

특집논문 (Special Paper)

방송공학회논문지 제29권 제4호, 2024년 7월 (JBE Vol.29, No.4, July 2024)

<https://doi.org/10.5909/JBE.2024.29.4.425>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

실사 공간 볼류메트릭 영상 취득 및 부호화 시스템

양이삭^{a)}, 류영일^{b)}, 박준형^{a)}, 최재열^{c)}, 정종범^{b)}, 김장현^{b)}, 류은석^{a)†}

Real-life Spatial Volumetric Video Acquisition and Encoding System

Isaac Yang^{a)}, Yeongil Ryu^{b)}, JunHyeong Park^{c)}, Jaeyeol Choi^{c)}, Jong-Beom Jeong^{b)},
Jang Hyun Kim^{b)}, and Eun-Seok Ryu^{a)†}

요약

본 논문은 고화질의 3차원 가상 현실 (virtual reality, VR) 몰입형 영상을 취득하기 위한 실사 공간의 볼류메트릭 비디오 취득 및 부호화 시스템을 제안한다. 제안된 시스템은 시그널링을 통해 동작하는 여러 카메라 중간 서버들과 이에 연결된 여러 카메라들을 활용하여 취득 대상이 되는 공간의 실사 영상을 취득하여 6 자유도 (six degrees of freedom, 6DoF) 몰입형 영상 제작을 위한 RGB raw 비디오로 저장한다. 이후 저장된 다시점 영상의 색상 보정을 통해 색상 균형을 맞춘 뒤 카메라 파라미터 및 깊이 맵 정보를 추정하여 렌더링을 통해 볼류메트릭 영상을 생성할 수 있으며, 전처리된 볼류메트릭 영상은 몰입형 영상 표현 표준인 MPEG immersive video (MIV), implicit neural visual representation (INVR) 등에 사용될 수 있다. 본 논문에서 제안한 시스템을 기반으로 취득된 영상을 사용하여 MIV와 INVR 표준 소프트웨어 기반 성능평가를 진행하였고, 카메라 및 취득 오브젝트 배치 조절을 통해 3차원 렌더링 시 주관적 화질 향상을 확인하였다.

Abstract

This paper proposes a volumetric video acquisition and encoding system for capturing high-definition 3D virtual reality (VR) immersive images in real-world spaces. The proposed system utilizes multiple camera intermediate servers and connected cameras, operating through signaling, to capture real-world images of the target space. These images are stored as RGB raw video for creating six degrees of freedom (6DoF) immersive content. After storing the multi-view images, color correction is applied to balance the colors. Camera parameters and depth map information are then estimated, allowing for volumetric video generation through rendering. The pre-processed volumetric video can be used in MPEG immersive video (MIV), implicit neural visual representation (INVR), and other immersive video representation standards. Using the images captured by the proposed system, performance evaluations based on MIV and INVR standard software were conducted, demonstrating improved subjective quality in 3D rendering through adjustments in camera and object placement.

Keyword : Virtual reality, MIV, 6DoF, Spatial computing, Volumetric capturing

Copyright © 2024 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

1. 서론

최근 VR과 공간 컴퓨팅 (spatial computing) 기술의 발전으로 가상 공간에서 자유롭게 실감적으로 감상할 수 있는 몰입형 비디오 및 볼류메트릭 비디오 콘텐츠에 대한 수요가 증가하고 있다. 대용량의 영상 데이터를 효율적으로 처리할 수 있는 영상 처리 기술과 가상 현실 공간을 구현하는 컴퓨터 그래픽 기술의 발전으로 실제 현실 공간에서 취득된 실사 영상을 가상 현실 공간에서 감상할 수 있는 볼류메트릭 비디오 콘텐츠로 제작하는 것이 가능하다. 이러한 볼류메트릭 비디오 시장은 2020년에 14억 달러 규모에 달하였고, 2025년까지 58억 달러 규모까지 성장하여 복합 연평균 성장률 32.8%를 기록할 것으로 예상되고 있다^[1]. 특히 스포츠 부문에서 3D 콘텐츠를 채택하는 경우가 증가하기 시작하면서 스포츠 경기뿐만 아니라 이를 홍보하기 위한 이벤트 및 창의적인 광고에서도 볼류메트릭 비디오 기술에 대한 요구가 증가하고 있으며, 이에 더해 콘서트와 같은 공연 문화, 라이브 콘텐츠, 커머스, 전시 등 실감미디어 콘텐츠를 수요로 하는 시장까지 볼류메트릭 비디오 시장의 저변이 확대되고 있다.

이러한 볼류메트릭 비디오는 2016년 10월 시작한 MPEG-I (ISO/IEC 23090 Coded Representation of Immersive Media) 프로젝트에서 제공하는 사용자 시점 자유도에 따라

three degrees of freedom (3DoF)와 6DoF로 구분하였고, 사용자 위치에서의 X, Y, Z 축을 중심으로 한 머리 회전만을 지원하는 3DoF와 달리 머리 회전에 더해 사용자가 시점 위치를 옮길 수 있어 더욱 몰입감 있는 시청 경험을 제공할 수 있는 6DoF를 구현하고 서비스하기 위해 MPEG-I는 6DoF 몰입형 비디오 및 볼류메트릭 비디오에 대하여 기술 표준화를 진행하고 있다^[2]. 또한 MPEG은 최근 주목을 받고 있는 암시적 신경망 영상 표현 기법인 INVR을 통한 영상 표현 방식의 표준화를 위해 2022년 7월부터 INVR ad-hoc group (AhG)을 구성하여 3D INVR을 통한 볼류메트릭 영상 표현에 관한 연구 및 표준화 활동을 진행하고 있다^[3]. 한편, 2018년부터 MPEG 3D graphics and haptics (3DGH) 그룹의 주도로 포인트 클라우드 압축 기술인 video-based point cloud compression (V-PCC)와 함께 동적 메쉬 압축 기술인 video-based dynamic mesh coding (V-DMC) 표준화가 진행되고 있다^[4].

이러한 볼류메트릭 영상 콘텐츠를 제작하기 위해서는 실제 현실 공간에서 취득하고자 하는 공간 및 대상에 대하여 동기화된 병렬적 다시점 실사 영상을 취득해야 한다. 하지만 다시점 영상 취득 시스템의 구현 난이도가 높고, 카메라 개수가 증가함에 따라 전체 시스템의 크기와 하드웨어 성능 요구치가 증가하며, 동기화된 영상 취득 과정의 구현을 위해 중간 카메라 서버의 연동과 함께 카메라 서버와 이에 연결된 여러 대의 카메라들 간 동기적인 제어가 필요하다^[5]. 이때 카메라 서버에 연결된 여러 대의 카메라들은 취득하고자 하는 공간의 빛의 종류 및 밝기에 따라 조도, 감도, 노출 시간, 광원 (light source), 화이트 밸런스 (white balance) 등 설정값을 수정해야 한다. 따라서 다시점 실사 영상 취득 시스템은 동기화된 병렬적인 실사 영상 취득 기능과 함께 취득 공간의 환경에 따라 카메라의 촬영 설정들을 모든 카메라에 동일하게 적용해야 한다.

본 논문에서는 실사 공간을 취득 대상으로 하는 볼류메트릭 영상 제작을 위한 다시점 영상 취득 및 볼류메트릭 영상 부호화 시스템을 제안한다. 제안된 시스템은 4대의 카메라 중간 서버 PC와 총 20대의 카메라로 영상을 취득하는 다시점 카메라 시스템으로 다시점 영상을 취득하고, UDP 멀티캐스팅 및 내장된 소프트웨어 동기화 기능을 통해 각 카메라에서 프레임 간 동기화 오차를 최소화하여 각각의

a) 성균관대학교 실감미디어공학과(Department of Immersive Media Engineering, Sungkyunkwan University)

b) 성균관대학교 컴퓨터교육학과(Department of Computer Science Education, Sungkyunkwan University)

c) 성균관대학교 인공지능융합학과(Department of Applied Artificial Intelligence, Sungkyunkwan University)

‡ Corresponding Author : 류은석(Eun-Seok Ryu)

E-mail: esryu@skku.edu

Tel: +82-2-760-0677

ORCID: <https://orcid.org/0000-0003-4894-6105>

※ This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the Graduate School of Metaverse Convergence support program(IITP-2024-RS-2023-00254129) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation)

※ This work was also supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.RS-2022-00167169, Development of Moving Robot-based Immersive Video Acquisition and Processing System in Metaverse)

· Manuscript May 31, 2024; Revised June 28, 2024; Accepted July 1, 2024.

영상으로 저장한다. 제안된 시스템은 성균관대학교 교내에 위치한 135.66㎡ 면적의 대형 강의실을 취득 공간으로 하여 다시점 영상 취득 및 볼류메트릭 영상 부호화 과정을 수행한다. 또한 제안된 시스템을 통해 취득되는 다시점 실사 영상은 MPEG-I의 6DoF 몰입형 영상 표준인 MIV의 참조 소프트웨어인 test model for immersive video (TMIV)와 3차원 장면 재구성을 위한 인공지능 기반 모델 중 하나인 3D gaussian splatting (3DGS) 기반 기술을 통해 볼류메트릭 영상 콘텐츠로 제작이 가능하다. 이후 취득된 볼류메트릭 영상은 부호화 시스템을 통해 렌더링 되고, 그 결과를 통해 제안된 시스템 성능을 검증 후 후속 연구의 주제 및 방향에 대하여 논의한다⁶⁾.

