

플렌옵틱 영상 서비스를 위한 적응적 저지연 전송 기술

□ 정준영, 현은희, 류관웅, 이제원 / 한국전자통신연구원

요약

플렌옵틱은 물리공간 내 여러 방향의 빛 정보를 한꺼번에 센서에서 획득하고, 이를 그대로 재현함으로써 사용자에게 실제와 동일한 경험을 제공한다. 플렌옵틱 영상에 대한 스트리밍 서비스를 위해서는 여러 시점에서 획득된 영상 정보를 사용자로 전달하여야 하며, 또한 사용자의 시청 시점 변화에 대해 즉각적인 시점 전환이 이루어져야 한다. 이는 단순히 전송할 정보량의 증가뿐만 아니라 거의 실시간에 가까운 초저지연의 전송 기술을 요구한다. 본 고에서는 이러한 플렌옵틱 영상 서비스를 위해 적응적 저지연 전송 기술을 적용함으로써 고려해야 할 사항들에 대해 살펴본다.

I. 서 론

최근 3D 영화와 가상 현실의 발달로 3차원 영상

기술에 대한 관심이 증가하고 있다[1][2]. 기존 2차원 영상은 라이트 필드에 존재하는 공간 정보를 샘플링 하지만 3차원 입체영상 표현에 필요한 각도 정보는 손실한다. 라이트 필드 고유의 각도 정보를 유지하기 위해 이동식 또는 다중 카메라 시스템을 사용할 수 있지만, 이러한 시스템은 종종 크기와 비용의 제약에 의해 제한을 받는다. 그러나 플렌옵틱 카메라는 한번의 촬영으로 모든 광선과 그 방향에 대한 4가지가 변체적 표현인 라이트 필드를 캡처할 수 있도록 만들어졌기 때문에 공간 및 각도 정보를 모두 획득할 수 있다[3][4]. 획득된 정보는 보정 및 후처리를 통해 3차원 모델링[5], 원근 변환[6], 물체 인식[7] 등과 같은 다양한 응용 분야에 활용될 수 있다. 최근 사용자들은 소비자의 위치에서 벗어나 스스로 콘텐츠를 제작, 공유 및 향유하므로 플렌옵틱 디바이스에서 촬영

※ 이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(2020-0-00920,(세부2) 중대형 공간용 초고해상도 비정형 플렌옵틱 영상 저장/압축/전송 기술 개발).

한 영상을 즉시적으로 수정, 편집 후 공유할 수 있는 플렌옵틱 저작 툴 개발이 요구된다.

한편 영상 스트리밍 분야에서는 스마트폰, 태블릿 PC와 같은 스마트 기기의 보급과 넷플릭스(Netflix), 유튜브(Youtube) 등과 같은 OTT(Over the Top) 플랫폼의 등장으로 인해 동영상 스트리밍의 수요가 폭발적으로 증가하고 있다. 이러한 영상 스트리밍을 효율적으로 전송하기 위해 HTTP(Hyper-Text Transfer Protocol) 기반 적응적 스트리밍 기술(HAS, HTTP Adaptive Streaming)이 주목받고 있다. TCP(Transmission Control Protocol) 기반의 HTTP 적응적 스트리밍 전송 시스템의 경우 대용량 콘텐츠 제공을 위한 요청-응답의 지속적 반복으로 전송 효율성이 낮다는 문제점이 있다. 최근에는 이러한 문제를 해결하기 위해 TCP에 비해 신뢰성은 낮지만 저지연(low latency)에 유리한 UDP(User Datagram Protocol) 기반의 새로운 프로토콜들이 제안되고 있으며, 일부 응용서비스를 위해 이미 활용단계에 있다.

본 고에서는 대용량의 극초저지연 전송이 요구될 것으로 예상되는 플렌옵틱 영상 스트리밍을 위해 전송 프로토콜 관점에서 고려되어야 할 사항들에 대해 소개한다. 또한 현재까지 전송 프로토콜 차원에서 제시되고 있는 속도 개선을 위한 노력들과 보다 안정적이고 향상된 QoE(Quality of Experience) 제공을 목표로 하는 적응형 스트리밍에 대해 살펴본다.

II. 플렌옵틱 영상 스트리밍

플렌옵틱 영상 스트리밍 서비스는 현재 사용중인 2D 영상 스트리밍이나 AR/VR/XR, 다시점 3D 영상 서비스 등과 마찬가지로 인터넷을 기반으로 제

공될 것으로 예상되며, 이를 통해 보편화라는 단계를 거쳐 활성화를 기대할 수 있을 것으로 보인다. 이미 실감형/몰입형 콘텐츠를 이용한 서비스가 대용량, 저지연 전송에 대한 필연성을 제기한지 오래되었고, 많은 논의 및 개선들이 이루어지고 있는 상황이다. 그러나 이러한 노력에도 불구하고 일반 사용자들이 보다 자유로운 플렌옵틱 영상 기반의 3차원 영상 서비스를 누리기 위해서는 달성해야 할 목표가 녹녹하지 않은 상황이다. 본 절에서는 플렌옵틱 영상 스트리밍을 위한 전송 지연 관점에서 고려되어야 할 사항들에 대해 알아본다.

1. 전송 지연 요소

플렌옵틱 디스플레이들은 플렌옵틱 카메라에 의해 획득된 광선들의 분포인 라이트 필드를 재현하기 위해 각 픽셀들에 방향성을 부여하는 광학 구조가 필요하고, 이를 통해 올바른 Accommodation(관측자 단안의 수정체 초점 조절에 의한 깊이 인지)과 연속 운동 시차의 표시를 위해서 광선들을 매우 조밀한 각도 간격으로(보통 0.5도 이하) 재현해야 하는데 이는 라이트 필드 디스플레이의 요구 정보량을 현격하게 증가시킨다. 예를 들어 10도(수평) x 10도(수직) 시야각에서 0.5도 간격으로 광선들을 재생하며 각 방향으로 FHD(Full High Density) 해상도를 구현한다면 요구되는 총 정보량은 FHD x (10/0.5) x (10/0.5)이어서 FHD 해상도의 400배가 된다[22]. 따라서 플렌옵틱 서비스의 실제 구현을 위해서는 아주 높은 정보량을 전송하고 처리할 수 있어야 한다.

