 표준화되었고 이러한 공통적인 규격은 V3C에 제정하고 있으며, part12에는 MIV 특화된 규격만 정의하고 있다. MIV 표준은 2021년 7월에 MIV FDIS(Final Draft International Standard)[5]로 발간되었으며 10월 초에 최종적으로 SC29에 공식적으로 접수되었다. 2021

년 10월 미팅에서 MIV 비트스트림의 적합성 시험(conformance test)을 위해 18개의 포인터를 발췌하였으며 이를 테스트하기 위해 8개의 비트스트림을 생성하였고, 테스트한 결과는 WD 문서(WD3 of ISO/IEC 23090-23:2021(E))로 발간되었다.

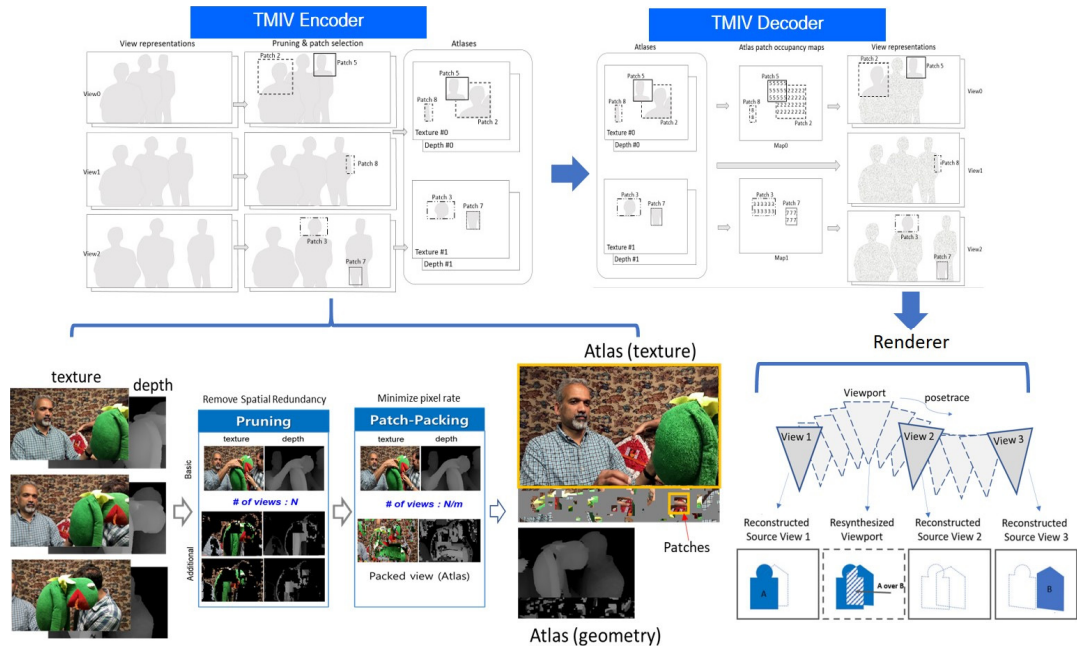
III. 이머시브 비디오 부호화를 위한 MIV 표준기술

SC29 WG4에서는 2019년 3월 회의 때 CfP(Call for Proposal)[8]에 대응하여 제안된 다섯 개의 기술을 기반으로 “TMIV(Test Model for Immersive Video)[9]”라는 테스트 모델을 통해 본격적인 표준 기술을 개발하기 시작하였다.

1. MIV 부복호화 표준 기술 개념

MIV 표준 테스트모델인 TMIV에서의 비디오 부호화 개념은 [그림 2]와 같다[9]. TMIV 인코더의 주된 역할은 소수의 일반적인 비디오 인코더를 통해 부호화 가능한 아틀라스를 생성하는 것이다. 이는 입력되는 다중시점 형태의 전방위 비디오로부터 3D 공간상에서 대응되는 픽셀들의 중복성을 줄이기 위해 가지치기라고 해석되는 푸루닝(pruning)과 패치 패킹(patch packing) 과정을 통해 주로 가능해진다. 이를 다르게 표현하자면, TMIV 인코더는 다수의 시점영상들을 기하정보들을 통해 공간상에 매핑하고 소스카메라의 위치에서 보여지는 픽셀들을 패치 형태로 추출한 후 아틀라스 영상포맷으로 변환한다. 푸루닝 과정에서는 기하정보를 이용해서 모든 소스 시점의 픽셀들을 3D 공간상으로 역투영하고 여기서 중복성을 제거하는 방식을 취하는데, 깊이정보가 정확하다면, 단말에서 뷰포터 렌더링에 필요한 픽셀들만 추출할 수 있다. 하지만 현실적으로 실사 영상에는 카메라 파라미터, 깊이맵 등의 기하 정보가 불완전하므로 약간의 공간적인 중복성을 허용함으로써, 비트율 측면에서는 증가가 있더라도 렌더링 관점에서 품질을 향상시키는 것이 필요하다.

