

Light Field 미디어 기술 개발 및 표준화 동향

□ 정원식, 이광순, 윤정일, 윤국진, 서정일 / 한국전자통신연구원

요약

최근 차세대 미디어 중 하나로 가상의 환경 내에서 극대화된 현실감 및 몰입감을 사용자에게 제공해 줄 수 있는 몰입형 미디어(Immersive Media)가 각광받고 있다. 이를 반영하듯 360도 전방위 영상을 이용한 VR미디어에 사용자의 움직임에 따라 자연스러운 사실감 및 입체감을 제공하는 Light Field 미디어 기술을 접목하려는 시도가 가속화되고 있으며, MPEG에서도 6DoF(Degrees of Freedom)기반의 LF 미디어를 서비스하기 위한 MPEG-I(Immersive) 표준화가 진행 중에 있다. 본 고에서는 LF 미디어 특성 및 이를 획득하고 재생하기 위한 요소기술 개발 현황을 살펴보고, LF 기술이 적용된 MPEG 표준화 동향을 소개한다.

I. 서 론

몰입형 미디어는 공간과 시간의 제약을 극복하고

인간에게 자연스러운 사실감과 몰입감을 제공하는 미디어로 UHD 이후의 차세대 미디어로 각광받고 있으며, MPEG과 같은 국제 표준화 단체에서도 몰입형 미디어 기술에 대한 표준화를 시작하고 있다. 시청자에게 자연스러운 사실감과 몰입감을 제공하기 위해서는 영상의 품질, 시청자 움직임에 맞는 영상 제공을 위한 인터렉션 및 실계와 같은 입체감 등이 필요하며, 시각 피로 없는 완전 입체 영상의 제공이 필수 요소라 할 수 있다.

완전한 입체감을 위해서는 인간이 입체감을 느끼는 물리적 요인을 모두 제공하여야 하며, 이는 크게 양안 시차, 주시, 운동 시차, 초점의 네 가지로 분류된다. 잘 알려진 안경식 3DTV의 경우에는 양안 시차와 주시 만을 제공하고 있어 입체감의 왜곡과 주시-초점 불일치로 인한 시각 피로가 발생하는 문제

※ 본 연구는 2018년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (2017-0-00072, 초실감 테라 미디어를 위한 AV부호화 및 LF미디어 원천기술 개발)

가 있다. 이에 비하여 네 가지 입체 요인을 모두 제 공함으로써 입체감의 왜곡이나 시각 피로가 없는 완전 입체 영상을 제공할 수 있는 기술로는 LF (Light Field) 기술이 있다. 현재까지 네 가지 입체 요인을 완전히 실현할 수 있는 기술로서 홀로그램이 가장 유력하나, 홀로그램을 상용화하기 위해서는 해결해야 할 기술적인 과제가 많으므로, 상용화 관점에서 다소 유리한 초다시점, 집적영상을 이용한 LF 기술이 대안으로 여겨지고 있다.

LF 기술은 카메라를 통해 빛의 세기와 방향 정보를 획득한 후 디스플레이를 통해 그대로 재현함으로써 시각적으로 왜곡이 없는 완전 입체의 실감 영상을 제공한다. LF 기술에 대한 연구는 대형 디스플레이에서 안경 착용 없이 완전 입체 영상을 볼 수 있도록 하기 위한 목적으로 시작되었다. 그러나, 대형 디스플레이에서 완전 입체 영상을 제공하기 위해서는 수백 개에 이르는 고화질의 영상을 동시에 디스플레이 하는 것이 필요하기 때문에 디스플레이의 해상도, 영상 데이터 처리 속도 등의 문제를 해결하여야 한다. 이에 비하여, HMD(Head Mounted Display)와 같은 개인용 단말에서 LF 기반의 완전 입체 영상을 제공한다면, 동시에 디스플레이 해야 하는 영상의 수를 크게 줄일 수 있어 디스플레이의 해상도와 데이터 처리 속도를 빠른 시

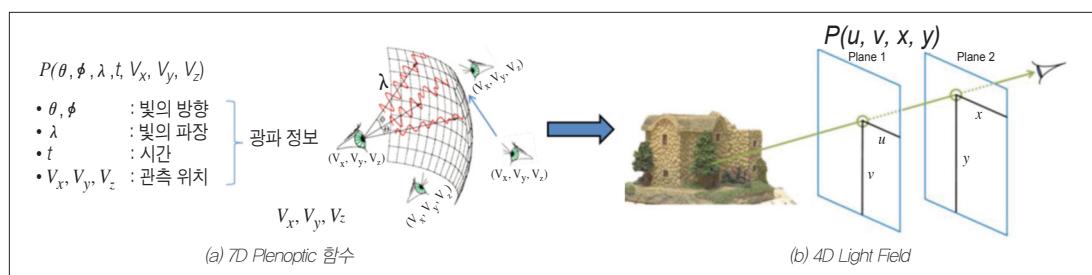
간 내에 구현 가능한 수준으로 줄일 수 있다. 따라서, 360도 전방위 영상에 LF 기술을 적용함으로써, HMD를 통하여 사실감과 몰입감이 크게 증가된 개인형 완전 입체 영상을 제공하기 위한 기술 개발이 활발히 진행 중이다. 또한, 이에 대한 기술 표준화를 위하여 MPEG에서는 MPEG-I 표준화를 시작하였다.

본 고에서는 먼저, LF 미디어 기술의 개념에 대하여 알아보고, LF영상의 획득, 처리 및 재생 기술에 대하여 알아본 뒤, LF 영상 기술이 360도 전방위 영상 분야에 어떻게 적용되고 있는지를 알아본다. 또한, LF 기술이 적용된 몰입형 미디어 기술 표준인 MPEG-I의 표준화 동향에 대하여 알아본다.

