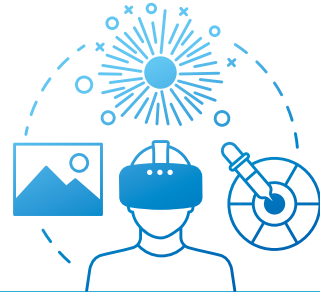


JPEG 플레노(pleno) 및 라이트 필드 코딩 소개

김영섭 _ 단국대학교 전자전기공학부 교수
박규섭 _ 단국대학교 전자전기공학부 연구원



1. 머리말

2차원 카메라 영상정보에 의존하는 기존의 영상처리 기술에 비해, 라이트 필드(light field) 영상은 공간상에서 임의의 방향으로 진행되는 빛의 정보를 제공한다. 즉, 2차원 영상을 구성하는 각 픽셀에 대해 샘플링된 방향에 대한 빛의 세기와 컬러 정보를 제공하기에 실제 공간에서 물체들을 사실적으로 묘사하는 데 아주 적합하다.

라이트 필드 기술은 광선의 분포를 재현하는 기술이다. 그런데 광선의 분포를 재현한다는 것은 일반적으로 생각하는 사진(정지 영상) 뿐만 아니라 실상 우리가 감상하는 동영상들 또한 실제로 초당 30번 이상 바뀌는 정지영상의 집합이기 때문에, 이러한 사진 기술은 현대 영상 기술의 근간이라고 표현할 수 있다.

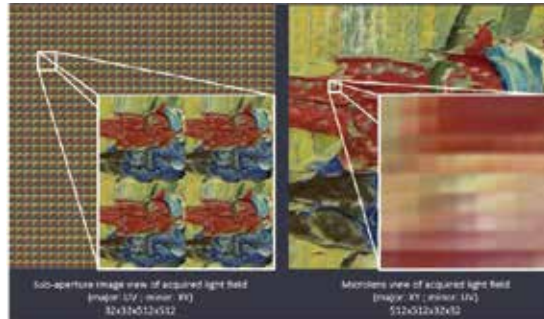
현재 사용되는 일반적인 이미징 기술보다 많은 부가가치를 창출할 수 있는 라이트 필드 기술을 활용하면 더 많은 빛 정보를 얻을 수 있으므로 텍스트 및 기하학적 정보에서 풍부한 시각적 장면 구조를 캡처할 수 있다. 더 많은 정보를 이용할 수 있게 된다면, 디지털 사진이나 비디오 제작 후에 현재 '고전'영상

방식을 활용하고 있는 사례에 대해 더 나은 성과를 달성하고 산업검사와 같은 새로운 사용사례를 다루는 것이 가능할 수 있다.

라이트 필드 데이터(광학 데이터)는 [그림 1]에서와 같이 공간의 모든 지점에서 모든 방향으로 빛의 양(방사성)을 기록한다. 이 광도는 일련의 카메라(광범위 기준 광원 데이터 결과) 또는 마이크로 렌즈를 사용하여 최종 영상에 기여하는 각 개별 광선을 샘플링하는 광원 카메라에 의해 근사화되고 캡처될 수 있다(좁은 기준 광원 데이터 결과). 또한, 광 필드 데이터는 컴퓨터 모델에서 동기적으로 생성될 수 있으며 캡처한 이미지와 포인트 클라우드 그리고 논리적으로 렌더링된 라이트 필드의 조합은 함께 융합될 수 있다.

2. JPEG의 라이트 필드 코딩

JPEG의 라이트 필드 코딩은 시스템 수준에서 고급 기능을 제공하면서 압축률을 높이기 위한 새로운 도구를 정의하는 것을 목표로 한다. 또한, 데이터 및 메타데이터 조작, 편집, 랜덤 액세스 및 상호작용, 개인정보보호, 소유권 및 기타 보안 메커니즘을 지



[그림 1] 라이트 필드 또는 플렌옵틱 이미지 및 마이크로 영상

원하는 것을 목표로 한다.