단말측에서는 메타데이터를 파싱하여 픽셀별로의 패치정보를 기록한 아틀라스 패치 점유 지도(Atlas Patch Occupancy Map)를 생성하는데, 이는 패치가 서로 겹쳤을 때 디코딩된 아틀라스 영상으로부터 픽셀별로 시점영상을 분리하는 데 이용된다. 렌더러(renderer)는 디코딩되어 분리된 패치들로부터 시청자가 바라보는 방향에 해당하는 중간시점의 영상을



〈자료〉 B. Salahieh, B. Kroon, J. Jung, M. Domanski(Eds.), "Test model 6 for Immersive Video," ISO/IEC JTC1/SC29/WG11, N19483, Jul. 2020.

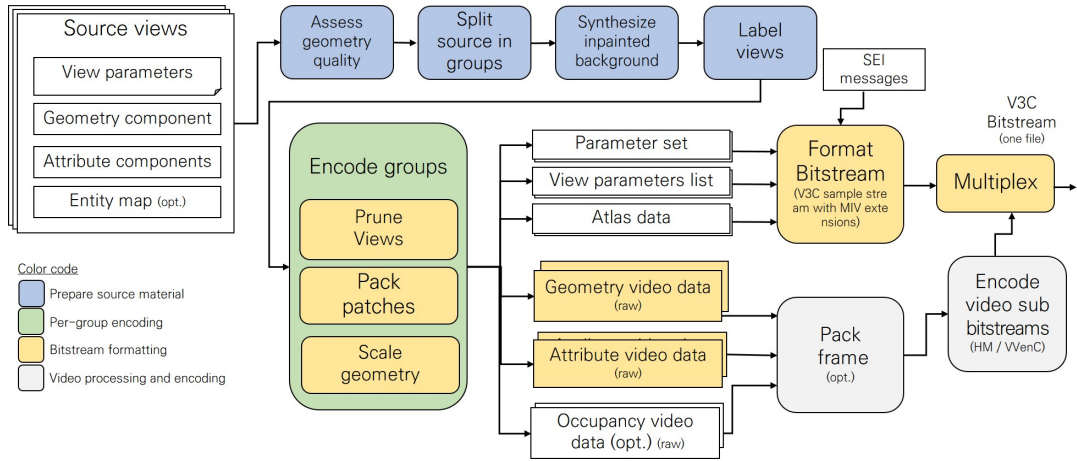
〈그림 2〉 MIV에서의 TMIV 인코더/디코더 개념

합성하며, 시청자가 선택한 시청경로에 해당하는 포스트레이스(posetrace)에 따라 연속적으로 뷰포트 영상을 제공하게 된다.

2. MIV 부호화에서의 TMIV 인코더 기술

〈그림 3〉과 같이 TMIV 인코더의 입력으로는 깊이맵으로 표현되는 기하정보와 칼라정보(texture), 투명(transparency) 정보 등으로 구성되는 속성정보(attribute component), 객체기반의 부호화를 위해 필요한 엔티티 맵(entity map), 내부/외부 카메라파라미터 정보 등으로 구성되는 시점파라미터(view parameter) 등이 될 수 있다. MIV에서의 압축 데이터량 및 렌더러의 품질은 전적으로 깊이 즉 기하 정보의 품질에 달려 있다. 따라서 TMIV에서는 기하정보의 품질을 자동으로 평가하여 좋은 품질인지 나쁜 품질인지에 대해 시그널링할 수 있으며 실사 콘텐츠 및 CG 콘텐츠를 구분하는 용도로 사용할 수 있다.

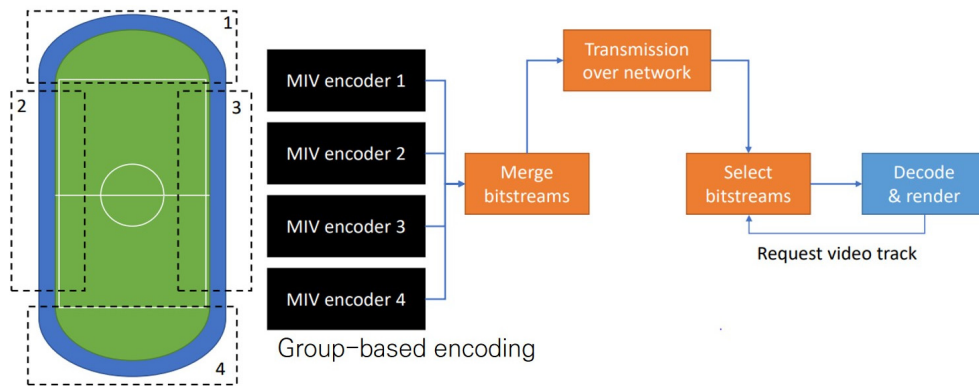
TMIV는 기본적으로 그룹 인코더(Group encoder)를 채택하고 있는데, 그룹의 개념은



〈자료〉 B. Salahieh, B. Kroon, J. Jung, M. Doma-ski(Eds.), "Test model 11 for Immersive Video," ISO/IEC JTC1/SC29/WG11, N0142, Oct. 2021.