II. Light Field 미디어 기술 개발

1. Light Field 미디어의 개념

LF는 피사체로부터 반사되는 빛의 세기와 방향을 3차원 공간상에 표현하기 위한 장으로서, LF 미디어는 LF를 영상으로 획득한 후 디스플레이를 통해 재현함으로써 시각적으로 왜곡이 없는 완전한 몰입형 실감 영상을 시청 가능하게 하는 미디어로 정의한다.



〈그림 1〉 공간상의 빛을 표현하기 위한 LF 개념

LF는 Plenoptic 함수를 통해 표현되며, 고밀도의 광선을 획득 및 재현할 수 있는 광학 소자, 이미지 센서, 디스플레이 소자 등 영상시스템을 통해 구현되고 있다. LF는 <그림 1(a)>에서와 같이 3차원 공간상에서 피사체로부터 반사된 빛은 모든 파장(λ)에 대해 전방향으로(θ, ϕ) 관측자의 위치(V_x, V_y, V_z)에 입사되는 빛의 세기의 시간적인 변화(t)를 나타내는 7D Plenoptic 함수, $P(\theta, \phi, \lambda, t, V_x, V_y, V_z)$ 로 표현된다. 따라서, 7D Plenoptic 함수를 구현하게 되면 특정 3차원 공간에 대한 모든 LF 정보를 얻을 수 있다[1][2]. 그러나, 현실적으로는 <그림 1(b)>에서와 같이 $P(x, y, u, v)$ 로 표현되는 4D 함수를 통해 두 개의 2차원 면을 통과하는 광선들의 세기와 방향으로 LF 영상을 충분히 표현할 수 있기 때문에 실제 LF 영상시스템의 구현 및 신호 처리를 위해서는 4D LF의 두 면을 절단한 형태인 2D LF로 해석하는 것이 용이하다.

이상으로부터, LF 영상은 피사체로부터 반사된 모든 광선들이 여러 각도의 카메라에 투영된 영상으로서, 완전한 입체, 연속적인 시점의 변화 및 깊이에 따른 초점의 변화를 지원하기 위해 고밀도의 방향에서 획득된 시점영상으로 정의하기로 한다.

2. LF 영상 획득 기술

LF 영상을 획득하기 위해서는 고밀도의 공간 및 방향 성분을 가지는 광선의 획득이 필요하다. LF 영상 획득 방법에는 카메라를 수평 및 수직으로 배열하여 구성된 다중 카메라를 통해 획득하는 방법과 마이크로렌즈 어레이(Microlens Arrays) 기반의 카메라로 획득하는 방법이 있다.

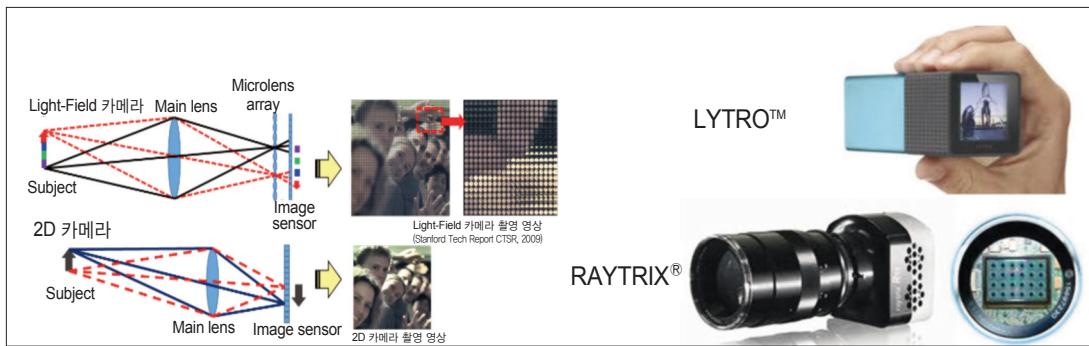
다중 카메라 기반의 LF 카메라는 최외각 시점 영상간의 베이스라인 폭을 넓게 할 수 있어 시야각이

넓고, 획득 가능한 깊이 범위가 깊으며, 시점 영상 당 해상도를 CCD의 해상도와 동일하게 유지할 수 있다는 장점이 있다. 하지만 <그림 2>에서 보는 바와 같이 많은 수의 카메라를 동시에 다루어야 하기 때문에 카메라 설치가 어렵고, 정렬 및 보정과 같은 촬영 영상에 대한 후처리가 어렵다는 단점이 있다[3].



<그림 2> 다중카메라 기반의 LF 카메라

반면, 마이크로렌즈 어레이 기반의 LF 카메라는 <그림 3>에서와 같이 대물렌즈와 마이크로렌즈 어레이로 구성되며, 대물렌즈에 의해 가상 이미지를 생성한 후 다시 마이크로렌즈 어레이에 의해 방향성 광선을 CCD 센서에 투영하는 방식으로 LF 영상을 획득한다. 따라서 단일 CCD 센서로 영상의 정보뿐만 아니라 방향성 정보를 동시에 획득할 수 있다는 장점이 있다[3]. 하지만, 획득 가능한 광선의 각도가 작고 컴퓨팅에 의해 생성된 시점 영상 중 최외각 간의 베이스라인 폭이 다중 카메라에 비해 좁아 획득 가능한 깊이 범위가 얕다는 것[4,5]과, 여러 시점 영상이 한정된 이미지 센서에 동시에 투영되므로 시점 수가 증가할수록, 즉 방향 해상도가 증가할수록 시점 당 공간 해상도가 감소한다는 단점이 있다. 따라서, 상용화를 위해서는 시점 영상의 해상도 및 획득 가능한 깊이 범위를 동시에 증대시키기



