

일반논문 (Regular Paper)

방송공학회논문지 제31권 제1호, 2026년 1월 (JBE Vol.31, No.1, January 2026)

<https://doi.org/10.5909/JBE.2026.31.1.106>

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

상용 비디오 코덱 호환성을 고려한 4D Gaussian Splatting 압축 프레임워크

이 혜 미^{a)}, 변 주 형^{a)}, 김 민 태^{a)}, 심 동 규^{a)†}

4D Gaussian Splatting Compression Framework Compatible with Commercial Video Codecs

Hyemi Lee^{a)}, Joohyung Byeon^{a)}, Mintae Kim^{a)}, and Donggyu Sim^{a)†}

요 약

본 논문에서는 정렬된 가우시안 속성 정보를 공간적으로 패킹하여, 상용 비디오 코덱과 호환되면서 두 개의 인 • 디코더만으로 효율적인 부호화 및 복호화가 가능한 4D Gaussian Splatting 압축 프레임워크를 제안한다. 제안하는 방법은 가우시안의 속성 정보를 장면 재구성에 필수적인 정보와 방향 의존적 색상 표현을 위한 구면 조화 함수 (Spherical harmonics, SH) 계수의 AC 성분으로 분리하고, 이를 두 개의 프레임으로 패킹하여 계층적으로 부호화한다. 이러한 구조를 통해 전송 환경에 따라 품질을 점진적으로 향상시킬 수 있는 프로그레시브 전송을 지원한다. 또한 기존 쿼터니언 기반 회전 표현을 로그맵 기반 3차원 벡터로 변환하여 프레임 간 회전 속성의 일관성을 유지하고 압축 효율을 향상시킨다. 제안 방법과 기존 비디오 코덱 기반 속성별 가우시안 스플랫 압축 기술을 동일 조건에서 비교한 결과, Peak Signal-to-Noise Ratio (PSNR) 기준 평균 -5.60%, Structural Similarity Index Measure (SSIM) 기준 평균 -3.72%의 BD-rate 개선을 보였으며, 이를 통해 시각적 품질을 유지하면서도 효과적으로 압축 효율을 향상시킬 수 있음을 확인하였다.

Abstract

In this paper, we propose a 4D Gaussian Splatting compression framework that spatially packs aligned Gaussian attribute information, enabling efficient encoding and decoding with only two encoders and decoders while maintaining compatibility with commercial video codecs. The proposed method separates the Gaussian attribute data into essential information for scene reconstruction and the AC components of spherical harmonics (SH) coefficients for direction-dependent color representation, which are packed into two frames and hierarchically encoded. This structure supports progressive transmission, allowing the rendering quality to be gradually improved according to transmission conditions. In addition, the conventional quaternion-based rotation representation is replaced with a log-map-based three-dimensional vector representation, which maintains temporal consistency of rotational attributes between frames and improves compression efficiency. Experimental results show that, compared with the video codec - based attribute-wise Gaussian Splat compression method, the proposed approach achieves average BD-rate reductions of -5.60% in terms of Peak Signal-to-Noise Ratio (PSNR) and -3.72% in terms of Structural Similarity Index Measure (SSIM), demonstrating that it effectively enhances compression efficiency while preserving visual quality.

Keyword : GSC, HEVC, Gaussian splat compression, Progressive transmission

Copyright © 2026 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

오랫동안 컴퓨터 비전과 그래픽스 분야에서는 3차원 장면 재구성 (3D scene reconstruction)이 핵심 연구 과제로 다뤄져 왔다. 기존의 3차원 재구성 기술은 Multi-View Stereo (MVS)^[1], Structure-from-Motion (SfM)^[2]과 같이 다중 시점 영상을 이용하는 기하학적 추정 기법을 기반으로 3차원 공간 정보를 복원해 왔다. 그러나 이러한 접근 방식은 다수의 시점 영상이 필요하고, 카메라나 피사체의 움직임에 따라 정합 오류가 발생하기 쉽다. 한편, Neural Radiance Field (NeRF)^[3]는 3차원 공간 내 좌표의 색상 및 밀도 값을 인공신경망 기반으로 학습해 정밀한 장면 재구성을 실현함으로써 3차원 장면 재구성 연구에 새로운 방향을 제시하였다. NeRF는 다중 시점 이미지로부터 3차원 장면의 복사휘도 필드 (radiance field)를 학습하고 이를 신경망 파라미터 내에 암시적으로 저장하는 방식으로 동작한다. 특정 시점의 2D 이미지를 생성하기 위해 가상의 광선을 투사하고, 그 경로 상에 위치한 여러 지점들의 색상과 밀도 값을 다층 퍼셉트론 (Multi-Layer Perceptron, MLP)으로 추론한 뒤 이를 종합하여 최종 픽셀 색상을 결정하는 볼륨 렌더링 (volume rendering) 과정을 수행한다. 이러한 접근은 제한된 수의 2D 영상으로부터 복잡한 장면을 사실적으로 재구성하고, 새로운 시점 이미지를 고품질로 생성하는데 뛰어난 성능을 보였다. 하지만 NeRF는 높은 시각적 품질을 확보하기 위해서 학습 및 렌더링 비용이 큰 신경망 구조가 필요하며, 속도를 개선하려면 필연적으로 품질 저하와의 트레이드오프 (trade-off)를 감수해야 하므로, 실시간 처리가 중요한 응용 분야에서는 실용성이 낮다는 한계

를 지닌다.

NeRF의 이러한 한계를 해결하기 위한 후속 연구들이 뒤따랐고, 3차원 공간을 다수 개의 타원 형태의 가우시안 확률분포와 그 속성값으로 명시적으로 표현하는 3D Gaussian Splatting (3DGS)^[4] 기술이 제안되었다. 이 기술은 NeRF처럼 고품질의 미분 가능한 볼륨 렌더링을 지원하면서도 1080p 기준 100fps 이상의 빠른 렌더링 속도를 제공한다. 3DGS의 각 가우시안은 기하학적 형태를 결정하는 속성과 외관을 표현하는 속성으로 정의된다. 기하학적 특성은 3차원 공간상의 각 가우시안의 중심 좌표 (Position)와 공분산 행렬로 결정되며, 공분산 행렬은 3개의 크기 (Scale) 파라미터와 4개의 회전 (Rotation) 파라미터로 표현된다. 외관 특성은 장면 내에서 가우시안의 시각적 기여도를 나타내는 불투명도 (Opacity)와 관측 방향에 따라 변화하는 색상 특성을 표현하기 위해 사용되는 구면 조화 함수 (Spherical harmonics, SH) 계수로 이루어진다. 일반적으로 SH 차수 (degree)가 3인 경우 시점에 따라 변하지 않는 색상을 표현하기 위한 3개의 계수 (SH-DC)와 시점에 따라 변화하는 색상을 표현하기 위한 성분 45개의 계수 (SH-AC)가 사용된다. 결과적으로 SH 차수가 3인 하나의 가우시안은 총 59개의 속성을 가지며, 각 속성이 32비트 실수형 형식으로 저장되기 때문에 매우 큰 저장 공간이 필요하고, 이로 인해 메모리 및 연산 자원이 제한된 모바일 기기나 웹 환경에서는 실제 응용에 제약이 따른다.

최근 Gaussian Splatting 기반 데이터 표현이 산업 전반으로 빠르게 확산되면서 Moving Picture Experts Group (MPEG)에서도 gaussian splat (GS) 기반 장면 데이터의 압축 및 전송을 위한 Gaussian Splat Coding (GSC) 표준화 논의가 활발히 진행되고 있다. 대표적인 접근 방식으로는 Self-organizing Gaussian Splats^[5]에서 소개된 Parallel Linear Assignment Sorting (PLAS) 알고리즘을 통해 고차원 가우시안 속성을 공간적으로 정렬하여 2차원 그리드로 배치하고, 이를 기존 비디오 코덱으로 부호화하는 비디오 코덱 기반 압축 방식^[6]과 Geometry-based Point Cloud Compression (G-PCC) 기반 압축 방식^[7] 등이 있다. 비디오 코덱 기반 압축 방식은 G-PCC 기반 방식보다 동일한 화질 수준에서 더 낮은 비트율을 달성하는 높은 압축 효율을 보였으며^[6], 상용 비디오 코덱을 그대로 활용할 수 있어 기존

a) 광운대학교 컴퓨터공학과(Department of Computer Engineering, Kwangwoon University)

✉ Corresponding Author : 심동규(Donggyu Sim)
E-mail: dgsim@kw.ac.kr
Tel: +82-2-940-6470

ORCID: <https://orcid.org/0000-0002-2794-9932>

※ 이 논문은 LG 전자의 실시간 3차원 입체 미디어 서비스를 위한 Gaussian splat 데이터 처리 기술 개발, 정부 (과학기술정보통신부)의 재원으로 정보통신기획평가원-학·석사연계 ICT 핵심인재양성 지원 (IITP-2025-RS-2022-00156215) 및 2025년도 정부 (교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업 (RS-2025-25432419)의 지원을 받아 수행된 연구임

• Manuscript November 17, 2025; Revised December 26, 2025; Accepted December 31, 2025.

인프라와의 호환성이 높다. 반면 G-PCC는 정적 포인트 클라우드 압축 방식으로 화면 간 예측을 지원하지 않아 시간적 중복성을 활용할 수 있는 4D Gaussian Splatting (4DGS) 압축에는 적합하지 않다. 또한 G-PCC 표준에 최적화된 하드웨어 및 산업적 인프라가 부재하여 상용화 측면에서도 제약이 존재한다. 이에 비해 비디오 코덱 기반 압축 방식은 시간적 중복성을 이용한 프레임 간 예측이 가능하다는 장점을 지니지만, 가우시안 속성 파라미터의 수가 많아 각 속성 채널을 개별 프레임으로 부호화할 경우 많은 수의 인코더 및 디코더를 필요로 한다는 한계가 있다.