JPEG의 라이트 필드 코딩은 데이터 속도, 계산 복잡성 및 전력 소비 측면에서 합리적인 리소스 요구 사항으로 최상의 품질 콘텐츠 표현을 보장하는 효율적인 압축 형식을 제공한다. 다음에 요구하는 사항 외에도 지원되는 기능에는 짧은 대기 시간, 레거시 JPEG 형식과의 호환성, 확장성, 랜덤 액세스, 오류 복구, 개인 정보보호, 소유권 및 보안, 병렬 및 분산 처리, 계층적 데이터 처리 및 디스플레이 또는 디스플레이 요소 간의 데이터 공유가 포함된다.

* 관련 파일 형식은 JPEG 시스템 사양을 준수하며 데이터의 캡처, 생성, 고정, 처리, 렌더링 및 편집 및 이러한 데이터와의 사용자 상호작용을 위한 관련 메타데이터에 대한 시그널링 구문을 포함할 것이다.

2.1 라이트 필드 코딩의 요구 사항

라이트 필드 코딩에서 요구되어야 할 것으로는 다음 표준 항목들에 대하여 하나 이상 충족되어야 한다.

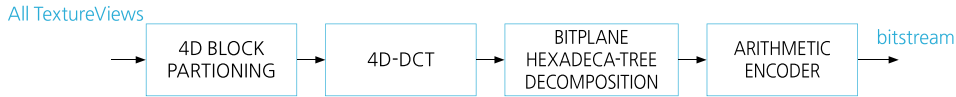
- 렌즈 이미지 시스템에 의해 생산된 LF 콘텐츠의 부호화 기술
- 고밀도 카메라 어레이 시스템에서 생산되는 LF 콘텐츠 코드화 기술
- 표준 영향을 미치는 LF 코딩 및 처리 기술과 관련된 시스템 레벨 솔루션

렌즈 카메라에는 좁은 기준 광장을 기록할 수 있는 마이크로 렌즈 광학 어레이가 내장되어 있다. 고밀도 카메라 어레이는 다른 공간 위치에서 장면을 그려내는 일반 카메라를 활용하는 로봇 작동 카메라 조작 시스템 또는 물리적으로 구별되는 여러 개의 카메라의 밀집 그리드를 통해 구현될 수 있다.

2.2 라이트 필드 코딩에서의 4D 변환모드

4D 변환 모드에서 라이트 필드는 [그림 2]와 같이 4단계 프로세스로 인코딩된다. 먼저, 4D 광장 데이터는 사전 정의된 고정 스케닝 순서에 따라 독립적으로 인코딩된 고정 크기의 4D 블록으로 분할된다. 이들 블록은 겹치지 않는 4D 서브 블록 세트로 더 분할 될 수 있으며, 최적 분할 파라미터는 R-D(rate-distortion) 기준에 기초하여 도출된다. 각각의 서브 블록은 가변 블록 크기 4D DCT 변환에 의해 독립적으로 변환된다. 후속적으로 변환된 블록은 헤사데카-트리 비트 평면 분해 및 적응형 산술 인코딩을 사용하여 양자화 및 엔트로피 코딩되어, 라이트 필드의 압축된 표현을 생성한다. 이 코딩 절차는 각 색상 성분에 독립적으로 적용된다.

효율적인 인코딩을 위해 4D 블록 파티셔닝과 변환 계수의 비트 평면 클러스터링은 트리 구조를 사용하



[그림 2] 4D 변환모드 인코딩 아키텍처

여 시그널링 된다. 서브-블록에서 4D 블록의 분할은 3진 플래그를 사용하여 이진트리로 시그널링 된다.

2.3 기능 및 잠재적 응용

라이트 필드 처리는 실제 세계와 합성 세계를 완벽하고 구별할 수 없는 방식으로 표현하기 위한 새로운 접근 방식의 가능성을 제시한다. 이를 통해 시각적 콘텐츠를 생성, 조작 및 상호작용할 수 있는 풍부한 사용자 환경이 제공되며, 이러한 새로운 경험은 결국 무한한 가능성을 제시할 것이다.

기존 이미지 코딩 형식으로는 구현하기 어려운 수 많은 새로운 응용 프로그램인 확장성, 랜덤 액세스, 오류 복원력 및 높은 압축 효율 등 기존 이미지 코딩 형식에서 이미 제공한 기능들 외에도 새로운 기능들을 제공함으로써 이를 용이하게 한다.