<그림 3> 마이크로렌즈 어레이 기반 LF 카메라의 구조 및 상용품

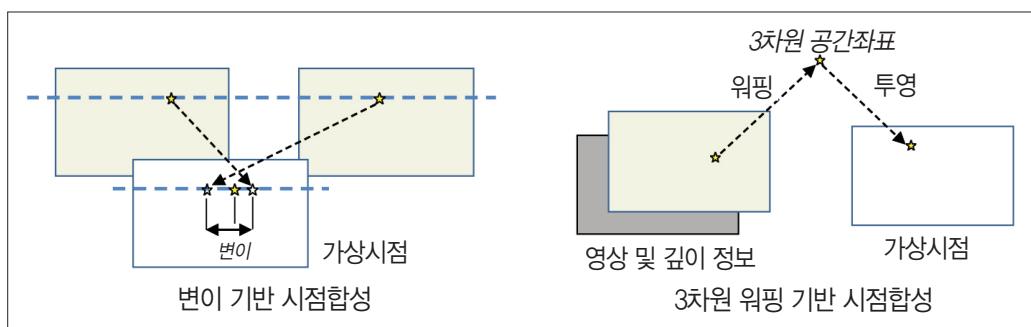
위한 기술 개발이 필요하다.

3. LF 영상 합성 및 가변 초점 재현 기술

LF 카메라로부터 촬영되거나 CG로부터 제작된 LF영상의 공통 포맷은 마이크로 영상, Full-parallax 시점영상 및 깊이 영상이 될 수 있으며, LF 디스플레이에 재현되어 시청자에게 완전 입체 영상을 보여줄 수 있다. LF 디스플레이에서 완전 입체 영상을 재현하기 위해서는 양안 시차와 더불어 자연스러운 시점의 변화 및 시청자가 주시하는 지점과 초점이 일치하도록 하여야 하며, 이를 위해서는 시청자의 동공 내에 여러 개의 시점 영상이 투영 가

능하도록 고밀도의 시점 영상을 재생하여야 한다. LF 카메라의 획득 가능한 시점 영상의 수와 대용량의 영상 데이터 량을 고려할 때, 고밀도의 시점 영상을 카메라로부터 직접 획득하는 많은 어려움이 따른다. 따라서, 저밀도로 획득한 LF 영상으로부터 가상 시점 영상 합성을 통하여 고밀도의 LF 영상을 생성하여 재현하는 기술이 필수적이라 할 수 있다.

가상 시점 영상 합성 방법에는 <그림 4>에서 보는 바와 같이 스테레오 영상에서 사용되는 변이 기반과 3차원 워핑(warping) 기반 가상 시점 합성 방법이 있다[6]. 변이 기반 합성은 깊이 추정 과정에서 얻어진 각 픽셀의 변이 값을 기준으로 두 영상 사이의 새로운 시점 영상의 픽셀 값을 계산하여 합성하는 방

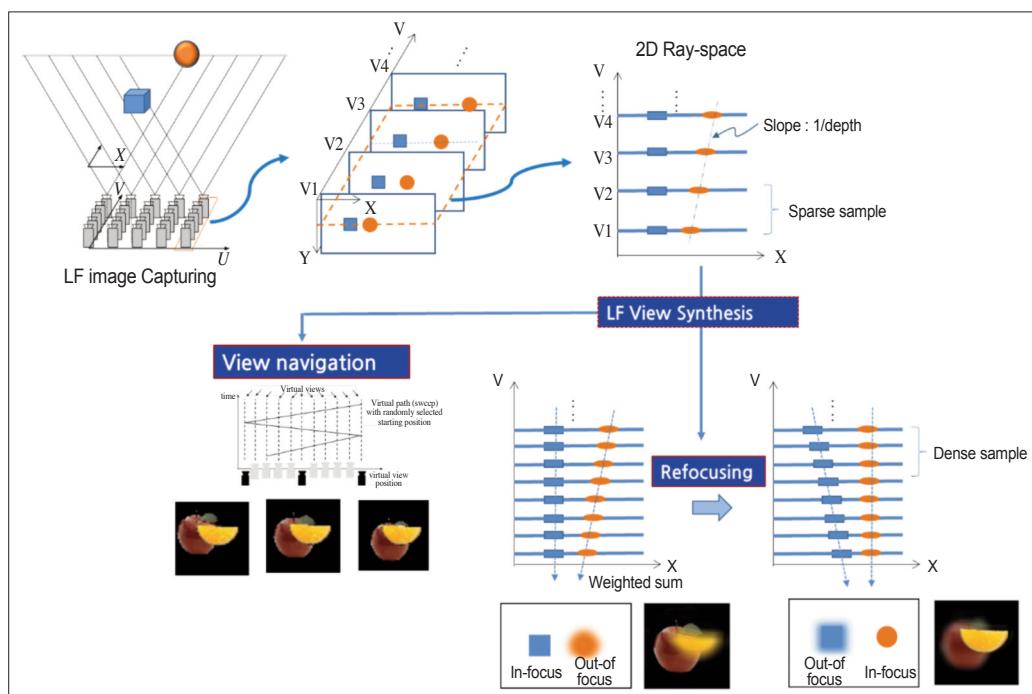


<그림 4> 변이 및 3차원 워핑 기반 시점 합성 방법 비교

법이다. 두 영상 간 변이 방향 사이의 제한적인 시점 영상만 합성 가능하나, 구현이 간단하고 처리 복잡도가 낮은 장점이 있다. 반면 3차원 워핑 기반 합성은 변이 값과 카메라 매개변수로부터 추정된 깊이 정보를 이용하여 각 픽셀을 3차원 점으로 워핑하고, 이를 새로운 시점의 영상으로 재투영하는 합성 방법이다. 워핑 및 재투영의 복잡한 처리 과정이 필요하나, 임의 위치의 가상 시점 합성이 가능하다. 기존 다시점 영상에서의 가상 시점 합성은 좌우 시차를 지원하기 위해 수평방향으로의 시점 합성을 고려하였다면, LF영상에서는 상하, 좌우, 대각 방향의 시차를 지원하여야 하고, 전방위 영상으로 확장했을 때는 전후 방향으로의 시차까지 지원할 수 있도록 가상 시점 영상이 생성되어야 한다.

