

ISOBMFF 대체 트랙 기반의 프로그래머블 G-PCC 미디어 캡슐화

정승혁*, 김동호**, 유동호***

ISOBMFF Alternative Track-based Encapsulation for Programmable G-PCC Media

Seunghyeok Jeong*, Dong Ho Kim**, and Dongho You***

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. RS-2023-00278925).
이 논문은 2026년도 교육부 및 서울특별시의 재원으로 서울엔지니어센터의 지원을 받아 수행된 지역성장
인재양성체계(ANCHOR)의 결과임(2026-ANCHOR-01-014-04).

요 약

본 논문은 다중 속성 변체를 포함하는 프로그래머블 G-PCC 미디어의 저장·전송 효율 개선을 목적으로, ISOBMFF 대체 트랙 구조 기반의 캡슐화 방식을 제안한다. 기존 독립 트랙 방식에서는 시각적 속성만 다른 다수의 변체가 기하 데이터를 각각 보유하여, 변체 수만큼 대용량 기하 데이터가 물리적으로 중복 적재되는 문제가 있었다. 이를 위해 SAM2 기반 전처리로 기하가 동일한 다중 변체를 생성하고, 트랙 참조 및 대체 그룹 메커니즘으로 단일 기하 트랙을 복수의 속성 트랙이 공유하는 파일 계층 구조를 구현하였다. 실험 결과, 제안 구조는 화질 손실 없이 저장 용량을 절감하였으며 그 이점은 변체 수가 증가할수록 커져 10개 기준 약 15% 절감되었고, 런타임에서 속성 트랙만 교체해 실시간 전환이 가능함을 확인하였다.

Abstract

This paper improves the storage and transmission efficiency of programmable G-PCC media with multiple attribute variations. Conventionally, each variation is stored as an independent track carrying its own copy of the geometry, so the geometry is duplicated as many times as there are variations. The proposed method generates geometry-identical variations through SAM 2 preprocessing and uses the ISOBMFF Track Reference and Alternate Group mechanisms so that a single geometry track is shared by multiple attribute tracks. Experiments show that the structure reduces storage without quality loss, with savings growing as variations increase and reaching about 15% for ten variations, and that real-time switching is achieved by replacing only the attribute track at runtime.

Keywords

programmable media, ISOBMFF, G-PCC, point cloud

* 국립공주대학교 컴퓨터교육과 박사과정
- ORCID: <https://orcid.org/0009-0005-5152-5856>
** 서울과학기술대학교 스마트ICT융합공학과 교수
- ORCID: <https://orcid.org/0000-0001-9136-8932>
*** 국립공주대학교 컴퓨터교육과 교수(교신저자)
- ORCID: <https://orcid.org/0000-0003-3724-3244>

• Received: Jul. 30, 2026, Revised: Aug. 13, 2026, Accepted: Aug. 16, 2026
• Corresponding Author: Dongho You
Dept. of Computer Education, Kongju National University, Kongju, Korea
Tel.: +82-41-850-8821, Email: dongho.you@kongju.ac.kr

I. 서 론

최근 실감 미디어는 사용자의 상호작용에 따라 콘텐츠가 실시간으로 재구성되는 프로그래머블 미디어(Programmable media)로 진화하고 있으며, 볼륨메트릭 방송, 전시, 원격 협업과 같이 3차원 객체의 외형(색상, 재질)을 실시간으로 전환하는 인터랙티브 서비스가 확산되고 있다[1]. 특히 점(Point)의 집합으로 3차원 공간을 정의하는 포인트 클라우드는 형상을 정의하는 기하 정보와 시각적 특성을 나타내는 속성 정보가 구조적으로 분리되어 있어, 실시간 속성 제어와 동적 시각화를 수반하는 프로그래머블 미디어 환경을 실현하기에 적합한 특성을 지닌다[2].

그러나 다수의 속성 변체를 독립 트랙 혹은 독립 파일로 저장하는 기존 방식에서는 형상을 정의하는 기하 데이터가 변체 수만큼 물리적으로 중복 적재된다. 기하 데이터의 압축 후 용량 비중이 작지 않으므로, 변체, 객체, 사용자 수가 늘어나는 대규모 인터랙티브 서비스일수록 저장, 전송 비용이 선형적으로 증가하는 문제가 발생한다[3].