또한 단순히 정보량의 증가뿐만 아니라 거의 실시간에 가까운 전송 지연의 실현이 불가피하게 된다. 인터렉션 형태로 발생되는 사용자의 뷰포인트 변화에 대해 거의 실시간에 가까운 속도로 대응해야

사용자가 불편함을 느끼지 않는 상태로 플렌옵틱 영상을 통한 자유시점 이동을 즐길 수 있게 된다. 플렌옵틱 영상의 스트리밍에 있어서 뷰포인트의 변경은 클라이언트 내에 영상 재생을 위해 버퍼링되어 있는 데이터 이외에 추가 데이터를 필요로 하고, 이러한 데이터들은 안정성이 보장되지 않은 인터넷 망을 거쳐 실시간으로 요청되고 전송되어야 한다.

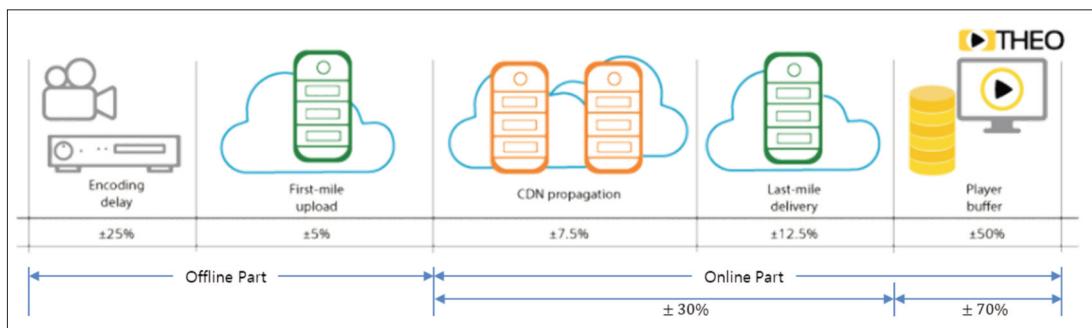
인터넷 망에서 일반적으로 영상 스트리밍을 제공하는 앤드-투-앤드 미디어 파이프라인은 복잡하며, 다양한 구성 요소를 지니게 된다. 또한 각 구성 요소들이 전체 서비스의 대기 시간에 영향을 준다. <그림 1>에서와 같이 영상 스트리밍의 공통 아키텍처를 기능에 따른 구간으로 구별하고 각 구간마다 전송 지연에 영향을 주는 요소들을 살펴보면, 첫 번째로 인코딩과 패키징 구간에서 해상도, 전송 비트레이트, 압축 표준 선택 및 패키징하는 시간과 주기 등에 따라 지연 요소들이 존재하게 된다. 두 번째로 인코딩된 영상은 네트워크를 통한 배포를 위해 인제스쳐단으로 전송해야 하는데 이 구간의 안정성이 실제 전체 스트리밍 서비스의 지연에 영향을 주게 된다. 세 번째로 콘텐츠 배포 구간으로 CDN (Content Delivery Network) 전파 구간에서 캐쉬 설정, 운용 절차 및 전

송 파라미터 설정 등을 통해 지역 요소들이 생기게 된다. 네 번째는 인터넷 망을 통한 전송 구간으로 사용되는 전송 프로토콜이나 실시간 적응적 스트리밍 매커니즘 등의 운용에 따른 지역 요소들이 존재하고, 마지막으로 수신단에서의 미디어 플레이어 버퍼링 및 유연한 재생을 위한 합성 등의 후처리 작업에 의해 지역 요소들이 발생하게 된다[9]. 물론 제공하고자 하는 서비스의 특성, 콘텐츠의 특성, 사용자들의 성향에 따라 각 구간에서의 기능 배분 및 리소스 할당 등에 대해 많은 튜닝 작업이 필요하다.

이와 같은 미디어 파이프라인에서 스트리밍 서버는 CDN 앞쪽에 위치하게 되며 라이브 스트리밍을 제외하면 서버로부터 클라이언트의 플레이어까지 가 온라인 구간이 된다. 즉, 플렌옵틱 카메라를 통해 촬영된 영상들에 대해 인코딩된 조각들이 스트리밍 서버에 존재하게 되고 클라이언트의 요청에 의해 해당 콘텐츠의 조각들이 전달되면서 디코딩과 렌더링을 거쳐 사용자에게 제공된다.

2. 전송 지연 최소화

몰입형 콘텐츠 중에 비교적 보편화된 가상 현실



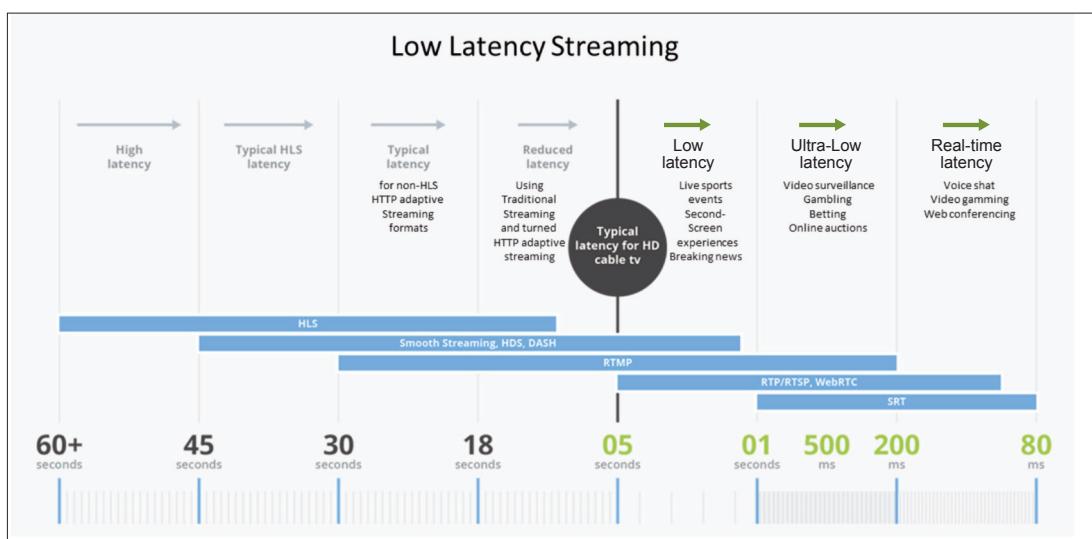
<그림 1> 스트리밍을 위한 공통 아키텍처[8]

콘텐츠의 경우 시청 시 사용자가 경험하게 되는 어지럼증과 멀미, 구토감을 제어하고 몰입감을 극대화하기 위한 기준이 제시되었는데, 2018년 TTA에 의해 발표된 VR 콘텐츠 제작 지침[10]에 따르면 시청자의 움직임(예를 들어 머리 돌리기)과 움직임 이후 나타나는 화면상의 배경이 업데이트되기까지 시간 지연을 의미하는 모션 투 포톤 지연시간(Motion to Photon Latency)은 20ms 이하로 유지되어야 하고 콘텐츠의 프레임률은 일반 영상 기반인 경우 초당 30프레임 이상을 권장하고 있다. 이러한 기준들을 참조해보면 플렌옵틱 영상에 대한 사용자의 뷰포인트 전환 시 이에 대한 반응 또한 해당 기준에 준하거나 그 이상이 되어야 할 것으로 보인다.

