

대형 가상현실 공연장을 위한 360 도 비디오 스트리밍 시스템 프로토타입 구현

류영일, 최이현, 류은석

성균관대학교

yeongilryu@skku.edu, ceh9141@skku.edu, esryu@skku.edu

Implementing 360-degree VR Video Streaming System Prototype for Large-scale Immersive Displays

Yeongil Ryu, YiHyun Choi, Eun-Seok Ryu

Sungkyunkwan University

요 약

최근 K-Pop 을 위시한 예술공연 콘텐츠에 몰입형 미디어를 접목한 온택트 (Ontact) 미디어 스트리밍 서비스가 주목받고 있는 가운데, 본 논문은 일반적으로 사용되는 2D 디스플레이 또는 HMD (Head-Mounted Display) 기반 VR (Virtual Reality, VR) 서비스에서 탈피하여, 대형 가상현실 공연장을 위한 360 도 VR 비디오 스트리밍 시스템을 제안한다. 제안된 시스템은 Phase 1, 2, 3 의 연구개발 단계를 밟아 6DoF (Degrees of Freedom) 시점 자유도를 지원하는 360 도 VR 비디오 스트리밍 시스템을 개발하는 것을 최종목표로 하고 있으며, 현재는 Phase 1: 대형 가상현실 공연장을 위한 3DoF 360 도 VR 비디오 스트리밍 시스템 프로토타입의 개발까지 완료되었다. 구현된 스트리밍 시스템 프로토타입은 서브픽처 기반 Viewport-dependent 스트리밍 기술이 적용되어 있으며, 기존 방식과 비교하였을 때 약 80%의 비트율 감소, 약 543%의 영상 디코딩 속도 향상을 확인하였다. 또한, 단순 구현 및 성능평가에서 그치지 않고, 실제 미국 UCSB 에 위치한 대형 가상현실 공연장 AlloSphere 에서의 시범방송을 수행하여, 향후 Phase 2, 3 연구단계를 위한 연구적 기반을 마련하였다.

1. 서론

2019 년부터 유행한 코로나바이러스감염증-19 (COVID-19)의 영향으로 교육, 근로, 외식과 같은 기본적인 사회활동뿐만 아니라, 영화 및 여행과 같은 엔터테인먼트 산업 전반적인 분야에서도 비접촉 커뮤니케이션 중심의 언택트 (Untact) 서비스들이 빠르게 개발되고 있다. 엔터테인먼트 산업 중에서도 다수의 관객과 공연자 간의 실시간 커뮤니케이션이 중요한 공연 분야 또한 온라인을 통한 온택트 서비스로 빠르게 전환되고 있지만, 기존 PC 나 스마트폰과 같은 전통적 2D 화면으로는 사용자들에게 충분한 현장감 및 몰입감을 제공하기 힘들다. 이를 극복하기 위하여 해외 유명 가수들 및 국내 K-Pop 가수들은 가상현실, 증강현실 (Augmented Reality, AR) 및 홀로그래피

(Holography) 기술들에 기반하여 몰입형 미디어를 활용한 공연을 새롭게 시도하고 있다.

몰입형 미디어는 일반적으로 전통적인 2D 비디오와는 다르게 사용자와의 상호작용을 통하여 해당 콘텐츠의 현장감 및 몰입감을 다양한 감각의 정보를 통하여 사실적으로 제공한다. 이러한 몰입형 미디어 기술은 학계 및 산업계로부터 엔터테인먼트 시장을 이끌어 갈 차세대 성장동력으로 주목받고 있으며, 국제 비디오 부호화 단체인 MPEG (Moving Picture Expert Group)은 몰입형 미디어 분야 전반에 걸친 핵심기술들을 MPEG-I (Immersive) 프로젝트를 통하여 2016 년부터 표준화 진행 중에 있다. 하지만 HMD 를 활용한 VR 서비스를 이용할 때 발생하는 VR 멀미[1], 몰입형 미디어의 매우 높은 데이터 처리

요구량 및 부호화/복호화 연산 복잡도 등의 이유로 기대만큼 몰입형 미디어 서비스의 시장은 확대되지 못하고 있다. 또한, 현재 서비스되고 있는 대부분의 몰입형 미디어 서비스의 경우, 촬영된 360 도 VR 영상을 단순히 HMD 로 재생하는 방식의 서비스로 국한되어 있으며, HMD 를 사용하지 않는 몰입형 VR 공연의 경우에는 몰입형 미디어가 공연예술의 주된 매체로써 활용되기 보다는 하나의 특색 있는 무대장치 수준에 머물러 있는 것이 현실이다. 이러한 문제를 해결하고 K-Pop 가수들 및 관련 예술공연 콘텐츠 생산자들에게 기존 PC 나 스마트폰과 같은 2D 화면이 아닌 대형 가상현실 공연 서비스라는 새로운 몰입형 미디어 Use-case 를 제시하기 위하여, 본 논문은 대형 가상현실 공연장의 특성을 고려한 360 도 VR 비디오 스트리밍 시스템을 제안한다.