본 논문에서는 이를 해결하기 위해 ISOBMFF

(ISO Base Media File Format)의 대체 트랙 구조(Alternative track structure)를 적용하여, 단일 기하 트랙을 다수의 속성 트랙이 공유하는 G-PCC 캡슐화 방식을 제안하고 구현한다[4].

제안 방식은 비표준 확장 없이 ISOBMFF에 이미 정의된 트랙 참조 및 대체 그룹 메커니즘만으로 구성되며, 기하 비트스트림을 변체 간 동일하게 공유하므로 화질 손실 없이 저장 용량을 절감하고 변체 수가 증가할수록 그 이득이 커진다. 다만 재생 중 전환이 트랙 교체만으로 이루어지도록 모든 속성 변체를 사전 복호화하여 메모리에 상주시키므로, 변체 수가 증가할수록 재생 중 메모리 사용량이 증가하는 트레이드오프가 존재한다.

II. 다중 속성 G-PCC 미디어의 통합 캡슐화 및 인터랙티브 재생 아키텍처

2.1 시스템 개요 및 전처리 과정

제안 시스템은 그림 1과 같이 다중 속성 변체 생성을 위한 전처리 및 ISOBMFF 캡슐화 단계로 구성된다.

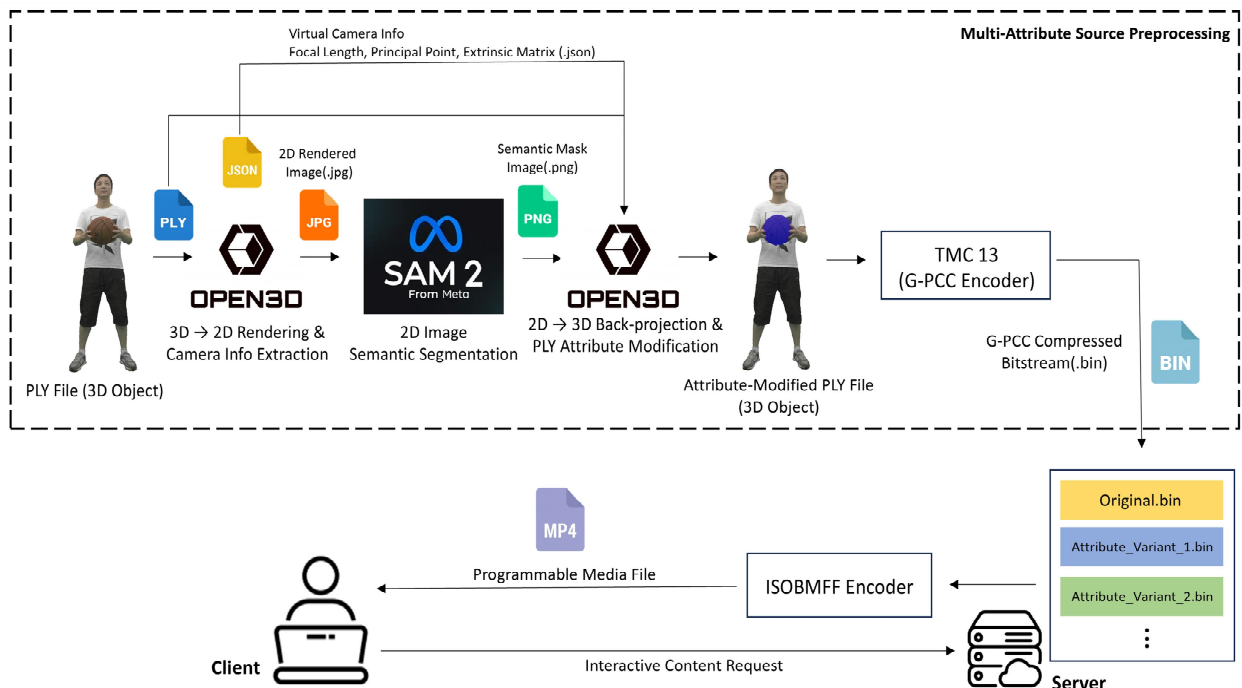


그림 1. 프로그래머블 미디어 구성을 위한 전체 시스템 아키텍처
Fig. 1. Overall system architecture for programmable media configuration