본 연구는 동적 장면을 다루기 위해 프레임 단위로 구성된 3DGS 모델들이 시간 순서대로 배열된 시퀀스 구조의 4DGS 데이터를 대상으로 한다. 특히 본 연구에서 사용하는 데이터는 특정 시간 구간 내에서 가우시안 스플랫의 개수가 프레임마다 동일하고, 프레임 간 동일 인덱스의 가우시안이 일관되게 대응되도록 스플랫의 순서가 보존되는 tracked GS 시퀀스^[8]이다. 본 논문에서는 이를 효과적으로 압축하기 위해, 정렬된 가우시안 속성 정보가 담긴 각 2D 그리드를 공간적으로 패킹하여, 두 개의 비디오 인·디코더만으로 4DGS를 부호화 및 복호화할 수 있는 상용 비디오 코덱 호환 압축 프레임워크를 제안한다. 제안 방식은 가우시안의 핵심 기하 정보와 기본 색상 정보를 포함하는 Essential 프레임과, 방향 의존적 색상 표현을 위한 SH-AC 프레임을 패킹하고 각각 부호화하는 계층적 압축 구조를 따른다. 이러한 구조는 디코더가 필요에 따라 필수 프레임만 선택적으로 복호화할 수 있는 프로그레시브 전송 (progressive transmission)을 지원하므로, 렌더링 품질과 비트율 간의 균형을 유연하게 조절할 수 있다. 또한 기존 쿼터니언 (Quaternion) 기반 회전 표현을 로그맵 (Log-map) 기반 3차원 벡터 표현으로 변환하여 파라미터 수를 줄이면서도 시간축에서 회전 파라미터의 안정적인 변화를 유지해 부호화 효율을 향상시킨다. 제안 방법과 기존 비디오 코덱 기반 속성별 가우시안 스플랫 압축 기술을 동일 조건에서 비교한 결과, Peak Signal-to-Noise Ratio (PSNR) 기준 평균 -5.60%, Structural Similarity Index Measure (SSIM) 기준 평균 -3.72%의 BD-rate 개선을 보여, 시각적 품질을 유지하면서도 압축 효율을 효과적으로 향상시킬 수 있음을

확인하였다.

본 논문의 구성은 다음과 같다. 2장에서는 기존 3DGS 압축 기술에 대해 설명하고, 3장에서 상용 비디오 코덱 호환성을 고려한 4DGS 압축 방법의 전체 구조와 주요 핵심 절차를 설명한다. 4장에서는 제안하는 방법의 성능을 기존 기술과 비교하여 평가하고, 5장에서 결론을 맺는다.

II. 관련 Gaussian Splatting 압축 기술

본 장에서는 3D 및 4D Gaussian Splatting 데이터의 효율적인 압축을 위한 기존 연구들을 살펴본다. 먼저, 가우시안 속성 그리드를 공간적으로 패킹하고 이를 비디오 코덱으로 부호화하는 2D 그리드 패킹 기반 3DGS 압축 방법을 소개한다. 이어서 시간축을 포함한 동적 장면의 4DGS 데이터를 대상으로, 정렬된 각 속성별 2차원 그리드를 개별적으로 비디오 코덱으로 부호화하는 기술의 구조를 설명하며, 두 접근법의 장점과 한계를 논의한다.

1. 2D 그리드 패킹 기반 3D Gaussian Splatting 압축

비디오 코덱 기반 3DGS 압축은 가우시안 스플랫 데이터를 PLAS 알고리즘을 이용해 2차원 그리드로 정렬한 뒤, 이를 비디오 코덱으로 압축하는 절차를 따른다. 이 방식은 G-PCC 기반 3DGS 압축보다 객관적/주관적 화질 평가에서 우수한 성능을 보였으나, 실제 응용에는 몇 가지 한계가 존재한다. 먼저, 모든 가우시안 속성 채널을 개별 프레임으로 부호화할 경우 총 59개의 비디오 스트림이 생성되어, 대부분의 상용 코덱이 동시에 처리할 수 있는 스트림 수를 초과한다. 따라서 속성 그리드를 순차적으로 부·복호화해야 하며, 그 과정에서 높은 지연 (latency)이 발생하는 문제가 뒤따른다. 또한 속성 데이터를 YUV 4:4:4 포맷으로 묶어 부호화할 경우, 상용 비디오 코덱 (예: HEVC^[9], VVC^[10])의 Main Profile에서 이를 지원하지 않아 호환성 문제가 발생한다. 이를 해결하기 위해 정렬된 가우시안 속성 정보들을 공간적으로 두 개의 프레임으로 패킹하고 이를 HEVC 참조 소프트웨어를 통해 압축하는 방법^[11]이 제안되었다. 제

안된 방법은 부호화 효율과 코덱 호환성을 고려하여 전처리 과정에서 쿼터니언 회전 정보를 세 개의 오일러 각 (Euler angles)으로 변환하고, SH 계수의 AC 성분은 RGB 4:4:4에서 YUV 4:2:0 포맷으로 변환하였다. 또한 양자화 단계에서는 위치 정보를 제외한 모든 속성값을 각 속성 채널의 최소값과 최대값을 기준으로 양자화하여 8비트 정수형 데이터로 변환하고, 위치 속성은 품질 저하를 방지하기 위해 16비트로 양자화한 후 상위 바이트 (Most Significant Byte, MSB)와 하위 바이트 (Least Significant Byte, LSB)로 분리하여 저장하였다.

이후 그림 1과 같이 두 개의 YUV 4:2:0 포맷 비디오 프레임으로 속성들을 패킹하였다. 첫 번째 프레임은 위치, 크기, 회전, SH-DC 성분을 포함하고, 두 번째 프레임은 불투명도와 SH-AC 성분을 포함하였다. 두 프레임은 HEVC 참조 소프트웨어로 압축되었으며, 슬라이스 단위로 속성별 특성에 따라 서로 다른 양자화 파라미터 (quantization parameter, QP)가 적용되었다. 복호화 과정은 부호화의 역순으로 진행되었고, 복원된 모든 속성들을 통합하여 최종 3DGS 모델을 재구성하고 렌더링한 후 원본과 비교해 화질을 평가하였다. 실험 결과에 따르면 2D 그리드 패킹 기반 압축 방식은 기존 방법 대비 최대 -17.11%의 BD-rate 향상을 달성하였다. 이는 기존 59개의 비트스트림이 2개로 줄어들면서 불필요한 하이 레벨 신택스 (High-Level Syntax, HLS) 전송이 감소하고, 전처리 단계에서 오일러 각 변환과 YUV 4:2:0 색상 변환이 효과적으로 동작한 결과로 분석되었다. 따라서 제안된 방법은 상용 비디오 코덱으로 3DGS

데이터를 효율적으로 압축할 수 있음을 입증하였다. 그러나 제안된 공간적 패킹 기법은 두 비트스트림이 모두 복호화된 이후에만 렌더링이 가능하다는 제약을 가지며, 본 방법에서 사용된 오일러 각 기반 회전 표현은 3DGS 환경에서는 압축 효율을 보이지만, 오일러 각의 각 성분의 주기적 특성으로 인해 회전 성분이 $\pm\pi$ 경계를 넘을 때 값이 급격하게 변하는 래핑 (wrapping) 현상이 발생하여, 실제로는 유사한 회전임에도 불구하고 수치적으로는 큰 차이를 갖는 값으로 표현될 수 있다. 이로 인해 시간축을 포함한 4DGS에 적용할 경우 프레임 간 회전 파라미터가 불연속적으로 변하고, 그 결과 비디오 코덱의 화면 간 예측 효율이 저하되는 한계가 존재한다.

2. 비디오 코덱 기반 4D Gaussian Splatting 압축

MPEG 내에서 진행 중인 GSC 연구는 주로 G-PCC 기반으로 기존 포인트 클라우드 압축 소프트웨어를 수정하고 확장하는 형태로 발전해 왔다. 그러나 G-PCC는 정적 포인트 클라우드를 대상으로 설계되어 화면 간 예측을 지원하지 않아 4DGS 데이터 압축에는 적합하지 않으며, G-PCC 표준에 최적화된 하드웨어와 인프라가 없다는 한계점을 지닌다. 이에 따라 Zhejiang University는 비디오 코덱을 활용한 GSC 방법^[6]을 MPEG에 기고하였다. 이 방법은 프레임 단위로 가우시안 스피틀릿 데이터를 가우시안 속성 (attribute)별 비디오 시퀀스로 재구성한 뒤, 이를 비디오 코덱으로 압축하는 방법으로 G-PCC 방법보다 객관적/주관적 품

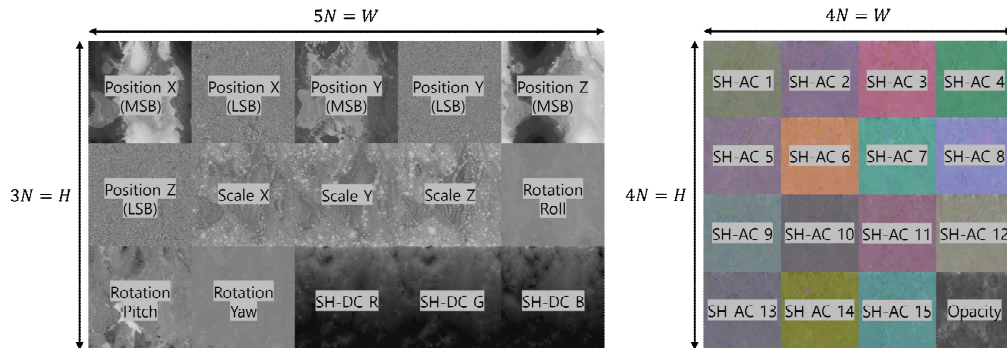


그림 1. 2D 그리드 패킹 기반 3D Gaussian Splatting 압축 방법의 가우시안 속성 정보 패킹 구조

Fig. 1. Packing structure of Gaussian attribute in the 2D grid packing-based 3D Gaussian Splatting compression method

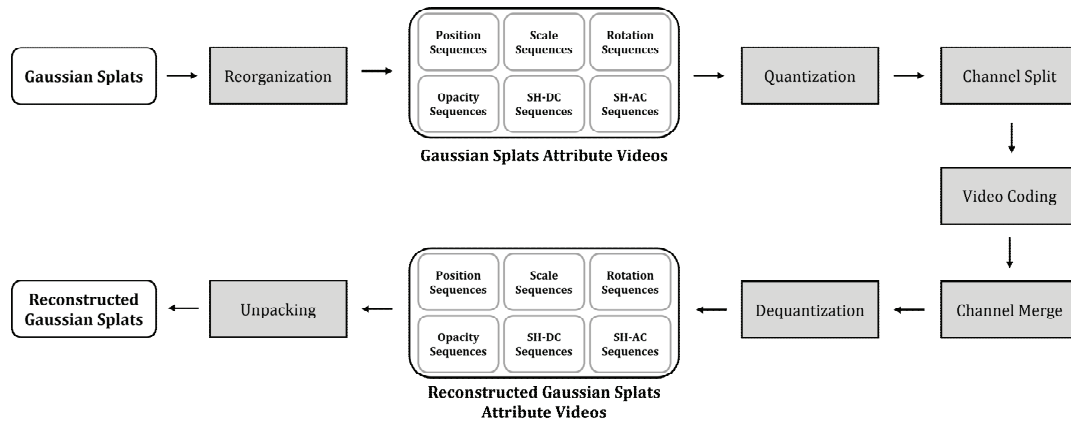


그림 2. 비디오 코덱 기반 GSC 부·복호화 구조도 (Zhejiang University, MPEG 기고)

Fig. 2. Block diagram of video codec-based GSC encoding and decoding processes (Zhejiang University, MPEG Contribution)

질 측면에서 우수한 성능을 보였다.