본 논문이 제안하는 대형 가상현실 공연장을 위한 360 도 VR 비디오 스트리밍 시스템은 최종적으로 6DoF 시점 자유도를 지원하는 것을 목표로 개발 진행 중이며, 현재 3DoF 시점 자유도를 지원하는 프로토타입 구현이 완료된 상태이다. 본 논문에서는 연구개발 중인 스트리밍 시스템의 설계 및 구현에 대하여 설명하고, 현재 구현 완료된 프로토타입의 성능 분석을 수행하여, 제안된 시스템의 우수성을 증명한다.

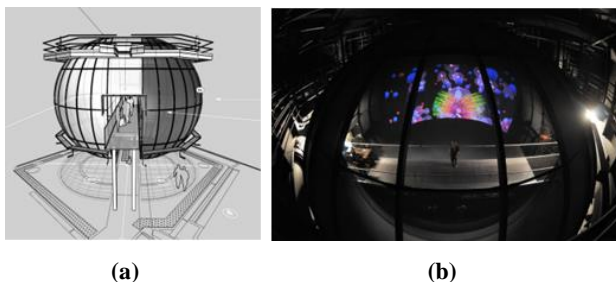


그림 1. AlloSphere : UCSB 의 대형 가상현실 공연장 (a) AlloSphere 가상 모델, (b) 위에서 내려다 본 AlloSphere

2. 대형 가상현실 공연장을 위한 360 도 VR 비디오 스트리밍 시스템 구성 및 설계

본 논문은 최신 MPEG-I 표준기술을 활용하여, 다수의 프로젝터 및 입체음향 장비로 구성되어 있는 구 형태의 대형 가상현실 공연장을 위한 360 도 VR 비디오 스트리밍 시스템을 제안한다. 본 논문이 제안하는 시스템은 한국에서 실시간으로 취득된 6DoF 360 도 VR 비디오를 해외 대형 가상현실 공연장으로 스트리밍하여 서비스하는 것을 최종목표로 하여 현재 진행 중인 연구 프로젝트이다. 현재는 미국 UCSB (University of California, Santa Barbara)의 MAT (Media Arts and Technology) 소속 JoAnn Kuchera-Morin 교수 연구팀과의 협업을 통해 대형 가상현실 공연장인 AlloSphere[2]를 활용한

3DoF 360 도 VR 비디오 스트리밍 시스템 프로토타입의 구현 및 시범방송 수행까지 완료되었다.

AlloSphere 는 그림 1 과 같이 천공된 알루미늄 스크린으로 이루어진 5m 반지름의 캡슐 형태의 소리 반향이 거의 없도록 설계된 대형 VR 시설이다. 약 2m 너비의 다리가 캡슐 내부의 중앙 부분을 연결하고 있으며, 최대 30 명의 유저들을 동시에 수용 가능하다. AlloSphere 는 시설 내부에 26 대의 프로젝터와 광섬유 네트워크로 연결된 17 대 이상의 워크스테이션, 그리고 55 대의 스피커 등으로 구성되어 있으며, 주요 워크스테이션 14 대 (GR01-GR14)와 26 대의 프로젝터가 360 도 영상을 AlloSphere 내벽에 출력하여 사용자에게 서비스를 제공한다. 14 대의 워크스테이션 중 13 대의 워크스테이션(GR02-GR14)은 각각 2 대의 프로젝터와 직접 연결되어 있으며, 전송받은 비디오 데이터를 디코딩하여 연결된 프로젝터를 통해 AlloSphere 로 출력하는 기능을 담당한다. 각 프로젝터들은 360 도 영상 내 일정 영역을 담당하여 출력하며, 프로젝터로부터 출력된 영상 간의 자연스러운 연결을 위하여 출력된 영상 간 일정부분 겹칠 수 있도록 프로젝터들이 배치되어 있다. 나머지 1 대의 워크스테이션 (GR01)은 미디어 서버로부터 전송받은 360 도 비디오 비트스트림을 13 대의 워크스테이션으로 전송하는 역할을 수행한다. 그 외에도 AlloSphere 는 컨트롤러 및 터치스크린 등 다양한 인터랙션 디바이스를 통하여 전방위 입체영상, 고해상도의 3D 공간 음향, 그리고 자연스러운 인터랙션을 통하여 고수준의 몰입도를 다수의 유저들에게 제공할 수 있다.