전처리에서는 Open3D로 원본 PLY를 다중의 가상 시점 영상으로 렌더링하며, 이때 초점 거리, 주점, 외부 파라미터 등 카메라 정보를 함께 추출한다. 렌더링된 2D 영상은 2D 영상 세그멘테이션 모델인 SAM 2로 분할되어 마스크로 변환되고, 저장된 카메라 파라미터를 기반으로 다시 3차원으로 역투영되어 각 점에 새로운 색상 속성이 부여된 변체 PLY를 생성할 수 있게 된다[5][6]. 이러한 과정을 여러 색상에 대해 반복하여 N개의 속성 변체를 확보한다. 이 과정에서 점의 위치 좌표는 변경하지 않으므로 모든 변체는 원본과 동일한 기하 구조를 공유하며, 이는 트랙 공유 캡슐화가 성립하기 위한 전제 조건이 된다. 이후 각 변체는 TMC13 인코더로 압축되어 ISOBMFF 멀티 트랙 구조로 통합된다.

2.2 대체 트랙 구조 기반 캡슐화

본 절에서 제안하는 구조의 핵심은, 비표준 확장 없이 표준에 정의된 트랙 참조와 대체 그룹만으로 단일 기하 트랙을 복수의 속성 트랙이 공유하도록 한 점에 있다. 그림 2와 같이 파일은 압축 데이터가 저장되는 mdat 영역과 각 트랙의 구성을 기술하는 moov 영역으로 구분되며, moov 내부에는 하나의

기하 트랙과 복수의 속성 트랙이 병렬로 존재한다. 제안 구조에서는 기하 트랙의 tref 박스 내 gpca 엔트리가 복수의 속성 트랙을 가리키는 하향식 참조 구조를 이루며 공통 GDU(Geometry Data Unit)가 N개의 ADU(Attribute Data Unit)와 데이터 중복 없이 결합된다. 기하 트랙과 속성 트랙은 시맨틱 정보로 역할이 분리되며 동시에 속성 변체 트랙들은 tkhd의 alternate_group 값을 공유하여 사용자가 특정 속성을 선택하면 그룹 내 나머지 트랙은 비활성화되고 선택된 트랙만이 재생된다[4][7]. 본 구조는 기하 데이터의 추가 변경 없이 신규 속성 트랙의 추가만으로 콘텐츠 변체를 확장할 수 있는 이점이 있다.

2.3 사전 복원 및 실시간 재생

그림 3의 재생 아키텍처는 ISOBMFF 디코더, TMC13 기반 G-PCC 디코더, 그리고 인터랙티브 제어 기능이 확장된 MPEG-3DG-Renderer 기반 OpenGL 렌더러로 구성된다. ISOBMFF 디코더는 초기 로드 단계에서 alternate_group 내 모든 변체를 사전 파싱하여 각 트랙의 위치와 대응 관계를 미리 확보하며, 이를 통해 전환 요청 시 파일을 다시 탐색하지 않고 즉시 대상 트랙에 접근할 수 있다.