II. 배경 연구

1. 다시점 영상 취득에서의 카메라 파라미터 추정

다시점 카메라로부터 취득된 영상들을 6DoF VR 몰입형 영상으로 제작하기 위해서 여러 단계와 그로부터 생성되는 결과물이 필요하고, 그 중 하나가 카메라의 위치 및 회전 각도 등을 정의하는 값 등의 카메라 파라미터이다. 현실 공간에서 배치된 카메라들로부터 취득된 2차원 이미지 혹은 영상만을 사용하여 3차원 재구성을 수행하기 위해 2차원 영상으로부터 3차원 정보를 추출하는 기술인 structure from motion (SfM)이 사용된다⁷⁾. SfM에서는 입력 이미지 혹은 영상으로부터 특징점을 추출하는데, 이때 이미지의 크기와 회전과 같은 불변하는 특징점을 추출하는 알고리즘인 scale-invariant-feature transform (SIFT) 알고리즘과 함

께 FAST 알고리즘을 개선해 특징점 검출 속도가 빠르고 회전이 적용되었거나 스케일이 다른 이미지에서도 특징점을 검출할 수 있는 oriented FAST and rotated BRIEF (ORB) 알고리즘 등을 사용한다⁸⁾. 즉 SfM 알고리즘의 주요 과정은 이미지의 변형을 고려하여 이미지의 고유한 특징점을 검출해 이미지 간 특징점들을 매칭하는 것이다.

이러한 SfM 알고리즘을 기반으로 하여 COLMAP 라이브리리가 제안되었고, 이를 통해 multiview stereo (MVS) 알고리즘으로 복수의 2차원 영상으로부터 3차원 영상을 재구성할 수 있다⁹⁾. COLMAP은 SfM 알고리즘을 통해 다시점 이미지들을 입력으로 받아 카메라 내부 및 외부 파라미터와 이미지들의 특징점으로 구성된 포인트 클라우드 데이터를 생성하고, 그 결과를 MVS 알고리즘의 입력값으로 전달하여 3D 모델로 재구성한다. 본 논문에서 제안하는 실사 공간의 볼류메트릭 영상 생성 시스템에서는 공간을 둘러싼 카메라들의 내부 및 외부 파라미터를 추정하기 위하여 COLMAP을 사용한다.

2. 다시점 영상 취득에서의 깊이 맵 추정

3차원 재구성을 목표로 하는 대상을 동기적으로 취득한 다시점 영상이 있을 때, 이를 3차원 재구성하기 위해서 다시점 영상으로부터 카메라 파라미터와 함께 깊이 맵 (depth map) 정보를 획득해야 한다. 영상에서 취득된 물체가 카메라의 시점으로부터 떨어져 있는 거리에 대한 정보를 깊이 맵 정보라고 하며, 이러한 깊이 맵 정보를 추정하기 위해서 light detection and ranging (LiDAR) 센서를 사용할 수 있다. LiDAR 센서가 장착된 카메라에서 LiDAR 센서가 빛을

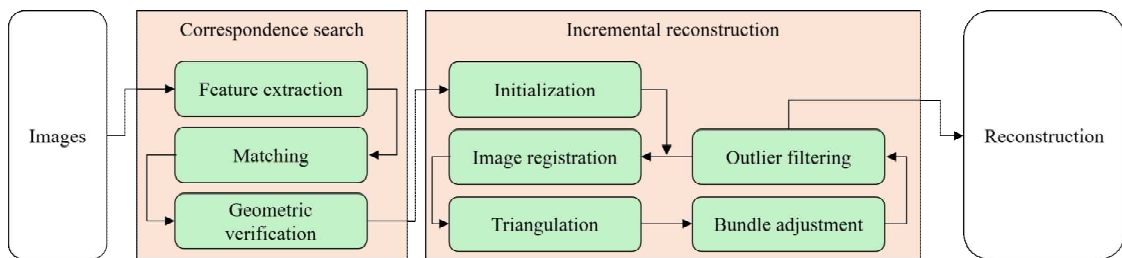


그림 1. Incremental SfM 알고리즘
 Fig. 1. Incremental structure-from-motion algorithm

방출하면 취득 대상이 되는 물체에 도달한 빛이 반사되어 다시 센서로 도달할 때까지 소요되는 시간에 빛의 속도를 곱하여 물체의 각 지점까지의 거리를 알아내는 **time of flight (ToF)**라 하고, 이 ToF 기법을 사용하는 LiDAR 센서를 통해 깊이 맵 정보를 다시점 영상 취득 시점에서 수행할 수 있다^[10]. LiDAR 센서를 통한 깊이 맵 정보 획득은 형태 인식이 가능하고 정밀도가 높다는 장점이 있다. 하지만, 다시점 영상 취득 시 여러 LiDAR 센서를 사용하기 때문에 신호 간섭이 발생하여 그로 인해 취득되는 깊이 맵 품질이 감소하고, 야외 실사 환경의 경우 기상 상황에 민감하며, LiDAR 센서의 가격이 높다는 단점이 있다.

LiDAR 센서를 사용하지 않고 취득된 다시점 영상으로부터 깊이 맵 정보를 추정할 수 있다. 인간의 시각 시스템이

양쪽 눈의 시차를 통해 3차원 정보를 인식하는 원리를 이용하여, 영상 내 대응점의 위치가 인접한 시점 사이에서 변화하는 정도를 통해 영상 내 물체의 거리를 계산하여 각 시점에서의 깊이 맵 정보를 추정할 수 있다. 취득된 다시점 영상을 입력으로 하면 SfM 알고리즘을 통해 추정한 카메라 파라미터로 각 카메라 시점들의 상대적인 위치를 알 수 있다. 위치가 주어진 각 시점에서 물체가 떨어져 있는 거리 정보가 포함된 깊이 맵 정보와 함께 각 시점에서 취득된 영상인 텍스처 정보가 합쳐져 3차원 영상으로 렌더링 될 수 있다. 즉, 3차원 볼류메트릭 영상으로의 렌더링 과정을 수행하기 위해 텍스처 정보의 역할을 수행할 각 카메라 시점에서 취득된 원본 영상과 이를 통해 각 시점의 카메라 파라미터 정보, 그리고 각 시점의 깊이 맵 정보가 입력으로 제공되어야 한다.

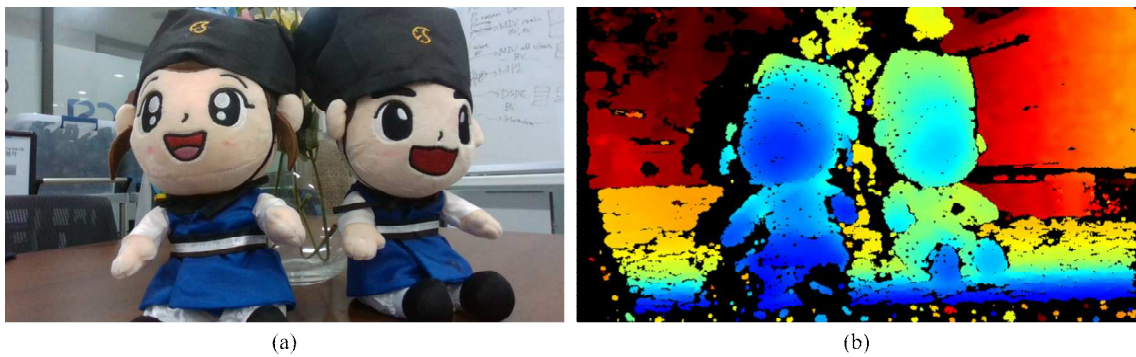


그림 2. LiDAR 센서가 장착된 Intel RealSense D435 카메라로 취득한 RGB 이미지와 깊이 맵 정보 (a) RGB 이미지 (b) 깊이 맵 정보
 Fig. 2. RGB image and depth map information acquired by Intel RealSense D435 camera module with LiDAR sensor (a) RGB image (b) depth map information

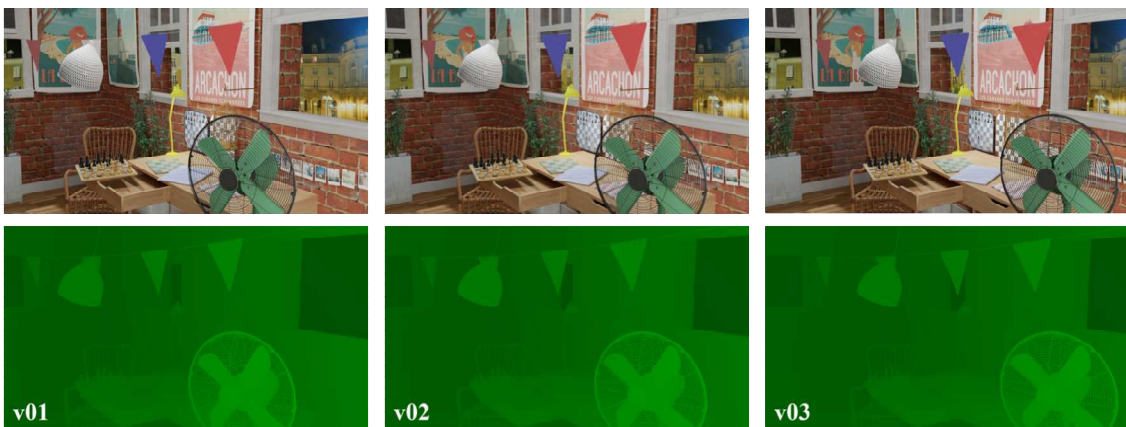


그림 3. 다시점 영상의 텍스처와 깊이 맵 정보
 Fig. 3. Texture and depth map of multiview video

3. MPEG Immersive Video (MIV, ISO/IEC 23090-12)

MPEG-I 표준 중 하나인 MIV 표준은 실제 또는 가상의 3D 장면을 다수의 실제 또는 가상의 카메라로 취득한 영상을 압축할 수 있는 기능을 제공한다^[11]. 360도 영상과 달리 시청자가 콘텐츠의 시점 위치와 방향을 제어할 수 있는 몰입형 미디어 영상은 제한된 동작 범위 내에서 시점 위치와 6DoF 렌더링을 제공하는데, MIV는 VR 및 증강 현실 (augmented reality), 혼합 현실 (mixed reality) 등의 다양한 use-case에 활용될 수 있다. 이에 MPEG은 제한된 범위 내에서 화면 전환 이동에 따른 운동 시차를 제공하는 것을 목표로 하여 exploration experiment (EE)를 진행하였으며, 이후 2019년 3월 MPEG 126차 회의에서는 MPEG-I에서 몰입형 영상 표준화 기술을 지원하는 테스트 모델인 TMIV를 정의하였다^[12].