[그림 3] MIV에서의 TMIV 인코더 구조

전방위 360 공간에서 획득된 텍스처와 기하정보를 공간적으로 분할하고 분할된 공간별로 TMIV 인코딩하여 아틀라스 영상을 생성한 후, 각각 비디오 인코딩하는 것이다. 그룹 개념이 필요한 이유는 시청자가 원하는 임의 공간으로의 직접적인 접근(spatial access)을 위해 비트스트림별로 부분 디코딩이 가능하게 하기 위함이다. 그룹인코더의 개수는 비트스트림(geometry, attribute)으로 인코딩될 아틀라스 영상의 개수에 따라 결정되며, 포맷 비트스



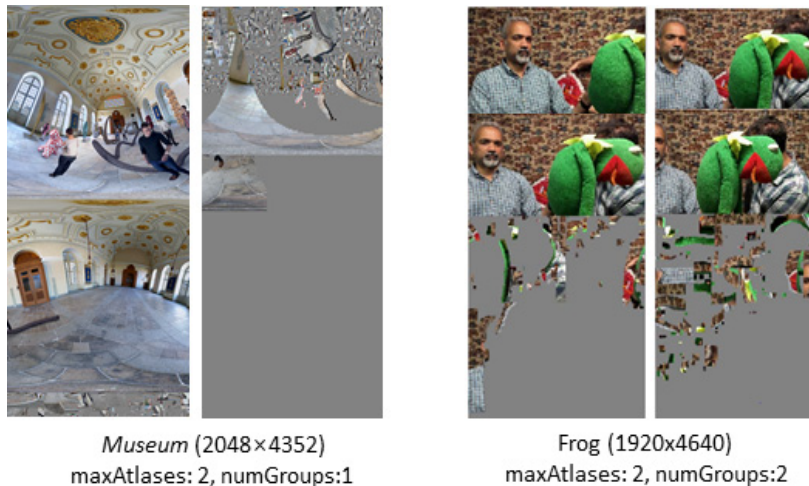
〈자료〉 B.Kroon, "MIV tutorial," 2021 International Conference on Visual Communications and Image Processing (VCIP), IEEE, 2021.

[그림 4] MIV 그룹 인코더를 이용한 이머시브 비디오 서비스 시나리오

트림(format bitstream)에서는 아틀라스, 시점, 패치 등과 관련된 메타데이터를 MIV 표준에 부합되도록 포맷팅하여 별도의 비트스트림으로 생성하며, 이 과정에서 SEI(supplemental enhancement information) 형태의 메타데이터도 포함할 수 있다. MIV 그룹 인코더를 이용한 이머시브 비디오 서비스의 대표적인 예는 [그림 4]와 같다. 축구장과 같이 넓은 공간일 경우, 단일 TMIV 인코더로 모든 공간에 대한 이머시브 비디오를 압축 전송하고 이를 한꺼번에 디코딩하는 데는 한계가 있다. 따라서, 공간을 적절히 나누고 분할된 공간별로 단일 TMIV 인코더로 인코딩한 후 생성된 비트스트림을 다중화하여 디코더에 전송하면, 디코더 측에서는 시청자가 선택한 뷰포트 영상에 필요한 비트스트림을 추출하여 디코딩 및 렌더링하게 된다.

3. TMIV 인코더에서의 세부 기술

TMIV 인코더는 입력되는 여러 시점의 영상으로부터 패치를 추출하여 패키징한 후 출력되는 아틀라스의 해상도를 자동으로 계산한다. 일반적인 비디오 코덱 표준에 정의된 레벨이 지원하는 최대 픽셀율을 기준으로 단말에서 가용한 디코더의 개수에 의해 계산되는데, 여기서 디코더의 개수는 아틀라스의 개수에 해당한다. 최종적으로 생성될 아틀라스의 수평해상도는

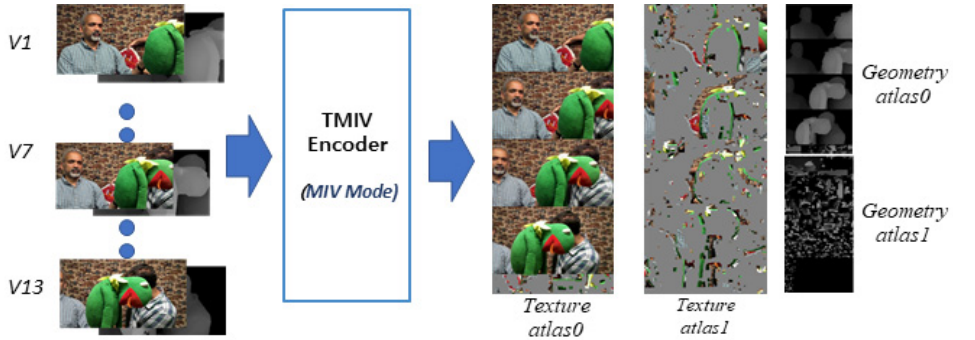


〈자료〉 B. Salahieh, B. Kroon, J. Jung, M. Doma-ski(Eds.), "Test model 11 for Immersive Video," ISO/IEC JTC1/SC29/WG11, N0142, Oct. 2021.