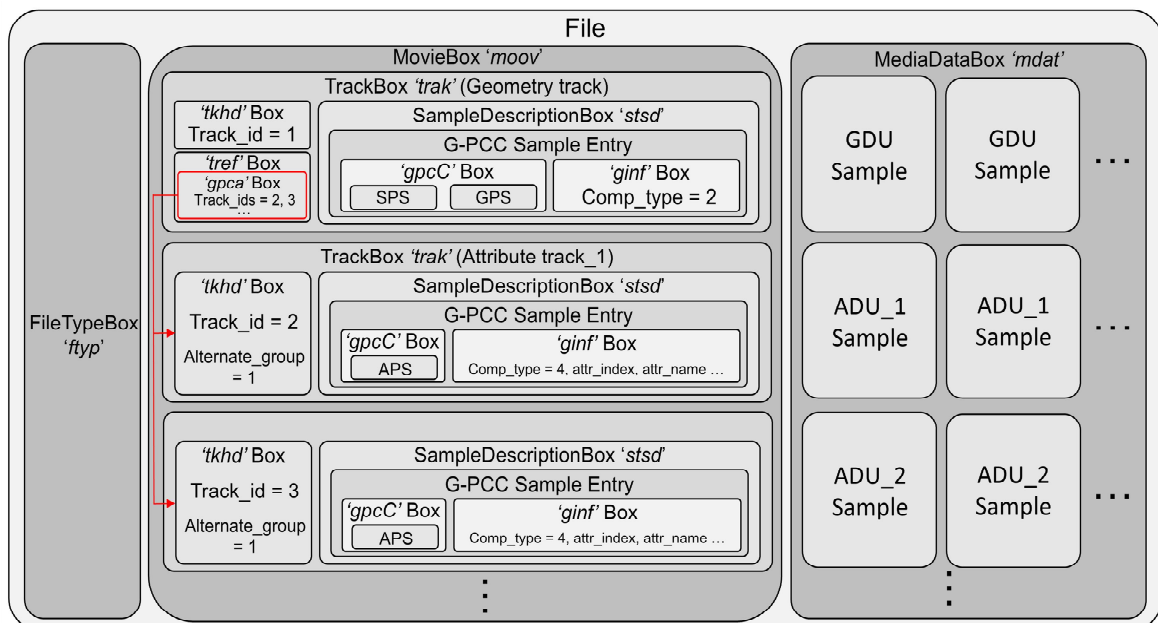


그림 2. 프로그래머블 G-PCC 미디어를 위한 ISOBMFF 기반 다중 속성 트랙 캡슐화 구조
Fig. 2. ISOBMFF-based multi-attribute track encapsulation structure for programmable G-PCC media

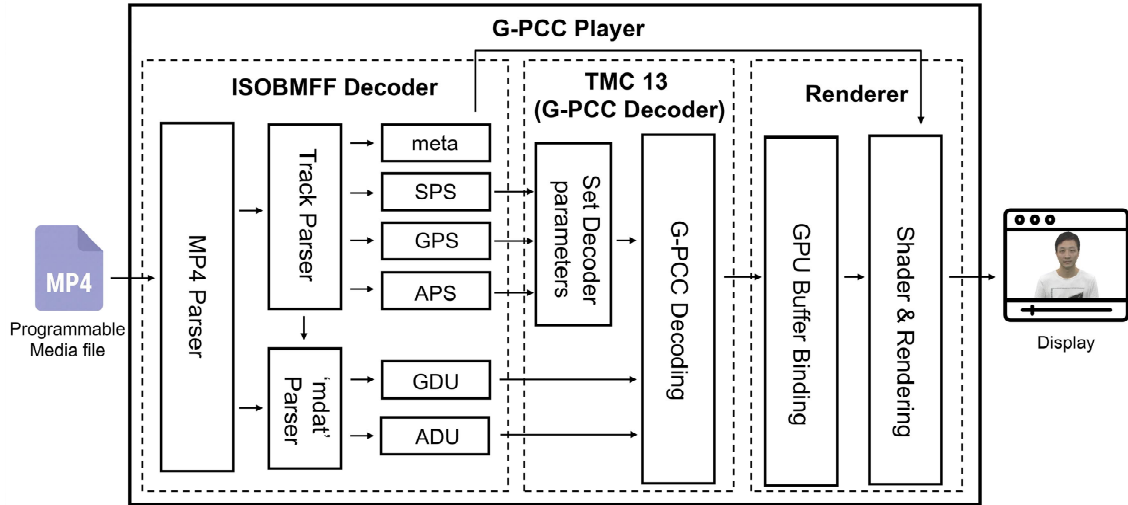


그림 3. 실시간 속성 전환을 위한 인터랙티브 재생 시스템 아키텍처

Fig. 3. Interactive playback system architecture for real-time attribute switching

G-PCC 복호화 단계에서는 모든 트랙을 복호화하여 해당 Point Cloud의 좌표 및 색을 메모리 버퍼에 상주시키고, 렌더링 시 사용자의 요청에 따라 색상 속성 버퍼의 포인트만 즉각 교체하여 실시간으로 화면을 갱신한다. 따라서 속성 전환은 추가 복호화나 파일 입출력을 동반하지 않으며, 연산 병목을 재생 단계와 완전히 분리하여 전환 시점에 추가 연산이 발생하지 않도록 설계하였다.

III. 실험 및 결과 분석

본 실험은 Ryzen 7 7800X3D CPU 및 RTX 4070 Super GPU 환경에서 C++17 표준을 기반으로 수행되었다. 시스템 구현을 위해 TMC13(v23), ISOBMFF 라이브러리, MPEG-3DG-Renderer 라이브러리 및 SAM 2를 활용하였으며, G-PCC 인코딩 시 기하 데이터의 압축 방식으로는 Octree 방식을, 속성 데이터 압축 방식으로는 RAHT 알고리즘을 적용하였다 [8]. 테스트 시퀀스로는 Basketball과 Longdress (vox10, 300프레임)를 사용하였다. 표 1은 두 시퀀스의 평균 점 개수와 데이터 유닛별 압축 크기를 정리한 것이다. Basketball은 압축 후 기하가 46,449KB, 속성이 229,853KB로 기하가 전체의 약 17%를, Longdress는 기하 60,153KB, 속성 404,278KB로 약 13%를 차지한다. 기존 방식에서는 변체마다 기하가 그대로 반복 저장되므로, 변체가 늘어날