제공되는 새로운 기능들 중 일부 라이트 필드 표현은 다음과 같이 표현된다.

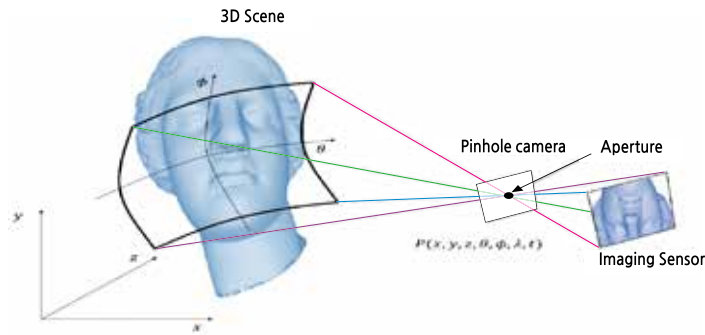
- 캡처 후 필드 깊이 변경
- 캡처 후 관심 대상에 초점을 변경하고 재집중
- 이미 촬영되거나 합성된 이미지의 조명변경
- 무료 내비게이션을 포함한 장면 내에서의 원근감과 관찰자의 위치를 변경
- 세분화, 수정, 심지어 특정 객체의 제거 또는 교체와 같은 장면 내 객체의 분석 및 조작성 향상

3. JPEG 플렌옵틱(plenoptic)

최근 JPEG 표준화위원회는 JPEG Pleno라는 이니셔티브를 만들었다. ‘pleno’는 ‘plenoptic’에 대한 것으로, 한 장면 내의 어떤 지점에 대한 정보뿐만 아니라 다른 위치에서 관찰될 때 어떻게 변하는지에 대한 수학적인 표현이다. ‘pleno’는 또한 ‘완성-완벽’을 뜻하는 라틴어로, JPEG 위원회가 오늘날 가능한 것 이상의 장면(이미지 기술)에 대한 더 완벽한 묘사를 제공하고자 하는 미래의 영상화에 대한 바람을 가리키는 말이다. JPEG Pleno 이니셔티브의 비전에 대한 이론적 근거를 논의하고 그것이 미래의 영상을 어떻게 재창조할 수 있는지를 제시한다.

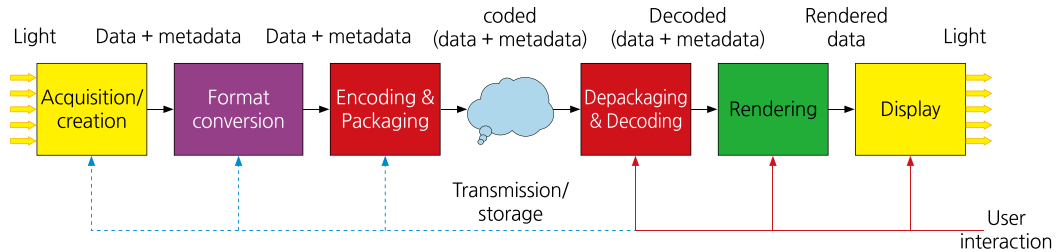
널리 연구된 플렌옵틱 표현은 시각 세계의 진정한 홀로그램 특성을 제공하여, 플렌옵틱 표현은 성가신 수학적 세부사항을 다루지 않고 표현이 무엇을 설명하는지에 대한 아이디어를 제공해야 한다.

핀홀 카메라로 장면을 관찰한다고 상상해 보라. 카메라는 시공간의 해당 시점에서 연필 P라고 하는 핀홀을 통과하는 광선 세트를 캡처한다. 카메라는 좌표 $P(h, U)$ 로 매개 변수화된 빛의 강도 분포를 등록한다. 색상이 파장 k 로 특징지어지는 컬러 카메라를 고려하면 연필 (h, U, k, t) 을 확장할 수 있으며, 동적 장면에서 t 는 시간을 나타낸다. 시각 세계의 전체적인 플렌옵틱 특성화는 핀홀 카메라를 공간 좌표 x, y, z 의 모든 시점에 위치시킴으로써 얻어지며, 그 결과 7차원 표현 $P(x, y, z, h, U, k, t)$ 로 묘사된다. [그



[그림 3] 이미지를 스캔하는 핀홀 카메라의 플렌옵틱 캡처

※ 핀홀을 통과하는 빛은 플렌옵틱의 값에 해당하며, 다른 방향 또한 가능함.