그림 4는 TMIV 렌더러에서의 렌더링 프로세스 과정을 나타낸다. MIV 표준은 high-efficiency video coding (HEVC), versatile video coding (VVC) 등과 같이 일반적인 2D 영상 코덱을 통해 사용자의 움직임에 따른 끊김 없는 운동 시차의 제공과 고품질 뷰포트 (viewport) 영상의 재현을 목표로 정

의 및 개발되었다^[13]. TMIV 인코더는 MIV 부호화 과정에서 핵심적인 전처리 기능을 수행하며, 일반적인 영상 인코더가 부호화 가능하도록 다수의 뷰포트 영상들로부터 소수의 아틀라스 영상을 생성한다. 입력되는 다시점 영상으로부터 3D 공간 상에 대응되는 픽셀들의 중복성을 줄이는 프루닝 (pruning)과 패치 패킹 (patch packing) 과정을 수행한다. TMIV 디코더는 인코딩 과정을 통해 생성된 MIV 표준 메타 데이터를 파싱하여 프루닝된 원본 시점의 영상을 복원하고, TMIV 렌더러가 디코더를 통해 분리된 패치들로부터 사용자가 바라보는 방향에 해당하는 중간 시점 영상을 합성하여 시청 경로에 해당하는 정보인 포즈 트레이스 (pose trace)에 따라 연속적인 뷰포트 영상을 렌더링한다^[14].

III. 실사 공간 볼류메트릭 영상 취득 및 부호화 시스템 파이프라인

그림 5는 제안된 시스템의 전체 파이프라인을 나타낸다. 제안된 시스템의 선행 조건으로는 다시점 카메라 배치 및 카메라 취득 파라미터 설정이 있다. 다시점 카메라 배치 단

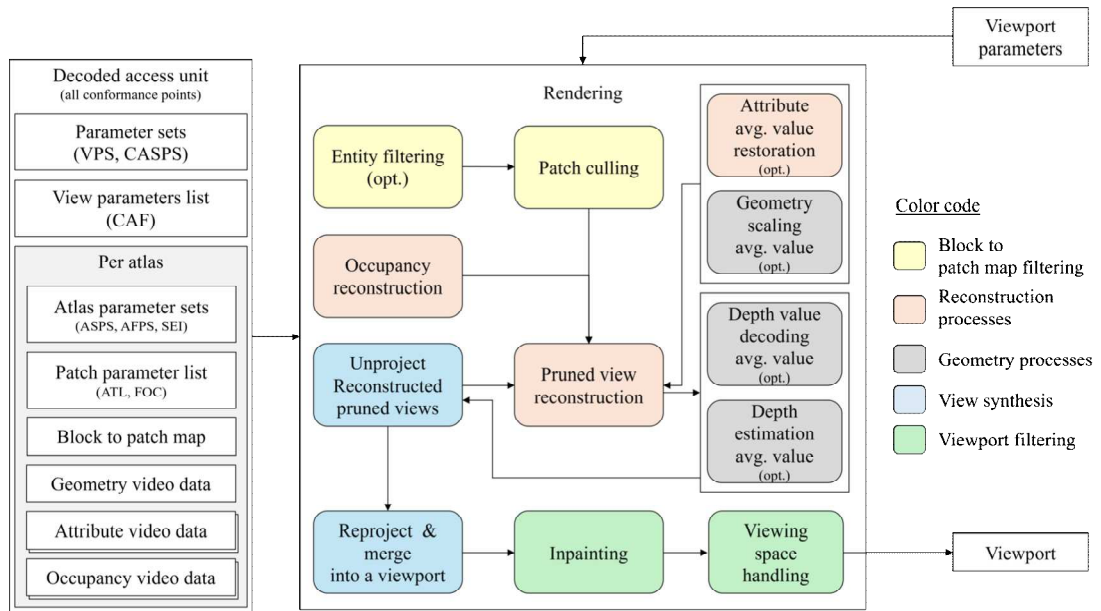


그림 4. TMIV 렌더링 프로세스 순서도
 Fig. 4. Process Flow for the TMIV Renderer

계에서는 취득하고자 하는 대상 공간의 주변으로 카메라를 원형 혹은 정사각형 형식과 같이 순환되는 형식의 구조로 카메라를 배치한다. 카메라 파라미터 설정 단계에서는 다시점 영상 취득 시 카메라에 설정되어야 할 파라미터인 카메라 취득 파라미터를 설정한다. 이후 동일한 취득 설정이 적용된 카메라들과 이들을 동기적으로 제어하는 카메라 중간 서버를 통해 다시점 영상 취득 시스템을 실행하여 취득 대상 공간의 다시점 실사 영상을 취득하고, 색상 보정 과정을 통해 각 시점의 영상들의 색상을 균일하게 설정한다. 색상 보정된 다시점 영상들을 통해 취득 시 사용된 카메라의 상대적인 위치 좌표 추정을 위해 카메라 파라미터 추정 단계를 거치고, 색상 보정된 다시점 영상과 함께 카메라 파라미터를 입력으로 하여 깊이 맵 정보를 추정한다. 이로써 출력된 깊이 맵 정보 영상은 색상 보정된 다시점 영상을 텍스처 맵 정보로 하여 카메라 파라미터 정보와 함께 렌더링 소프트웨어의 입력으로 주어질 수 있으며, 이를 통해 볼류메트릭 영상을 생성할 수 있다^{[15][16]}.

IV. 다시점 영상 취득 단계

본 절에서는 제안된 시스템의 파이프라인 중 카메라 취

득 관련 설정 및 동기적 카메라 취득 모듈에 대하여 설명한다. 다시점 카메라의 배치가 완료된 후 다시점 영상을 취득하기 위해서는 카메라 별로 상이한 설정들을 통일할 필요가 있다. 이는 같은 모델의 카메라와 렌즈를 사용하더라도 카메라 별로 색상 영역에서 차이가 발생할 수 있으며, 또한 취득 공간에서의 조명 환경이 다를 수 있기 때문이다. 따라서 다시점 영상 취득 이전에 취득 시 시스템을 이루는 카메라들의 취득 관련 설정들을 최대한 균일하게 통일하는 전처리 단계가 선행되어야 한다. 본 논문에서는 카메라 및 렌즈의 영상 취득 관련 설정 및 파라미터들을 카메라 취득 파라미터라 칭한다.

이러한 카메라 취득 파라미터에는 시스템을 구성하는 각각의 카메라들의 렌즈 초점과 조리개 설정, 그리고 동기적 취득 시스템의 입력 값으로 설정할 수 있는 총 취득 프레임 수 (total frame number), 네트워크 인터페이스 IP 주소 (network interface IP address), 노출 시간 (exposure time), 목표 프레임 레이트 (resulting frame rate), 광원 (light source), 화이트 밸런스 (white balance), 모자이크 제거 필터 (de-mosaicking filter), 노이즈 제거 필터 파라미터 (noise reduction parameter), 디테일 향상 파라미터 (sharpness enhancement parameter)가 있다. 이는 카메라 모델 및 배치 구조에 따라 상이한 색상 영역과 조명 환경을 최대한 균일하게

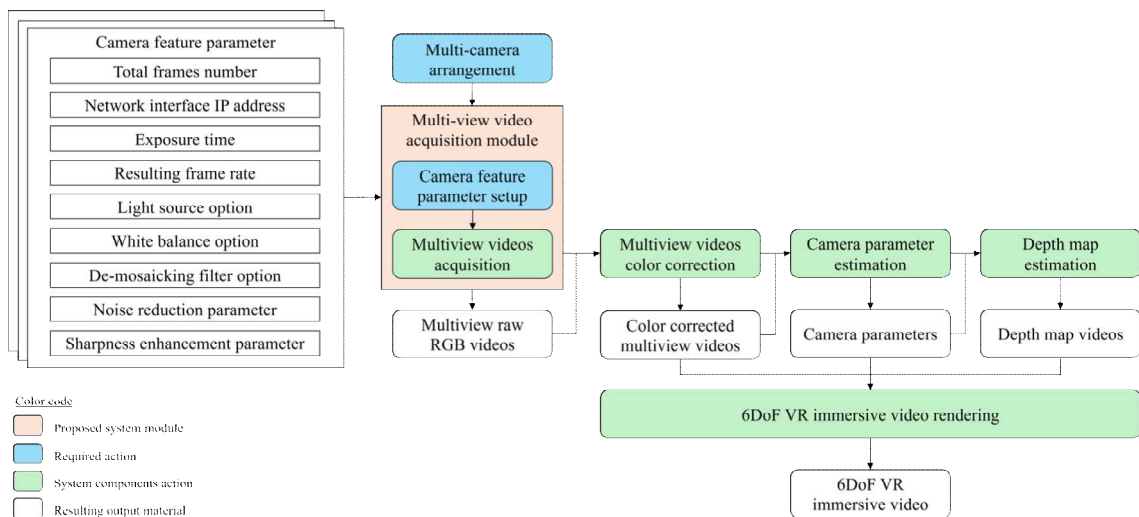


그림 5. 실사 공간 볼류메트릭 영상 취득 및 부호화 파이프라인
 Fig. 5. Real-life spatial volumetric video acquisition and decoding pipeline

통일시킨 후 후처리 단계에서의 색상 보정 과정을 보다 수월하게 하기 위함이다. 이를 위해 실내 공간의 다시점 영상 취득 시 조명 환경으로 인해 발생하는 영상 내 색수차와 모션 블러 현상을 최대한 감소시키면서 실제 공간 내 밝기와 유사한 밝기의 영상을 취득하기 위해 다시점 영상 취득 시스템 실행 시 노출 시간, 라이트 소스, 화이트 밸런스 등의 설정을 `cfg` 파일 형식의 입력으로 받아 카메라 중간 서버에 연결된 카메라들에게 일괄적으로 적용한다.