여기서 제시된 20ms의 지연시간 기준에 스트리밍 네트워크를 통한 플렌옵틱 영상 데이터의 전송을 포함하면 실현 불가능한 목표치가 될 것이다. 예를 들어 임의의 시청자가 플렌옵틱 영상 시청 중 시선 변경이나 움직임에 의해 뷰포인트를 이동하였을

때 이를 인지하여 새로운 뷰포인트를 재생하기 위한 추가 데이터를 요청하고 이에 대한 응답을 받아 디코딩, 렌더링을 거쳐 재생하는 과정을 모두 20ms 이내에 처리하는 것은 현재로서는 불가능해 보이며 가까운 미래에 해결될 수 있을 것 같지 않다. 결국 20ms의 목표치는 클라이언트에 이미 존재하고 있는 데이터들을 이용하여 변경된 뷰포인트 재생을 위한 처리에 사용되어야 할 것으로 보인다. 이를 위해 뷰포인트 이동 후 해당 뷰포인트의 영상 재생을 위한 정보를 받아 재생하는 방식이 아니라 예측이나 별도의 로직에 의해 사용자의 요청 시점 이전에 후보 정보들이 클라이언트에 이미 도달하여 있고, 사용자의 뷰포인트 이동을 감지하면 기획된 정보들을 이용한 디코딩과 렌더링을 거쳐 재생하면서 동시에 추가로 필요한 데이터들을 전송하는 방식이 타당해 보인다.

하지만 아무리 사용자의 뷰포인트 이동을 대비하여 필요할 수 있는 정보들을 미리 전송한다 하더라



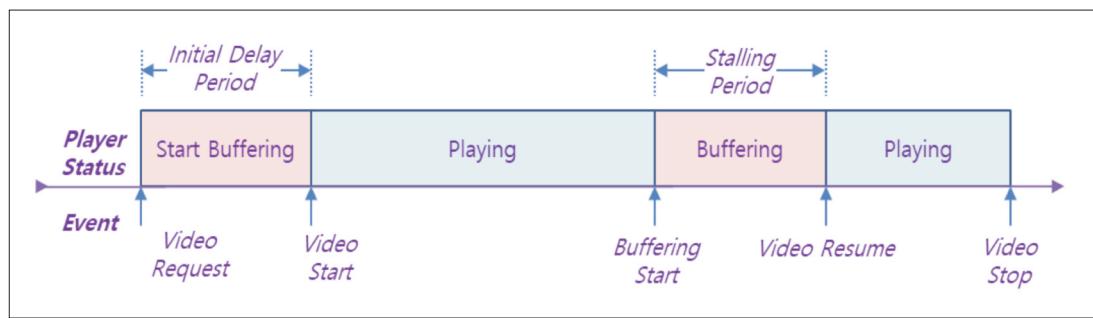
<그림 2> 응용 서비스 별 레이턴시[11][23]

도, 사용자의 뷔포인트의 이동은 화면 내에서 랜덤한 위치로 빈번히 발생될 것이므로 보다 많은 양의 데이터를 보다 빠르게 전송하기 위한 방법들에 대한 요구는 그 목표치를 계속 높여갈 것으로 보인다.

전송에 있어서 신속성은 레이턴시를 최소화하는 것으로 주문형 영상에 대한 인터랙티브 스트리밍 환경에서 레이턴시는 클라이언트에서 사용자의 서비스 요청이나 부가 액션을 인지한 후 이에 대한 정보가 서버로 전송되고 해당 액션에 따라 재생될 영상 조각들이 클라이언트로 전송되어 재생될 때까지의 시간으로 정의할 수 있다. 실제 제공되는 서비스의 종류에 따라 허용 가능한 레이턴시 기준이 상이하게 되는데, 허용되는 레이턴시에 따른 서비스 유형들을 살펴보면 <그림 2>에서와 같이 일반적인 대규모 시청자들이 있는 단방향 비디오 서비스의 경우 18초에서 30초 사이의 일반적인 레이턴시를 허용하며, 라이브 뉴스 및 스포츠 중계 서비스는 18초에서 5초 사이의 보다 단축된 레이턴시 기준이 요구된다. 일방적인 전달에 의한 서비스가 아니라 사용자 인터랙션이 포함된 양방향 특성을 가지는 e스포츠(e-sports) 같은 게임 중계 이벤트 등을 포함하는 스포츠 생중계 등의 서비스를 위해서는 5초 이하 1초까지의 저지연(Low Latency) 조건이 필요

하며, 실시간에 가깝게 느껴지는 더 나은 상호작용이 필요한 비디오 감시, 온라인 경매 등의 서비스를 위해서는 1초 이하의 초저지연(Ultra Low Latency) 스트리밍이 가능하여야 한다. 더 나아가 완전히 실시간에 준하는 인터렉션을 포함된 화상회이나 스트리밍 게임, 웹 컨퍼런스 등의 서비스는 0.2초 이하의 레이턴시를 요구하게 된다[23]. 여기에 더해 시청자의 반응에 실시간으로 반영하여 실시간 스트리밍이 가능하여야 하는 AR/VR/MR, 다초점 영상 전송, 360 비디오 전송, 플렌옵틱 영상 전송 등의 몰입형 콘텐츠 서비스의 경우 수십 밀리초에서 수 밀리초 이하로 거의 지연이 없는 극초저지연 실시간 전송이 가능하여야 한다.

레이턴시를 증가시키는 요소는 먼저 재생(Play) 버튼을 누른 후 혹은 자동 재생이 시작된 후 첫 번째 영상이 화면에 표시될 때까지의 초기 지연(<그림 3>의 Initial Delay 구간)과 영상 시청 도중 추가적인 후속 영상 조각이나 시청 환경, 리액션 등을 반영하기 위한 신규 영상 조각들의 리버퍼링 때문에 발생되는 일시적인 재생 멈춤(<그림 3>의 영상 재생 중 버퍼링 구간)에 의한 지연이 있을 수 있다[9][12]. 초기 전송 지연은 클라이언트의 영상 재생기가 화면 노출을 위해 서버로부터 일정 개수의 영상 세그



<그림 3> 스트리밍 지연

먼트를 전송받아 버퍼링하는 절차 때문에 발생하고, 이는 하나의 세그먼트가 포함하고 있는 영상 길이에 버퍼링 하여야 할 세그먼트의 개수를 곱한 시간만큼 걸리게 된다.