본 논문이 제안하는 360 도 비디오 스트리밍 시스템은 Phase 1: 대형 가상현실 공연장을 위한 3DoF 360 도 VR 스트리밍 시스템 프로토타입 개발, Phase 2: 대형 가상현실 공연장을 위한 3DoF VR 스트리밍 시스템 개발, 그리고 최종적으로는 Phase 3: 대형 가상현실 공연장을 위한 6DoF VR 스트리밍 시스템 개발을 목표로 연구 중에 있으며, 현재는 Phase 1 까지 연구개발이 완료되었다. Phase 1 에서 구현된 스트리밍 시스템 프로토타입은 최신 비디오 압축 표준 기술인 VVC (Versatile Video Coding)[3]에 포함되어 있는 서브픽처 (Subpicture)에 기반한 Viewport-dependent 스트리밍 기법을 활용한다. 서브픽처는 영상을 비트스트림 수준에서 추출 및 병합이 가능한 다수의 영역으로 분할하는 기능을 지원하는 비디오 코딩 툴이다.

구현된 스트리밍 시스템 프로토타입은 다음과 같은 방식으로 동작한다. 1) 다수의 서브픽처로 분할되어 있는 360 도 비디오 비트스트림을 미디어 서버에 위치시키고, 2) 미디어 서버는 저장된 360 도 비디오 비트스트림을 미국 UCSB 에 위치한 AlloSphere 내부 워크스테이션 GR01 으로 전송한다. 3) GR01 은 전송받은 비디오 비트스트림으로부터 각 프로젝터가 출력하는

360 도 영상 내 영역들을 고려하여 각 워크스테이션 GR02-GR14 가 영상 렌더링하는 데 필요한 서버픽처를 추출한다. 그 후, 13 개의 비트스트림을 재구성하여 각 워크스테이션 GR02-GR14 로 전송한다. 4) 2 개의 프로젝터들과 직접 연결되어 있는 워크스테이션 GR02-GR14 는 전송받은 비트스트림을 디코딩 및 렌더링하여 연결된 프로젝터에 출력한다. 그림 2 는 이러한 스트리밍 시스템 프로토타입의 시스템 아키텍처를 보여준다. 해당 시스템 프로토타입은 워크스테이션 GR02 부터 GR14 까지의 13 개 워크스테이션에 360 도 비디오 데이터 전부를 전송하는 것이 아니라, 각 워크스테이션이 렌더링하도록 정의되어 있는 360 도 영상 내 영역을 포함한 서버픽처들만을 전송한다. 이를 통하여 중복된 비디오 데이터의 전송을 최소화하고, 각 워크스테이션에서의 비디오 디코딩 및 렌더링 속도를 향상시킨다.