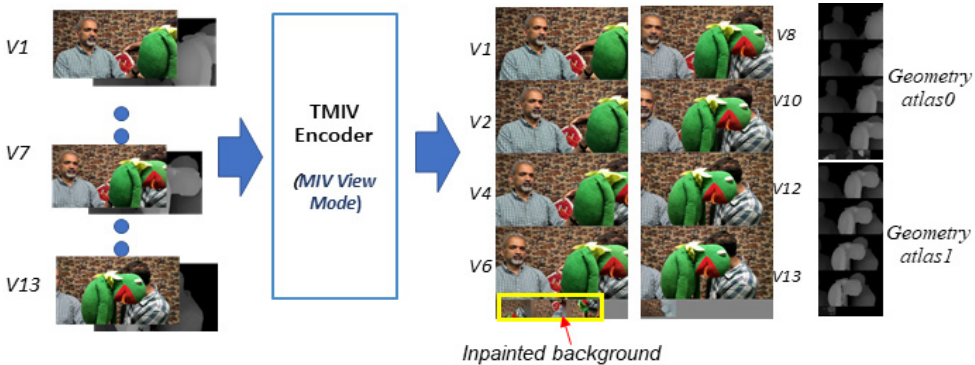
[그림 5] MIV에서 아틀라스 비디오 포맷 예

소스해상도의 수평해상도와 동일하게 유지되고 수직해상도는 전체적으로 4K 혹은 8K 해상도의 픽셀율에 맞도록 가변적으로 결정된다. 즉, 아틀라스는 바로 디스플레이하기 위한 영상이 아니라 기본시점 및 부가시점들을 전달하기 위한 컨테이너로서의 역할만 하므로 해상도는 자유롭게 설정되고 그 크기는 메타데이터를 통해 시그널링된다. [그림 5]는 아틀라스 영상 포맷의 2가지 예를 보여주고 있는데, 좌측은 그룹의 개수가(numGroups) 한 개이고 아틀라스 개수(maxAtlases)가 두 개인 경우, 우측은 그룹의 개수가 두 개이면서 아틀라스 개수가 두 개인 경우이다. 여기서 아틀라스의 수평해상도는 소스해상도의 수평해상도와 동일하고, 수직방향으로 4K 해상도의 픽셀율에 맞도록 가변적으로 결정되었음을 알 수 있다.

다시점 영상 합성에서 난해한 과정 중의 하나는 폐색영역을 채우는 인페인팅(inpainting)



(a) MIV 모드



(b) MIV View 모드

〈자료〉 J.Jung, B.Kroon, "Common Test Conditions for MPEG Immersive Video," ISO/IEC JTC1/SC29/WG4, N0143, Oct. 2021,

[그림 6] MIV에서 TMIV 인코더 동작 모드

기술로서 TMIV에서는 [그림 3]의 합성 인페인티드 배경(Synthesize inpainted background) 블록을 통해 인코더에서 일부 처리한다. 디코더에서도 기본적인 인페인팅을 하지만 인코더 측에서 이를 수행함으로써, 푸루닝되기 전의 모든 소스영상으로부터 폐색영역을 보간할 수 있으며 딥러닝 등의 복잡한 연산을 사용함으로써 인페인팅의 품질을 향상시킬 수 있다. TMIV 인코더는 MIV 모드 및 MIV 뷰(view) 모드로 동작할 수 있다. [그림 6]과 같이 MIV 모드는 푸루닝을 허용한 패치들을 포함하고 있고, MIV 뷰 모드는 푸루닝 연산 없이 모든 시점영상으로부터 공간 영상 재생에 필요한 몇 개의 기본시점을 선택하고 인페인팅된 배경과 함께 패키징한다. 즉, MIV 모드는 기하정보가 정확한 소스영상으로부터 푸루닝에 필요한 공간정보를 비교적 정확히 추출할 수 있는 환경에서 비실시간 인코딩을 위해 사용되며, MIV 뷰 모드는 그렇지 않은 환경에서 고속 인코딩을 위해 사용될 수 있다. [그림 6]과 같이 MIV에서는 기하정보 즉 깊이정보를 인코딩할 때, 데이터양 특히 픽셀울의 최소화를 위해 기하 아틀라스의 크기를 일정 비율만큼 축소한다. 깊이맵 영상은 텍스처 영상에 비해 낮은 주파수 성분으로 구성되어 있어, 원래의 크기로 복원했을 때 상대적으로 정보 손실이 덜 발생한다. TMIV 인코더는 축소 작업에 2×2 크기의 “최대 풀링(max pooling)” 필터를 사용하여 렌더링 품질에 중요한 전경 객체의 정보를 최대한 보존하도록 했다. 이때 디코더에서 복원에 필요한 부가 정보는 기하 업스케일링 파라미터(Geometry upscaling parameters) SEI 정보로 보내진다.