이러한 지연 요소들이 사용자의 QoE에 미치는 영향에 대한 연구[12]에 따르면 버퍼링을 위한 두 가지 구간에 대해 사용자들의 반응이 다르게 나타나는데 먼저 초기 지연에 관해서는, 이용자들이 해당 어플리케이션에 대한 일상적인 사용 경험으로부터 예상되는 대기시간을 감안하게 되므로 해당 서비스에 대한 만족도에 영향을 주지 않았으나 이미 시청 중이던 영상 서비스의 중단에 대해서는 서비스 이탈이 확대되는 것으로 파악되었다. 물론 다양한 형태의 영향요인들을 반영하여 추가적인 연구가 진행되어야 하겠지만 초기 지연에 비해 일시적인 재생 중단을 줄이기 위한 노력들이 진행되어야 할 것으로 보인다.

결국, 플렌옵틱 영상 스트리밍에 있어서도 사용자의 뷰포인트 이동에 따른 리버퍼링에 의한 지연을 최소화하는 것이 전송에서 해결해야 할 요소로 보인다. 궁극적으로 <그림 4>에서와 같이 전송을 위한 세그먼트 길이(첨크 단위)를 작게 하여 초기 지연을 최소화하고 적절한 전송 매커니즘을 개발하

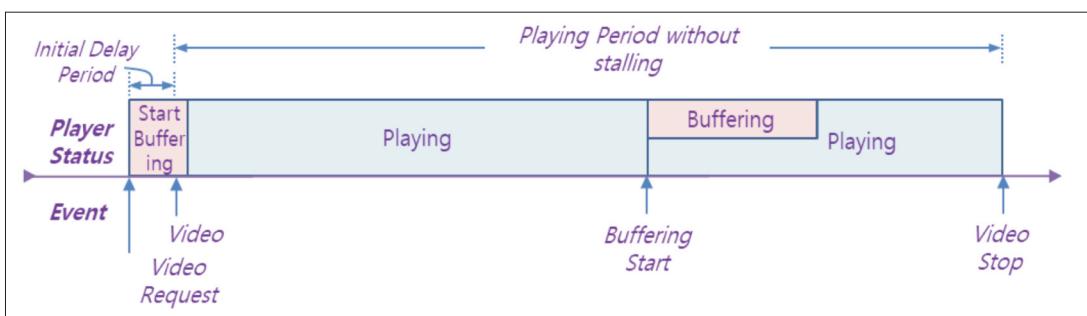
여 재생 중 리버퍼링 시 발생 가능한 지연 및 재생 중단이 사용자의 QoE를 만족할 수 있다면 최적의 솔루션이 될 것으로 보이며 단순한 트랜스포트 계층과 어플리케이션 계층 프로토콜의 성능 개선과 더불어 적절한 방식의 적응형 전송 기법에 대한 개발이 필요할 것으로 보인다.

III. 플렌옵틱 영상 전송을 위한 적응형 스트리밍 기술 & QoE 연구

본 절에서는 플렌옵틱 영상 전송을 위한 동적 적응형 스트리밍 개념 및 구조를 제시하고, 기존 적응형 스트리밍 전송기법들을 기반으로 하는 플렌옵틱 영상 전송 기술과 관련된 연구 결과들에 대해서 알아본다. 또한, 플렌옵틱 영상에 대한 QoE 평가에 있어서 객관적인 평가 지표와 표준 등이 존재하지 않으므로 기존 연구[13]에서 제시한 주관적인 QoE 평가 방법에 대해 살펴본다.

1. 동적 적응형 스트리밍 기술 연구

플렌옵틱 영상 전송을 위한 적응형 스트리밍 프



<그림 4> 스트리밍 지연 최소화 방안

로토콜의 목적은 효율적이고 사용자 친화적인 미래 서비스를 구축하기 위해 인식된 시각화 품질의 측면을 고려하면서 보다 적은 중단(Interruption)을 보장하는 QOE 중심의 재생을 가능하게 하는 것이다.

현재 플렌옵틱 관련 연구는 주로 영상 콘텐츠 전송보다 스틸 이미지 시각화에 더 초점을 맞추고 있다. JPEG Pleno 프레임워크[14]는 표준화를 목적으로 라이트 필드 포맷을 개발하고 있으며 MPEG-I는 2018년도 라이트 필드 영역 주소를 지정하기 시작했다.

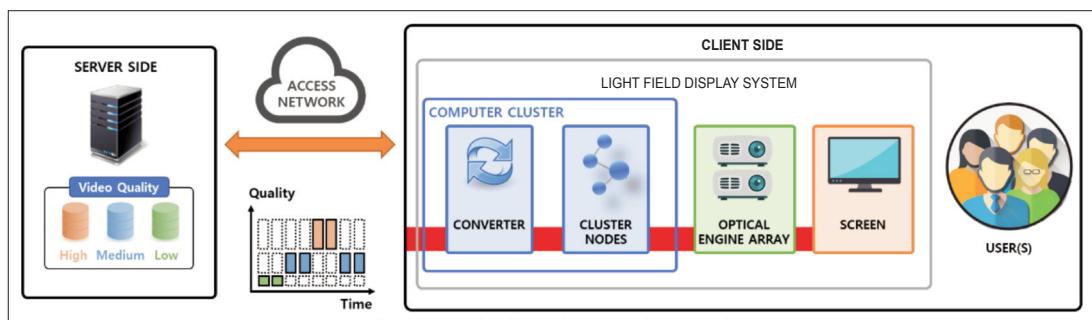
플렌옵틱 영상 스트리밍의 경우 퀄리티 전환이 얼마나 허용 가능할 수 있는지 그리고 재생 중에 이벤트를 중지하는 것보다 각 해상도를 줄이는 것 이 더 수용 가능한지 여부 등에 대한 고려 사항들이 있을 수 있다. 요약하자면, 서로 다른 공간 해상도(Spatial Resolution)나 각 해상도(Angular Resolution)를 갖는 표출 기반의 동적(Dynamic) 적응형 스트리밍의 개념이 플렌옵틱 영상 스트리밍에도 적용 가능한지가 증명되어야 한다.