그림 2는 비디오 코덱 기반 GSC 방법의 전체 구조도를 나타낸다. 해당 방법에서는 먼저 PLAS 알고리즘을 이용하여 첫 번째 프레임의 가우시안들을 공간적으로 정렬하고, 이를 기준으로 생성된 인덱스 매핑 테이블을 이후 프레임에 순차적으로 적용해 모든 프레임에서 일관된 정렬 구조를 유지하였다. PLAS는 데이터를 정사각형 형태의 2D 그리드로 매핑하므로 가우시안 스플랫의 수가 정수의 제곱 형태가 되도록 부족한 경우 더미 스플랫 (dummy splats)을 추가하여 처리하였다. 정렬된 결과는 시간축으로 스택킹 (temporal stacking)되어 각 속성별로 독립된 가우시안 속성 정보 비디오를 구성하였다. 속성 비디오는 위치, 크기, 회전, 불투명도, SH 계수에 대한 시퀀스로 구성되었다. 부호화 전에 양자화 및 채널 분리 등의 전처리 과정이 수행되었으며, 각 속성 채널에 대해 전체 프레임을 대상으로 채널별 최소값과 최대값을 계산한 뒤, 이를 기준으로 최소-최대 (min-max) 기반 균등 양자화 (uniform quantization)가 적용되었다. 위치 속성은 시각적 민감도를 고려해 16비트로, 그 외 속성은 8비트로 양자화되었고, 회전 속성은 양자화 전에 단위 벡터로 정규화되었다. 위치 속성의 16비트 데이터는 상위 바이트와 하위 바이트로 분리되어 각각 8비트로 저장되었다. 채널 분리 단계에서는 쿼터니언 $[w, x, y, z]$ 를 단일 채널 w 와 3채널 $[x, y, z]$ 로 분리하였으며, 쿼터니언 w 와 불투명도 속성과 같은 단일 채널 데이터는 YUV 4:0:0 포

맷, 다채널 속성 데이터는 YUV 4:4:4 포맷으로 부호화되었다. 부호화에는 ffmpeg의 x265 라이브러리가 사용되었으며, SH 차수 3 기준으로 총 22개의 비트스트림이 생성되었다. 복호화 과정에서는 각 비트스트림을 디코딩한 후, 위치 속성의 MSB와 LSB를 병합하고, 회전 속성의 단일 채널 w 와 3채널 $[x, y, z]$ 를 결합하여 복원하였다. 이후 저장된 최소/최대값 및 비트 깊이 정보를 이용해 역양자화를 수행하여 실수형 데이터를 복원하였으며, 속성별 시퀀스는 다시 프레임 단위의 가우시안 스플랫 형태로 언패킹 (unpacking)되어 렌더링 가능한 상태로 복원되었다. 성능 검증을 위해 실험을 수행한 결과, 비디오 코덱 기반 GSC 방법은 G-PCC 기반 3DGS 압축 방법보다 낮은 비트율 구간에서 더 우수한 비트율-화질 성능 (rate-distortion performance)을 보였다. Zhejiang University에서 제안한 이 방법은 높은 압축 효율을 보이는 동시에 기존 비디오 코덱 인프라를 활용할 수 있어 빠른 상용화가 가능하다는 장점을 가진다. 그러나 본 방법은 상용 비디오 코덱을 이용해 압축하기에는 몇 가지 한계를 보였다. 우선, 사용된 YUV 4:4:4 포맷은 대부분의 상용 비디오 복호화기가 지원하는 Main Profile에서 제한되기 때문에 호환성 문제가 발생한다. 또한 제안된 방법은 총 22개의 비디오 스트림을 생성하는데, 이는 대부분의 상용 코덱이 동시에 처리할 수 있는 스트림 수를 초과하여 부호화 및 복호화 과정을 순차적으로 수행해야 하고, 그 결과 처리 지연이 발생한다는 한계가 존재한다.

III. 제안하는 4D Gaussian Splatting 압축 프레임워크

본 논문에서는 정렬된 가우시안 속성 정보를 공간적으로 패킹하여, 두 개의 비디오 인·디코더를 이용해 4DGS의 부호화 및 복호화가 가능한 상용 비디오 코덱 호환 4DGS 압축 프레임워크를 제안한다. 제안 방식은 가우시안의 속성 정보 그리드를 핵심 기하 정보와 기본 색상 정보를 포함하는 Essential 프레임과, 방향 의존적 색상 표현을 위한 SH-AC 프레임으로 분리하여 패킹하는 계층적 압축 구조를 따른다. 이러한 구조는 디코더가 필요에 따라 필수 프레임만 선택적으로 복호화할 수 있는 프로그레시브 전송을 지원하여, 렌더링 품질과 비트율 간의 균형을 유연하게 조절할 수 있다. 또한 기존 쿼터니언 기반 회전 표현을 로그맵 기반 3차원 벡터 표현으로 변환하여 회전 파라미터 수를 4개에서 3개로 줄이고, 4DGS 회전을 오일러 각으로 표현할 때 나타나는 래핑 현상에 따른 프레임 간 불연속성 문제를 해결함으로써 패킹된 프레임 간 시간적 일관성을 유지하면서 압축 효율을 향상시킨다. 제안하는 방법은 표준 비디오 코덱 환경 내에서 동작하므로 기존 상용 인프라와의 높은 호환성을 가지며 효율적인 4DGS 압축이 가능하다. 본 절에서는 먼저 제안하는 방법의 전체 과정을 설명하고, 이어서 회전 속성의 쿼터니언 로그맵 변환과 프로그레시브 전송을 지원하는 패킹 설계를 자세히 설명한다.

그림 3은 제안하는 방법의 전체 부호화 및 복호화 과정을 나타낸 블록도이다. 전체 과정은 각 프레임의 3DGS 속성들을 각각 그리드로 구성한 후, 전처리와 양자화를 거쳐 프

레이م 단위로 패킹하고, 이를 HEVC 참조 소프트웨어를 통해 압축하는 순서로 이루어진다. 먼저, 입력으로 주어진 가우시안 스플랫 시퀀스의 3DGS 모델로부터 각 가우시안의 위치, 크기, 회전, 불투명도, 그리고 구면 조화 함수 계수와 같은 속성들을 PLAS 알고리즘을 사용하여 속성들의 채널별로 2D 그리드 형태로 정렬한다. 본 연구에서는 특정 시간 구간 내에서 프레임 간 가우시안의 개수와 인덱스가 동일하게 유지되는 tracked GS 시퀀스를 입력으로 사용한다. 해당 시퀀스는 시간 구간 내 모든 프레임의 ground truth (GT) 영상을 사용하여 학습한 기준 3DGS 모델을 초기값으로 하여 생성된다. 이후 각 프레임에 대해 기준 3DGS 모델에서 시작해 프레임별 GT 영상에 맞도록 가우시안의 추가 및 제거는 제한하고, 개수와 인덱스를 고정된 상태에서 가우시안의 속성을 파인튜닝 (fine-tuning)한다. 이를 통해 모든 프레임에서 동일 인덱스의 가우시안이 서로 대응되며, 프레임 간 큰 불연속 없이 속성이 변화하도록 유도되어 시간적 일관성이 확보된다. 이에 따라 3DGS 데이터가 두 개 이상의 프레임으로 구성된 경우, 첫 번째 프레임을 기준으로 공간 정렬을 수행하고, 생성된 인덱스 매핑 테이블을 후속 프레임에 순차적으로 적용하여 모든 프레임 간의 정렬 일관성을 유지한다. 이때 2D 그리드를 정사각형 형태로 유지하기 위해 필요한 경우 더미 스플랫을 추가하여 패딩을 수행한다. 이렇게 정렬된 결과는 시간축으로 스택킹 되어 각 속성 채널별로 가우시안 스플랫 속성 비디오를 형성한다. 회전 속성은 네 개의 값을 갖는 쿼터니언 형식으로 저장되어 있으나, 압축 효율을 높이기 위해 이를 세 개의 로그맵 회전 표현으로 변환한다. 대부분의 상용 복호화기가 지원하는

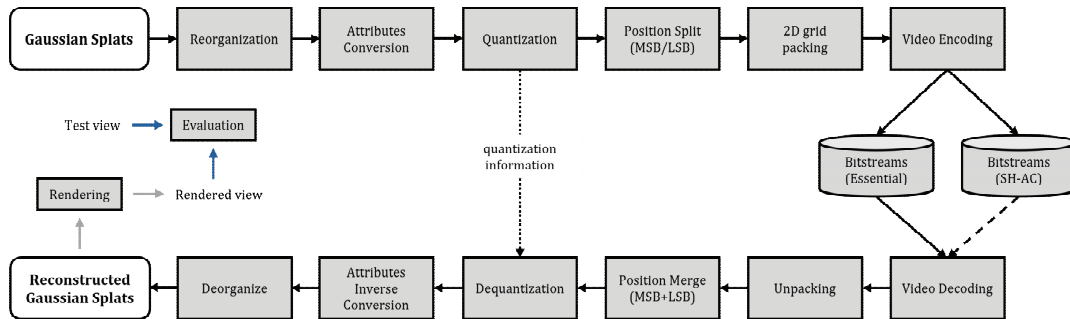


그림 3. 제안하는 4DGS 부·복호화 구조도

Fig. 3. Proposed 4DGS encoding and decoding processes

YUV 4:2:0, 8비트 영상 포맷으로 패킹을 수행하기 위해 SH 계수의 AC 성분은 ITU-R BT.709 계수를 기반으로 RGB 색공간에서 YUV 색공간으로 변환되며, 이후 U와 V 성분은 2×2 블록 평균을 통해 서브샘플링되어 4:2:0 포맷으로 변환된다. 모든 실수형 속성값은 식 (1)에 따라 균등 양자화를 수행하여 정수형 데이터로 양자화된다. 각 속성은 하나 이상의 채널로 구성될 수 있으며, 각 채널은 전 프레임에 걸쳐 계산된 최소값과 최대값을 기준으로 양자화된다. SH 계수의 AC 성분은 차수가 3일 경우 총 45개 ($15\text{개 계수} \times 3\text{채널}$)로 구성되며, 차수와 채널별로 독립적인 양자화 범위가 사용된다. 위치 정보를 제외한 속성은 8비트 깊이로 양자화되고, 시각적 민감도가 높은 위치 속성은 품질 저하를 방지하기 위해 16비트로 양자화된 후 상위 바이트와 하위 바이트로 분리되어 각각 8비트 단위로 저장된다.