그림 6은 제안된 시스템의 모듈 중 동기적으로 다시점 실사 영상을 취득하는 다시점 영상 취득 모듈의 다이어그램을 나타낸다. 시스템을 시작하면 UDP 소켓 및 멀티캐스트 `time to live (TTL)` 등과 같은 UDP 멀티캐스트 통신을 위한 값들을 설정하고 소켓 구조체 초기화 및 네트워크 인터페이스를 설정한다. 이후 멀티캐스트 그룹을 생성하고 가입한 뒤, 카메라 시스템 실행 전 입력으로 설정한 카메라 취득 설정들을 카메라 중간 서버에 연결된 카메라들에게 일괄적으로 적용하면서 카메라의 동기적 제어를 위한 카메라 소프트웨어 트리거를 등록한다. 모든 설정이 완료되면 코드를 통해 카메라 모듈을 열고 모든 소프트웨어 트리거가 실행되는 즉시 영상을 취득할 수 있도록 카메라를 대기 상태로 설정한다.

Sender 및 receiver PC 역할을 하는 카메라 중간 서버들의 UDP 멀티캐스트 통신이 성공적으로 연결되면 영상 취득을 위한 쓰레드 (thread)를 생성한다. 제안된 시스템의 다시점 영상 취득 모듈은 멀티 쓰레드를 통해 카메라 중간 서버들을 병렬적으로 제어하고, 이를 통해 동기적이고 병렬적인 다시점 영상 취득 시스템 제어를 가능하게 한다. Sender는 동일한 멀티캐스트 그룹에 가입된 receiver들에게 영상 취득 시작을 알리는 버퍼 메시지를 전송하고, 메시지를 전송한 sender와 수신한 receiver들은 동기적으로 영상 취득을 위한 멀티 쓰레드 모듈을 반복적으로 수행하게 된다. 메인 쓰레드는 카메라들의 소프트웨어 트리거를 실행하는데, sender의 메인 쓰레드는 매 프레임마다 소프트웨어 트리거 실행 메시지 전송 역할 또한 수행하고, receiver의 메인 쓰레드는 sender가 전송한 메시지를 수신하는 즉시 소프트웨어 트리거를 실행하는 역할 또한 수행한다. 카메라 소프트웨어 트리거 실행 후 이미지 취득 쓰레드에서 카메라가 취득한 이미지 데이터를 이미지 버퍼에 저장하고, 영상 저장 쓰레드에서 버퍼에 저장된 이미지 데이터를 RGB 포맷의 원본 영상 파일에 저장한다. 이 과정은 사용자가 취득하고자 하는 총 취득 프레임 수만큼 반복하면서 매 프레임마다 수행되고, 로그 파일 작성을 위해 매 프레임마다 이미지 취득 쓰레드에서 이미지 취득과 동시에 카메라의 타임

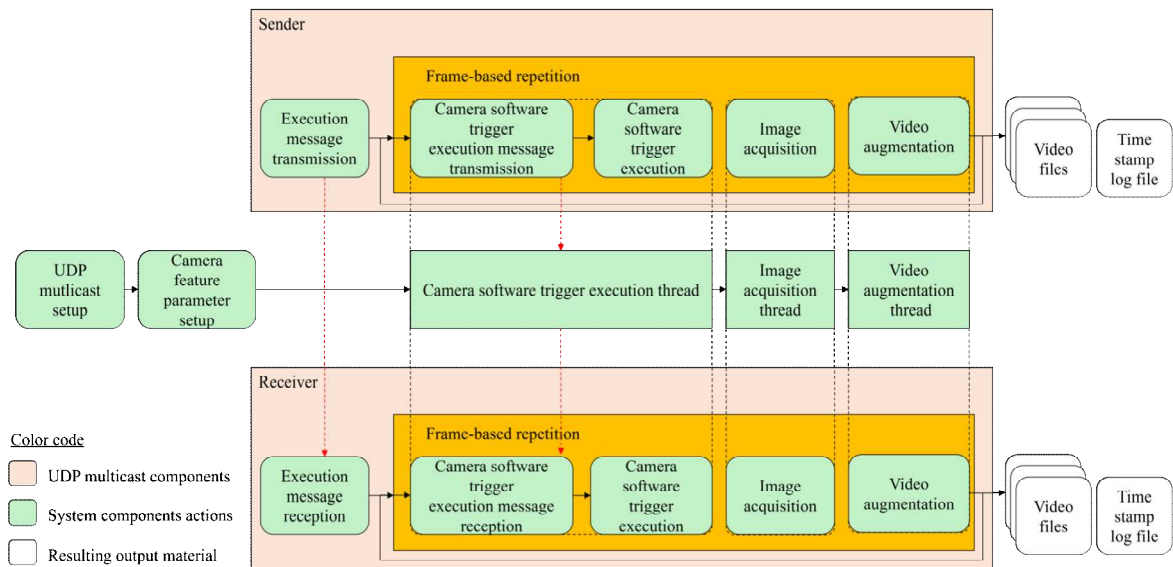


그림 6. 다시점 영상 취득 모듈 다이어그램
 Fig. 6. Multiview video acquisition module diagram

스태มป์를 저장해 로그 파일로 저장한다. 마지막 프레임까지 취득을 완료하게 되면 비디오 저장 쓰레드가 종료되어 RGB 포맷 원본 영상 파일 저장이 종료되고, 취득 시 측정된 카메라 별 취득 시점에서의 타임 스탬프, 프레임 간 타임 스탬프 시간차, 프레임 별 인접 시점의 카메라들 간 실행 시간 오차를 계산하여 로그 파일에 작성한 뒤 취득 모듈을 종료한다.

V. 볼류메트릭 영상 생성 단계

1. 색상 보정

본 절은 제안된 시스템의 색상 보정 단계를 설명하며, 그림 7은 제안된 시스템의 색상 보정 단계를 나타낸다. 다시점 영상은 각 시점을 이루는 여러 카메라들로부터 취득된 영상으로, 같은 모델을 사용하여 다시점 영상을 취득하더라도 각 카메라마다 고유한 색상 영역과 취득 시점에서의 조명 환경 등 영상 취득에 영향을 주는 요소들이 다를 수 있다. 이러한 카메라 고유 설정의 불균형은 균일한 다시점 영상 취득에 악영향을 끼치며, 이는 볼류메트릭 영상 합성의 성능을 저하시킬 수 있다. 색상 보정은 이러한 카메라 고유 설정의 불균형을 고려하여 다시점 영상에서의 색상 특성의 일관성을 증가시키는 단계로, 서로 다른 조명 조건으로 인해 색상이 일관되지 않은 시점의 영상에 대한 깊이 맵의 품질을 향상시키는 것을 목표로 한다. 이러한 목표를 달성하기 위해 ISO/IEC JTC1 SC29 WG4 MPEG video coding 표준화 그룹에서 채택한 색상 보정 기술은 다시점 영상에서 모든 입력 시점을 독립적으로 처리하여 시간에 따른 색상을 통일해 시간적 색상 일관성을 구현하고, 각 입력 시점의 색상 특성을 중앙 참조 시점과 맞추어 시점 간

색상 일관성을 구현하였다¹⁷⁾.

상용 소프트웨어인 Davinci Resolve는 look up table (LUT) 기능을 통해 영상의 색상을 보정하는데, 이는 색상 그레이드와 톤 매핑을 빠르고 효율적으로 수행하기 위한 기능으로 각 색상 값과 대응하는 출력 색상 값을 매핑하기 위한 표인 LUT를 통해 입력 영상을 필터링한다. 이러한 방법은 높은 신뢰도와 빠른 수행 속도를 보이며, 취득 공간의 조명 및 카메라 배치 환경이 동일한 경우 각 카메라 시점에 대하여 LUT을 한 번 생성하면 이를 재사용할 수 있어 색상 보정 단계의 자동화가 가능하다는 장점이 있다. 이러한 장점으로 인해 제안된 시스템은 Davinci Resolve의 LUT 기능과 함께 다시점 영상 취득 시 LUT 색상을 위한 컬러 차트로서 Datacolor 사의 Spyder Checkr 컬러 차트를 활용하여 색상 보정을 수행한다. 컬러 차트는 데이터베이스 형식으로 저장된 컬러 차트의 원본 색상과 영상 내에서 취득된 컬러 차트의 색상을 비교하여 취득 영상의 색상을 보정할 수 있게 한다. 다시점 영상 취득 시 컬러 차트를 포함한 이미지를 취득하고 이를 바탕으로 Davinci Resolve에서 색상 보정을 위한 LUT을 생성하게 되는데, 이때 1D LUT과 3D LUT을 생성할 수 있다. 1D LUT은 휘도와 톤만을 보정할 수 있는 반면 3D LUT은 X, Y, Z 축의 공간 좌표로 LUT을 표현해 영상의 색상과 채도 값을 보정할 수 있다. 제안된 시스템은 각 입력 시점에 대하여 3D LUT을 생성하며, 이후 영상 인코딩 및 디코딩을 위한 오픈 소스 소프트웨어인 fast forward MPEG (FFmpeg)을 통해 각 입력 시점에 대하여 생성된 3D LUT을 모든 입력 시점 영상에 적용한다.