〈그림 5〉는 동적 적응형 스트리밍 구성도를 나타낸다. 콘텐츠의 경우 카메라 어레이를 이용하거나 컴퓨터에서 가상 장면으로부터 렌더링 될 수 있다. 이러한 뷰들은 서로 다른 공간과 각 해상도를 가지

고 비디오 스트리밍 데이터베이스에 저장되므로 품질 속성과 저장 및 대역폭 요구사항이 다르다. 클라이언트는 지정된 품질 관련 파라미터를 사용하여 콘텐츠 세그먼트를 요청한다. 라이트 필드 디스플레이 시스템이 해당 세그먼트를 수신하면 이를 광학 엔진(Optical Engine)을 통해 3D 라이트 필드를 렌더링 하는 클러스터 내의 컴퓨터에 분산시켜 화면에 빛을 투사한다.

컨버터(Converter)는 서버에서 전송된 영상 세그먼트를 수신해서 클러스터 노드(Cluster Node)가 광학 엔진의 기능과 개수를 기초해서 콘텐츠가 동일한 각과 공간 해상도를 가지고도록 변환시키는데 필요한 데이터를 전달한다. 이 때, 컨버터에서 수신된 화질에 따라 클러스터 노드로 전달되는 영상의 화질은 차이가 있지만 그 크기는 화질과 상관없이 동일한 크기를 가진다.

시각화 프로세스를 빠르게 하기 위해서 서버에서 서비스 중인 클라이언트의 디스플레이 정보를 안다는 가정하에 최적화된 품질의 콘텐츠를 전송할 수 있지만 문제점이 발생한다. 우선, 서버는 모든 클라이언트의 디스플레이 정보를 저장하고 있어야 하며 보내고자 하는 콘텐츠 화질이 디스플레이 화질보다 낮은 경우 콘텐츠를 업스케일링해서 전송해야 하기



〈그림 5〉동적 적응형 스트리밍 구성도

때문에 전송량이 증가하게 된다.

우선적으로 기존 스트리밍 서비스를 위해서 연구되었던 방식들이 플렌옵틱 영상 스트리밍 서비스에도 적용 가능한지에 대한 사용자의 주관적인 평가들에 대한 연구가 진행되고 있는 중이다. 플렌옵틱 영상을 시각화하는 동안 광선이 홀로그래픽 스크린에서 불규칙한 위치에 도달하여 기존의 픽셀 개념은 적용되지 않지만, 픽셀 개념은 여전히 소스 데이터, 즉 변환될 렌더링 된 장면의 이산 뷰(Discrete Views)에 적용될 수 있다. 예를 들어, 기존 2D에서 적용된 공간 해상도는 여전히 적용 가능하다.

디스플레이의 각 해상도는 주어진 고정값으로, 시스템의 광학 엔진의 배치와 파라미터에 의해 결정된다[15]. 콘텐츠의 각 해상도는 FOV(Field of View) 크기에 대한 소스 뷰의 수로부터 계산된다. 예를 들어 디스플레이의 FOV가 45도이고 뷰의 개수가 90개일 때, 도당 뷰 수는 2가 되고 각 해상도는 0.5로 계산된다. 플렌옵틱 영상에서 각 해상도가 공간 해상도 보다 더 중요한 평가 척도이다. 앞서 정의한 공간 및 각 해상도를 줄임으로써 데이터 전송량을 줄일 수 있을 것으로 판단되지만 적절한 해상도를 결정하는데 있어 콘텐츠의 종류와 시청자의 시각 인식 정도가 중요한 파라미터가 될 것이다.

선행연구[13]에서는 HoloVizio 80WLT[16]를 사용하여 최대 180도의 FOV를 지원하는 콘텐츠를 이용해서 정적 관찰자가 인식 가능한 FOV가 얼마나 되는지 실험하였다. 개인 관찰자 입장에서는 FOV 135도 이상에서는 어떤 유의미한 차이를 느끼지 못한다는 결과를 보여줬다. 이것은 180도 대신 135도 FOV를 제공하면 25%의 데이터 감소를 유도하는 동시에 사용자의 요구를 충족시킬 수 있다는 것을 의미한다. 하지만 다수의 관찰자가 동시에 보기 원하는 전시 사용 사례의 경우는 135도 이상의 더 큰

FOV가 요구된다. 따라서, FOV 조절을 통한 전송 데이터 양을 감소시키는 방식은 시청 환경이 주요 고려사항이 됨을 알 수 있다. 활용 중인 FOV의 일부에 대해서만 플렌옵틱 영상 데이터를 전송하면 전송 속도가 현저히 감소할 수 있지만 관찰자 위치에 대한 실시간 정보가 필요하다. FOV 내에서 사용자 위치 추적이 가능하다면 타일 기반 SRD(Tile-based Spatial Relationship Description) 스트리밍에서[17] 사용되는 알고리즘을 이용하여 데이터 전송량을 대폭 줄일 수 있다.

보간 기법(Interpolation Techniques)은 원래의 뷰들 사이에 중간 뷰를 생성하며, 이를 통해 전체 뷰 수와 각 해상도를 높일 수 있는 방법이다. [18]의 저자들은 7점 비교 척도를 사용하여 관찰자가 보간된 영상과 그에 대응되는 높은 각 해상도에서의 직접 렌더링 된 영상에 대해 주관적 품질 평가를 수행했다. 도당 1개 뷰를 입력으로 넣고 보간법을 이용해서 각 해상도를 증가시키는 방법은 QoE를 상당히 높이지만 도당 0.4뷰일 경우 보간법이 유의미한 성능을 내지 못했다. 보간 기법의 경우 연산 요구사항이 너무 높아 실시간 솔루션에 적합한 렌타임을 가능하게 할 수 없기 때문에 실시간 적응형 스트리밍에 아직은 적합하지 않다.

동적 적응형 플렌옵틱 영상 스트리밍 전송은 클라이언트 측에 변환 모듈이 있어야 의미가 있다. 현재 변환 과정이 실시간으로 가능한지는 논쟁의 여지가 있다. 그 이유는 변환 과정에서 공간 및 각 해상도가 높을수록 더 많은 시간이 소요되기 때문이다. [19]의 라이브 캡처 시스템과 같이 18개의 소스 뷰만 사용한다면 10ms 정도의 변환 시간이 소요되므로 시간이 중요한 활용 사례 시나리오에 적합할 수 있다.

2. 플렌옵틱 영상 QoE 평가 연구

현재 시스템과 서비스의 성공은 근본적으로 그들이 제공하는 사용자 경험에 달려 있다. 플렌옵틱 영상 시각화의 QoE는 다른 형태의 시각화에는 적용되지 않는 독특한 시각적 현상 때문에 그 특성들을 파악해야 할 뿐만 아니라, 그에 상응하는 주관적 품질 평가 기법이 아직 표준화되어 있지 않기 때문에 이에 대한 연구가 필요하다.