영상 합성 기술과 더불어 LF영상기반 실감미디

어서비스를 가능케 하는 핵심 기술 중 하나로 가변 초점 재현 기술을 들 수 있다. 가변 초점 재현 기술은 시청자가 주시하고 있는 객체에 초점을 맞춤으로써 주시-초점 불일치로 인한 시각 피로를 감소시킬 수 있는 기술로서, 영상 합성 기술과는 <그림 5>에서와 같은 상호 연관성을 가진다. 이 그림은 에피폴라 이미지(Epipolar Image) 개념으로 LF 영상 합성과 가변 초점 기술 개념을 설명하고 있다. <그림 5>에서 어레이 방식의 LF 카메라를 통해 획득된 LF 영상 중에서 수평 또는 수직 방향의 시점 영상을 4D LF로 표현하면 수평축은 공간 성분, 수직축은 방향 성분에 해당하는 시점의 변화로 표현된다. 이 시점 영상들에서 각 수평 라인을 분리하여 재배열하면 2D ray space상의 광선 샘플로 표현되며, 이를 에피폴라 이미지라 한다. <그림 5>에서 3D 공간상 객



<그림 5> LF영상 합성 및 가변 초점 기술 개념

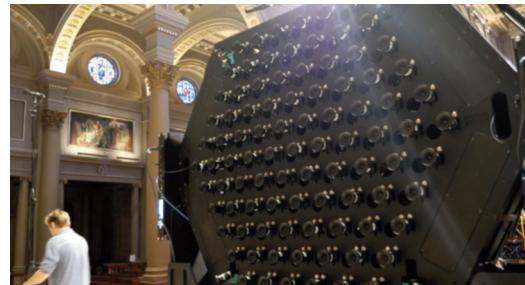
체들을 시점 방향(v)의 움직임으로 표현하면, 기울기는 해당 객체의 깊이를 나타내며, 저밀도로 획득된 시점 영상의 경우에는 수직축(v) 방향으로 저주파의 광선 샘플 분포를 나타내게 된다. 저주파의 광선 샘플 분포는 시점의 변화 혹은 초점의 변화와 같은 LF영상 재현 시, 광선 신호처리 관점에서는 엘리어싱, 영상 처리 관점에서는 불연속 시점의 변화 및 블록킹이 있는 가변 초점 영상 재생으로 나타난다. 이러한 문제를 해결하기 위해서는 고주파의 방향성 광선 즉, 고밀도의 시점 영상을 필요로 한다. 이는 Nyquist 이론을 만족할 정도로 공간상에서 조밀하게 샘플링 된 광선이 필요하다는 것을 의미하며, 시청자 관점에서 보면 동공 크기가 샘플링 주기라고 가정할 때 이 보다 더 높은 주파수의 시점 영상이 필요하다는 것을 의미한다. LF 영상 합성에 의해 고밀도로 합성된 수평 및 수직 방향의 시점이 재현된다면, <그림 5>와 같이 사용자 선택에 의한 연속적인 시점의 변화(View Navigation)가 가능하며, 엘리어싱 현상이 없는, 즉 블록킹 현상이 없는 초점 변화가 가능하게 된다.

4. LF 기반 360 전방위 영상 획득 및 재현 기술

HMD를 이용한 가상현실이 차세대 ICT 플랫폼 기술로 전세계적으로 주목을 받으며 급부상함에 따라, 다중 카메라를 전방향으로 배치하여 공간정보를 획득하고, 사용자의 6자유도의 움직임을 지원하여, 시점 뷰(view) 변화에 따라 운동시 차와 양안 시차가 제공되는 영상을 재현하는 기술이 개발되고 있다. LF 기반 360 전방위 영상 획득 및 재현 기술은 가상 현실의 몰입감을 극대화하기 위한 필수 기술로, 이전의 LF 영상은 4D LF를 기반으로 한정된 각도의 광

선 만을 다루었다면, 몰입형 미디어에 적용되는 LF는 360도 전방위의 광선을 다루는 기술이라 할 수 있다. 이에 대한 연구개발이 연구소, 대학, 기업 등 여러 기관에서 진행되고 있으며, 주요 ICT 기업들도 LF 기반의 360 전방위 영상 획득 및 재현 기술 개발을 추진하고 있다. 본 절에서는 대표적으로 Lytro와 Facebook에서 공개한 기술을 살펴본다[7,8].

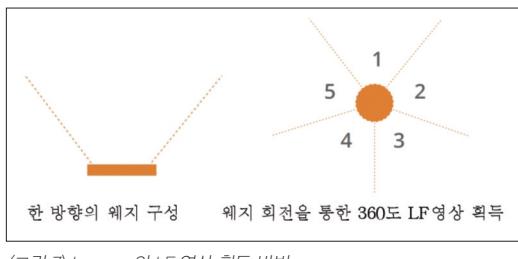
Lytro는 초기에 설계하였던 구 형태의 카메라 리그 구조를 수정하여, 육각형 모양의 평면에 정면 방향 11개 층으로 평행하게 배치된 91대 카메라와 좌우측 상하단 촬영을 위하여 배치된 4대 카메라를 포함한 총 95대의 다중 카메라로 구성된 새로운 Immerge 카메라를 공개하였다.



〈그림 6〉 Lytro사의 Immerge 카메라

Immerge 카메라는 120도 시야각의 웨지(wedge) 형태의 다중 카메라 어레이 평면을, 72도 간격으로 회전하면서 360도 전방위의 LF 영상을 획득한다. 즉, 360도 전방위 영상을 얻기 위해서는 카메라 회전에 따른 5번의 촬영이 필요하다. 따라서, 전방위를 동시에 촬영하지 않으므로, 동기화된 전방위 영상획득이나, 72도의 촬영 영역 경계를 넘어서는 움직임을 가지는 객체에 대한 촬영이 불가능하다.