$$x_{\text{quantized}} = \frac{x - \min}{\max - \min} \times (2^{\text{bitdepth}} - 1) \quad (1)$$

수식 (1)에서 x 는 실수형 속성값, $\min$ 과 $\max$ 는 해당 속성의 채널별 최소값과 최대값, bitdepth 는 양자화에 사용되는 비트 수를 의미한다. 양자화된 속성값이 담긴 2D 그리드는 두 개의 YUV 4:2:0 포맷 비디오 프레임에 패킹되고, 각 프레임은 HEVC 참조 소프트웨어를 사용하여 압축된다. 각 속성은 타일 (tile) 및 슬라이스 (slice) 단위로 구성되어 병렬 부호화가 가능하며, 속성별로 다른 양자화 파라미터를 적용할 수 있다. 복원 단계에서는 압축된 비트스트림을 디코딩한 뒤, 역패킹을 통해 공간적으로 패킹되어 있던 속성들을 각 속성별 2D 그리드로 복원한다. 위치 속성은 분리된 MSB와 LSB를 결합하여 복원되고, 모든 양자화된 정수형 속성은 저장된 최소/최대값과 비트 깊이 정보를 이용해 역양자화되어 실수형으로 복원된다. SH 계수의 AC 성분은 YUV에서 RGB로 변환되며, 4:2:0 포맷의 U, V 성분은 하나의 샘플을 2×2 영역에 복사하는 방식으로 업샘플링된다. 로그맵 형식으로 저장된 회전 정보는 다시 쿼터니언 형식으로 변환되고, 이렇게 복원된 모든 가우시안 속성들을 통합하여 프레임마다 최종 3DGS 모델을 재구성한다. 마지막으로 해당 모델을 렌더링하여 원본과의 시각적 유사도를 평가하며, 학습에 사용하지 않은 테스트 시점에

대한 평균 PSNR와 SSIM을 측정한다.

1. 로그맵 기반의 3차원 회전 표현

3DGS에서 각 가우시안의 회전은 일반적으로 4차원 단위벡터인 쿼터니언으로 표현된다. 쿼터니언은 회전축과 회전각을 하나의 4차원 단위벡터로 표현함으로써, 오일러 각에서 발생하는 짐벌락 (gimbal lock) 문제 없이 학습 과정에서 회전을 안정적으로 표현할 수 있다. 하지만 각 가우시안의 회전은 3차원 회전군 $SO(3)$ 의 원소임에도 불구하고, 쿼터니언은 4개의 파라미터를 필요로 하므로 학습 이후 압축·전송 효율 측면에서는 비효율적이다. 이를 해결하기 위해 쿼터니언으로 표현된 가우시안의 회전을 오일러 각으로 변환하여 효과적으로 3DGS를 압축할 수 있다. 쿼터니언과 오일러 각 표현 간의 관계는 식 (2)와 같이 정의된다.

$$\begin{aligned} \text{roll} &= \text{atan2}(2(wx + yz), 1 - 2(x^2 + y^2)) \\ \text{pitch} &= \text{asin}(2(wy - zx)) \end{aligned} \quad (2)$$

$$\text{yaw} = \text{atan2}(2(wz + xy), 1 - 2(y^2 + z^2))$$

오일러 각 기반 회전 표현은 쿼터니언의 4차원 회전 정보를 3차원으로 축소하여, 렌더링 화질의 손상 없이 비트율을 감소시킬 수 있는 효율적인 방식이다. 그러나 오일러 각 변환을 시간축을 포함하는 4DGS 환경에 적용하면 프레임 간 회전 파라미터가 주기적 성질에 따라 불연속적으로 바뀔 수 있어, 비디오 코덱의 화면 간 예측 효율이 저하되는 문제가 발생한다. 이러한 불연속성은 오일러 각의 각 성분이 주기적 특성을 가지며, 서로 다른 값이 동일한 실제 회전을 나타내는 특성에서 기인한다. 예를 들어 roll 과 yaw 는 $-\pi$ 와 π 가 동일한 회전을 의미하며, pitch 는 $-\frac{\pi}{2}$ 와 $\frac{\pi}{2}$ 에서 roll 과 yaw 의 조합에 따라 같은 회전을 표현할 수 있어 중복 표현이 존재한다. 그림 4는 Bartender 시퀀스의 frame 0에서 쿼터니언으로 학습된 회전 정보를 오일러 각으로 변환한 결과의 성분별 분포를 나타낸 것이다. pitch 는 $\pm \frac{\pi}{2}$ 부근에서의 분포가 거의 없는 반면, roll 과 yaw 는 $\pm \pi$ 근처에서 높은 빈도를 보인다. 따라서 동일한 가우시안이

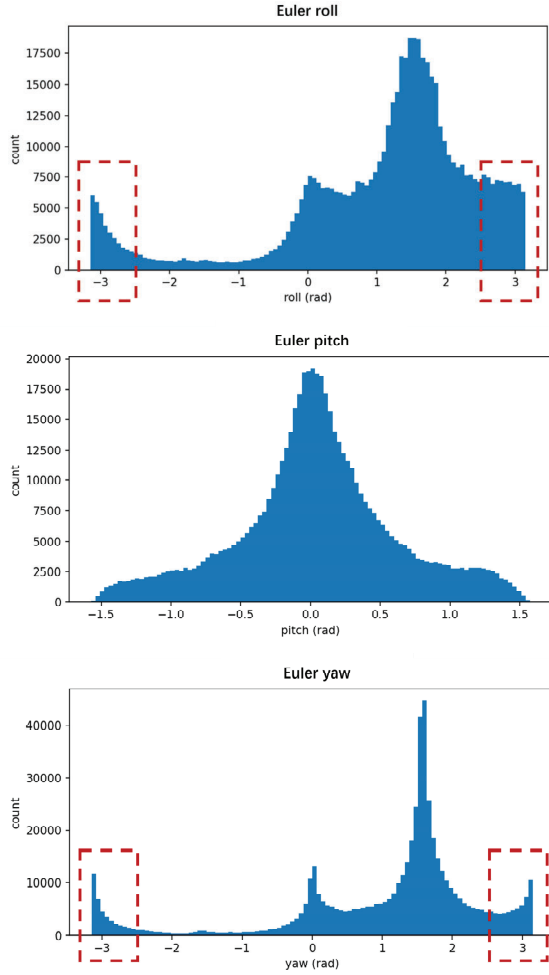


그림 4. Bartender 시퀀스 frame 0의 쿼터니언 회전 정보로부터 변환된 Euler 각 성분의 분포

Fig. 4. Distributions of Euler angle components converted from quaternion rotations in Bartender frame 0

시간적으로 유사한 회전을 갖더라도, 오일러 각으로 변환되는 과정에서 *roll*과 *yaw* 값이 $\pm\pi$ 경계에서 래핑 되어 $-\pi$ 와 π 사이를 반복적으로 오가며 시간축을 따라 회전 파라미터 간 큰 값의 차이가 발생할 수 있다. 결과적으로 오일러 각 표현 방법은 4DGS 환경에서 프레임 간 회전 파라미터의 시간적 불연속성을 초래해, 비디오 코덱의 화면 간 예측 효율을 저하시킬 수 있다.

이에 본 연구에서는 4차원 쿼터니언 회전 정보를 3차원

으로 축소하면서도 패킹된 프레임 간 일관성을 유지하기 위해 쿼터니언을 로그맵 기반의 3차원 벡터 형식으로 변환하는 방법을 제안한다. 쿼터니언은 회전축 $\mathbf{n}$ 과 회전각 θ 로 분해될 수 있으며, 이때 회전축 방향으로 크기 θ 를 갖는 벡터 $\mathbf{r} = \theta\mathbf{n}$ 이 로그맵 벡터로 정의된다. 이 변환을 통해 가우시안 스플랫의 회전 정보를 표현하는 데 필요한 파라미터 수를 4개에서 3개로 줄여 압축 효율을 높일 수 있다. 쿼터니언으로 표현된 회전 정보를 로그맵 기반의 3차원 벡터로 변환하는 과정은 다음과 같다. 먼저 각 프레임의 쿼터니언을 단위 벡터로 정규화한 뒤, $\mathbf{q} = (w, x, y, z)$ 에 로그맵 변환을 적용한다. 이때 로그맵 벡터 $\mathbf{r} = (r_x, r_y, r_z)$ 는 식 (3)과 같이 정의된다. $\|\mathbf{v}\|$ 가 매우 작은 경우에는 수치적 안정성을 위해 로그맵 벡터를 근사적으로 표현하며, 이때는 식 (4)와 같이 정의된다.

$$\|\mathbf{v}\| = \sqrt{x^2 + y^2 + z^2}, \theta = 2 \times \text{atan2}(\|\mathbf{v}\|, w), \quad (3)$$

$$\mathbf{r} = \theta \times \frac{\mathbf{v}}{\|\mathbf{v}\|} = (r_x, r_y, r_z)$$

$$\mathbf{r} \approx 2\mathbf{v} = (2x, 2y, 2z) \quad (4)$$

복호화 시에는 식 (5)를 이용하여 로그맵 벡터를 다시 쿼터니언으로 변환한다. 이때 $\|\mathbf{r}\|$ 이 매우 작은 경우에는 수치적 안정성을 위해 식 (6)을 사용하여 근사적으로 계산할 수 있다.

$$\theta = \|\mathbf{r}\| = \sqrt{r_x^2 + r_y^2 + r_z^2}, \quad (5)$$

$$\mathbf{n} = \frac{\mathbf{r}}{\|\mathbf{r}\|}, w = \cos\left(\frac{\theta}{2}\right), (x, y, z) = \mathbf{n} \times \sin\left(\frac{\theta}{2}\right)$$

$$(x, y, z) \approx \left(\frac{r_x}{2}, \frac{r_y}{2}, \frac{r_z}{2}\right), w \approx 1 - \frac{\theta^2}{8} \quad (6)$$

그림 5는 Bartender 시퀀스의 frame 0에서 쿼터니언으로 학습된 회전 정보를 회전축-회전각 (axis-angle) 형태로 변환했을 때의 회전각 θ 의 분포를 나타낸다. 회전각 θ 는 0과 2π 일 때 동일한 회전을 의미하며, 단위 쿼터니언 공간 $\mathcal{S}^3$ 에서는 이 두 값이 부호만 다른 동일한 쿼터니언에 대응한다. 로그맵에서 회전각 θ 는 0과 2π 에서 래핑이 발생하지

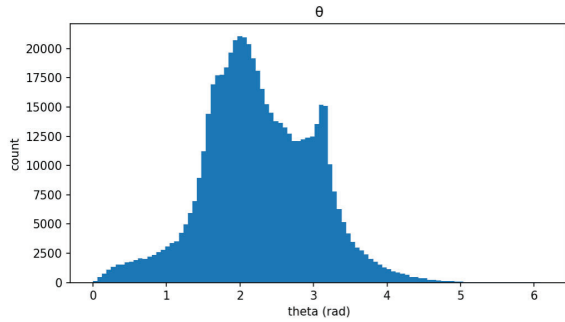


그림 5. Bartender 시퀀스 frame 0의 쿼터니언 회전 정보로부터 변환된 회전 각 θ 분포

Fig 5. Distributions of rotation angle θ from quaternion rotations in Bartender frame 0

만, 학습 과정에서 쿼터니언의 부호가 프레임 간 완전히 반전되는 경우는 거의 없기 때문에 학습된 쿼터니언으로부터 얻은 회전각 θ 의 분포는 0과 2π 근처에서 낮게 나타나며, 전체적으로 중간 회전 각도에 집중되는 경향을 보인다. 또한 쿼터니언의 부호 반전이 드물다는 점은 회전축 $\mathbf{n}$ 의 방향이 프레임 간 일관되게 유지됨을 의미하므로, 로그맵 표현에서 회전축 역시 시간적으로 안정적으로 변화한다. 이러한 특성으로 인해 로그맵 회전 표현 $\mathbf{r} = \theta\mathbf{n}$ 은 오일러 각 표현과 달리 래핑으로 인한 급격한 값 변화가 거의 발생하지 않아, 시간축을 따라 회전 파라미터 값이 부드럽고 안정적으로 변화한다.