몰입형 3D 경험은 연속적인 수평 운동 시차(Horizontal Motion Parallax)에서 비롯된다. 이는 관찰자가 측면으로 전환하는 동안 시차 효과가 때 끄럽게 나타나며, 보이는 이산 뷰가 없다는 것을 의미한다. 각 해상도가 높을수록 수평 시차는 더 부드러워지지만 전송 속도의 요구사항도 높아진다. 즉, 소스 뷰가 많을수록 더 많은 데이터를 전송해야 함을 의미한다. 따라서, 지원 가능한 총 데이터 크기를 유지하면서 탁월한 사용자 경험을 갖기 위해서는 충분히 높은 각 해상도를 제공하는 것이 중요하다. 관찰자가 정적 상태에서 각 해상도에 따른 QoE 결과와 동적 상태일 때의 실험 결과는 다르다. 동적 상태일 때와 비교해 정적 상태일 때 관찰자가 낮은 각 해상도에서 영상에 대한 이질감을 덜 느낀다.

현재 라이트필드 디스플레이는 주로 수평 시차(Horizontal Parallax Only, HPO) 전용 모델로 개발되고 있다. 수평 이동 시차 축의 유연성(Smoothness)이 인식 품질 지표로 고려된다. 각 해상도가 높을수록 유연도가 높아진다. 충분히 높은 각 해상도를 제공하지 못하면 3D 연속성이 감소할 뿐만 아니라, 크로스톡(Crosstalk) 효과가 발생한다. 여기서 크로스톡은 콘텐츠의 인접한 이산 뷰가 서로 겹쳐져 동시에 보이는 현상을 의미한다.

평균적으로 각 해상도가 1도당 소스 뷰가 2개 이

상일 때 관찰자가 수평 모션 시차의 부드러움에 만족하는 경향이 있었으며 1도당 소스 뷰가 1개 미만일 때 콘텐츠의 품질이 크게 저하되는 것으로 지각하였다. 다만, 콘텐츠의 내용에 따라 1도당 소스 뷰가 1개일 때도 관찰자가 여전히 보기 충분한 수준이라 지각하였다.

또한, 낮은 공간 해상도로 인한 흐림(Blur) 효과가 콘텐츠의 각 해상도가 충분하지 않아 발생하는 시각적 품질 저하를 실제로 줄일 수 있다는 실험 결과가 있다[20]. 이 결과는 품질 전환 중에 공간 해상도의 손실은 동시에 시각화된 콘텐츠를 흐리게 하여 QoE를 낮출 가능성이 있지만 불충분한 각 해상도를 보상하여 시차 효과의 저하를 줄여 개선할 수 있음을 의미한다.

다음은 [13]에서 제시한 동적 적응형 플렌옵틱 영상 스트리밍의 QoE 평가 방법을 소개한다. 비디오 콘텐츠는 품질 전환과 스톤링(Stalling) 이벤트로 생성했으며 여기서 스톤링은 화면 정지와 같은 현상을 의미한다. 품질 전환은 공간 및 각 해상도를 조절하고 스톤링은 지속시간을 조절하여 콘텐츠를 생성했다. 시험 조건은 1) 공간 해상도를 감소시키는 품질 전환, 2) 각 해상도를 감소시키는 품질 전환, 3) 공간 해상도와 각 해상도 모두를 감소시키는 품질 전환, 4) 스톤링 이벤트로 4가지였다. 품질 전환의 경우 높은 해상도에서 낮은 해상도로 수행됐다. 실험은 하나의 품질 전환 또는 스톤링 이벤트만 포함된 형태

<표 1> 테스트 콘텐츠 별 공간 해상도

콘텐츠	고 해상도	저 해상도
Red	1024 X 768	640 X 480
Yellow	1024 X 768	800 X 600
Ivy	960 X 540	640 X 360
Tesco	1280 X 720	640 X 360
Gears	1920 X 1080	640 X 360



<그림 6> 테스트 콘텐츠 영상 화면(Red, Yellow, Ivy, Tesco, Gears)

<표 2> 테스트 콘텐츠 파라미터

콘텐츠	프레임율(FPS)	재생시간(초)	인코딩 속도(kbps)	압축 방식
Red	25	14.4	130,240	H.264
Yellow		13.6	704,160	
Ivy		10	38,000	
Tesco		12.5	76,800	
Gears		7.2	286,000	

로 진행되었고 이벤트는 항상 콘텐츠 재생시간 중간에 발생하도록 하였다. <표 1>은 콘텐츠 소스로 사용된 2가지 품질의 공간 해상도를 보여준다.

콘텐츠의 각 해상도는 도당 1개 또는 2개의 소스 뷰를 사용했다. 콘텐츠 FOV를 45도로 고정했기 때문에 45개 또는 90개의 소스 뷰가 사용되었다고 할 수 있다. 스톰링 지속시간은 500ms 또는 1,500ms로 정했는데 그 이유는 관련 연구들에서 500ms는 관찰자가 눈으로 확인은 되지만 용인 가능한 수준의 시간이며 1,500ms는 관찰자가 견디기 힘들고 상당한 자연시간으로 간주될 수 있는 시간으로 실험 결과를 보여주고 있기 때문이다[21]. 스톰링 이벤트에는 일반적인 회전하는 원형 그래픽 표시가 없는데 이는 관찰자에게 자연 시간을 매우 명시적으로 만들 가능성이 있기 때문이다. <그림 6>은 QoE 측정을 위해 사용된 HoloVizio C80 라이트 필드 시네마 시스템에 표현된 소스 뷰를 보여준다.

실험에서 사용된 콘텐츠의 프레임률은 25FPS이다. <표 2>는 테스트 콘텐츠별로 정의된 파라미터 값을 보여준다.

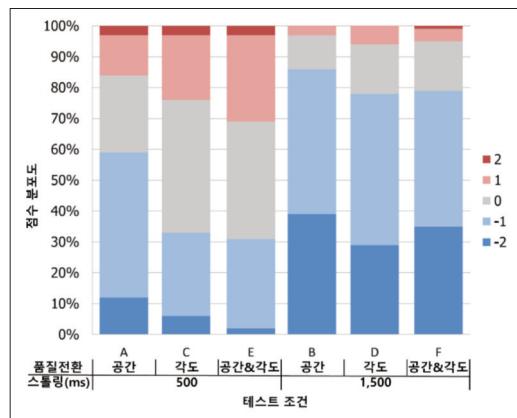
주관적 품질 평가를 위해서 3개의 품질 전환과 2개의 스톰링 이벤트를 쌍으로 평가하였다. <표 3>은 평가하는 테스트 쌍을 보여준다.