2. 카메라 파라미터 추정

그림 8은 제안된 시스템의 카메라 파라미터 추정 과정을

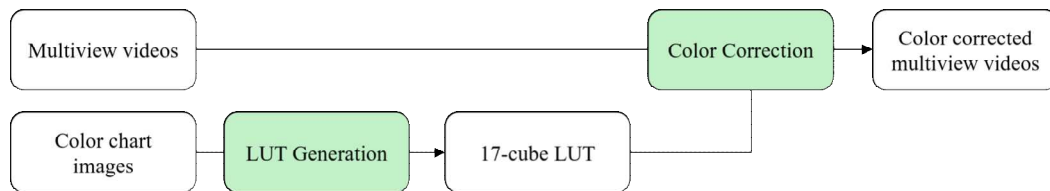


그림 7. 제안된 시스템의 look up table (LUT) 기반 색상 보정 다이어그램
Fig. 7. LUT based color correction diagram of proposed system

나타낸다. 제안된 시스템은 MVS 알고리즘을 위한 범용 라이브러리인 COLMAP을 활용하여 카메라 파라미터를 추정하였다. 카메라의 초점 거리, 주점, 비대칭계수를 내부 파라미터, 회전 및 평행 이동 변환 등 카메라 좌표계와 월드 좌표계 사이의 변환 관계를 설명하는 파라미터를 카메라 외부 파라미터라고 한다. 제안된 시스템은 COLMAP을 통해 카메라 내부 및 외부 파라미터를 추정한다. 색상 보정된 다시점 영상의 첫 번째 프레임들을 저장한 각 시점들의 이미지를 카메라 파라미터 추정을 위한 입력으로 제공하며, COLMAP을 통해 각 시점에 대하여 특징점을 추출한 뒤 이를 바탕으로 카메라 내부 및 외부 파라미터를 추정한다. 카메라 파라미터 추정 시 추출한 특징점을 바탕으로 주어진 다시점 이미지에서의 카메라 파라미터 추정을 위한 최적의 알고리즘을 COLMAP에서 선택할 수 있으며, 선택된 알고리즘을 바탕으로 카메라 파라미터 추정이 완료되면 COLMAP 좌표계로 표현된 데이터베이스 파일로 내보낸 뒤 이를 MPEG-I 파트2 표준인 omnidirectional media format (OMAF) 좌표계로 변환하여 깊이 맵 추정을 위한 입력으로 제공할 수 있다.

3. 깊이 맵 추정

제안된 시스템은 깊이 맵 정보 추정을 위해 MPEG의 깊이 맵 추정을 위한 소프트웨어인 immersive video depth estimation (IVDE)를 사용한다^[18]. 양안 시차를 활용한 깊이 맵 정보 추정은 다시점 영상이 입력으로 주어졌을 때 인접

시점 간 대응점의 거리를 바탕으로 각 점에서의 거리인 깊이를 계산하는 과정인 스테레오 정합 (stereo matching) 과정을 수행하는데, 이 스테레오 정합은 지역 정합 (local matching)과 전역 정합 (global matching) 방식으로 구분된다. 지역 정합은 일정 크기의 윈도우 내 픽셀 정보만을 사용하여 스테레오 정합을 수행하는 방법으로, 연산 복잡도가 낮아 단순하고 빠른 연산이 가능하지만 전역 정합 방식에 비해 정확도가 낮은 단점이 있다. 반면 전역 정합 방식은 영상의 전체적인 정보를 사용해 스테레오 정합을 수행하는 방법으로 메모리 요구량과 연산 복잡도가 높지만 지역 정합 방식에 비해 높은 성능을 기대할 수 있다. IVDE는 전역 정합 방식으로 스테레오 정합을 수행하며, 최소 그래프 컷 (graph-cut) 기법을 통해 영상 분할을 수행한다. 이를 통해 에너지 함수를 최적화하고 연산 복잡도를 낮추면서 입력 시점 간 시간적 일관성을 보존할 수 있다^[19]. IVDE는 다시점 영상과 함께 다시점 영상의 파일 및 디렉토리 정보, 카메라 파라미터 정보, 깊이 맵 추정 파라미터 정보를 저장한 설정 파일을 입력으로 제공받아야 한다. 이에 따라 제안된 시스템은 색상 보정된 다시점 영상과 함께 OMAF 좌표계로 변환된 카메라 파라미터 정보 및 디렉토리 정보, 깊이 맵 추정 파라미터 정보를 저장한 *cfg* 파일을 입력으로 하여 IVDE를 통해 깊이 맵 정보를 추정한다.

4. 렌더링

볼류메트릭 영상 생성을 위한 방법 및 소프트웨어에는

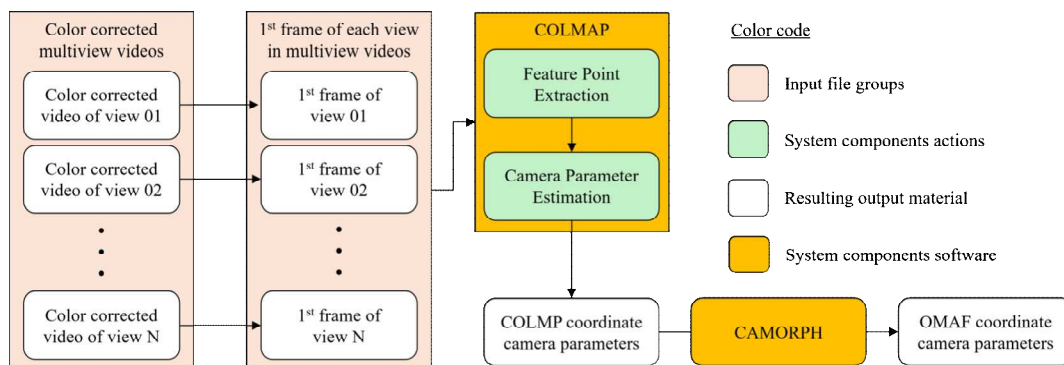


그림 8. 제안된 시스템의 카메라 파라미터 추정 다이어그램
 Fig. 8. Camera parameters estimation diagram of proposed system

여러 가지 종류가 있으나, 이들은 공통적으로 다시점 영상 혹은 이미지가 반드시 필요하다는 것과 함께 렌더링 과정이 필요하다는 점이 있다. 하지만 TMIV와 같이 깊이 맵 정보가 필요한 소프트웨어가 있는 반면, 인공 신경망 기반 방사형 필드 모델인 *neural radiance fields (NeRF)* 모델 혹은 가우시안 분포 표현 방식인 *3DGS*와 같이 학습된 모델을 바탕으로 렌더링을 할 수 있는 내재적 방식의 렌더링 모델 혹은 소프트웨어는 다시점 영상만을 입력으로 제공받아 볼류메트릭 영상을 생성할 수 있다^{[20][21]}.

제안된 시스템은 볼류메트릭 영상 생성을 위해 TMIV 렌더러를 사용하며, 색상 보정된 다시점 영상과 함께 COLMAP을 통해 추정된 카메라 파라미터, 카메라 파라미터를 통해 생성한 포즈 트레이스 정보, IVDE를 통해 추정된 깊이 맵 정보를 TMIV 렌더러의 입력으로 제공한다^[22]. 포즈 트레이스 정보는 TMIV 렌더러가 볼류메트릭 영상을 생성

할 때 입력 시점의 카메라 파라미터 중 시작 지점 좌표부터 종료 지점 좌표까지 이동하는 방향 및 거리를 지정하여 생성된 볼류메트릭 영상에서 카메라가 이동하는 방식을 정의한다.

VI. 실험 결과

1. 실험 조건

그림 9는 제안된 시스템을 위한 시스템 구축 환경을 나타낸다. 다시점 영상 취득을 위한 실제 공간을 주변으로 총 20대의 카메라 모듈과 4대의 카메라 서버 모듈을 설치하였으며, 5대의 카메라 모듈과 1대의 카메라 서버가 하나의 카메라 서버 그룹을 구성한다. 총 4개의 카메라 서버 그룹은 동일한 UDP

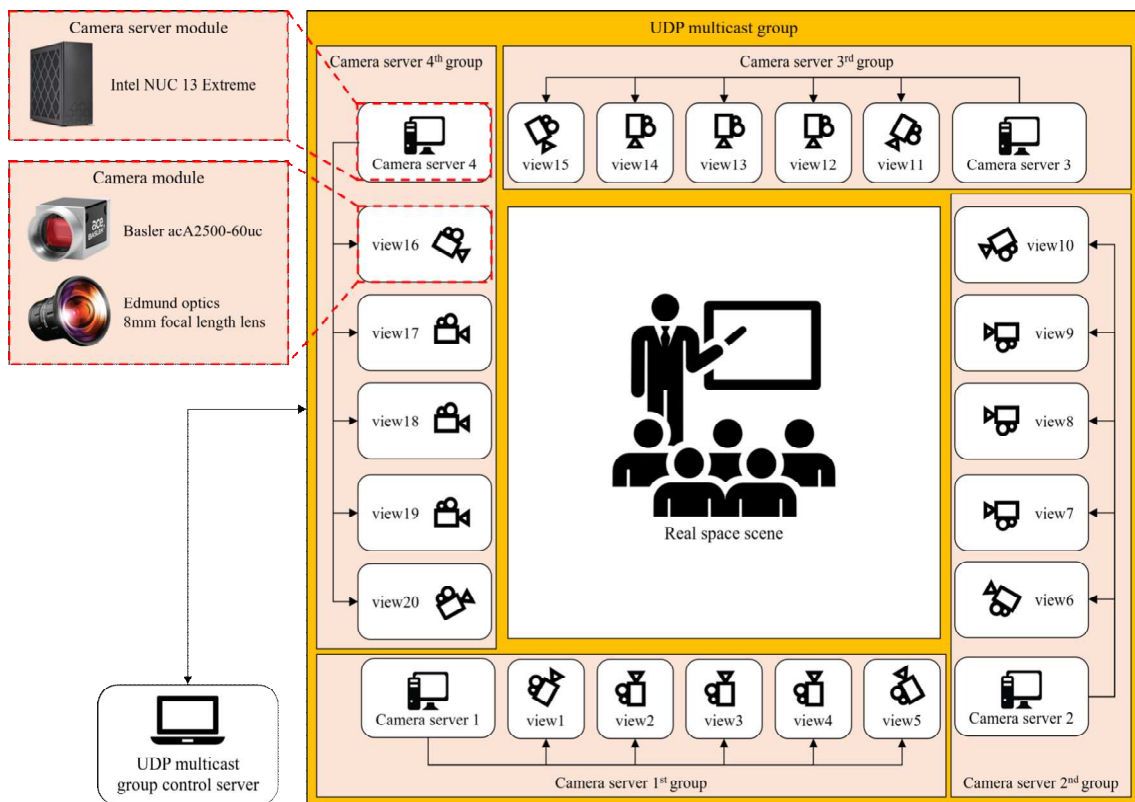


그림 9. 제안된 시스템 구축 환경
Fig. 9. System construction environment of proposed system

멀티캐스트 그룹에 연결되어 UDP 멀티캐스트 그룹 제어 서버를 통해 다시점 영상 취득 과정을 수행한다. 제안된 시스템은 카메라 모듈로서 Basler 사의 acA2500- 60uc, 렌즈로서 Edmund Optics 사의 8mm series fixed focal length lens f/1.8 86-569, 카메라 중간 서버로서 Intel NUC 13 Extreme 을 사용하였다. 픽셀 사이즈 $4.8\mu\text{m}\times 4.8\mu\text{m}$, 최대 해상도 2592×2048 , 최대 프레임 레이트 60fps, 동기적 영상 취득을 위한 소프트웨어 트리거 기능을 지원하는 제안된 시스템

하드웨어 구성을 통해 해상도 1952×1088 , 프레임 레이트 30fps, 노출 시간 50ms, 색온도 2800K, 지속적 화이트 밸런스 적용 등의 설정을 적용하여 다시점 영상을 취득하였다.