수록 해당 비중만큼의 용량이 불필요하게 누적된다. 따라서 표 1의 기하-속성 비율은 제안 구조가 회수할 수 있는 용량의 상한을 결정하며, 이는 뒤에서 제시하는 절감률과 직접 연관된다.

표 1. 실험 데이터 셋의 특성

Table 1. Characteristics of experimental data sets

Object name	Average number of point	Compressed GDU size	Compressed ADU size
Basketball	788,713	46,449KB	229,853KB
Longdress	834,315	60,153KB	404,278KB

그림 4는 속성 변체 개수 증가에 따른 전체 파일 크기의 누적 변화를 기존 방식과 제안 방식에 대해 비교한 것이다. 기존 방식은 변체마다 기하가 반복 적재되어 용량이 선형으로 증가하는 반면, 제안 구조는 기하를 한 번만 저장하여 그 상승폭을 억제한다. 특히 두 곡선의 간격이 변체 수 증가에 따라 점점 벌어진다는 점은, 콘텐츠 다양성이 큰 실감 미디어 서비스일수록 제안 구조의 저장 효율 이점이 극대화된다는 본 연구의 핵심 주장을 뒷받침한다. 동일한 표준 준수 조건에서 변체 10개 기준 Basketball은 15.1%, Longdress는 11.7%의 파일 크기가 절감되었으며, 두 시퀀스의 절감률 차이는 표 1에서 관찰한 기하-속성 용량 비율 차이에 기인한다. 즉 전체 데이터에서 기하가 차지하는 비중이 큰 콘텐츠일수록 중복 제거로 회수되는 용량이 커진다.

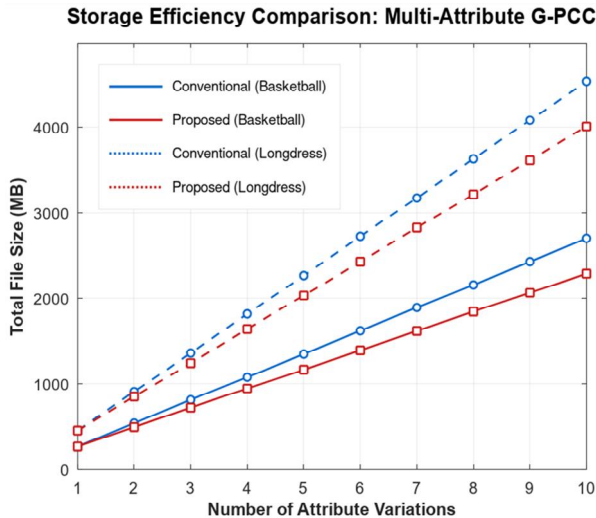


그림 4. 다중 속성 G-PCC의 저장 효율 비교
Fig. 4. Multi-attribute G-PCC storage efficiency



그림 5. 실시간 인터랙티브 속성 전환 결과
Fig. 5. Results of real-time interactive attribute switching

그림 5는 실제 구현된 플레이어의 재생 화면으로, 사용자의 선택에 따라 동일 기하 위에서 색상

속성이 즉시 전환되는 결과를 보인다. 메모리에 상주한 속성 버퍼의 포인터만 교체함으로써 추가 연산 부하 없이 전환이 이루어지며, 이는 단일 기하를 공유하는 제안 구조에서 속성 전환이 화질 손상 없이 이루어짐을 보이는 결과이다.