<표 3> 평가하는 테스트 조건 쌍

테스트 ID	품질 전환	스톰링 지속시간(ms)
A	공간 해상도	500
B	공간 해상도	1,500
C	각 해상도	500
D	각 해상도	1,500
E	공간 및 각 해상도	500
F	공간 및 각 해상도	1,500

평가는 5점 비교 척도를 사용했다. 관찰자는 쌍의 두 번째 콘텐츠(스톰링 이벤트)와 첫 번째 콘텐츠(품질 전환 이벤트)를 비교했다. 여기서 양의 값 (+1, +2)은 스톰링 이벤트가 품질 전환 이벤트보다 견딜만하다는 선호도를 나타내며, 반대로 음의 값 (-1, -2)은 품질 전환이 스톰링 이벤트보다 견딜만하다는 선호도를 나타낸다.

평가를 위해서 사용된 라이트필드 디스플레이인



<그림 7> 시험 조건 별 비교 점수 분포

HoloVizio C80 라이트 필드 시네마로 테스트를 위해서 디스플레이에는 45도의 FOV로 보정되었다. 테스트 환경 조명의 조건은 20lx이다. 테스트 참가자는 디스플레이 중앙에서 시청 거리 4.6m를 떨어져 시청하며 약 1m의 측면 이동도 가능했다. 총 20명의 테스트 참가자가 실험에 참여했다. <그림 7>은 시험 조건 별 비교 점수 분포를 보여준다.

1,500ms 높은 스톤링 이벤트와 품질 전환과 비교할 때 일반적으로 대부분의 관찰자들은 품질 전환에 대한 선호도가 더 높았다. 스톤링 이벤트 선호도의 경우 세 가지 조건(B,D,F)에서 모두 6% 이하였다. 500ms 낮은 스톤링 조건(A,C,E)에서의 선호도 결과를 보면 조건 A에서 관찰자의 59%가 품질 전환에 대한 선호도가 더 높았으며 16%만 스톤링

<표 4> 공간 및 각 해상도 변화 따른 요구 대역폭 감소량

콘텐츠	공간 해상도(%)	각 해상도(%)	공간&각 해상도(%)
Red	39		19.5
Yellow	61		30.5
Ivy	44.4	50	22.2
Tesco	25		12.5
Gears	11		5.5

이벤트를 선택했다. 조건 C와 E의 경우 품질 전환과 스톤링 이벤트에 대한 선호도가 비교적 균등한 것을 알 수 있다. 이 결과에서 각 해상도만 감소된 품질 전환과 각 및 공간 해상도 모두 감소된 품질 전환 간의 차이는 크지 않다는 것을 알 수 있다. <표 4>는 공간 및 각 해상도를 줄임으로써 얻은 대역폭 요구 감소량을 나타낸다. <표 4>에서 25%는 해당 고해상도 콘텐츠 대비 요구 대역폭이 1/4 만큼 줄어든다는 것을 의미한다.

플렌옵틱 콘텐츠에 대한 QoE 연구 평가 결과는 긴 스톤링 이벤트 대신 공간 및 각 해상도 조절을 통한 품질 전환이 분명히 더 선호되며 짧은 스톤링 이벤트와 비교해서도 더 선호될 수 있다는 것을 보여준다. 비록 본 실험에서는 감소된 각 해상도가 1도였지만 QoE 측면에서 스톤링 이벤트와 비교해서 공간 및 각 해상도가 모두 감소된 품질 전환과 각 해상도만 감소된 품질 전환 간의 선호도 차이는 미비하다는 것을 보여준다. 즉, 콘텐츠의 특성에 따라서 공간 및 각 해상도 모두를 감소시켜 전송하는 방법이 전송 측면에서는 유리할 수 있음을 보여준다.

현재 QoE 연구는 플렌옵틱 영상에 대한 객관적인 품질 메트릭이 존재하지 않기 때문에 주로 주관적인 평가들로 진행되고 있다. 또한 앞서 살펴본 QoE 평가의 경우 콘텐츠당 하나의 이벤트가 발생하는 제한적인 환경에서 수행되었다. 하나의 콘텐츠 품질 평가에서 품질 전환 또는 스톤링 이벤트 등 평가 항목들이 다양하게 존재하는 환경에서의 QoE 평가가 추가적으로 이루어져야 할 것이다.

IV. 결 론

플렌옵틱 기술은 라이트 필드 기반으로 실제 물

리계와 동일한 시각 경험을 제공하기 위해 필요한 영상 취득, 편집, 가공, 활용하는 기술이다. 현재 플렌옵틱 영상 분야는 영상 획득 기술에 대한 연구 개발이 활발히 진행되고 있으며 이를 위한 해결 수단으로 마이크로 렌즈를 포함한 구성 최적화 기술이 집중적으로 연구되고 있다. 또한 플렌옵틱 영상에 대한 재초점 및 템스 추정, 초해상도 플렌옵틱 영상 처리 및 저작 기술에 대한 연구도 진행되고 있다.

그러나 플렌옵틱 영상 스트리밍을 위한 전송 기술에 대한 연구는 이제 시작 단계이다. 본 고에서는 플렌옵틱 영상 스트리밍을 위한 전송 프로토콜 관

점에서 고려되어야 할 사항들과 전송 프로토콜 차원에서 제시되고 있는 속도 개선을 위한 기술들에 대해 살펴보았다.

또한 안정적이고 향상된 QoE 제공을 목표로 하는 동적 적응형 스트리밍 기술에 대해 검토하였으며 QoE 관점에서 전송 대역에 관련된 연구결과들을 살펴보았다. 앞에서 언급한 바와 같이 플렌옵틱 서비스의 대중화를 위해서는 대용량 초저지연 전송이라는 매우 어려운 기술적 문제를 해결하여야 할 것이다. 이와 관련된 연구 개발에 더 많은 노력이 요구된다.