2. 다시점 영상 취득

그림 10은 제안된 시스템의 동기적 다시점 영상 취득 모듈로 취득한 강의실의 다시점 실사 영상을 나타낸다. 30fps



그림 10. 제안된 시스템의 동기적 다시점 영상 취득 모듈로 취득한 다시점 실사 영상
Fig. 10. Multiview video acquired by multiview video acquisition module of proposed system



그림 11. 제안된 시스템으로 취득한 다시점 실사 영상의 색상 보정 영상
Fig. 11. Color corrected video of multiview video acquired by proposed system



그림 12. 제안된 시스템의 색상 보정 수행 과정
 Fig. 12. Color correction process of proposed system

의 프레임 레이트로 총 100 프레임을 취득하였을 때 프레임 동기화를 통해 시간적 동기화를, 카메라 동기화를 통해 공간적 동기화를 이루었음을 확인하였다. 하지만 카메라의 고유 색상 설정 및 실내 조명 환경으로 인해 취득된 다시점 영상의 시점 간 색상 불균형을 확인할 수 있다.

3. 색상 보정

다시점 영상 취득 후 나타나는 색상의 불균형을 해결하기 위해 색상 보정 단계를 수행해야 한다. 그림 11은 전체 입력 시점에 대하여 색상 보정을 수행한 다시점 영상을 나

타내며, 다시점 영상 취득 시 발생하는 색상 불균형 문제가 색상 보정 후 해결되었음을 확인할 수 있다.

그림 12는 제안된 시스템의 색상 보정 수행 과정을 나타낸다. 색상 보정을 위해 모든 입력 시점에 대하여 컬러 차트를 포함한 이미지를 취득한 후, 취득한 컬러 차트 이미지를 이용해 Davinci Resolve에서 3D cube LUT 파일을 생성한다. FFmpeg을 통해 다시점 영상의 모든 입력 시점에 대하여 각 입력 시점 영상에 LUT를 적용하면 그림 11과 같은 색상 보정된 다시점 영상을 생성할 수 있다. 이때 Davinci Resolve에서는 17, 33, 65 point 3D cube LUT 파일을 생성할 수 있으나, FFmpeg에서는 17 point 3D cube LUT 파일만을

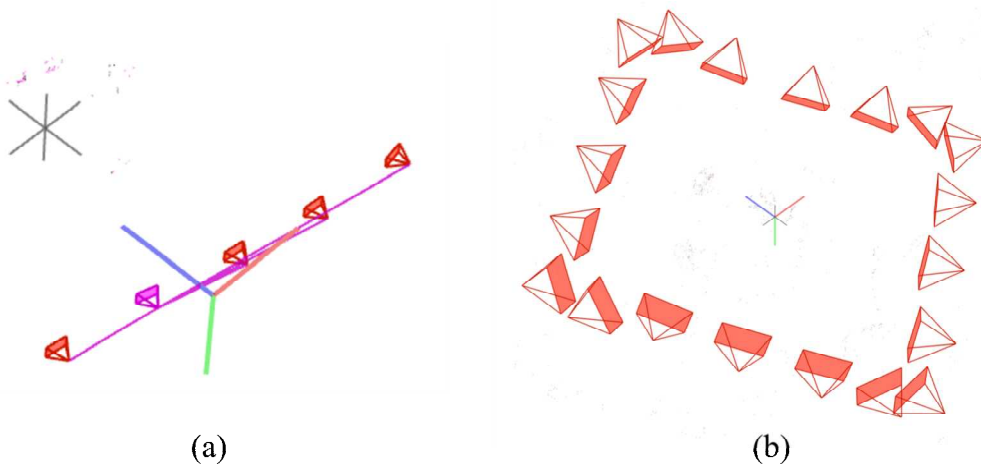


그림 13. COLMAP을 통한 카메라 파라미터 추정을 위한 카메라 배치 수정 (a) 정면을 바라보도록 배치한 카메라 시점에서의 카메라 파라미터 추정 (b) 각도 조정 후 카메라 파라미터 추정
 Fig. 13. Camera alignment modification for camera parameter estimation (a) camera parameter estimation from a camera point of view placed facing the front (b) camera parameter estimation after adjustment

지원하기 때문에 제안된 시스템은 17 point 3D cube LUT을 생성한다.

4. 카메라 파라미터 추정

제안된 시스템은 그래픽 환경 COLMAP을 사용하여 카메라 파라미터를 추정하였다. 색상 보정된 다시점 영상의 모든 입력 시점에 대하여 첫 번째 프레임만을 이미지로 저장해 COLMAP의 특징점 추출 알고리즘의 입력으로 사용한다. 특징점 추출 후 재구성 단계에서 현재 입력된 이미지와 추출된 특징점을 바탕으로 최적의 알고리즘을 발견하는 automatic reconstruction 기능을 통해 카메라 파라미터 추정을 수행할 수 있고, 수행이 완료되면 그림 13과 같이 카메라 위치 및 특징점으로 이루어진 포인트 클라우드 (point cloud)를 표시한다. 이를 데이터베이스 파일로 내보내면 COLMAP 좌표계로 표현된 카메라 파라미터 데이터베이스 파일이 생성되는데, 깊이 맵 추정 및 TMIV 렌더러 실행을 위해 OMAF 좌표계로 변환하여야 한다. 제안된 시스템은 독일 Fraunhofer HHI 연구소에서 제작한 카메라 파라미터

좌표계 변환 소프트웨어인 CAMORPH를 통해 OMAF 좌표계로의 변환을 수행하였다^[23].

제안된 시스템의 다시점 실사 영상 취득 환경과 같이 순환하는 형식의 카메라 배치는 카메라 파라미터 추정 정확도가 떨어지는데, 이는 IVDE의 깊이 맵 정보 추정의 퀄리티를 감소시키는 원인이 된다. 또한 제안된 시스템의 다시점 실사 영상 취득 환경과 같이 카메라 배치가 직각을 이루는 경우 인접 시점 간 공유하는 화면의 크기가 현저하게 떨어지기 때문에 특징점 추출에 실패하여 모든 입력 시점에 대한 카메라 파라미터 추정에 실패할 수 있고, 깊이 맵 정보 추정의 퀄리티가 대폭 감소할 수 있다. 이를 해결하기 위해 제안된 시스템은 카메라 배치 시 직각을 이루는 부분인 시점 1, 5, 6, 9, 10, 11, 15, 16, 20에서는 그림 16-(b)와 같이 45도에 가까운 각도로 카메라 각도를 수정하여 인접 시점 간 공유하는 화면의 크기를 증가시켰고, 나머지 시점 카메라는 정면을 바라보도록 하였다. 모든 시점 카메라가 정면을 바라보는 경우 모든 입력 시점에 대하여 카메라 파라미터 추정에 실패하였으나, 이와 같은 카메라 배치 및 각도 수정을 통해 카메라 파라미터 추정에 성공하였다.

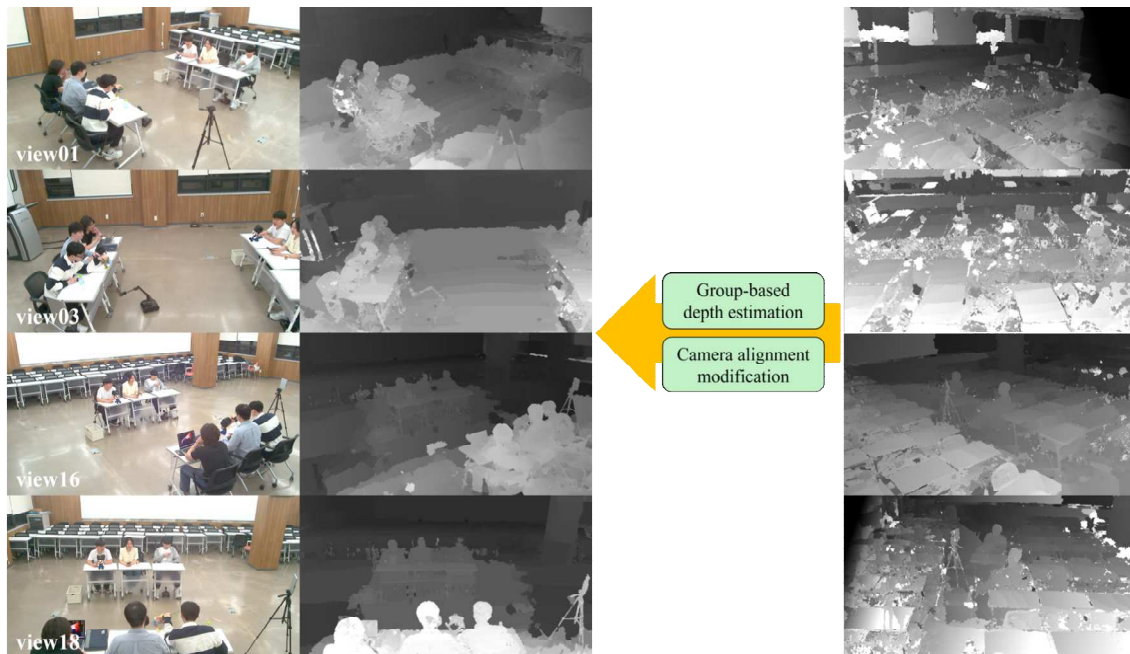


그림 14. IVDE를 통해 추정된 깊이 맵 정보
 Fig. 14. Depth map information estimated by IVDE

5. 깊이 맵 추정

그림 14는 IVDE를 통해 추정된 깊이 맵 정보를 나타낸다. 색상 보정된 다시점 영상과 함께 OMAF 좌표계로 변환된 카메라 파라미터 파일을 입력으로 사용하는 IVDE는 그림 14와 같이 입력 시점에서 물체가 가까울수록 밝게, 멀수록 어둡게 깊이 맵을 표현한다. IVDE를 통해 추정된 깊이 맵 정보는 YUV420 포맷의 영상 파일로 저장되며, 이는 이후 렌더링 단계에서 TMIV 렌더러의 입력으로 사용된다. 제안된 시스템에서의 깊이 맵 정보 생성 결과 그림 14와 같이 오브젝트 주변으로 노이즈가 발생하거나 가까운 곳에 있는 오브젝트의 깊이 정보가 어둡게 표현되는 것과 같이 깊이 정보가 반전되는 등과 같은 아티팩트가 발생함을 확인할 수 있다. 이러한 아티팩트는 모든 입력 시점에 대하여 다시점 영상을 입력으로 사용해 깊이 맵 정보를 추정하였을 때 더욱 빈번하고 확연하게 발생한다. 이를 해결하기 위