[그림 4] 엔드-투-엔드(끝과 끝) 플렌옵틱 처리 흐름

※ 처리 흐름은 핵심 아키텍처를 나타내며, 전형적인 플렌옵틱 콘텐츠 코딩 및 디코딩의 컴포넌트.

림 3은 플렌옵틱 카메라의 기본적인 영상처리 과정이다.

장면의 2D 및 3D 스캔을 기반으로 한 현재 시각적 정보의 표현 모델은 방금 설명한 플렌옵틱 표현에 포함된 정보의 일부만 해당된다. 이러한 콘텐츠는 추가로 이산화, 윈도우화, 서브 샘플링 및 양자화 된다.

이러한 콘텐츠는 핀홀 카메라 위치의 수, 강도 분포의 해상도, 색상 정보(빨간색, 녹색, 파란색 등)를 나타내는 구성 요소 수, 장면의 역학에 대처하기 위해 초당 캡처된 프레임 수 등 다양한 차원을 따라 더욱 이산화되고, 창이 열리고, 서브 샘플링되고, 정량화된다. 결과적으로 마지막 휴가 모임에서 사진을 찍던 카메라는 단지 사진을 찍은 장소의 전체적 표현

에 대한 하나의 근사치(많은 것 중)를 만들어낸다.

3.1 JPEG 플렌옵틱 처리 흐름

무한한 수의 접근법을 사용하여 플렌옵틱 콘텐츠를 캡처, 처리, 전달, 저장, 소비, 상호작용할 수 있지만, 이러한 시스템은 모두 플렌옵틱 프로세싱 흐름을 정의하기 위해 상호작용하는 별개의 구성 요소로 구성된 체인이라 볼 수 있다. 엔드-투-엔드(끝과 끝) 플렌옵틱 데이터 흐름을 보다 잘 이해하기 위해 [그림 4]의 프로세스에 주요 구성 요소를 보면, 콘텐츠 수집/창조로 시작으로 궁극적으로 디스플레이 및 사용자 상호작용으로 이어지는 프로세스의 주요 구성 요소를 보여준다.

3.1.1 수집/창조

플렌옵틱 처리 흐름의 제1 구성 요소는 적절한 센서를 이용하여 광 데이터 획득을 다룬다. 마이크로 렌즈를 장착한 2D 일반 카메라 어레이 및 장치에서부터 고성능 센서(depth sensors)에 이르기까지 다양한 센서를 사용할 수 있다. 각 센서는 특정 데이터 획득 모델에 의존하는데 예를 들어, 정기적인 카메라 어레이 또는 불규칙한 샘플링 그리드(포인트 클라우드 등)를 사용할 수 있다. 센서가 자연 및 실제 빛을 포착하지만 플렌옵틱 콘텐츠는 컴퓨터가 생성될 수 있다. 즉, 물리적 장면에서 직접 획득하지 못함을 의미한다.

전산화된 창작물은 자연수집과 동일한 모델을 사용하여 데이터를 생성할 수 있으며, 출처가 다른 데이터는 이음새를 없애는 방식으로 단일 콘텐츠에서 매끄럽게 혼합될 수 있다. 획득 및 생성 프로세스 모두 관련 메타데이터를 시각 데이터에 관련 형식으로 추가할 수 있으며, 이 메타데이터는 매우 다른 목적을 가지고 있을 수 있다. 센서, 렌즈, 카메라 배열을 설명하거나 의미 검색을 지원하거나 사생활 보호, 소유권 권리 또는 보안 통제를 지원하는 데 사용할 수 있다. 센서의 예로는 기존의 2D 카메라, 일반 및 임의 배열의 카메라, 마이크로 렌즈가 장착된 라이트 필드 카메라 및 깊이 센서가 있다.