IV. 결 론

본 논문은 다중 속성 변체를 포함하는 G-PCC 콘텐츠의 저장·전송 효율 개선을 위해 단일 기하 트랙을 공유하는 ISOBMFF 기반 캡슐화 구조를 제안·구현하였다. 표준에 정의되어 있으나 이 용도로는 활용되지 않던 트랙 참조·대체 그룹 메커니즘을 적용하여, 화질 손실 없이 기하 중복만을 제거함으로써 저장 용량을 절감하였고, 변체 수가 증가할수록 그 이득이 커짐을 확인하였다. 다만 표준 참조 소프트웨어 제약으로 복호화 단계에서 기하 정보의 중복 연산이 잔존하므로, 향후에는 복호화된 데이터를 효율적으로 재사용할 수 있는 전용 디코더 아키텍처의 연구가 필요하다. 이러한 개선이 이루어진다면 제안하는 캡슐화 구조의 이점이 스토리지 효율을 넘어 차세대 실감 미디어 시스템의 가용성 향상으로 이어질 것으로 기대된다. 특히 제안 구조는 표준을 그대로 준수하므로 별도의 시스템 변경 없이 적용할 수 있으며, 색상뿐 아니라 반사율, 시점 등 다양한 속성으로도 확장될 수 있다. 이는 변체와 사용자 규모가 지속적으로 증가하는 대규모 인터랙티브 실감 미디어 서비스에서 제안 구조의 활용 가치가 더욱 커질 것임을 시사한다.

References

- [1] What is Programmable Media, <https://cloudinary.com/blog/whats-new-programmable-media>. [accessed: Jul. 20, 2026]
- [2] S. Schwarz, et al., "Emerging MPEG standards for point cloud compression", IEEE Journal on Emerging and Selected Topics in Circuits and Systems, Vol. 9, No. 1, pp. 133-148, Mar. 2019. <https://doi.org/10.1109/JETCAS.2018.2885981>.

- [3] S. Jeong, S. Kang, E. Kim, D. Kim, and D. You, "Dynamic attribute transformation for programmable G-PCC media", Proc. Summer Conf. of the Korean Institute of Broadcast and Media Engineers (KIBME), Jeju, Korea, pp. 138-140, Jun. 2026.
- [4] ISO/IEC, "Carriage of geometry-based point cloud compression data", ISO/IEC 23090-18, 2024.
- [5] A. Kirillov, et al., "Segment Anything", Proc. IEEE/CVF Int. Conf. on Computer Vision (ICCV), Paris, France, pp. 4015-4026, Oct. 2023. <https://doi.org/10.1109/ICCV51070.2023.00371>.
- [6] E. Kalogerakis, M. Averkiou, S. Maji, and S. Chaudhuri, "3D shape segmentation with projective convolutional networks", Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, pp. 6630-6639, Jul. 2017. <https://doi.org/10.1109/CVPR.2017.702>.
- [7] S. Jeong, G. Jeong, D. Kim, and D. You, "ISOBMFF encapsulation for carriage of G-PCC bitstream: a tutorial and guidelines", Journal of KICS, Vol. 50, No. 12, pp. 1917-1936, Dec. 2025. <https://doi.org/10.7840/kics.2025.50.12.1917>.
- [8] ISO/IEC, "Geometry-based point cloud compression", ISO/IEC 23090-9, 2023.

김 동 호 (Dong Ho Kim)



1997년 2월 : 연세대학교
전자공학과(공학사)
1999년 2월 : KAIST
전기전자공학과(공학석사)
2004년 2월 : KAIST
전기전자공학과(공학박사)
2004년 3월 ~ 2007년 2월 :

삼성기술원 삼성전자 책임 연구원
2007년 3월 ~ 현재 : 서울과학기술대학교
스마트ICT융합공학과 교수
관심분야 : 무선통신시스템, 실감미디어통신

유 동 호 (Dongho You)



2012년 2월 : 서울과학기술대학교
매체공학(공학사)
2014년 2월 : 서울과학기술대학교
미디어IT공학(공학석사)
2018년 2월 : 서울과학기술대학교
방송정보통신융합공학(공학박사)
2018년 9월 ~ 2021년 2월 : 독일

드레스덴공과대학교 도이치텔레콤연구그룹 선임연구원
2021년 3월 ~ 2025년 8월 : 한남대학교 정보통신공학과
조교수
2025년 9월 ~ 현재 : 국립공주대학교 컴퓨터교육과
부교수
관심분야 : 실감미디어 통신, 다중감각 통신, 저지연 통신
네트워크

저자소개

정 승 혁 (Seunghyeok Jeong)



2025년 2월 : 한남대학교
정보통신공학과(공학사)
2026년 2월 : 한남대학교
정보통신공학과(공학석사)
2026년 3월 ~ 현재 :
국립공주대학교 컴퓨터교육과
박사과정

관심분야 : 실감미디어 통신, 데이터 캡슐화, 미디어
스트리밍