해 시스템을 구성하는 카메라들을 4개의 그룹으로 분할하여 각 그룹 별로 인접 그룹의 양 끝 시점을 공유하도록 설정한 뒤 깊이 맵 정보를 추정하였다. 그 결과 그림 14와 같이 깊이 맵 정보 주관적 품질을 향상시킬 수 있으나, 다시점 영상 취득 시 선명도가 낮아 특징점 추출의 성능이 낮아져 카메라 파라미터 추정의 정확도가 저하되면 깊이 정보가 반전되어 표현될 수 있음 또한 확인하였다.

6. 렌더링

그림 15는 제안된 시스템의 TMIV 렌더링 결과를 나타낸다. 색상 보정된 다시점 실사 영상과 함께 OMAF 좌표계로 변환된 카메라 파라미터 정보, IVDE로 생성된 깊이 맵 정보 영상, 모든 입력 시점에 대한 포즈 트레이스 정보를 입력으로 사용하여 TMIV 렌더링을 수행하면 그림 15와 같이 모든 입력 시점에 대하여 중간 시점을 합성해 볼륨메트릭



그림 15. 제안된 시스템의 TMIV 렌더링 결과 (a) 하얀 책상을 다수 배치한 환경 (b) 하얀 책상을 최소화하여 배치한 환경
 Fig. 15. TMIV rendering result of proposed system (a) environment with a lot of white desks (b) environment in which white desks are minimally arranged



그림 16. 제안된 시스템의 3DGS 렌더링 결과
 Fig. 16. 3DGS rendering result of proposed system

영상을 생성할 수 있다. 이때 그림 15-(a)와 같이 밝은 색상의 오브젝트가 다수 배치된 환경에서 IVDE와 TMIV를 거쳐 깊이 맵 정보 추정 및 렌더링 과정을 수행하면 오브젝트를 중심으로 아티팩트가 발생함을 확인할 수 있다. 이는 하얗거나 밝은 색상의 오브젝트가 영상의 색상 보정 및 특징점 추출 성능을 저하시키기 때문이며, 그림 15-(b)와 같이 이러한 요소를 최소화하여 렌더링을 수행했을 때 각 시점에서 영상 전체적인 노이즈가 감소하였음을 확인할 수 있다.

3DGS는 TMIV와 달리 깊이 맵 정보 및 카메라 파라미터 정보를 입력으로 사용하지 않고 색상 보정된 다시점 영상만을 입력으로 사용한다. 3DGS의 학습 모델을 통해 카메라 내부 및 외부 파라미터를 학습하고, 학습한 파라미터를 바탕으로 가우시안 분포를 생성해 3차원 재구성을 표현한다. 그림 16은 이러한 3DGS를 통해 3차원 재구성하였을 때 생성된 중간 시점 이미지를 나타낸다. 하지만 3DGS는 2차원 이미지를 3차원 정지 장면 (scene)으로 재구성하기 때문에 제안된 시스템과 같이 영상 기반 3차원 재구성 시스템에 적용하기에 제한이 있으나, 현재 정지 이미지가 아닌 영상의 가우시안 분포를 표현해 3차원 재구성을 하는 모델인 동적 3DGS 모델로서 4D gaussian splatting (4DGS)가 발표되어 관련 연구가 활발히 진행 중이기 때문에, 이를 활용한

다시점 실사 영상의 3차원 재구성을 기대할 수 있을 것이다.

VII. 결론

본 논문은 실제 환경에서의 대형 공간을 대상으로 다시점 실사 영상을 취득하고 이를 볼류메트릭 영상으로 부호화 및 3차원 재구성을 하는 시스템을 구현하여 시스템 성능을 확인하였다. 자연광이 충분하여 색상 보정 및 특징점 추출이 원활한 야외 환경과 달리, 그림 17-(a)에서 확인할 수 있듯 조명 환경이 서로 다르고 밝거나 어두운 색상의 오브젝트가 다수 배치된 실내 강의실 환경을 대상으로 다시점 실사 영상을 취득하였기 때문에 색상 보정 및 특징점 추출에서 시행착오를 겪었으나, 카메라 및 오브젝트 배치 조정을 통해 깊이 맵 추정과 볼류메트릭 영상 부호화 성능을 향상시킬 수 있었다. 하지만 여전히 오브젝트 주변으로 노이즈가 발생하는 등 아티팩트가 발생하는 현상을 확인하였다. 향후 연구에서는 이를 해결하기 위해 다시점 실사 영상 취득 시 해상도 조정 혹은 초점 변경 등의 방법을 통해 선명도 높은 다시점 실사 영상을 취득하고, 카메라 파라미터 및 깊이 맵 정보 추정 품질을 향상시키기 위한 추가 연구를 진행하고

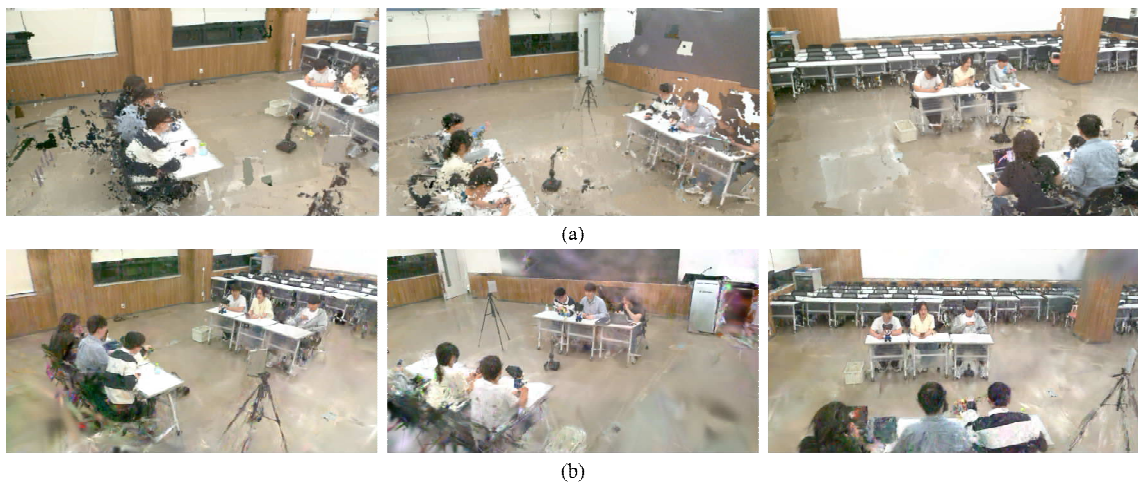


그림 17. 제안된 시스템으로 취득한 다시점 영상에 대한 볼류메트릭 영상으로의 재구성 결과 (a) TMIV 렌더러를 통한 중간 시점 합성 결과 (b) 3DGS를 통한 렌더링 결과

Fig. 17. Results of reconstruction into volumetric videos for multiview images acquired by proposed system (a) intermediate view synthesis results through TMIV renderer (b) rendering results through 3DGS

자 한다. 또한 MIV 표준 이외에도 현재 표준화가 활발히 진행되고 있는 MPEG INVR 표준 기반의 렌더링 참조 소프트웨어를 활용한 볼류메트릭 영상 부호화 실험을 통해 제안된 시스템을 보완하고 use-case를 확장하고자 한다. 현재 논문은 LED 조명이 설치된 실내 공간에서 취득한 영상에 대하여 볼류메트릭 영상 취득 및 부호화 시스템의 성능을 평가하였으나, 크기 및 조명 환경 등과 같은 환경 조건을 다양하게 구성하여 여러 취득 대상 공간에서의 시스템 성능 평가가 필요하다. 이에 따라 향후 연구에서는 취득 대상 공간의 다양한 환경 조건을 고려하여 업데이트될 볼류메트릭 영상 취득 및 부호화 시스템의 성능을 평가하고자 한다.