참고 문헌

- [1] Z. Pan, Y. Zhang, and S. Kwong, "Efficient motion and disparity estimation optimization for low complexity multiview video coding," *IEEE Trans. Broadcast.*, vol. 61, no. 2, pp. 166-176, Jun. 2015
- [2] J. Lei, S. Li, C. Zhu, M.-T. Sun, and C. Hou, "Depth coding based on depth texture motion and structure similarities," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 2, pp. 275-286, Feb. 2015
- [3] Y. Fang, G. Ding, J. Li, and Z. Fang, "Deep3DSaliency: Deep stereoscopic video saliency detection model by 3D convolutional networks," *IEEE Trans. Image Process.*, vol. 28, no. 5, pp. 2305-2318, May 2019
- [4] J. Lei, X. He, H. Yuan, F. Wu, N. Ling, and C. Hou, "Region adaptive R-_ model-based rate control for depth maps coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 6, pp. 1390-1405, Jun. 2018
- [5] M. M. Corral, A. Dorado, H. Navarro, A. Llavorad, G. Saavedra, and B. Javidi, "From the plenoptic camera to the-at integral-imaging display," *Proc. SPIE*, vol. 9117, Jun. 2014, Art. no. 91170H
- [6] G. Wu, B. Masia, A. Jarabo, Y. Zhang, L. Wang, Q. Dai, T. Chai, and Y. Liu, "Light _eld image processing: An overview," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 7, pp. 926-954, Oct. 2017
- [7] A. Ghasemi and M. Vetterli, "Scale-invariant representation of light field images for object recognition and tracking," *Proc. SPIE*, vol. 9020, pp. 15-23, Mar. 2014
- [8] Steven Tielaemans , "The importance of low latency in video streaming", <https://www.linkedin.com/pulse/importance-low-latency-video-streaming-steven-tielaemans>, 2018.10.10.
- [9] Eric Kim, "라이브 비디오 서비스 구축을 위한 노하우", <https://www.popit.kr/라이브-비디오-서비스-구축을-위한-노하우-2편/> , 2019.01.21.
- [10] TTA, "멀미 저감을 위한 머리 장착형 영상 장치 기반 가상현실 콘텐츠 제작지침", TTA.KO-10.1030/R1, 2018.12.19.
- [11] GM Cusiac, "Why SRT, HLS, and MPEG-DASH are the future of streaming", <https://www.epiphan.com/blog/why-srt-hls-mpegdash-streaming/>, June 30, 2020
- [12] T. Hossfeld, S. Egger, R. Schatz, M. Fiedler, K. Masuch, C. Lorentzen, "Initial delay vs. interruptions: Between the devil and the deep blue sea", 2012 Fourth International Workshop on Quality of Multimedia Experience, 2012.7.5.
- [13] Kara, Peter A., et al. "Evaluation of the concept of dynamic adaptive streaming of light field video." *IEEE Transactions on Broadcasting* vol. 64. No. 2, pp. 407-421, 2018

- [14] T. Ebrahimi, S. Foessel, F. Pereira, and P. Schelkens, "JPEG Pleno: Toward an efficient representation of visual reality," *IEEE Multimedia*, vol. 23, no. 4, pp. 1420, Oct./Dec. 2016
- [15] P. T. Kovacs, A. Boev, R. Bregovic, A. Gotchev, "Quality measurements of 3D light-field displays," Eighth International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM), 2014
- [16] Holografika. (Jun. 2017). Holovizio 80WLT Light Field Display. [Online]. Available: <http://www.holografika.com/Documents/80WLTempalysize.Pdf>
- [17] L. D'Acunto, J. van den Berg, E. Thomas, and O. Niamut, "Using MPEG DASH SRD for zoomable and navigable video," in Proc. ACM 7th Int. Conf. Multimedia Syst., Klagenfurt, Austria, 2016, Art. no. 34
- [18] A. Cserkaszky, P. A. Kara, A. Barsi, M. G. Martini, "To Interpolate or not to Interpolate: Subjective Assessment of Interpolation Performance on a Light Field Display," IEEE International Conference on Multimedia and Expo (ICME) 8th Workshop on Hot Topics in 3D Multimedia (Hot3D), 2017
- [19] V. K. Adhikarla, F. Marton, T. Balogh, and E. Gobbetti, "Real-time adaptive content retargeting for live multi-view capture and light field display," *Vis. Comput.*, vol. 31, nos. 68, pp. 1023-1032, 2015
- [20] P. A. Kara et al., "The interdependence of spatial and angular resolution in the quality of experience of light field visualization," in Proc. IEEE Int. Conf. 3D Immersion (IC3D), 2017, pp.18
- [21] P. A. Kara, W. Robitz, M. G. Martini, C. T. Hewage, and F. M. Felisberti, "Getting used to or growing annoyed: How perception thresholds and acceptance of frame freezing vary over time in 3D video streaming," in Proc. IEEE Int. Conf. Multimedia Workshops (ICMEW), pp.16, 2016
- [22] 박재형, "Light field display 기술", 한국정보디스플레이학회지, 제16권 4호, 2015, 24p~34p
- [23] Nicolas Weil, Romain Bouqueau, "Ultra Low Latency with CMAF", https://parisvideotech.com/wp-content/uploads/2017/07/Bouqueau-Weil_UltraLowLatencyWithCMAF.pdf, Paris Video Tech Meetup #4, July, 2017

필자소개



정준영

- 1999년 2월 : 영남대학교 전자공학 학사
- 2001년 2월 : KAIST 전자공학 석사
- 2010년 8월 : 충남대학교 정보통신공학 박사
- 2001년 2월 ~ 현재 : 한국전자통신연구원 미디어연구본부 책임연구원
- 주관심분야 : 디지털 방송 시스템, 멀티미디어 스트리밍, 방송 서비스 기술



현은희

- 1986년 2월 : 충남대학교 전자계산학 학사
- 1988년 2월 : 중앙대학교 전자계산학 석사
- 1988년 2월 ~ 현재 : 한국전자통신연구원 미디어연구본부 책임연구원
- 주관심분야 : 멀티미디어 스트리밍, 방송 서비스 기술

필자소개

류관웅



- 1997년 2월 : 영남대학교 전자공학 학사
- 1999년 2월 : 영남대학교 정보통신공학 석사
- 2006년 2월 : 영남대학교 정보통신공학 박사
- 2004년 2월 ~ 2005년 1월 : NTT DoCoMo YRP 연구소 인턴쉽
- 2006년 7월 ~ 2009년 1월 : 주XRONET 책임연구원
- 2009년 7월 ~ 현재 : 한국전자통신연구원 미디어연구본부 선임연구원
- 주관심분야 : 방송전송시스템, 통신시스템, 디지털신호처리

이제원



- 2011년 2월 : 충남대학교 전기정보통신공학부 전자전파정보통신 학사
- 2013년 2월 : 충남대학교 전자전파정보통신공학과 석사
- 2019년 2월 : 충남대학교 전자전파정보통신공학과 박사
- 2015년 11월 ~ 현재 : 한국전자통신연구원 미디어연구본부 연구원
- 주관심분야 : 케이블 방송 및 데이터 통신, 멀티미디어전송 기반 기술