그림 6은 Bartender 시퀀스의 frame 0과 frame 1에서 추

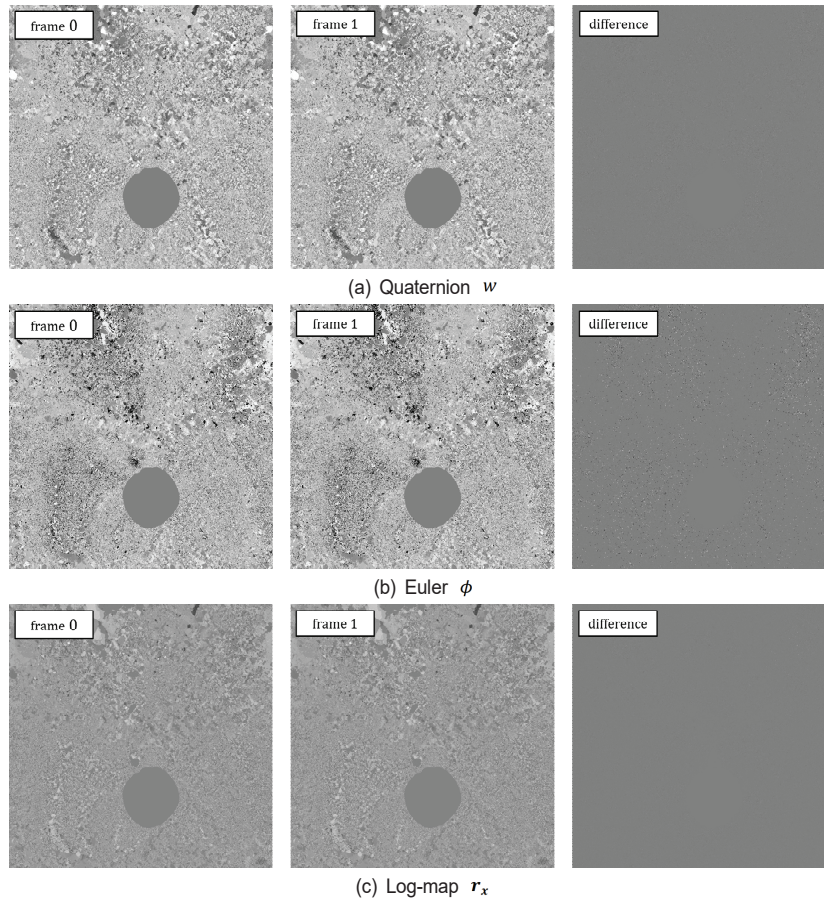


그림 6. 가우시안 회전 속성 표현 방식에 따른 프레임 간 회전 속성 그리드의 질댓값 차이

Fig. 6. Absolute frame-to-frame differences of rotation attribute grids for different Gaussian rotation representations

출한 회전 속성 정보를 (a) 쿼터니언 w , (b) 오일러 $roll$, (c) 로그맵 r_x 기준으로 각각 시각화하고, 두 프레임 간 회전 속성값의 절대값 차이 (absolute frame-to-frame difference) 그리드를 비교한 결과를 나타낸다. 쿼터니언 표현의 경우 w 속성 그리드에서 프레임 간 차이가 거의 나타나지 않는다. 동일한 가우시안이 시간적으로 유사한 회전을 가지더라도, 오일러 각 표현에서는 래핑 특성으로 인해 프레임 간 큰 값 차이가 발생하는 가우시안이 다수 존재한다. 이러한 특징은 $roll$ 속성 그리드의 프레임 간 차이에서도 뚜렷하게 나타난다. 반면, 제안하는 로그맵 기반 표현은 래핑 구간에 해당하는 값의 분포가 매우 낮아 r_x 속성 그리드의 프레임 간 차이가 쿼터니언과 유사하게 작게 유지된다. 이러한 프레임 간 회전 그리드의 일관성은 4DGS 환경에서 부호화 효율을 향상시키는 데 유리하게 작용한다.

2. 점진적 전송이 가능한 가우시안 스플랫 속성의 공간적 패킹 방법

제안하는 방법은 단일 디코더 환경에서도 점진적인 전송과 복원이 가능한 가우시안 스플랫 속성의 공간적 패킹 방식을 따른다. 본 방법은 3DGS 속성 데이터를 장면 재구성에 필수적인 정보와 방향 의존적 색상 표현을 위한 SH 계수의 AC 성분으로 분리하고, 각각을 하나의 프레임으로 패킹하여 총 두 개의 비디오 스트림으로 압축한다. 이러한 구

조는 네트워크 환경이나 디코더 수에 따라 Essential 프레임만 복호화하거나 SH-AC 프레임을 추가적으로 복호화함으로써 품질을 점진적으로 향상시킬 수 있다. 기존 연구 (그림 1)는 3DGS 속성 데이터를 2D 프레임으로 패킹해 비디오 코덱으로 압축하였으나, 장면 재구성에 필수적인 속성과 부가적인 속성을 구분하지 않아 두 프레임을 모두 복호화해야만 완전한 3DGS 모델을 복원할 수 있었다. 이로 인해 한 프레임만 수신된 상태에서는 렌더링을 시작할 수 없으며, 두 프레임의 복호화가 순차적으로 수행되어야 하므로 단일 디코더 환경에서는 처리 지연이 발생하고, 네트워크 상황에 따른 유연한 전송 제어가 어려웠다. 제안하는 방법은 이러한 한계를 해결하기 위해 양자화된 속성 그리드를 그림 7과 같이 3DGS 속성을 장면 재구성에 필수적인 정보와 SH 계수의 AC 성분으로 분리해 두 개의 프레임으로 패킹하고, HEVC 참조 소프트웨어를 이용해 압축한다. 제안하는 방법은 가장 널리 사용되는 HEVC Main Profile 디코더에서 복호화가 가능하도록 YUV 4:2:0 색상 포맷으로 가우시안 속성들을 패킹하였으며, 양자화 에러가 시각적 품질에 큰 영향을 미치는 위치 속성은 16비트로 양자화한 뒤 상위 바이트와 하위 바이트로 분리해 8비트로 패킹하고, 그 외 나머지 속성은 8비트 정밀도로 양자화하여 패킹 후 부호화한다. 구체적으로 Essential 프레임은 위치, 크기, 회전, 불투명도, SH 계수 DC 성분 정보를 포함하고, SH-AC 프

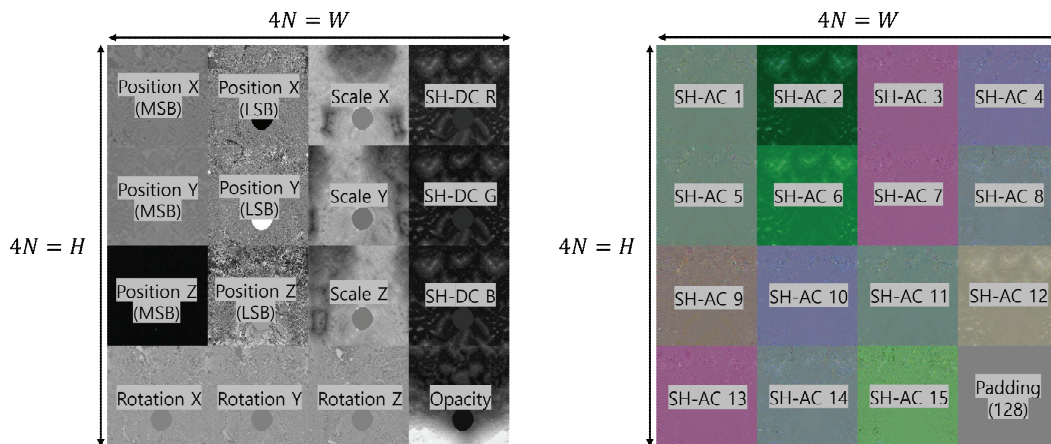


그림 7. 제안하는 가우시안 속성 정보 패킹 구조
Fig. 7. Proposed packing structure of Gaussian attributes

레이은 SH 계수 AC 성분 정보를 포함한다. SH 차수가 3인 경우 총 45개의 AC 성분이 존재하며, 각 채널에는 15개의 AC 속성 그리드가 배치된다. 이때 사각형 구조를 유지하기 위해 남은 한 블록은 상수 128 값의 Padding 영역으로 채운다. MPEG Bartender, Breakfast, Cinema 데이터셋에 대해 각 속성 그리드의 크기는 각각 768×768 , 768×768 , 704×704 이며, 이를 $4N \times 4N$ 형태로 패킹한 결과 프레임의 해상도는 3072×3072 , 3072×3072 , 2816×2816 으로 상용 비디오 디코더가 무리 없이 처리할 수 있는 해상도 범위에 해당한다. 패킹된 프레임에서 각 속성 그리드는 하나의 타일이자 슬라이스로 구성되어 타일 단위로 병렬 부호화 및 복호화가 가능하며, 슬라이스 단위로 서로 다른 양자화 파라미터를 적용할 수 있어 속성의 특성에 따라 양자화율을 유연하게 조정할 수 있다. 두 프레임은 서로 독립적으로 복호화되므로 네트워크 환경이나 단말기의 디코더 개수에 따라 전송 및 복원 방식을 유연하게 조정할 수 있다. 예를 들어, 두 개의 디코더를 병렬로 사용할 경우 관측 방향에 따른 색 변화까지 표현하는 SH 계수 AC 성분을 포함한 고품질의 3DGS 장면을 재구성할 수 있다. 반면, 네트워크 대역폭이 제한되거나 단일 디코더를 사용하는 환경에서는 Essential 프레임만 수신하여 즉시 렌더링을 수행할 수 있으며, 수신되지 않은 SH-AC 성분 값은 0으로 대체해도 3DGS 시퀀스 복원이 가능하다. 또한 Essential 프레임을 먼저 복호화하여 초기 렌더링을 수행하고, 렌더링과 동시에 SH-AC 프레임을 복호화해 렌더링 결과를 갱신함으로써 품질을 점진적으로 향상시키면서 전체 처리 지연을 최소화할 수 있다.