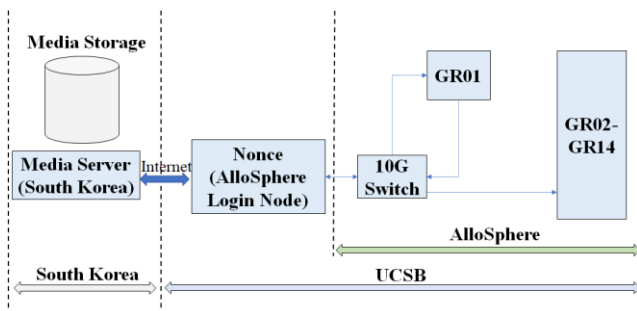


그림 2. 프로토타입 시스템 아키텍처

3. 시스템 프로토타입 구현 및 성능분석

제안된 대형 가상현실 공연장을 위한 360 도 VR 비디오 스트리밍 시스템 프로토타입은 미국 UCSB 의 AlloSphere 에서의 시범 방송을 위하여 AlloSphere 시설의 하드웨어 및 소프트웨어 환경에 호환되도록 구현되었다. 구현된 시스템 프로토타입은 대한민국에 위치한 미디어 서버에서 동작하는 *ServerApp*. 미국 UCSB 에 위치한 워크스테이션 GR01 에서 동작하는 *SubpicDistributorApp*. 그리고 26 개 프로젝터와 연결된 13 개 워크스테이션 GR02-GR14 각각에서 동작하는 *VVC-PlayerApp-AlloSphere*. 이렇게 3 종류의 소프트웨어로 구성된다. *ServerApp* 은 *SubpicDistributorApp* 과 서버-클라이언트 구조로 통신하며, *SubpicDistributorApp* 이 클라이언트로서 접속하면, 다수의 서버픽처로 분할되어 인코딩 된 8K 초고해상도 3DoF 360 도 VR 비디오 비트스트림을 *SubpicDistributorApp* 에게 전송한다. *SubpicDistributorApp* 은 전송받은 비디오 비트스트림을 파싱하여 분할된 각 서버픽처를 비트스트림 수준에서 추출하며, 추출된 서버픽처들을 AlloSphere 에서 동작하는 13 개 워크스테이션 GR02-GR14 를 위한 13 개 비트스트림으로 재구성하여, 각 워크스테이션에서 동작하는 *VVC-PlayerApp-AlloSphere* 로 전송한다. 이 때, 재구성되는

서버픽처들은 각 13 개 워크스테이션의 필요 서버픽처만으로 구성되어, 최적화된 전송효율을 달성할 수 있도록 구현되었다. *VVC-PlayerApp-AlloSphere* 는 전송받은 비디오 비트스트림의 디코딩, 픽셀 데이터의 YUV to RGBA 변환, 서버픽처 위치 재조정 작업 수행 후 프로젝터로 영상 렌더링하는 기능을 수행한다. 예를 들어, GR12 의 경우 프로젝터 #3, #4 와 직접 연결되어 있고, 프로젝터 #3, #4 가 360 도 VR 공간 내에서 렌더링을 담당하는 영역이 서버픽처 0, 1, 2, 8, 9 에 포함되어 있다면, *SubpicDistributorApp* 은 GR12 에서 동작하는 *VVC-PlayerApp-AlloSphere* 로 영상 전체를 스트리밍 하는 것이 아니라, 서버픽처 0, 1, 2, 8, 9 로만 구성된 비트스트림을 생성하여 이를 전송한다. 이러한 방식을 통하여 360 도 VR 영상 전체가 동시에 렌더링되는 대형 가상현실 공연장의 특성을 고려하여, 서버픽처 기반 Viewport-dependent 스트리밍 기술을 스트리밍 시스템에 적용하였다.

앞에서 설명한 3 개의 소프트웨어 모두 C++ 프로그래밍 언어로 작성되었으며, 리눅스 환경에서 동작하는 소프트웨어다. *SubpicDistributorApp* 은 VVC 표준의 오픈소스 참조 소프트웨어인 VTM (VVC Test Model)를 기반으로 작성되었으며, *VVC-PlayerApp-AlloSphere* 는 Fraunhofer HHI 에서 배포하는 VVDeC 을 기반으로 작성되었다. VVDeC 은 VTM 소스코드를 최적화한 오픈소스 VVC 디코더이며, 4k 및 8k 와 같은 초고해상도 영상의 실시간 디코딩을 지원한다.

본 논문은 구현된 스트리밍 시스템 프로토타입의 성능평가와 시범방송을 위하여 VVC 표준 CTC (Common Test Condition)에 포함되어 있는 *Trolley* 를 테스트 시퀀스로 사용하였다. 8192x4096 의 해상도를 갖는 ERP (Equirectangular) 360 도 영상인 *Trolley* 테스트 시퀀스는 Viewport-dependent 스트리밍 기법의 적용 위하여, 8x8, 총 64 개의 서버픽처로 분할되어 VVC 인코더로 코딩되었다.

표 1 은 13 개 워크스테이션 GR02-GR14 에게 각 워크스테이션에 연결된 프로젝터의 Viewport 를 고려하지 않고, 360 도 VR 영상 전체를 전송하는 기존 방식인 Viewport-independent 스트리밍 시스템과 최신 비디오 압축 표준인 VVC 의 서버픽처를 활용한 Viewport-dependent 스트리밍 기법을 적용하여 구현된 스트리밍 시스템 프로토타입의 성능을 비교한 표이다. 성능 비교 결과, 기존에 대형 가상현실 공연장에서 쓰이는 Viewport-independent 스트리밍 기술에 비하여, 구현된 서버픽처 기반 Viewport-dependent 스트리밍 기술을 적용하였을 때, 약 80%의 비트율 감소, 약 543%의 영상 디코딩 속도 향상을 확인할 수 있다. 각 워크스테이션마다 비트율 감소 정도나 디코딩 속도의 향상 정도가 다른 것은 각 워크스테이션에 연결된 프로젝터 Viewport 에 포함되는 서버픽처

개수와 각 서브픽처 내부 영상 데이터의 복잡도가 다르기 때문이다.