Immerge를 통하여 획득한 LF 영상의 처리 및 렌더링 방법에 관해서는 상세 내용이 공개되지 않



〈그림 7〉 Immerge의 LF 영상 획득 방법

았으나, 일부 알려진 단편적인 정보들을 종합적으로 분석해 보면, 각 방향에서 촬영된 다수 카메라의 시점 영상을 이용하여 깊이 정보 및 텍스쳐 정보를 추출하고, 자체의 LF 데이터 포맷으로 변환한 후, 이를 3차원 공간에 구성하여 직경 1미터의 뷰잉 볼륨(Viewing Volume) 내에서 사용자의 6DoF 움직임에 따른 시점 영상이 재현되는 것으로 파악된다.

Facebook은 2017년도 자사의 개발자 컨퍼런스 F8에서 차기 버전의 Surround 360 카메라인 x24와 x6를 공개하였으며, Adobe, Otoy, Foundry, Mettle 등 주요 포스트 프로덕션(Post Production) 업체들과 협력 관계를 맺고 6DoF 가상현실 콘텐츠 제작에 필요한 획득, 압축, 전송 및 재현을 위한 툴체인(Toolchain) 개발을 주도하고 있다. x24와 x6 카메라는 각각 24대와 6대의 카메라를 구 중심에서 전방향으로 배치한 구조로, 픽셀 단위로 RGB 영상 및 깊이 정보 획득이 가능하여, 이를 활용한 3차원

모델 기반의 6DoF 서비스를 지원한다고 한다. 특히 장면 구성 및 렌더링 툴 간에 전달되는 장면과 관련된 모든 데이터 입출력을 지원하는 패키지 기술인 Otoy의 ORBX는 MPEG-I 표준화 그룹에서 6DoF 지원을 위한 컨테이너 포맷(Container Format)으로 제안되어 논의되고 있다.



〈그림 9〉 Facebook의 차기 Surround 360 카메라로 공개된 x24와 x6

국내에서는 한국전자통신연구원에서 2017년도부터 추진 중인 “초실감 테라미디어를 위한 AV 부호화 및 LF 미디어 원천기술 개발” 사업 내에서 LF 기반 전방위 입체 영상 획득 및 재현에 필요한 핵심 원천기술 개발이 진행되고 있다. 이를 통해 HMD를 착용한 사용자의 6DoF 움직임에 따라 최대 4K의 초고해상도 시점 영상을 제공하여 실사 VR 콘텐츠의 공간감과 몰입감을 극대화시키는 것이 기술



〈그림 8〉 사용자의 6DoF 움직임에 따른 시점 영상 재현 개념

개발의 최종 목표이다. LF 기술은 가능한 많은 시점의 빛의 세기를 기록했다가 사람이 보는 방향으로 여러 시점의 빛의 세기를 동시에 보여줌으로써 자연스러운 운동 시차와 완전 입체 정보를 제공하기 위한 기술이지만, 현실적으로 많은 시점의 영상 획득도 어렵고, 이를 재현하기 위한 데이터 처리 및 디스플레이 제작도 어렵다는 문제가 있다. 하지만 LF 기술을 HMD에 접목하면 한 사람의 시점만 고려하여 처리하고 재생하면 되기 때문에, 제한된 시점의 영상을 기반으로 사용자의 움직임에 따른 가상의 시점 영상을 동적으로 빠르게 합성하여 완전 입체 영상을 제공하는 것이 이 기술의 주요 특징이다. 또한 HMD를 이용할 때 발생하는 초점 불일치로 인한 시각 피로를 감소하기 위한 기술도 함께 연구되고 있어 실사 VR 콘텐츠를 시청할 때 느껴지는 시각 정보 불일치에 따른 이질감을 줄이는 효과를 얻을 수 있을 것으로 기대된다.

III. Light Field 미디어 표준화

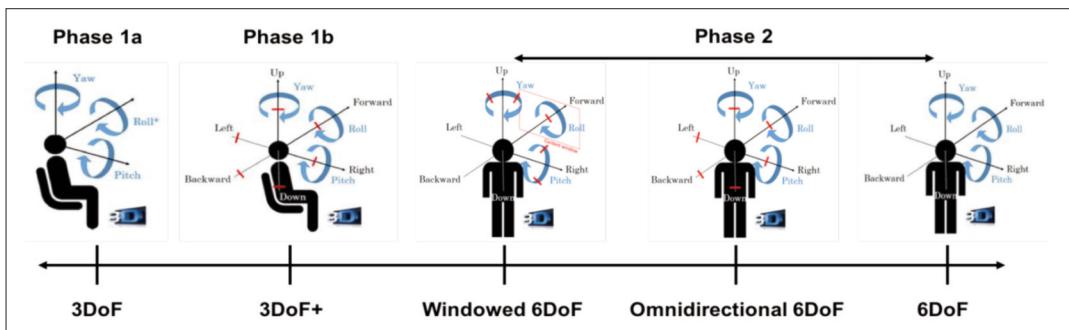
MPEG은 2016년 1월부터 360VR을 필두로

Light Field, Free Viewpoint Television 및 Point cloud 등을 포함하는 몰입형 미디어 표준화를 MPEG-I라는 프로젝트 명으로 수행 중에 있다. MPEG-I에서는 몰입형 미디어를 3DoF, 3DoF+ 및 6DoF의 세 가지로 분류하고, 2021까지 순차적으로 표준화한다는 계획을 가지고 있다.