IV. 실험 결과

본 논문에서 제안하는 방법의 성능을 평가하기 위해 Zhejiang University가 MPEG에 기고한 비디오 코덱을 활용한 속성별 가우시안 스피릿 압축 방법과의 결과 비교 실험을 수행하였다. 두 방법은 동일한 조건에서 평가되었으며, 압축 과정에서는 HEVC 참조 소프트웨어 HM-16.26^[12]을 사용하고, 속성별 양자화 비트 깊이와 비디오 코덱 양자화 파라미터를 동일하게 설정하였다. 모든 실험형 속성값

은 각 속성 채널의 최소값과 최대값을 기준으로 균등 양자화를 수행하여 8비트 정수형 데이터로 변환하였다. 위치 속성의 경우 일정 비트 깊이 이하로 양자화될 때 시각적 품질 저하가 두드러지는 특성이 있으므로, 16비트 정밀도로 양자화한 뒤 상위 바이트와 하위 바이트로 분리하여 각각 8비트 단위로 저장하였다. 표 1은 실험에 사용된 네 개의 레이트 포인트별 각 속성의 비디오 양자화 파라미터를 나타내며, 이러한 파라미터는 속성의 특성에 따른 손실이 최종 렌더링 품질에 미치는 영향을 고려하여 설정하였다. 제안하는 방법의 성능을 평가하기 위한 실험은 HEVC CTC (Common Test Condition)의 LD (Low Delay) 조건에서 진행하였으며, 실험 데이터셋은 Bartender, Breakfast, Cinema 데이터이고, 본 논문에서는 각 데이터의 첫 16개의 프레임을 사용하였다. 객관적 화질 평가는 압축 및 복원 과정을 거쳐 프레임마다 재구성된 3DGS 모델을 학습에 사용되지 않은 테스트 시점에서 렌더링한 영상과 원본 영상을 비교하는 방식으로 수행하였다. 각 프레임별로 PSNR과 SSIM을 계산하였으며, 모든 테스트 시점과 프레임에 대한 평균 값을 최종 성능 지표로 사용하였다. 실험은 AMD Ryzen 7 5800X3D 8-Core 프로세서, 64G RAM, Ubuntu 22.04 운영체제 환경에서 진행하였다.

표 1. 가우시안 속성별 양자화 파라미터 (QP)

Table 1. Quantization parameters (QP) for gaussian attributes

	Rate 1	Rate 2	Rate 3	Rate 4
Scale	21	16	11	6
Rotation	27	22	17	12
Opacity	31	26	21	16
SH-DC	29	24	19	14
SH-AC 1~3	37	32	27	22
SH-AC 4~8	43	38	33	28
SH-AC 9~15	49	44	39	34
Position (MSB)	lossless	lossless	lossless	lossless
Position (LSB)	lossless	lossless	lossless	lossless

표 2. 비디오 코덱 기반 속성별 GS 압축 방법 대비 제안하는 방법의 BD-rate 및 디코딩 시간 성능

Table 2. BD-rate and decoding time performance comparison between the proposed method and video codec-based per-attribute GS compression method

Sequence	BD-rate [%] (PSNR)	BD-rate [%] (SSIM)	Decoding time [%]
Bartender	-6.71%	-5.31%	101.31%
Breakfast	-6.77%	-3.78%	102.20%
Cinema	-3.33%	-2.07%	100.61%

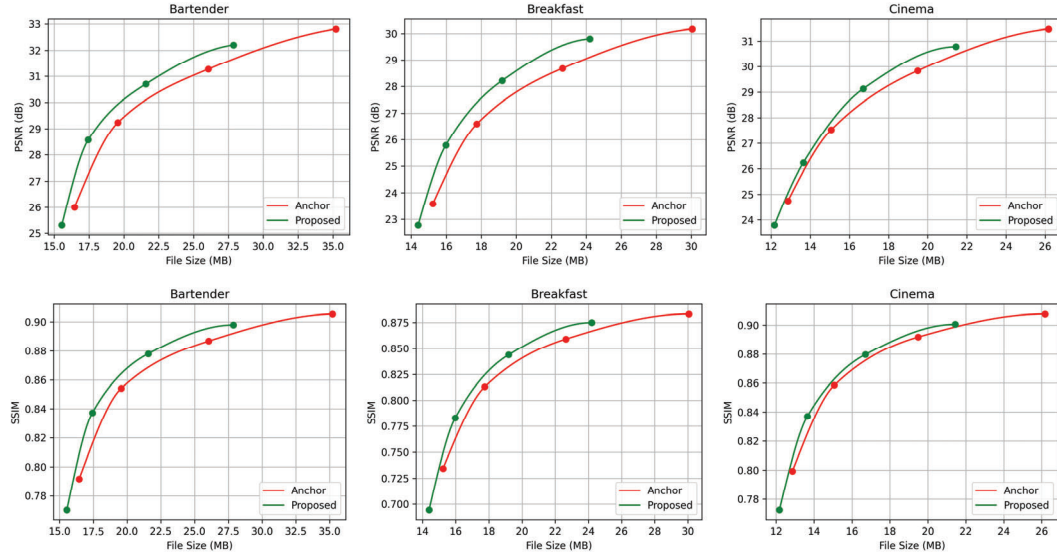


그림 8. 비디오 코덱 기반 속성별 GS 압축 방법과 제안 방법의 RD-curve 비교

Fig. 8. RD-curve comparison between the proposed method and video codec-based per-attribute GS compression method

표 2와 그림 8은 각각 비디오 코덱을 활용한 속성별 가우시안 스플랫 압축 방법 대비 제안하는 방법의 테스트 시퀀스별 BD-rate 및 디코딩 시간 결과와 RD-curve를 나타낸다. 그 결과, Bartender, Breakfast, Cinema 장면에 대해 PSNR 기준으로 각각 -6.71%, -6.77%, -3.33%의 BD-rate 향상 효과를, SSIM 기준으로는 각각 -5.31%, -3.78%, -2.07%의 향상 효과를 확인하였다. 이는 제안하는 방법에서 회전 표현의 로그맵 변환과 YUV 4:2:0 포맷으로의 색상 변환이 효과적으로 동작하였음을 의미한다. 또한 비디오 코덱 기반 속성별 가우시안 스플랫 압축 방법에서 22개의 비트스트림에 대해 전송되던 하이 레벨 신택스가 2개로 축소되어 비트량이 감소하였다. 비디오 코덱 기반 속성별 가우시안 스플랫 압축 방법에서는 각 속성에 대해 별도의 비트스트림이 생성된다. 구체적으로, 회전 표현을 위한 쿼터니언 $\mathbf{w}$ 성분과 opacity 성분은 4:0:0 포맷으로 부호화되어 2개의 비트스트림이 생성되고, 나머지 성분들은 4:4:4 포맷으로 부호화되어 추가로 20개의 비트스트림이 생성된다. 이때 4:0:0 비트스트림의 하이 레벨 신택스 크기는 스트림당 91 bytes, 4:4:4 비트스트림의 하이 레벨 신택스 크기는 스트림당 92 bytes이므로, 총 22개의 비트스트림에 대한 하

이 레벨 신택스 크기는 2,022 bytes에 이른다. 반면, 제안하는 방법에서는 Essential 프레임과 SH-AC 프레임의 두 비트스트림을 4:2:0 포맷으로 부호화하며, 각 프레임의 하이 레벨 신택스 크기는 96 bytes로 총 192 bytes에 불과하다. 결과적으로 제안하는 방법은 비디오 코덱 기반 속성별 가우시안 스플랫 압축 방법 대비 하이 레벨 신택스 크기를 약 90.5% 감소시키며, 이러한 하이 레벨 신택스 오버헤드 감소가 전체 비트량 감소에 기여하여 제안하는 방법은 시각적 품질을 유지하면서도 기존 방법보다 우수한 압축 효율을 달성하였다. 한편, 표 2의 디코딩 시간은 단일 디코더가 전체 비트스트림을 순차적으로 처리하는 환경을 전제로 측정되었다. 제안하는 방법은 비디오 코덱 기반 속성별 가우시안 스플랫 압축 방법과 비교했을 때 Bartender, Breakfast, Cinema 시퀀스에서 각각 약 101.31%, 102.20%, 100.61%의 디코딩 시간을 기록하여 디코딩 시간이 소폭 증가하였으나 그 차이는 크지 않다. 또한 상용 HEVC 디코더 및 하드웨어 인프라가 4K 해상도 영상의 실시간 디코딩을 충분히 지원하고 있으므로, 제안하는 방법 또한 실제 응용 환경에서 무리 없이 실시간 디코딩이 가능하다.