IV. MIV 표준 프로파일

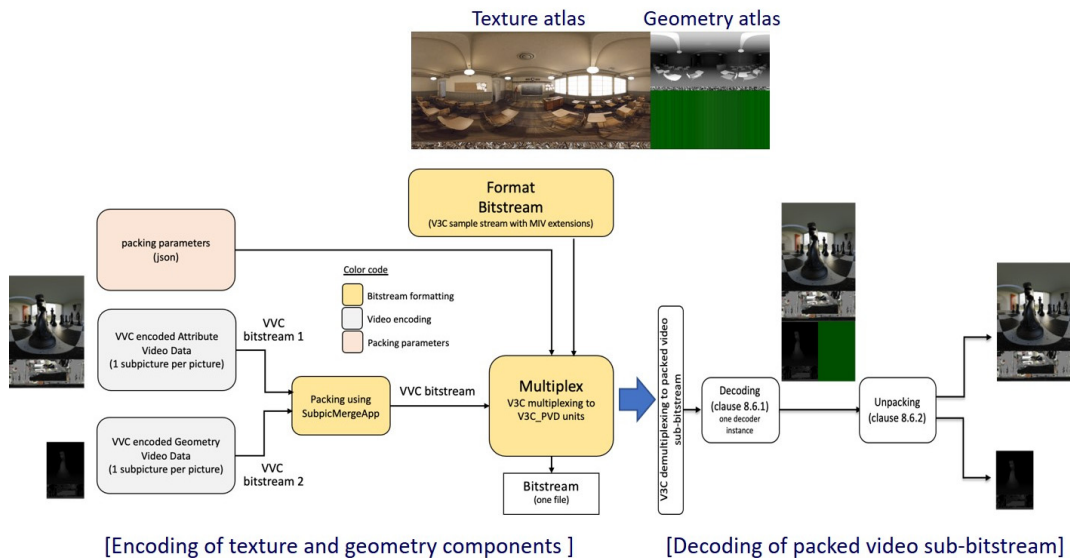
MIV 표준에서 규정한 프로파일과 이의 대표적인 선택스는 [표 1]과 같다. MIV Extended 프로파일은 아틀라스 프레임 패키징 기능, 별도의 점유지도 인코딩, 2개의 속성 정보 인코딩 기능 등을 포함하고 있다. MIV에서의 아틀라스 프레임 패키징(frame packing)은 텍스처, 기하 및 점유지도 아틀라스들을 단일 아틀라스로 패키징함으로써, 디코더에서의 동기화를 용이하게 하고 실행되는 디코더의 개수를 줄일 수 있게 한다. 즉, TMIV에서의 프레임 패키징은 [그림 7]과 같이 VVC에서 지원하는 서브픽처 병합(subpicture merging) 툴을 이용하여 여러 개의 아틀라스를 인코딩한 VVC 비트스트림을 단일 VVC 비트스트림으로 병합함으로써 구현될 수 있다.

[표 1] MIV 표준 프로파일

Key Syntax element	MIV Main	MIV Extended	MIV Extended Restricted Geometry (MPI)	MIV Geometry Absent
vps_packing_information_present_flag	0	0, 1	0, 1	0, 1
vps_occupancy_video_present_flag	0	0, 1	0	0
vps_geometry_video_present_flag	1	0, 1	0	0
vme_embedded_occupancy_enabled_flag	1	0, 1	0	0
al_attribute_count	0.1	0, 1, 2	2	0, 1
al_attribute_type_id	ATTR_TEXTURE	ATTR_TEXTURE, ATTR_TRANSPARENCY	ATTR_TEXTURE, ATTR_TRANSPARENCY	ATTR_TEXTURE

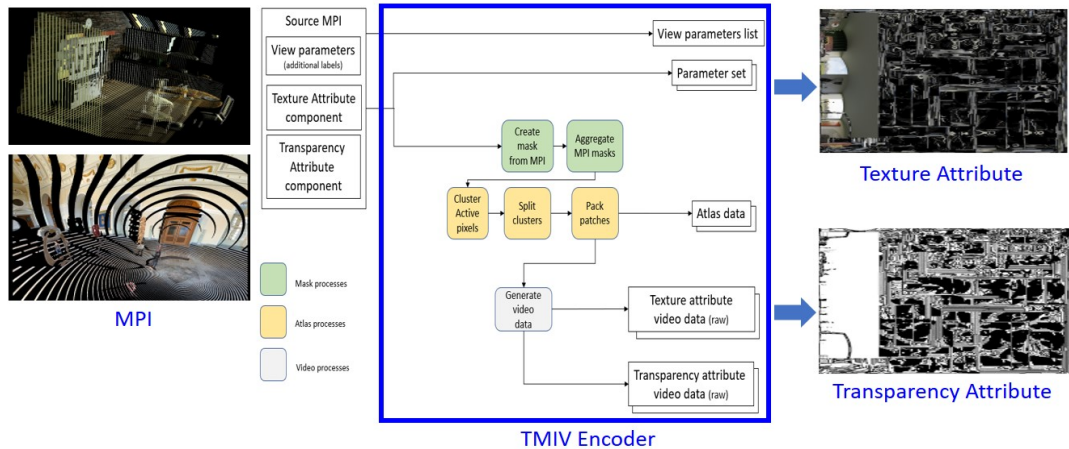
〈자료〉 J.Boyce, V.K.Malamalkital, B.Chupeau, "Information technology – Coded representation of immersive media – Part 12: MPEG Immersive video", ISO/IEC JTC1/SC29/WG4, N00111, Oct. 2021.