1. MPEG-I 표준화 개요

MPEG-I 프로젝트는 몰입형 및 전방위 비디오를 위한 구조, 포맷, 부호화 및 포인트 클라우드 등의 기술에 대한 표준을 제정하기 위한 것으로, 〈그림 10〉에서와 같이 phase 1a, phase 1b 및 phase 2의 총 3단계에 걸친 표준화 로드맵을 제시하고 있다. Phase 1a는 고정된 중심을 기준으로 회전운동만을 지원(3DoF)하고, Phase 1b는 3DoF에 계층 기반의 깊이 감을 제공함으로써 제한된 병진 운동(3DoF+)을 추가적으로 제공하며, phase 2는 사용자 움직임에 따라 회전 및 병진 운동을 모두 지원(6DoF)함으로써 몰입감을 극대화하기 위한 것이다.

MPEG-I 프로젝트는 총 8개의 파트로 나뉘어 표준화 작업이 진행 중에 있다[9,10]. 각각의 파트를



〈그림 10〉 MPEG-I 표준화 계획

간단히 설명하면 다음과 같다.

- Part 1: Technical Report on Immersive Media

실제적인 기술적 표준화가 이루어지는 부분이 아닌 MPEG-I의 표준화 방향성과 진행사항을 전체적으로 정리하는 파트이며, 표준화 로드맵, 사용자 요구사항, 유스케이스(Usecase) 및 단계별 구조 등을 포함

- Part 2: Omnidirectional MediA Format (OMAF)

단일 전방위 비디오 및 오디오 데이터로 구성된 몰입형 미디어의 효율적인 저장, 전송, 재현, 교환을 위한 파일 포맷 표준을 개발하는 것으로, 현재 3DoF 지원을 목표로 하는 OMAF 버전 1의 표준화가 완료 단계에 있으며, 초고해상도, 계층기반 전방위 비디오 및 깊이 감 제공을 위한 3DoF+에 대한 요구사항 정의 중

- Part 3: Immersive Video

3DoF 지원을 위한 단일 전방위 비디오 부호화 및 LF, FTV 미디어 등을 포함하는 6DoF에 대한 부호화 표준 기술 개발

- Part 4: Immersive Audio

기 제정된 표준들이 이미 3DoF 및 3DoF+ 오디오를 지원하고 있는 것으로 판단하고 있으며, 6DoF 지원 몰입형 오디오의 부호화 및 재생 관련 표준 기술 개발

- Part 5: Point Cloud Compression

LF 미디어와 함께 시청자가 자유롭게 움직일 수 있는 환경에서 시점의 제한이 없는 6DoF 지원 몰입형 미디어 구현에 핵심 기술로 고려되고 있는 포인트 클라우드 데이터의 부호화 표준 기술 개발

- Part 6: Immersive Media Metrics

몰입형 미디어 품질평가를 위한 객관적인 데이터 포맷 개발

- Part 7: Immersive Media Metadata

HMD 등의 단말에서 복호화된 몰입형 미디어 데이터를 재현하기 위하여 요구되는 메타데이터 표준 개발

- Part 8: Interfaces for Network Media Processing

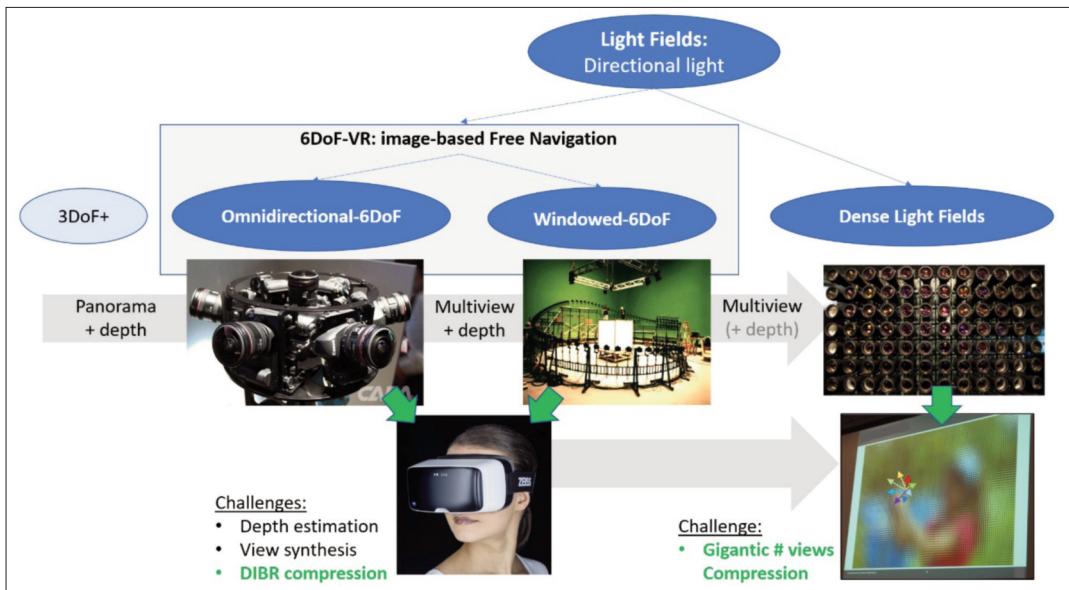
높은 컴퓨팅 파워가 요구되는 대용량 미디어 데이터의 재현 및 처리 과정을 클라우드 서버에 나누어 처리할 수 있도록 네트워크 연동을 위한 인터페이스 표준 개발

지금까지 MPEG-I phase 1a 단계와 관련한 논의는 대부분 Part 2인 OMAF을 중심으로 진행되었으며, 2017년 10월 MPEG 미팅에서 FDIS(Final Draft International Standard) 문서를 발행하였다 [11]. Phase 1a 표준화가 마무리 된 현재 MPEG-I의 관심사는 Phase 2 단계의 6DoF 서비스에 관한 것으로, 이를 표현-압축-전송-재현까지의 전반적인 상황에 대하여 시스템 및 비디오 서브그룹을 중심으로 표준화가 진행 중에 있다.