그림 9는 각 테스트 시퀀스의 첫 번째 프레임에 대해 레

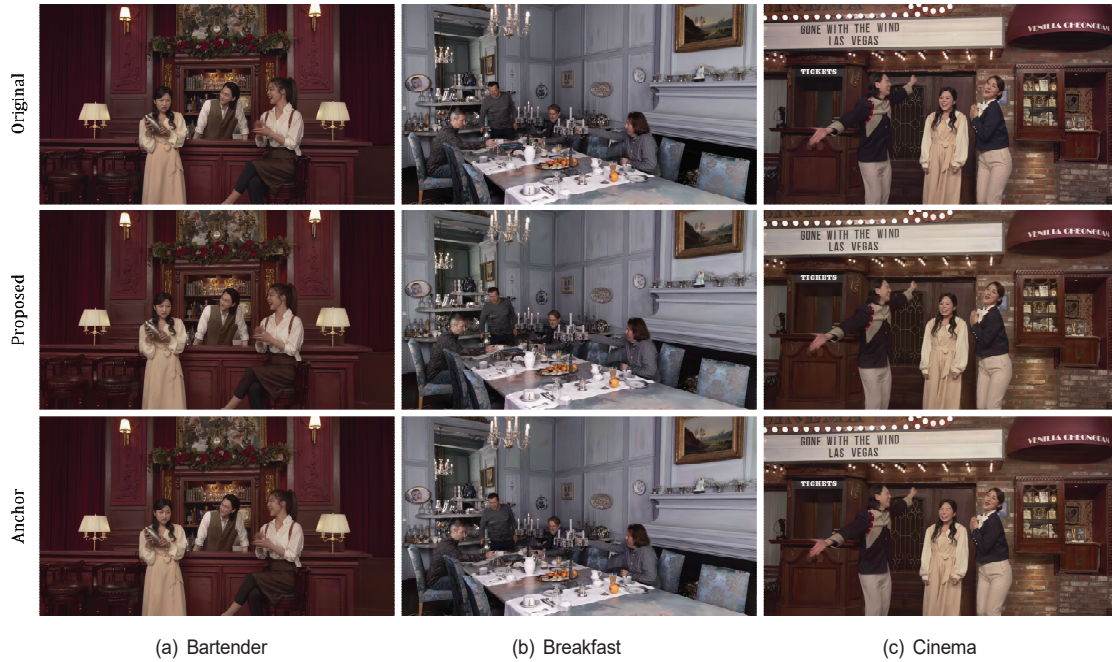


그림 9. 원본 모델과 제안하는 방법 기반 복원 모델과 비디오 코덱을 활용한 속성별 GS 압축 방법 기반 복원 모델의 렌더링 결과 비교
Fig. 9. Comparison of rendering results of the original model, the reconstruction model based on the proposed method, and the reconstruction model based on the video codec-based per-attribute GS compression method

이트 포인트 4에서의 렌더링 결과를 비교한 것이다. 비교 영상은 모두 모델 학습에 사용되지 않은 테스트 시점에서 렌더링되었으며, 원본 모델, 제안하는 방법, 그리고 비디오 코덱 기반의 속성별 압축 방법의 복원 결과를 나열하였다. 동일 레이트 포인트 기준, 제안하는 방법의 객관적 화질 지표는 비디오 코덱 기반의 속성별 압축 방법 대비 소폭 낮게 측정되었다. 그러나 주관적 화질 평가 측면에서는 원본 모델의 렌더링 결과와 매우 높은 유사도를 보여 우수한 복원 성능을 확인할 수 있다.

표 3과 그림 10은 제안하는 4DGS 부·복호화 구조에서 가우시안 스피클렛의 회전 정보를 오일러 각으로 변환한 경우와 로그맵 기반 3차원 벡터로 변환한 경우를 비교한 테스트 시퀀스별 BD-rate와 RD-curve를 나타낸다. 실험 결과, Bartender, Breakfast, Cinema 시퀀스에서 PSNR 기준으로 각각 -8.97%, -7.13%, -11.36%의 BD-rate 개선을, SSIM 기준으로는 각각 -6.08%, -4.51%, -5.33%의 개선을 확인하였다. 기존의 2D 그리드 패킹 기반 3D Gaussian Splatting

표 3. 오일러 각 기반 방법 대비 로그맵 기반 회전 표현 방법의 BD-rate 성능
Table 3. BD-rate performance comparison between Euler angle-based and Log-map-based rotation representation methods

Sequence	BD-rate [%] (PSNR)	BD-rate [%] (SSIM)
Bartender	-8.97%	-6.08%
Breakfast	-7.13%	-4.51%
Cinema	-11.36%	-5.33%

압축 연구에서는 가우시안의 회전을 쿼터니언에서 세 개의 오일러 각으로 변환하여 부호화하였다. 그러나 이를 4DGS 환경에 동일하게 적용한 결과, 오일러 각의 주기적 특성으로 인해 특히 *roll*과 *yaw* 성분에서 $\pm\pi$ 부근의 불연속적 값 변화가 크게 나타났고, 프레임 간 회전 속성 그리드의 변화가 커 부호화 효율이 감소하는 한계가 확인되었다. 이는 제안하는 로그맵 기반 회전 표현이 오일러 각 표현보다 4DGS 압축에서 더 높은 부호화 효율을 제공함을 보여준다.

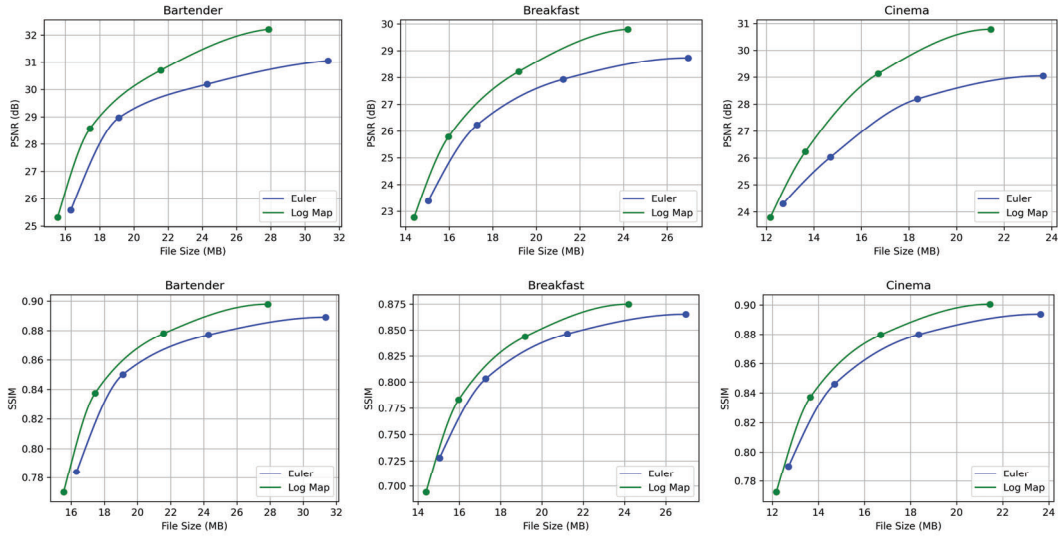


그림 10. 오일러 각 표현을 사용한 방법과 로그맵 기반 회전 표현을 적용한 방법의 RD-curve 비교
Fig. 10. RD-curve comparison between Euler angle-based and Log-map-based rotation representation methods

그림 11은 제안하는 방법에서 Essential 프레임만 복호화한 경우와, SH-AC 프레임을 추가로 복호화한 경우의 RD-curve를 비교한 결과이다. 이를 통해 Essential 프레임 만으로도 장면의 구조와 전반적인 색상 정보를 정상적으로

재구성할 수 있음을 확인하였으며, SH-AC 프레임을 함께 복호화하면 비트량은 약간 증가하더라도 빛의 방향에 따른 색상 변화가 반영되어 PSNR과 SSIM이 향상되고 시각적 품질이 개선됨을 확인하였다. 그림 12는 모델 학습에 사용

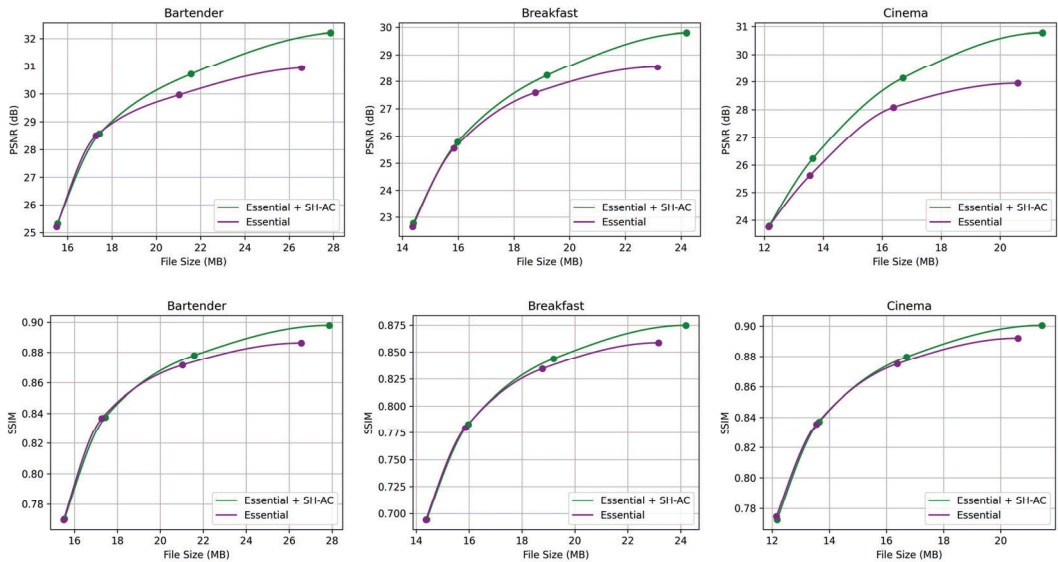


그림 11. 필수 프레임과 SH-AC 성분 프레임 복호화 구성에 따른 RD-curve 비교
Fig. 11. RD-curve comparison according to decoding configurations of Essential and SH-AC frames