MIV Extended Restricted Geometry 프로파일은 [그림 8]과 같이 MPI(Multiple Plane Image) 포맷 형태의 이머시브 비디오를 부호화하기 위한 것으로서, 인코딩되는 속성을 투명정보(transparency)를 포함하여 2개로 고정하며, 계층 정보를 표현하기 위해서 연속적인 기하정보를 불연속적인 기하 값(restricted geometry)으로 인코딩한다. MPI는 3D



〈자료〉 B. Salahieh, B. Kroon, J. Jung, M. Doma-ski(Eds.), "Test model 11 for Immersive Video," ISO/IEC JTC1/SC29/WG11, N0142, Oct. 2021.

[그림 7] MIV에서의 아틀라스 프레임 패키징

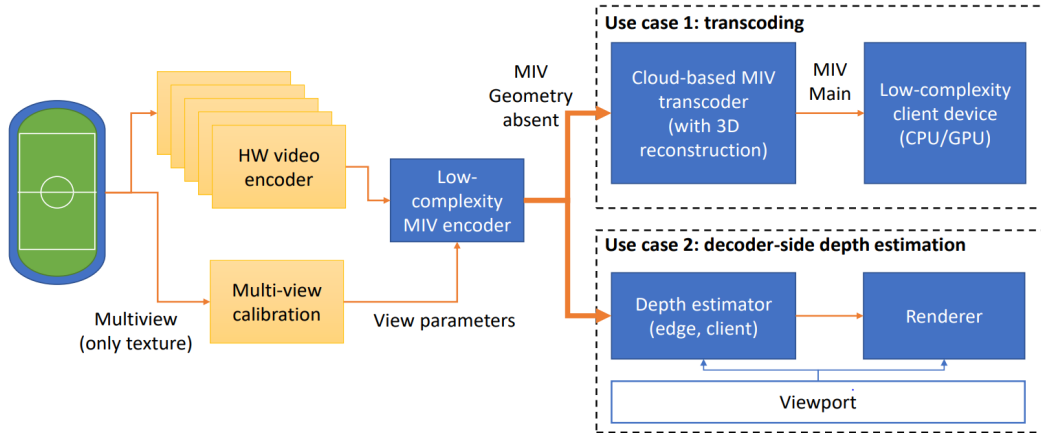


〈자료〉 B. Salahieh, B. Kroon, J. Jung, M. Domansk(Eds.), "Test model 6 for Immersive Video," ISO/IEC JTC1/SC29/WG11, N19483, Jul. 2020.

[그림 8] MIV에서 MPI 기반 이머시브 비디오 부호화

공간을 깊이 방향으로 계층화하는데 이는 중앙 카메라를 기준으로 일정 간격의 계층화된 면에 텍스처를 재투영하여 계층화된 깊이 영상(layered depth image)으로 표현하는 것이다. 이 계층화된 텍스처 및 투명 정보를 MIV의 아틀라스로 패키징하여 압축 전송하면 단말에서는 비교적 간단한 연산을 통해 렌더링할 수 있게 된다.

MIV Geometry Absent 프로파일은 중간 시점영상 합성에 필요한 기하 정보를 인코더로부터 전송하지 않고 디코더에서 직접 추출하게 하는 기술들을 포함한다. 이때 인코딩하는 아틀라스는 기본시점 영상만으로 패키징된 MIV 뷰 모드를 사용함으로써 저사양의 인코더를 사용할 수 있다는 장점이 있지만, 디코더를 상당히 복잡하게 하는 단점이 있다. 따라서 디코더의 복잡도를 줄이기 위해, 깊이 추출에 필요한 특징들을 미리 서버에서 추출하여 기하 보조(Geometry assistance) SEI를 통해 전송하고 이 특징들을 이용하여 디코더에서 고속으로 깊이 추출을 할 수 있게 한다. MIV Geometry Absent 프로파일을 이용한 응용 시나리오는 [그림 9]와 같이 우선 다중 시점영상을 여러 대의 비디오 인코더를 통해 실시간으로 획득한다. 이때 멀티뷰 캘리브레이션 틀에서는 3D 기하정보 추출을 위한 카메라 정보를 추출하여 MIV 인코더를 통해 다중화하여 전송한다. 그러면 첫 번째 시나리오로서 복잡한 연산이 가능한 클라우드 서버에서 깊이정보를 추출하고 3D 복원을 함으로써, 시청자가 선택한 장면을 실시간으로 생성한다. 두 번째 시나리오로서 MIV인코더에서 보내진 특징들을 이용