본 논문은 구현된 프로토타입의 성능검증 뿐만 아니라 미국 UCSB 에 위치한 AlloSphere 에 직접 구현된 프로토타입을 적용하여 시범방송을 수행하였다. 시범방송을 통하여 실제 환경에서도 구현된 프로토타입이 원활히 동작함을 확인하였고, Phase 2, 3 로 해당 연구를 더 고도화 할 수 있는 연구적 기반을 마련하였다. 그림 3 은 AlloSphere 에서의 시범방송 중, 동작하는 스트리밍 시스템 프로토타입의 모습을 보여준다.

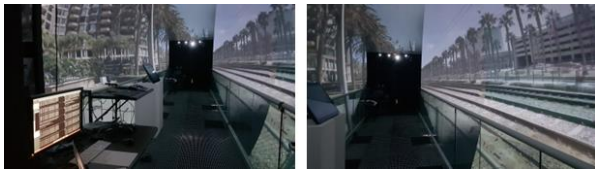


그림 3. AlloSphere 에서의 시범방송

4. 결론

본 논문은 일반적으로 사용되는 2D 디스플레이 또는 HMD 기반 VR 서비스에서 탈피하여, 대형 가상현실 공연장을 위한 360 도 VR 비디오 스트리밍 시스템을 제안한다. 제안된 시스템은 Phase 1, 2, 3 의 연구개발 단계를 밟아 6DoF 시점 자유도를 지원하는 360 도 VR 비디오 스트리밍 시스템을 개발하는 것을 최종목표로 하고 있으며, 현재는 Phase 1: 대형 가상현실 공연장을 위한 3DoF 360 도 VR 비디오 스트리밍 시스템 프로토타입의 개발까지 완료하였다.

본 논문에서 구현된 스트리밍 시스템 프로토타입의 성능을 분석한 결과, 기존 방식에 비하여 약 80%의 비트를 감소, 약 543%의 디코딩 속도 향상 효과를 확인하였다. 또한, 단순 구현 및 성능평가에서 그치지 않고, 실제 미국 UCSB 에 위치한 대형 가상현실 공연장 AlloSphere 에서의 시범방송을 통하여, 향후 Phase 2, 3 을 위한 연구적 기반을 마련하였다.

Acknowledgement

이 논문은 2022 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2020-0-00920-003,(세부 2)중대형 공간용 초고해상도 비정형 플렌옵틱 영상 저장/압축/전송 기술 개발).

참 고 문 헌 (References)

- [1] Ryu, Yeongil, and Eun-Seok Ryu, "Overview of Motion-to-Photon Latency Reduction for Mitigating VR Sickness," KSI Transactions on Internet and Information Systems (TIIS) vol. 15, no. 7, pp. 2531-2546, Jul. 2021.
- [2] Amatriain, Xavier, et al. "The allosphere: Immersive multimedia for scientific discovery and artistic exploration," IEEE MultiMedia, vol. 16, no. 02, pp. 64-75, Apr. 2009.
- [3] Bross, Benjamin, et al. "Overview of the versatile video coding (VVC) standard and its applications." IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 10, pp. 3736-3764, Oct. 2021.

표 1. 기존 스트리밍 방식과 구현된 3DoF 360도 VR 비디오 스트리밍 시스템 간의 성능비교

Graphic renderer	Conventional (Viewport-independent)		Proposed (Subpicture-based viewport-dependent)		Bitrate reduction (%)	Decoding speed enhancement (%)
	Bitrate (kbps)	Decoding speed (fps)	Bitrate (kbps)	Decoding speed (fps)		
GR02	26,638.14	13.45	8,379.78	43.54	68.54	323.72
GR03			12,819.24	38.91	51.88	289.29
GR04			3,121.89	45.26	88.28	336.51
GR05			7,159.99	79.05	73.12	587.73
GR06			7,131.92	79.49	73.23	591.00
GR07			5,397.77	89.18	79.74	663.05
GR08			4,754.86	137.68	82.15	1023.64
GR09			3,887.87	79.57	85.40	591.60
GR10			2,639.61	124.90	90.09	928.62
GR11			4,671.68	74.68	82.46	555.24
GR12			2,324.59	65.52	91.27	487.14
GR13			3,466.86	45.41	86.99	337.62
GR14			2,901.32	47.92	89.11	356.28
Average					80.17	543.96