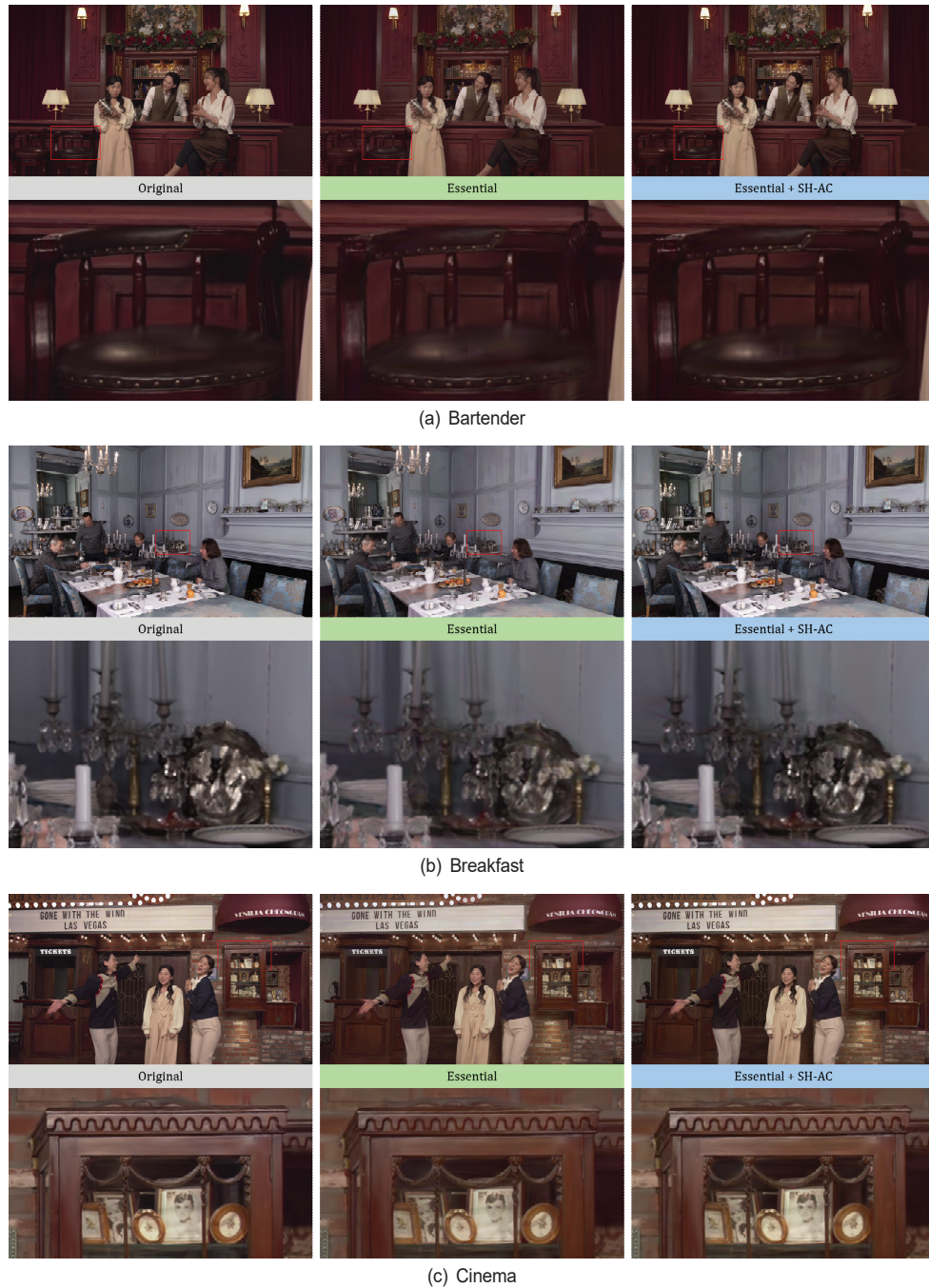


그림 12. 원본 모델과 제한하는 필수 프레임 기반 복원 모델 및 SH-AC 프레임 포함 복원 모델의 렌더링 결과 비교

Fig. 12. Comparison of rendering results for the original model, the proposed Essential frame-based reconstruction model, and the reconstruction model including SH-AC frame

되지 않은 테스트 시점에서 각 시퀀스의 첫 번째 프레임을 렌더링한 결과로, 레이트 포인트 4 환경에서의 세 가지 복원 영상을 비교하여 보여준다. 각 데이터셋마다 원본 가우시안 스플랫 모델의 렌더링 결과와 제안하는 방법으로 복원된 두 가지 구성 (① Essential 프레임만 복호화한 경우, ② Essential 프레임과 SH-AC 프레임 모두 복호화한 경우)의 렌더링 결과를 시각적으로 비교하였다. SH-AC 성분을 포함하여 복호화한 경우에는 시점 변화에 따른 색상 변화나 반사 효과가 사실적으로 재현되어, 재질감과 조명 효과가 원본에 가까운 모습을 보였다. Essential 프레임만 복호화한 경우에도 주관적인 화질 열화가 크지 않고, 전반적으로 자연스럽게 안정적인 장면 복원이 가능하였다. 이러한 결과는 제안하는 방법이 필수 속성과 SH-AC 성분을 분리한 프레임 기반의 프로그래시브 전송 구조를 가지고 있으며, 이 구조가 전송 효율 측면에서 효과적임을 보여준다. 즉, 디코더는 네트워크 대역폭이나 응용 목적에 따라 SH-AC 프레임을 선택적으로 수신할 수 있으며, Essential 프레임만 복호화하더라도 큰 화질 손실 없이 비트량을 절감하면서 안정적인 장면 복원이 가능하다. 따라서 제안하는 방법은 전송 환경에 따라 품질-비트를 절충이 가능한 유연하고 효율적인 4DGS 부호화 구조를 제공한다.

V. 결 론

본 논문에서는 정렬된 가우시안 속성 정보를 공간적으로 패키징하여, 상용 비디오 코덱 환경에서 두 개의 인·디코더만으로 효율적인 부호화 및 복호화가 가능한 4D Gaussian Splatting 압축 프레임워크를 제안하였다. 제안하는 방법에서는 가우시안의 속성 정보를 핵심 기하 정보와 기본 색상 정보를 포함하는 프레임과 방향 의존적 색상 표현을 위한 SH 계수의 AC 성분 정보를 포함하는 프레임으로 분리하여 패키징하는 계층적 압축 구조를 사용함으로써 프로그래시브 전송을 지원한다. 또한 회전 속성 그리드의 프레임 간 일관성을 유지하며 부호화 효율을 향상시키기 위해 기존 쿼터니언 기반 회전 표현을 로그맵 기반 3차원 벡터 표현으로 변환하였다. 제안하는 방법은 비디오 코덱 Main Profile이 지원하는 표준 포맷 내에서 압축이 가능하므로 기존 상용

코덱 인프라와의 호환성이 높다는 장점을 가진다. 실험 결과, 제안하는 방법은 비디오 코덱을 활용한 속성별 가우시안 스플랫 압축 방법 대비 LD 환경에서 PSNR 기준 BD-rate가 평균 -5.60%, SSIM 기준 BD-rate가 평균 -3.72% 향상된 성능을 보였으며, 이를 통해 시각적 품질을 유지하면서도 압축 효율을 효과적으로 개선함을 확인하였다. 향후 연구에서는 tracked GS 시퀀스뿐만 아니라 프레임 간의 정보 연속성이 보장되지 않는 non-tracked GS 시퀀스를 대상으로도 동작할 수 있는 범용적인 4DGS 압축 기술로 본 프레임워크를 발전시키고자 한다.

참 고 문 헌 (References)

- [1] M. Goesele, B. Curless, and S. M. Seitz, "Multi-View Stereo Revisited," 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), pp. 2402-2409, June, 2006.
doi: <https://doi.org/10.1109/CVPR.2006.199>
- [2] Johannes L. Schönberger and Jan-Michael Frahm, "Structure-from-Motion Revisited," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4104-4113, June, 2016.
doi: <https://doi.org/10.1109/CVPR.2016.445>
- [3] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis," European Conference on Computer Vision (ECCV), Lecture Notes in Computer Science, Vol. 12346, pp. 405-421, November, 2020.
doi: https://doi.org/10.1007/978-3-030-58452-8_24
- [4] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, "3D Gaussian Splatting for Real-Time Radiance Field Rendering," ACM Transactions on Graphics, Vol. 42, No. 4, pp. 1-14, July, 2023.
doi: <https://doi.org/10.1145/3592433>
- [5] W. Morgenstern, F. Barthel, A. Hilsmann, and P. Eisert, "Compact 3D Scene Representation via Self-Organizing Gaussian Grids," European Conference on Computer Vision (ECCV), Lecture Notes in Computer Science, Vol. 15143, pp. 18-34, November, 2024.
doi: https://doi.org/10.1007/978-3-031-73013-9_2
- [6] Sicheng Li, Yiyi Liao, and Lu Yu, "[GSC][JEE2] A Potential Video-based Anchor for Gaussian Splats Coding," ISO/IEC JTC 1/SC 29/WG 4/m72063, The 152nd MPEG meeting, Online, March, 2025.
- [7] Kyohei Unno, Diego Fujii, Keisuke Nonaka, and Kei Kawamura, "[JEE6] Preliminary software implementation based on G-PCC 1st edition for anchor generation of 3D Gaussian coding," ISO/IEC JTC 1/SC 29/WG 7/m70095, The 148th MPEG meeting, Kemer, TR, November, 2024.
- [8] Jun Young Jeong, Reagan Koo, Kwan-Jung Oh, Hong-Chang Shin, Gwangsoon Lee, "[GSC][JEE 6.1-related] Training Method for

- Generating Temporally Consistent Per-frame I-3DGS Dataset,” ISO/IEC JTC 1/SC 29/WG 4/m71763, The 150th MPEG meeting, Online, March, 2025.
- [9] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, “Overview of the High Efficiency Video Coding (HEVC) Standard,” IEEE Transactions on Circuits and Systems for Video Technology, Vol. 22, No. 12, pp. 1649 - 1668, December, 2012.
doi: <https://doi.org/10.1109/TCSVT.2012.2221191>
- [10] M. Lee, H. Song, J. Park, B. Jeon, J. Kang, J. Kim, Y. Lee, J. Kang, and D. Sim, “Overview of Versatile Video Coding (H.266/VVC) and Its Coding Performance Analysis,” IEIE Transactions on Smart Processing and Computing, Vol. 12, No. 2, pp. 122-154, April, 2023.
doi: <https://doi.org/10.5573/IEIESPC.2023.12.2.122>
- [11] Hyemi Lee, Joohyung Byeon, Mintae Kim, Donggyu Sim, “3D Gaussian Splatting Compression via 2D Grid Packing”, 2025 The Korean Institute of Broadcast and Media Engineers Summer Conference, pp. 912-915, 2025, June.
- [12] HEVC Reference Software, Version 16.26. https://vcgit.hhi.fraunhofer.de/jvet/HM/-/tree/HM-16.26?ref_type=tags

저 자 소 개



이 혜 미

- 2025년 2월 : 광운대학교 컴퓨터공학과 학사
- 2025년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 석사
- ORCID : <https://orcid.org/0009-0004-0367-4630>
- 관심분야 : 3D데이터압축, 영상압축, 컴퓨터비전



변 주 형

- 2019년 2월 : 광운대학교 컴퓨터공학과 학사
- 2021년 2월 : 광운대학교 컴퓨터공학과 석사
- 2021년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- ORCID : <https://orcid.org/0000-0002-6165-9189>
- 관심분야 : 3D데이터압축, 영상압축, 컴퓨터비전



김 민 태

- 2023년 2월 : 광운대학교 컴퓨터공학과 학사
- 2025년 2월 : 광운대학교 컴퓨터공학과 석사
- 2025년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- ORCID : <https://orcid.org/0009-0007-9100-8567>
- 관심분야 : 3D데이터압축, 영상압축, 컴퓨터비전



심 동 규

- 1993년 2월 : 서강대학교 전자공학과 공학사
- 1995년 2월 : 서강대학교 전자공학과 공학석사
- 1999년 2월 : 서강대학교 전자공학과 공학박사
- 1999년 3월 ~ 2000년 8월 : 현대전자 선임연구원
- 2000년 9월 ~ 2002년 3월 : 바로비전 선임연구원
- 2002년 4월 ~ 2005년 2월 : University of Washington Senior research engineer
- 2005년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 교수
- ORCID : <https://orcid.org/0000-0002-2794-9932>
- 관심분야 : 영상신호처리, 영상압축, 컴퓨터비전