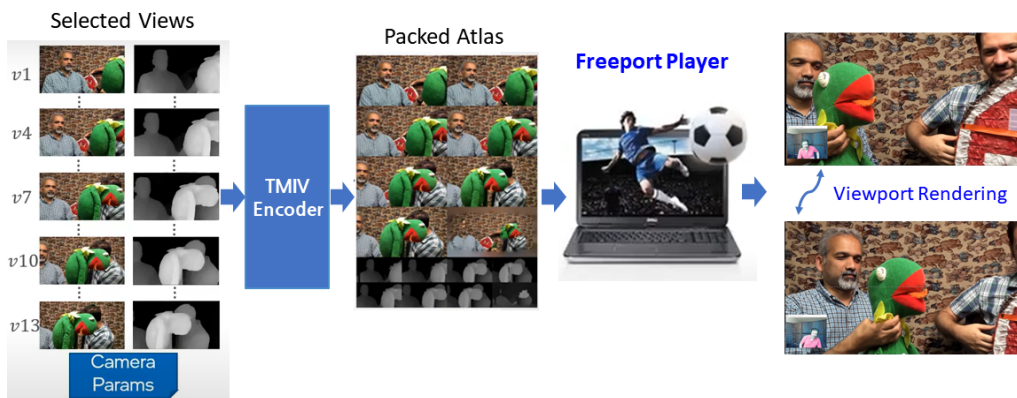
〈자료〉 B. Kroon, "MIV tutorial," 2021 International Conference on Visual Communications and Image Processing (VCIP). IEEE, 2021.

[그림 9] MIV Geometry Absent 프로파일 기반의 이머시브 비디오 서비스 시나리오

하여 고속으로 깊이를 추출함으로써 디코더에서 직접 시청자가 바라보는 뷰포트 영상을 합성하여 재현할 수도 있다.

V. MIV 표준 개발 사례 및 향후 전망

2021년 1월 MPEG 표준화 회의에서 인텔(Intel)은 자사 GPU 칩이 탑재된 노트북을 통해



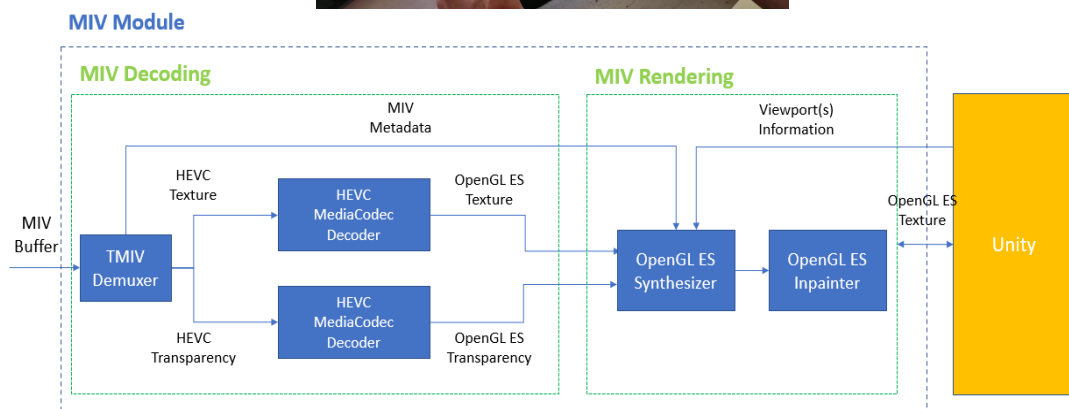
〈자료〉 B. Salahieh, et al., ISO/IEC JTC 1/SC 29/WG 4 m55800, Jan. 2021.

[그림 10] Intel사의 MIV 기반 Freeport player

MIV extended 프로파일로 압축된 비트스트림의 실시간 디코딩과 렌더링 기술을 선보였다. [그림 10]과 같이 다시점 영상 형태의 이머시브 비디오를 TMIV 인코더를 통해 인코딩한 후, 노트북에서 듀얼 HEVC 디코더 칩으로 디코딩하였다. 동시에 딥러닝 기반의 헤드 트래킹 기술을 이용하여 시청자의 시청방향을 추적하고 그 방향으로 시점영상을 실시간으로 렌더링함으로써, 일반적인 모니터에서도 이머시브 비디오 서비스가 가능함을 선보였다.