2. MPEG-I Phase 2

MPEG-I Phase 2의 목표는 시청자가 자유롭게 움직일 수 있는 환경에서 6DoF 지원 몰입형 미디어를 제공하는 것으로, Phase 2의 가장 중요한 개발 요소는 6DoF 지원 몰입형 미디어를 전송 및 재현하기 위한 부호화, 컨테이너 포맷 및 대용량 데이터를 처리하기 위한 원격 처리기술이다.

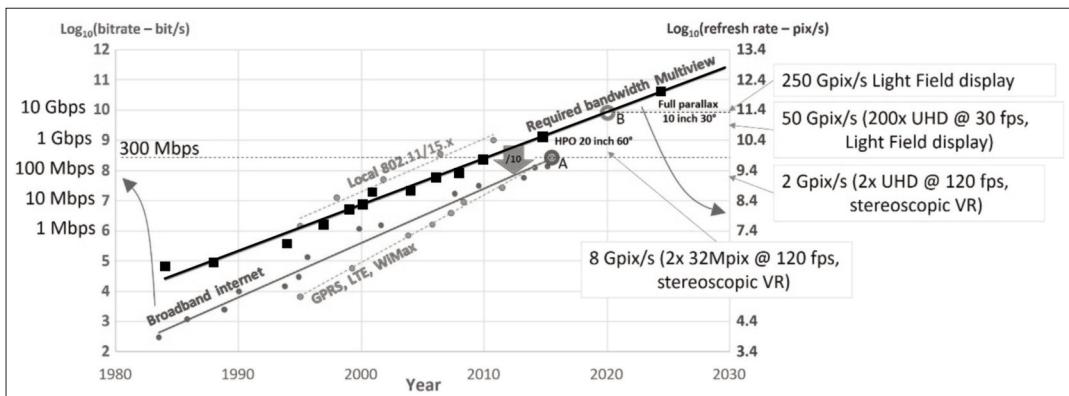
MPEG-I Phase 2를 위한 비디오의 경우, 6DoF 와 Light Field가 주요 키워드라 할 수 있으며, LF



〈그림 11〉 6DoF와 Light Field에 대한 서비스 시나리오 및 발전 전망[12]

에 기반한 6DoF 지원 몰입형 미디어의 발전 방향 및 서비스 시나리오는 〈그림 11〉에서와 같다[12]. 6DoF를 지원하기 위해서는 시청자의 움직임에 따라, 움직임에 맞는 영상을 제공하는 것이 필수적이며, 이를 위해서는 수 많은 시점 영상이 필요하다. Omnidirectional 6DoF와 Windowed 6DoF는 다수

의 컬러 영상과 깊이 정보를 이용하여 시청자의 시점에 맞는 가상 시점 영상을 생성하여 제공하는 것으로, 영상 부호화 관점에서는 컬러 영상과 깊이 영상의 동시 부호화와 가상 시점으로부터의 예측 부호화 등이 중요한 요소라 할 수 있다. 또한, Dense LF는 많은 수의 조밀하게 배열된 카메라로부터 얻



〈그림 12〉 Light Field 디스플레이의 refresh rate과 대역폭과의 관계[12]

은 컬러 영상을 이용하여 HMD나 안경 착용 없이 주시-조점이 일치된 시각 피로 없는 완전 입체 영상을 제공하는 것으로, 카메라로부터 얻어지는 영상의 수가 LF 재현에 충분한 경우에는 가상 시점 합성이 필요 없으므로 깊이 정보 역시 필요 없지만, 부호화해야 하는 영상의 수가 크게 증가한다.

MPEG-I Phase2의 비디오 분야의 목표는 앞에서 살펴본 LF 기반의 6DoF 지원 영상을 효율적으로 부호화하기 위한 표준 개발이라 할 수 있다. 이 경우, 앞에서 살펴본 바와 같이 부호화해야 할 영상의 수가 기준에 비하여 월등히 많기 때문에 압축율 뿐만 아니라 부호화된 영상의 총 비트율이 중요하다 할 수 있다.

〈그림 12〉는 연도에 따른 LF 디스플레이에서 필요로 하는 refresh rate (오른쪽 y축)과 대역폭(왼쪽 y축)과의 관계를 나타낸다. 이 그림에서 검정선은 디스플레이가 발전됨에 따라 정상적인 디스플레이를 위해 필요한 대역폭을 의미하며, 회색선은 브로드밴드 망의 가능 대역폭을 나타낸다. 여기서, 필요 대역폭 계산은 현 시점에서의 최신 코덱인 HEVC를 이용하여 압축부호화 하는 것을 가정하였다. 이 그림에서 필요 대역폭과 가능 대역폭 간의 약 10배 정도의 차이가 유지됨을 확인할 수 있으며, 이는 LF 영상 압축을 위해서는 HEVC의 성능이 10배 향상되어야 함을 의미한다. 따라서, 단순히 HEVC의 확장으로는 목표 성능을 확보하기 힘들 가능성이 크며, 새로운 접근의 부호화 기술이 필요한 시점이라 할 수 있다. 현 시점에서의 MPEG-I Phase2를 위한 부호화 기술에 대한 표준화 활동은 요구사항을 정의하고, 테스트 영상을 모집하고 있는 표준화 초기 단계에 있다.

Phase 2 시스템 표준화와 관련해서, 6DoF지원 몰입형 미디어의 대용량데이터의 처리, 전송 및 재

현을 위하여 HNSS(Hybrid Natural/Synthetic Scene Data Container) 포맷과 NBMP(Network Based Media Processing) 기술에 대한 표준화가 진행 중에 있다. HNSS는 포인트 클라우드, LF기반의 실사 및 CGI기술을 통해 가상으로 생성한 데이터 등 다양한 포맷의 미디어 데이터를 저장, 전송 및 재현하기 위한 컨테이너 포맷으로 현재 요구사항[13] 및 CfP(Call for Proposal) 문서[14]가 발생되었다. NBMP는 고성능의 컴퓨팅 파워가 요구되는 대용량 미디어의 전처리 및 재현 과정을 단말이 아닌 네트워크 기반의 원격으로 처리하는 기술로 2017년 7월 NBMP AhG이 발족되었으며 현재 요구사항[15] 및 CfP 문서[16]가 발행되었다.