3.1.2 형식 변환

수집/생성 데이터 모델은 크게 센서 및 획득 제약 조건에 의해 결정된다. 따라서 압축, 디스플레이 또는 사용자 상호작용 요구사항과 같은 전반적인 응용 프로그램 요구사항을 고려할 때 데이터를 표현하는 가장 적절한 방법이 아닐 수 있다. 획득한 (또는 합성으로 생성된) 플렌옵틱 데이터를 보다 적합한 모델로 변환하려면 형식 변환이 필요할 수 있다. 상호 운

용성의 이유로 표현 모델의 수는 가능한 적어야 하며, 단일 모델이 가장 이상적이다.

그러나 다양한 응용 프로그램 시나리오 및 관련 요구 사항을 고려할 때 다소 제한된 수의 모델을 가정하고 필요에 따라 대체 모델이 효율적으로 변환하는 것이 더 합리적이다.

이 변환은 데이터와 메타데이터 양방 모두 발생할 수 있지만, 데이터 변환은 변환 가정 및 제약 조건에 따라 더 높은 복잡성을 수반하고 최종 사용자 경험에 더 큰 영향을 미친다. 관련 변환의 예는 텍스처 세트 및 심화 뷰(depth views) 세트를 RGB 값이 있는 포인트 클라우드 세트나 연관된 텍스처가 있는 메시로 변환하는 것이다.

3.1.3 인코딩 및 패키징

인코더는 랜덤 액세스, 확장성 및 여러 복원성과 같은 관련 기능들의 추가 세트를 충족시키면서 조밀성 측면에서 데이터의 보다 효율적인 표현을 목표로 한다. 인코딩 과정은 인간의 지각 특성을 고려할 때 가장 일반적으로 손실이 있지만, 손실이 없거나 손실이 있을 수 있다. 이상적인 단일 코딩 형식은 상호 운용성을 극대화하지만, 코딩 형식은 선택한 표현 모델의 수에 따라 달라진다. 배경 및 전경과 같은 장면의 다른 부분은 다른 표현의 모델을 채택할 수 있으므로 동일한 장면에서 서로 다른 코딩 형식이 공존해야 한다.

다시 말하지만, 인코딩 프로세스는 데이터 및 메타데이터 모두에 적용되며 결국 약간 다른 요구 사항을 고려한다. 인코딩 후에 코딩된 데이터 및 메타데이터는 식별된 시스템-레벨 기능들을 제공할 수 있는 적절한 파일 형식으로 패키징된다.

3.1.4 디패키징 및 디코딩

전송 또는 저장 후 패키징 된 데이터는 디코딩을 위해 디패키징 되어, 채택된 표현 모델에서 데이터 및 메타데이터를 복구한다. 손실 인코더가 사용되는 경우 디코딩된 데이터는 해당 원본 데이터와 비교할 때 코덱 자체, 데이터 특성 및 사용된 코딩 속도에 따라 일정량의 왜곡을 갖는다.

3.1.5 렌더링

렌더링은 디코딩된 데이터에서 적절한 데이터를 추출하여 사용 가능한 디스플레이로 최상의 화질을 제공하는 기술이다. 디코딩된 데이터에서 대부분의 기존 이미지 및 비디오 응용 프로그램에서는 필터링 또는 대비 향상과 같은 품질 및 미관을 개선하기 위해서, 후처리로 렌더링을 사용할 수 있지만 렌더링 없이 디코딩된 콘텐츠가 디스플레이에 직접 제공되기도 한다. 그러나 대부분의 경우, 예를 들어, 특정 시점 또는 초점 평면을 얻기 위해 관련 데이터를 추출하려면 풍부한 가용 장면 표현을 ‘마이닝’해야 한다. 또한, 플렌 옵틱 콘텐츠는 헤드-마운트 디스플레이의 움직임을 통한 것과 같은 명시적인 명령 상호작용 또는 간접 상호작용을 통해 사용자 제어하에 있게 되기 때문에, 렌더링은 이제 다소 복잡하고 정교해져야 한다.