2021년 MPEG 표준화 4월 회의에서 인터디지털은 MIV MPI가 스마트 단말에서와 같은 저사양 디바이스에서 구현 가능함을 선보였다. MPI는 단말에서 가장 간단하게 렌더러를 구현할 수 있는 MIV 프로파일이기 때문에 [그림 11]과 같이 태블릿에서 실시간으로 MV 콘텐츠를 렌더링할 수 있었다. 최종적으로 사용자의 터치 조작에 의해 렌더링되는 시점이 변화하여 몰입감이 증대되는 것을 선보였다.

현재 MIV 표준화는 2022년 초에 MIV IS의 발간을 목표로 함에 따라 적합성 시험을 진행 중이며, 동시에 MIV 표준기술의 우수성을 검증하고 홍보하기 위해 검증 시험(verification test)을 진행하고 있다. 2021년 4월 회의 때부터 MIV 표준의 확장을 목표로 MIV edition



<자료> J. Fleureau, et al., ISO/IEC JTC 1/SC 29/WG 4 m 56727, Apr. 2021.

[그림 11] Interdigital사의 MIV player

-2를 위한 유저케이스(use-case) 및 요구사항을 작성하였으며, 현재까지의 유저케이스는 총 8건이다[12]. 여기서 MIV 대비 재현 가능한 시청 공간의 확장, 조명에 의해 반사되는 영역의 렌더링 품질 고도화 등의 요구사항 만족을 목표로 하고 있다. 이를 위해 MIV 표준을 확장하기 위한 MIV edition-2의 표준화는 2022년 4월부터 본격적으로 시작하여 2023년 7월에 완료한 후 FDIS를 발간할 예정이다.

MIV는 V3C를 통해 포인터클라우드 부호화 표준인 V-PCC와의 공통적인 부분을 공유하고 있어 포인터클라우드와 연동된 다양한 이머시브 미디어서비스를 제공할 수 있을 것으로 기대되고 있다. 특히, MIV edition-2 유저케이스에 포함된 객체 기반 MIV 콘텐츠 합성 및 부호화(Object-based MIV content composition and compression)는 MIV를 포함하여 V-PCC를 위해 제작된 포인터클라우드, 메시(mesh) 등의 이종 객체미디어를 MIV를 통해 제공함으로써, 실사와 CG 객체가 융합된 다양한 메타버스 서비스 구현을 목표로 하고 있다. 즉 현재 메타버스는 CG 위주로의 비디오가 제공되고 있지만, 다양한 환경에서 제작된 실사, CG 모델링 객체 등이 MIV edition-2를 통해 제공된다면 다양한 비즈니스 모델의 개발이 활성화될 것으로 기대되고 있다.

● 참고문헌

- [1] "MPEG-I Use Cases for omnidirectional 6DoF, windowed 6DoF, and 6DoF," ISO/IEC JTC1/SC29/WG11, w16768, Apr. 2017.
- [2] M. Wien, J. M. Boyce, T. Stockhammer, and W.-H. Peng, "Standardization Status of Immersive Video Coding," IEEE Jour. Emerg. Select. Topics Circuits Syst., Vol.9, No.1, Mar. 2019, pp.5-17.
- [3] ISO/IEC 23090-2, Information technology - Coded Representation of Immersive Media - Part 2: Omnidirectional Media Format(OMAF) 2nd Edition.
- [4] Broxton, Michael, et al., "Immersive light field video with a layered mesh representation," ACM Transactions on Graphics(TOG) 39.4, 2020, 86-1.
- [5] J. Boyce, V.K.Malamalkital, B.Chupeau, "Information technology - Coded representation of immersive media - Part 12: MPEG Immersive video," ISO/IEC JTC1/SC29/WG4, N00111, Oct. 2021.
- [6] 이광순, 서정일, "MPEG 이머시브 비디오 표준기술 동향과 전망", IITP, 주간기술동향 1969호, 2020. 10. 21.
- [7] Information Technology - Coded Representation of Immersive Media - Part 5: Visual Volumetric

- Video-Based Coding(V3C) and Video-Based Point Cloud Compression(V-PCC), Standard ISO/IEC23090-5, N19329, May 2020.
- [8] Call for Proposals on 3DoF+ Visual, ISO/IEC JTC1/SC29/WG11 MPEG/N18145, Jan. 2019, Marrakesh, Morocco.
- [9] B. Salahieh, B. Kroon, J. Jung, M. Doma-ski (Eds.), "Test model 11 for Immersive Video," ISO/IEC JTC1/SC29/WG11, N0142, Oct. 2021.
- [10] ISO/IEC 23008-2: Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 2: High efficiency video coding.
- [11] Boyce, Jill M., et al., "MPEG Immersive Video Coding Standard," Proceedings of the IEEE 2021.
- [12] J.Jung, B.Kroon, "Common Test Conditions for MPEG Immersive Video," ISO/IEC JTC1/SC29/WG4, N0143, Oct. 2021.
- [13] V.K.Malamalkital, B.Kroon, "Draft use-cases and requirements for MIV - edition 2," ISO/IEC JTC1/SC29/WG4, N0143, Oct. 2021.