IV. 결 론

본 고에서는 사용자에게 향상된 몰입감과 자연스러운 움직임 자유도를 제공하기 위한 LF 미디어 기술 및 관련 MPEG 표준화 동향에 대하여 살펴보았다. LF 미디어는 카메라를 통해 빛의 세기와 방향 정보를 획득한 후 디스플레이를 통해 그대로 재현함으로써, 시각적으로 왜곡 없이 현실 세계를 가장 근접하게 재현할 수 있는 새로운 미디어 형태 중 하나로 인식되고 있다. 더욱이 VR미디어와 같은 몰입형 미디어의 성장이 예상됨에 따라 LF미디어 기술과 결합된 6DoF지원 몰입형 미디어를 제공하기 위한 노력이 가속화되고 있으며, MPEG는 이를 서비스하기 위한 요소기술 표준화를 추진 중에 있다. 이러한 변화 흐름 속에 본 고가 LF 미디어 기술의 연구 개발 방향을 예측하고 몰입형 미디어 시장에서의 기술 경쟁력을 높이는데 조금이나마 도움이 되기를 기대한다.

참고 문헌

- [1] M. Levoy and P. Hanrahan, "Light Field Rendering," Proc. ACM SIGGRAPH '96, pp. 31-42, 1996.
- [2] S.J. Gortler, R. Grzeszczuk, R. Szeliski, and M.F. Cohen, "The Lumigraph," Proc. 23rd Ann. Conf. Computer Graphics and Interactive Techniques (SIGGRAPH '96), pp. 43-54, 1996.
- [3] Y. Taguchi, T. Koike, K. Takahashi, and T. Naemura, "TransCAIP: A Live 3D TV System Using a Camera Array and an Integral Photography Display with Interactive Control of Viewing Parameters" IEEE Trans. on Visualization and Computer Graphics, vol. 15, no. 5, 2009.
- [4] <http://www.raytrix.de/>
- [5] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, P. Hanrahan, "Light field photography with a hand-held plenoptic camera," Stanford Tech Report CTSR 2005-02
- [6] 호요성 외 3인, "실감형 다시점 3차원 영상의 획득 및 처리 기술," 진샘미디어, 2016.
- [7] <http://www.lytro.com/immerse>
- [8] <http://developers.facebook.com/videos/f8-2017/surround-360-beyond-stereo-360-cameras/>
- [9] ISO/IEC JTC1/SC29/WG11, "Working Draft 1.0 of TR: Technical Report on Architectures for Immersive Media," N17060, July 2017.
- [10] 호요성, "MPEG-I 표준과 360도 비디오 콘텐츠 생성," 전자공학회지, 44권 8호, pp.52-57, 2017.
- [11] ISO/IEC JTC1/SC29/WG11, "Text of ISO/IEC FDIS 23090-2 Omnidirectional Media Format," N17235, October 2017.
- [12] ISO/IEC JTC1/SC29/WG11, "MPEG-I Visual activities on 6DoF and Light Fields," N17285, October 2017.
- [13] ISO/IEC JTC1/SC29/WG11, "Requirements for MPEG-I hybrid natural/synthetic scene data container(V.1)," N17064, July 2017.
- [14] ISO/IEC JTC1/SC29/WG11, "Draft Call for Proposals for MPEG-I hybrid natural/synthetic scene data container(V.1)," N17065, July 2017.
- [15] ISO/IEC JTC1/SC29/WG11, "Requirements for network based media processing(v1)," N17062, July 2017.
- [16] ISO/IEC JTC1/SC29/WG11, "Draft Call for Proposals on Network-Based Media Processing," N17263, October 2017.

필자 소개

정원식



- 1992년 : 경북대학교 공과대학 전자공학과 졸업 (공학사)
- 1994년 : 경북대학교 대학원 전자공학과 졸업 (공학석사)
- 2000년 : 경북대학교 대학원 전자공학과 졸업 (공학박사)
- 주관심분야 : 디지털방송, MPEG, VR/AR/MR

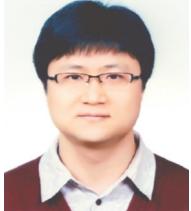
이광순



- 1993년 : 경북대학교 공과대학 전자공학과 졸업 (공학사)
- 1995년 : 경북대학교 대학원 전자공학과 졸업 (공학석사)
- 2004년 : 경북대학교 대학원 전자공학과 졸업 (공학박사)
- 주관심분야 : LF 비디오 획득/재현, VR/AR

필자소개

윤정일



- 1996년 : 전북대학교 제어계측공학과 학사
- 1998년 : 광주과학기술원 기전공학과 석사
- 2005년 : 광주과학기술원 기전공학과 박사
- 2005년 ~ 현재 : 한국전자통신연구원 테라미디어연구그룹 책임연구원
- 주관심분야 : 디지털 방송 시스템, 방통융합 멀티미디어 기술

윤국진



- 1999년 : 전북대학교 컴퓨터공학과 학사
- 2001년 : 전북대학교 컴퓨터공학과 석사
- 2016년 : 경희대학교 전자전파공학과 박사
- 2001년 ~ 현재 : 한국전자통신연구원 테라미디어연구그룹 책임연구원
- 주관심분야 : VR/AR, MPEG, 디지털방송 기반기술

서정일



- 1994년 : 경북대학교 전자공학과 학사
- 1996년 : 경북대학교 전자공학과 석사
- 2005년 : 경북대학교 전자공학과 박사
- 2000년 ~ 현재 : 한국전자통신연구원 테라미디어연구그룹 그룹장
- 주관심분야 : 실감미디어, UMW, 360 Video, 공간미디어, MPEG