참 고 문 헌 (References)

- [1] Y. -H. Seo. Volumetric Natural Contents 4D Video Technology, Broadcast and Media Magazine, Vol. 26, No. 2, pp.56-66, 2021.
- [2] K. Choi, G. Bang. MPEG-I Video Standardization Trend, Broadcast and Media Magazine, Vol. 23, No. 4, pp.353-369, 2018.
- [3] G. Bang. Neural Video Representation and MPEG-INVR Standardization Trend, Broadcast and Media Magazine, Vol. 28, No. 3, pp.45-54, 2023.
- [4] J. -S. Lee. Introduction to Video-based Dynamic Mesh Coding Standardization Trend, Broadcast and media Magazine, Vol. 28, No. 3, pp.23-32, 2023.
- [5] Y. -S. Ho, K. -J. Oh. Overview of Multi-view Video Coding. 2007 14th International Workshop on Systems, Signals and Image Processing and 6th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services, pp.5-12,2007.
doi: <https://doi.org/10.1109/IWSSIP.2007.4381085>
- [6] N. Sabater, G. Boisson, B. Vandame, P. Kerbirou, F. Babon, M. Hog, R. Gendrot, T. Langlois, O. Burller, A. Schubert, V. Allie. Dataset and Pipeline for Multi-View Light-Field Video. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1743-1753, 2017.
doi: <https://doi.org/10.1109/CVPRW.2017.221>
- [7] J. L. Schonberger, J. Frahm. Structure-from-Motion Revisited. Proceedings of the IEEE conference on computer vision and pattern recognition, pp.4104-4113, 2016.
doi: <https://doi.org/10.1109/CVPR.2016.445>
- [8] R. Mur-Artal, J. M. M. Montiel, J. D. Tardos. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. IEEE Transactions on Robotics, Vol. 31, No. 5, pp. 1147, 2015.
doi: <https://doi.org/10.1109/TRO.2015.2463671>
- [9] Y. Liang, Y. Yang, X. Fan, T. Cui. Efficient and Accurate Hierarchical SfM Based on Adaptive Track Selection for Large-Scale Oblique Images. Remote Sensing, Vol. 15, No. 5, pp. 1374-1398, 2023.
doi: <https://doi.org/10.3390/rs15051374>
- [10] Y. Kang, Y. Ho. High-resolution Depth Generation using Multi-view Camera and Time-of-Flight Depth Camera. Journal of the Institute of Electronics Engineers of Korea, Vol. 48, No. 6, pp.749-755, 2011.
- [11] J. -B. Jeong, S. Lee, J. Choi, G. Lee, S. Kwak, W. -S. Cheong, B. Lee, E. -S. Ryu. Towards Group-based Adaptive Streaming for MPEG Immersive Video. Journal of Broadcast Engineering, Vol. 28, No. 2, pp.194-212, 2023.
doi: <https://doi.org/10.5909/JBE.2023.28.2.194>
- [12] J. -B. Jeong, S. Lee, E. -S. Ryu, DATRA-MIV: Decoder-Adaptive Tiling and Rate Allocation for MPEG Immersive Video, ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), Vol. 20, No. 7, pp. 1-22, 2024.
doi: <https://doi.org/10.1145/3648371>
- [13] J. -B. Jeong, S. Lee, E. -S. Ryu. Implementing Geometry Packing for MPEG Immersive Video. Journal of Broadcast Engineering, Vol. 27, No. 6, pp. 861-871, 2022.
doi: <https://doi.org/10.5909/JBE.2022.27.6.861>
- [14] J. M. Boyce, R. Dore, A. Dziembowski, J. Fleureau, J. Jung, B. Kroon, B. Salaheih, V. K. M. Vadakital, L. Yu. MPEG Immersive Video Coding Standard. Proceedings of the IEEE, Vol. 109, pp.1521-1536, 2021.
doi: <https://doi.org/10.1109/JPROC.2021.3062590>
- [15] B. Syzdelko, D. Mieloch, A. Dziembowski, J. Stankowski, D. Kloska, J. Y. Jeong, G. Lee, ECPC - versatile multicamera system calibration framework for immersive video applications. SoftwareX, Vol. 26, pp. 1-7, 2024.
doi: <https://doi.org/10.1016/j.softx.2024.101669>
- [16] J. Choi, J. -B. Jeong, S. Lee, E. -S. Ryu. Overview of the Volumetric Video Capturing System for Immersive Media. International Conference on Information and Communication Technology Convergence (ICTC), pp.574-577, 2022.
doi: <https://doi.org/10.1109/ICTC55196.2022.9953020>
- [17] A. Dziembowski, D. Mieloch, S. Rozek, M. Domanski. Color Correction for Immersive Video Applications. IEEE Access Access, Vol. 9, pp.75626-75640, 2021.
doi: <https://doi.org/10.1109/ACCESS.2021.3081870>
- [18] J. Y. Jeong, M. U. Kakli, G. Lee, J. Seo. [MIV] Usage of Optical Flow for Improving IVDE's Temporal Enhancement Functionality. Standard ISO/IEC JTC1/SC29/WG4, MPEG/m58196, 2021.
- [19] V. Kolmogorov, R. Zabih. Multi-camera Scene Reconstruction via Graph Cuts. Computer Vision - ECCV 2002 - 7th European Conference on Computer Vision, Proceedings, pp.82-96, 2002.
doi: https://doi.org/10.1007/3-540-47977-5_6
- [20] J. -H. Park, J. -B. Jeong, J. Choi, E. -S. Ryu, Comparison of Neural Radiance Fields-based models for 3D Scene Reconstruction. The Korean Institute of Broadcast and Media Engineers (KIBME) Summer Conference, pp. 93-96, 2023.
- [21] J. Choi, J. -H. Park, J. -B. Jeong, E. -S. Ryu. [INVR]EE2.1-Related: 3D Gaussian Splatting for Visual Representation. Standard ISO/IEC JTC1/SC29/WG4, MPEG/m66420, 2024.
- [22] Y. Ryu, M. Jang, E. -S. Ryu. Multiview Camera System-based 6DoF Immersive Media Content Acquisition and Compression. The Korean Institute of Broadcast and Media Engineers (KIBME) Summer

Conference, pp. 290-290, 2023.
[23] B. Brand, M. Batz, J. Keinert. CAMORPH: A Toolbox for Conversion
Between Camera Parameter Conventions. The International Archives

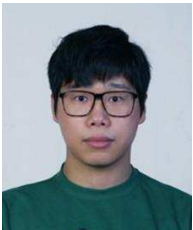
of the Photogrammetry, Remote Sensing and Spatial Information
Sciences, Vol. 48, pp. 29-36, 2022.
doi: <https://doi.org/10.5194/isprs-archives-XLVIII-2-W1-2022-29-2022>

저 자 소 개



양 이 삭

- 2016년 3월 ~ 2023년 8월 : 성균관대학교 컴퓨터공학과 학사
- 2023년 9월 ~ 현재 : 성균관대학교 실감미디어공학과 석박통합과정
- ORCID : <https://orcid.org/0009-0007-3176-6972>
- 주관심분야 : 멀티미디어 통신 및 시스템, 실감미디어, 볼류메트릭 비디오, 로보틱스



류 영 일

- 2016년 2월 : 가천대학교 컴퓨터공학과 학사
- 2018년 2월 : 가천대학교 컴퓨터공학과 석사
- 2021년 3월 ~ 현재 : 성균관대학교 컴퓨터교육학과 박사과정
- ORCID : <https://orcid.org/0000-0002-9210-6245>
- 주관심분야 : 멀티미디어 통신 및 시스템, 비디오 압축 표준



박 준 형

- 2018년 3월 ~ 2024년 2월 : 성균관대학교 영상학과 학사
- 2024년 3월 ~ 현재 : 성균관대학교 실감미디어공학과 석사과정
- 2023년 7월 ~ 2023년 8월 : 한국전자통신연구원 학생연구원
- ORCID : <https://orcid.org/0009-0000-6524-1559>
- 주관심분야 : 실감미디어, 인공지능, 그래픽스, 멀티미디어 통신 및 시스템



최 재 열

- 2018년 3월 ~ 2024년 2월 : 성균관대학교 컴퓨터교육학과 학사
- 2024년 3월 ~ 현재 : 성균관대학교 인공지능융합학과 석사과정
- 2023년 1월 ~ 2023년 2월 : 한국전자통신연구원 학생연구원
- ORCID : <https://orcid.org/0009-0009-2923-1252>
- 주관심분야 : 실감미디어, 인공지능, 그래픽스, 멀티미디어 통신 및 시스템

저 자 소 개



정 종 범

- 2018년 8월 : 가천대학교 컴퓨터공학과 학사
- 2018년 9월 ~ 2019년 8월 : 가천대학교 컴퓨터공학과 석사과정
- 2019년 9월 ~ 현재 : 성균관대학교 컴퓨터교육학과 석박통합과정
- 2020년 1월 ~ 2020년 3월 : University of California, Santa Barbara 방문연구원
- 2021년 8월 ~ 2022년 1월 : Purdue University 방문연구원
- 2022년 9월 ~ 현재 : 성균관대학교 글로벌융합학부 강사
- ORCID : <https://orcid.org/0000-0002-7356-5753>
- 주관심분야 : 멀티미디어 통신 및 시스템, 비디오 압축 표준, MPEG immersive video, video-based dynamic mesh coding



김 장 현

- 2011년 8월 : 연세대학교 전기전자공학과 박사
- 2016년 1월 ~ 2018년 12월 : 연세대학교 연구교수
- 2018년 1월 ~ 2021년 12월 : ㈜퍼시픽나인 대표
- 2021년 1월 ~ 현재 : 성균관대학교 사범대학 산학교수
- ORCID : <https://orcid.org/0000-0002-4829-5958>
- 주관심분야 : 인공지능, 인공지능교육, 로보틱스, 홀로그래픽 시스템



류 은 석

- 1999년 8월 : 고려대학교 컴퓨터학과 학사
- 2001년 8월 : 고려대학교 컴퓨터학과 석사
- 2008년 2월 : 고려대학교 컴퓨터학과 박사
- 2008년 3월 ~ 2008년 8월 : 고려대학교 연구교수
- 2008년 9월 ~ 2010년 12월 : 조지아공대 박사후과정
- 2011년 1월 ~ 2014년 2월 : InterDigital Labs Staff Engineer
- 2014년 3월 ~ 2015년 2월 : 삼성전자 수석연구원/파트장
- 2015년 3월 ~ 2019년 8월 : 가천대학교 컴퓨터공학과 조교수
- 2019년 9월 ~ 현재 : 성균관대학교 컴퓨터교육과 부교수
- 2023년 9월 ~ 현재 : 성균관대학교 실감미디어공학과 학과장
- ORCID : <https://orcid.org/0000-0003-4894-6105>
- 주관심분야 : 멀티미디어 통신 및 시스템, 비디오 코딩 및 국제 표준, HMD/VR 응용분야