3.1.6 디스플레이

최종 디스플레이는 사용자 경험뿐만 아니라 다른 구성 요소에 대한 기술 솔루션 및 요구 사항도 결정된다. 최근 센서의 발전과 더불어 디스플레이는 자동 입체, 라이트 필드 및 Head mounted Displays(HMD)를 포함하여 상당히 진화하였다. 당연히 플렌옵틱 디스플레이가 이용 가능하다면 사용자는 보다 풍부한 플렌옵틱 표현으로부터 온전히 이

익을 얻을 수 있다.

3.1.7 사용자 상호작용

상호작용은 전송 측으로 돌아가서 더 자세하게 캡처하거나 캡처해야 하는 기본 관점을 정의해야 할 필요성이 있다. 시간이 지나면서 플렌옵틱 콘텐츠와 사용자 상호작용이 실제 세계의 상호작용만큼 자연스럽게 강력해질 것으로 기대된다.

이러한 처리 흐름으로 인해 시각적 콘텐츠 모델 및 처리에 새로운 문을 열어주지만, 디패키징 및 디코딩, 렌더링 및 디스플레이 구성요소(그림 4) 오른쪽) 특히 디코딩 구성 요소의 패키지 제거 및 디코딩, 렌더링 및 표시 광범위한 응용 프로그램 및 서비스 채택에서 가장 중요한 역할을 하는데, 이는 JPEG Pleno가 중요한 역할을 하는 곳이 바로 여기에 있다.

4. JPEG pleno의 프레임워크 비전

아리스토텔레스는 “전체는 그 부분의 합보다 크다.”라고 말했다. 이것은 JPEG Pleno의 기본 목표이다. 빛과 다양한 관련 양식을 가진 플렌옵틱 표현이라는 통일된 뿌리 실체가 있다는 것을 진지하게 그리고 효과적으로 고려하기 위한 도구 이상을 목표로 하기 때문이다. 이것은 JPEG Pleno의 개념 현장으로, 텍스처플러스 깊이(Texture-plus Depth), 라이트 필드, 포인트 클라우드 및 홀로그램 이미징과 같은 새로운 이미징 방식을 나타내는 표준 프레임워크를 제공하는 것을 목표로 한다. 또한 이러한 이미징은 어느 모델이 전체 또는 일부를 캡처 또는 생성했는지에 관계없이 플렌옵틱 기능에서 영감을 얻은 빛 표현으로 이해되어야 한다.


이미지 조작, 메타데이터, 랜덤 액세스 및 상호작용 및 다양한 파일 형식에 대한 고급 기능 지원을 제

공하는 완전히 통합된 시스템으로 이해되는 표현 프레임워크이다. 또한 프라이버시 보호, 소유권, 보안을 제공해야 한다.

JPEG Pleno 프레임워크는 실제 또는 합성된 세계에서 복제된 세계까지 모든 필요한 도구를 서로 다른 양식, 요구 사항 및 기능을 고려하면서 동일한 현실을 나타내기 위해 단일 시스템에 조화롭게 통합하는 데 초점을 맞추고 있다.

5. 맺음말

JPEG Pleno의 로드맵은 2015년 시작된 초기 이후 플로우를 따르고 있으며 기존 JPEG 형식이 이전 응용 프로그램과 동일한 오늘날의 디지털 이미징 응용 프로그램 유형의 영향을 미치기 위해 2020년 이후에도 계속될 것이다.

잠재적인 응용, 사용 측면에서 기여를 이끌어낸 체계적인 토론 회의 및 워크숍을 통한 탐색사례, 요구 사항 및 원하는 기능과 관련하여 JPEG 표준화위원회는 2016년에 공식적으로 새로운 작업 항목을 시작했으며, 2017년 하반기에 JPEG Pleno 사양에 대한 첫 번째 작업 초안을 작성하고 2018년부터 첫 번째 국제표준을 제안하는 제안이 진행 중에 있다. 

[참고문헌]

- [1] JPEG Pleno: Toward an Efficient Representation of Visual Reality
- [2] ISO/IEC JTC1/SC29/WG1N74014 'JPEG Pleno Call for Proposals on Light Field'
- [3] ISO/IEC AWI 21794 'JPEG Pleno Call for Proposals on Light Field Coding'